

Train to Opportunity: the Effect of Infrastructure on Intergenerational Mobility

Julián Costas-Fernández
José-Alberto Guerra
Myra Mohnen

Documento CEDE

#48

Diciembre de 2020

Serie Documentos Cede, 2020-48 ISSN 1657-7191 Edición electrónica. Diciembre de 2020

© 2020, Universidad de los Andes, Facultad de Economía, CEDE. Calle 19A No. 1 – 37 Este, Bloque W. Bogotá, D. C., Colombia Teléfonos: 3394949- 3394999, extensiones 2400, 2049, 2467

infocede@uniandes.edu.co

<http://economia.uniandes.edu.co>

Impreso en Colombia – Printed in Colombia

La serie de Documentos de Trabajo CEDE se circula con propósitos de discusión y divulgación. Los artículos no han sido evaluados por pares ni sujetos a ningún tipo de evaluación formal por parte del equipo de trabajo del CEDE. El contenido de la presente publicación se encuentra protegido por las normas internacionales y nacionales vigentes sobre propiedad intelectual, por tanto su utilización, reproducción, comunicación pública, transformación, distribución, alquiler, préstamo público e importación, total o parcial, en todo o en parte, en formato impreso, digital o en cualquier formato conocido o por conocer, se encuentran prohibidos, y sólo serán lícitos en la medida en que se cuente con la autorización previa y expresa por escrito del autor o titular. Las limitaciones y excepciones al Derecho de Autor, sólo serán aplicables en la medida en que se den dentro de los denominados Usos Honrados (Fair use), estén previa y expresamente establecidas, no causen un grave e injustificado perjuicio a los intereses legítimos del autor o titular, y no atenten contra la normal explotación de la obra.

Universidad de los Andes | Vigilada Mineducación Reconocimiento como Universidad: Decreto 1297 del 30 de mayo de 1964. Reconocimiento personería jurídica: Resolución 28 del 23 de febrero de 1949 Minjusticia.

Documento CEDE

Los documentos CEDE son producto de las investigaciones realizadas por al menos un profesor de planta de la Facultad de Economía o sus investigadores formalmente asociados.

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

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

Abstract

Can transport infrastructure promote intergenerational mobility? This paper estimates the causal impact of 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 100% censuses of 1851, 1881 and 1911. By geolocating individuals down to the street level, we measure access to the railroad network using the distance to the nearest train station. To address the non-random access to the railroad network, we create a hypothetical railway map based solely on geographic cost consideration. We find that sons who grew up one standard deviation (roughly 5 km) closer to the train station are 6 percentage points more likely to work in a different occupation than their father and 5 percentage points more likely to be upward mobile. Access to the railroad network benefitted families at the top and bottom of the occupational ranking. Through a decomposition exercise, we find that the majority of upward mobility is driven by improvements 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.

[†]CReAM, University College London; j.costas-fernandez@ucl.ac.uk

[‡]Department of Economics, Universidad de los Andes; ja.guerra@uniandes.edu.co

[§]Department of Economics, University of Ottawa and University of Essex; Research Associate, Centre for Economic Performance, LSE; mmohnen@uottawa.ca

Tren hacia la Oportunidad: Efecto de la Infraestructura sobre la Movilidad Intergeneracional*

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

Resumen

¿Puede la infraestructura de transporte promover la movilidad intergeneracional? Este documento estima el impacto causal de la red ferroviaria en la movilidad ocupacional intergeneracional en Inglaterra y Gales durante el siglo XIX. Creamos una nueva base de datos que empareja padres e hijos presentes en los censos de los años 1851, 1881 y 1911. Al geolocalizar la residencia de las personas hasta el nivel de la calle, medimos el acceso a la red ferroviaria según la distancia entre la residencia y la estación de tren más cercana. Dado que el acceso a la red ferroviaria no es aleatorio, creamos una red ferroviaria hipotética basada únicamente en consideraciones geográficas y costo efectivas. Encontramos que los hijos que crecieron una desviación estándar (aproximadamente 5 km) más cerca de la estación de tren tienen una probabilidad mayor, de 6 puntos porcentuales, de trabajar en una ocupación distinta a la de sus padres y 5 puntos porcentuales más de probabilidad de presentar una movilidad ascendente. El acceso a la red ferroviaria benefició a las familias pertenecientes a la parte baja y alta de la distribución del ranking ocupacional. A través de un ejercicio de descomposición, encontramos que la mayor parte de la movilidad ascendente está explicada por mejoras en las oportunidades laborales locales.

Palabras Clave: Movilidad intergeneracional, infraestructura, movilidad espacial

Clasificación JEL: H54, J62, N13

* Agradecemos a la Economic History Society por el apoyo financiero, al UK Data Service y al Cambridge Group for the History of Population and Social Structure por compartir generosamente sus datos.

[†] CReAM, University College London; j.costas-fernandez@ucl.ac.uk

[‡] Universidad de los Andes, ja.guerra@uniandes.edu.co

[§] Department of Economics, University of Ottawa and University of Essex, mmohnen@uottawa.ca.

I Introduction

The role of infrastructure in fostering economic prosperity goes back to Adam Smith’s *Wealth of Nations*, which listed “the duty of erecting and maintaining certain public works” among the three core obligations of the sovereign. Transport infrastructures can also improve the economic opportunity of individuals by connecting residents to job opportunities further away and attracting better options locally. Large transport infrastructure has been widely used by policy-makers with this objective in mind. For instance, the high speed railway linking up London, the Midlands, the North and Scotland (HS2) lists as one of its aim to “make Britain better better connected, rebalancing the UK economy and bring jobs and investment to the Midlands and North.”¹ China has proposed an ambitious high-speed rail (HSR) program by 2030 to connect all provincial capitals (excluding Lhasa) and large cities with more than half million people. 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.²

While there is a growing belief that transport infrastructure can affect long run labour market opportunities, there is a lack of empirical evidence to support this belief. The empirical literature on the evaluation of transport infrastructure has largely focused on growth of productivity and aggregate income. Especially in the present context of increasing levels of inequality in many countries, assessing the role of transport infrastructure on individual outcome is important. In particular, the degree to which economic and social outcomes are passed along generations is key to assess the persistence of inequality across societies. Can infrastructure promote economic opportunities and break the link between parents and their children’s economic status?

This paper seeks to fill this gap. We estimate the effect of the railroad network access on intergenerational mobility. We exploit the expansion of the railroad network 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 to residents by increasing their geographic mobility and changing the economic landscape. Thanks to this historical setting, we are able to examine long term effects of this large public investment on the intergenerational mobility patterns.

We create a new and unique dataset which combines several historical datasets. We use the newly digitised 100% historical censuses of 1851, 1881 and 1911 for England and Wales, developed by the I-CeM project (Schürer and Higgs, 2014). We link individual across censuses and identify close to 1 million father-sons

¹<https://www.hs2.org.uk/why/connectivity/>

²<https://www.nytimes.com/2009/12/23/world/asia/23iht-letter.html>

pairs. Following previous papers (Boberg-Fazlic, Sharp et al., 2013; Clark and Cummins, 2015; Ferrie, 2005; Long and Ferrie, 2013; Olivetti and Paserman, 2015), we measure intergenerational mobility through occupation. We geographically locate individuals down to the street level based on their address in the census. This dataset permits the analysis of a large and representative sample at a more geographically disaggregated level than was previously feasible. We document striking spatial patterns in intergenerational mobility. Using the digitised railroad network (Alvarez, Bogart, Satchell, Shaw-Taylor and You, 2017), we are also able to measure access to the railroad network at an individual level based on the geographical proximity between the place of residence and the nearest train station.

To determine the effect of railroad access on intergenerational mobility, we exploit the spatial and temporal variation in the expansion of the rail network between 1851 and 1881. We compare the intergenerational occupation mobility of individuals who grew up closer to a railroad station to those who grew up further away. We address the non-random proximity to the railroad network by creating 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 baseline identifying assumption is that individuals who grew up closer to a railroad station experienced changes in economic opportunities only through the better access to railroad network, conditional on county and census year fixed effects, and a set of control variables including proximity to historical centers and travel routes and household characteristics.

Our main results present the causal effect of railroad access on intergenerational mobility. Growing up closer to a train station led to a break between the occupation 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 as their father. They were also 5 percentage points more likely to be upward mobile. These effects appear to be driven not only by a significant move out of farming activities, but also transitions into industrial and commercial activities. The railroad also had distributional consequences. It resulted in a shift in the distribution of occupational ranking where sons with better access to the railroad network moved away from the middle towards to top and bottom 25 percentage point 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 possible channels: the change in local labour market opportunities, the ease of geographic mobility, and the change in the relative benefit from moving. Better connected sons were 9 percentage points more likely to move. To examine the returns to moving for intergenerational mobility patterns, we first take the decision to move as given. Conditional on moving away from the county, better connected sons were 4 percentage points more likely to be upward mobile but less likely to be downward mobile. On the other hand, conditional on staying, better connected sons were also 4 percentage points more likely to be upward mobile but also more likely to be downward mobile. This points to changes in labour market opportunities and consequently opportunity costs of moving. We then estimate the returns to geographical mobility taking into account the selection by looking at brothers who grew up in the same household. Specifically, we compare the sons who moved with their brothers who stayed in the county. Based on back-of-the-envelope estimates, 90% of upward mobility is driven by labour market opportunities, while only 7% come from geographical mobility. We also show significant heterogeneity in this decomposition based on the main historical activity of cities. Manufacturing and port towns experienced even greater local opportunities brought by the railroad. This decreased the relative benefit from moving away. In contrast, resort and mining towns experienced lower local labour market opportunities.

Taken together, our results demonstrates that social mobility issues can be tackled by place-based policies at a local level. There is significant evidence across countries that lower-income populations tend to suffer from restricted transport options. Transport infrastructure creates economic opportunities that have long term effect on social mobility. This is consistent with the growing evidence that place matters for upward mobility. Chetty and Hendren (2018); Chetty, Hendren, Kline and Saez (2014) find that geographic isolation, as measured by lengthy commute times, was a significant factor in people's ability to leave poverty. Other research demonstrates that the presence of public transit improves access to employment at all levels and that transit reduces the geographic mismatch between households and employment. Moreover, when evaluating the effectiveness of transport infrastructure investments, focusing on spatial mobility and disregarding local effects would provide an underestimate of the impact.

This paper contributes to several strands of the literature. Starting with the work of Aschauer (1989), a vast analytical and empirical literature has been concerned with the effects of infrastructure development on income growth, productivity and welfare. Recent work has investigated the impact of transport networks on economic outcomes such as growth, trade, and urbanization (Banerjee et al., 2020; Campante and Yanagizawa-Drott, 2018; Donaldson, 2018; Donaldson and Hornbeck, 2016; Duranton and Turner,

2012; Faber, 2014; Lipscomb, Mobarak and Barham, 2013; Pascali, 2017). Chatterjee and Turnovsky (2012) find that government spending on infrastructure has increased regional disparities within Europe, and Artadi and Sala-i Martin (2003) point to excessive public investment as a contributing factor to rising income inequality in Africa. These studies have focused either on relatively aggregate outcomes or on cross sectional individual-level outcomes.³ We complement the existing literature by providing micro-level evidence on how individual respond to these large-scale interventions.

We also contribute to the intergenerational mobility literature (Solon (2002) and Black and Devereux (2011) provide extensive surveys of the empirical economics studies). Many economic historians have documented the intergenerational mobility and long-run trends in distribution in Britain (Jackson, 1987; Lindert, 2000). Long (2013) used census data to suggest that accounting for differences in the life-cycle of fathers and sons suggests that social mobility in nineteenth century Britain was somewhat greater what previously documented. Clark and Cummins (2015) however show persistence of social status based on the analysis of surnames. In this paper, we use the newly digitised full census of 1851 to 1911 and employ a novel linking method suggested by (Abramitzky, Mill and Pérez, 2019) to construct intergenerational occupational mobility. Using a representative sample of close to 1 million father-son pairs, we document for the first time the spatial clustering of intergenerational mobility in nineteenth century in England and Wales.

While the literature documents differences in intergenerational mobility across regions within countries and changes in intergenerational mobility over time, the factors that determine changes and regional 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; Chetty, Hendren and Katz, 2016; Guerra and Mohnen, 2020; Long and Ferrie, 2013). These factors shape access to physical capital and accumulation of human capital. Different types of economic shocks, including economic downturns or upturns, natural resource booms, and technological changes, may well reinforce or break the transmission of economic status, and the resulting changes may persist for one or more generations. Alesina, Hohmann, Michalopoulos and Papaioannou (2019) find that colonial investments in the transportation network and missionary activity are associated with upward mobility. Another closely related paper is the one by Perez (2017). Similarly to our paper, 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

³The broader literature on the impacts of transport infrastructure is summarized in Redding and Turner (2015).

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. We investigate England and Wales during the second Industrial Revolution. This important episode in history has been much debated and can provide important insights into the nature of growth and social mobility. Moreover, our data allows us to measure connectivity at the individual level. We find that the railroad network broke the link between father-son occupations and generated upward mobility for sons who grew up closer to the railroad network. We see important transitions not only out of farming but into commercial and industrial activities. We show that this implied a shift in the distribution of occupation with winners and losers. Finally, our decomposition exercise shows that the majority of the changes brought by the railroad comes from local labour market opportunities.

The rest of the paper is organised as follows. Section 2 paints the historical background of the railway system in the nineteenth century. It also describes our newly constructed datasets by linking several historical sources. In section 3 we present descriptives on intergenerational mobility patterns 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

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

The main period of railway construction was 1825-1914. 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, authorised by Act of Parliament in 1826, 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 enterprises. During the Railway Mania, 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.⁴

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.⁵ In 1840, the Board of Trade with its Railway Department was created. It was the first government department to assume responsibility for railways. 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 which was affecting other industries such as mining, manufacturing, and shipping. Railways facilitated the formation of a steam-based 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. It was exported through ports to the west of Carlisle. Southampton became one of the most successful Atlantic ports, thanks to the London and Southampton Railway (which developed into the London and South Western Railway). They also gave a great stimulus to industry by reducing the freight costs of heavy materials such as coal and minerals, as well as reducing costs of transporting finished goods around the country.

