The impact of commuting and mass transport on the London labour market: Evidence from the New Survey of London Life and Labour

Jessica Bean (Denison University)

Andrew Seltzer (Royal Holloway, LSE, IZA, CEH, ICS)

Jonathan Wadsworth (Royal Holloway, CEP, IZA, LOWER)

*** Preliminary: Please do not quote ***

We have benefitted from comments from Simon Abernathy, Bruce McClung and participants at the Economic History Society Conference in Telford, the Midwestern Economics Association Meeting in Evanston, and the Markets and States in History Conference at Warwick University. Remaining errors are our own.

The growth of public transportation networks played an important role in the development of London's labour markets in the first half of the twentieth century. Although the railways, the Underground, trams, and buses all existed in some form at the start of the twentieth century, over the next 30 years the cost of using public transport declined and there were dramatic improvements in network coverage, speed, and comfort of public transport. Writing in 1930, G. Ponserby and S. K. Ruck argued that the improvements to public transport dramatically changed the lives of working-class Londoners.[1] Improved public transport increased opportunities for shopping and leisure. More importantly, it dramatically changed labour market opportunities. In the late Victorian era, most working class Londoners lived nearby their workplace. The absence of the ability to commute to work restricted employment opportunities, reduced workers' bargaining power, and led to overcrowding in working class areas. By 1930, there were dramatic improvements to the coverage, frequency, speed, and cost of the public transport networks. This made it possible for virtually all working class Londoners to commute to work.

In this paper, we focus on the effect of public transport on London labour markets circa 1930. We examine commuting patterns and wage outcomes for working class Londoners. We use individual-level data on earnings and worker characteristics drawn from the New Survey of London Life and Labour. The New Survey was a large random sample of approximately two per cent of London's working class households covering the 28 metropolitan boroughs and 9 adjacent municipal boroughs of the city. The survey, which was conducted between 1928 and 1932, asked a range of questions on individual demographics, household structure, and labour market characteristics (occupation, earnings, and hours worked). Importantly for our purposes the New Survey also asked respondents two questions relevant to commuting: their weekly expenditure on travel and their exact home and workplace addresses. We have used *London A to Z* to obtain the GIS coordinates of workers' residences and workplaces. We have also obtained the GIS coordinate of all London Underground and railway stations using Wikipedia. We use the GIS coordinates to calculate the distance each worker in the New Survey travelled to work and the distance between workers' homes and the nearest Underground and railways stations.

Our analysis of the impact of public transport infrastructure is informed by a large literature on the effects of commuting on earnings and other labour market outcomes. This literature identifies two main mechanisms by which public transport can increase earnings. First, the ability to search over and commute to a wide geographic area implies low labour market frictions. Access to public transport made it feasible for workers to travel further in search of better job matches and higher wages. Secondly, access to public transport can increase workers' bargaining power with local employers. Commuting costs create a wedge between wages paid by an employer and those received by an employee. This wedge gives employers

1

a degree of monopsony power because employees cannot credibly threaten to take a position at more distant employers paying slightly higher wages. Increasing access to public transport or reducing its cost lowers the cost to workers of changing jobs and thus increases their bargaining power, regardless of whether they actually use public transport to commute.

In our empirical analysis we examine first the determinants of commuting and secondly worker earnings. Our approach to examining the determinants of commuting is to run a series of probit regressions on whether individual workers commute distances <1 kilometre, 1-3 kilometres, 3-7 kilometres, and 7-50 kilometres. We control for a variety of personal and neighbourhood characteristics and for access to the Underground, measured by the distance from home to the nearest Underground station. Contemporaries argued that the Underground was used primarily for middle distance commutes, with foot, bus, or tram used for shorter commutes and trains used for longer commutes. If this was the case, we would expect that being close to an underground station would increase the likelihood that an individual commutes 3-7 kilometres and correspondingly reduces the likelihood of them commuting other distances. An alternative hypothesis is that the distance to the nearest underground station is strongly correlated with unmeasured local labour market activity, and thus workers living near a station would be more likely to work locally than workers living further away.

Our approach to examining earnings is as follows. We estimate Mincer-type wage regressions on weekly earnings. As independent variables we include worker characteristics (age and its square, gender, a series of birthplace dummies, a series of occupation, hours worked), a series of workplace borough dummies, the centrality of the worker's home (distance from Charing Cross), and the commuting variables (distance commuted, distance to nearest underground and train stations, weekly expenditure on transport). The regressions have a potential endogeneity issue. Higher incomes give workers more choices of residence, and it is possible that higher income individuals chose to live closer to work. We address this problem by separately considering heads of households and secondary earners, who were much less likely to drive household location decisions.

Our findings show an important role for commuting in London's labour market. We find strong evidence that workers living near an Underground station were more likely to commute middle distances and weak evidence that they were less likely to work locally. We estimate that each kilometre commuted was associated with a 0.4-0.9 percent increase in earnings. The estimated effect for non-heads of household is approximately twice as large as that of household heads. In addition, we estimate that a one kilometre increase in the distance to the nearest London Underground station reduced earnings by on the order of one percent for heads of households, but had no significant effect on others.

The outline for the remainder of this paper is as follows. After the introduction, the second section provides a brief summary of London public transport circa 1930 and improvements since the first household survey was undertaken in the 1890s. The third section outline our theoretical approach to understanding the relationship between public transportation and earnings. The four section describes the New Survey data and the linkages to GIS data. The fifth section presents our empirical results. Finally, the sixth section concludes.

II. The London Public Transport System

The industrial revolution resulted in London becoming the world's largest city in the 19th century. In 1801 the population of London was slightly under 1.1 million, with approximately 87.5 percent residing in the inner boroughs. By 1901 the population had grown to approximately 6.5 million with approximately 69.7 percent living in the inner boroughs. By 1931 these figures were 8.1 million people and 54.2 percent. The dramatic growth of London's population led to an increase in the developed area and created an enormous demand for transportation. Consequently, over the course of the nineteenth century London developed the worlds' most comprehensive public transport network. The four main forms of public transport that we consider in this study are railways, the London Underground, trams, and buses.[2] The railway network dates back to 1836, with the opening of the London and Greenwich Railway. Horse drawn trams date back to 1860 and electrified trams to 1901. The tram network was fully electrified by the eve of the First World War. The London Underground is the oldest underground metro system in the world, with the first underground passenger trains opening in 1863 and the first deep-level electrical railway line opening in 1890. Most of the modern underground lines had been built by 1910. Horse drawn omnibuses had been used in London since 1829. Motor driven buses first appeared in 1902 and were universal by 1914. Although the London's public transportation was the most extensive in the world in the late-nineteenth and early-twentieth centuries, the networks were still patchy. Public transport was also infrequent, unreliable, slow, and expensive. As a consequence, most working class Londoners lived a short walk from their place of employment.

