SERC Spatial Economics Research Centre

SERC DISCUSSION PAPER 120

Driving Up Wages: The Effects of Road Construction in Great Britain

Rosa Sanchis-Guarner (SERC, The Grantham Research Institute on Climate Change and the Environment, London School of Econmics and Political Science)

October 2012

This work is part of the research programme of the independent UK Spatial Economics Research Centre funded by a grant from the Economic and Social Research Council (ESRC), Department for Business, Innovation & Skills (BIS) and the Welsh Assembly Government. The support of the funders is acknowledged. The views expressed are those of the authors and do not represent the views of the funders.

Driving up Wages: The Effects of Road Construction in Great Britain

Rosa Sanchis-Guarner*

October 2012

* SERC, The Grantham Research Institute on Climate Change and the Environment, London School of Economics & Political Science

Acknowledgements

The work in this paper is based on data from BSD, ASHE and LFS, produced by the Office for National Statistics (ONS) and supplied by the Secure Data Service (SDS) at the UK Data Archive and by the Virtual Microdata Laboratory (VML). The data is Crown copyright and reproduced with the permission of the controller of HMSO and Queen's Printer for Scotland. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates. All the results have been granted final clearance by the staff of the VML and by the staff of the SDS. Additionally, any interpretations or opinions expressed in this presentation are those of the authors and do not necessarily reflect the views of the Department for Transport who also provided data to the study.

I thank Henry Overman, Steve Gibbons, Teemu Lyytik inen, Thierry Mayer, Guy Michaels, Gilles Duranton, Gabriel Ahlfeldt, Steve Ross, Ferdinand Rauch and Felix Weinhardt for comments and suggestions. I am also grateful to conference and seminar participants at the LSE Geography Economics Cluster seminar, the Urban Economics Association Meeting 2011, the LSE Labour Markets Workshop, the Finnish Government Institute for Economic Research (VATT), the University of Valencia, the Society of Labour Economics Meeting 2012, the 2nd IEB Workshop in Urban Economics 2012 and the 2nd European-UEA Meeting 2012 for comments. Giulia Faggio, Richard Welpton and the staff at the VML and the SDS kindly helped with the data. All errors and opinions are my own. This is an incomplete draft, please do not cite without permission. Comments and suggestions are welcomed. I acknowledge the Bank of Spain, the LSE and the ERSC grant ES/J007382/1 "Transport investments and spatial economic performance" for financial support.

Abstract

This paper estimates the effects of road construction on individual labour market outcomes using micro-data from Great Britain between 2002 and 2008. To capture these effects, I use a measure of accessibility to employment through the road network at a very detailed geographical level. I test the effect of accessibility changes on weekly wages and hours worked. In order to tackle potential sources of bias, I use an instrumental variable which exploits the variation in employment accessibility stemming only from changes in minimum travel times between locations. I argue that, conditional on controls, small scale spatial variation in the accessibility impact of road construction can be considered to be exogenous because road schemes are aimed to improve connectivity and reduce congestion for wider and more distant areas. I further use home and work location specific individual fixed-effects to control for endogenous sorting of workers and I also restrict the sample to workers who are located very close to the projects. I find a positive impact of accessibility from work location on weekly wages and total hours worked but no effect of accessibility from home neither on wages nor hours, conditional on commuting time. These effects are not due to selection into employment as a result of road construction. I also find evidence of accessibility from home reducing commuting travel time. Increased spatial competition or agglomeration externalities are potential explanations for the findings.

Keywords: Job accessibility, labour markets, roads, spatial sorting

JEL Classifications: J31, 018, R12

1. Introduction

Road network is a hugely important part of infrastructure in all countries, including Great Britain. Roads is the most important transportation mode for both goods and passengers. In particular, according to Transport Statistics Great Britain, between 1998 and 2008, over 90% of passenger transport was done by road. In 2001 over 80% of the commutes in the UK were done by motor vehicles (cars, vans, buses or motorcycles), and this percentage increased in city size [1]. Moreover, in these years, around 65% of goods transport was also done by roads. New transport projects require large amounts of public (and private) investment and they sometimes generate great controversy about their placement. Between 2001 and 2008, investment in road infrastructure increased in Great Britain around 55% in nominal terms and around 40% in per capita terms. Investment in roads accounted for over 40% of total infrastructure investment (which includes roads, trains, ports and airports). During the same period, the trunk and motorway network was extended by around 175 kilometres [2].

