The Making of the Modern Metropolis: Evidence from London*

Stephan Heblich†
University of Bristol

Stephen J. Redding‡
Princeton University, NBER and CEPR

Daniel M. Sturm§
London School of Economics and CEPR

September 7, 2018

Abstract

Modern metropolitan areas involve large concentrations of economic activity and the transport of millions of people each day between their residence and workplace. We use the revolution in transport technology from the invention of steam railways, newly-constructed spatially-disaggregated data for London from 1801-1921, and a quantitative urban model to provide evidence on the role of these commuting flows in supporting such concentrations of economic activity. Steam railways dramatically reduced travel times and permitted the first large-scale separation of workplace and residence. We show that our model is able to account for the observed changes in the organization of economic activity, both qualitatively and quantitatively. In counterfactuals, we find that removing the entire railway network reduces the population and the value of land and buildings in Greater London by 20 percent or more, and brings down commuting into the City of London from more than 370,000 to less than 60,000 workers.

Keywords: Agglomeration, Urbanization, Transportation

JEL Classification: O18, R12, R40

*We are grateful to Bristol University, the London School of Economics, and Princeton University for research support. Heblich also acknowledges support from the Institute for New Economic Thinking (INET) Grant No. INO15-00025. We would like to thank Victor Couture, Jonathan Dingel, Ed Glaeser, Vernon Henderson, Petra Moser, Leah Platt-Boustan, Will Strange, Claudia Steinwender, Jerry White, Christian Wolmar and conference and seminar participants at Berkeley, Canadian Institute for Advanced Research (CIFAR), Center for Economic Policy Research (CEPR), Dartmouth, EIEF Rome, German Economic Association, Harvard, MIT, National Bureau of Economic Research (NBER), University College London (UCL), Urban Economics Association (UEA), Vienna, Yale, Zoom Urban Seminar, and Zurich for helpful comments. We would like to thank David Green for sharing printed copies of the Henry Poole data and T. Wangyal Shawa for his help with the GIS data. We would also like to thank the Cambridge Group for the History of Population and Social Structure, the British Library (BL), the British Library of Political and Economic Science (BLPES) at the London School of Economics, the Guildhall Library, London Metropolitan Archives (LMA), and the Omnibus Society for their help with data. Finally, we are grateful to Charoo Anand, Iain Bamford, Horst Bräunlich, Dennis Egger, Andreas Ferrara, Ben Glaeser and Florian Trouvain for excellent research assistance. The usual disclaimer applies.

†Dept. Economics, Priory Road Complex, Priory Road, Clifton, BS8 1TU. UK. Tel: 44 117 3317910. Email: stephan.heblich@bristol.ac.uk.

‡Dept. Economics and WWS, Julis Romo Rabinowitz Building, Princeton, NJ 08544. Tel: 1 609 258 4016. Email: reddings@princeton.edu.

§Dept. Economics, Houghton Street, London, WC2A 2AE. UK. Tel: 44 20 7955 7522. Email: d.m.sturm@lse.ac.uk.

1 Introduction

Modern metropolitan areas include vast concentrations of economic activity, with Greater London and New York City today accounting for around 8.4 and 8.5 million people respectively. These intense population concentrations involve the transport of millions of people each day between their residence and workplace. Today, the London Underground alone handles around 3.5 million passenger journeys per day, and its trains travel around 76 million kilometers each year (about 200 times the distance between the earth and the moon). Yet relatively little is known about the role of these commuting flows in sustaining dense concentrations of economic activity. On the one hand, these commuting flows impose substantial real resource costs, both in terms of time spent commuting and the construction of large networks of complex transportation infrastructure. On the other hand, they are also central to the creation of predominantly commercial and residential areas, with their distinctive characteristics for production and consumption.

In this paper, we use the mid-19th-century transport revolution from the invention of steam railways, a newly-created, spatially-disaggregated dataset for Greater London from 1801-1921, and a quantitative urban model to provide new evidence on the contribution of the separation of workplace and residence to agglomeration. The key idea behind our approach is that the slow travel times achievable by human or horse power implied that most people lived close to where they worked when these were the main modes of transportation. In contrast, the invention of steam railways dramatically reduced the time taken to travel a given distance, thereby permitting the first large-scale separation of workplace and residence. This separation enabled locations to specialize according to their comparative advantage in production and residence. Using both reduced-form and structural approaches, we find substantial effects of steam passenger railways on city size and structure. We show that our model is able to account both qualitatively and quantitatively for the observed changes in the organization of economic activity within Greater London.

Methodologically, we develop a new structural estimation procedure for the class of urban models characterized by a gravity equation for commuting flows. Although we only observe these bilateral commuting flows in 1921 at the end of our sample period, we show how this framework can be used to estimate the impact of the construction of the railway network. In a first step, we use our bilateral commuting data for 1921 to estimate the parameters that determine commuting costs. In a second step, we combine these parameter estimates with historical data on population, land values and the evolution of the transport network going back to the early-19th century. Using a combined commuter and land market clearing condition in the model, we solve for the implied unobserved historical values for employment by workplace and commuting patterns. Although we estimate the model's commuting parameters using 1921 information alone, we show in overidentification checks that its predictions provide a good approximation to the available historical data. In particular, we find that the model captures the sharp divergence between night-time and day-time population in the City of London from the mid-19th century onwards, and replicates the property of the early commuting data that most people lived close to where they worked at the dawn of the railway age.

We show that our methodology holds in an entire class of urban models, because it uses only the assumptions of gravity in commuting and land market clearing, together with the requirements that payments for residential floor space are proportional to residential income and that payments for commercial floor space are proportional to workplace income. An implication of this property is that our results hold under a range of assumptions about other model components, such as the costs of trading goods, the determinants of productivity and amenities including the strength of agglomeration forces, the supply elasticity for floorspace, and the reservation level of utility in the wider economy. Given the data for the initial equilibrium in 1921, we show that the observed changes in population and land values are sufficient statistics in the model for determining historical workplace employment and commuting patterns, and control for other potential determinants of the spatial distribution of economic activity.

While our baseline quantitative analysis controls for these other forces, another key question of interest is the counterfactual of what would have been the effect of the new commuting technology in the absence of any other changes. To address this question, we make additional assumptions about these other model components, and pick one quantitative model within our class in order to solve for a counterfactual equilibrium. In particular, we choose an extension of the canonical urban model of goods trade and commuting, which is particularly tractable, and enables us to explore alternative assumptions about structural parameters in a transparent and flexible way. Holding the supply of floor space and productivity and amenities constant, we find that removing the entire railway network reduces the total population and rateable value of Greater London by 30 and 22 percent respectively, and decreases commuting into the City of London from more than 370,000 in 1921 to less than 60,000. By comparison, removing only the underground railway network diminishes total population and rateable values for Greater London by 8 and 6 percent respectively, and brings down commuting into the City of London to just under 300,000 workers. In both cases, the increase in the net present value of land and buildings substantially exceeds historical estimates of the construction cost of the railway network. Allowing for a positive floor space supply elasticity or introducing agglomeration economies magnifies these effects. Using a calibrated floor space supply elasticity of 2.86 and elasticities of productivity and amenities to employment density of 0.05, in line with empirical estimates, we find that much of the aggregate growth of Greater London can be explained by the new transport technology of the railway.

London during the 19th century is arguably the poster child for the large metropolitan areas observed around the world today. In 1801, London's built-up area housed around 1 million people and spanned only 5 miles East to West. This was a walkable city of 60 squares and 8,000 streets that was not radically different from other large cities up to that time. In contrast, by 1901, Greater London contained over 6.5 million people, measured more than 17 miles across, and was on a dramatically larger scale than any previous urban area. By the beginning of the 20th-century, London was the largest city in the world by some margin (with New York City and Greater Paris having populations of 3.4 million and 4 million respectively at this time) and London's population exceeded that of several European countries.[1] Therefore, 19th-century London provides a natural testing ground for assessing the empirical relevance of theoretical models of city size and structure.

Our empirical setting also has a number of other attractive features. First, during this period, there is a revolution in transport technology in the form of the steam locomotive, which dramatically increased travel speeds from around 6 to 21 mph. Steam locomotives were first developed to haul freight at mines (at the Stockton to Darlington Railway in 1825) and were only later used to transport passengers (with the London and Greenwich Railway in 1836 the first to be built specifically for passengers).[2] Second, in contrast to other cities, such as Paris, London developed through a largely haphazard and organic process. Until the creation of the Metropolitan Board of Works (MBW) in 1855, there was no municipal authority that spanned the many different local jurisdictions that made up Greater London, and the MBW's responsibilities were largely centered on infrastructure. Only in 1889 was the London County Council (LCC) created, and the first steps towards large-scale urban planning for Greater London were not taken until the Barlow

[1] London overtook Beijing's population in the 1820s, and remained the world's largest city until the mid-1920s, when it was eclipsed by New York. By comparison, Greece's 1907 population was 2.6 million, and Denmark's 1901 population was 2.4 million.

[2] Stationary steam engines have a longer history, dating back at least to Thomas Newcomen in 1712, as discussed further below.

Commission of 1940. Therefore, 19th-century London provides a setting in which we would expect both the size and structure of the city to respond to decentralized market forces.

We contribute to several strands of existing research. Our paper connects with the theoretical and empirical literatures on agglomeration, including Henderson (1974), Fujita, Krugman, and Venables (1999), Fujita and Thisse (2002), Davis and Weinstein (2002), Lucas and Rossi-Hansberg (2002), Davis and Dingel (2012) and Kline and Moretti (2014), as reviewed in Rosenthal and Strange (2004), Duranton and Puga (2004), Moretti (2011) and Combes and Gobillon (2015). A key challenge for empirical work is finding exogenous sources of variation to identify agglomeration forces. Rosenthal and Strange (2008) and Combes, Duranton, Gobillon, and Roux (2010) use geology as an instrument for population density, exploiting the idea that tall buildings are easier to construct where solid bedrock is accessible. Greenstone, Hornbeck, and Moretti (2010) provide evidence on agglomeration spillovers by comparing changes in total factor productivity (TFP) among incumbent plants in "winning" counties that attracted a large manufacturing plant and "losing" counties that were the new plant's runner-up choice. In contrast, we exploit the transformation of the relationship between travel time and distance provided by the invention of the steam locomotive.

Our paper is also related to a recent body of research on quantitative spatial models, including Redding and Sturm (2008), Allen and Arkolakis (2014), Ahlfeldt, Redding, Sturm, and Wolf (2015), Redding (2016), Allen, Arkolakis, and Li (2017), Caliendo, Parro, Rossi-Hansberg, and Sarte (2018), Desmet, Nagy, and Rossi-Hansberg (2018), Monte (2018), and Monte, Redding, and Rossi-Hansberg (2018), as reviewed in Redding and Rossi-Hansberg (2017). All of these papers focus on time periods for which modern transportation networks by rail and/or road exist, whereas we exploit the dramatic change in transport technology provided by the steam locomotive. As in Ahlfeldt, Redding, Sturm, and Wolf (2015), we model heterogeneity in worker commuting decisions using an extreme value distribution, following McFadden (1974). Our main contributions relative to these previous studies are as follows. First, we show that an entire class of quantitative urban models features the same combined land and commuter market clearing condition, which enables us to develop results that are robust across this class of models, and to control for a range of unobserved determinants of the spatial distribution of economic activity. Second, we develop a new structural estimation methodology for this class of quantitative urban models, which uses bilateral commuting flows for a baseline year (in our case 1921) and undertakes comparative statics from this baseline year (in our case backwards in time). We show that this estimation procedure can be used to recover unobserved historical employment and commuting data (prior to 1921) from the bilateral commuting data for our baseline year and historical data on population and land rents.[3] This procedure is applicable in other contexts, in which historical data are incomplete or missing, but bilateral commuting data are available for a baseline year. Third, we show that a quantitative urban model can explain both qualitatively and quantitatively the dramatic changes in the organization of economic activity in Greater London following the transition from human/horse power to the steam railway.

Our paper also contributes to the empirical literature on the relationship between the spatial distribution of economic activity and transport infrastructure, as reviewed in Redding and Turner (2015). One strand of this literature has used variation across cities and regions, including Chandra and Thompson (2000), Michaels (2008), Duranton and Turner (2011), Duranton and Turner (2012), Faber (2014), Duranton, Morrow, and Turner (2014), Donaldson and Hornbeck (2016), Donaldson (2018) and Baum-Snow, Henderson, Turner, Zhang, and Brandt (2018). A second group

[3] Our use of bilateral commuting data for our baseline year enables us to encompass this class of quantitative urban models and recover the unobserved historical employment and commuting data. In contrast, Ahlfeldt, Redding, Sturm, and Wolf (2015) did not have access to bilateral commuting data between blocks, and instead used employment by residence and workplace to predict commuting between blocks.

of studies have looked within cities, including Warner (1978), Jackson (1987), McDonald and Osuji (1995), Fogelson (2003), Gibbons and Machin (2005), Baum-Snow and Kahn (2005), Billings (2011), Brooks and Lutz (2018), and Gonzalez-Navarro and Turner (2018). A third set of papers has examined the welfare effects of transport infrastructure, as in Allen and Arkolakis (2017) and Fajgelbaum and Schaal (2017). Within this literature, our work is most closely related to research on suburbanization and decentralization, including Baum-Snow (2007), Baum-Snow, Brandt, Henderson, Turner, and Zhang (2017), and Baum-Snow (2017). Our contributions are again to use the large-scale variation from the transition from human/horse power to steam locomotion and to show that our model can account both qualitatively and quantitatively for the observed changes in city structure.

The remainder of the paper is structured as follows. Section 2 discusses the historical background. Section 3 summarizes the data sources and definitions. Section 4 presents reduced-form evidence on the role of transport infrastructure improvements in shaping the organization of economic activity within Greater London. Section 5 introduces our theoretical framework. Section 6 undertakes our baseline quantitative analysis. Section 7 undertakes counterfactuals and compares the economic impact of the construction of the railway network to its construction cost. Section 8 concludes. A separate web appendix establishes isomorphisms between a class of urban models and contains the proofs of propositions, supplementary empirical results, and further details on the data.

2 Historical Background

London has a long history of settlement that dates back to before the Roman Conquest of England in 43 CE. We distinguish four main definitions of its geographical boundaries, where each subsequent region is a subset of the previous one. First, we consider London together with the Home Counties that surround it, which contain a 1921 population of 9.61 million and an area of 12,829 kilometers squared, and encompass large parts of South-East England.[4] Second, we examine Greater London, as defined by the modern boundaries of the Greater London Authority (GLA), which includes a 1921 population of 7.39 million and an area of 1,595 kilometers squared. Third, we consider the historical County of London, which has a 1921 population of 4.48 million and an area of 314 kilometers squared. Fourth, we examine the City of London, which has a 1921 population of 13,709 and an area of around 3 kilometers squared, and whose boundaries correspond approximately to the Roman city wall. From medieval times, the City of London acted as the main commercial and financial center of what became the United Kingdom, with the neighboring City of Westminster serving as the seat of Royal and Parliamentary government.[5]

Data are available for these four regions at two main levels of spatial aggregation: boroughs and parishes. Greater London and its surrounding Home Counties together encompass 257 boroughs and 1,161 parishes; Greater London alone contains 99 boroughs and 283 parishes; the County of London comprises 29 boroughs and 183 parishes; and the City of London includes 1 borough and 111 parishes.[6] In Figure 1, we show the outer boundary of the surrounding Home Counties with a thick black line; the boundary of Greater London with a thick red line; the boundary of the County of London with a thick purple line; and the boundary of the City of London with a thick green line (barely visible). Borough boundaries are shown with medium black lines; parish boundaries are indicated using medium gray lines; and the River Thames is denoted by the thick blue line. As apparent from the figure, our data permit a high-level

[4]The Home Counties that are contiguous with London include Essex, Hertfordshire, Kent, Middlesex and Surrey.

[5]For historical discussions of London, see Porter (1995), Ball and Sunderland (2001), White (2007, 2008, 2012), and Kynaston (2012).

[6]Parish boundaries in the population census change over time. We use the constant definitions of mappable units (henceforth referred to for simplicity as parishes) provided by Shaw-Taylor, Davies, Kitson, Newton, Satchell, and Wrigley (2010), as discussed further below.

of spatial resolution, where the median parish in Greater London has a 1901 population of 1,515 and an area of 0.97 kilometers squared, while the median borough in Greater London has a 1901 population of 26,288 and an area of 11.41 kilometers squared.

In the first half of the 19th-century, there was no municipal authority for the entire built-up area of Greater London, and public goods were largely provided by local parishes and vestries (centered around churches). As a result, in contrast to other cities, such as Paris, London's growth was largely haphazard and organic.[7] In response to the growing public health challenges created by an expanding population, the Metropolitan Board of Works (MBW) was founded in 1855. However, its main responsibilities were for infrastructure, and many powers remained in the hands of the parishes and vestries.[8] With the aim of creating a central municipal government with the powers required to deliver public services effectively, the London County Council (LCC) was formed in 1889. The new County of London was created from the Cities of London and Westminster and parts of the surrounding counties of Middlesex, Surrey and Kent.[9] As the built-up area continued to expand, the concept of Greater London emerged, which was ultimately reflected in the replacement of the LCC by the Greater London Council (GLC) in 1965. Following the abolition of the GLC in 1985 by the government of Margaret Thatcher, Greater London again had no central municipal government, until the creation of the Greater London Authority (GLA) in 1999.

At the beginning of the 19th-century, the most common mode of transport was walking, with average travel speeds in good road conditions of around 3 miles per hour (mph). The state of the art technology for long distance travel was the stage coach, but it was expensive because of the multiple changes in teams of horses required over long distances, and hence was relatively infrequently used. Even with this elite mode of transport, poor road conditions limited average long distance travel speeds to around 5 mph (see for example Gerhold 2005 and Bogart 2017). Given these limited transport options, most people lived close to where they worked, as discussed in the analysis of English 18th-century time use in Voth (2001). With the growth of urban populations, attempts to improve existing modes of transport led to the introduction of the horse omnibus from Paris to London in the 1820s. Its main innovation relative to the stage coach was increased passenger capacity for short-distance travel. However, the limitations of horse power and road conditions ensured that average travel speeds remained low at around 6 mph.[10] A further innovation along the same lines was the horse tram (introduced in London in 1860), but average travel speeds again remained low, in part because of road congestion (again at around 6 mph).[11]

Against this background, the steam passenger railway constituted a major transport innovation, although one with a long and uncertain gestation. The first successful commercial development of a stationary steam engine was by Thomas Newcomen in 1712 to pump mine water. However, the development of the separate condenser and rotary motion by James Watt from 1763-75 substantially improved its efficiency and expanded its range of potential applications. The first commercial use of mobile steam locomotives was to haul freight from mines at the Stockton and Darlington railway in 1825. However, in part as a result of fears about the safety of steam locomotives and the

[7] The main exceptions are occasional Royal interventions, such as the creation of Regent Street on the initiative of the future George IV in 1825.

[8] See Owen (1982). These public health challenges included cholera outbreaks, as examined in Ambrus, Field, and Gonzalez (2015). The MBW's main achievements were the construction of London's Victorian sewage system and the Thames embankment, as discussed in Halliday (1999).

[9] The LCC continued the MBW's infrastructure improvements, including some new road construction through housing clearance (e.g. Kingsway close to the London School of Economics), and built some social housing. The first steps towards large-scale urban planning for Greater London were not taken until the Barlow Commission in 1940, as discussed in Foley (1963).

[10] See Section G.6 of the web appendix, Barker and Robbins (1963) and London County Council (1907).

[11] A later innovation was the replacement of the horse tram with the electric tram (with the first fully-operational services starting in 1901). But average travel speeds remained low at around 8 mph, again in part because of road congestion, as shown in Section G.6 of the web appendix. Private car use was negligible with registered cars per person in the County of London in 1920 equal to 0.01 (London County Council 1921).

dangers of asphyxiation from rapid travel, it was not until 1833 that carriages with passengers were hauled by steam locomotives at this railway. Only in 1836 did the London and Greenwich railway open as the first steam railway to be built specifically for passengers. The result was a dramatic transformation of the relationship between travel time and distance, with average travel speeds using this new technology of around 21 mph.[12]

Railway development in London, and the United Kingdom more broadly, was undertaken by private companies in a competitive and uncoordinated fashion.[13] These companies submitted proposals for new railway lines for authorization by Acts of Parliament. In response to a large number of proposals to construct railway lines through Central London, a Royal Commission was established in 1846 to investigate these proposals. To preserve the built fabric of Central London, this Royal Commission recommended that railways be excluded from a central area delineated by the Euston Road to the North and the Borough and Lambeth Roads to the South.[14] A legacy of this recommendation was the emergence of a series of railway terminals around the edge of this central area, which led to calls for an underground railway to connect these terminals. These calls culminated in the opening of the Metropolitan District Railway in 1863 and the subsequent development of the Circle and District underground lines. While these early underground railways were built using "cut and cover" methods, further penetration of Central London occurred with the development of the technology for boring deep-tube underground railways, as first used for the City and South London Railway, which opened in 1890, and is now part of the Northern Line.[15]

In Figures 2, 3 and 4, we show the overground and underground railway networks in Greater London for 1841, 1881 and 1921 respectively, where a complete set of maps for each census decade is found in Sections G.4 and G.5 of the web appendix. The parts of Greater London outside the County of London are shown in white; the parts of the County of London outside the City of London are displayed in blue; and the City of London is indicated in gray. Overground railway lines are displayed in black and underground railway lines are indicated in red. In 1841, which is the first population census year in which any overground railways are present, there are only a few railway lines. These radiate outwards from the County of London, with a relatively low density of lines in the center of the County of London. Four decades later in 1881, the County of London is criss-crossed by a dense network of railway lines, with greater penetration into the center of the County of London, in part because of the construction of the first underground railway lines. Another four decades later in 1921, there is a further increase in the density of both overground and underground lines.