Between the late-nineteenth century and the early-1930s, both supply-side and demand-side factors led to dramatic increases in the number of workers who commuted. As a result of these improvements, the total number of commuters using the four main modes of public transport increased by approximately 285 percent between 1*** and 19**. The growth of the outer suburbs, meant that many Londoners had little access to local employment. The standard working day declined from ten hour to eight, leaving workers more time to commute. At the same time, a series of technological changes improved the opportunities to commute. Electrified and motorised transport replaced horse-drawn vehicles. The number of

[2] There also existed public ferries. In the 1930s ferries comprised such a small portion of travel that we can safely ignore them for our study.

track miles for the Underground and trams increased. The number of road miles covered by buses increased dramatically. The number of vehicles run over each of the networks increased by even more than the extent of coverage. The average speed of travel increased. The introduction of workmen's fares and continued technological improvements made commuting more affordable. The extent of improvements in coverage, usage, cost, and speed between 1900 and 1929 for the four different modes of transport is shown in Table 1.

An important feature of transport infrastructure was that workers' preferred mode of transport differed depending on the distance travelled. For very short distance commutes, most workers walked or cycled. A large majority (83 percent) of workers in the New Survey who resided within one kilometre of their workplace report no transport expenditures.[3] Slightly under half (46.6 percent) of those commuting between one and two kilometres report no transport expenditures. A substantial majority (89.2 percent) of those commuting over two kilometres report non-zero transport expenditures. The stark contrast between those commuting less than one kilometre and those commuting over two kilometres highlights the importance of public transport. Walking had near zero monetary costs, but was far slower. A two kilometre commute would take about 24 minutes each way, assuming the widely quoted average walking speed of five kilometres per hour. Using the speeds shown in Table 1 and assuming that a commuter had to walk 500 metres from home to the bus and from the bus to their workplace place and that the average waiting time for the bus was five minutes, a bus would save about five minutes each way over a distance of 2 kilometres, and ** minutes over a distance of 4 kilometres.

Short-distance commuters (between 1 and about 3-5 kilometres) typically used either the buses or trams. Ponsoby and Ruck argue that these modes of transport were used fairly interchangeably, depending on convenience. Buses and trams tended to cost less than the other forms of transport on a per trip basis.[4] However, they also made more frequent stops and had to share the roads with other users, and were thus considerably slower than the trains or underground. The undergrounds were the primary form of medium-distance travel. The trains were typically used for longer distance travel. Virtually, all workers in the New Survey catchment area lived within a few hundred metres of a bus or tram stop. By contrast, the Underground and railways relied on very capital-intensive infrastructure investment and coverage was far from universal. The near-universal access to short-distance transport has important implications for our estimation strategy. We are not able to identify the returns to access to buses/trams because of the absence of variation in our data. This implies that our

[3] The true figure is likely to be higher, as a disproportionate number of these respondents have missing values for transport expenditures. We believe that it is likely that a large number of missing values correspond to either commuting by bicycle (and thus no expenditure) or to the original enumerator entering "-" instead of "0".

[4] The cost figures in table 1 are on a per-mile rather than a per-trip basis. Ponserby and Ruck, Travel, p. ** report the distribution of fares for the four different modes of travel. *******

estimated returns to access to the Underground and rail networks should be interpreted as additional returns, relative to having access to buses and trams only.

III. A Simple Model of Commuting

This section develops a simple framework modelling the decision of where to work for an individual with a fixed residence, and discuss how we implement this framework empirically using the NSLLL data.

Figure 1 provides a simple diagrammatic treatment of the worker's commuting problem and the impact of public transport. As a starting point, consider the case where the worker's residence is fixed; i.e. they cannot move home in response to changes in the labour market. This assumption makes it easy to highlight the impact of public transportation and is perhaps reasonable for the short-run, although highly problematic in the long-run, even for a housing market with severe supply constraints, such as inter-war London.[5] We address the possibility of joint home and workplace location decisions later. The horizontal axis represents physical location, with the origin being the worker's home and moves to the right reflecting increasing commuting distance. The vertical axis is the net wage earned by the worker, e.g. wage – commuting costs (where commuting costs include both monetary and time costs).[6] The net wage curves for skilled and unskilled workers are shown by W_S and W_U, respectively. For simplicity, we assume that all employers pay the same money wages for workers of a given skill level, thus the slope of the net wage curve is determined by commuting cost, which is a function of the distance between home and work. Note that a worker does not have the choice of working at every point on the horizontal axis, only those where employers are actually located.

The worker's reservation (net) wage is W*, below this, he chooses not to work. The reservation wage is heterogeneous across workers, and is typically higher for skilled workers, W_S^* in figure 1, than unskilled workers, W_U^*, as skilled workers are likely to be more productive at home as well as in the workplace.

We begin by considering the unskilled and skilled workers' problem of where to work, given a fixed residence. Assuming that jobs are homogeneous in their attributes other than distance from home, an unskilled worker maximizes their net wage by taking the closest available unskilled job or, if there are no jobs available inside distance d1, by not working. Unlike the

[5] Llewellyn Smith and Marsh, p. * discuss the housing shortage faced by working class Londoners.
[6] We use the term "commuting costs" to refer to any costs associated with distance between work and home. Much of the literature on spatial labour markets has focussed on search costs, which are also assumed to increase with distance (Cramton, 1999; van Ommeron, et. al., 1997; and Wasmer and Zenou, 2006).

unskilled worker, a skilled worker can choose to take either a skilled job or a closer unskilled job at a lower wage. In the absence of public transport, the skilled worker maximizes their net wage by taking the closest skilled job between home and d4. If there are no skilled jobs available in this range, they then compare available skilled jobs between d4 and d5 and unskilled jobs between home and d0. If there are no unskilled jobs inside distance d0 or skilled jobs inside distance d5, they choose not to work.

Building an Underground line along the space represented by the horizontal axis lowers the cost of commuting and thus shifts up the net wage function. This is shown with the new wage curves W_U' and W_S'.[7] If a worker is employed close to their home, they will still choose to walk (or take the bus or tram) to work, and so the presence of an Underground line does not change their net wage or influence their decision about where to work. If they are employed further from home, the Underground reduces their commuting costs (and thus increases their net wage) at a rate that increases with the distance travelled. Unskilled workers now earn more than their reservation net wage at jobs located between d1 and d3. Skilled workers now earn more than their reservation wage at unskilled jobs between d0 and d2 and at skilled jobs between d5 and d6. Note that the effect of the Underground line is direction-specific, the net wage curves in different directions from home do not change in response to the building of a new line.