Steam facilitated speed, and speed attracted passenger traffic as well as freight. The potential of speed en-

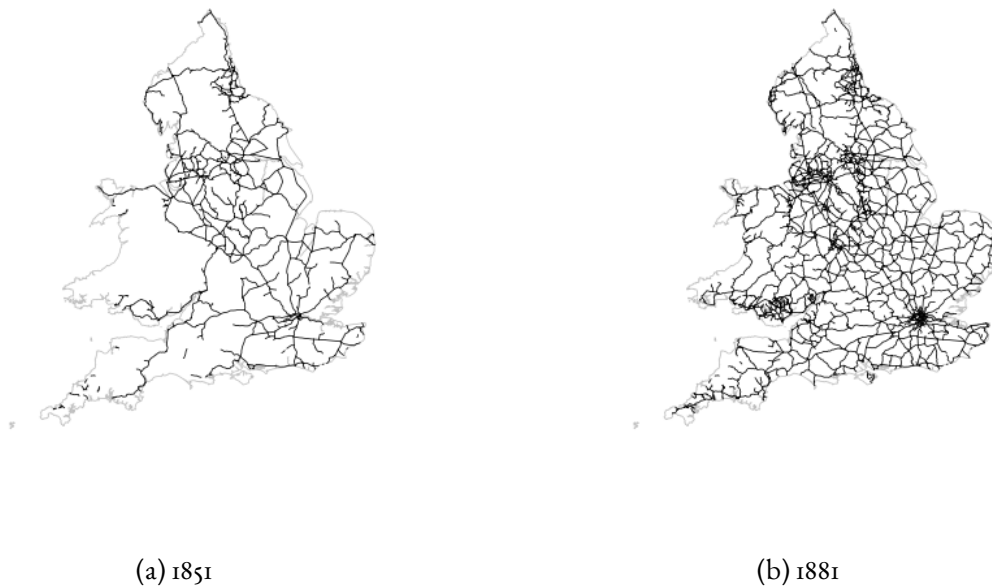
⁴The Railway Mania was brought to an end when the government announced closure for depositing schemes. The Railway Grouping of 1923 terminated much of the inter-company competition that characterized the 1825-1914 period, as the Big Four companies was created.

⁵https://www.mtholyoke.edu/courses/rschwartz/ind_rev/rs/denault.htm

couraged 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.

Figure 1: Railway Network, 1851-1881



2.2 Intergenerational Mobility

2.2.1 Linking Individuals Across Censuses

We combine several datasets to link intergenerational mobility and access to the railroad network. 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 (beginning of April). The only economic outcome available in our data is self-reported occupation. There are over 400 occupations such as physician, cook, stable keeper, cabinet maker or farmer.⁶

To link individuals across censuses, we rely on 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 changed their surname due to marriage, we focus on men between the ages of 40 and 52. We link men across two consecutive censuses. Records were only compared in the linking process if they had an exact match on parish of birth. The age and full first and last names variables, on the other hand, were permitted to have some variation. Age was allowed to be up to two years higher or lower than would be expected. 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. This procedure allows us to link individuals classified as “sons” in the 1851 (1881) census to the men in 1881 (1911) census.⁷ As the censuses record households together along with the household structure, we identify the sons and/or their fathers of these linked men (see Appendix 8.2.1 for further details). We link 652,192 father-sons pairs in 1851-1881 and 1,227,324 in 1881-1911. This represents approximately 42-49% of the population. As a point of comparison, match rates in other studies are between 7-42% (see Table 11).⁸ 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 fathers to be between 20 and 65 years old. This restriction

⁶There are 1.4 million occupational strings in the 1881 census.

⁷The details of the matching procedure and the representativeness are described in the Appendix 8.2.1.

⁸The 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) and Long and Ferrie (2018) also match men English and Welsh census data from 1851 to 1911 and achieving 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.

guarantees that we are looking at men during their working years.⁹

2.2.2 Intergenerational Occupational Mobility

The linked data allows us to observe an individual's occupation as an adult (between 40 and 52 years old) and his father's occupation during his youth (between 10 and 22 years old). The 30 year interval allows the occupation information for both generations to be observed at a similar age. We measure intergenerational mobility through occupations as is commonly done in historical setting (Boberg-Fazlic 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 of the principal difficulties in the study of earnings mobility is obtaining a true measure of permanent income in the face of frequent transitory income shocks. As shocks often occur without job changes, occupation should be less affected by such disturbances. Further, occupation and social class capture 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 labor, place of work, and so on. Finally, it allows us to compare our results to previous findings.

Occupations are ranked based on HISCAM (version 1.3.1 GB) which gives a score to each occupation based on their position in the social stratification structure (Lambert, Zijdemann, Van Leeuwen, Maas and Prandy, 2013). Scores assigned to occupations represent the relative positions of those employed in each occupation, as revealed by the social interaction patterns.¹⁰ 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's occupation has higher (lower) HISCAM score than the occupation of his father and

⁹In 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.

¹⁰Some 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, majors...). 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).

the difference between father and son HISCAM scores is higher than one standard deviation of the son's distribution.¹¹ To properly measure changes in occupation between father-son, the HISCAM ranking we use is constant over time. However, it is important to note that 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 between father-son, we would need a ranking in which an occupation's position changes over time. We perform our analysis using such a ranking as a robustness check (see Figure 15).

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. 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.¹² The second classification recodes historical occupational codes into a class scheme, a status scale and a division into economic sectors. These include professional, industrial, commercial, domestic, and agriculture (Woollard, 1998).¹³ We refer to these as Woollard occupations. The third classification is the HISCLASS which categorized occupational into 12 groups (Van Leeuwen and Maas, 2011). The classification is based on rank associated to each group increases with the skill level starting from unskilled farm workers (rank 1) to higher professional (rank 12). We aggregate these groups into four larger groups: farmers, higher managers, skilled workers, and lower skilled workers.¹⁴ Finally, we create a binary variable for the literacy of an individual based on Armstrong (1972)'s measure of the literacy requirement by occupation.

¹¹More formally, let H^s be the HISCAM score of the son, with standard deviation $\sigma^s = \sqrt{Var(H^s)}$. Then 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 15).

¹²"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.

¹³"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.

¹⁴"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.

2.2.3 Geolocating Individuals

To geo-reference individuals, we use the Great Britain addresses (GB1900) (Southall, Aucott, Fleet, Pert and Stoner, 2017), the digitised parish and county boundaries provided by the UK Data Service (Satchell, Kitson, Newton, Shaw-Taylor and Wrigley, 2017). We first 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. We then measure individual access to the railroad network based on the shortest straight line distance between their address and the nearest train station. 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.

3 Patterns of Intergenerational Mobility

Table 1 presents descriptive statistics of our linked 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 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 is strong inequality between individuals at that time with very few individuals at the top of the distribution. We also see a clear ranking with professional occupations having on average a higher HISCAM score than other occupations and agriculture occupations having on average the lowest HISCAM scores. Nevertheless, occupational ranking and the occupation categories contain different information. Within each occupation categories, we observe a range of occupational ranking. For instance, within professional occupations monks have high occupational ranking while soldiers have low occupational ranking. It is therefore important to examine both the occupational ranking and occupation categories when exploring intergenerational mobility. Figure 8 in the Appendix present similar density plots for HISCO and HISCLASS. Figure 11 in the Appendix also shows the relationship between distance to the nearest train station and the father's HISCAM. We see that fathers living within 5km to the nearest train station belong to the full range of HISCAM ranking, while those living further away tend to work in

Table 1: Descriptive Table

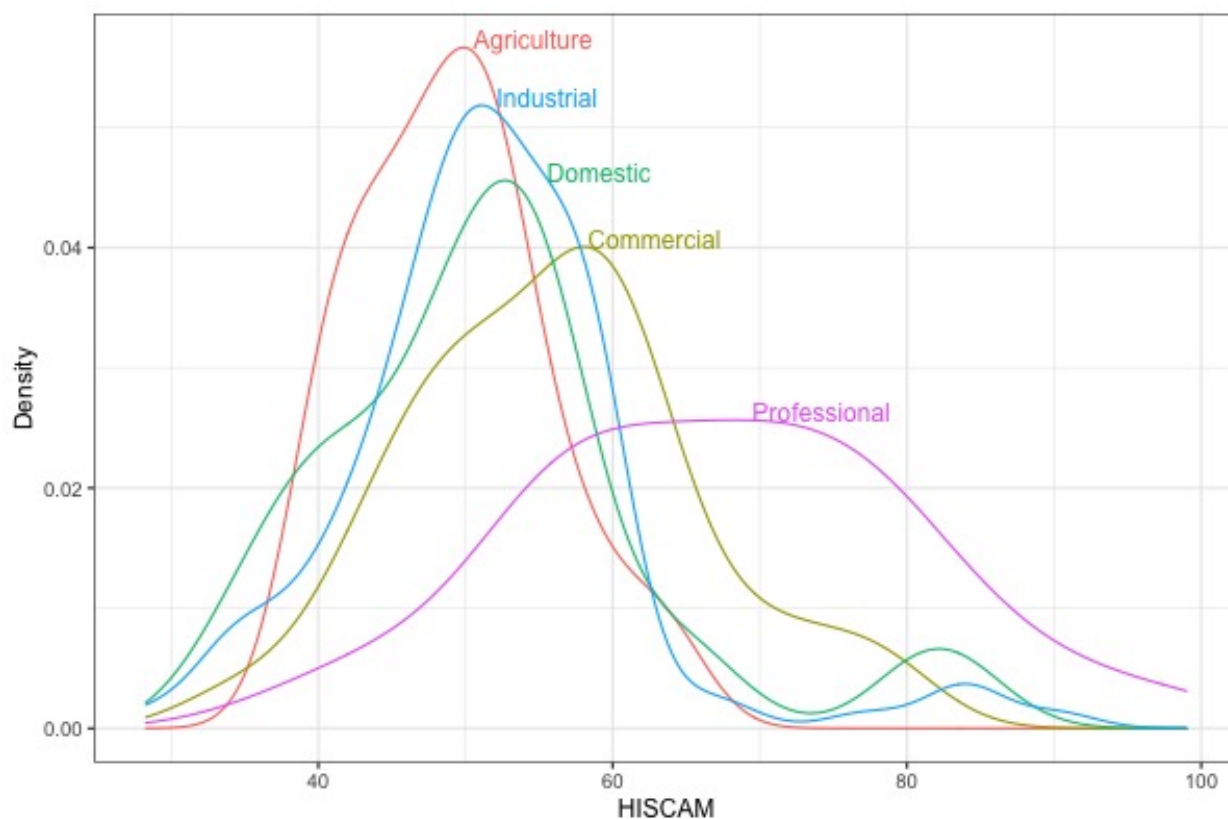
	Mean	St. Dev.	Min.	Max.
SONS				
Age	44.64	3.51	40	52
Foreign-born	0.02	0.15	0	1
Urban resident	0.39	0.49	0	1
Literate	0.32	0.47	0	1
HISCAM occupation rank	50.10	10.11	28.28	99
$HISCO^{son} \neq HISCO^{father}$	0.80	0.40	0	1
$ HISCAM^{son} - HISCAM^{father} $	8.03	8.39	0	71
Upward mobility	0.18	0.39	0	1
Downward mobility	0.15	0.36	0	1
County mover	0.32	0.47	0	1
Dist. moved county mover	98.56	96.98	0.02	633.24
FATHERS				
Age	46.67	7.61	20	65
Foreign-born	0.05	0.22	0	1
Urban resident	0.39	0.49	0	1
Household size	6.76	2.15	0	45
Number of sons	4.63	2.09	0	17
Number of servants	0.17	0.65	0	54
HISCAM occupation rank	49.41	9.12	28.28	99
Literate	0.31	0.46	0	1
Dist. to nearest train station (in km)	3.25	5.41	0.01	106.57
COUNTY				
Number of father-son pairs	17,622.60	20,678.72	498	110,755
Area (km ²)	617.392	1,344.112	0.001	7,134.007
Population	172,660.70	245,294.20	6,633	1,448,853
Avg. HISCAM	49.60	1.91	44.00	53.37
Avg. dist. to train station (in km)	5.62	5.39	0.81	26.32

Note: The sample consists of 955,291 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.

lower ranked occupations. The colour highlights the intergenerational mobility patterns. Upward and downward mobilities occurs at all distances. Naturally upward mobility is correlated to the father's HIS-

CAM and only possible if the father belongs to the lower HISCAM ranking.

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



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

The correlation between father-son HISCAM occupational rank is 0.28. Table 2 provides a cross-classification of son and father's 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.¹⁵

A new feature of our dataset is the ability to geographically locating individuals. Figure 3 shows the connectivity by county. Most individuals lived within 5 to 10km away from the nearest train station. Residents of Wales and Cornwall (West and South West) 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 (Northwest and Midlands) 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 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 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

¹⁵Long (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.