3 Data

We construct a new spatially-disaggregated dataset for London for the period 1801-1921, which is discussed in further detail in Section G of the web appendix. Our main data source is the population census of England and Wales, which

[12]Consistent with this difference in travel speeds, railways were more frequently used for longer-distance travel, while omnibuses and trams were more important over shorter distances (including from railway terminals to final destinations), which tended to make these alternative modes of transport complements rather than substitutes. The share of railways in all passenger journeys by public transport was 49 percent in 1867 (the first year for which systematic data are available) and 32 percent in 1921 (see London County Council 1907). From 1860 onwards, Acts of Parliament authorizing railways typically included clauses requiring the provision of "workmen's trains" with cheap fares for working-class passengers, as ultimately reflected in the 1883 Cheap Trains Act (see for example Abernathy 2015).

[13]For further historical discussion of railway development, see for example Croome and Jackson (1993), Kellet (1969), and Wolmar (2009, 2012).

[14]This parliamentary exclusion zone explains the location of Euston, King's Cross and St. Pancras railway terminals all on the Northern side of the Euston Road. Exceptions were subsequently allowed, often in the form of railway terminals over bridges coming from the south side of the Thames at Victoria (1858), Charing Cross (1864), Cannon Street (1866), and Ludgate Hill (1864), and also at Waterloo (1848).

[15]When it opened in 1863, the Metropolitan District Railway used steam locomotives. In contrast, the City and South London Railway was the first underground line to use electric traction from its opening in 1890 onwards.

begins in 1801, and is enumerated every decade thereafter. A first key component for our quantitative analysis of the model is the complete matrix of bilateral commuting flows between boroughs, which is reported for the first time in the 1921 population census.[16] Using this matrix, we find that commuting flows between other parts of England and Wales and Greater London were small in 1921, such that Greater London was largely a closed commuting market.[17] Summing across rows in the matrix of bilateral commuting flows for Greater London, we obtain employment by workplace for each borough (which we refer to as "workplace employment"). Summing across columns, we obtain employment by residence for each borough (which we refer to as "residence employment"). We also construct an employment participation rate for each borough in 1921 by dividing residence employment by population.

We combine these data on bilateral commuting flows in 1921 with historical population data from earlier population censuses from 1801-1911. Assuming that the ratio of residence employment to population is stable for a given borough over time, we use the 1921 value of this ratio and the historical population data to construct residence employment for earlier census years.[18] Parish and borough boundaries are relatively stable throughout most of the 19th century, but experience substantial change in the early-twentieth century. For our reduced-form empirical analysis using the parish-level data, we construct constant parish boundary data for the period 1801-1901 using the classification provided by Shaw-Taylor, Davies, Kitson, Newton, Satchell, and Wrigley (2010), as discussed in Section G.1 of the web appendix. For our quantitative analysis of the model using the borough-level data, we use constant borough definitions throughout our sample period based on the 1921 boundaries. For years prior to 1921, we allocate the parish-level data to the 1921 boroughs by weighting the values for each parish by its share of the geographical area of the 1921 borough. Given that parishes have a much smaller geographical area than boroughs, most parishes lie within a single 1921 borough.

We measure the value of floor space using rateable values, which correspond to the annual flow of rent for the use of land and buildings, and equal the price times the quantity of floor space in the model. In particular, these rateable values correspond to "The annual rent which a tenant might reasonably be expected, taking one year with one another, to pay for a hereditament, if the tenant undertook to pay all usual tenant's rates and taxes ... after deducting the probable annual average cost of the repairs, insurance and other expenses" (see London County Council 1907). With a few minor exceptions, they cover all categories of property, including public services (such as tramways, electricity works etc), government property (such as courts, parliaments etc), private property (including factories, warehouses, wharves, offices, shops, theaters, music halls, clubs, and all residential dwellings), and other property (including colleges and halls in universities, hospitals and other charity properties, public schools, and almshouses). The main exemptions include roads, canals, railways, mines, quarries, Crown property occupied by the Crown, and places of divine worship, as discussed further in Section G.2 of the web appendix. Rateable values were assessed at the parish level approximately every five years during our sample period. All of the above categories of properties were assessed, regardless of whether or not their owners were liable for income tax.

These rateable values have a long history in England and Wales, dating back to the 1601 Poor Relief Act, and were originally used to raise revenue for local public goods. Different types of rateable values can be distinguished,

[16]The population census for England and Wales is the first to report bilateral commuting data. In the United States, the 1960 population census is the first to report any commuting information, and the matrix of bilateral commuting flows between counties is not reported until 1990.

[17]In the 1921 population census, 96 percent of the workers employed in Greater London also lived in Greater London. Of the remaining 4 percent, approximately half lived in the surrounding Home Counties, and the remainder lived in other parts of England and Wales. As residence is based on where one slept on Census night, while workplace is usual place of work, some of this 4 percent could be due to business trips or other travel.

[18]Empirically, we find relatively little variation in employment participation rates across boroughs in 1921.

depending on the use of the revenue raised: Schedule A Income Taxation, Local Authority Rates, and Poor Law Rates. Where available, we use the Schedule A rateable values, since Schedule A is the section of the national income tax concerned with income from property and land, and these rateable values are widely regarded as corresponding most closely to market valuations.[19] Where these Schedule A rateable values are not available, we use the Local Authority rateable values, Poor Law rateable values, or property valuations for income tax. For years for which more than one of these measures is available, we find that they are highly correlated with one another across parishes. The overall level of the property valuations for income tax is somewhat lower, which is consistent with the rateable values including all properties, regardless of the income tax liability of their owners. To address this difference, we use a consistent time-series on Schedule A rateable values for the County of London constructed by London County Council. In particular, we compute the ratio for the County of London between each of our other measures of property values and the consistent Schedule A rateable values, and use this ratio to adjust our other measures for all parishes. We thus construct data on the value of land and buildings for each parish for 1815, 1843, 1847, 1852, 1860, 1881, 1896, 1905, 1911 and 1921. We use linear interpolation between these years to create a time-series on rateable values for each parish for each census decade from 1831-1921.

Data on workplace employment are not available prior to 1921. Therefore, in our structural estimation of the model, we use our bilateral commuting data for 1921, together with our data on residence employment and rateable values for earlier years, to generate model predictions for workplace employment for earlier years. In overidentification checks, we compare these model predictions to the data on the "day population" of the City of London that are available from the Day Censuses of 1866, 1881, 1891 and 1911.[20] In the face of a declining residential population ("night population"), the City of London Corporation undertook these censuses of the "day population" to demonstrate its enduring commercial importance. The day population is defined as "... every person, male or female, of all ages, residing, engaged, occupied, or employed in each and every house, warehouse, shop, manufactory, workshop, counting house, office, chambers, stable, wharf, etc, and to include all persons, of both sexes and all ages, on the premises during the working hours of the day, whether they sleep or do not sleep there."[21] Therefore, the "day population" includes both those employed in the City of London and those resident in the City of London and present during the day (e.g. because they are economically inactive). We generate an analogous measure of day population in the model, which equals our model's prediction for workplace employment plus our data on economically-inactive residents (population minus residence employment).

In additional overidentification checks, we compare the model's predictions for commuting distances with historical commuting data based on the residence addresses of the employees of the tailor Henry Poole, which has been located at the same workplace address in Savile Row in the City of Westminster since 1822.[22] Finally, we combine these data with a variety of other Geographical Information Systems (GIS) data on the evolution of the transport network over time, as discussed further in Sections G.4 and G.5 of the web appendix. We distinguish four modes of transport based on the historical evidence on average travel speeds discussed in Section G.6 of the web appendix: (a) overground railways; (b) underground railways, (c) omnibus and tram, and (d) walking.

[19]For example, Stamp (1922) argues that "It is generally acknowledged that the income tax, Schedule A, assessments are the best approach to the true values." (page 25) After the Metropolis Act of 1869, all rateable values for the County of London are assessed based on Schedule A.

[20]For further discussion of the City of London Day Censuses, see Section G.3 of the web appendix.

[21]Salmon (1891), page 97.

[22]For further discussion of the Henry Poole data, see Section G.8 of the web appendix and Green (1988).

4 Reduced-Form Evidence

In subsection 4.1, we present time-series evidence on the evolution of the organization of economic activity within Greater London over our long historical time period. In subsection 4.2, we provide further evidence on the role of railways in shaping this reorganization of economic activity. In subsection 4.3, we report a non-parametric specification that estimates a separate railway treatment for each parish in Greater London.

4.1 City Size and Structure over Time

We first illustrate the dramatic changes in the spatial organization of economic activity within Greater London between 1801 and 1921. In Figure 5, we display total population over time for the City of London (left panel) and Greater London (right panel). In each case, population is expressed as an index relative to its value in 1801 (such that 1801=1). In the first half of the 19th century, population in the City of London was relatively constant (at around 130,000), while population in Greater London grew substantially (from 1.14 million to 2.69 million). From 1851 onwards (shown by the red vertical line), there is a sharp drop in population in the City of London, which falls by around 90 percent to 13,709 by 1921, with this rate of decline slowing over time towards the end of the 19th century and beginning of the 20th century. Over the same period, the population of Greater London as a whole continues to grow rapidly from 2.69 million in 1851 to 7.39 million in 1921.[23] Therefore, the rapid expansion in population for Greater London throughout the 19th century goes hand in hand with a precipitous drop in population in its most important commercial center from the mid-19th century onwards.[24]

In Figure 6, we contrast the evolution of "night" and "day" population for the City of London. The "night" population data are the same as those shown in the left panel of Figure 5 and are taken from the population census (based on residence on census night). In contrast, the "day" population data are from the City of London Day Censuses, as discussed above, except for the 1921 figure, which is constructed as workplace employment plus economically-inactive residents (i.e. population minus residence employment). Figure 6 shows that the sharp decline in night population in the City of London from 1851 onwards coincides with a sharp rise in its day population. This pattern suggests that the combination of population decline in the City of London and population expansion for Greater London as a whole is explained by the City of London increasingly specializing as a workplace rather than as a residence. This evidence of an increased separation of workplace and residence is also consistent with a sharp rise in public transport journeys per head of population from around 7 in 1834 to just under 400 in 1921, as shown in Figure F.1 in Section F of the web appendix. Extrapolating the day population series further back in time to the 1850s (not shown in Figure 6) would suggest that night and day population were approximately equal to one another at this time, which is consistent with most people living close to where they worked in the early decades of the 19th century.

Additional evidence that the sharp drop in night population in the City of London from the mid-19th century onwards is not part of an economic decline in this location comes from our rateable value data. In Figure 7, we display the City of London's share of total rateable value in Greater London. In the early-19th century, this share declines from around 14 to 9 percent, which is consistent with a geographical expansion in the built-up area of Greater London.

[23] Although the second decade of the twentieth century spans the First World War from 1914-18, the primitive nature of aircraft and airship technology at that time ensured that Greater London experienced little bombing and destruction (see for example White 2008).

[24] While London as an empirical setting has the advantage of rich historical data before and after the arrival of the railway, together with bilateral commuting information for 1921, this pattern of a declining central population and expanding metropolitan population with improvements in transport technology is also observed for other metropolitan areas, as described for New York in Angel and Lamson-Hall (2014).

As this expansion occurs, and undeveloped land becomes developed, the share of already-developed land in overall land values tends to fall, because the denominator of this share increases. In contrast to this pattern, in the years after 1851 when the City of London experiences the largest declines in residential population, its rateable value share *increases* from 9 to 11 percent. Finally, in the decades at the end of the 19th century, the pattern of a decline in this rateable value share again reasserts itself, consistent with the continuing geographical expansion in the built-up area of Greater London. Therefore, the steep population decline in the City of London in the decades immediately after 1851 involves an *increase* rather than a decrease in the relative value of this location.

A comparison of Figures 5 and 6 to the evolution of the railway network over time in Figures 2-4 above already suggests that the observed changes in night population, day population and land value are likely to be related to the innovation in transport technology. The timing of the response in the City of London's population in Figure 5 (from 1851 onwards) relative to the first opening of a steam railway (in 1836) is consistent with the railway network becoming increasingly valuable as more locations are connected to it and with it taking time for firms and workers to relocate in response to the new transport technology. The sharpest declines in the population of the City of London from 1851-1881 in Figure 5 correspond closely to the greatest increases in the penetration of overground and underground railways into the center of the County of London in Figure 3. We provide further evidence in the next section on this timing of the population response to the new transport technology using a difference-in-differences specification.

4.2 Difference-in-Differences Specification

We now use our spatially-disaggregated parish-level data for Greater London from 1801-1901 to provide reduced-form evidence on the role played by railways in this reorganization of economic activity. In our baseline specification, we combine both overground and underground railways, although we find a similar pattern of results in robustness checks using only overground railways. The main identification challenge is that railways are unlikely to be randomly assigned, because they were constructed by private-sector companies, whose stated objective was to maximize shareholder value. As a result, parishes in which economic activity would have grown for other reasons could be more likely to be assigned railways. We address this identification challenge by considering specifications that include both a parish fixed effect and a parish time trend, and examining the relationship between the timing of deviations from these parish trends and the arrival of the railway. We start by estimating a common average treatment effect for Greater London as a whole, before later exploring heterogeneity within Greater London, as suggested by our findings for the City of London in the previous subsection. Before including our full set of controls, we consider the following baseline specification:

$$\log R_{jt} = \alpha_j + \sum_{\tau=0}^{\tau=T} \beta_\tau \left(\mathbb{S}_j \times \mathbb{I}_\tau \right) + d_t + u_{jt} \tag{1}$$

where j indexes parishes; t indicates the census year; R_{jt} is parish population; α_j is a parish fixed effect; d_t is a census year dummy; $\mathbb{S}_j$ is an indicator variable that equals one if a parish has an overground or underground railway station in at least one census year during our sample period; τ is a treatment year indicator, which equals census year minus the last census year in which a parish had no railway;[25] and $\mathbb{I}_\tau$ is an indicator variable that equals one for parishes that are treated with a railway in treatment year τ and zero otherwise. We cluster the standard errors on boroughs, which allows the error term to be serially correlated within parishes over time, and to be correlated across parishes

[25] Therefore, $\tau = 0$ corresponds to the last year in which a parish had no railway and positive values of τ correspond to post-treatment years. For example, if the railway arrives in a parish in 1836, census year 1831 corresponds to $\tau = 0$, and census year 1841 corresponds to $\tau = 10$.

within boroughs.[26]

The inclusion of the parish fixed effects (α_j) controls for the non-random assignment of railways based on the level of parish population or other time-invariant factors. Therefore, we allow for the fact that parishes treated with a railway could have had higher population levels in all years (both before and after the railway). The census year dummies (d_t) control for secular changes in population across all parishes. The key coefficients of interest (β_τ) are those on the interaction terms between the railway indicator ($\mathbb{S}_j$) and the treatment year indicator ($\mathbb{I}_\tau$), which capture the treatment effect of a railway on parish population in treatment year τ. They have a "difference-in-differences" interpretation, where the first difference compares treated to untreated parishes, and the second difference undertakes this comparison before and after the arrival of the railway. The main effect of $\mathbb{S}_j$ is captured in the parish fixed effects and the main effect of $\mathbb{I}_\tau$ is captured in the census year dummies. We include six interaction terms for decades from 10 to 60+ years after a parish receives a railway station. We aggregate treatment years greater than 60 into the 60+ category to ensure that this final category has a sufficient number of observations.[27]

In Column (1) of Table 1, we report the results of estimating this baseline specification from equation (1). We find positive and statistically significant treatment effects of the railway on parish population, which range from around 60-270 percent (up to the log approximation). In Column (2), we augment this specification with parish time trends, which allows for the non-random assignment of railways based on trends in parish population over time.[28] In this specification, we allow for the fact that parishes treated with a railway could have had higher trend population growth in all years (both before and after the railway). We now identify the treatment effect of the railway solely from deviations from these parish time trends after the arrival of the railway. Again we find positive and statistically significant treatment effects, which are now somewhat smaller but still substantial, ranging from 11-133 percent.

In Column (3), we check whether or not the timing of the deviation from these parish time trends coincides with the arrival of the railway. In particular, we augment the specification from Column (2) with interaction terms between the railway dummy ($\mathbb{S}_j$) and dummies for treatment years before the arrival of the railway ($\mathbb{I}_\tau$ for $\tau < 0$). The excluded category is $\tau = 0$. We consider a symmetric time window, in which we include six interaction terms for decades from 10 to 60+ years before and after a parish receives a railway station. As apparent from Column (3), we find no evidence of statistically significant deviations from the parish time trends before the arrival of the railway. But we continue to find large and statistically significant deviations from these parish time trends in the years after the arrival of the railway. Therefore, this specification supports the interpretation of our estimates in Column (2) that the change in parish population growth rates occurs immediately after the arrival of the railway.

While we have so far focused on estimating a common treatment effect across all boroughs within Greater London, we now explore heterogeneity in this treatment, and provide further evidence connecting the decline in population in the City of London in Figure 5 to the arrival of the railway. In particular, we allow the railway treatment effect to differ between the City of London and other parts of Greater London by augmenting our baseline specification from

[26] We also experimented with Heteroskedasticity and Autocorrelation Consistent (HAC) standard errors following Conley (1999), and found these to be typically smaller than the standard errors clustered on borough that are reported below.

[27] As the first railway was constructed in 1836, the maximum possible value for τ is 70 for parishes that receive a railway station before 1841, for which $\tau = 10$ in 1841 and $\tau = 70$ in 1901. Similarly, as our parish-level sample ends in 1901, the minimum possible value for τ is -90 for parishes that receive a railway after 1891, for which $\tau = 10$ in 1901 and $\tau = -90$ in 1801. Of our 3,113 (283 $\times$ 11) parish-year observations, 1,408 involve parishes that have a railway station in at least one census year during the sample period. The distribution of these 1,408 observations across the treatment years is: $\tau <= -60$ (133); $\tau = -50$ (83) ; $\tau = -40$ (106); $\tau = -30$ (128); $\tau = -20$ (128); $\tau = -10$ (128); $\tau = 0$ (128); $\tau = 10$ (128); $\tau = 20$ (119); $\tau = 30$ (104); $\tau = 40$ (96); $\tau = 50$ (60); and $\tau >= 60$ (67).

[28] One of the parish time trends is collinear with the year dummies and hence is omitted without loss of generality.

equation (1) (as reported in Column (1) of Table 1) with a three-way interaction term between the railway dummy, the treatment year dummy, and a dummy for the City of London:

$$\log R_{jt} \quad = \alpha_j + \sum_{\tau=0}^{\tau=T} \beta_\tau \left(\mathbb{S}_j \times \mathbb{I}_\tau\right) + \sum_{\tau=0}^{\tau=T} \gamma_\tau \left(\mathbb{S}_j \times \mathbb{I}_\tau \times \mathbb{I}_j^{\text{City}}\right) + d_t + u_{jt}, \tag{2}$$

where $\mathbb{I}_j^{\text{City}}$ is an indicator variable that equals one for parishes in the City of London and zero otherwise; all other variables are defined as above; the railway treatment effect for parishes in the City of London is now given by $(\beta_\tau + \gamma_\tau)$; and the railway treatment effect for other parts of Greater London remains equal to β_τ. A legacy of the parliamentary exclusion zone is that relatively few parishes within the City of London are treated with overground railways. Furthermore, the treatments for underground railways occur relatively late in the sample period. As a result, there is a relatively short interval after the arrival of the railway for these parishes. Therefore, we only include three City-of-London interaction terms $\{\gamma_\tau\}$ for decades 10 to 30+ years after a parish receives a railway station.[29]

In Column (4) of Table 1, we estimate this specification from equation (2). Again we find positive and statistically significant treatment effects of the railway on parish population for other parts of Greater London (as captured by β_τ), which remain of around the same magnitude as in Column (1). However, we find substantially and statistically significantly smaller treatment effects of the railway on parish population for the City of London (as reflected in large negative and statistically significant estimates of γ_τ). Furthermore, the estimated γ_τ are larger in absolute magnitude than the estimated β_τ, implying an overall negative and statistically significant treatment effect of the railway on the population of parishes in the City of London $(\beta_\tau + \gamma_\tau)$.

In Column (5), we augment this specification with parish time trends to allow parishes that are treated with a railway to have different trend rates of population growth in all years (both before and after the railway). Even in this specification, where we identify the railway treatment effect solely from deviations from parish time trends, we find the same pattern of negative and statistically significant treatment effects for parishes in the City of London and positive and statistically significant treatment effects for parishes in other parts of Greater London.

In Column (6), we again check whether the timing of these deviations from parish time trends coincides with the arrival of the railway. We augment the specification in Column (5) by including interactions with treatment years before the arrival of the railway for both sets of coefficients (β_τ and γ_τ). The excluded category is again the treatment year ($\tau = 0$), and we again consider symmetric time windows before and after the arrival of the railway. We find no evidence of statistically significant deviations from the parish time trends before the arrival of the railway, whether for the City of London $(\beta_\tau + \gamma_\tau)$ or for other parts of Greater London (β_τ). However, we continue to find large and statistically significant deviations from the parish time trends in the years after the arrival of the railway, which are negative for the City of London $(\beta_\tau + \gamma_\tau < 0)$ and positive for other parts of Greater London $(\beta_\tau > 0)$. This specification provides further support for the idea that the reduction in population growth in the City of London and its increase further from the center of Greater London is closely connected to the railway. Indeed, it is hard to think of confounding factors that are timed to coincide precisely with the arrival of the railway, and are structured to have exactly the same pattern of opposite effects on population in central versus outer Greater London.

[29] Of the 1,221 (111×11) parish-year observations for the City of London, only 154 of these observations involve parishes that have a railway station in at least one census year during the sample period. The distribution of these 154 observations across the treatment years is $\tau <= -30$ (73); $\tau <= -20$ (14) ; $\tau <= -10$ (14); $\tau = 0$ (14); $\tau = 10$ (14); $\tau = 20$ (11) and $\tau >= 30$ (14).