The information needed to estimate the wage curves in figure 1 is readily available from the NSLLL. The intercepts of the W_S and W_U curves (Wm_S and Wm_U) are simply the monetary wages received by skilled and unskilled workers, respectively. This is reported in the data for virtually all working individuals. The slopes of these curves are determined by the cost of transport. The NSLLL also contains information about the monetary costs of travelling to work for a large portion of the sample. We assume that the non-monetary costs of transportation are a linear function of commuting time, and we can construct plausible estimates of these costs using commuting distance and accessibility of railway and Underground stations.

The approach described above yields two predictions that are testable in the NSLLL data.

1) Workers residing near public transportation will commute greater distances and have higher monetary wages than those living further from a station. Workers residing near a station will spend less time walking to the station, and thus have lower time costs of commuting to a given destination. This implies higher net earnings for a given level of

[7] The Underground stops at discrete intervals and thus there are additional costs incurred from walking from the Underground stop to the workplace. Graphically, this means that the curves W'_U and W'_S should be step functions with local maxima at the Underground stations, and decreases in net wages with increasing distance from the station. This does not affect any of our analysis above and for expediency, we represent the curves as straight lines.

monetary earnings. It also implies that there are some distant commutes that will be undertaken by workers residing near a station that would not be undertaken by a worker living further from the station. This can be seen in figure 1, where living further from the station (and thus facing a higher time cost of walking to the station) can be represented by a downwards parallel shift of the W' curve. The worker with the shorter walk to the station will have a higher W' curve which, by definition, intersects with the W* curve further from home. Thus there will be some relatively long commutes that they will be willing to make that an otherwise identical worker residing further from the station would not make. The higher average monetary earnings associated with proximity to a station is a result of the fact that some skilled workers who live close to a station would be willing to commute to a distant skilled job whereas an otherwise similar worker living further from the station will take a local unskilled job.

2) Skilled workers will be more willing to commute, and thus there will be a positive relationship between skill and commuting distance. In addition, the wages of skilled workers will be more sensitive to proximity to public transport. One reason for the greater willingness of skilled workers to commute is that wages are higher in the skilled sector, and thus the worker will be willing to incur greater commuting costs in order to attain a skilled job. In Figure 1, if we assume that skilled and unskilled workers face the same commuting costs and same reservation wage, the higher wage earned by skilled workers implies that the W curve intersects the W* curve further from the residence of the worker, implying a willingness to commute greater distances.[8] The literature in regional economics and international economics provides a second, supply-side, reason why skilled workers typically commute further. Skilled jobs are likely to be geographically highly concentrated due to economies of agglomeration, for example due to knowledge spillovers (Marshall, 1920; Romer, 1986; and Krugman, 1991). Unskilled jobs, on the other hand, are likely to be geographically disperse, as they typically have little or no economies of agglomeration, and thus unskilled workers are likely to work close to their home.

The impact of public transport, as discussed above, will be to induce some workers to switch to more distant jobs. The set of possible effects for skilled workers includes switching from unemployment to remote skilled jobs, switching from unemployment to relatively remote unskilled jobs, switching from local unskilled jobs to remote skilled jobs, and switching from jobs that are not convenient to public transport to jobs that are. Unskilled workers also may switch from unemployment to work or from jobs that are not convenient to public transport to jobs that are, although these effect are likely to be relatively small if unskilled employment is geographically disperse.

[8] Note that in figure 1, the reservation net wage is higher for the skilled worker. Nevertheless, the mechanism described above is still relevant if the difference in wages between skilled and unskilled workers is greater than the difference in the reservation wage plus any differences in the opportunity cost of time spent commuting.

One problem with the approach described above is that it implicitly assumes perfect competition in the labour market, and thus the only effect of public transport is to increase net wages in proportion to distance actually travelled. However, Bhaskar and To (1999); Rotemberg and Saloner (2000); and Bhaskar, Manning, and To (2002) argue that labour markets may be characterised by localised monopsony power created by search frictions, such as commuting distance. Firms are discretely located in space, and for each worker some employers will be located much closer to home than others. The existence of commuting costs means that there is a difference between the wage paid by the employer and the net wage received by the worker, and this difference will be a function of distance. In Figure 1 this is simply the vertical distance between W_{MS} (or W_{MU}) and the relevant net wage curve. There may exist large differences in the net wage between the nearest employer and next closest distant alternative, even if all employers pay the same money wage. Consequently, a wage decrease at the closer employer or a wage increase at the more distant employer will not prompt a worker to switch jobs unless the change is sufficiently large to make up for higher commuting costs. This wedge between wages paid and net wages received gives employers localised bargaining power, and the ability to pay less than the market wage without losing all of their workers. This implies that the supply curve faced by individual firms is upward sloping, and wages are less than marginal factor cost.

An implication of monopsonistic competition is that public transport has two separate effects on labour markets. First, it has the standard competitive market effect shown in figure 1; namely reducing the cost of commuting and thus increasing the net wage at more distant employers. Secondly, public transport reduces the extent of monopsony power that firms possess over local workers, as its existence lowers the cost of search and commuting. In other words, an increase in the worker's the ability to commute reduces the wedge between wage and net wage and between wage and marginal factor cost, and thus decreases employers' ability to pay less than the market wage. This second impact of public transport is not a function of the actual distance commuted by individual workers, as the net wages of even a worker who has no commute may increase as a result of the existence of public transport. This suggests that the appropriate way to measure the impact of public transportation on wages is by its effect on *ability to commute*, rather than distance actually commuted. In the empirical analysis that follows, we measure ability to commute using the distance between workers' homes and the nearest London Underground station. One concern with this approach is that Underground station locations are not randomly distributed and are likely to be near centres of economic activity. In other words, proximity to the underground may captures unmeasured *local* job opportunities. In the empirical section we test for this possibility by examining whether proximity to the underground increased the likelihood that a worker was employed locally.