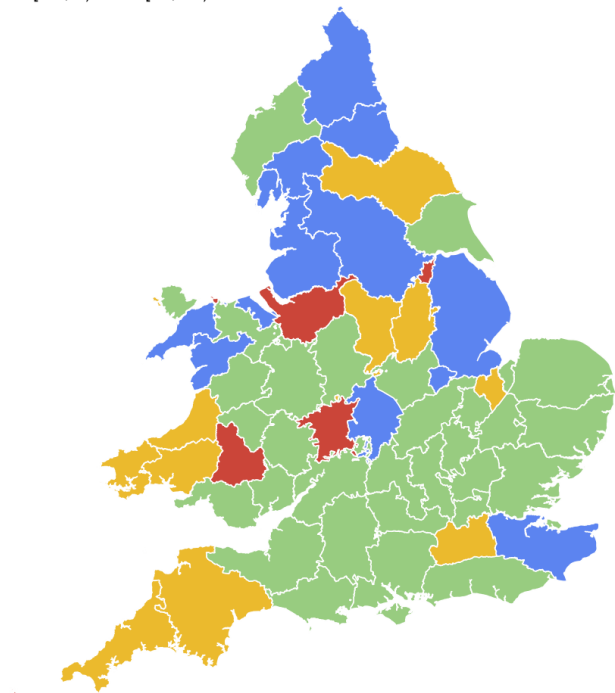
Table 2: Mobility Matrix for Connected Sons

Son	Connected Father				Total
	Manager	Skilled Workers	Lower Skilled	Farmers	
Manager	0.402 {37,385}	0.161 {30,935}	0.147 {47,097}	0.122 {17,065}	132,482
Skilled Workers	0.192 {17,858}	0.415 {79,630}	0.189 {60,692}	0.127 {17,709}	175,889
Lower Skilled	0.348 {32,341}	0.379 {72,772}	0.609 {195,106}	0.358 {50,054}	350,273
Farmers	0.057 {5,304}	0.044 {8,521}	0.055 {17,565}	0.393 {54,974}	86,364
Total	92,888	191,858	320,460	139,802	745,008
Son	Non-connected Father				Total
	Manager	Skilled Workers	Lower Skilled	Farmers	
Manager	0.336 {3,553}	0.120 {3,733}	0.117 {4,228}	0.099 {8,001}	19,515
Skilled Workers	0.193 {2,035}	0.483 {15,007}	0.178 {6,460}	0.104 {8,403}	31,905
Lower Skilled	0.290 {3,066}	0.282 {8,762}	0.534 {19,310}	0.283 {22,867}	54,005
Farmers	0.181 {1,914}	0.114 {3,539}	0.171 {6,196}	0.513 {41,447}	53,096
Total	10,568	31,041	36,194	80,718	158,521

Note: 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.

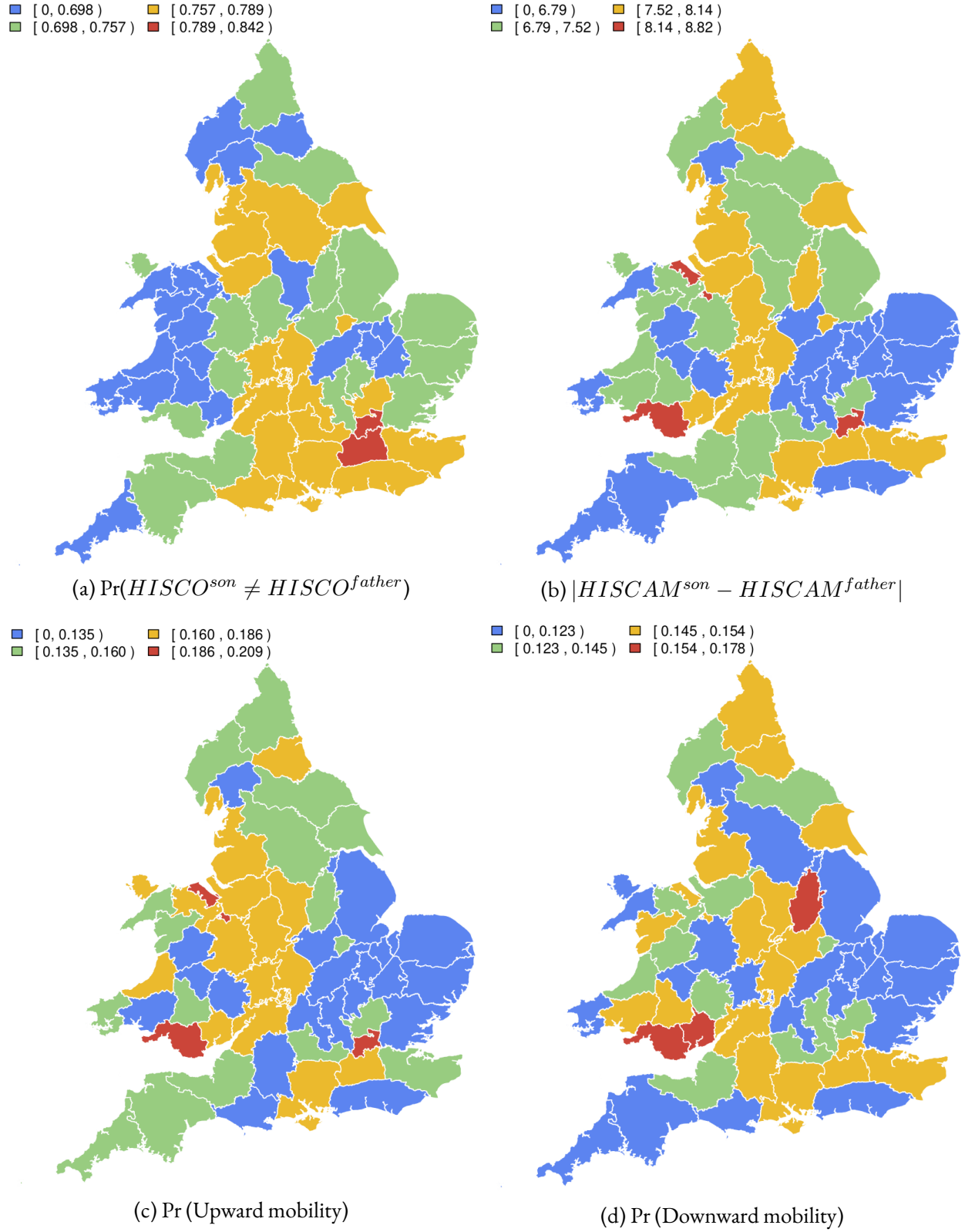
many wealthy estate owners.

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 $t - 1$



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:

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

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)¹⁶, (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, $Proximity_{i,c,t-1}$, based on 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 based on an indicator variable of whether there is 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 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 become a head of household yet. In robustness checks, we also 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.

Finally, we also 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

¹⁶Given the level of detail of the HISCAM classification, we include the additional restriction of being larger than one standard deviation. This makes it possible for the son to stay in the same broad occupation as his father. We provide robustness checks looking at different thresholds.

including wealth, land suitability and local industries. Consequently, for two sons growing up in the same county and the same year, the parameter α_1 captures the effect of growing up one standard deviation closer to the nearest train station on intergenerational mobility. The error term $\epsilon_{i,c,t-1}$ could be correlated across areas that were connected to a similar part of the network. We therefore cluster standard errors at the level of the parish of residence measured at time $t - 1$. We estimate this equation by OLS.

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, within a county the distance to the nearest train station is unlikely to be random. We know from anecdotal evidence that bills proposing new lines and stations to Parliament were often rejected. For instance, the Great Western Railway presented a bill in 1837 connecting Oxford to the main London to Bristol lines. This was to be the first of no less than three bills before the line was finally granted permission six years later. Land owners, the Canal Company and the University objected to the plan. Therefore the location of train stations and consequently the proximity to it may bias our results. 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.

To address the endogenous proximity to the train station, we use an instrumental variable approach or “inconsequential place IV approach” (Alvarez et al., 2017; Banerjee et al., 2020; Duflo and Pande, 2007;

Faber, 2014; Lipscomb et al., 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. This least-cost predicted railroad network allows us to isolate the portion of the variation that is attributable to exogenous cost considerations. Moreover, the hypothetical network addresses the endogeneity in the location of the train station within a county as it is uncorrelated to local characteristics such as land value and town centres. 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.

We proceed in three steps. In the first step, we identify major towns in 1801.¹⁷ 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 network edges connecting larger towns.¹⁸ The remaining least cost path network connecting all pairs of major towns is the late projected line. The resulting dynamic least cost path (DLCP) network presented in Figure 5 is a function of the location of the 1801 population and geographic features of England and Wales.

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 to the DLCP network. Therefore, the first stage equation is defined as:

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

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

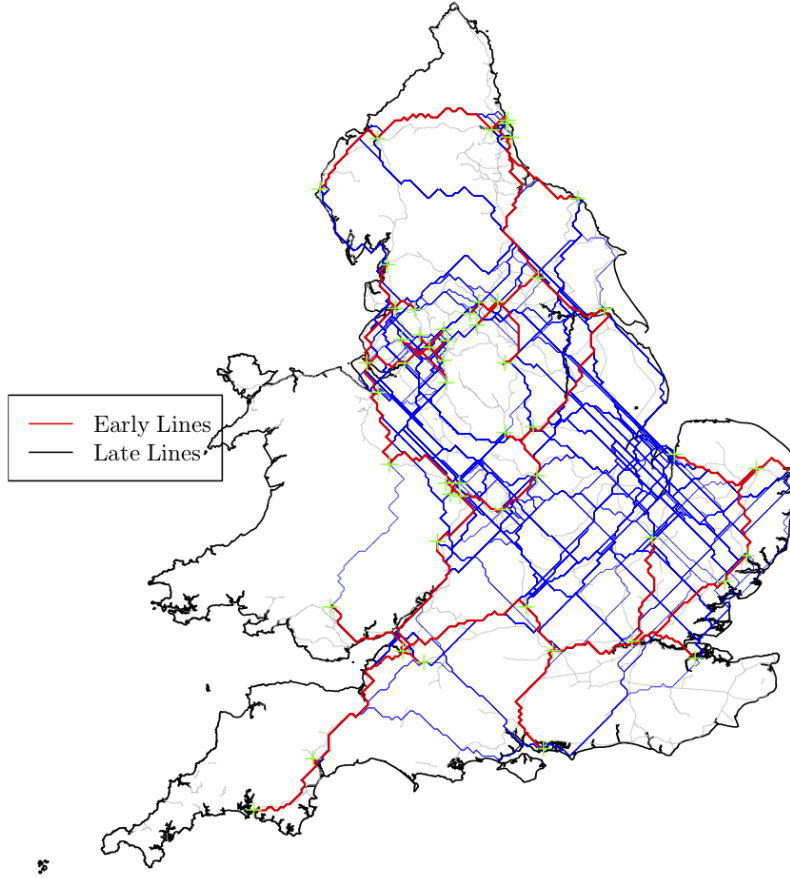
¹⁷ Within all towns in 1801, we consider a town a major town if it belongs to the top 10% of the population distribution. This represents towns with at least 9,172 inhabitants in 1801. There is a total of 53 towns in the top 10%.

¹⁸ For an edge connecting towns p and q the cost of implementing it is:

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

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

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 line and blue lines are the “late” 1881 lines.

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.

First, we used 1801 major towns as nodes in our hypothetical network. This means that any town 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.¹⁹ These variables proxy for the importance of the place of residence as a traffic junction and a likely stop for the railroad.

Second, the dynamic least cost path is likely to follow pre-existing historical travel routes between cities. Any effects we attribute to the road network could 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).

Third, the place of residence may be correlated with underlying intergenerational mobility patterns. To the extent that the initial wealth of a family determines their place of residence and experience different intergenerational mobility patterns, 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 all our specifications, we estimate regressions before and after including the set of additional controls. Therefore, the baseline identifying assumption is that individuals residing along the dynamic least cost path railroad network affects experience changes in economic outcomes from one generation to the next only through the railroad connection, conditional on control variables, 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.

¹⁹The 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 .

Table 3: First stage regressions

	(1)	(2)	(3)
Dep. var.:	Proximity _{<i>i,c,t-1</i>}		
Proximity to DLCP network _{<i>i,c,t-1</i>}	0.016*** (0.001)	0.014*** (0.001)	0.014*** (0.001)
Obs.		969,243	
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 _{<i>t-1</i>} 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 $t - 1$ in parenthesis. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

5.2 Main Result

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 which includes all control variables in addition to county and census year fixed effects, 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 in a different occupation than their father. They were also 5 percentage points more likely to be upward mobile.²⁰ 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).

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 railroads and train stations. OLS regressions underestimate the gains from connectivity. This is consistent with the railroad locations targeting areas with limited upward mobility. These could be struggling areas or areas where poorer families were more likely to live. 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.

²⁰The 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.

Table 4: The effect of railroad connection on intergenerational mobility

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS			IV		
$HISCO^{son} \neq HISCO^{father}$	0.020*** (0.001)	0.011*** (0.001)	0.010*** (0.001)	0.089*** (0.010)	0.065*** (0.010)	0.062*** (0.010)
$ HISCAM^{son} - HISCAM^{father} $	0.370*** (0.023)	0.238*** (0.022)	0.228*** (0.022)	1.249*** (0.132)	1.102*** (0.147)	1.056*** (0.144)
Upward Mobility	0.014*** (0.001)	0.009*** (0.001)	0.008*** (0.001)	0.051*** (0.006)	0.050*** (0.006)	0.049*** (0.006)
Downward Mobility	0.005*** (0.001)	0.003*** (0.001)	0.003*** (0.001)	0.012*** (0.004)	0.007 (0.005)	0.006 (0.005)
Obs.	969,243					
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 coefficient represents the coefficient of the standardised 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 $t - 1$ in parenthesis. *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.

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

where $\Pr(Occ_{i,c,t}^{son} = k | Occ_{i,c,t-1}^{father} = m)$ is the probability that a son with a father in occupation category m would work in occupational category k . The control variables are the same as in the previous specification.

Table 5 presents the results from Equation 3 for HISCO occupations. 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 14 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. Given the importance of farming, we also examine the subsample of sons of non-farms and farmers separately in Table 21 in the Appendix. In both subsamples, intergenerational mobility patterns are the same as in the baseline although sons are farmers have a higher intergenerational mobility.