4.3 Non-parametric Specification

To provide further evidence on the heterogeneity in railway treatment effects, we now report a non-parametric specification, in which we estimate a separate railway treatment for each parish. We show that the difference in estimated treatment effects between the City of London and the rest of Greater London in the previous subsection reflects a more general pattern, in which the estimated railway treatment varies systematically with distance from the center of Greater London. As in the previous subsection, we consider a "difference-in-differences" specification, in which the first difference is across parishes, and the second difference is across time.

In a first step, we compute the relative population of parishes, by differencing the log population for each parish in each year from the mean across parishes in that year. This differencing from mean log population in each year removes any secular trend in population across all Greater London parishes over time, which allows us to control for the fact that different parishes are treated with the railway in different census years. In a second step, we compute the growth in the relative population of each parish over the thirty-year period before the arrival of the railway, where this difference over time differences out any fixed effect in the level of log relative parish population. We focus on a narrow thirty-year window to ensure a similar time interval over which population growth is computed for all parishes. We cannot compute this difference for parishes that are never treated with the railway, and hence we drop these parishes. All other parishes have at least thirty years before the arrival of the railway, because our sample begins in 1801, and the first railway in Greater London is built in 1836.

In a third step, we compute the growth in the relative population of each parish over the thirty-year period after the arrival of the railway, where this difference over time again differences out any fixed effect in the level of log relative parish population. We again focus on a narrow thirty-year window and drop any parish with less than thirty years between its treatment year and the end of our parish-level sample in 1901. In a fourth and final step, we compute the "difference-in-difference," namely the change in each parish's growth in relative population between the thirty-year periods before and after the arrival of the railway. By taking this difference between the growth rates before and after the arrival of the railway, we difference out any parish time trend that is common to these two periods. Therefore, we again focus on deviations from parish trends, as in the previous subsection.

In Figure 8, we display these double differences in relative population growth for each parish against the straight-line distance from its centroid to the Guildhall in the center of the City of London. We indicate parishes in the City of London by hollow red circles, while parishes in the other parts of Greater London are denoted by solid blue circles. We also show the locally-weighted linear least squares regression relationship between the two variables as the solid black line. We find a sharp non-linear relationship between the railway treatment and distance from the Guildhall. For parishes within five kilometers of the Guildhall, we find negative average estimated treatment effects (an average of -0.56 log points), particularly for those parishes inside the City of London. In contrast, for parishes beyond five kilometers from the Guildhall, we find positive average estimated treatment effects (an average of 0.19 log points), with these treatment effects becoming smaller for more peripheral parishes. These substantial differences between the two groups are statistically significant at conventional critical values.

Therefore, in this non-parametric specification that allows for heterogeneous treatment effects across parishes, we again find evidence of a systematic reorganization of economic activity. Following the arrival of the railway, we find a reduction in relative population growth in parishes close to the center of Greater London, and an increase in relative population growth in parishes further from the center of Greater London.

5 Theoretical Framework

We now develop a theoretical framework to explain these changes in the organization of economic activity within Greater London. We show that this theoretical framework encompasses an entire class of quantitative urban models that satisfy the following three properties: (i) a gravity equation for bilateral commuting flows; (ii) land market clearing, such that income from the ownership of floor space equals the sum of payments for residential and commercial floor space; (iii) payments for residential floor space are a constant proportion of residential income (the total income of all residents) and payments for commercial floor space are a constant proportion of workplace income (the total income of all workers). Within this class of models, workplace incomes are sufficient statistics for the demand for commercial floor space, while residential incomes are sufficient statistics for the demand for residential floor space, and commuting costs regulate the difference between workplace and residential incomes.

In Section D of the web appendix, we develop a number of isomorphisms, in which we show that this class of quantitative urban models encompasses a wide range of different assumptions about consumption, production and goods' trade costs, including (i) the canonical urban model with a single final good and costless trade (as in Lucas and Rossi-Hansberg 2002 and Ahlfeldt, Redding, Sturm, and Wolf 2015); (ii) multiple final goods with costly trade and Ricardian technology differences (as in Eaton and Kortum 2002 and Redding 2016); (iii) final goods that are differentiated by origin with costly trade (as in Armington 1969 and Allen, Arkolakis, and Li 2017); (iv) horizontally-differentiated firm varieties with costly trade (as in Helpman 1998, Redding and Sturm 2008, and Monte, Redding, and Rossi-Hansberg 2018); (v) extensions of each of these frameworks to incorporate non-traded services. All of these specifications accommodate commuting between residence and workplace; they comprise both perfectly and imperfectly competitive market structures; and they encompass both constant and increasing returns to scale production technologies. Regardless of which of these specifications is chosen, our quantitative approach remains the same, because it only relies on the three properties discussed above, which hold across all of these specifications.

We consider a city (Greater London) that is embedded within a wider economy (the United Kingdom). The city consists of a discrete set of locations $\mathbb{N}$ (boroughs in our data). Time is discrete and is indexed by t (census decades in our data). Workers are assumed to be geographically mobile and choose between the city and the wider economy. Population mobility implies that the expected utility from living and working in the city equals the reservation level of utility in the wider economy $\bar{U}_t$. While we take this reservation level of utility as given in our baseline specification, we endogenize it in Sections D.2-D.4 of the web appendix, and show that our baseline quantitative approach remains unchanged. If a worker chooses the city, she chooses a residence n and a workplace i from the set of locations $n, i \in \mathbb{N}$ to maximize her utility.[30] We allow locations to differ from one another in terms of their attractiveness for production and residence, as determined by productivity, amenities, the supply of floor space, and transport connections, where each of these location characteristics can evolve over time.[31]

[30] Motivated by our empirical finding above that net commuting into Greater London is small even in 1921, we assume prohibitive commuting costs between Greater London and the wider economy. Therefore, a worker cannot live in the city and work in the wider economy or vice versa.

[31] To ease the exposition, we typically use n for residence and i for workplace, except where otherwise indicated.

5.1 Preferences

Worker preferences are defined over consumption goods and residential floor space. We assume that these preferences take the Cobb-Douglas form, such that the indirect utility for a worker ω residing in n and working in i is:[32]

$$U_{ni}(\omega) = \frac{b_{ni}(\omega) w_i}{\kappa_{ni} P_n^{\alpha} Q_n^{1-\alpha}}, \qquad 0 < \alpha < 1, \tag{3}$$

where we suppress the time subscript from now onwards, except where important; P_n is the price index for consumption goods; Q_n is the price of floor space, w_i is the wage, κ_{ni} is an iceberg commuting cost, and $b_{ni}(\omega)$ is an idiosyncratic amenity draw that captures all the idiosyncratic factors that can cause an individual to live and work in particular locations within the city.[33] Each of the isomorphisms in Section D of the web appendix involves a different specification for the consumption price index (P_n). We show below that our baseline quantitative approach holds regardless of which of these specifications for the consumption price index (P_n) is chosen.

We assume that idiosyncratic amenities ($b_{ni}(\omega)$) are drawn from an independent extreme value (Fréchet) distribution for each residence-workplace pair and each worker:

$$G_{ni}(b) = e^{-B_n b^{-\epsilon}}, \qquad B_n > 0,\ \epsilon > 1, \tag{4}$$

where the Fréchet scale parameter B_n determines the average desirability of location n as a residence (e.g. leafy streets and scenic views).[34] Our quantitative approach allows these residential amenities (B_n) to be potentially endogenous to the surrounding concentration of economic activity. The Fréchet shape parameter ϵ regulates the dispersion of idiosyncratic amenities, which controls the sensitivity of worker location decisions to economic variables (e.g. wages and the cost of living). The smaller ϵ is, the greater the heterogeneity in idiosyncratic amenities, and the less sensitive work location decisions are to economic variables.

Conditional on choosing to live in Greater London, the specification for preferences in equations (3) and (4) implies that the probability a worker chooses to reside in n and work in i is:

$$\lambda_{ni} \equiv \frac{L_{ni}}{\bar{L}} = \frac{B_n w_i^{\epsilon} \left(\kappa_{ni} P_n^{\alpha} Q_n^{1-\alpha}\right)^{-\epsilon}}{\sum_{r \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} B_r w_{\ell}^{\epsilon} \left(\kappa_{r\ell} P_r^{\alpha} Q_r^{1-\alpha}\right)^{-\epsilon}}, \tag{5}$$

where L_{ni} is the measure of commuters from n to i; $\bar{L}$ is total city employment, which equals total city residents $\bar{R}$.[35]

A first key implication of our extreme value specification for idiosyncratic amenities is that bilateral commuting flows in equation (5) satisfy a gravity equation. Therefore, the probability of commuting between residence n and workplace i depends on the characteristics of that residence n, the attributes of that workplace i and bilateral commuting costs ("bilateral resistance"). Furthermore, this probability also depends on the characteristics of all residences r, all workplaces ℓ and all bilateral commuting costs ("multilateral resistance").

Summing across workplaces, we obtain the probability that an individual lives in each location ($\lambda_n^R = R_n/\bar{L}$),

[32] For empirical evidence using US data in support of the constant housing expenditure share implied by the Cobb-Douglas functional form, see Davis and Ortalo-Magné (2011).

[33] Although we model commuting costs in terms of utility, they enter the indirect utility function (3) multiplicatively with the wage, which implies that there is a closely-related formulation in terms of the opportunity cost of time spent commuting.

[34] This Fréchet scale parameter enters worker choice probabilities isomorphically to a Hicks-neutral amenities shifter in the indirect utility function (3). While we assume that this parameter varies by residence n to capture residential amenities, it is straightforward to also allow it to vary by workplace i, or to allow it to vary by both residence n and workplace i.

[35] While we assume that workers make location decisions, we allow each worker to have non-working dependents, where we choose the ratio of workers to non-working dependents to match the ratio of residence employment to population in our census data.

while summing across residences, we have the probability that an individual works in each location ($\lambda_n^L = L_i/\bar{L}$):

$$\lambda_n^R = \frac{\sum_{\ell \in \mathbb{N}} B_n w_\ell^\epsilon \left(\kappa_{n\ell} P_n^\alpha Q_n^{1-\alpha}\right)^{-\epsilon}}{\sum_{r \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} B_r w_\ell^\epsilon \left(\kappa_{r\ell} P_r^\alpha Q_r^{1-\alpha}\right)^{-\epsilon}}, \qquad \lambda_i^L = \frac{\sum_{r \in \mathbb{N}} B_r w_i^\epsilon \left(\kappa_{ri} P_r^\alpha Q_r^{1-\alpha}\right)^{-\epsilon}}{\sum_{r \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} B_r w_\ell^\epsilon \left(\kappa_{r\ell} P_r^\alpha Q_r^{1-\alpha}\right)^{-\epsilon}}, \tag{6}$$

where R_n denotes employment by residence in location n and L_i denotes employment by workplace in location i.

A second key implication of our extreme value specification for idiosyncratic amenities is that expected utility is equalized across all pairs of residences and workplaces within Greater London and equal to the reservation level of utility in the wider economy:

$$\bar{U} = \delta \left[\sum_{r \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} B_r w_\ell^\epsilon \left(\kappa_{r\ell} P_r^\alpha Q_r^{1-\alpha}\right)^{-\epsilon} \right]^{\frac{1}{\epsilon}}, \tag{7}$$

where $\delta = \Gamma((\epsilon - 1)/\epsilon)$ and $\Gamma(\cdot)$ is the Gamma function.

The intuition for this second result is that bilateral commutes with attractive economic characteristics (high workplace wages and low residence cost of living) attract additional commuters with lower idiosyncratic amenities, until expected utility (taking into account idiosyncratic amenities) is the same across all bilateral commutes and equal to the reservation utility. A closely related implication is that workplaces and residences face upward-sloping supply functions in real wages for workers and residents respectively (as captured in the choice probabilities (5)). To obtain additional workers, a location must pay higher wages to attract workers with lower realizations for idiosyncratic amenities for that workplace. Similarly, to acquire additional residents, a location must offer a lower cost of living to entice residents with lower realizations for idiosyncratic amenities for that residence.

5.2 Production

We assume that consumption goods are produced using a Cobb-Douglas technology, with labor and commercial floor space as inputs, and Hicks-neutral productivity differences across locations. We allow these productivity differences to be potentially endogenous to the surrounding concentration of economic activity. Cost minimization and zero profits imply that factor payments are constant shares of revenue (Y_i):

$$w_i L_i = \beta Y_i, \qquad Q_i H_i^Y = (1-\beta) Y_i, \tag{8}$$

where H_i^Y is commercial floor space use. As a result, payments for commercial floor space are proportional to workplace income ($w_i L_i$):

$$Q_i H_i^Y = \frac{1-\beta}{\beta} w_i L_i. \tag{9}$$

Each of the isomorphisms in Section D of the web appendix involves a different specification for consumption, production and trade costs that affects the determination of revenue (Y_i). We show below that our baseline quantitative approach holds regardless of which of these different specifications is used to determine revenue.

5.3 Commuter and Land Market Clearing

Commuter market clearing implies that total employment in each location (L_i) equals the number of workers choosing to commute to that location:

$$L_i = \sum_{n \in \mathbb{N}} \lambda_{ni|n}^C R_n, \tag{10}$$

where $\lambda^C_{ni|n}$ is the probability of commuting to workplace i conditional on living in residence n:

$$\lambda^C_{ni|n} = \frac{\lambda_{ni}}{\lambda^R_n} = \frac{(w_i/\kappa_{ni})^{\epsilon}}{\sum_{\ell \in \mathbb{N}} (w_\ell/\kappa_{n\ell})^{\epsilon}}. \tag{11}$$

Comparing these conditional commuting probabilities to the unconditional probabilities in equation (5), all characteristics of residence n (namely B_n, Q_n and P_n) have cancelled from the conditional probabilities in equation (11). The reason is that these residence characteristics do not vary across workplaces for a given residence. Therefore, conditional on the measure of residents in a location, the flows of commuters to each workplace depend solely on wages (w_i) and commuting costs (κ_{ni}).

We assume that floor space is owned by landlords, who receive payments from the residential and commercial use of floor space, and consume only consumption goods. As we observe positive residence employment, positive workplace employment and a single rateable value for each borough, we assume that all locations are incompletely specialized in commercial and residential activity, and no-arbitrage ensures a common price of floor space (Q_n) for both uses. Land market clearing implies that total income from the ownership of floor space equals the sum of payments for residential and commercial use:

$$\mathbb{Q}_n = Q_n H_n = (1-\alpha) v_n R_n + \left(\frac{1-\beta}{\beta}\right) w_n L_n, \tag{12}$$

where v_n is the average per capita income of the residents of a location, as determined below; H_n is the quantity of floor space; $\mathbb{Q}_n$ is the product of the price and quantity of floor space (Q_n and H_n respectively), which corresponds to the rateable value in our data.

As apparent from this land market clearing condition (12), a key property of this class of quantitative urban models is that payments for residential floor space are a constant proportion of residential income ($v_n R_n$), and payments for commercial floor space are a constant proportion of workplace income ($w_n L_n$). Additionally, per capita residential income (v_n) is a weighted average of the wages in all locations, where the weights are determined by the conditional commuting probabilities:

$$v_n = \sum_{i \in \mathbb{N}} \lambda^C_{ni|n} w_i. \tag{13}$$

Finally, the supply of floor space (H_n) depends on both geographical land area (K_n) and the density of development as measured by the ratio of floor space to land area (h_n). Following Saiz (2010), we allow the supply of floor space to respond endogenously to changes in its price:

$$H_n = h_n K_n, \qquad h_n = h Q_n^{\mu} \tag{14}$$

where h is a constant; $\mu \geq 0$ is the floor space supply elasticity; and $\mu = 0$ corresponds to the special case of a perfectly inelastic supply of floor space.

6 Baseline Quantitative Analysis

We now implement our baseline quantitative analysis for the entire class of quantitative urban models considered in Section D of the web appendix. Starting at the initial equilibrium in our baseline year in 1921, we use a combined land and commuter market clearing condition in this class of models to undertake a comparative static for removing the railway as a commuting technology. Conditioning on the observed historical changes in residence employment and rateable values back to the early-19th century, we generate predictions for the unobserved historical changes in workplace employment and commuting patterns.

Our approach uses the recursive nature of this class of models. In particular, once we condition on the observed historical changes in residence employment and rateable values, our predictions for historical workplace employment and commuting from this combined land and commuter market clearing condition do not depend on the other equilibrium conditions of the model. Therefore, we are able to generate our predictions without having to take a stand on a range of other model components, such as the costs of trading goods, whether goods are homogeneous or differentiated, whether productivity or amenities are endogenous or exogenous, the functional form and strength of agglomeration economics, the elasticity of the supply of floor space, and the reservation utility in the wider economy. Using our combined land and commuter market clearing condition, the observed historical changes in residence employment and rateable values are sufficient statistics that control for the unobserved changes in these other determinants of the spatial distribution of economic activity.

Only when we undertake counterfactuals in Section 7 are we required to make assumptions about these other model components and pick one model from this class in order to solve for a counterfactual equilibrium. Throughout the remainder of this section, we undertake our baseline quantitative analysis in a number of steps, where each step involves the minimal set of assumptions, before making additional assumptions to move to the next step.

6.1 Commuting and Employment (Step 1)

In our first step, we simply use the observed data on bilateral commuting flows (L_{nit}) from the population census in our baseline year $t = 1921$ to directly compute the following variables in that baseline year: (i) total city employment, $L_t = \sum_{n \in \mathbb{N}} \sum_{i \in \mathbb{N}} L_{nit}$; (ii) the unconditional commuting probability, $\lambda_{nit} = L_{nit}/L_t$; (iii) workplace employment, $L_{it} = \sum_{n \in \mathbb{N}} L_{nit}$; (iv) residence employment, $R_{nt} = \sum_{i \in \mathbb{N}} L_{nit}$; (v) the conditional probability of commuting to workplace i conditional on living in residence n, $\lambda^C_{nit|n} = L_{nit}/R_{nt}$.

6.2 Wages and Expected Income in the Initial Equilibrium (Step 2)

In our second step, we solve for wages (w_{nt}) and per capita residential income (v_{nt}) in the initial equilibrium in year $t = 1921$ using the observed workplace employment (L_{nt}), residence employment (R_{nt}) and rateable values ($\mathbb{Q}_{nt} = Q_{nt} L_{nt}$). We assume central values for the utility and production function parameters. In particular, we set the share of housing in consumer expenditure $(1-\alpha)$ equal to 0.3, consistent with consumer expenditure survey data (e.g. Davis and Ortalo-Magné 2011). We set the share of labor in production costs (β) equal to 0.6, in line with historical data on the labor share (e.g. Matthews, Feinstein, and Odling-Smee 1982).

Given values for these parameters, we use the land market clearing condition (12) and the expression for per capita residential income in terms of the conditional commuting probabilities (13), which together imply:

$$\mathbb{Q}_{nt} = (1-\alpha) \left[\sum_{i \in \mathbb{N}} \lambda^C_{nit|n} w_{it} \right] R_{nt} + \left(\frac{1-\beta}{\beta} \right) w_{nt} L_{nt}. \tag{15}$$

This combined land and commuter market clearing equation (15) provides a system of N equations that determines unique values for the N unknown wages (w_{nt}) in each location for our baseline year. Using these solutions for wages (w_{nt}) and the observed conditional commuting probabilities ($\lambda^C_{ni|n}$), we can immediately recover per capita residential income (v_{nt}) in each location for our baseline year from equation (13).

Lemma 1 *Suppose that ($\mathbb{Q}_{nt}$, L_{nit}, L_{it}, R_{nt}) are observed and the parameters (α, β) are known. Assuming that the rows of the observed matrix of bilateral commuting flows (L_{nit}) are linearly independent, the vector of wages (w_{nt}) for each location is uniquely determined by the combined land and commuter market clearing condition (15).*

Proof. See Section B of the web appendix. ■

Equation (15) makes explicit why our approach holds in the entire class of quantitative trade models examined in Section D of the web appendix. First, we are not required to make assumptions about whether amenities (B_{nt}) are exogenous or endogenous or the determinants of the consumption goods price index (P_{nt}) such as the costs of trading goods, because the total income of all residents ($v_{nt}R_{nt}$) is a sufficient statistic for the demand for residential floor space, where we observe residents (R_{nt}) and can solve for per capita residential income (v_{nt}) in the initial equilibrium. Second, we are not required to make assumptions about whether productivity is exogenous or endogenous or other determinants of revenue (Y_{nt}) including the costs of trading goods, because the total wage bill for all workers ($w_{nt}L_{nt}$) is a sufficient statistic for the demand for commercial floor space, where we observe employment (L_{nt}) and can solve for wages (w_{nt}) in the initial equilibrium. Finally, we are not required to make assumptions about the supply of floor space (H_{nt}), because it is captured in the observed rateable values ($\mathbb{Q}_{nt} = Q_{nt}H_{nt}$). Using the structure of this class of models, the observed variables ($\mathbb{Q}_{nt}$, L_{nit}, L_{it}, R_{nt}) are sufficient statistics to determine unique values for the unobserved wages (w_{nt}) and per capita residential income (v_{nt}) in the initial equilibrium.

6.3 Historical Predictions for Workplace Employment and Commuting (Step 3)

In our third step, we use these solutions for wages and per capita residential income in our baseline year of 1921, together with our combined land and commuter market clearing condition, to undertake our comparative static for removing the railway as a commuting technology, going backwards in time to the early-19th century.

In particular, we follow an "exact hat algebra" approach similar to that used in the quantitative international trade literature following Dekle, Eaton, and Kortum (2007), in which the relative change in a variable between an earlier year $\tau < t$ and our baseline year t is denoted by $\hat{x}_{nt} = x_{n\tau}/x_{nt}$. We begin by rewriting the land market clearing condition (12) for an earlier year $\tau < t$ in terms of relative changes ($\hat{x}_t = x_\tau/x_t$) and baseline values for each variable in year t (x_t):

$$\hat{\mathbb{Q}}_{nt}\mathbb{Q}_{nt} = (1-\alpha)\hat{v}_{nt}v_{nt}\hat{R}_{nt}R_{nt} + \left(\frac{1-\tilde{\beta}}{\tilde{\beta}}\right)\hat{w}_{nt}w_{nt}\hat{L}_{nt}L_{nt}, \tag{16}$$

where ($\hat{\mathbb{Q}}_{nt}$, $\hat{R}_{nt}$, $\mathbb{Q}_{nt}$, v_{nt}, R_{nt}, w_{nt}, L_{nt}) are observed or have been solved for; only the relative changes in expected residential income ($\hat{v}_{nt}$), employment ($\hat{L}_{nt}$) and wages ($\hat{w}_{nt}$) are unobserved.