A second concern about the model shown in Figure 1 is the potential endogeniety of residential location. If salaries in some areas are higher than in others, there will be a tendency for workers to move to the high-wage areas over time. A correlation between wages

and commuting distance could be driven by the ability of wealthier workers to locate closer to their jobs, rather than by the sort of optimisation described above. Urban economists have typically modelled place of residence and place of work as a joint optimization decision constrained by the fixed costs of changing jobs and moving home (Crampton, 1999; Kain, 1962; and Steinnes, 1977). This decision has typically been empirically estimated using systems of simultaneous equations examining moves between residences and jobs. However, examining moves requires multiple observations of the same individuals, and panel surveys were not first undertaken until well after the period of our study.

The possibility of reverse causality in the wage/commuting distance relationship implies that to obtain unbiased estimates of the impact of commuting it is necessary for us to be able identify cases where place of residence can plausibly be described as exogenous. Following Kain (1962), O'Reagan and Quigley (1993), and Rees and Shultz (1970) we assume that the home/work location decision revolves around the primary earner in the household, and thus place of residence is exogenous for secondary earners. In the empirical analysis we define secondary earners in two distinct ways. First, the New Survey directly identifies the head of each household and thus everyone not listed as the head can be classified as a secondary earner. Secondly, we identify secondary earners using the earnings data, defining them to be everyone other than the highest earner within a household.

IV. Data

Our data is drawn from the New Survey of London Life and Labour (NSLLL), a household survey overseen by the statistician Arthur Lyon Bowley, conducted between 1928 and 1932.[9] The sample for the NSLLL comprised approximately 2 percent of London working class households, defined by a maximum income of £250.[10] The New Survey was intended as a follow up to the 1886-1903 Survey into Life and Labour in London (SLLL), and thus many of the features of the NSLLL, such as its coverage, income limits, and information recorded, were designed to ensure compatibility with the earlier survey. Unlike the SLLL, most of the original records of the NSLLL have survived largely intact, and were encoded in the 1990s by the team of Roy Bailey, Dudley Baines, Timothy Hatton, Paul Johnson, Anna Leith, and Angela Raspin.[11] All totalled, the computerised records contain 26,915 households, 94,137 individuals, and 49,445 income earners.

[9] Bailey and Leith, Computerising and coding and Bailey and Leith, New Survey.
[10] The original survey also included middle class and Jewish households. These records were not systematically encoded and are not part of our data.
[11] The original cards from the boroughs of Walthamstow and Tottenham have been lost and thus these boroughs are not part of the computerised records. The adjacent boroughs of Leyton and Hornsey comprise slightly over 3 percent of the NSLLL sample and thus it is likely that the share of records lost was fairly small.

The coded records of the NSLLL contain considerable individual-level and household-level information. As our research focusses on labour markets, our main objective is to explain the determinants of earnings. This paper uses a variety of information similar to that used in other studies of earnings, including: age, gender, place of birth, relationship to head of household, occupation, industry, and hours worked. The NSLL also contains household information that is absent in many data sets. For our purposes, the most important of this information pertains to location. The data contain exact street addresses for places of residence and employment. We have used Google Maps to obtain GIS coordinates for these addresses. About * percent of addresses are no longer in existence, in most cases due to catastrophic damage from Second World War bombings. In these cases, we have been able to use historical street directories to obtain an approximate location for ** percent of these observations. We have used Wikipedia to obtain GIS coordinates for all extant Undergrounds stations in 1931.[12]

The sample includes a total of 49,445 individuals listed as wage earners. However, only 26,235 of these observations contain earnings in the previous week. Our final sample also excludes workers who are missing information about hours worked (1,133 individuals), sex (23 individuals), occupation (42 individuals), or workplace area (2,831 individuals). We also excluded income earners with a reported age of 0 (2,824 individuals) and those who lived more than 50 kilometres from Charing Cross (5 individuals) or who commuted more than 100 miles to work (16 individuals). As some individuals fall into more than one exclusion category, we are left with a usable sample of 20,106 individuals.

To assess the impact of commuting on London labour markets, it is necessary to include individual-level commuting information in our analysis. The NSLLL data contains one direct measure of commuting, namely weekly expenditure on public transportation. We believe that this is likely to be a fairly noisy measure of commuting for several reasons. First, many respondents did not supply easily quantifiable answers and there are a lot of missing observations for this variable.[13] Secondly, it is unclear whether reported transport expenditures included only commuting costs.[14] Thirdly, even if it were possible to resolve these data-related issues, the cost variable does not capture either the time costs of commuting or identify the separate effects of commuting and ability to commute.

---

[12] https://en.wikipedia.org/wiki/List_of_London_Underground_stations and https://en.wikipedia.org/wiki/List_of_former_and_unopened_London_Underground_stations.

[13] When coding transport costs we have handled missing observations in two different manners. Our first approach is to simply leave all missing values as missing, and drop the observation from the analysis. Our second approach is to recode missing values to zero if the individual's residence and workplace were less than a kilometre apart. We generally prefer this approach, as relatively most individuals (about 83 per cent) who resided within a kilometre of their workplace reported zero transport costs. The results reported in the next section use the latter approach, but the results excluding all missing values are qualitatively similar.

[14] The space for travel expenses lies of the work section of the enumerators' cards, and thus it is clear that the expenses were meant to be for commuting to work. In practice, we believe that virtually all reported transport expenditures in the data were work related. A substantial majority (81.5 percent) of reported positive transport expenditures were divisible by 12 pence and virtually all (96.4) were divisible by 6 pence, as would be expected if all reported travel costs were for round trips between work and home. Nevertheless, it is likely that there is measurement error in the data due to inclusion of non-work-related transport expenditures.

Because of these issues with transport expenditures, the main indicators of commuting in most of our analysis is based on linear distances. We compute commuting distances using the home and work addresses in the original data.[15] We compute each worker's *ability to commute* as the distance between their home and the nearest London Underground station. Future versions of this paper will also include distance to the nearest railway station, as this was the main alternative to the Underground for longer distance travel.[16] We do not compute distances to bus and tram stops, because virtually all London workers had nearby access to at least one of these.[17] This universal access implies that the NSLLL data is likely to contain little variation in distances to the nearest bus or tram stop. Finally, we compute each worker's proximity to the centre of London as the distance between their residence and Charing Cross Station.[18]

We have used the GIS coordinates to calculate these three distances for each worker using the Great Circle Distance formula:

$$d = R*\arccos(\sin(lat\_a)*\sin(lat\_b) + \cos(lat\_a)*\cos(lat\_b)*\cos(d\_lon))$$

where R = Radius of the earth (6365 km)

lat_a, lat_b = Latitudes of the two points,

d_lon = Difference in longitude between the two points.