It has been shown that railroad lead to higher school enrolment and increase skill premia in the local

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 (0.028)	0.114*** (0.042)	0.073*** (0.019)	0.009 (0.010)	-0.010 (0.009)	0.004 (0.003)	0.009*** (0.002)	0.012*** (0.003)
Managerial	-0.015 (0.014)	0.034 (0.037)	-0.002 (0.018)	0.001 (0.007)	0.025*** (0.008)	-0.002 (0.003)	0.001 (0.002)	-0.0001 (0.001)
Clerical	0.041** (0.018)	0.016 (0.036)	-0.002 (0.031)	0.026*** (0.010)	-0.001 (0.013)	0.005* (0.003)	0.003 (0.003)	0.007** (0.003)
Sales	0.056** (0.023)	-0.015 (0.047)	-0.021 (0.031)	-0.063*** (0.024)	0.001 (0.015)	0.001 (0.005)	-0.005 (0.005)	0.004 (0.005)
Services	-0.034 (0.021)	-0.072* (0.038)	-0.026 (0.024)	0.009 (0.008)	0.025 (0.018)	0.002 (0.004)	0.012** (0.005)	0.007* (0.004)
Farm	-0.025 (0.020)	-0.230*** (0.078)	-0.040** (0.018)	-0.054*** (0.014)	-0.063*** (0.017)	-0.111*** (0.017)	-0.051*** (0.008)	-0.135*** (0.015)
Labourer	-0.023 (0.037)	0.153** (0.075)	0.019 (0.044)	0.074*** (0.024)	0.022 (0.025)	0.100*** (0.015)	0.031*** (0.010)	0.105*** (0.012)
Obs.	22,269	15,285	21,358	75,318	34,226	226,466	574,321	969,243

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 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 and household characteristics including the number of servants, household size and whether the father is born outside England and Wales. Current parish clustered standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01

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 census do not have information on education level. Instead we measure skill and literacy based on occupations. In Table 13 in the Appendix, we see that better connected sons are 8 percentage points more likely to be literate and 6 percentage points more likely to work in a high-skilled occupation.²¹ When conditioning on fathers being illiterate or unskilled, we see that sons are upward mobile in terms of these skills. Table 15 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 from lower skilled workers to skilled workers and

²¹Skill level is defined as an indicator variable equal to one if the HISCLASS occupational ranking is "manager", "skilled worker" or "lower skilled".

vice versa induced by the railroad. Better connected sons of lower skilled workers moving up to skilled worker while sons of 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

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

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

Table 6 presents the causal 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.²²

²²We further disaggregate the distribution of HISCAM by quantile in Figure 13 in the Appendix. We see that 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.

Table 6: Distributional Consequences

	(1)	(2)	(3)	(4)
	Father			
	Bottom	Middle	Top	Any
Bottom	0.011 (0.022)	0.044*** (0.010)	-0.012 (0.010)	0.035** (0.014)
Middle	-0.061*** (0.019)	-0.095*** (0.012)	-0.137*** (0.017)	-0.116*** (0.014)
Top	0.050*** (0.012)	0.051*** (0.009)	0.149*** (0.018)	0.081*** (0.011)
Obs.	235,909	446,906	286,428	969,243

Notes: Each coefficient represents the coefficient of $\text{Proximity}_{i,c,t-1}$ 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 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 and household characteristics including the number of servants, household size and whether the father is born outside England and Wales. Current parish clustered standard errors in parenthesis. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

5.5 Threats to the Instrumental Variable and 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 9.2.

First, the estimator would be biased if residents and firms move over time along the same spatial lines of the DLCP network. For instance, fathers who have high ambition for their sons may decide to live

not only in connected places but also closer to predicted train stations. Table 19 explores the possibility of self-selection by estimating 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). Our results remain similar in magnitude.

Second, we show our baseline results are similar than results obtained when using alternative measures of proximity to the railroad network and intergenerational mobility. In Figure 14, 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. All coefficients on proximity are similar in magnitude. In Figure 15, 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 addition, we remove occupations specific to the railroad such as train conductor or controller which would mechanically increase with the expansion of the railroad network in Table 20. Finally, farming occupations may have a particular role in the transition during the Industrial Revolution. In Table 21 we split the sample between fathers who are in farming activities and all other activities. Sons of farmers also experienced intergenerational mobility. However they also experienced a significant 3 percentage points increase in downward mobility.

Third, we examine an alternative specification including polynomials for the control variables and parish fixed effect in Figure 16. 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.

Fourth, we explore possible 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 22, we see that the baseline results remain and the effects are similar in magnitude.

Finally, we consider whether our results are driven by a particular census year, county, rural/urban divide, age and/or place of birth. We see that the results are not driven by a particular census year in Table 23. Our results are also robust to excluding one county at a time in Figure 17. We see similar patterns

of intergenerational mobility for sons who grew up in urban and rural areas in Table 24.²³ 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 (Haider and Solon, 2006). We therefore explore whether age at which the father's occupation is measured affects our results in Table 25. Similarly, we look at the age of sons in Table 26 to examine whether the time at which they get access to the railroad affects our results. Sons aged 12-16 years old experience a slightly larger mobility patterns induced by the railroad network than sons aged 17-22 years old. We also look at social mobility by migration status in Table 27. Recent work by Abramitzky, Boustan and Eriksson (2012, 2014) shows that migration status is an important factor for intergenerational mobility patterns. We find that our results are mainly driven by natives.

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:

$$\begin{aligned} \Pr(up|train) &= \Pr(up|stay, train) \times \Pr(stay|train) \\ &\quad + \Pr(up|move, train) \times \Pr(move|train) \end{aligned} \tag{5}$$

where *up* stands for upward 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 their childhood county and those who

²³Individuals living in an urban area is defined as those who grew up within 2.5km of a 1801 town.

have stayed respectively. Taking the total derivative with respect to train, we obtain:

$$\begin{aligned}
\Delta \Pr(up|train) = & \underbrace{\Delta \Pr(up|stay, train)}_{\text{change in upward mobility induced by the train}} \\
& + \underbrace{[\Delta \Pr(up|move, train) - \Delta \Pr(up|stay, train)]}_{\text{change in the benefit from moving induced by the train}} \times \underbrace{\Pr(move|train)}_{\text{baseline geographic mobility}} \\
& + \underbrace{[\Pr(up|move, train) - \Pr(up|stay, train)]}_{\text{baseline benefit from moving}} \times \underbrace{\Delta \Pr(move|train)}_{\text{ease of geographic mobility from the train}}
\end{aligned} \tag{6}$$

The train therefore affect the change in upward mobility through three channels: (1) the change in upward mobility induced by being closer to the train station, (2) the change in the relative benefit from moving, and (3) the ease of geographic 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.

6.1 Intergenerational mobility by spatial mobility pattern

Guided by the Equation 6 above, we first estimate the ease of geographical mobility from the access to the train. Railroads have been shown to facilitate the spatial mobility for connected individuals by reducing the travel time and cost (Morten and Oliveira, 2014). 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 possibility of spatial mobility, we first look at the probability of sons' moving away from their childhood parish and further examine whether geographic mobility was targeted towards connected parishes. We use the the following equation

$$\Pr(move_{i,c,t}) = \beta_1 Proximity_{i,c,t-1} + \beta_2 X_{i,c,t-1} + \gamma_t + \rho_c + \epsilon_{i,c,t} \tag{7}$$

where $move_{i,c,t}$ 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 7 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 up 5km closer to the train station. We find that better connected sons have a higher likelihood of moving regardless of the original family class (bottom, middle or top). 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. In Table 18 in the Appendix, we provide descriptive statistics on the characteristics of sons distinguishing between those who decided to move and those who stayed. We do not see significant differences in these characteristics which include father's occupational ranking and whether he was literate.

Table 7: Railroad and Geographic Mobility

Dep. var.	$\Pr(\text{move}_{i,c,t})$			
	(1)	(2)	(3)	(4)
Sample	All	Bottom ^{father}	Middle ^{father}	Top ^{father}
$\text{Proximity}_{i,c,t-1}$	0.091*** (0.021)	0.108*** (0.026)	0.074*** (0.021)	0.094*** (0.023)
Avg. dep. var.	0.32	0.30	0.31	0.35
Obs.	969243	235909	446906	286428

Notes: The dependent variable is an indicator variable which switches to one if the son moved away from his childhood county (column 1), conditional on being the son of a father in the bottom 25% (column 2), middle 26-74% (column 3) and top 25% (column 4) of the HISCAM distribution. All regressions include childhood parish_{t-1} fixed effects and year fixed effects. Additional controls include the historical importance of town and historical travels routes consisting of the distance to the closest 1801 town, its populations and the population in the surrounding areas weighted by distance, the distance to the closest Roman road and port, household characteristics including the number of servants, household size and whether the father is born outside England and Wales. Standard errors clustered at the parish in year $t - 1$ in parenthesis. *p<0.1; **p<0.05; ***p<0.01.

We estimate the relative benefits of having better access to the railroad network by movers and stayers, taken the decision to move as given. In Table 8 we examine the baseline specification conditioning on sons who are movers or stayers. Being better connected to the railroad network increased the probability

of being upward mobile by 4 percentage points for both movers and stayers. However, it resulted in different effects on the probability of working in a different occupation as their father and being downward mobile. The railroad decoupled the tie in occupation between fathers and sons for stayers, but not for movers. Stayers are also more likely to experience an increase in their probability of being downward mobile whereas movers experienced the opposite. This would be consistent with moving only if the expected returns are higher from moving than staying.

Table 8: Social Mobility Pattern by Geographic Mobility

	(1)	(2)	(3)	(4)	(5)	(6)
	Movers			Stayers		
$HISCO^{son} \neq HISCO^{father}$	0.018*** (0.006)	0.010 (0.007)	0.008 (0.007)	0.111*** (0.014)	0.081*** (0.015)	0.077*** (0.014)
$ HISCAM^{son} - HISCAM^{father} $	0.379*** (0.120)	0.455*** (0.143)	0.409*** (0.141)	1.467*** (0.167)	1.230*** (0.184)	1.183*** (0.182)
Upward Mobility	0.040*** (0.006)	0.047*** (0.007)	0.045*** (0.007)	0.049*** (0.006)	0.044*** (0.008)	0.042*** (0.007)
Downward Mobility	-0.007 (0.005)	-0.008 (0.006)	-0.009 (0.006)	0.019*** (0.005)	0.014** (0.006)	0.013** (0.006)
Obs.	308,910	308,910	308,910	660,333	660,333	660,333
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 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 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 includes only sons who have moved away from their childhood county of residence (columns 1 to 3) and those who have not moved away (columns 4 to 6). All regressions include childhood county and census year fixed effects. Additional controls include the historical importance of town and historical travels routes consisting of the distance to the closest 1801 town, its 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 (column 3 and 6). Standard errors clustered at the parish in year $t - 1$ in parenthesis. *p<0.1; **p<0.05; ***p<0.01

If we ignore the endogeneity in the decision to move and where to go, and abstract from any “macro” level

effects²⁴ we can decompose the effect of the railroad into the three channels at work. Based on a back-of-the-envelope calculation, the change in the local opportunities induced by the train account for roughly 91% of the upward mobility, while the ease of geographic mobility accounts for 7% and the change in the relative benefit from moving accounts for 2%.²⁵ Upward mobility was mainly driven by changes in the local labour market opportunities. When estimating the effect of the railroad network, those who have stayed constitute the majority of the effect.

6.2 Effects by specialisation of cities

The effects of access to the railroad might have been different depending on the specialisation of a city. We categorise cities by their historical main activities based on the British Parliamentary Papers (see Appendix 9.3). Cities fall in one of five categories: administrative, manufacturing, port, resort or mining. We refer to cities in any of these categories as “places of opportunities”. All other cities are categorised as “other”. Tables 16 and 17 in the Appendix present the results disaggregating by the categories of cities. Sons who grew up closer to the train station in administrative, manufacturing, and port towns and other cities are more likely to be upward mobile. Only in port and other cities do we see a significant break in the occupation of fathers and sons. Better connected sons in port and other cities are also significantly more likely to move away. The propensity to move induced by the train is not significantly changed for sons in all other cities. When distinguishing between movers and stayers, we uncover different patterns. The propensity of sons to work in a different occupation as their fathers is different for the sample of sons who decided to move and those who decided to stay. Conditional on moving away, better connected sons in administrative, manufacturing, and port towns are more likely to work in the same occupation as their father. In contrast, the same sons who decided to stay in their county of birth were more likely to work in a different as their father.

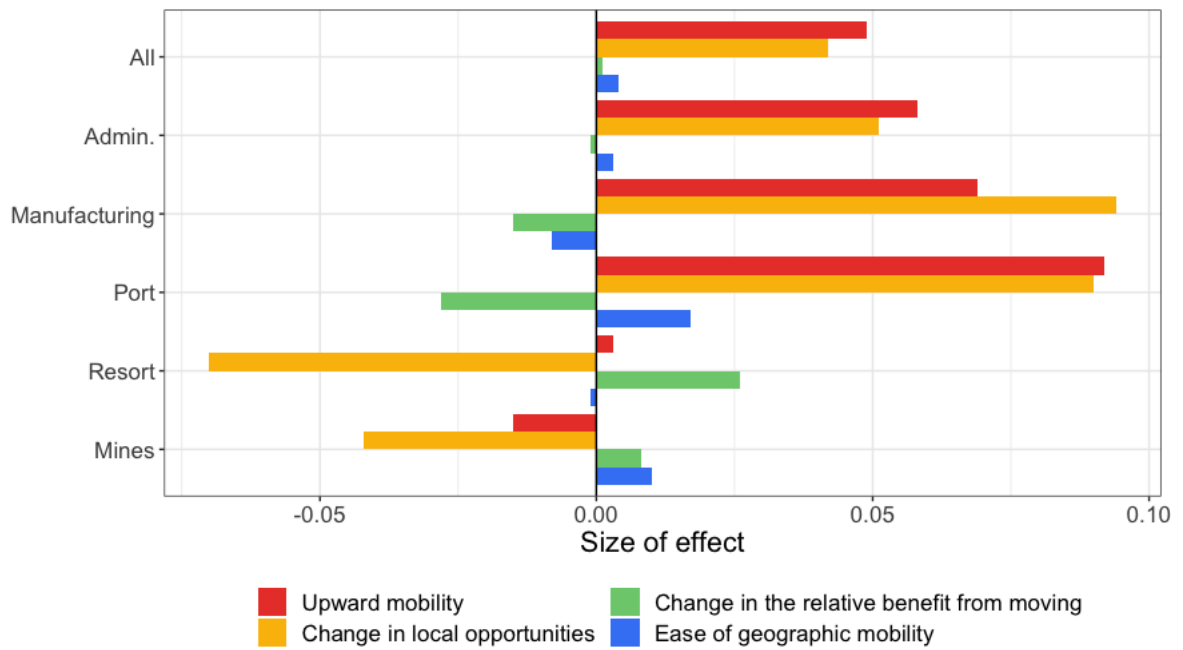
Figure 6 presents the decomposition of upward mobility by the different channels distinguishing between the main historical activity of cities. We uncover stark differences by the main historical activity

²⁴Estimating equation 6 is not straightforward. First, the decision to stay or move is endogenous and correlated with the railroad access. Highly skilled individuals or individuals from certain background are more likely to move. This will lead to a selection-bias. Second, if one decides to move, the decision on where to move is also endogenous and likely correlated with the skill set of an individual and complementarities with the labour opportunity at destination. Therefore the relative benefit of moving should take into account the specific place of origin and destination. Finally, 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 generates positive and negative spillovers to other areas.