We next show that the unobserved relative changes in expected residential income ($\hat{v}_{nt}$) and employment ($\hat{L}_{nt}$) can be written in terms of the unobserved relative change in wages ($\hat{w}_{nt}$) and observables. From equation (13), the relative change in expected residential income ($\hat{v}_{nt}$) for an earlier year $\tau < t$ must satisfy:

$$\hat{v}_{nt}v_{nt} = \sum_{i\in\mathbb{N}} \frac{\lambda^C_{nit|n}\hat{w}^{\epsilon}_{it}\hat{\kappa}^{-\epsilon}_{nit}}{\sum_{\ell\in\mathbb{N}}\lambda^C_{n\ell t|n}\hat{w}^{\epsilon}_{\ell t}\hat{\kappa}^{-\epsilon}_{n\ell t}}\hat{w}_{it}w_{it}, \tag{17}$$

where ($\lambda^C_{nit|n}$, w_{it}, v_{nt}) are observed or have been solved for; we estimate the change in commuting costs ($\hat{\kappa}^{-\epsilon}_{nit}$) below; and only the relative change in wages ($\hat{w}_{nt}$) is unobserved.

Additionally, from the commuting market clearing condition (10), the relative change in employment ($\hat{L}_{it}$) for an earlier year $\tau < t$ must satisfy:

$$\hat{L}_{it}L_{it} = \sum_{n \in \mathbb{N}} \frac{\lambda^C_{nit|n} \hat{w}^{\epsilon}_{it} \hat{\kappa}^{-\epsilon}_{nit}}{\sum_{\ell \in \mathbb{N}} \lambda^C_{n\ell t|n} \hat{w}^{\epsilon}_{\ell t} \hat{\kappa}^{-\epsilon}_{n\ell t}} \hat{R}_{nt} R_{nt}. \tag{18}$$

where ($\lambda^C_{nit|n}$, w_{it}, L_{it}, R_{nt}, $\hat{R}_{nt}$) are observed or have been solved for; we estimate the change in commuting costs ($\hat{\kappa}^{-\epsilon}_{nit}$) below; and again only the relative change in wages ($\hat{w}_{nt}$) is unobserved.

Using equations (17) and (18) to substitute for the terms in expected residential income ($\hat{v}_{nt}v_{nt}$) and employment ($\hat{L}_{nt}L_{nt}$) in our combined land and commuter market clearing condition (16) for year $\tau < t$, we obtain:

$$\begin{aligned} \hat{\mathbb{Q}}_{nt}\mathbb{Q}_{nt} &= (1-\alpha)\left[\sum_{i \in \mathbb{N}} \frac{\lambda^C_{nit|n} \hat{w}^{\epsilon}_{it} \hat{\kappa}^{-\epsilon}_{nit}}{\sum_{\ell \in \mathbb{N}} \lambda^C_{n\ell t|n} \hat{w}^{\epsilon}_{\ell t} \hat{\kappa}^{-\epsilon}_{n\ell t}} \hat{w}_{it} w_{it}\right] \hat{R}_{nt} R_{nt} \\ &+ \left(\frac{1-\tilde{\beta}}{\tilde{\beta}}\right) \hat{w}_{nt} w_{nt} \left[\sum_{i \in \mathbb{N}} \frac{\lambda^C_{int|i} \hat{w}^{\epsilon}_{nt} \hat{\kappa}^{-\epsilon}_{int}}{\sum_{\ell \in \mathbb{N}} \lambda^C_{i\ell t|i} \hat{w}^{\epsilon}_{\ell t} \hat{\kappa}^{-\epsilon}_{i\ell t}} \hat{R}_{it} R_{it}\right], \end{aligned} \tag{19}$$

where ($\hat{\mathbb{Q}}_{nt}$, $\hat{R}_{nt}$, $\lambda^C_{ni|nt}$, $\mathbb{Q}_{nt}$, v_{nt}, R_{nt}, w_{nt}, L_{nt}) are observed or have been solved for; we estimate the change in commuting costs ($\hat{\kappa}^{-\epsilon}_{nit}$) below; and only the relative changes in wages ($\hat{w}_{nt}$) are unobserved.

This combined land and commuter market clearing condition (19) provides a system of N equations that determines unique values for the N unknown relative changes in wages in each location ($\hat{w}_{nt}$).

Lemma 2 *Suppose that ($\hat{\mathbb{Q}}_{nt}$, $\hat{R}_{nt}$, L_{nit}, $\lambda^C_{ni|nt}$, $\mathbb{Q}_{nt}$, v_{nt}, R_{nt}, w_{nt}, L_{nt}) are observed or have been solved for. Given the parameters {α, β} and an estimated change in bilateral commuting costs $\hat{\kappa}_{nit}$, the combined land and commuter market clearing condition (19) determines a unique vector of relative changes in wages ($\hat{w}_{nt}$) in each location.*

Proof. See Section C of the web appendix. ∎

Using these solutions for the relative change in wages ($\hat{w}_{nt}$), we can immediately recover the unique relative change in per capita residential income ($\hat{v}_{nt}$) from equation (17). Similarly, we can solve for the unique relative change in employment ($\hat{L}_{nt}$) from the commuter market equilibrium condition (18). Finally, we can obtain the unique relative change in commuting flows ($\hat{L}_{nit}$) using the conditional commuting probabilities (11):

$$\hat{L}_{nit}L_{nit} = \frac{\lambda^C_{nit|n} \hat{w}^{\epsilon}_{it} \hat{\kappa}^{-\epsilon}_{nit}}{\sum_{\ell \in \mathbb{N}} \lambda^C_{n\ell t|n} \hat{w}^{\epsilon}_{\ell t} \hat{\kappa}^{-\epsilon}_{n\ell t}} \hat{R}_{nt} R_{nt}, \tag{20}$$

where we have solved for $\hat{w}_{it}$; we estimate $\hat{\kappa}^{-\epsilon}_{nit}$; and we observe ($L_{nit}$, $\hat{R}_{nt}$, R_{nt}, $\lambda^C_{nit|n}$).

Equation (19) again makes clear why our approach holds in the entire class of quantitative trade models examined in Section D of the web appendix. Our combined land and commuter market clearing condition uses the requirement from land market clearing that income from the ownership of floor space equals the sum of payments for residential and commercial floor space use. It also uses the implication of commuter market clearing that the measure of workers in each location equals the measure of residents choosing to commute to that location. Given the observed values of variables in the initial equilibrium and the observed historical changes in rateable values and residents, this combined land and commuter market clearing condition contains enough information to back out the implied changes in wages, employment and commuting patterns back to the early-19th century. As long as payments for residential floor space are proportional to residential income and payments for commercial floor space are proportional to workplace income, we are not required to make additional assumptions about goods' trade costs, productivity, amenities, the floor space supply elasticity, consumption goods price indexes, market structure or returns to scale in production, or the reservation level of utility in the wider economy. Using our combined land and commuter market clearing condition, the observed changes in rateable values and residents are sufficient statistics that control for unobserved changes in these other variables over our sample period.

6.4 Estimating Commuting Costs (Step 4)

In our fourth step, we estimate the impact of the railway on commuting costs using our 1921 data on bilateral commuting flows and the model's gravity equation predictions for commuting. First, we assume a relationship between commuting costs and travel times. Second, we use our maps of the transport network to construct measures of bilateral travel times between boroughs. Third, we estimate the elasticity of commuting flows with respect to these travel times, taking into account the endogeneity of the transport network using instrumental variables estimation.

For all pairs of boroughs with positive commuting flows, we assume that bilateral commuting costs are a constant elasticity function of bilateral travel times and a stochastic error:

$$-\epsilon \log \kappa_{nit} = -\epsilon\phi \log d_{nit}^{W} + u_{nt} + u_{it} + u_{nit}, \tag{21}$$

where d_{nit}^{W} is our measure of travel time; $\epsilon\phi$ is a composite elasticity of commuting flows to travel times that captures both the elasticity of commuting flows to commuting costs ($-\epsilon$) and the elasticity of commuting costs to travel times (ϕ); u_{nt} is an unobserved residence component of commuting costs; u_{it} is an unobserved workplace component of commuting costs; and u_{nit} is an unobserved component of commuting costs that is specific to a workplace-residence pair. For all pairs of boroughs with zero commuting flows, our model implies prohibitive commuting costs ($\kappa_{nit} \to \infty$), and we make this assumption to ensure that the model is consistent with the observed data.

We measure bilateral travel times by distinguishing four transport networks based on historical average travel speeds, as discussed further in Section G.6 of the web appendix: (a) overground railways (21 mph); (b) underground railways (15 mph), (c) omnibuses and trams (6 mph), and (d) walking (3 mph). Following Donaldson (2018) and Allen and Arkolakis (2014), we model travel times between boroughs as the solution to a least-cost path problem using these four transport networks. In particular, we discretize Greater London into a raster of over a million grid points, and overlay the transport network on this grid. We assign a cost or weight to traveling across each grid point based on the average travel speed for the transport network on which it is located.[36] Normalizing the cost for overground railways to one, we assume the following costs for the other modes of transport: underground railways 1.4 (=21/15); horse omnibuses 3.5 (=21/6); and walking 7 (=21/3).[37] If a grid point is located on more than one transport network, we assign it the cost for the transport network with the highest average travel speed.

Using these assumed costs for each grid point, we compute the weighted distance (which we refer to as bilateral travel time) between each pair of boroughs $n \neq i$ as the sum of the costs of traveling across the grid along the least-cost path between their centroids. To avoid making assumptions about the travel time of each borough with itself, we assume that own travel times within boroughs are unaffected by the expansion of the railway network ($\hat{\kappa}_{nnt} = 1$), which is consistent with the idea that many workers who lived and worked within the same borough may have walked or used omnibuses and trams to commute to work. To abstract from own travel times, we also

[36] The density of stations in London is high relative to the size of the boroughs used in our quantitative analysis of the model. Therefore, for simplicity, we abstract from the role of stations as points of interchange with the transport network.

[37] These assumed relative costs are close to those estimated for 19th-century rail and land transport in Donaldson (2018).

estimate the commuting cost parameters below using bilateral commuting flows to other boroughs $n \neq i$. We denote the sets of points connected to each transport network at time t by $\Im_t^{\mathrm{OR}}$, $\Im_t^{\mathrm{UR}}$, and $\Im_t^{\mathrm{OT}}$, where the superscripts OR, UR, and OT indicate overground railways, underground railways, and omnibuses and trams respectively. Similarly, we denote our vector of assumed costs for each transport network by $\boldsymbol{\delta} = [1\ \delta^{UR}\ \delta^{OT}\ \delta^{WA}]$, where the superscript WA indicates walking. Using this notation, we can write the bilateral travel times between boroughs n and i at time t as $d_{nit}^{W} = d_{nit}^{W}(\Im_t^{\mathrm{OR}}, \Im_t^{\mathrm{UR}}, \Im_t^{\mathrm{OT}}, \boldsymbol{\delta})$, where the superscript W indicates the weighting by transport mode.

Taking logarithms in equation (5), and using our assumed relationship between commuting costs and travel times in equation (21), we obtain the following gravity equation for the log unconditional commuting probability for all bilateral pairs with positive flows to other boroughs $n \neq i$ in our baseline year t:

$$\log \lambda_{nit} = \mathcal{W}_{it} + \mathcal{B}_{nt} - \epsilon\phi \log d_{nit}^{W} + u_{nit} \tag{22}$$

where the workplace fixed effect is $\mathcal{W}_{it} = u_{it} + \epsilon \log w_{it}$ and the residence fixed effect is $\mathcal{B}_{nt} = u_{nt} + \log B_{nt} - \epsilon\alpha \log P_{nt} - \epsilon(1-\alpha)\log Q_{nt}$.

A challenge in estimating equation (22) is that the transport network is endogenous, because overground and underground railways, omnibuses and trams were operated by private-sector companies. Therefore, the sets of grid points connected to each transport network ($\Im_t^{\mathrm{OR}}$, $\Im_t^{\mathrm{UR}}$, $\Im_t^{\mathrm{OT}}$) and hence the bilateral travel times (d_{nit}^{W}) are non-random. In particular, bilateral pairs that have more commuters for unobserved reasons captured by the error term (u_{nit}) could have more bilateral transport connections, and hence lower bilateral travel times (d_{nit}^{W}). To address this concern, we instrument the bilateral travel times based on the transport network (d_{nit}^{W}) using bilateral travel times from a world in which walking is assumed to be the only mode of transport (each grid point has a cost of 7 (=21/3)), so that bilateral travel times depend solely on straight-line distance. We denote these bilateral travel times in the absence of other modes of transport by d_{nit}^{S}, where the superscript S is a mnemonic for straight-line distance. Conditional on the workplace and residence fixed effects, our identifying assumption is that the unobserved factors that affect commuting in the error term (u_{nit}) are orthogonal to the straight-line distance (d_{nit}^{S}) between locations. In our empirical setting, Greater London is relatively homogeneous in terms of other economic and geographic features that could be correlated with straight-line distance, and we provide empirical checks on this identifying assumption below.

In Column (1) of the top panel of Table 2, we begin by estimating the gravity equation (22) using ordinary least squares (OLS). We find an elasticity of commuting flows with respect to travel times of around 3.02, which is statistically significant at the 1 percent level, and a regression R-squared of more than 0.8. In Column (2), we estimate the same specification using two stage least squares (2SLS), instrumenting our log bilateral travel times with log straight-line distance. We find a somewhat larger elasticity of commuting flows with respect to travel times of around 3.35, which is again statistically significant at the 1 percent level. This marginal increase in the coefficient between the OLS and IV specifications suggests that a greater incentive to invest in routes with more commuters for unobserved reasons in the error term may have been offset by other factors. In particular, the historical literature emphasizes the non-cooperative behavior of the private-sector railways, and their attempts to carve out geographical territories of dominance through a proliferation of branch lines. This struggle for areas of geographic dominance could have led to over-investment in routes that were less attractive in terms of their unobserved characteristics in the error term, thereby resulting in a marginally larger IV coefficient. As shown in the first-stage estimates reported in the bottom panel of Table 2, we find that travel time increases less than proportionately with straight-line distance (with an elasticity of 0.640), because railways reduce travel time by more over longer straight-line distances. As also shown in this bottom panel, straight-line distance is a powerful instrument for travel times, with a first-stage F-statistic well above the conventional threshold of 10. Consistent with this, we comfortably reject the null hypothesis of underidentification in the Kleibergen-Paap underidentification test reported in the top panel.

In Column (3), we provide a first check on our identifying assumption, using the idea that the relationship between travel times and straight-line distance could be non-linear in logs. In particular, conditional on there existing a railway connection between a pair of locations, railways reduce travel times by more over longer straight-line distances (as reflected in the elasticity of less than one above). However, there are likely to be fewer railway connections over longer straight-line distances, which implies that this relationship could be convex, with higher elasticities of travel times with respect to straight-line distance over longer distances.[38] To explore this idea, Column (3) augments the specification in Column (2) by using both the log straight-line distance and the square of this log straight-line distance as instruments. As reported in the bottom panel of the table, both instruments are individually statistically significant at conventional critical values, and the F-statistic for their joint significance is well above the conventional threshold of 10. As also shown in the bottom panel, we find a positive estimated coefficient on the square of log straight-line distance, consistent with a convex relationship. As this specification is now overidentified, we also report the results of the Hansen-Sargan overidentification test in the top panel, and are unable to reject the null hypothesis of the model's overidentifying restrictions.

In the remaining columns of the table, we further probe our identifying assumption. We begin by exploring the concern that there could be unobserved factors that affect bilateral commuting flows in Central London relative to other parts of Greater London. To address this concern, we define nine bilateral types of flows within Greater London based on whether the origin or destination is in the following three areas: (i) Central London (the City of London, the City of Westminster and Holborn); (ii) The rest of the County of London; (iii) the rest of Greater London. In Column (4), we augment the specification in Column (3) by including fixed effects for these nine bilateral types of flows. Even relying solely on variation in straight-line distance within each type of flow to identify the estimated coefficient on bilateral travel times, we continue to find a similar pattern of results. Both instruments remain highly significant and we again pass the Hansen-Sargan overidentification test. Finally, we examine the concern that there could be unobserved factors that affect bilateral commuting flows over very long distances relative to those over very short distances. To address this concern, Column (5) further augments the specification in Column (4) with fixed effects for quintiles of straight-line distance. Therefore, in this final specification, we identify the estimated coefficient on bilateral travel times solely from variation in straight-line distance within these quintiles and within our nine bilateral types of flows. Using this more limited variation, we find that the instruments are individually less significant, although the first-stage F-statistic for their joint significance remains well above the conventional threshold of 10, and we continue to pass the Hansen-Sargan overidentification test.

In Figure 9, we display the estimated workplace and residence fixed effects ($\mathcal{W}_i$ and $\mathcal{B}_n$ respectively) from our baseline specification in Column (3) against distance from the Guildhall. The relative value of these fixed effects provides a measure of a location's comparative advantage as a workplace or residence, after controlling for bilateral commuting costs. As apparent from the figure, the City of London and the other central boroughs have a substantial comparative advantage as a workplace, with the workplace fixed effect exceeding the residence fixed effect by up to

[38]This elasticity of travel times to straight-line distances converges to one as the number of transport connections approaches zero.

6 log points. In contrast, for boroughs more than 5 km from the Guildhall, the average residence fixed effect exceeds the average workplace fixed effect by around 0.75 log points. As reductions in commuting costs enable boroughs to specialize according to their comparative advantage as a workplace or residence, this increases real income for a given population in the model. This higher real income in turn attracts a larger population and bids up the price of floor space, until expected utility is again equal to the reservation utility in the wider economy.

As a final check on the within-sample fit of our baseline specification, Figure F.2 in Section F of the web appendix shows the conditional correlation between the log unconditional commuting probabilities to other boroughs $n \neq i$ and our estimates of bilateral commuting costs, after removing workplace and residence fixed effects. Consistent with our model's predictions, we find a tight and approximately log linear relationship between bilateral commuting flows and our estimates of bilateral commuting costs, with a conditional correlation of over 0.7. Therefore, while our parametrization necessarily abstracts from many idiosyncratic factors that could affect bilateral commuting costs, it appears to provide a reasonable approximation to observed cross-section bilateral commuting flows. We report further overidentification checks on our commuting cost estimates below, where we compare our model's predictions for workplace employment and commuting patterns back to the 19th century with historical data.

Using our estimates of the commuting cost elasticity ($\epsilon\phi$), we construct measures of the change in commuting costs from the removal of the railway network between our baseline year of $t = 1921$ and an earlier year $\tau < t$:

$$\hat{\kappa}_{nit}^{-\epsilon} = \left(\frac{\kappa_{ni\tau}}{\kappa_{nit}}\right)^{-\epsilon} = \left(\frac{d_{ni\tau}^{W}(\Im_{\tau}^{\text{OR}}, \Im_{\tau}^{\text{UR}}, \Im_{\tau}^{\text{OT}}, \boldsymbol{\delta})}{d_{nit}^{W}(\Im_{t}^{\text{OR}}, \Im_{t}^{\text{UR}}, \Im_{t}^{\text{OT}}, \boldsymbol{\delta})}\right)^{-\epsilon\phi}, \qquad n \neq i, \tag{23}$$

where $\hat{\kappa}_{nnt}^{-\epsilon} = 1$ and we assume that the unobserved components of commuting costs (u_{nt}, u_{it}, u_{nit}) are constant over time. Therefore, changes in commuting costs ($\hat{\kappa}_{nit}^{-\epsilon} \neq 1$ for $n \neq i$) are driven by changes in the set of grid points connected to the transport network ($\Im_{t}^{\text{OR}}$, $\Im_{t}^{\text{UR}}$, $\Im_{t}^{\text{OT}}$), given our assumed travel speeds ($\boldsymbol{\delta}$) and elasticity estimate ($\epsilon\phi$). We separate the elasticity of commuting flows with respect to commuting costs (ϵ) from the elasticity of commuting costs with respect to travel times (ϕ) by assuming a value for the Fréchet shape parameter of $\epsilon = 3$ following Bryan and Morten (2018) and Galle, Yi, and Rodriguez-Clare (2018), which implies an elasticity of commuting costs with respect to travel times of $\phi = 1.12$.

Using these estimates of the change in commuting costs from the removal of the railway network in equation (23), we undertake our comparative static for the removal of the transport network. In our combined land and commuter market clearing condition (19), these estimated changes in commuting costs ($\hat{\kappa}_{nit}^{-\epsilon}$) are multiplied by the conditional commuting probabilities ($\lambda_{nit|n}$) in the initial equilibrium in our baseline year of $t = 1921$. These initial conditional commuting probabilities control for unobserved differences in the initial level of commuting costs (as captured in u_{nt}, u_{it}, u_{nit}). As we remove the railway network in earlier years $\tau < t$, bilateral commuting flows necessarily remain zero for all pairs of boroughs with zero flows in our baseline year ($\lambda_{nit|n} = 0$), consistent with the more primitive commuting technology in these earlier decades.

Taking the results of this section as a whole, we find that our estimated bilateral commuting costs have substantial explanatory power for observed bilateral commuting flows, with an elasticity of around -3.35.

6.5 Overidentification Checks (Step 5)

Using these estimated commuting cost parameters (ϵ, ϕ) and our combined land and commuter market clearing condition (19), we implement our comparative static for the removal of the railway network, as discussed in Section 6.3

above. In Step 5, we now report overidentification checks, in which we compare the predictions of this class of models to separate data not used in the estimation of the commuting cost parameters.