Table 2 shows some summary statistics about commuting. The overall impression from table 2 is that, consistent with descriptive evidence in the companion volumes, commuting was an important influence on the London labour market during the inter-war period. The first six rows provide summary statistics about distance commuted. The mean distance in the sample is just under four kilometres. Approximately 20 percent worked within a kilometre of home,

[15] We expect that distance commuted and transport costs will be closely correlated. In the NSLLL data the correlation is a strong, but far from perfect 0.51.

[16] Ponsoby and Ruck, Travel, pp. 186-8 argue that short-distance commuters used buses and trams interchangeably based on local access and route. Middle distance commuters tended to use the Underground, which was considerably faster than either buses or trams. Long distance commuters generally used the trains.

[17] Ponsoby and Ruck, Travel contains several insert maps showing the networks for each of the different modes of public transport. It is evident from these maps that there existed comprehensive bus and tram networks throughout the NSLLL catchment area. Very few individuals in our sample would have been more than a few hundred meters away from a bus or tram route. This universal access implies that, unlike the Underground and trains, there is likely very little variation in distances to the nearest bus or tram stop the NSLLL data. In addition, obtaining data on the locations of bus and tram stops is much more difficult than for the trains or Underground, as the stops did not have the same large scale supporting infrastructure. Moreover, buses, unlike the other forms of transport, did not travel along fixed routes. Routes and stops frequently were changed in response to changes in demand. The absence of fixed routes means that any attempt to obtain GIS coordinates would require matching dates of routes to dates of New Survey interviews.

[18] Charing Cross Station is the only main train station in central London. It is also only a few meters away from the original Charing Cross, which is used by Ordinance Survey as the geographic centre of the city.

54 percent worked within three kilometres of home, and 92 percent worked within ten kilometres of home. While we are unaware of data on analogous figures for earlier periods, the companion volume to the New Survey argues that in the late-nineteenth century it was the norm to live very close to work.[19] The next five rows show statics on the centrality of residences and workplaces. As might be expected, workplaces were on average closer to central London than residences, implying a net inward commute. This is confirmed in the next five rows, which show the direction commuted (i.e. the difference in centrality between home and work). While commuting towards the city centre was the norm, about half as many workers "reverse commuted" away from the centre. Reverse commuting was particularly prevalent for workers in a handful of industries that were disproportionately located in the outer boroughs of the survey area such as metals; electrical apparatus; food, drink, and tobacco; wood and furniture; and construction-related trades. Finally, the last three rows show transportation expenditures. Approximately 63.5 percent of income earnings in the NSLLL data who reported expenditures, reported non-zero values.[20] Expenditures averaged about 19.5 pence per week across all individuals and 30.7 pence per week among those reporting non-zero costs.

V. Empirical Results

The underlying empirical approach to estimating the effects of commuting on wages is the Mincer regression (Mincer, 1958 and 1974). The Mincer approach models wages as a function of worker and workplace attributes. Specifically, we estimate regressions of the form:

$$\ln(W_i) = a + BX_i + b_1 D_{W,i} + b_2 D_{CX,i} + b_3 D_{U,i} + e_i$$

The dependent variable in our regressions is the natural log of the weekly wage (in hundredths of pence); X is a vector of control variables (age, age squared, sex [male=1], weekly hours worked, whether they were born outside the United Kingdom [1=yes], whether they were born in the British Isles outside of England [1=yes], whether London born [1=yes], whether living in the same borough where they were born [1=yes], whether living in an adjacent borough to to the one in which they were born [1=yes], and, in some specifications, dummies for occupation, skill level, and workplace location). Table 3 shows summary statistics for the regression variables separated by head of household status and the expected signs of the regression coefficients. Most of the control variables are fairly standard in wage regressions, and we expect the standard relationships between the controls and earnings; i.e.

[19] Llwellyn-Smith, pp. 6-7 and Ponsoby and Ruck, Travel, pp. 189-91. Also see Booth, Life and Labour, p. 568.
[20] This increases to 72 per cent if we exclude the recoded observations where transport expenditures are missing and the worker commutes less than one kilometre.

wages will increase at a decreasing rate with age, will be higher for men than women, will increase in hours worked, and will be higher in skilled jobs than others.

The birthplace dummies are used to capture the effects of migration on earnings. We use the migration status variables as controls in the regressions, but are guided by a long literature on the effects of migration on earnings. Broadly speaking, the literature has identified three separate effects of migration status on earnings: selection, discrimination, and local knowledge. An extensive literature (see Roy, 1951 and Abromitzky et. al., 2012) argues that migrants may be either positively or negatively selected depending on the dispersion of wages and rewards to skills in the source and destination locations. A second effect is discrimination, which is likely to lead to lower earnings for migrants. Finally, local labour market knowledge and connections are also likely to lead to lower earnings for migrants. Because these factors operate simultaneously, the overall effect of migration status on earnings is theoretically ambiguous.

Our primary interest is the coefficients on the three distance variables: D_W, distance commuted to work; D_{CX}, distance from home to Charing Cross Station (a measure of centrality); D_U, distance from home to the nearest London Underground station. The model presented in Section II implies that the coefficient on D_W will be positive, in other words, individuals will commute further only if they have a higher paying job to offset the cost of commuting. We expect the coefficient on D_{CX} to be positive, in other words workers outside of central London will earn more all else equal. This is counterintuitive in a modern context; however, during the 1930s, working class jobs in inner London were, on average, lower quality and lower paid than those in the city outskirts. In the NSLLL sample, wages were, on average, about 5 per cent higher outside a four kilometre radius from Charing Cross than within the radius (565.3 vs. 536.6 pence per week). Moreover, outside the four kilometre radius a higher proportion of jobs were "skilled" under the Armstrong classification (52.5 vs. 46.2). We expect the coefficient on D_U to be negative, because, as noted in Section II, being located close to an Underground station would have increased workers' possible outside options and thus increased their bargaining power with their current employers. In some specifications we added two additional independent variables related to commuting: the interaction of D_{CX} and D_U and the natural log of the (monetary) cost of transport. The interaction of D_{CX} and D_U captures the effect of different levels of job density on worker's gains from commuting (or threatening to commute).[21] Individuals located in areas with a high density of local jobs did not have to commute large distances to exercise their outside option. In other words, access to the Underground is likely to have been more important further from

[21] We calculated job density as follows. First, we calculated the distance to Charing Cross for each workplace using the GIS coordinates. Second, we calculated the number of sample jobs in each one kilometre wide "ring", centred on Charing Cross. Third, we summed up the total number of jobs in central (within four kilometres of Charing Cross) and outer London (4-10 kilometres from Charing Cross). Fourth, we divide the number of jobs by the area of the ring (about 50 km^2 for the centre and 264 km^2 for the outer ring). Job density peaked in the ring 1-2 kilometres from Charing Cross and steadily declined further out. We calculate that there were approximately 7 times as many jobs per square kilometre within 4 kilometres of Charing Cross as beyond.

the city centre where there were fewer local jobs. Thus we expect the interaction of D_{CX} and D_U to have a negative coefficient. We also include transportation costs in addition to distances in one specification. We expect that the coefficient on ln(transport costs) will be positive, as our model predicts that workers will only spend more on transport if they are able to access higher-wage jobs.