²⁵0.049 is the change in upward mobility induced by the train (Table 4). The change in local opportunity is 0.042 (coefficients in Table 8). The ease of geographic mobility is $(0.21 - 0.17) \times 0.19$ (coefficient in Table 7 and the average upward mobile conditional on movers and stayers) and the change in the relative benefit from moving is $(0.045 - 0.042) \times 0.32$ (coefficients in Tables 8 and 7) and the average probability of moving).

of cities. Better connected sons who grew up in administrative, manufacturing and port towns experienced higher upward mobility than the average connected sons. In contrast, sons in resort towns were less likely to be upward mobile if they grew up closer to the train station. In all cases, the change in the local opportunities induced by the train is the main driving force of the upward mobility.

Figure 6: Heterogeneity by main historical activity of city



7 Conclusion

Can transport infrastructure alter social mobility and break the link between parents and their children's economic status? 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 1911 and geographically locate them down to the street level. This level of disaggregation allows us to measure access to 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.

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. This 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 local labour market induced by the railroad. We also find stark differences depending on the specialisation of a town. Our results 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 helping families move to higher opportunity areas.

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., 2019. Intergenerational Mobility in Africa. Technical Report. National Bureau of Economic Research.
- 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.
- 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 .
- Artadi, E.V., Sala-i Martin, X., 2003. The economic tragedy of the 20th Century: Growth in Africa. Technical Report. National Bureau of Economic Research.
- Aschauer, D.A., 1989. Is public expenditure productive? *Journal of Monetary Economics* 23, 177–200.
- 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.
- Bennett, R.J., 2012. Urban population database, 1801-1911. [data collection] Robson, B., University of Manchester, Department of Geography. SN: 7154, <http://doi.org/10.5255/UKDA-SN-7154-1> .
- Black, S.E., Devereux, P.J., 2011. Recent developments in intergenerational mobility. *Handbook of Labor Economics* 4, 1487–1541. doi:10.1016/S0169-7218(11)02414-2.
- 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.

- Campante, F., Yanagizawa-Drott, D., 2018. Long-range growth: Economic development in the global network of air links. *The Quarterly Journal of Economics* 133, 1395–1458.
- 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 ii: County-level estimates. *The Quarterly Journal of Economics* 133, 1163–1228.
- Chetty, R., Hendren, N., Katz, L.F., 2016. The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review* 106, 855–902.
- 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. doi:10.1093/qje/qju022 .Advance, arXiv:arXiv:1011.1669v3.
- Clark, G., Cummins, N., 2015. Intergenerational wealth mobility in england, 1858–2012: surnames and social mobility. *The Economic Journal* 125, 61–85.
- 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.

- 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).
- Haider, S., Solon, G., 2006. Life-cycle variation in the association between current and lifetime earnings. *American Economic Review* 96, 1308–1320.
- Herzog, I., 2013. The potential and limits of optimal path analysis. *Computational approaches to archaeological spaces* 186.
- Jackson, R.V., 1987. The structure of pay in nineteenth-century Britain. *The Economic History Review* 40, 561–570.
- 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., Zijdemann, R.L., Van Leeuwen, M.H., Maas, I., Prandy, K., 2013. The construction of HIS-CAM: 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.
- Lindert, P.H., 2000. Three centuries of inequality in Britain and America. *Handbook of income distribution* 1, 167–216.
- 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.

- 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.
- Long, J., Ferrie, J., 2018. Grandfathers matter (ed): Occupational mobility across three generations in the US and Britain, 1850-1910. *The Economic Journal* .
- 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.
- Pascali, L., 2017. The wind of change: Maritime technology, trade, and economic development. *American Economic Review* 107, 2821–54.
- Perez, S., 2017. Railroads and the rural to urban transition: Evidence from 19th Century Argentina. unpublished , 1–103.

- 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.
- Redding, S.J., Turner, M.A., 2015. Transportation costs and the spatial organization of economic activity, in: *Handbook of regional and urban economics*. Elsevier. volume 5, pp. 1339–1398.
- Rostow, W.W., 1959. The stages of economic growth. *The Economic History Review* 12, 1–16.
- Sanderson, E., Windmeijer, F., 2016. A weak instrument F-test in linear IV models with multiple endogenous variables. *Journal of Econometrics* 190, 212–221.
- Satchell, A., Kitson, P., Newton, G., Shaw-Taylor, L., Wrigley, E., 2017. 1851 England and Wales census parishes, townships and places. [data collection]. UK Data Service. SN: 852816, <http://doi.org/10.5255/UKDA-SN-852232>.
- 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.
- Solon, G., 2002. Cross-country differences in intergenerational earnings mobility. *Journal of Economic Perspectives* 16, 59–66.
- 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.

8 Appendix

8.1 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 both national and universal scales. The national scale has been computed using data from Great Britain and is constant for the 1800-1938 period.²⁶ 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

²⁶More information about the computation of the scales can be found in Lambert et al. (2013). And in the website <http://www.camsis.stir.ac.uk/hiscam/>.

Dunn, 2019)

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 7: 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.

8.2 Data Construction

8.2.1 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, standardise 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 9: 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
Linkage rate combined		25	
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
Linkage rate combined		29	

Note: Step 1 is the number of unique individuals with at least one potential match, Step 2 is the number of unique individuals with unique matches, Step 3 is the number of unique individuals with unique matches after dropping second best match with sufficient age difference, and Step 4 is the number of unique individuals after doing the within cleaning and merging matches from step 2 and step 3.

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 10: Linkage Statistics for men between 40-52 years of age

	1851-1881	1881-1911
Pop	652,192	1,227,324
Linkage rate	42	49
Avg. age distance	0.54	0.41
Avg. surname JW-distance	0.01	0.01
Avg. name JW-distance	0.00	0.00

8.2.2 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.

Table 11: Comparison with other studies using linked data

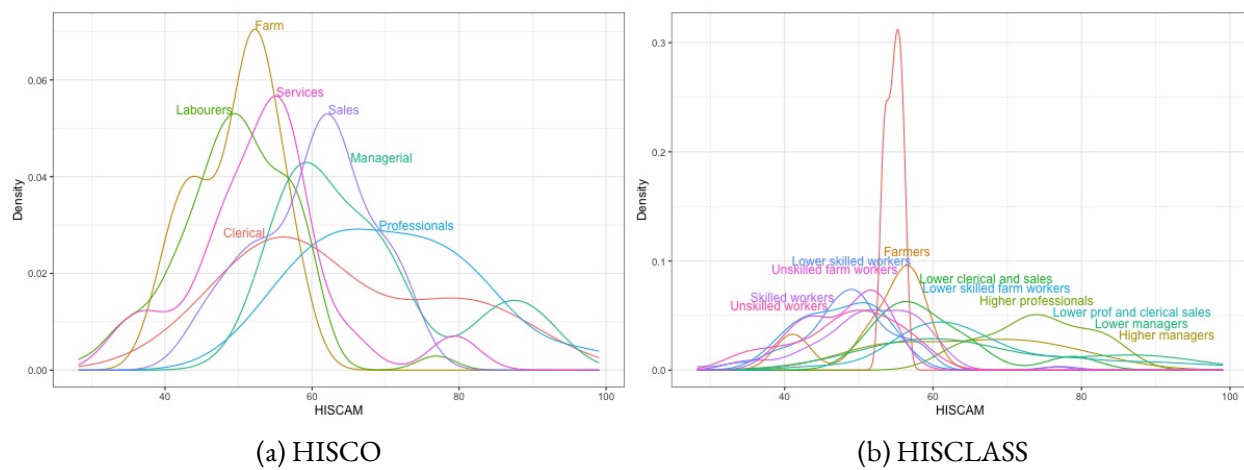
Article	Source	Match rate	Number linked
Costas Fernandez et al. (2020)	1881 England and Wales Census to 1911 England and Wales Census (Full, Men 40-52)	49%	1,227,324
Costas Fernandez et al. (2020)	1851 England and Wales Census to 1881 England and Wales Census (Full, Men 40-52)	42%	652,192
Guerra and Mohnen (2020)	1851 London (Full census) to 1881 London (Full, Men 43-49)	33%	263,264
Milner (2019)	1851 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 1911 England and 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)

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 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.

8.2.3 Occupation classification

Figure 8: HISCAM density by occupation classification



8.2.4 Geolocating individuals

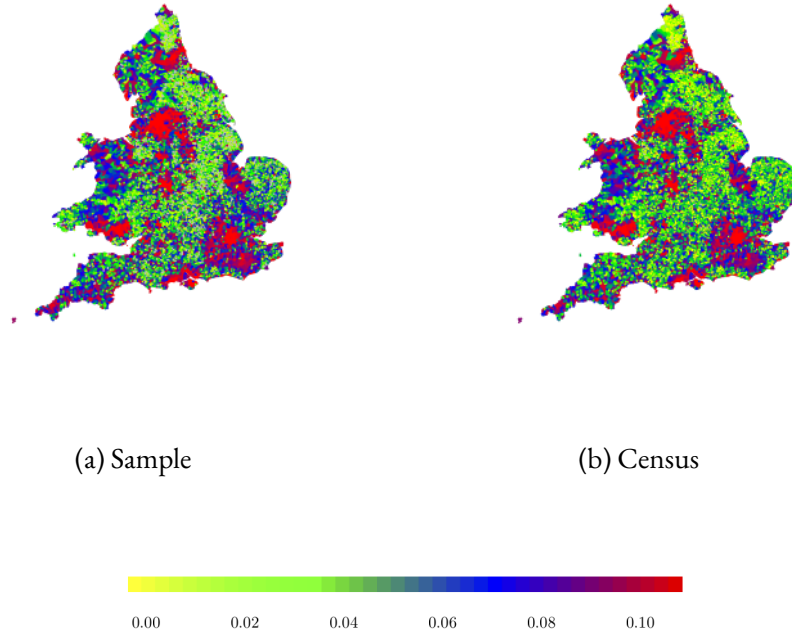
We geo-locate individuals at two levels: the parish and the address. We geolocate addresses by matching the address provided in the I-CEM data for each individual with the address database put together by the GB1900 team Southall et al. (2017).²⁷ To improve the quality of the match we split the UK into 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.²⁸

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>.)

²⁷The GB1900 final raw gazetteer data dump can be accessed from <http://www.visionofbritain.org.uk/>

²⁸A further refinement that one could apply is to also condition on a minimum distance between first and second best match candidate.

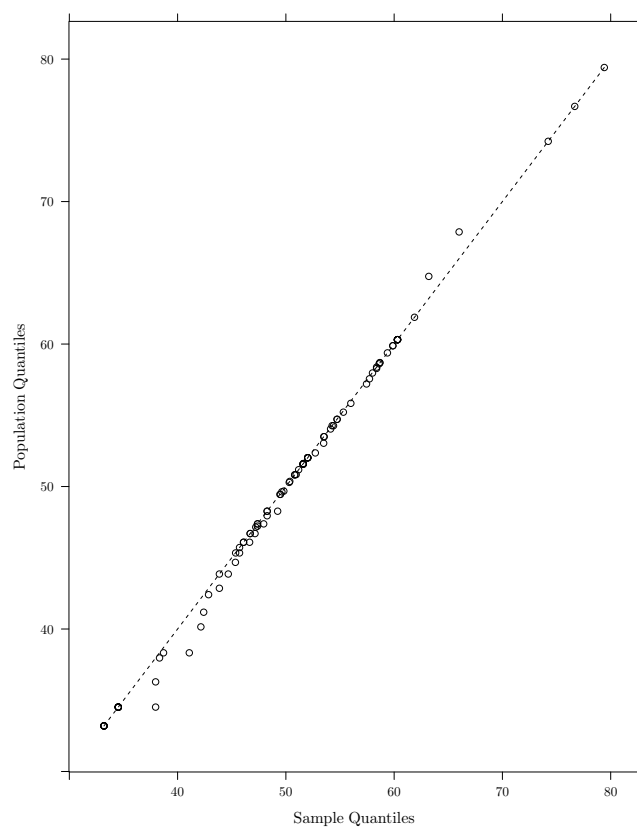
Figure 9: Size of the sample and population



Note: Figure 9a displays 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 9b displays parish populations for males 40-52 pooling data from 1881-1911. Sizes are represented as percentage of the total. The legend covers the 1 to 99 percentile with a knot at every percentile. Parishes that could not be uniquely matched across Censuses are in grey.