In Figure 10, we compare the the model's predictions for day population in the City of London with the data from the City of London Day Censuses. As discussed in Section 3 above, we construct day population in the model in the same way as measured in the data, namely as equal to the model's prediction for workplace employment plus economically-inactive residents (population times one minus the constant employment participation rate for each borough). In Figure 10, we plot residence (night) population in the data using the solid gray line (without markers); we show day population in the data using the solid black line (with triangle markers); and we display day population in the model using the dashed black line (with circle markers).

Our quantitative analysis conditions on the observed historical changes in population, which implies that residential (night) population is identical in the model and data for all years. Similarly, our quantitative analysis conditions on workplace employment in the initial equilibrium in our baseline year, which implies that our model's predictions for day population in 1921 necessarily coincide with the data. However, for years prior to 1921, day population in the model and data can diverge from one another. As apparent from the figure, our model is not perfect, and slightly under-predicts day population in the closing decades of the 19th-century. Nonetheless, our model is strikingly successful in capturing the sharp trend increase in day population between the 1860s and the early decades of the 20th century. By 1831, before the construction of any railways, the model predicts a day population in the City of London only just above its night population in the data (around 136,000 compared to about 128,000), which is consistent with the idea that most people lived close to where they worked in an era when the main modes of transport were by human or horse power. Therefore, our model is successful in replicating the large-scale changes in the evolution of the day population relative to the night population in the City of London over our sample period.

In Figure 11, we compare the model's predictions for commuting patterns in the 19th-century with our data on commuting distances from the personnel ledgers of Henry Poole bespoke tailors, as discussed further in Section G.8 of the web appendix. We focus on the model's predictions for commuters into the workplace of Westminster, in which this company is located. In the left panel, we compare the model's predictions for 1861 with the commuting distances of workers who joined Henry Poole between 1857 and 1877.[39] In the right panel, we compare the model's predictions for 1901 with commuting distances of workers who joined Henry Poole between 1893 and 1914. In these comparisons, there are several potential sources of discrepancies between the model and data, including the fact that this company is located in a specific site within Westminster, whereas the model covers all of that borough. Nevertheless, we find that our model is remarkably successful in capturing the change in the distribution of commuting distances between these two time periods. In the opening decades of the railway age, most workers in Westminster in both the model and the data lived within 5 kilometers of their workplace. In contrast, by the turn of the twentieth-century, we find substantial commuting over distances ranging up to 20 kilometers in both the model and data.

From the results of both of these overidentification checks, our model is able to explain the observed reorganization of economic activity within Greater London, not only qualitatively but also quantitatively.

[39] As discussed in more detail in section G.8 of the Appendix, we use the residential address of each worker at the time they joined Henry Poole to compute their commuting distance.

6.6 Historical Employment Workplace and Residence (Step 6)

Having validated our model's predictions using historical data for the City of London, we now examine its predictions for workplace employment for the other boroughs for which such historical data are not available. In Figure 12, we display the shares of boroughs in workplace employment (black) and residence employment (gray) for 1831 (left panel) and 1921 (right panel). Whereas both variables are observed in 1921, only residence employment is observed in 1831, and workplace employment in 1831 is predicted by the model. Comparing the two panels, we find that workplace and residence employment are much closer together in 1831 than in 1921. Therefore, boroughs are much less specialized by workplace and residence in the era before the railway, consistent with short commuting distances at that time. Even in 1831, the City of London was by far the largest net importer of workers. However, its net imports of around 8,000 workers at that time are a pale shadow of those of over 370,000 workers in 1921. Whereas the City of Westminster and Holborn are both small net exporters of residents in 1831, they are both net importers of workers in 1921. Therefore, as the geographical boundaries of Greater London and its commercial center expanded outwards, the specialization of individual locations within Greater London evolved over time.

7 Counterfactuals

An important advantage of our baseline quantitative analysis is that it uses only the combined land and commuter market clearing equation and conditions on the observed historical changes in residence employment and rateable values. Using this approach, we do not need to take a stand on other components of the model, and report results that are robust across an entire class of models. Furthermore, this approach allows us to control for a range of other factors that could have affected the spatial distribution of economic activity, such as changes in goods' trade costs, productivity, amenities, the supply of floor space, and the reservation utility in the wider economy.

In this section, we consider another key question of interest, namely the counterfactual of what would have been the effect of the change in commuting technology in the absence of any of these other changes. We report two sets of counterfactuals: (i) gravity-based counterfactuals that use only the gravity equation and (ii) model-based counterfactuals, in which we solve for a counterfactual equilibrium of the model. As the gravity-based counterfactuals use only the gravity equation, they hold in the entire class of models considered in Section D of web appendix. In contrast, the model-based counterfactuals require us to choose one quantitative model from this class, in order to solve for general equilibrium effects and welfare. In practice, we find that our gravity-based and model-based counterfactuals are strongly correlated with one another. Therefore, the gravity-based counterfactuals provide a powerful diagnostic tool for policy-makers seeking to predict the impact of transport infrastructure. Furthermore, the strong predictive power of gravity alone suggests that alternative models that feature this gravity relationship are likely to exhibit relatively similar predictions. Notwithstanding this information in the gravity-based counterfactuals, we find that our model-based counterfactuals have greater explanatory power for key moments in the data, suggesting that there is additional insight from solving for equilibrium in the model.

For each of these specifications, we report counterfactuals for (a) removing the entire railway network and (b) eliminating only the underground railway network. This second underground counterfactual further isolates a pure change in commuting costs, because the underground network is exclusively used to transport people. In both cases, we hold omnibus and tram routes constant at the 1921 network structure. In principle, one could allow for changes in omnibus and tram routes in response to the removal of the railway network, although the direction of this response is unclear. On the one hand, some of the increase in commuting costs from the removal of the railway network could be offset by an expansion of the omnibus and tram network. On the other hand, these two modes of transport are imperfect substitutes, with quite different average travel speeds. Indeed, over our sample period, omnibuses and trams were largely complementary to railways, expanding in tandem with them, and being more important for shorter journeys (including from railway terminals to final destinations). Therefore, to focus on the direct effect of the removal of the railway network on commuting costs, we hold the omnibus and tram network constant.

7.1 Gravity-based Counterfactuals

Our gravity-based counterfactuals use the gravity equation to generate mechanical predictions for the effect of the removal of the railway network, in which we change commuting costs, but hold all other exogenous and endogenous variables constant at their values in the initial equilibrium. In particular, using our empirical gravity equation (22), we can rewrite the unconditional commuting probability in equation (5) as follows:

$$\lambda_{nit} = \frac{\mathbb{B}_{nt}\mathbb{W}_{it}\left(d_{nit}^{W}\right)^{-\epsilon\phi}u_{nit}}{\sum_{r\in\mathbb{N}}\sum_{\ell\in\mathbb{N}}\mathbb{B}_{rt}\mathbb{W}_{\ell t}\left(d_{r\ell t}^{W}\right)^{-\epsilon\phi}u_{r\ell t}} \tag{24}$$

where the residence fixed effect is $\mathbb{B}_{nt} = \exp\left(\mathcal{B}_{nt}\right) = \exp(u_{nt})B_{nt}P_{nt}^{-\epsilon\alpha}Q_{nt}^{-\epsilon(1-\alpha)}$ and the workplace fixed effect is $\mathbb{W}_{it} = \exp\left(\mathcal{W}_{it}\right) = \exp(u_{it})w_{it}^{\epsilon}$.

Holding these fixed effects and the error term ($\mathbb{B}_{nt}$, $\mathbb{W}_{it}$, u_{nit}) constant, we can use equation (24) to re-write the workplace and residence choice probabilities (λ_{it}^{L} and λ_{nt}^{R} respectively) after a change in commuting costs as:

$$\lambda_{it}^{L\prime} = \frac{\sum_{r\in\mathbb{N}}\lambda_{rit}\left(\hat{d}_{rit}^{W}\right)^{-\epsilon\phi}}{\sum_{r\in\mathbb{N}}\sum_{\ell\in\mathbb{N}}\lambda_{r\ell t}\left(\hat{d}_{r\ell t}^{W}\right)^{-\epsilon\phi}}, \qquad \lambda_{nt}^{R\prime} = \frac{\sum_{\ell\in\mathbb{N}}\lambda_{n\ell t}\left(\hat{d}_{n\ell t}^{W}\right)^{-\epsilon\phi}}{\sum_{r\in\mathbb{N}}\sum_{\ell\in\mathbb{N}}\lambda_{r\ell t}\left(\hat{d}_{r\ell t}^{W}\right)^{-\epsilon\phi}}, \tag{25}$$

where we use a prime to denote the level of a variable in the counterfactual equilibrium and a hat to denote the relative change in this variable between the counterfactual and actual equilibria. Therefore, λ_{it}^{L} is the actual workplace choice probability in 1921; $\lambda_{it}^{L\prime}$ is the counterfactual workplace choice probability; and $\hat{\lambda}_{it}^{L} = \lambda_{it}^{L\prime}/\lambda_{it}^{L}$ is the relative change in the workplace choice probability between the counterfactual and actual equilibria.

For each census year back to 1831, we remove the parts of the railway network constructed after that census year, and compute the implied change in commuting costs relative to our baseline year of 1921 ($\left(\hat{d}_{rit}^{W}\right)^{-\epsilon\phi}$). Using these changes in commuting costs, we generate counterfactual workplace and residence choice probabilities from equation (25). Holding the total population of Greater London constant, we use these counterfactual probabilities to generate workplace and residence employment for each borough within Greater London. In Figure 13, we display the implied net commuting into the City of London in these counterfactuals for removing the entire railway network (left panel) and eliminating only the underground railway network (right panel). The solid black line (no markers and labelled "baseline") shows net commuting into the City of London in our baseline quantitative analysis of the model in Section 6. The solid gray line (no markers and labelled "Gravity") displays the results of our mechanical predictions from equation (25). In the right panel for the underground railway, these predictions are flat before 1861, because the first underground line does not open until 1863.

In these gravity-based counterfactuals, we find substantial predicted changes in the organization of economic activity within Greater London. At the end of our sample period in 1921, observed net commuting into the City of

London stands at over 370,000 workers. In our baseline quantitative analysis from Section 6, which both changes the transport network and also allows unobserved location characteristics to evolve over time to match the observed data, these net commuting flows fall to just over 8,000 workers before the first railway in 1831 (left panel). In comparison, when we change only commuting costs and hold all other workplace and residence characteristics constant, we find that net commuting into the City of London falls to just under 30,000 workers in 1831 (left panel). This pattern of results suggests that the change in commuting costs alone, holding all else constant, can explain much of the divergence between between workplace and residence employment in the City of London from the mid-19th century onwards. Even removing only the underground network back to 1861 reduces the City of London's net imports of commuters to just under 300,000 workers (right panel), highlighting the role of the underground in enabling workers to commute into the most central parts of Greater London.

7.2 Model-based Counterfactuals

We now choose one of the quantitative models from the class considered in Section D of the web appendix to solve for the full equilibrium response to the change in the railway network. Therefore, we now take into account not only the reallocation of workers and residents across boroughs, but also endogenous changes in wages and prices and the total population of Greater London as a whole. Again we follow an "exact hat algebra" approach, similar to that used in the quantitative international trade literature. The key difference from our earlier baseline quantitative analysis in Section 6 is as follows. In our baseline quantitative analysis, we allow other unobserved location characteristics (such as productivity, amenities and the supply of floor space) to change in an unrestricted way over time to rationalize the observed changes in residence employment and rateable values in the data. In contrast, in our counterfactuals, we isolate the effect of a change in commuting costs, by either holding these unobserved location characteristics constant, or parameterizing the way in which they change with commuting costs. Therefore, the model's counterfactual predictions for residence employment and rateable values need not equal the observed values in the data.

We undertake our counterfactuals using the canonical urban model from the class of models considered in Section D of the web appendix, which is particularly tractable, and enables us to consider alternative assumptions in a transparent and flexible way. We first solve for a counterfactual equilibrium assuming that productivity, amenities and the supply of floor space are exogenous and constant. We next introduce a positive floor space supply elasticity. Finally, we further generalize the analysis to incorporate both a positive floor space supply elasticity and agglomeration forces, such that the supply of floor space, productivity and amenities all respond endogenously to changes in the organization of economic activity within Greater London. In the paper, we focus on an "open-city" specification, in which there is population mobility with the wider economy. Therefore, following the change in commuting costs, the total population of Greater London adjusts until expected utility equals the reservation level of utility in the wider economy. In Section E.3 of the web appendix, we show that the predictions of the model for specialization by workplace and residence are robust to considering a "closed-city" specification, in which the total population of Greater London is held constant, and changes in commuting costs lead to changes in worker expected utility.

7.2.1 Preferences, Production and the Supply of Floor Space

The canonical urban model imposes additional structure on the specifications of preferences, production and the supply of floor space that were introduced in Section 5 above. In preferences (3), we now add the assumption of a single homogeneous final consumption good that is costlessy traded within Greater London, such that $P_{nt} = P_t$ for all $n \in \mathbb{N}$. We allow for trade costs between Greater London and both the rest of the United Kingdom and the rest of the world, where changes in these trade costs over time are reflected in changes in the price for the final good received by firms in Greater London (P_t). We also allow residential amenities to have an endogenous component from agglomeration forces, in the form of residential externalities that depend on residence employment:

$$B_{nt} = b_{nt} R_{nt}^{\eta}, \tag{26}$$

where b_{nt} captures exogenous residential fundamentals (e.g. scenic views); and η parameterizes the strength of residential externalities. In line with evidence of the rapid spatial decay of agglomeration forces and the relatively large area of our boroughs, we assume that these residential externalities depend on own borough residence employment.

In the production technology (8), we now make the assumption that the single final consumption good is produced under conditions of perfect competition using labor and floor space:

$$P_t = \frac{1}{A_{nt}} w_{nt}^{\beta} Q_{nt}^{1-\beta}, \qquad 0 < \beta < 1, \tag{27}$$

where A_{nt} is final goods productivity.[40] We allow productivity to have an endogenous component from agglomeration forces, which take the form of production externalities that depend on workplace employment:

$$A_{nt} = a_{nt} L_{nt}^{\psi}, \tag{28}$$

where a_{nt} captures exogenous production fundamentals (e.g. access to natural water); ψ parameterizes the strength of production externalities; again consistent with existing evidence of the rapid spatial decay of agglomeration forces and the relatively large area of our boroughs, we assume that these production externalities depend on own borough workplace employment.

Using the supply of floor space (14), the price and quantity of floor space (Q_n and H_n respectively) are related to observed rateable values ($\mathbb{Q}_n$) as follows:

$$Q_{nt} = \left(\frac{\mathbb{Q}_{nt}}{hK_n} \right)^{\frac{1}{1+\mu}}, \qquad H_{nt} = (hK_n)^{\frac{1}{1+\mu}} \mathbb{Q}_{nt}^{\frac{\mu}{1+\mu}}, \tag{29}$$

which implies that the relative changes in the price and quantity of floor space are power functions of the observed relative changes in rateable values, such that $\hat{Q}_{nt} = \hat{\mathbb{Q}}_{nt}^{\frac{1}{1+\mu}}$ and $\hat{H}_{nt} = \hat{\mathbb{Q}}_{nt}^{\frac{\mu}{1+\mu}}$.

Although our baseline quantitative analysis in Section 6 does not require any assumptions about production externalities (ψ), residential externalities (η) and the floor space supply elasticity (μ), we are now required to make assumptions about these parameters in order to solve for a counterfactual equilibrium. Using our rateable values data and separate information on office rents in the City of London in the second half of the 19th century from Devaney (2010), we calibrate a floor space supply elasticity of 2.86, as discussed in Section G.9 of the web appendix.[41] For the production and residential externality parameters, we compare the model's counterfactual predictions with no agglomeration forces ($\psi = \eta = 0$), and with values for these parameters in line with the range of estimates reviewed in Rosenthal and Strange (2004) ($\psi = \eta = 0.05$).

[40] In Section D.5 of the web appendix, we show that this final goods technology (27) can be interpreted as a composite of the technologies for a tradable consumption good and non-tradeable services. London had substantial employment in both industry and services during our sample period. It was one of the main industrial centers in the United Kingdom, with manufacturing accounting for over 25 percent of employment in Greater London in the population census of 1911.

[41] This elasticity implies 69 percent growth in floor space in the City of London between 1871 and 1911, which is consistent with the estimate in Turvey (1998) that floor space in the City of London grew by at least 50 percent during this period. Consistent with London's rapid 19th-century growth predating the planning regulations that were introduced after the Second World War, this floor space supply elasticity of 2.86 is also in line with the values reported for US metropolitan areas with relatively light planning regulations in Saiz (2010).

7.2.2 Model Inversion

In our baseline quantitative analysis in Section 6, we use the observed variables to control for unobserved changes in other factors over time. Now that we have chosen one quantitative model from within the class considered in Section D of the web appendix, we can explicitly solve for unique values for the unobserved relative changes in production fundamentals ($\hat{a}_{nt}$), residential fundamentals ($\hat{b}_{nt}$) and the supply of floor space ($\hat{H}_{nt}$) across locations, as shown in Propostion E.1 in Section E of the web appendix. We use the values of the observed variables in the initial equilibrium in our baseline year to solve for *changes* in these unobserved location characteristics, where the observed bilateral commuting flows in the baseline year control for the initial level of commuting costs. Only the *relative* changes in these location characteristics are identified, because changes in the absolute level of production fundamentals across all locations enter the model isomorphically to a change in the common price of the final good across all locations. Similarly, changes in the absolute level of residential fundamentals across all locations enter the model isomorphically to a change in the common reservation level of utility for all locations.

7.2.3 General Equilibrium

General equilibrium is referenced by the four vectors of the workplace choice probability (λ^L_{nt}), the residence choice probability (λ^R_{nt}), the price of floor space (Q_{nt}), the wage (w_{nt}) and the scalar of total city population ($\bar{L}$). These five variables solve the following five equations: the workplace choice probability (λ^L_{nt} in equation (6)), the residence choice probability (λ^R_{nt}) in equation (6), the population mobility condition (7), the land market clearing condition (12), and the zero-profit condition (27). In the special case in which productivity, amenities and the supply of floor space are exogenous $\psi = \eta = \mu = 0$, there are no agglomeration forces and the supply of land is perfectly inelastic, which ensures the existence of a unique equilibrium, as shown in Proposition E.2 in Section E of web appendix. Therefore, our counterfactuals yield unique predictions for the impact of the change in commuting costs on the spatial distribution of economic activity. In the presence of agglomeration forces ($\psi > 0$ and $\eta > 0$) and an elastic supply of land ($\mu > 0$), whether or not the equilibrium is unique depends on the strength of these agglomeration forces relative to the model's congestion forces and the exogenous differences in production and residential fundamentals across locations. For the values of agglomeration forces considered in our counterfactuals, we obtain the same counterfactual equilibrium regardless of our starting values for the model's endogenous variables.

Rewriting the conditions for general equilibrium in the counterfactual equilibrium in terms of relative changes, the general equilibrium vector of counterfactual changes ($\hat{\lambda}^L_{nt}$, $\hat{\lambda}^R_{nt}$, $\hat{Q}_{nt}$, $\hat{w}_{nt}$, $\hat{\bar{L}}_t$) in response to the change in commuting costs ($\hat{\kappa}_{nit}$) solves the following system of five equations for land market clearing (30), zero-profits in production (31), workplace choices (32), residential choices (33) and population mobility (34):

$$\hat{Q}^{1+\mu}_{nt}\mathbb{Q}_{nt} = \left\{(1-\alpha)\left[\sum_{i\in\mathbb{N}}\frac{\lambda^C_{nit|n}\hat{w}^{\epsilon}_{it}\hat{\kappa}^{-\epsilon}_{nit}}{\sum_{\ell\in\mathbb{N}}\lambda^C_{n\ell t|n}\hat{w}^{\epsilon}_{\ell t}\hat{\kappa}^{-\epsilon}_{n\ell t}}\hat{w}_{it}w_{it}\right]\hat{\lambda}^R_{nt}\lambda^R_{nt} + \left(\frac{1-\tilde{\beta}}{\tilde{\beta}}\right)\hat{w}_{nt}w_{nt}\hat{\lambda}^L_{nt}\lambda^L_{nt}\right\}\hat{\bar{L}}\bar{L}_t, \tag{30}$$

$$\hat{Q}_{nt} = \hat{A}^{1/(1-\beta)}_{nt}\hat{w}^{-\beta/(1-\beta)}_{nt}, \tag{31}$$

$$\hat{\lambda}^L_{nt}\lambda^L_{nt} = \frac{\sum_{n\in\mathbb{N}}\lambda_{n\ell t}\hat{B}_{nt}\hat{w}^{\epsilon}_{\ell t}\hat{\kappa}^{-\epsilon}_{n\ell t}\hat{Q}^{-\epsilon(1-\alpha)}_{nt}}{\sum_{r\in\mathbb{N}}\sum_{\ell\in\mathbb{N}}\lambda_{r\ell t}\hat{B}_{rt}\hat{w}^{\epsilon}_{\ell t}\hat{\kappa}^{-\epsilon}_{r\ell t}\hat{Q}^{-\epsilon(1-\alpha)}_{rt}}, \tag{32}$$

$$\hat{\lambda}^R_{nt}\lambda^R_{nt} = \frac{\sum_{\ell\in\mathbb{N}}\lambda_{n\ell t}\hat{B}_{nt}\hat{w}^{\epsilon}_{\ell t}\hat{\kappa}^{-\epsilon}_{n\ell t}\hat{Q}^{-\epsilon(1-\alpha)}_{nt}}{\sum_{r\in\mathbb{N}}\sum_{\ell\in\mathbb{N}}\lambda_{r\ell t}\hat{B}_{rt}\hat{w}^{\epsilon}_{\ell t}\hat{\kappa}^{-\epsilon}_{r\ell t}\hat{Q}^{-\epsilon(1-\alpha)}_{rt}}, \tag{33}$$

$$1 = \left[\sum_{r \in \mathbb{N}} \sum_{\ell \in \mathbb{N}} \lambda_{r\ell t} \hat{B}_{nt} \hat{w}_{\ell t}^{\epsilon} \hat{\kappa}_{r\ell t}^{-\epsilon} \hat{Q}_{rt}^{-\epsilon(1-\alpha)} \right]^{\frac{1}{\epsilon}}. \tag{34}$$

where our specifications of productivity (28) and amenities (26) imply:

$$\hat{A}_{nt} = \hat{L}_{t}^{\psi}, \tag{35}$$

$$\hat{B}_{nt} = \hat{R}_{t}^{\eta}. \tag{36}$$

Recall that $\hat{\lambda}_{nt}^{L} = \lambda_{nt}^{L\prime}/\lambda_{nt}^{L}$; we observe or have solved for the initial equilibrium values of the endogenous variables ($\mathbb{Q}_{nt}$, λ_{nt}^{L}, λ_{nt}^{R}, $\bar{L}_{t}$, $\lambda_{nit|n}^{C}$, λ_{nit}, w_{it}); we assume that production and residential fundamentals, the price of the final good, and the reservation level of utility in the wider economy remain constant ($\hat{a}_{nt} = 1$, $\hat{b}_{nt} = 1$, $\hat{P}_{t}$ and $\hat{\bar{U}}_{t} = 1$); we have used $\hat{\mathbb{Q}}_{n} = \hat{Q}_{n}^{1+\mu}$; and we have chosen units in which to measure expected utility such that $\bar{U}_{t} = 1$.