We ran several specifications of the regression model to ensure the robustness of the main results.[22] To address the potential endogeneity of workplace and residential location, we ran the regressions separately for the full sample and split by head of household status. As noted in section 2, it is likely that the choice of residence was driven much more by the workplace of primary income earners than others. Home location is plausibly exogenous for secondary earners within households, and thus we expect distance to work to have a larger coefficient for non-heads than heads. A second reason for running separate regressions for heads and non-heads is that bargaining power may be very different across these two groups. A large literature on wage bargaining has shown that both willingness to bargain and bargaining power is greater for relatively high income earners, males, and older workers.[23] Individuals identified as heads of households were much more likely to be in each of these categories, and thus to have been more responsive to the existence of outside options.[24] Thus we expect a larger negative coefficient on D_U for heads than for non-heads. As outlined in our model in Section II, one possible consequence of access to public transport is that workers may switch from a nearby low-skilled job to a more distant skilled job. This implies that occupation and skill-level are endogenous outcomes of commuting, rather than simply control variables as in the specification above. Thus as an additional robustness check, we have run our regressions excluding occupation, skill level, and workplace location controls. Because this specification effectively allows switching to higher paid jobs to be an outcome of commuting, we expect that the coefficient on the distance commuted will be greater in this specification.

The regression results are reported in Table 4.[25] The first specification covers the full sample and includes dummy variables for occupation and place of work. The second through fifth specifications split the sample into heads of household and non-heads. In the second and third specification the head of household is defined directly from the Survey. In the fourth and fifth

[22] We ran one additional set of robustness tests that are not reported in the paper. Abernathy (2016) raises concerns about one interviewer, G. E. Bartlett, who interviewed 18.5 per cent of the households in the data, about 5 times as many as the next most prolific reviewer. Abernathy argues that this level of activity is implausibly high, and suggests that Bartlett may have subcontracted out some of his questionnaires or even filled them in himself. We have run all of our regressions omitting workers interviewed by Bartlett. The results are substantively the same as those reported, although, in some specifications, the estimated effects of commuting are slightly larger than those reported in the paper.

[23]

[24] Relative to non-heads, heads of households on average earned 98 percent more (£151.44 per year vs. £76.62), were nearly twice as old (42.4 years vs. 22.1 years), and were twice as likely to be male as non-heads (92.5% vs. 46.9%).

[25] Due to space constraints we have not reported the coefficients for the occupation or workplace location dummies. A full set of regression results is available from the authors.

specifications, head of household is defined as the highest income earner in the household.[26] The remaining specifications present additional robustness checks. The sixth specification drops the occupation and workplace location dummies. The seventh specification replaces the occupation dummies with dummies for Armstrong skill classification (professional occupations are the omitted category). The eighth specification adds the interaction of D_{CX} and D_U as an independent variable. Finally, the ninth column adds ln(transport cost).

The regressions are strongly significant overalland the coefficients on the control variables all have the expected signs. The coefficients on these variables are also fairly robust to model specification. Except for the second specification (heads of households), living in the same (or adjacent) borough to the borough of birth has a positive effect on earnings. This suggests that either migrants were negatively selected or the effects of local networks or knowledge outweighed other factors.

The coefficients on the distance variables are largely consistent with our model and suggest an important role for commuting and public transport in the operation of London labour markets. The coefficient on the distance from Charing Cross is positive in all specifications, although not significant for non-heads of household. Even after controlling for personal characteristics and for occupation, workers in the city centre had lower earnings than those on the outskirts. The coefficient on distance commuted to work is positive and strongly significant in all specifications. As would be expected, the size of the coefficient on distance increases when the occupation and workplace location dummies are excluded from the regression and decreases when the pecuniary costs of transport is included in the regression. The coefficient on distance to the Underground is negative and statistically significant in the regressions on the full sample (when the interaction of D_{CX} and D_U is excluded). When the interaction of D_{CX} and D_U is included in the regression, the coefficient on D_U becomes insignificant, but the coefficient on the interaction is negative and strongly significant. This suggests that proximity to public transport only mattered for outer London residents, who had fewer local job opportunities.

The relative size of the coefficients on D_W and D_U in the regressions split by head of household status (specifications 2, 3, 4, and 5) reveals an interesting difference in the determinants of pay. For heads of household, both coefficients have the expected signs and are statistically significant. For non-heads, the coefficient on D_U is statistically insignificant, while the coefficient on D_W is strongly significant and more than twice the size of the coefficient in the heads of household regression. While we cannot formally test the reasons behind these differences with the available data, one plausible interpretation concerns different willingness to use voice/exit mechanisms to obtain higher wages. In other words,

[26] The correlation between the two definitions of head of household is 0.67.

heads are relatively more likely to bargain with a nearby employers for higher wages, whereas non-heads are more likely to commute in search of higher pay.

Table 1
Public transportation statistics for 1900, 1913, and 1929

		1900	1913	1929
Local and Underground railways	Route miles	95.2	110.6	120.6
	Train miles (1,000,000s)	5.01	18.45	24.13
	Passengers (1,000,000s)	214.5	474.7	648.8
	Average fare per mile (in 1930 pence)	1.143	0.989	0.790
	Average scheduled speed (in kilometres per hour)	24.1	28.8	31.1
Mainline railways	Route miles	561.2	558.2	534.1
	Passengers (1,000,000s)	233	250	415
Buses	Route miles	300	467	1,170
	car miles (1,000,000s)	42.0	110.0	215.7
	average scheduled speed (in kilometres per hour)	8.0	13.7	15.3
	passengers (1,000,000s)	264.5	735.7	1,912.1
	Average fare per mile (in 1930 pence)	1.879	1.061	0.960
	average seats per vehicle	23	34	50
Tramways	miles of roadway	221.7	350.3	345.5
	car miles (1,000,000s)	47.9	95.9	104.3
	average scheduled speed (in kilometres per hour)	N.A.	14.2	16.1
	passengers (1,000,000s)	340.2	812.1	1,076.3
	Average fare per mile (in 1930 pence)	1.062	0.899	0.680
	average seats per vehicle	38	67	67

Sources: Ponserby and Ruck, p. 194 and Munby, p. 537.