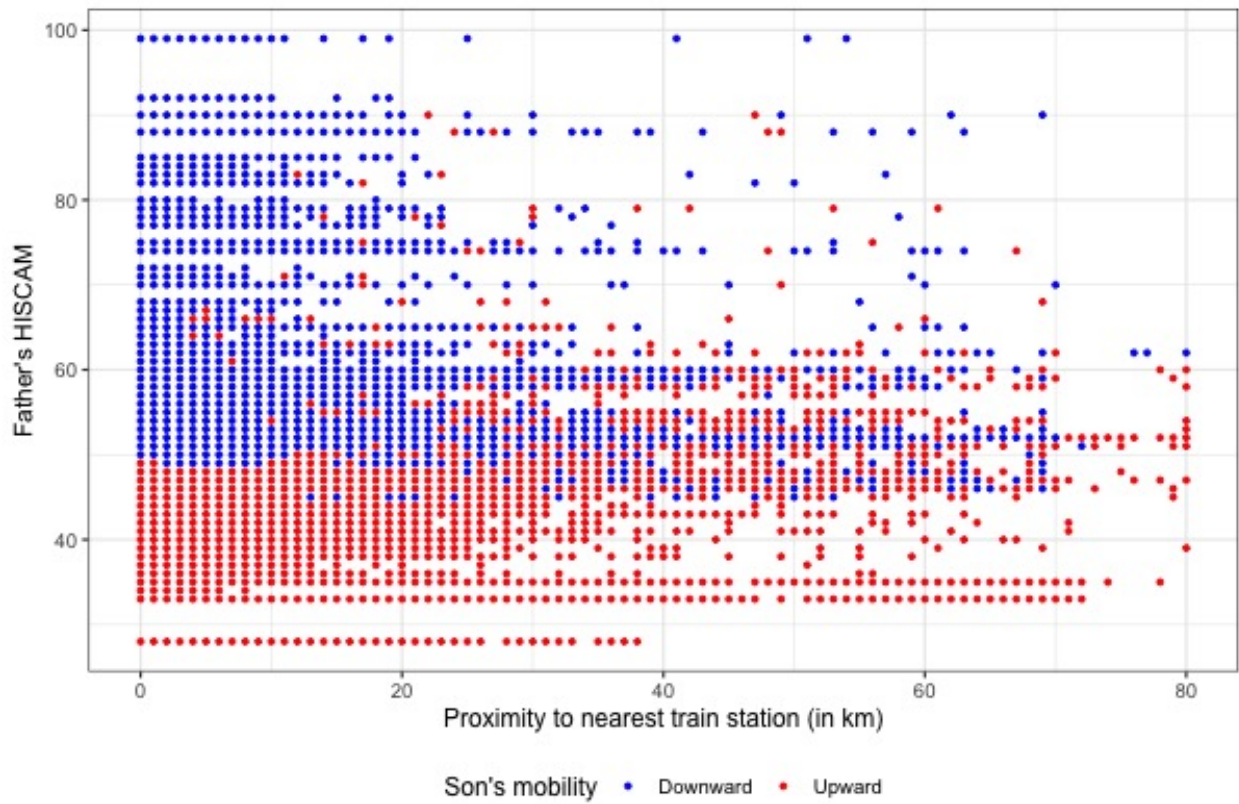
9 Additional Descriptives

Figure 10: HISCAM Distribution
Census vs Matched Sample



Note: From left to right dots are the 1 to 99 percentiles in our estimation sample against the same quantiles in the Census for males 40-52 with a valid occupation code that is matched to HISCAM. Both distributions are constructed by pooling the 1881 and 1911 data.

Figure 11: Proximity to train station and 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 12: Share of non-unique matches

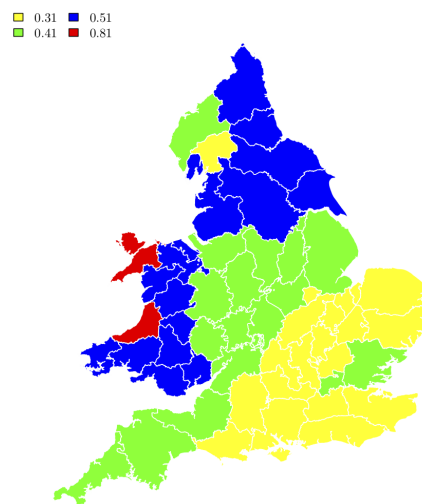


Table 12: Role of railroad network access on linking individuals

Dep. var.: Share of linked individuals among the parish population aged 10-22				
	(1)	(2)	(3)	(4)
	Distance to hypothetical rail line		Distance to nearest train station	
Proximity	0.002 (0.001)	-0.001 (0.001)	-0.002** (0.001)	-0.002** (0.001)
Obs.	24,450			
Year FE	Yes	Yes	Yes	Yes
County _{t-1} FE	Yes	Yes	Yes	Yes
Historical importance of town	No	Yes	No	Yes
Historical travel routes	No	Yes	No	Yes

Notes:

9.1 Additional Results

Table 13: Skill level based on occupation

	(1)	(2)	(3)
	Father illiterate	Father unskilled	All
Son literate	0.070*** (0.010)	0.104*** (0.017)	0.082*** (0.015)
Son skilled	0.075*** (0.013)	0.052*** (0.011)	0.060*** (0.013)
Obs.	661,407	594,975	966,733

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 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). 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 $t - 1$ in parenthesis. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 14: 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 (0.028)	0.011** (0.004)	0.007 (0.013)	0.015 (0.022)	-0.001 (0.004)	0.012*** (0.005)
Industrial	0.044 (0.031)	0.011 (0.011)	0.044 (0.032)	0.073* (0.040)	0.092*** (0.015)	0.099*** (0.014)
Commercial	0.025 (0.024)	0.023*** (0.007)	0.047 (0.029)	0.057* (0.030)	0.013 (0.008)	0.025*** (0.007)
Domestic	-0.001 (0.009)	-0.006** (0.003)	-0.018* (0.010)	-0.040 (0.030)	0.020*** (0.005)	-0.003 (0.003)
Agriculture	-0.048*** (0.015)	-0.047*** (0.007)	-0.088*** (0.023)	-0.111*** (0.030)	-0.125*** (0.018)	-0.139*** (0.014)
Obs.	30,023	613,244	90,235	16,270	216,839	969,243

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 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, controls for 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 and household characteristics including the number of servants, household size and whether the father is born outside England and Wales. Current parish clustered standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01

Table 15: The effect of rail connection by HISCLASS occupations

	(1)	(2)	(3)	(4)	(5)
	Father				
	Manager	Skilled Worker	Lower Skilled	Farmer	Any
Manager	0.026 (0.023)	0.002 (0.007)	0.003 (0.008)	0.009 (0.007)	0.012* (0.007)
Skilled Workers	0.025* (0.015)	-0.026** (0.013)	0.028** (0.014)	0.027*** (0.008)	0.043*** (0.010)
Lower Skilled	0.031* (0.018)	0.062*** (0.014)	0.046** (0.020)	0.072*** (0.012)	0.084*** (0.013)
Farmers	-0.082*** (0.018)	-0.039*** (0.007)	-0.077*** (0.012)	-0.107*** (0.017)	-0.139*** (0.015)
Obs.	103,456	222,899	356,654	220,520	928,122

Notes: Each coefficient represents the coefficient of the standardised 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, controls for 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 and household characteristics including the number of servants, household size and whether the father is born outside England and Wales. Current parish clustered standard errors in parenthesis. * $p < 0.1$; ** $p < 0.05$; ***

Figure 13: Effect of railroad connection by HISCAM quantile

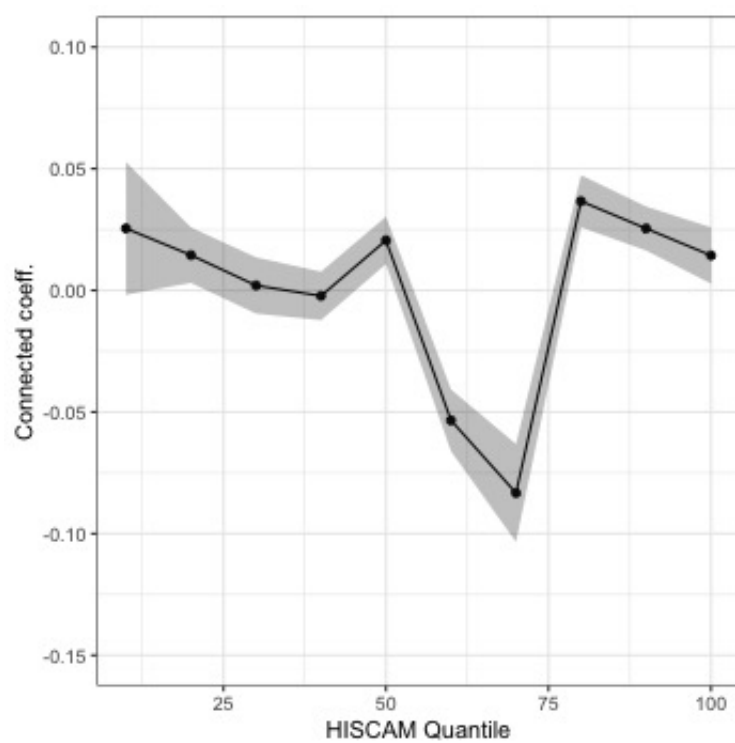


Table 16: Main historical activities of cities

	(1)	(2)	(3)	(4)	(5)	(6)
	Admin.	Manufacturing	Port	Resort	Mining	Other
$Pr(move)$	0.095 (0.097)	-0.021 (0.064)	0.569*** (0.113)	-0.031 (0.092)	0.120 (0.145)	0.071*** (0.023)
$HISCO^{son} \neq HISCO^{father}$	0.007 (0.017)	0.092* (0.052)	0.084*** (0.023)	0.068 (0.047)	-0.082 (0.135)	0.089*** (0.023)
$ HISCAM^{son} - HISCAM^{father} $	0.448 (0.344)	1.307** (0.574)	1.524*** (0.441)	0.876 (0.671)	0.898 (1.767)	1.292*** (0.314)
Upward Mobility	0.058*** (0.016)	0.069** (0.029)	0.092*** (0.025)	0.003 (0.042)	-0.015 (0.077)	0.055*** (0.014)
Downward Mobility	-0.011 (0.011)	0.005 (0.015)	-0.007 (0.017)	-0.007 (0.027)	-0.030 (0.087)	0.013 (0.009)
Obs.	135,593	213,453	135,517	12,242	25,136	579,740

Notes: Each coefficient represents the coefficient of the standardised 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 childhood parish's main history activity: administrative (column 1), manufacturing (column 2), port (column 3), resort (column 4), mine (column 5), or other (column 6). All regressions include fixed effects for census year and childhood county_{*t-1*}, controls for 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, household characteristics including the number of servants, household size and whether the father is born outside England and Wales. Standard errors clustered at the parish in year $t - 1$ in parenthesis. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 17: Social mobility by historical characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
	Mover					
	Admin.	Manufacture	Port	Resort	Mines	Other
$HISCO^{son} \neq HISCO^{father}$	-0.021*	-0.012	-0.041	0.024	-0.156	0.021*
	(0.013)	(0.024)	(0.032)	(0.058)	(0.148)	(0.012)
$ HISCAM^{son} - HISCAM^{father} $	0.301	0.259	-0.702	1.243	-0.920	0.275
	(0.325)	(0.464)	(0.852)	(1.148)	(2.531)	(0.280)
Upward Mobility	0.036	0.033*	0.023	0.143***	-0.011	0.053***
	(0.022)	(0.019)	(0.035)	(0.037)	(0.115)	(0.015)
Downward Mobility	0.014	0.019	-0.041	-0.108***	-0.011	-0.023*
	(0.016)	(0.019)	(0.029)	(0.040)	(0.062)	(0.012)
Obs.	57,730	50,240	56,497	4,224	6,458	186,919
	(7)	(8)	(9)	(10)	(11)	(12)
	Stayer					
	Admin.	Manufacture	Port	Resort	Mines	Other
$HISCO^{son} \neq HISCO^{father}$	0.002	0.173*	0.092***	0.101	-0.067	0.123***
	(0.023)	(0.094)	(0.033)	(0.073)	(0.150)	(0.034)
$ HISCAM^{son} - HISCAM^{father} $	-0.023	2.132**	2.009***	0.710	1.351	1.755***
	(0.393)	(1.033)	(0.497)	(0.802)	(2.092)	(0.434)
Upward Mobility	0.051***	0.094**	0.090***	-0.070	-0.042	0.051***
	(0.018)	(0.044)	(0.027)	(0.056)	(0.098)	(0.016)
Downward Mobility	-0.039**	-0.001	0.016	0.052	-0.025	0.032**
	(0.015)	(0.021)	(0.024)	(0.043)	(0.106)	(0.013)
Obs.	77,863	163,213	79,020	8,018	18,678	392,821

Notes: Each coefficient represents the coefficient of the standardised 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. Subsamples are based on the main history activity of the childhood parish: administrative (columns 1 and 7), manufacturing (columns 2 and 8), port (columns 3 and 9), resort (columns 4 and 10), mine (columns 5 and 11), or other (columns 6 and 12). All regressions include fixed effects for census year and childhood county_{t-1}, controls for 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, household characteristics including the number of servants, household size and whether the father is born outside England and Wales. Standard errors clustered at the parish in year $t - 1$ in parenthesis. *p<0.1; **p<0.05; ***p<0.01.

Table 18: Descriptives by movers and stayers

	Son mover				Son stayer			
	Mean	St. Dev.	Min	Max	Mean	St. Dev.	Min	Max
Age father	46.42	7.61	20	65	46.79	7.60	20	65
Foreigner born father	0.06	0.24	0	1	0.05	0.21	0	1
Urban area	0.39	0.49	0	1	0.39	0.49	0	1
# of sons	4.58	2.09	0	15	4.66	2.09	0	17
# of servants	0.21	0.74	0	50	0.15	0.61	0	54
HISCAM father	50.21	9.46	28.28	99.00	49.04	8.94	28.28	99.00
Literate father	0.31	0.46	0.00	1.00	0.32	0.46	0.00	1.00

9.2 Robustness Checks

Stayers vs. movers 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 19, we see that intergenerational mobility patterns can be seen for both stayers and movers.