We begin by solving for a counterfactual equilibrium for the special case of the model with no agglomeration forces ($\psi = \eta = 0$) and a perfectly inelastic supply of land ($\mu = 0$). In the two panels of Figure 13, we show net commuting into the City of London in these counterfactuals using the gray dashed line with circle markers (labelled "Inelastic No Agglomeration"). We find that removing the entire railway network back to 1831 reduces net commuting into the City of London to around 58,000, while removing the underground network back to 1861 reduces these net commuting flows to just under 300,000. One striking feature of the figure is the similarity between these counterfactual equilibrium predictions, in which we hold productivity, amenities and the supply of floor space constant, and the predictions of our baseline quantitative analysis (black solid line with no markers), in which we allow productivity, amenities and the supply of floor space to evolve over time to rationalize the observed data. Therefore, a first key takeaway is that much of the increased separation of workplace and residence in the late-19th century can be explained by the change in commuting costs from the new transport technology rather than by other factors.

In these model-based counterfactuals, we find somewhat smaller effects on net commuting into the City of London than in the gravity-based counterfactuals in the previous section, because we now allow wages, prices and total city population to adjust to the change in commuting costs. As a result, the removal of the entire railway network back to 1831 reduces net commuting to 58,000 compared to 29,000. Despite these differences, another notable feature of Figure 13 is the similarity between our model-based counterfactual predictions (dashed gray line with circle markets) and our gravity-based predictions (solid gray line with no markers). Therefore, a second key takeaway is for policy makers predicting the impact of future changes in transport infrastructure. Our findings suggest that policy makers can obtain a reasonable approximation to the impact of such changes in transport infrastructure using only observed commuting flows in the initial equilibrium and estimates of the change in commuting costs.

As another check on our counterfactual predictions, we compare the predicted changes in residence employment to the observed changes in residence employment in the data. We do not expect these correlations to be perfect, because our counterfactual focuses solely on a change in commuting costs, holding all else constant, whereas in the data other factors are also likely to have changed over time, as controlled for in our baseline quantitative analysis of the model in Section 6. Nevertheless, we find a strong, positive and significant correlation of 0.69 between our counterfactual predictions and the observed changes in residence employment. This strong correlation again suggests that much of the observed reorganization of economic activity in Greater London reflects the direct effect of the change in commuting costs. The corresponding correlation between our gravity-based predictions from the previous subsection and the observed changes in residence employment is 0.48, suggesting that there is additional information to be gained from using the model to solve for a counterfactual equilibrium.

Introducing a positive floor space supply elasticity ($\mu = 2.86$) magnifies the predicted impact of the change in commuting costs, as shown by the black dashed line with circle markers (labelled "Elastic No Agglomeration"). In our counterfactual for removing the entire railway network back to 1831, net commuting into the City of London falls to 23,000 (left panel), while removing only the underground network back to 1861 reduces these commuting flows to 249,000 (right panel). This pattern of results highlights a complementarity between expansions in the supply of floor space and improvements in transport infrastructure. Finally, introducing agglomeration forces ($\psi = \eta = 0.05$) further enhances the predicted impact of the change in commuting costs, as shown by the gray dashed line with triangle markers (labelled "Elastic Agglomeration"). Removing the entire railway network back to 1831 reduces net commuting into the City of London to 12,000 (left panel), while eliminating only the underground network back to 1861 reduces these commuting flows to 215,000 (right panel). The figure of 12,000 net commuters from removing the entire railway network is close to our earlier finding of 8,000 net commuters from our baseline quantitative analysis of the model, in which we allow the unobserved characteristics of locations to change to rationalize the observed changes in the data. Therefore, the model with an elastic supply of floor space and agglomeration economies can explain almost all of the sharp increase in the separation of workplace and residence in the City of London in the late-19th century following the invention of the steam railway. In both specifications with a positive floor space supply elasticity, we again find a strong correlation across boroughs between the model's counterfactual predictions for changes in residence employment from removing the entire railway network and the observed changes in the data. This correlation is 0.67 with a positive floor space supply elasticity and 0.64 with both a positive floor space supply elasticity and agglomeration forces.

In Figure 14, we examine the counterfactual predictions of the model for the total population of Greater London, where population is computed using the model's predictions for residence employment and our constant employment participation rate for each borough. Again the left panel shows results for removing the entire railway network, while the right panel gives results for eliminating only the underground network. The solid black line shows the observed population for Greater London, which rises from 1.92 to 7.39 million between 1831 and 1921. The two dashed gray lines and the dashed black line show the model's counterfactual predictions. As the removal of the railway network increases commuting costs, boroughs specialize less according to their comparative advantages as workplaces or residences, which reduces real income in Greater London for a given population. This in turn leads to a population outflow until expected utility in Greater London is again equal to the reservation level of utility in the rest of the United Kingdom.[42] In our counterfactual with an inelastic supply of floor space and exogenous productivity and amenities, we find that removing the entire railway network back to 1831 reduces the total population of Greater London to 5.15 million, as shown by the dashed gray line with circle markers (left panel).

Introducing a positive floor space supply elasticity ($\mu = 2.86$) again magnifies the predicted impact of the change in commuting costs, as the decline in the demand for floor space leads to an endogenous contraction in the supply of floor space, further reducing the attractiveness of London as a location. In our counterfactual for removing the

[42] To the extent that the overground railway network was also removed in other parts of the United Kingdom, and this was reflected in a fall in the reservation level of utility ($\bar{U}$), the effects on Greater London's population would be reduced. Nevertheless, the degree of separation of workplace and residence in Greater London is much greater than in other less-densely-populated locations, implying that Greater London would be more adversely affected by the removal of overground railways than these other locations. Furthermore, the underground railway network is largely specific to Greater London, because Glasgow is the only other city in the United Kingdom with an underground line.

entire railway network back to 1831, total population in Greater London falls to 3.10 million, as shown by the dashed black line with circle markers (left panel). Finally, introducing both a positive floor space supply elasticity ($\mu = 2.86$) and agglomeration forces ($\psi = \eta = 0.05$) further enhances the predicted impact of the change in commuting costs, as the reduction in workplace and residence employment leads to an endogenous decline in productivity and amenities. In our counterfactual for the removal of the entire railway network back to 1831, Greater London's total population falls to 2.04 million, as shown by the dashed gray line with triangle markers (left panel). Even in our counterfactuals for removing only the underground network back to 1861, we find substantial effects on the total population of Greater London ranging from 575,000 to 1.97 million (right panel). Therefore, these findings suggest that endogenous changes in the supply of floor space and agglomeration forces are important amplification mechanisms for the impact of transport infrastructure investments. For empirically-reasonable values for the floor space supply elasticity and agglomeration forces, we find that the model is capable of explaining much of the aggregate growth of Greater London.

In Section E.3 of the web appendix, we report a robustness check using a closed-city specification, in which the total population of Greater London is fixed, and the removal of the railway network leads to a decline in expected utility in Greater London. We find the model's predictions for a sharp decline in net commuting into the City of London are strikingly similar across these two different specifications. This robustness of the results is again consistent with the idea that the sharp increase in the separation of workplace and residence in late-19th century Greater London is largely driven by the direct effect of the change in commuting costs. In this closed-city specification, we find substantial changes in expected utility from the new transport technology. Across each of our specifications, we find that removing the entire railway network reduces expected utility by around 18 percent, while removing the underground network decreases expected utility by about 5 percent.

7.3 Economic Impacts Relative to Construction Cost

We now compare the counterfactual economic impact of the overground and underground railway networks with historical estimates of their construction costs. We focus on our open-city specification, in which population mobility implies that expected utility for each worker is equal to the reservation level of utility and remains unchanged following the change in commuting costs. As in the classical approach to valuing public goods using land values following George (1879), the welfare gains from the new transport technology are experienced by landlords through changes in the value of floor space. Therefore, we compute the counterfactual change in the aggregate value of floor space in Greater London from the removal of the overground and underground railway networks. As local policy makers are often concerned about broader measures of local economic activity, we also evaluate the counterfactual change in aggregate revenue or Gross Domestic Product (GDP), which equals aggregate income.

Whereas most existing analyses of transport infrastructure changes are concerned with marginal changes to an existing network (e.g. building an extra underground line), we consider a major new transport technology, which permitted the first large-scale separation of workplace and residence. Therefore a key feature of our analysis is that we model the reorganization of economic activity that occurs as locations specialize in commercial or residential activity. Another distinctive feature of our analysis relative to conventional cost-benefit approaches is that we take into account both the endogenous response of the supply of floor space and endogenous changes in productivity and amenities through agglomeration forces. To avoid making assumptions about railway operating costs, market structure, environmental externalities etc., as would be required for a full cost-benefit analysis, we focus on comparing the impact of the new transport technology on the net present value of economic activity with its construction cost.

To measure construction costs, we distinguish between shallow "cut-and-cover" underground railways, deep "bored-tube" underground railways, and overground railways. We measure the length of each of these types of railways in Greater London, classifying the parts of an underground railway company's lines that run above ground as overground railways. We measure total construction costs as the length of each type of line times the construction cost per mile for that type of line. As discussed further in Section G.10 of the web appendix, we measure construction cost per mile using historical estimates of authorized capital per mile for the private-sector companies that built these lines, which yields estimates of 555,000 per mile for bored-tube underground railways, 355,000 per mile for cut-and-cover underground railways, and 60,000 per mile for overground railways (all in 1921 prices).

In Table 3, we report the results of our comparisons of economic impact to construction costs. In the top panel, we consider the removal of the entire railway network, while in the bottom panel, we examine the elimination of the underground network. In the first column of each panel, we assume an inelastic supply of floor space ($\mu = 0$) and exogenous productivity and amenities ($\psi = \eta = 0$). In the second column of each panel, we allow a positive floor space supply elasticity ($\mu = 2.86$), while continuing to assume exogenous productivity and amenities ($\psi = \eta = 0$). In the third column, we present results incorporating both a positive floor space supply elasticity ($\mu = 2.86$) and agglomeration forces ($\psi = \eta = 0.05$). For each of these specifications, we solve for the counterfactual equilibrium and compute the change in the rateable value and revenue of Greater London as a whole. We convert these changes in the flow of rateable values and revenue into net present values assuming an infinite lifetime and 3 or 5 percent discount rates. In each panel, we also report the ratio of these reductions in the net present value of rateable value or revenue to the savings in construction costs.

As shown in Table 3, regardless of whether we assume a 3 or 5 percent discount rate, and irrespective of whether we consider the entire railway network or only the underground railway network, we find ratios of reductions in the net present value of rateable value or revenue to construction costs that are substantially greater than one. Therefore these findings suggest that the large-scale investments in the construction of the 19th century railway network in Greater London can be rationalized in terms of their effects on the net present value of economic activity. Comparing the first and second columns, we find that allowing for a positive floor space supply elasticity substantially increases the economic impact of the railway network. This finding again highlights the role of complementary expansions in the supply of floor space. Of course, these expansions themselves come at a real cost, but our analysis highlights the way in which they enhance the economic impact of the construction of the railway network. Comparing the second and third columns, we also find that incorporating agglomeration economies further magnifies the economic effects of the railway network. This finding highlights the relevance of endogenous changes in productivity and amenities for the evaluation of improvements in transport infrastructure.

These results comparing economic impact and construction costs come with a number of caveats, including our abstraction from railway operating costs, market structure, and environmental externalities etc. Nonetheless, our findings highlight the substantial economic benefits from the construction of London's 19th-century railway network, and the role of complementary expansions in the supply of floor space and agglomeration economies in magnifying these economic benefits.

8 Conclusions

We show that the separation of workplace and residence plays a central role in the concentration of economic activity in large metropolitan areas, using the revolution in transport technology from the invention of steam railways, newly-constructed spatially-disaggregated data for London from 1801-1921, and a class of quantitative urban models.

While the population of Greater London as a whole continued to grow rapidly throughout the 19th-century, we show that the population of its commercial center in the City of London fell by around 90 percent between 1851 and 1921. This precipitous decline in night population is combined with a sharp rise in day population and a marked increase in the City of London's share of the rateable value of Greater London. We provide reduced-form regression evidence connecting these changes to the new transport technology. We find no evidence of deviations in population growth rates from parish time trends before the arrival the railway and substantial deviations from these parish time trends after its arrival. We show that these deviations vary systematically by geographical location, with reductions in population growth relative to trend in the most central parts of Greater London, and increases in population growth relative to trend further from the center of Greater London.

Methodologically, we develop a new structural estimation procedure for the class of urban models characterized by a gravity equation for commuting flows. Although we only observe these bilateral commuting flows in 1921 at the end of our sample period, we show that this framework can be used to estimate the impact of the construction of the railway network going back to the early-19th century. We use a combined land and commuter market clearing condition and condition on the observed historical changes in residents and rateable values, which enables us to control for a range of other unobserved factors that could have affected the spatial distribution of economic activity. In overidentification checks, we show that our model successfully captures the sharp divergence between night and day population in the City of London in the mid-19th century onwards, and replicates the property of early commuting data that most people lived close to where they worked at the dawn of the railway age.

Having validated our model, we undertake counterfactuals for what would have been the effect of the new commuting technology on economic activity in Greater London in the absence of any other unobserved changes. Holding the supply of floor space and productivity and amenities constant, we find that removing the entire railway network reduces the total population and rateable value of Greater London by 30 and 22 percent respectively, and decreases commuting into the City of London from more than 370,000 to less than 60,000 workers. Allowing for a positive floor space supply elasticity or introducing agglomeration economies magnifies these effects. Using a calibrated floor space supply elasticity of 2.86 and elasticities of productivity and amenities to employment density of 0.05 in line with empirical estimates, we find that much of the aggregate growth of Greater London can be explained by the new transport technology of the railway. Across all of our specifications, we find increases in the net present value of land and buildings that are well above historical estimates of construction costs,

Taken together, we find that a class of quantitative urban models is remarkably successful in explaining the large-scale changes in the organization of economic activity observed during 19th-century London, and our findings highlight the role of modern transport technologies in sustaining dense concentrations of economic activity.

References

Abernathy, S. T. (2015): "Opening up the Suburbs: Workmen's Trains in London 1860-1914," *Urban History*, 42(1), 70–88.

Ahlfeldt, G., S. Redding, D. Sturm, and N. Wolf (2015): "The Economics of Density: Evidence from the Berlin Wall," *Econometrica*, 83(6), 2127–2189.

Allen, T., and C. Arkolakis (2014): "Trade and the Topography of the Spatial Economy," *Quarterly Journal of Economics*, 129(3), 1085–1140.

——— (2017): "The Welfare Effects of Transportation Infrastructure Improvements," Dartmouth College, mimeograph.

Allen, T., C. Arkolakis, and X. Li (2017): "Optimal City Structure," Yale University, mimeograph.

Ambrus, A., E. Field, and R. Gonzalez (2015): "Loss in the Time of Cholera: Long-run Impact of a Disease Epidemic on the Urban Landscape," Duke University, mimeograph.

Angel, S., and P. Lamson-Hall (2014): "The Rise and Fall of Manhattan's Densities 1800-2010," Marron Institute of Urban Management Working Paper No. 18.

Armington, P. S. (1969): "A Theory of Demand for Products Distinguished by Place of Production," *IMF Staff Papers*, 16, 159–178.

Ball, M., and D. Sunderland (2001): *An Economic History of London 1800-1914*. Routlege, Abingdon.

Barker, T. C., and M. Robbins (1963): *A History of London Transport: The Nineteenth Century*, vol. 1. George Allen & Unwin, London.

Baum-Snow, N. (2007): "Did Highways Cause Suburbanization?," *Quarterly Journal of Economics*, 122(2), 775–780.

Baum-Snow, N. (2017): "Urban Transport Expansions, Employment Decentralization and the Spatial Scope of Agglomeration Economies," University of Toronto.

Baum-Snow, N., L. Brandt, J. Henderson, M. A. Turner, and Q. Zhang (2017): "Roads, Railroads and Decentralization of Chinese Cities," *Review of Economics and Statistics*, 99(2), 435–448.

Baum-Snow, N., V. Henderson, M. A. Turner, Q. Zhang, and L. Brandt (2018): "Does Investment in National Highways Help or Hurt Hinterland City Growth?," University of Toronto.

Baum-Snow, N., and M. E. Kahn (2005): "Effects of Urban Rail Transit Expansions: Evidence from Sixteen Cities, 1970-2000," *Brookings-Wharton Papers on Urban Affaires: 2005*, 1(4), 147–197.

Billings, S. B. (2011): "Estimating the Value of a New Transit Option," *Regional Science and Urban Economics*, 41(6), 525–536.

Bogart, D. (2017): "The Turnpike Roads of England and Wales," *The Online Historical Atlas of Transport, Urbanization and Economic Development in England and Wales c.1680–1911.*

Brooks, L., and B. Lutz (2018): "Vestiges of Transit: Urban Persistence at a Micro Scale," *Review of Economics and Statistics*, forthcoming.

Bryan, G., and M. Morten (2018): "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia," *Journal of Political Economy*, forthcoming.

Caliendo, L., F. Parro, E. Rossi-Hansberg, and P.-D. Sarte (2018): "The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy," *Review of Economic Studies*, forthcoming.

Chandra, A., and E. Thompson (2000): "Does Public Infrastructure affect Economic Activity? Evidence from the Rural Interstate Highway System," *Regional Science and Urban Economics*, 30(4), 457–490.

Combes, P.-P., G. Duranton, L. Gobillon, and S. Roux (2010): "Estimating Agglomeration Economies with History, Geology, and Worker Effects," in *Agglomeration Economies*, ed. by E. L. Glaeser, chap. 1, pp. 15–66. Chicago University Press, Chicago.

Combes, P.-P., and L. Gobillon (2015): "The Empirics of Agglomeration Economies," in *Handbook of Regional and Urban Economics*, ed. by G. Duranton, J. V. Henderson, and W. C. Strange, vol. 5A, chap. 5, pp. 247–348. Elsevier North Holland, Amsterdam.

Conley, T. G. (1999): "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, 92(1), 1–45.

Croome, D. F., and A. A. Jackson (1993): *Rails Through the Clay: a History of London's Tube Railways*. Capital Transport Publishing, London.

Davis, D. R., and J. Dingel (2012): "A Spatial Knowledge Economy," NBER Working Paper, 18188.

Davis, D. R., and D. E. Weinstein (2002): "Bones, Bombs, and Break Points: The Geography of Economic Activity," *American Economic Review*, 92(5), 1269–1289.

Davis, M. A., and F. Ortalo-Magné (2011): "Housing Expenditures, Wages, Rents," *Review of Economic Dynamics*, 14(2), 248–261.

Dekle, R., J. Eaton, and S. Kortum (2007): "Unbalanced Trade," *American Economic Review*, 97(2), 351–355.

Desmet, K., D. K. Nagy, and E. Rossi-Hansberg (2018): "The Geography of Development," *Journal of Political Economy*, 126(3), 903–983.

Devaney, S. (2010): "Trends in Office Rents in the City of London: 1867-1959," *Explorations in Economic History*, 47(2), 198–212.

Donaldson, D. (2018): "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, 108(4-5), 899–934.

Donaldson, D., and R. Hornbeck (2016): "Railroads and American Economic Growth: A Market Access Approach," *Quarterly Journal of Economics*, 131(2), 799–858.

Duranton, G., P. Morrow, and M. A. Turner (2014): "Roads and Trade: Evidence from the US," *Review of Economic Studies*, 81(2), 681–724.

Duranton, G., and D. Puga (2004): "Micro-Foundations of Urban Agglomeration Economies," in *Handbook of Regional and Urban Economics*, ed. by J. V. Henderson, and J.-F. Thisse, vol. 4, pp. 2063–2117. Elsevier, Amsterdam.

Duranton, G., and M. A. Turner (2011): "The Fundamental Law of Road Congestion: Evidence from US Cities," *American Economic Review*, 101(6), 2616–52.

——— (2012): "Urban Growth and Transportation," *Review of Economic Studies*, 79(4), 1407–1440.

Eaton, J., and S. Kortum (2002): "Technology, Geography, and Trade," *Econometrica*, 70(5), 1741–1779.

Faber, B. (2014): "Integration and the Periphery: The Unintended Effects of New Highways in a Developing Country," *Review of Economic Studies*, 81(3), 1046–1070.