Notes: The figures for the first column are from 1900, 1905, 1906, and 1907. See Ponserby and Ruck, p. 194 for details. Figures for fares are reported in current prices and converted into 1930 prices using O'Donoghue, Goulding, and Allen. The reported scheduled speed for the London Underground is for the Metropolitan and District Line.

Table 2
Some Commuting Statistics

	Mean	**St. Dev.**
Distance between work and home	3.96	4.48
% commuting < 1 km	20.6	
% commuting 1-3 km	33.6	
% commuting 3-5 km	18.2	
% commuting 5-10 km	20.0	
% commuting 10+ km	7.6	
Distance between home (work) and Charing Cross	4.6 (5.2)	3.7 (2.4)
% Living (working) < 1 km to Charing Cross	1.0 (4.7)	
% Living (working) 1-3 km to Charing Cross	17.2 (36.3)	
% Living (working) 3-5 km to Charing Cross	39.0 (22.4)	
% Living (working) 5-10 km to Charing Cross	38.6 (31.3)	
Distance commuted in	0.59	3.6
% commuting in 1+ km	38.1	
% commuting in 3+ km	15.8	
% commuting out 1+ km	18.2	
% commuting out 3+ km	7.4	
Transport expenditure (pence per week)	19.51	22.6
% recording zero expenditure	28.1	
% with zero expenditure (including recodes)	36.5	

Source: New Survey of London Life and Labour.

Notes: Distance commuted in is defined as (distance between home and Charing Cross - distance between work and Charing Cross). Commuting out implies that this value is negative.
Observations with missing transport expenditure and commuting distance under 1 kilometre have been recoded to zero.

Table 3
Summary Statistics and Expected Signs of Regression Coefficients

	full sample	HH heads	non-heads	expected sign
Ln(weekly earnings)	10.72	11.14	10.35	
	(0.62)	(0.39)	(0.54)	
Age	31.67	42.41	22.12	+
	(14.09)	(11.25)	(8.30)	
Age^2	1201.40	1924.88	558.14	-
	(1048.11)	(1009.52)	(538.50)	
Sex (Male = 1)	0.68	0.93	0.47	+
	(0.47)	(0.26)	(0.50)	
Born and currently residing in the same borough (1 = yes)	0.320	0.328	0.312	?
	(0.47)	(0.47)	(0.46)	
Born in an adjacent borough to current residence (1 = yes)	0.095	0.137	0.058	?
	(0.29)	(0.34)	(0.23)	
Hours worked	46.26	46.84	45.75	+
	(7.92)	(8.43)	(7.41)	
Distance commuted (Km)	3.87	3.92	3.82	+
	(4.32)	(4.57)	(4.08)	
Distance to Charing Cross (Km)	5.11	5.24	5.00	+
	(2.44)	(2.57)	(2.30)	
Distance to Underground (Km)	1.07	1.11	1.04	-
	(1.22)	(1.34)	(1.10)	
$D_U * D_{CX}$	7.25	7.92	6.65	-
	(25.23)	(30.94)	(18.73)	
Transport cost	4.99	4.90	5.07	+
	(3.81)	(3.86)	(3.77)	
Sample Size	20,106	9,463	10,643	

Source: New Survey of London Life and Labour.
Note: Weekly earnings reported in hundredths of pence.

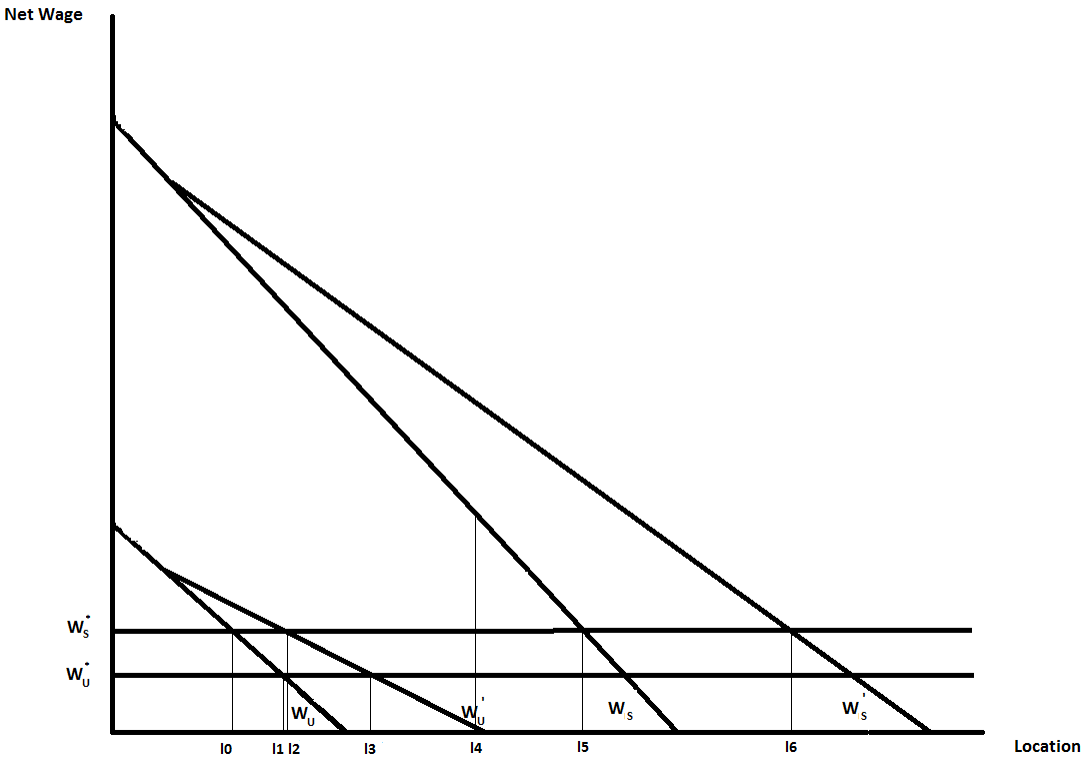
Table 4
Regressions on the determinants of earnings