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 14 shows that our baseline results are conservative.

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 20. We see that the our results are robust.

Farming activities In Table 21 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.

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 15. We first use the HISCAM occupation ranking that takes into account changes in the ranking of occupations over time. In our baseline 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.

Alternative specification Figure 16 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.

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 22 we see that our results are robust to potential measurement error.

Year Table 23 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.

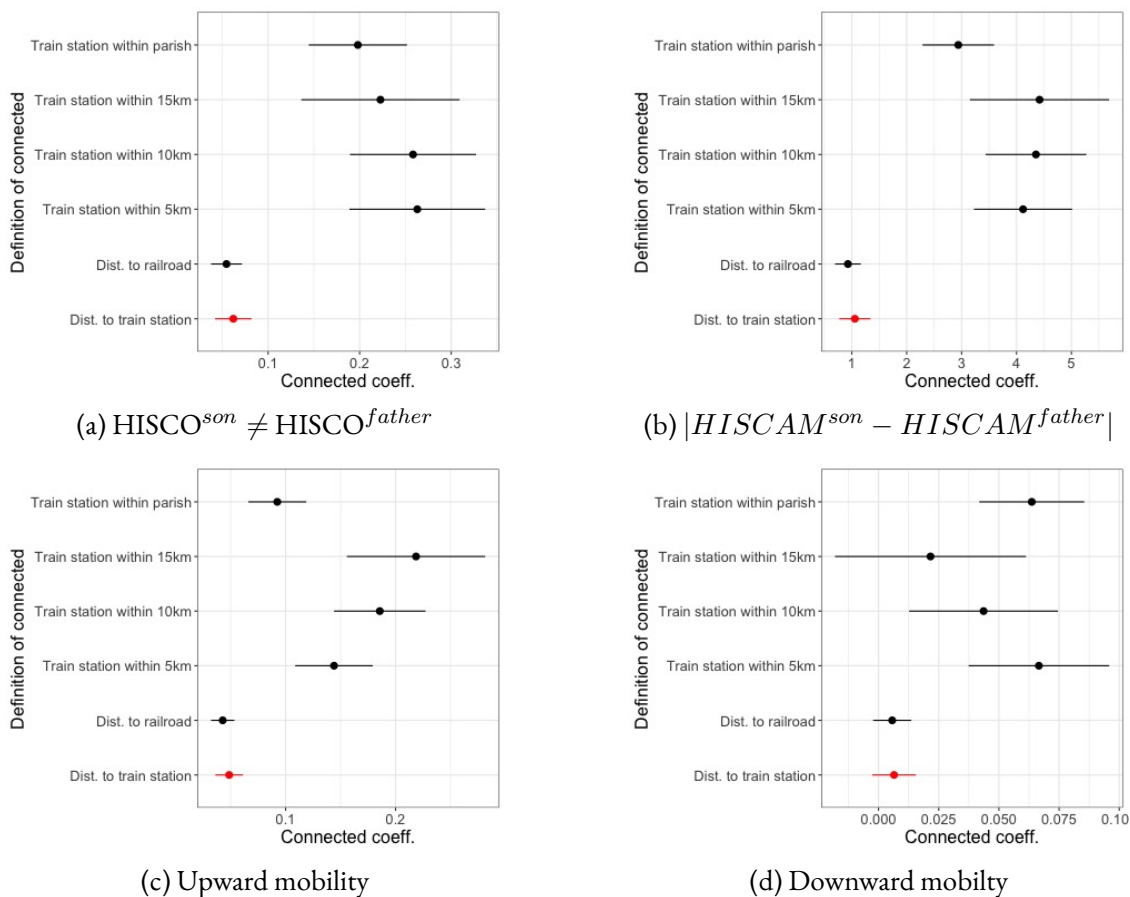
Excluding one region at a time We show that the results are robust to excluding one county at a time. Figure 17 shows us that our findings are not confined to a single region.

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 24. We do not observe large differences between the two groups.

Age 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 25 we do not see differences in the effects of having access to the railroad network by the age of the father. Similarly in Table 26 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 In Table 27 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 see that better connected foreigners experienced large upward mobility.

Figure 14: Alternative definition of proximity



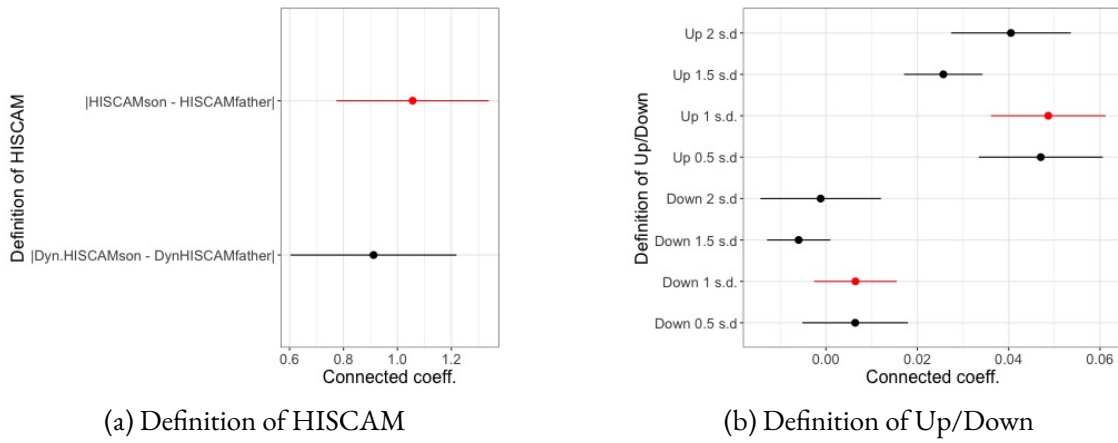
Notes: Each coefficient represents the coefficient of the standardised Proximity $_{i,c,t-1}$, instrumented by the proximity to the DLCP network. Proximity is defined as the distance to the nearest train station (red dot), indicator if the parish has a train station (first black dot), indicator if the 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 $_{t-1}$, controls include the historical importance of town and historical travels routes, and controls for household characteristics. Standard errors clustered at the parish in year $t - 1$ in parenthesis. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 19: Father as Stayer or Mover

	(1)	(2)	(3)	(4)	(5)	(6)
	Stayers			Movers		
$HISCO^{son} \neq HISCO^{father}$	0.036*** (0.007)	0.020** (0.008)	0.016** (0.008)	0.119*** (0.015)	0.091*** (0.015)	0.088*** (0.015)
$ HISCAM^{son} - HISCAM^{father} $	0.743*** (0.117)	0.673*** (0.143)	0.617*** (0.141)	1.455*** (0.180)	1.264*** (0.193)	1.229*** (0.190)
Upward Mobility	0.049*** (0.006)	0.052*** (0.007)	0.050*** (0.007)	0.046*** (0.007)	0.041*** (0.008)	0.039*** (0.008)
Downward Mobility	-0.001 (0.004)	-0.005 (0.006)	-0.006 (0.006)	0.021*** (0.005)	0.016** (0.006)	0.016** (0.006)
Obs.	405,743	405,743	405,743	563,500	563,500	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 coefficient 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. 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), household characteristics including 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 $t - 1$ in parenthesis. *p<0.1; **p<0.05; ***p<0.01.

Figure 15: Alternative definition of occupation ranking



Notes: Each coefficient 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 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 $_{t-1}$, controls for the historical importance of town and historical travels routes and the controls for household characteristics. Standard errors clustered at the parish in year $t - 1$ in parenthesis. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 20: Main results without rail related occupations

	(1)	(2)	(3)
$HISCO^{son} \neq HISCO^{father}$	0.091*** (0.010)	0.067*** (0.011)	0.064*** (0.010)
$ HISCAM^{son} - HISCAM^{father} $	1.259*** (0.135)	1.132*** (0.152)	1.085*** (0.149)
Upward Mobility	0.051*** (0.006)	0.051*** (0.007)	0.049*** (0.006)
Downward Mobility	0.013*** (0.004)	0.009* (0.005)	0.008* (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

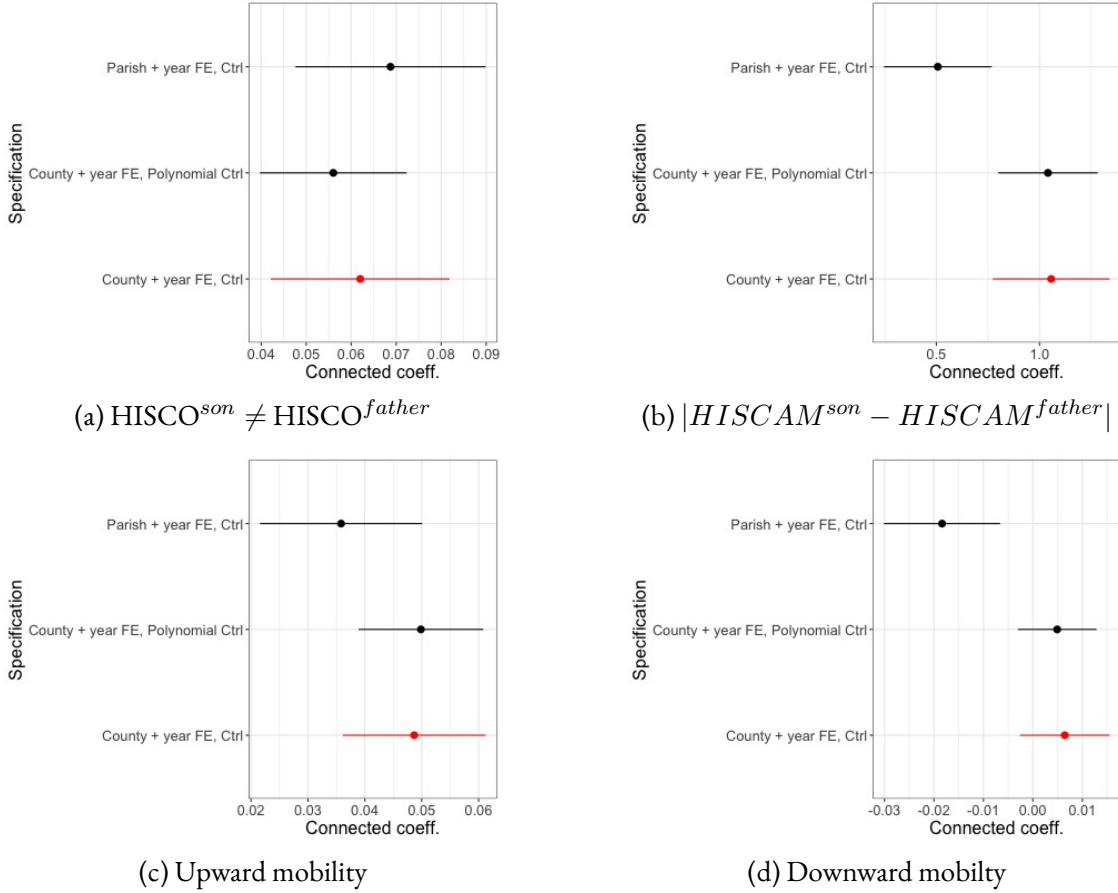
Notes: Each coefficient represents the coefficient of the standardised Proximity_{i,c,t-1}, instrumented by the proximity to the DLCP network. The dependent variables are an indicator variable which switches to one if the son does not work in the same occupation category as his father (row 1), the absolute 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 $t - 1$ in parenthesis. *p<0.1; **p<0.05; ***p<0.01.

Table 21: Farming occupations

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

Notes: Each coefficient 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 (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 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, household characteristics including the number of servants, household size and whether the father is born outside England and Wales. Standard errors clustered at the parish in year $t - 1$ in parenthesis. *p<0.1; **p<0.05; ***p<0.01.

Figure 16: Alternative specification



Notes: Each dot 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 (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 in parenthesis. *p<0.1; **p<0.05; ***p<0.01.

Table 22: Measurement error in geolocation

	(1)	(2)	(3)	(4)	(5)	(6)
	Address			Parish centroid		
$HISCO^{son} \neq HISCO^{father}$	0.089*** (0.010)	0.065*** (0.010)	0.062*** (0.010)	0.091*** (0.011)	0.074*** (0.011)	0.071*** (0.011)
$ HISCAM^{son} - HISCAM^{father} $	1.249*** (0.132)	1.102*** (0.147)	1.056*** (0.144)	1.264*** (0.136)	1.138*** (0.138)	1.085*** (0.136)
Upward Mobility	0.051*** (0.006)	0.050*** (0.006)	0.049*** (0.006)	0.052*** (0.006)	0.049*** (0.006)	0.047*** (0.006)
Downward Mobility	0.012*** (0.004)	0.007 (0.005)	0.006 (0.005)	0.012*** (0.004)	0.010** (0.005)	0.009* (0.005)
Obs.	969,243			969,243		
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 coefficient 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 $t - 1$ in parenthesis. *p<0.1; **p<0.05; ***p<0.01.