Given these facts, understanding the relationship between transport construction and economic outcomes is essential to the design of transport policy, and given the importance of road transportation for the movements of people and goods, the correct identification of the impact of road investments is also important for economic policy as a whole. Yet, the number of empirical studies that have tried to quantitatively assess these effects is still limited. The objective of this paper is to provide causal estimates of the effects of transport policy on labour markets outcomes, using micro data on firms and workers for Great Britain between 2002–2008.

Transport policy, in particular road construction, can affect labour markets through different channels. First of all, new transport links potentially reduce transport and commuting times and commuting costs. Therefore transport improvements bring employers and workers closer together and hence affect the effective size of the labour markets. For unemployed or inactive workers reduced commuting costs decrease search costs and reservation wages, and in this way, they can help mitigate frictional unemployment and increase the employment probabilities of those who are jobless. According to the so called "spatial mismatch hypothesis" (Inhanfeldt, 2006), physical proximity to jobs is an important determinant of labour market participation. Secondly, improvements in the transport network can increase the geographical scope of the agglomeration economies, as for a given physical distance employers and employees are nearer to each other. Better transport connections change the proximity between locations and, for given economic sizes, have a direct effect on effective density [3]. Additionally, in tighter labour markets increased competition for jobs and workers can foster productivity, as we could see more and better job matches taking place in these enlarged labour markets. Finally, transport investments can be also capitalised into non-labour prices and affect congestion, which will in turn affect location decisions of firms and workers. Given the unclear net theoretical predictions, the size and direction of the effects of transport policy on labour market outcomes remains mainly an empirical question (Gibbons & Machin, 2006).

In this paper I make use of rich datasets and a careful empirical strategy in order to gain insights on these relationships. Specifically, I estimate the impact of transport improvements, in particular road construction, on individual labour market outcomes. The effect of road construction is cap-

[1] UK Census 2001.

[2] Transport Statistics Great Britain.

[3] Rosenthal & Strange (2003) and Jofré-Monseny (2009) are some of the few studies which investigate the geographical scope of agglomeration economies.

tured using a measure of accessibility to employment (effective density). It captures the amount of employment which is reachable using the road network from a given location, inversely weighed by the travel time to reach other locations. New constructed links in the road network may affect travel times between locations, and thus directly affect effective density. I use microdata from several British datasets for the period 2002–2008 and test the effect of changes in accessibility on several outcomes, namely nominal wages, hourly earnings and hours worked. I further investigate the effect of accessibility on the probability of being employed and on commuting time. I follow an empirical strategy which is based on the use of individual micro-datasets and small geography. To be able to infer a causal interpretation the estimates I use a instrumental variables strategy that exploits small scale spatial variation in the way road construction changes accessibility to employment. To address endogeneity issues, I use the variation in accessibility stemming from road construction and compare individuals which are located in wards very close to the projects and which vary in the timing and the intensity of the transport treatment [4].

To construct the accessibility measures I use a novel data set of geo-located road projects carried out in Great Britain at different points in time during the period 2002–2008. I combine this data with road network data and use Geographical Information System (GIS) network analysis tools to calculate optimal minimum travel times by road between locations for every year in the sample. I use this information, jointly with information on the exact location of plants and employment, to calculate accessibility to employment in different years.

In contrast to many studies which use density of roads or connectivity, I measure the effect of transport policy using an index of accessibility. One advantage of using this measure is that it is not constrained to artificial geographical boundaries like measures based on the density of roads. More importantly, this measure has the advantage of capturing small improvements changes in the network, caused by the construction of new links. It weights the importance of the different locations based on their economic size and on their position in the road network relative to the rest of the country. Even if the location was previously connected to the road network, any changes in the network can affect the relative "connectivity" of the location, when the improvements reduced the travel time necessary to reach other points of the network. This measure is more appropriate in the UK context than other measures used in previous empirical studies.