	(1) Full Sample	(2) Household Heads	(3) Non-heads	(4) Highest earner	(5) Others	(6) Full Sample	(7) Full Sample	(8) Full Sample	(9) Full Sample
Age	0.111	0.028	0.149	0.076	0.119	0.108	0.105	0.111	0.109
	(95.14)**	(13.23)**	(43.31)**	(45.38)**	(37.40)**	(88.89)**	(90.17)**	(95.15)**	(85.24)**
Age^2	-0.001	-0.0003	-0.002	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(77.39)**	(13.37)**	(31.30)**	(41.72)**	(28.95)**	(72.87)**	(73.67)**	(77.40)**	(69.92)**
Sex	0.334	0.505	0.261	0.446	0.209	0.394	0.400	0.334	0.341
	(46.11)**	(27.32)**	(32.35)**	(38.94)**	(21.87)**	(58.03)**	(60.08)**	(46.12)**	(43.06)**
Same borough	0.071	-0.038	0.097	0.015	0.121	0.066	0.078	0.071	0.074
	(13.17)**	(5.98)**	(11.47)**	(2.46)*	(12.07)**	(11.88)**	(14.40)**	(13.16)**	(12.52)**
Adjacent borough	0.035	-0.041	0.030	0.002	0.045	0.023	0.043	0.035	0.039
	(4.07)**	(4.75)**	(1.79)	(0.21)	(2.09)*	(2.52)*	(4.94)**	(4.09)**	(4.26)**
D_W	0.006	0.004	0.009	0.005	0.009	0.009	0.007	0.007	0.005
	(8.67)**	(5.07)**	(7.74)**	(6.63)**	(6.36)**	(11.86)**	(8.59)**	(9.15)**	(5.90)**
D_{CX}	0.011	0.012	0.004	0.012	0.006	0.011	0.013	0.012	0.011
	(6.98)**	(7.06)**	(1.85)	(7.43)**	(2.02)*	(8.25)**	(8.04)**	(7.50)**	(6.29)**
D_U	-0.008	-0.011	0.003	-0.013	0.009	-0.007	-0.009	0.002	0.002
	(2.44)*	(3.25)**	(0.52)	(4.19)**	(1.30)	(2.52)*	(2.75)**	(0.55)	(0.53)
Hours worked	0.017	0.016	0.016	0.013	0.020	0.017	0.017	0.017	0.016
	(31.28)**	(20.70)**	(21.11)**	(19.65)**	(23.29)**	(30.13)**	(29.62)**	(31.31)**	(28.52)**
$D_{CX} * D_U$								-0.001	-0.0005
								(6.21)**	(3.87)**
Ln (Transport cost + 1)									0.008
									(9.64)**
Occupation dummies	YES	YES	YES	YES	YES	NO	NO	YES	YES
Skill dummies	NO	NO	NO	NO	NO	NO	YES	NO	NO
Workplace location dummies	YES	YES	YES	YES	YES	NO	YES	YES	YES
Constant	7.657	9.253	7.498	8.482	7.811	7.572	7.815	7.650	7.636
	(122.66)**	(107.90)**	(67.73)**	(128.64)**	(37.62)**	(237.46)**	(70.37)**	(122.52)**	(102.85)**
Observations	20,106	9,463	10,643	12,210	7,896	20,106	20,106	20,106	16,811
R-squared	0.70	0.53	0.58	0.60	0.51	0.66	0.69	0.70	0.69

Source: New Survey of London Life and Labour.

Notes: Robust t-statistics in parentheses. * = significant at 5%; ** = significant at 1%.

Figure 1
The Commuter's Problem



References

Abernethy, Simon T. (2013). “Deceptive data? The New Survey of London Life and Labour 1928 – 31”, Cambridge Working Paper in Economics and Social History, 16.

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. "Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration." *The American economic review* 102.5 (2012): 1832-1856.

Bailey, R. and Leith, A. (1998), Computerising and coding the New Survey of London Life and Labour. A Companion Paper for the Codebook (Release 2.0).

Bailey, R. and Leith, A. (1998), New Survey of London Life and Labour. Codebook for the Data Files (Release 2.0).

Bhaskar, V. and To, Ted (1999). "Minimum wages for Ronald McDonald monopsonies: A theory of monopsonistic competition." *Economic Journal*, 109(455), 190-203.

Bhaskar, V., Manning, Alan, and To, Ted (2002). "Oligopsony and monopsonistic competition in labor markets." *Journal of Economic Perspectives*, 16(2), 155-74.

Crampton, Graham R. (1999). "Urban labour markets." in Cheshire, Paul and Mills, Edwin S. eds. *Handbook of Regional and Urban Economics* 3, Amsterdam: Elsivier, 1499-1557.

Kain, John F (1962). "The journey-to-work as a determinant of residential location." *Papers in Regional Science*, 9(1), 137-60

Krugman, Paul R. (1991). *Geography and trade*. Cambridge: MIT Press.

Llewellyn Smith, Hubert (1930). “Introduction”. In Llewellyn Smith, Hubert (ed.) The New Survey of London Life and Labour (London: P. S. King), pp. 1-57.

Marshall, Alfred (1920). *Principles of economics*. London: Macmillan.

Mincer, Jacob (1958). "Investment in Human Capital and Personal Income Distribution". Journal of Political Economy 66 (4): 281–302.

Mincer, J. (1974). Schooling, Experience and Earnings. New York: National Bureau of Economic Research.

O'Regan Katherine M. and Quigley John M. (1993). "Family networks and youth access to jobs," *Journal of Urban Economics*, 34(2), 230-48.

Ponsoby, G. and Ruck, S. K. “Travel and Mobility”. In Llewellyn Smith, Hubert (ed.) The New Survey of London Life and Labour (London: P. S. King) pp. 171-99.

Romer, Paul M. (1986). "Increasing returns and long-run growth." *The Journal of Political Economy*, 94(5), 1002-37.

Rees, Albert and George Schultz (1970). *Workers in an urban labor market*. Chicago: University of Chicago Press.

Rotemberg, Julio J. and Saloner, Garth (2000). "Competition and human capital accumulation: a theory of interregional specialization and trade." *Regional Science and Urban Economics,* 30(4), 373-404.

Roy, Arthur (1951): “Some thoughts on the distribution of earnings.” *Oxford Economic Papers,* 3(2), 135-146.

Steinnes, Donald N. (1977) "Causality and intraurban location." *Journal of Urban Economics* 4(1), 69-79.

van Ommeren, Jos, Rietveld, Piet, and Nijkamp, Peter (1997). "Commuting: in search of jobs and residences," *Journal of Urban Economics*, 42(3), 402-21.

Wasmer, Etienne and Zenou, Yves (2006). "Equilibrium search unemployment with explicit spatial frictions," *Labour Economics*, 13(2), 143-65.