Table 23: Subsample by year

	(1)	(2)	(3)	(4)	(5)	(6)
	1851-1881			1881-1911		
$HISCO^{son} \neq HISCO^{father}$	0.064*** (0.007)	0.038*** (0.008)	0.035*** (0.007)	0.115*** (0.023)	0.076** (0.031)	0.070** (0.031)
$ HISCAM^{son} - HISCAM^{father} $	0.948*** (0.105)	0.790*** (0.124)	0.746*** (0.120)	2.128*** (0.328)	2.419*** (0.441)	2.328*** (0.433)
Upward Mobility	0.034*** (0.005)	0.031*** (0.006)	0.029*** (0.006)	0.083*** (0.013)	0.097*** (0.019)	0.094*** (0.019)
Downward Mobility	0.012*** (0.003)	0.010** (0.004)	0.009** (0.004)	0.034*** (0.010)	0.042*** (0.015)	0.040*** (0.015)
Obs.	273,844	273,844	273,844	695,399	695,399	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 coefficient 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 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 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 (column 3 and 6). Standard errors clustered at the parish in year $t - 1$ in parenthesis. *p<0.1; **p<0.05; ***p<0.01

Figure 17: Excluding one county at a time

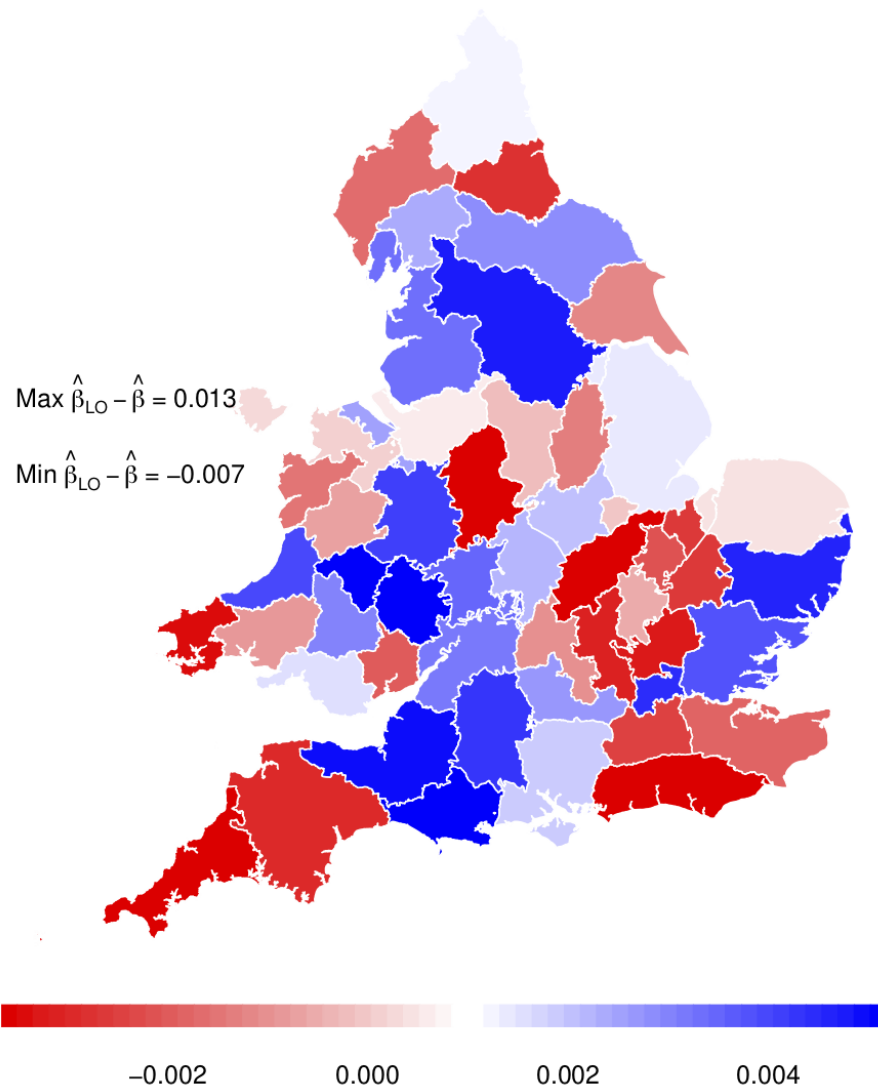


Table 24: Social Mobility Pattern by Urban-Rural

	(1)	(2)	(3)	(4)	(5)	(6)
	Urban			Rural		
$HISCO^{son} \neq HISCO^{father}$	0.101*** (0.030)	0.066*** (0.022)	0.064*** (0.022)	0.067*** (0.009)	0.051*** (0.011)	0.049*** (0.011)
$ HISCAM^{son} - HISCAM^{father} $	1.137*** (0.436)	0.916** (0.378)	0.880** (0.373)	1.145*** (0.141)	1.159*** (0.158)	1.130*** (0.155)
Upward Mobility	0.066*** (0.021)	0.054*** (0.017)	0.053*** (0.017)	0.045*** (0.006)	0.052*** (0.007)	0.051*** (0.007)
Downward Mobility	-0.006 (0.013)	0.003 (0.013)	0.002 (0.013)	0.012*** (0.004)	0.007 (0.005)	0.007 (0.005)
Obs.	380,281	380,281	380,281	588,962	588,962	588,962
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 coefficient 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. 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), household characteristics including 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 $t - 1$ in parenthesis.

*p<0.1; **p<0.05; ***p<0.01.

Table 25: 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*** (0.011)	0.014 (0.039)	0.066*** (0.013)	0.069*** (0.012)	0.081*** (0.012)
$ HISCAM^{son} - HISCAM^{father} $	1.085*** (0.136)	0.540 (0.765)	1.021*** (0.181)	1.110*** (0.158)	1.101*** (0.180)
Upward Mobility	0.047*** (0.006)	0.069* (0.036)	0.062*** (0.009)	0.049*** (0.007)	0.034*** (0.008)
Downward Mobility	0.009* (0.005)	-0.037 (0.037)	-0.004 (0.007)	0.009 (0.006)	0.020*** (0.006)
Obs.	969,243	6,068	221,143	451,323	290,709

Notes: Each coefficient 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 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, household characteristics including the number of servants, household size and whether the father is born outside England and Wales. Standard errors clustered at the parish in year $t - 1$ in parenthesis. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 26: 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*** (0.010)	0.052*** (0.010)	0.044*** (0.013)	0.073*** (0.013)	0.079*** (0.013)
$ HISCAM^{son} - HISCAM^{father} $	1.056*** (0.144)	0.879*** (0.184)	0.941*** (0.207)	1.279*** (0.210)	0.930*** (0.190)
Upward Mobility	0.049*** (0.006)	0.048*** (0.008)	0.058*** (0.010)	0.054*** (0.009)	0.029*** (0.009)
Downward Mobility	0.006 (0.005)	0.003 (0.007)	-0.006 (0.008)	0.010 (0.008)	0.018** (0.008)
Obs.	969,243	230,383	176,214	220,808	234,814

Notes: Each coefficient 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 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, household characteristics including the number of servants, household size and whether the father is born outside England and Wales. Standard errors clustered at the parish in year $t - 1$ in parenthesis. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 27: Social Mobility Pattern : Natives vs. Foreigners

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Natives			1st generation immigrant			2nd generation immigrant		
$HISCO^{son} \neq HISCO^{father}$	0.087*** (0.010)	0.063*** (0.010)	0.061*** (0.010)	-0.028 (0.045)	-0.074 (0.084)	-0.076 (0.085)	0.119*** (0.028)	0.089** (0.036)	0.089** (0.036)
$ HISCAM^{son} - HISCAM^{father} $	1.169*** (0.128)	1.022*** (0.146)	1.002*** (0.144)	0.776 (1.165)	1.017 (1.950)	1.016 (1.956)	1.735*** (0.595)	1.413* (0.761)	1.413* (0.760)
Upward Mobility	0.049*** (0.006)	0.047*** (0.006)	0.046*** (0.006)	0.133* (0.074)	0.234* (0.120)	0.237** (0.120)	0.069*** (0.024)	0.067** (0.033)	0.066** (0.033)
Downward Mobility	0.011*** (0.004)	0.007 (0.005)	0.007 (0.005)	-0.113** (0.048)	-0.161** (0.078)	-0.164** (0.078)	0.010 (0.021)	-0.001 (0.029)	-0.001 (0.029)
Obs.	904,689	904,689	904,689	7,865	7,865	7,865	40,877	40,877	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 coefficient represents the coefficient of 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), household characteristics including 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 $t - 1$ in parenthesis.

*p<0.1; **p<0.05; ***p<0.01.

9.3 Main historical activities of cities

Table 28: Principal towns and cities in England, Wales and Scotland in 1831

Rank town or City	Principal functions	Total population
London, Middlesex	Admin, Port	1,358,200
Manchester	Manufacturing	270,961
Glasgow	Manufacturing, Port	202,426
Liverpool with Toxteth	Manufacturing	189,242
Edinburgh	Admin, Port	162,156
Birmingham	Manufacturing	146,486
Leeds	Manufacturing	123,393
Bristol + Clifton	Admin, Port	103,886
Sheffield	Manufacturing	91,692
Bradford	Manufacturing	76,996
Plymouth	Port	75,534
Norwich	Admin, Manufacturing	61,166
Aberdeen	Port, Manufacturing	58,019
Paisley	Manufacturing	57,466
Nottingham	Manufacturing, Admin	50,680
Portsea + Portsmouth	Port	50,389
Newcastle-upon-Tyne with Gateshead	Port, Mines	47,837
Hull (with Sculcoates)	Port	46,426
Dundee	Manufacturing, Port	45,355
Bolton	Manufacturing	41,195
Brighton	Resort	40,634
Leicester	Manufacturing	39,063

Continued on next page

Table 28 – *Continued from previous page*

Rank town or City	Principal functions	Total population in 1831
Bath	Resort	38,063
Stoke-on-Trent	Manufacturing	37,220
Preston	Manufacturing	36,336
Ashton-under-Lyne	Manufacturing	33,597
Exeter	Admin	32,404
Oldham	Manufacturing	32,381
Sunderland (with Bishop Wearmouth)	Port, Mines	31,522
Huddersfield	Manufacturing	31,041
Eccles	Manufacturing	28,995
Jarrow (with South Shields)	Port, Mines	27,995
Greenock	Manufacturing	27,571
Blackburn	Manufacturing	27,091
Coventry	Manufacturing	27,070
Rochdale	Manufacturing	26,404
Stockport	Manufacturing	25,469
York	Admin	25,359
Tynemouth	Port	24,778
Wolverhampton	Manufacturing	24,732
Greenwich	Port	24,553
Wakefield	Manufacturing	24,538
Derby	Manufacturing	23,607
Macclesfield	Manufacturing	23,129
Dudley	Manufacturing	23,042
Cheltenham	Resort	22,942

Continued on next page

Table 28 – *Continued from previous page*

Rank town or City	Principal functions	Total population in 1831
Lancaster	Admin, Port	22,294
Merthyr	Mines	22,083
Chester	Admin	21,363
Shrewsbury	Admin	21,227
Yarmouth	Port	21,115
Cambridge	Admin	20,917
Kidderminster	Manufacturing	20,865
Sedgley, Staffs	Mines, Manufacturing	20,577
Ipswich	Admin, Port	20,454
Oxford	Admin	20,434
Perth	Admin, Port	20,016
Carlisle	Admin	20,006
Dewsbury	Manufacturing	19,854
Deptford	Port	19,795
Southampton	Port	19,324
Warrington	Manufacturing	19,155
Worcester	Manufacturing, Admin	18,610
Kilmarnock	Manufacturing	18,093
Woolwich	Manufacturing, Port	17,661
Dunfermline	Manufacturing, Admin	17,068
Barnsley (Silkstone)	Mines, Manufacturing	16,561
Chatham	Port, Manufacturing	16,485
Colchester	Admin	16,167
Saddleworth	Manufacturing	15,986

Continued on next page

Table 28 – *Continued from previous page*

Rank town or City	Principal functions	Total population in 1831
Reading	Admin	15,595
Mottram	Manufacturing	15,536
Maidstone	Admin	15,387
Halifax	Manufacturing	15,382
Northampton	Manufacturing, Admin	15,351
West Bromwich	Manufacturing	15,327
Kingswinford	Mines; Manufacturing	15,156
Bury	Manufacturing	15,086
Walsall	Manufacturing	15,066
Tipton (Wolverhampton)	Manufacturing	14,951
Bilston (Wolverhampton)	Manufacturing	14,492
Canterbury	Admin, Resort	14,463
Middleton (Rochdale)	Manufacturing	14,379
Inverness	Admin	14,324
Ormskirk	Manufacturing	14,053
Leyland (Preston)	Manufacturing	13,871
Swansea	Port	13,694
Kings Lynn	Port	13,370
Bedminster (Bristol)	Manufacturing	13,130
Falkirk	Manufacturing	12,743
Burslem (Stoke)	Manufacturing	12,741
Gosport (with Alverstoke)	Port	12,637
Croydon	Admin	12,447
Frome	Manufacturing	12,240

Continued on next page

Table 28 – *Continued from previous page*

Rank town or City	Principal functions	Total population in 1831
Montrose	Port	12,055
Gloucester	Admin, Port	11,933
Dover	Port	11,924
Lincoln	Admin	11,892
Dumfries	Admin	11,606
West Ham	Manufacturing	11,580
Doncaster	Admins, Resort	11,572
Bury St. Edmunds	Admin	11,436
Wrexham	Manufacturing, Mines	11,408
Whitehaven	Port	11,393
Batley	Manufacturing	11,335
Boston	Port	11,240
Taunton	Admin	11,139
Loughborough	Manufacturing	10,969
Trowbridge	Manufacturing	10,863
Bedwelty (S. Wales)	Mines	10,637
Rotherham	Manufacturing, Mines	10,417
Tonbridge	Admin	10,380
Margate	Resort	10,339
Hereford	Admin	10,280
Pontypool (with Trevethon)	Mines	10,280
Bradford-on-Avon (Wilts)	Manufacturing	10,102
Hastings	Resort	10,097
Kirkheaton	Manufacturing	10,020

Continued on next page

Table 28 – *Continued from previous page*

Rank town or City	Principal functions	Total population in 1831
Kendal	Admin, Manufacturing	10,015

Note: British Parliamentary Papers (HC348, 1831) Population:

Comparative Account of the Population of Great Britain in the Years 1801, 1811, 1821, and 1831.