In the empirical exercise, I focus on changes in accessibility due to changes in travel times (proximity). In particular, optimal travel times between locations change over time due to road construction. Over all possible improvements through which transport policy can affect road networks I focus on construction for two reasons. First of all, among all the different type of road interventions (e.g. dualling, resurfacing, improving lighting and signalling, etc), construction of new links is a clear source of variation which is easily identifiable and measurable. Construction of new road links affects optimal routes along the road networks and it constitutes the source which is likely to have the larger impact on the variation in travel times between locations. It also is a well-defined tool which can be used by policy to influence proximity between locations. The second reason is data availability. The collection of reliable data on transport improvements and projects over a length of time is demanding and their measurement can be problematic. The methodology I use to generate variation in optimal travel times needs a clear definition of changes in the road network and addition

[4] The methodology and the data of this paper are largely based on Gibbons et al. (2012), who estimate the effects of road construction on firm and ward outcomes using British data for the period 1998-2008.

new links to the network provides a reasonably source of variation (see Sections A.1.2., A.1.3. and A.1.4. in the Appendix for more information). Nevertheless, one must bear in mind that other type transport improvements might also affect optimal travel times between location and the analysis is limited from that perspective.

To carry out the empirical analysis I use three large workers and firms microdata datasets, which were provided by the UK Office for National Statistics (ONS). I calculate employment at the ward level using data from the census of establishments (Business Structure Database - BSD), a rich dataset which provides information of the universe of British plants by reporting their location (postcode), their sector of activity and on their number of employees. For the individual labour market outcomes I use two additional datasets: the Annual Survey of Hours and Earnings (ASHE) and the Labour Force Survey (LFS). The first survey is a panel of employees, and it provides information on their earnings, hours worked, job characteristics, and the postcodes of their workplace and their residential locations. The second dataset contains quarterly information on a sample of households and includes information on a wide range of labour market outcomes, as well as a large list of personal and household characteristics.

The geographical unit used as the basis for the calculation of the accessibility index is the electoral ward. Wards (as defined in 1998) are quite small areas and there are over 10,500 in Great Britain [5]. Having such a large number of units implies that there is a fair amount of variation in the values of accessibility changes due to road construction even when focusing on areas close to where the road construction took place. Therefore, the fine geographical detail of the data allows us to implement a careful identification strategy.

In order to infer causality from the estimates, I exploit micro-data and small spatial scale in order to address these potential sources of bias. When estimating the impact of accessibility to employment on individual labour market outcomes, for example weekly wages, there are three endogeneity issues that need to be tackled. Firstly, cross-sectional estimates of the effect of accessibility on labour market outcomes could be biased if the model does not capture underlying time-invariant factors (such as place specific productive advantages) that affect both effective density and economic outcomes. I use a fixed-effects estimation method to address this problem. When I use individual fixed-effects I exploit the within individual variation over time. However, workers might be sorting spatially (changing home or workplace) in order to take advantage of the accessability changes or due to other unobservable reasons. To address this issue I further use individual-home-job fixed-effects which exploit the variation in accessibility for an individual while he is keeping his home-work location pair fixed.

When using these individual-home-job fixed-effects, changes in accessibility can arise because of road construction which affects the travel times among locations and because employment relocates. Therefore, the second endogeneity issue is that accessibility changes due to relocation of employment may be partly driven by the outcome variable studied or be correlated with the same unobserved shocks. To address this source of bias, I construct an instrument which uses accessibility changes stemming only from the road construction. In practice, I use an alternative accessibility index for which employment is fixed at the beginning of the period and only the travel times between wards, which vary annually due to road construction, change.

[5] The average area of British wards is 21 square kilometres (they are smaller in England, 15 km^2, and much bigger in Scotland, 62 km^2) and the average population estimate in 2001 (for England and Wales) is around 6,000 people.

Finally, when deciding where to make transport investments, the government could be targeting specific areas because these have specific productivity or wage trends. In order to reduce the possible bias caused by the endogeneity of the placement of the transport investments, I focus on individuals which are located in wards within 10, 20 or 30 kilometres of road schemes. This way, I compare firms and locations which are close to the projects and exploit the fact that the impact of the road construction varies considerably even within the distance band. It is quite unlikely that the new links are aimed at specific individuals in wards within those narrowly defined distance bands, specially after controlling for different growth trends around the schemes.