Fajgelbaum, P. D., and E. Schaal (2017): "Optimal Transport Networks in Spatial Equilibrium," *NBER Working Paper*, 23200.

Fogelson, R. M. (2003): *Downtown: Its Rise and Fall, 1880-1950.* Yale University Press, New Haven CT.

Foley, D. L. (1963): *Controlling London's Growth: Planning the Great Wen.* University of California Press, Berkeley.

Fujita, M., P. Krugman, and A. J. Venables (1999): *The Spatial Economy: Cities, Regions, and International Trade.* MIT Press, Cambridge MA.

Fujita, M., and J.-F. Thisse (2002): *Economics of Agglomeration: Cities, Industrial Location and Regional Growth.* Cambridge University Press, Cambridge.

Galle, S., M. Yi, and A. Rodriguez-Clare (2018): "Slicing the Pie: Quantifying the Aggregate and Distributional Consequences of Trade," University of Calirfornia, Berkeley, mimeograph.

George, H. (1879): *Progress and Poverty.* Doubleday, Page & Company, New York.

Gerhold, D. (2005): *Carriers and Coachmasters: Trade and Travel Before the Turnpikes.* Phillimore, Bodmin, Cornwall.

Gibbons, S., and S. Machin (2005): "Valuing Rail Access Using Transport Innovations," *Journal of Urban Economics*, 57(1), 148–169.

Gonzalez-Navarro, M., and M. A. Turner (2018): "Subways and Urban Growth: Evidence from Earth," Brown University, mimeograph.

Green, D. R. (1988): "Distance to Work in Victorian London: A Case Study of Henry Poole, Bespoke Tailors," *Business History*, 30(2), 179–194.

Greenstone, M., R. Hornbeck, and E. Moretti (2010): "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings," *Journal of Political Economy*, 118(3), 536–598.

Halliday, S. (1999): *The Great Stink of London: Sir Joseph Bazalgette and the Cleansing of the Victorian Metropolis*. Sutton Publishing, Stroud, Gloucestershire UK.

Helpman, E. (1998): "The Size of Regions," in *Topics in Public Economics: Theoretical and Applied Analysis*, ed. by D. Pines, E. Sadka, and I. Zilcha, pp. 33–54. Cambridge University Press, Cambridge.

Henderson, J. V. (1974): "The Sizes and Types of Cities," *American Economic Review*, 64(4), 640–56.

Jackson, K. T. (1987): *Crabgrass Frontier: The Suburbanization of the United States*. Oxford University Press, Oxford.

Kellet, J. R. (1969): *The Impact of Railways on Victorian Cities*. Routlege and Kegan Paul, London.

Kline, P., and E. Moretti (2014): "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority," *Quarterly Journal of Economics*, 129(1), 275–331.

Krugman, P. (1991): "Increasing Returns and Economic Geography," *Journal of Political Economy*, 99(3), 483–499.

Kynaston, D. (2012): *City of London: The History*. Vintage, London.

London County Council (1907): *London Statistics*. London County Council, London.

——— (1921): *London Statistics*. London County Council, London.

Lucas, R. E., and E. Rossi-Hansberg (2002): "On the Internal Structure of Cities," *Econometrica*, 70(4), 1445–1476.

Matthews, R. C., C. H. Feinstein, and J. C. Odling-Smee (1982): *British Economic Growth 1856-1973*. Clarendon Press, Oxford.

McDonald, J. F., and C. I. Osuji (1995): "The Effect of Anticipated Transportation Improvement on Residential Land Values," *Regional Science and Urban Economics*, 25(3), 261–278.

McFadden, D. (1974): "The Measurement of Urban Travel Demand," *Journal of Public Economics*, 3(4), 303–328.

Michaels, G. (2008): "The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System," *Review of Economics and Statistics*, 90(4), 683–701.

Monte, F. (2018): "The Local Incidence of Trade Shocks," Georgetown University, mimeograph.

Monte, F., S. Redding, and E. Rossi-Hansberg (2018): "Commuting, Migration and Local Employment Elasticities," *American Economic Review*, forthcoming.

Moretti, E. (2011): "Local Labor Markets," in *Handbook of Labor Economics*, ed. by D. Card, and O. Ashenfelter, vol. 4b, pp. 1238–1303. Elsevier North Holland, Amsterdam.

Owen, D. (1982): *The Government of Victorian London: The Metropolitan Board of Works, the Vestries and the City Corporation*. Belknap Press, Cambridge MA.

Porter, R. (1995): *London: A Social History*. Harvard University Press, Cambridge MA.

Redding, S. J. (2016): "Goods Trade, Factor Mobility and Welfare," *Journal of International Economics*, 101, 148–167.

Redding, S. J., and E. Rossi-Hansberg (2017): "Quantitative Spatial Models," *Annual Review of Economics*, 9, 21–58.

Redding, S. J., and D. M. Sturm (2008): "The Costs of Remoteness: Evidence from German Division and Reunification," *American Economic Review*, 98(5), 1766–1797.

Redding, S. J., and M. A. Turner (2015): "Transportation Costs and the Spatial Organization of Economic Activity," in *Handbook of Regional and Urban Economics*, ed. by G. Duration, J. V. Henderson, and W. Strange, chap. 20, pp. 1339–1398. Elsevier, Amsterdam.

Rosenthal, S. S., and W. C. Strange (2004): "Evidence on the Nature and Sources of Agglomeration Economics," in *Handbook of Regional and Urban Economics*, ed. by J. V. Henderson, and J. Thisse, vol. 4. Elsevier North Holland, Amsterdam.

——— (2008): "The Attenuation of Human Capital Spillovers," *Journal of Urban Economics*, 64(2), 373–389.

Saiz, A. (2010): "The Geographic Determinants of Housing Supply," *Quarterly Journal of Economics*, 125(3), 1253–1296.

Salmon, J. (1891): *Ten Years' Growth of the City of London*. Simkin, Marshall, Hamilton Kent & Company, London.

Shaw-Taylor, L., R. S. Davies, P. M. Kitson, G. Newton, A. E. M. Satchell, and E. A. Wrigley (2010): "The Occupational Structure of England and Wales c.1817-1881," Work in Progress. For details of the wider project see http://www.geog.cam.ac.uk/research/projects/occupations/.

Turvey, R. (1998): "Office Rents in the City of London, 1867-1910," *The London Journal*, 23, 53–67.

Voth, H.-J. (2001): *Time and Work in England 1750-1830*. Oxford University Press, Oxford.

Warner, S. B. (1978): *Streetcar Suburbs: The Process of Growth in Boston, 1870-1900*. Harvard University Press, Cambridge MA.

White, J. (2007): *London in the 19th Century: A Human Awful Wonder of God*. Vintage, Reading.

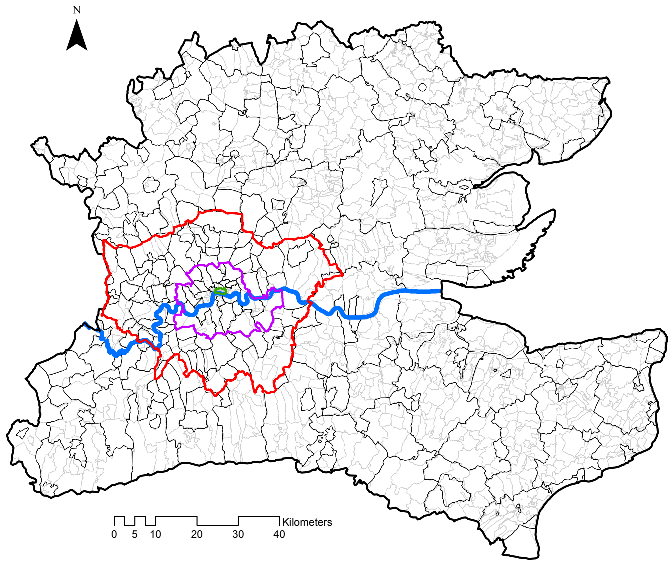
——— (2008): *London in the 20th Century: A City in the 20th Century*. Vintage, Reading.

——— (2012): *London in the Eighteenth Century: A Great and Monstrous Thing*. Bodley Head, London.

Wolmar, C. (2009): *Fire and Steam: How the Railways Transformed Britain*. Atlantic, London Books.

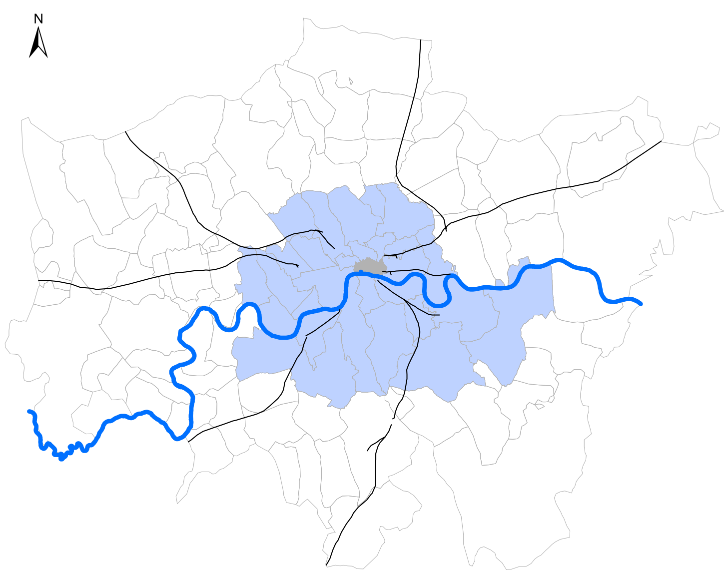
——— (2012): *The Subterranean Railway: How the London Underground was Built and How it Changed the City Forever*. Atlantic, London Books.

Figure 1: Administrative Boundaries



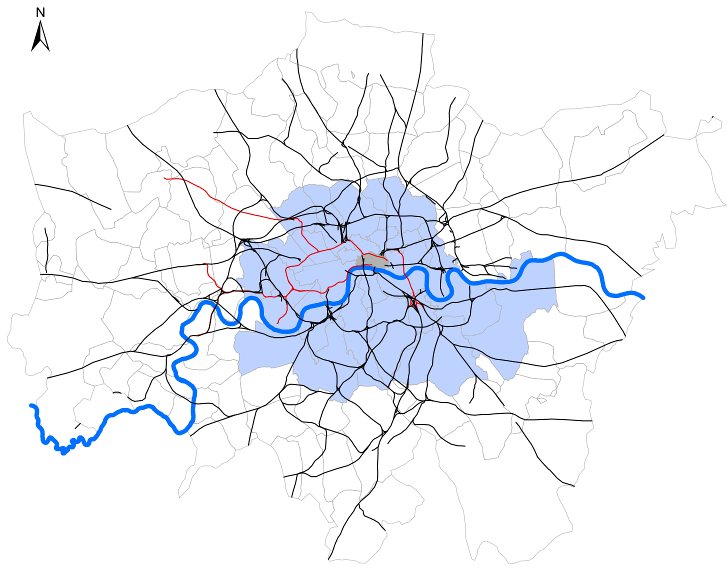
Note: Home counties surrounding London (thick black outer boundary); Greater London Authority (GLA) referred to as Greater London (red outer boundary); London County Council (LCC) (purple outer boundary); City of London (green outer boundary); River Thames (thick blue); boroughs (medium black lines); and parishes (medium gray lines).

Figure 2: Overground Railway Network in Greater London 1841



Note: Greater London outside County of London (white background); County of London outside City of London (blue background); City of London (gray background); River Thames shown in blue; overground railway lines shown in black.

Figure 3: Overground and Underground Railway Network in Greater London 1881

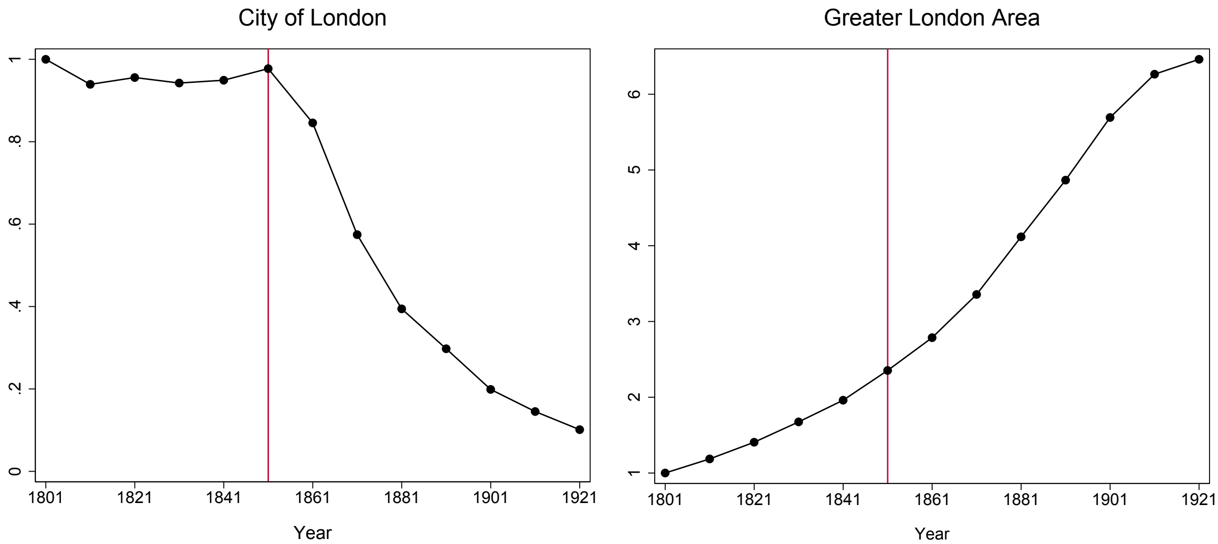


Note: Greater London outside County of London (white background); County of London outside City of London (blue background); City of London (gray background); River Thames shown in blue; overground railway lines shown in black; underground railway lines shown in red.

Figure 4: Overground and Underground Railway Network in Greater London 1921

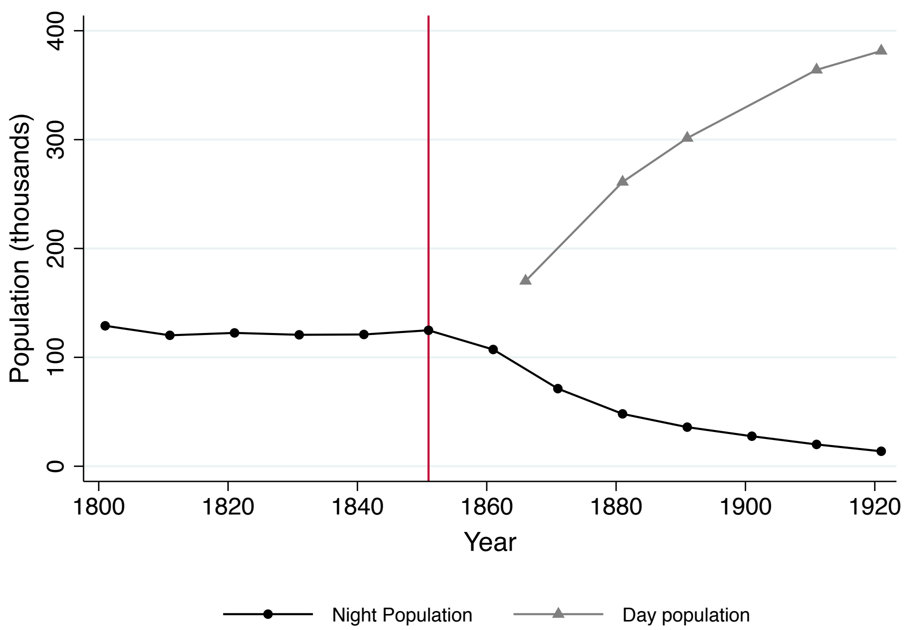
Note: Greater London outside County of London (white background); County of London outside City of London (blue background); City of London (gray background); River Thames shown in blue; overground railway lines shown in black; underground railway lines shown in red.

Figure 5: Population Indexes Over Time (City of London and Greater London, 1801 equals 1)



Note: Indices of residence (night) population from the population census over time.

Figure 6: Day and Night Population Over Time (City of London)

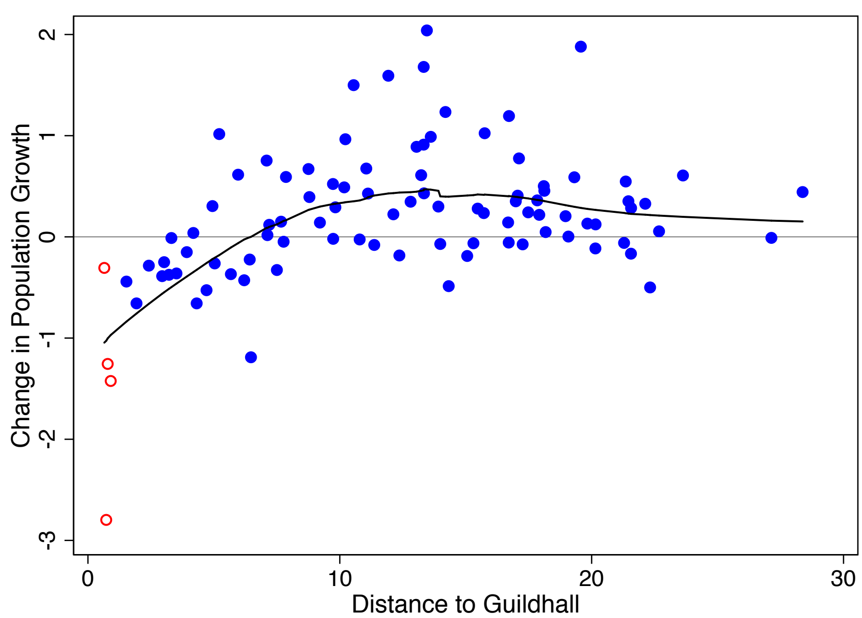


Note: Residence (night) population from the population census and day population from the City of London day census.

Figure 7: City of London's Share of Total Rateable Value in Greater London

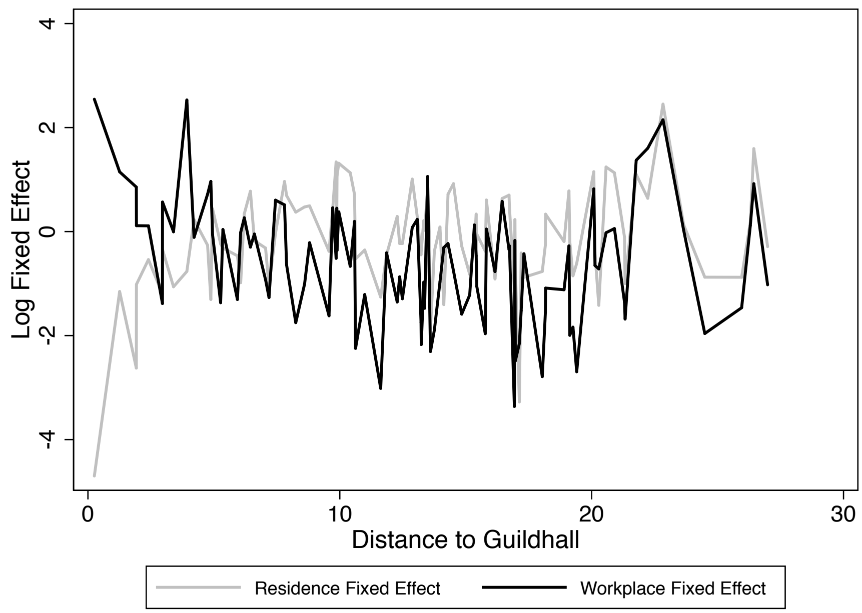
Note: City of London's Share of Total Rateable Value (value of land and buildings) in Greater London.

Figure 8: Non-parametric Railway Treatment Estimates for each Parish in Greater London



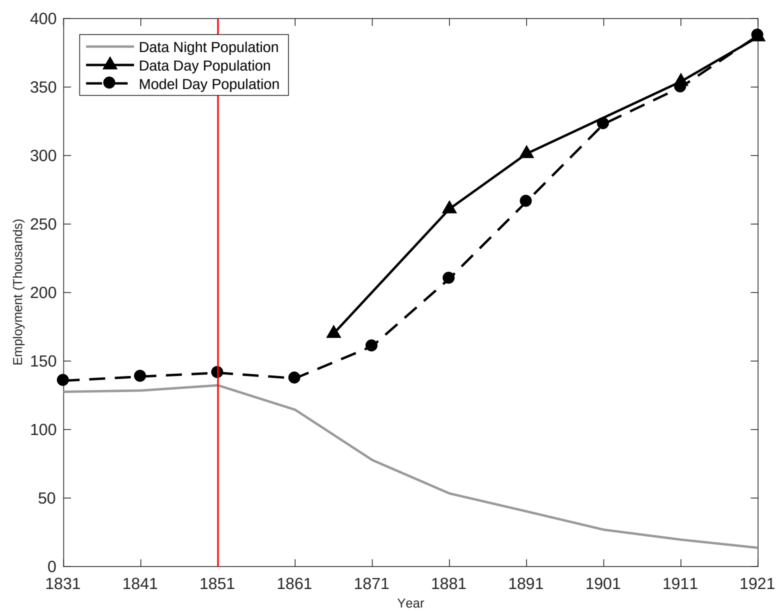
Note: Difference in mean population growth between the thirty-year periods before and after the arrival of a railway station in a parish; hollow red circles denote parishes in the City of London; solid blue circles denote parishes in other parts of Greater London.

Figure 9: Estimated Workplace and Residence Fixed Effects from Bilateral Commuting Probabilities in 1921



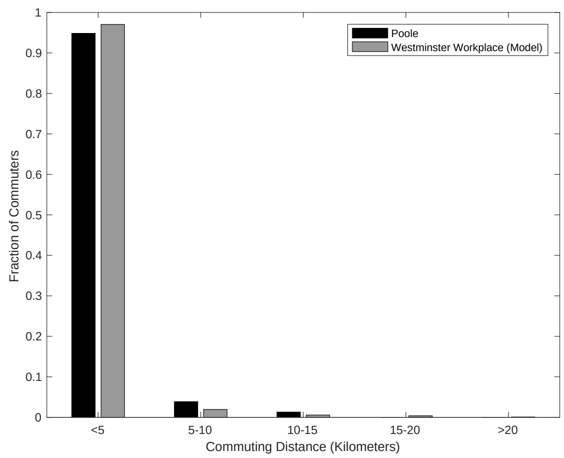
Note: Estimated workplace and residence fixed effects from the gravity equation (22) for log unconditional commuting probabilities (log λ_{nit}).

Figure 10: Night and Day Population in the City of London (Model and Data) 1831-1921

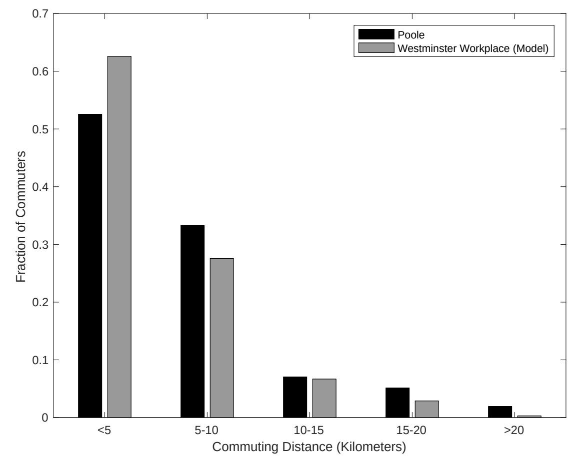


Note: "Data Night Population" is residential population from the population census, which equals residence employment plus non-economically active residents; "Data Day Population" is day population from the City of London Day Census, which equals workplace employment plus non-economically active residents; "Model Day Population" is the model's prediction for day population, which equals the model's prediction for workplace employment, plus non-economically active residents in the data, as measured by observed population divided by the employment participation rate; the model prediction's uses the estimated change in commuting costs from removing the railway network and conditions on the changes in population and rateable values in the data.

Figure 11: Commuting Distances in the Model and the Henry Poole Data



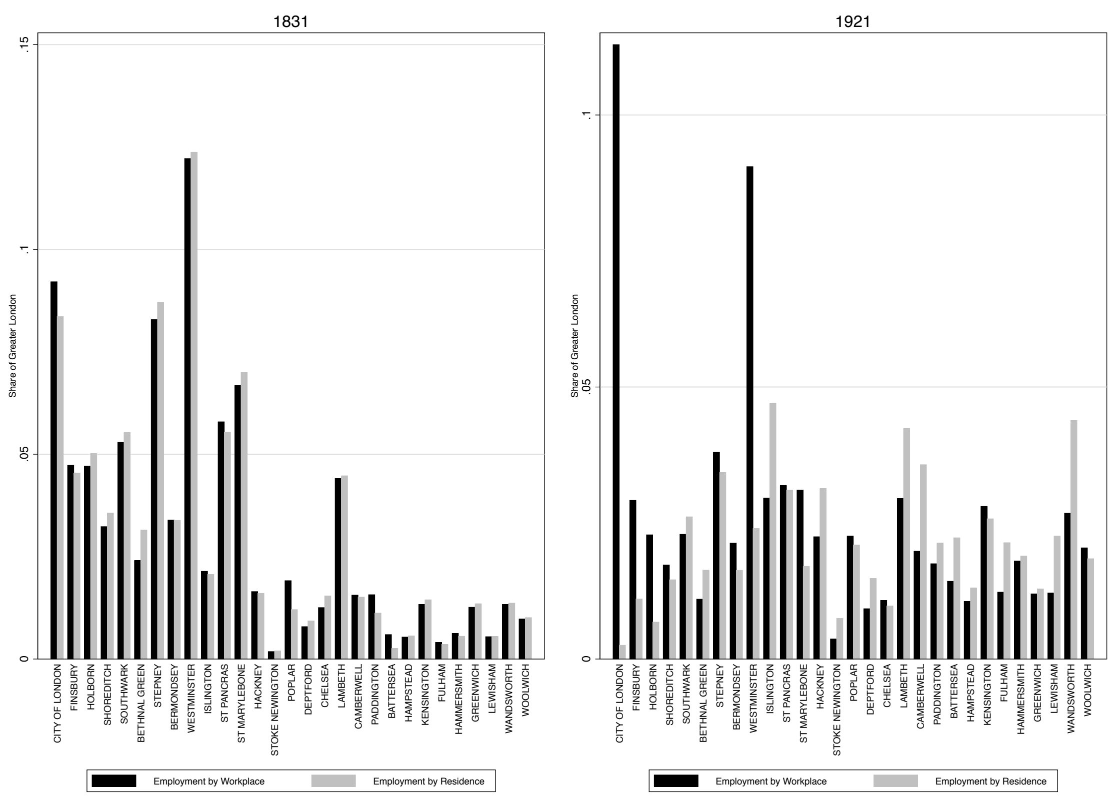
(a) 1857-1877



(b) 1893-1914

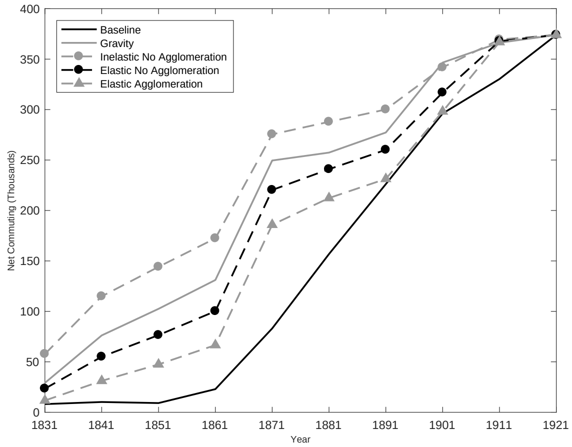
Note: Shares of workers by commuting distance for all workers employed in the borough of Westminster in the model and for workers employed by Henry Poole, Westminster. Model predictions are for 1861 and 1901. Henry Poole data are for workers hired in 1857-1877 and 1893-1914.

Figure 12: Shares of Employment in Greater London by Workplace and Residence in 1831 and 1921

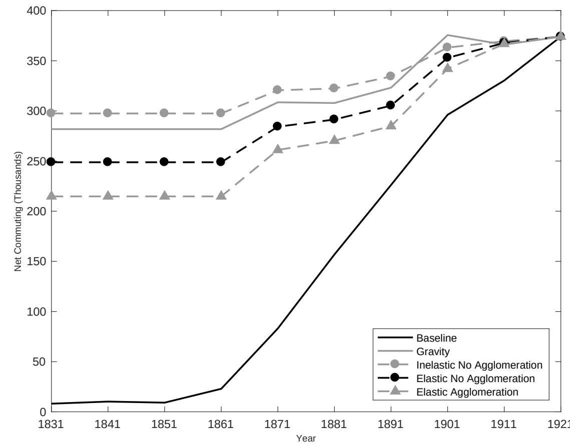


Note: Employment by workplace and residence for 1921 are observed. Employment by residence for 1831 is observed. Employment by workplace for 1831 is a prediction of the model. For legibility, we only display these shares of employment in Greater London for boroughs in the County of London. Boroughs are sorted by distance to the Guildhall in the center of the City of London.

Figure 13: Counterfactual Net Commuting into the City of London 1831-1921 (Open-City Specification)



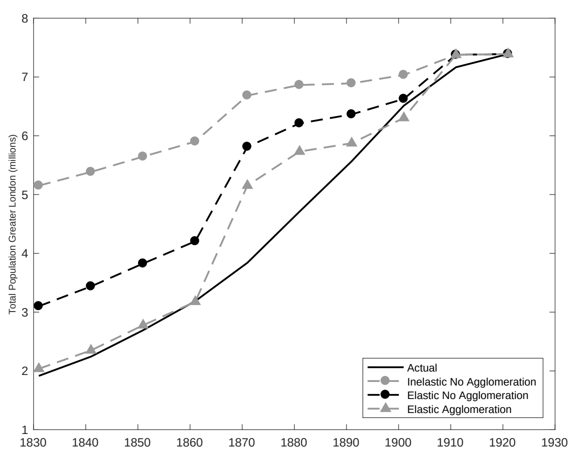
(a) All Rail Network



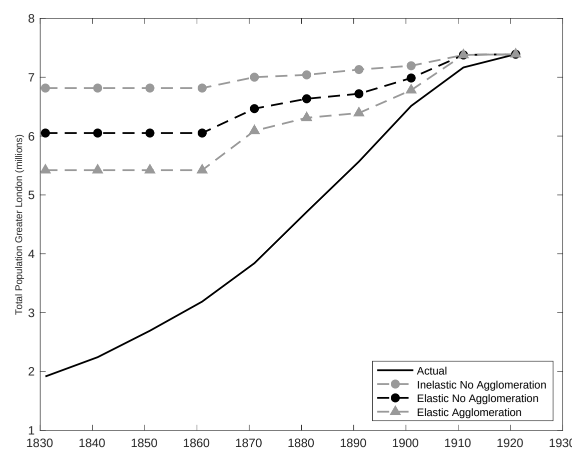
(b) Underground Rail Network

Note: "Baseline model prediction" shows net commuting from our baseline quantitative analysis from Section 6; "Gravity" shows net commuting in our gravity-based counterfactuals in Section 7.1; "Inelastic No Agglomeration" shows net commuting in our model-based counterfactual with a perfectly inelastic supply of floor space ($\mu = 0$) and exogenous productivity and amenities ($\psi = \eta = 0$) from Section 7.2; "Inelastic No Agglomeration" shows net commuting in our model-based counterfactual with a positive floor space supply elasticity ($\mu = 2.86$) and exogenous productivity and amenities ($\psi = \eta = 0$) from Section 7.2; "Elastic Agglomeration" shows net commuting in our model-based counterfactual with a positive floor space supply elasticity ($\mu = 2.86$) and positive production and residential externalities ($\psi = \eta = 0.05$) from Section 7.2. Model-based counterfactuals assume an open-city specification, with population mobility between Greater London and the wider economy.

Figure 14: Actual and Counterfactual Total Population of Greater London 1831-1921 (Open-City Specification)



(a) All Rail Network



(b) Underground Rail Network

Note: Actual is the observed population in the population census; "Inelastic No Agglomeration" shows population in our model-based counterfactuals with a perfectly inelastic supply of floor space ($\mu = 0$) and exogenous productivity and amenities ($\psi = \eta = 0$) from Section 7.2; "Inelastic No Agglomeration" shows population in our model-based counterfactuals with a positive floor space supply elasticity ($\mu = 2.86$) and exogenous productivity and amenities ($\psi = \eta = 0$) from Section 7.2; "Elastic Agglomeration" shows population in our model-based counterfactuals with a positive floor space supply elasticity ($\mu = 2.86$) and positive production and residential externalities ($\psi = \eta = 0.05$) from Section 7.2. Model-based counterfactuals assume an open-city specification, with population mobility between Greater London and the wider economy.

Table 1: Treatment Effects from Overground and Underground Railways for Greater London Parishes from 1801-1901

	(1)	(2)	(4)	(5)	(6)	(7)
	$\log R_{\ell t}$	$\log R_{\ell t}$	$\log R_{\ell t}$	$\log R_{\ell t}$	$\log R_{\ell t}$	$\log R_{\ell t}$
$\beta_{\tau=10}$	0.625***	0.114***	0.104*	0.776***	0.164**	0.149
	(0.138)	(0.041)	(0.060)	(0.251)	(0.071)	(0.091)
$\beta_{\tau=20}$	0.948***	0.262***	0.273**	1.131***	0.365***	0.378*
	(0.185)	(0.065)	(0.120)	(0.332)	(0.129)	(0.195)
$\beta_{\tau=30}$	1.327***	0.478***	0.511**	1.547***	0.632***	0.676**
	(0.256)	(0.107)	(0.210)	(0.447)	(0.223)	(0.338)
$\beta_{\tau=40}$	1.739***	0.704***	0.758**	1.909***	0.855***	0.928**
	(0.331)	(0.141)	(0.289)	(0.490)	(0.257)	(0.417)
$\beta_{\tau=50}$	2.365***	1.073***	1.154***	2.413***	1.224***	1.335**
	(0.528)	(0.223)	(0.424)	(0.586)	(0.353)	(0.561)
$\beta_{\tau=60}$	2.722***	1.326***	1.438***	2.744***	1.500***	1.654**
	(0.624)	(0.290)	(0.524)	(0.673)	(0.438)	(0.689)
$\beta_{\tau=0}$	—	—	—	—	—	—
$\beta_{\tau=-10}$			-0.048			-0.065
			(0.066)			(0.073)
$\beta_{\tau=-20}$			-0.085			-0.109
			(0.125)			(0.132)
$\beta_{\tau=-30}$			-0.092			-0.129
			(0.166)			(0.181)
$\beta_{\tau=-40}$			-0.109			-0.150
			(0.224)			(0.238)
$\beta_{\tau=-50}$			-0.084			-0.113
			(0.268)			(0.273)
$\beta_{\tau=-60}$			-0.093			-0.100
			(0.322)			(0.313)
$\gamma_{\tau=10}$				-1.205***	-0.287***	-0.299***
				(0.140)	(0.066)	(0.040)
$\gamma_{\tau=20}$				-1.789***	-0.809***	-0.833***
				(0.144)	(0.087)	(0.075)
$\gamma_{\tau=30}$				-2.658***	-1.459***	-1.501***
				(0.137)	(0.106)	(0.118)
$\gamma_{\tau=0}$	—	—	—	—	—	—
$\gamma_{\tau=-10}$						0.045
						(0.039)
$\gamma_{\tau=-20}$						-0.003
						(0.068)
$\gamma_{\tau=-30}$						0.033
						(0.085)
Parish time trends		Yes	Yes		Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Parish fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3113	3113	3113	3113	3113	3113
R-squared	0.916	0.981	0.981	0.924	0.981	0.981

Note: Observations are parishes and years; railway treatment is defined based on whether a parish has an overground or underground railway station; β_τ is the rail treatment for treatment year τ; γ_τ allows the rail treatment for treatment year τ to differ between the City of London and other parts of Greater London; τ is a treatment year indicator, which equals census year minus the last census year in which a parish had no railway, so that positive values of τ correspond to post-treatment years. For example, if the railway arrives in a parish in 1836, census year 1831 corresponds to $\tau = 0$, census year 1841 corresponds to $\tau = 10$ and census year 1821 corresponds to $\tau = -10$; standard errors are clustered on boroughs; * denotes statistical significance at the 10 percent level; ** denotes statistical significance at the 5 percent level; *** denotes statistical significance at the 1 percent level.

Table 2: Gravity Estimation Using 1921 Bilateral Commuting Data

	(1)	(2)	(3)	(4)	(5)
Second-stage Regression					
	$\log \lambda_{nit}$	$\log \lambda_{nit}$	$\log \lambda_{nit}$	$\log \lambda_{nit}$	$\log \lambda_{nit}$
$\log d^{W}_{nit}$	-3.016^{***}	-3.352^{***}	-3.352^{***}	-3.449^{***}	-3.502^{***}
	(0.043)	(0.048)	(0.048)	(0.052)	(0.087)
Workplace fixed effects	yes	yes	yes	yes	yes
Residence fixed effects	yes	yes	yes	yes	yes
Zone-pair fixed effects				yes	yes
Distance-grid-cell fixed effects					yes
Kleibergen-Paap (p-value)		0.000	0.000	0.000	0.000
Hansen-Sargen (p-value)			0.59	0.78	0.21
Estimation	OLS	IV	IV	IV	IV
Observations	3023	3023	3023	3023	3023
R-squared	0.834	-	-	-	-

	(1)	(2)	(3)	(4)	(5)
First-stage Regression					
		$\log d^{W}_{nit}$	$\log d^{W}_{nit}$	$\log d^{W}_{nit}$	$\log d^{W}_{nit}$
$\log d^{S}_{nit}$		0.640^{***}	0.318^{***}	0.297^{***}	0.317^{*}
		(0.006)	(0.112)	(0.112)	(0.185)
$\left(\log d^{S}_{nit}\right)^2$			0.023^{***}	0.024^{***}	0.023^{*}
			(0.008)	(0.008)	(0.014)
Workplace fixed effects		yes	yes	yes	yes
Residence fixed effects		yes	yes	yes	yes
Zone-pair fixed effects				yes	yes
Distance-grid-cell fixed effects					yes
First-stage F-statistic		9737.90	5624.82	3984.31	1468.84
Observations		3023	3023	3023	3023
R-squared		0.932	0.933	0.933	0.933

Note: λ_{nit} is the unconditional commuting probability from equation (5); d^{W}_{nit} is our least-cost-path travel cost measure with the following weights: overground railways 1; underground railways 1.4 (21/15); omnibus and tram 3.5 (21/6); and walking 7 (21/3); d^{S}_{nit} is straight-line distance; Kleibergen-Paap is the p-value for the Kleibergen-Paap underidentification test; Hansen-Sargen is the p-value for the Hansen-Sargan overidentification test; OLS refers to ordinary least squares; the second-stage R-squared is omitted from the instrumental variables (IV) specifications, because it does not have a meaningful interpretation; First-stage F-statistic is the F-statistic for the joint significance of the excluded exogenous variables in the first-stage regression; Heteroskedasticity robust standard errors in parentheses: * p<0.10 ** p<0.05 *** p<0.01.

Table 3: Counterfactuals for Removing the Entire Railway Network and the Underground Railway Network, Starting from the Initial Equilibrium in our Baseline Year of 1921

Removing Entire Railway Network			
Economic Impact			
	Inelastic Floor Space No Agglomeration	Elastic Floor Space No Agglomeration	Elastic Floor Space Agglomeration
Rateable Value	$-£14,215,049$	$-£35,655,748$	$-£46,572,012$
Revenue	$-£24,508,706$	$-£61,475,428$	$-£80,296,572$
NPV Rateable Value (3 percent)	$-£473,834,976$	$-£1,188,524,945$	$-£1,552,400,395$
NPV Rateable Value (5 percent)	$-£284,300,986$	$-£713,114,967$	$-£931,440,237$
NPV Revenue (3 percent)	$-£816,956,855$	$-£2,049,180,939$	$-£2,676,552,406$
NPV Revenue (5 percent)	$-£490,174,113$	$-£1,229,508,563$	$-£1,605,931,443$
Construction Costs			
Cut-and-Cover Underground		$-£9,961,911$	
Bored-tube Underground		$-£22,897,849$	
Overground Railway		$-£33,189,858$	
Total Construction Costs		$-£66,049,618$	
Ratio Economic Impact / Construction Cost			
$\frac{\text{NPV Rateable Value (3 percent)}}{\text{Construction Cost}}$	7.17	17.99	23.50
$\frac{\text{NPV Rateable Value (5 percent)}}{\text{Construction Cost}}$	4.30	10.80	14.10
$\frac{\text{NPV Revenue (3 percent)}}{\text{Construction Cost}}$	12.37	31.02	40.52
$\frac{\text{NPV Revenue (5 percent)}}{\text{Construction Cost}}$	7.42	18.61	24.31

Removing Underground Railway Network			
Economic Impact			
	Value		
Rateable Value	$-£3,768,307$	$-£11,477,548$	$-£17,608,379$
Revenue	$-£6,497,081$	$-£19,788,876$	$-£30,359,274$
NPV Rateable Value (3 percent)	$-£125,610,239$	$-£382,584,937$	$-£586,945,956$
NPV Rateable Value (5 percent)	$-£75,366,143$	$-£229,550,962$	$-£352,167,574$
NPV Revenue (3 percent)	$-£216,569,377$	$-£659,629,201$	$-£1,011,975,787$
NPV Revenue (5 percent)	$-£129,941,626$	$-£395,777,521$	$-£607,185,472$
Construction Costs			
Cut-and-Cover Underground		$-£9,961,911$	
Bored-tube Underground		$-£22,897,849$	
Total Construction Costs		$-£32,859,760$	
Ratio Economic Impact / Construction Cost			
$\frac{\text{NPV Rateable Value (3 percent)}}{\text{Construction Cost}}$	3.82	11.64	17.86
$\frac{\text{NPV Rateable Value (5 percent)}}{\text{Construction Cost}}$	2.29	6.99	10.72
$\frac{\text{NPV Revenue (3 percent)}}{\text{Construction Cost}}$	6.59	20.07	30.80
$\frac{\text{NPV Revenue (5 percent)}}{\text{Construction Cost}}$	3.95	12.04	18.48

Note: Counterfactuals start in our baseline year of 1921 and hold the omnibus and tram network constant at their values in that year. The top and bottom panels report counterfactuals for removing the entire railway network and the underground railway network respectively. We recover population in the counterfactual equilibrium from our solutions for residence employment and our constant employment participation rate for each borough. "Inelastic Floor Space No Agglomeration" is our model-based counterfactual with a perfectly inelastic supply of floor space ($\mu = 0$) and exogenous productivity and amenities ($\psi = \eta = 0$) from Section 7.2; "Inelastic Floor Space No Agglomeration" is our model-based counterfactual with a positive floor space supply elasticity ($\mu = 2.86$) and exogenous productivity and amenities ($\psi = \eta = 0$) from Section 7.2; "Elastic Floor Space Agglomeration" is our model-based counterfactual with a positive floor space supply elasticity ($\mu = 2.86$) and positive production and residential externalities ($\psi = \eta = 0.05$) from Section 7.2. All specifications assume an open-city, with population mobility between Greater London and the wider economy. Net present values are evaluated over an infinite lifetime, assuming either 3 or 5 percent discount rate. Construction costs are based on capital issued per mile for cut-and-cover, bored-tube and surface railway lines and the length of lines of each type in Greater London in 1921, as discussed further in Section G.10 of the web appendix. Pound sterling values are reported in 1921 prices.