In the preferred results, I find positive effects of accessibility from work on earnings (weekly basic and total wages) and total hours worked. I find no evidence that accessibility from home affects individual wages or hours worked. Increased spatial competition for jobs or agglomeration externalities are suitable potential explanations of these findings. In additional results, I also find some evidence that increasing accessibility from home reduces time traveled to workplace, but changes in accessibility do not have any effect on the probability of being employed. This suggests that even if road construction affects labour market outcomes by decreasing commuting costs, it does not affect selection into employment.

This research contribute to the existing evidence in three ways. First, I use very rich worker microdata and a novel dataset on road projects, which allows to study the relationship between transport policy and individual labour market outcomes at a very detailed geographical scale and to study several labour market outcomes. To my knowledge, there exists almost no previous evidence on the effect of road construction on individual labour market outcomes [6]. Secondly, given the quality of the data used, I can adopt a careful empirical strategy which allows to tackle several identification issues which could undermine the causal interpretation of the results. This improves the validity of the results in order to extract policy recommendations. Finally, I provide robust empirical evidence on some strands of the existing theoretical urban labour economics literature.

My results are able to shed some light on the channels through which transport policy affects individual economic outcomes. The findings add to existing evidence of the effects of agglomeration and market access on nominal wages (for example Mion & Naticchioni, 2005; Combes et al., 2008), although the focus is on changes in one specific channel which affects the level of agglomeration, e.g. road construction. Most of the evidence on the effect of market access on workers wages, even if causal, cannot separate the impacts attributable to changes in the economic size of locations from impacts attributable to changes in "proximity" between these locations. As explained in detail below, I exploit on the variation of accessibility (effective density) which stems from road construction. If there are any potential causal benefits to workers arising from changes in accessibility, these could be induced through transport policy. Therefore, the estimates inform not only on the overall effect of accessibility on worker economic outcomes but also on the channel through which policy can potentially impact these outcomes. Finally, my findings also inform about some of the predictions of the spatial mismatch literature (see Inhanfeldt, 2006, for a review) and of the thin labour markets theory (Manning, 2003).

The remainder of the paper is structured as follows. Section 2. sets out the channels through which transport policy can affect labour market outcomes. Section 3. reviews the related theoretical and empirical literature. The empirical strategy is explained in detail in Section 4.. The data sources

[6] The only exception is Aslund et al. (2010).

and the construction of the variables used in the empirical analysis are also described in this section. The results are fully explained in sections 5. and 6., and robustness checks and an interpretation of the coefficients is also provided. Finally, section 7. concludes and explains the future steps of this research.

2. Theoretical framework

The role of transportation in the spatial distribution of economic activity and economic performance has become of increased interest to researchers in the last years. Decreasing transport costs are considered to be a central driver of economic integration and of the rise of agglomeration externalities, but solid empirical evidence on the channels through which these effects operate is still needed.

Transport policy affects labour markets through multiple channels (for a review see Gibbons & Machin, 2006). While most of trade and location theory regards transport investments as having an effect on (goods) transport costs (Michaels, 2008, for example), urban labour theories (Zenou, 2009) consider transport policy as having an effect on commuting costs and therefore mostly operating through the labour market. Glaeser & Kohlhase (2003) document the declining role of goods-transportation costs in developed countries and highlight the increasingly important role of the mobility of workers in modern economies.

If road construction reduces commuting costs [7], employers and employees come closer together (spatial competition) and the effective size of the labour markets (scope) can be affected. As a consequence, reduced commuting cost can affect search intensity and reservation wages of workers, and therefore the number of job matches and observed wages.

The tightness of the labour market, i.e. the ratio of unemployed relative to the number of job vacancies, is affected by the accessibility and the size of the labour market. If labour markets are more accessible to unemployed workers residing outside the labour market area, or the labour market becomes effectively bigger because it is better connected, for a given number of vacancies the number of potential candidates would increase. This increases competition for jobs and might have two effects. Firstly, unemployed workers living further away from job might become employed due to the increase in accessibility. Secondly, due to increased competition, we could see an increase in the quality of job matches which could be translated into higher productivity and wages.

If employers require workers with specific characteristics or skills, employees can bargain to improve their labour market conditions (wages, hours worked, occupation) to compensate for longer commutes. However, if labour markets are thin (Manning, 2006), i.e. workers have access to a limited number of potential employers, longer commutes could not be fully capitalised into nominal wages. Longer commutes can also have an effect on wages through their effect on productivity, if shorter commutes are related to healthier or more motivated workers (van Ommeren & Gutiérrez-i-Puigarnau, 2011).

Related to the these points, the spatial mismatch hypothesis predicts that longer "physical" distance [8] between residential location and job location can have negative effects on the labour market outcomes of those living further away from employment centres. A reduction in commuting travel

[7] Commuting costs involve both time and monetary costs. As the transport improvements studied are driven by road construction, I focus on the effects of reduced commuting travel times as a proxy for commuting costs.

[8] Distance includes geographical distance but could also include bad public transport provision or poor access to private transport (car).

times induced by transport investments would thus affect search costs and reservation wages, therefore impacting observed wages and unemployment rates (see Phillips, 2012, for some evidence on these channels). All these effects might impact labour market outcomes like employment status (employed versus unemployed), unemployment spells, unemployment duration and nominal wages.

Transport policy can also affect labour supply, e.g participation and hours worked. It can weaken barriers to participation in the labour market, encouraging the entry of disadvantages groups (female or low-skilled) into job search. It can affect hours worked if wages increase due to transport improvements, making work more appealing; or even if increased competition in denser areas induce young professionals to behave in a more rivalrous manner (the "rat race" argument in Rosenthal & Strange, 2008). Reduced commuting costs could also affect hours worked if the changes in travel times and wages affect the optimal labour-leisure choice.

The emergence and scope of agglomeration externalities might also have positive effects on workers earnings if urbanisation economies [9] and proximity to markets (Krugman, 1991; Krugman & Venables, 1995) give raise to productivity gains which are capitalised into nominal wages. If effective density increases when locations become closer due to road construction, the scope of the agglomeration externalities could be affected. As explained in Rosenthal & Strange (2004), the extent of the externality is its scope. In particular, we expect that reduced commuting cost would affect the geographic scope of agglomeration externalities, reducing their attenuation, as for a given physical distance the "effective distance" (measured in travel time) decreases due to road construction [10]. If agents are closer, there is more potential for interaction. Ahlfeldt & Wendland (2011), using the construction of the railway in Berlin at the beginning of last century, show how distance becomes less important and the scope of spatial interaction increases as transport infrastructure improves.

Additionally, if better connectivity is regarded as an amenity of locations, improved infrastructure could be capitalised into house prices (some evidence of this is provided in Gibbons & Machin, 2005; Ahlfeldt, 2011b). In this sense transport policy would affect real wages, which would change because of changes in nominal wages and in local prices. Net changes in real wages would affect individuals real labour earnings, but if there is capitalisation of accessibility into house values, for home owners there could be a positive effect on their total wealth. Real earning changes could in turn affect residential mobility and the commuting versus migrating choice. Higher earners would find it profitable to commute longer if they can afford better quality housing further away from their jobs. In addition, transport investment could affect residential sorting if capitalisation into house rents force low income workers (renters) to move away from areas where transport improvements have taken place.

3. Empirical evidence

Even if some authors have explicitly included the role of transportation into spatial economic analysis (Combes & Lafourcade, 2001; Puga, 2002; Behrens et al., 2004; Venables, 2007), there is still need to empirically establish the causal link from transportation infrastructure to spatial economic performance, especially with regard to the effects on individual outcomes.

[9] Sharing, matching and learning in Duranton & Puga (2004) terminology; input-output linkages, labour market pooling and knowledge spillovers in the classical Marshall (1890) terminology.

[10] Rosenthal & Strange (2003) and Rosenthal & Strange (2008) document the attenuation of agglomeration externalities with distance.

There exists a substantial amount of empirical evidence on the effects of infrastructure (including transport) on economic outcomes at the macro-level (for a review see Gramlich, 1994; Straub, 2011). This literature has focused on the impacts of investment in roads and public infrastructure on several outcomes, and the authors generally estimate an aggregated Cobb-Douglas production function where infrastructure or roads are treated as a factor of production (García-Milá et al., 1996).

A number of recent papers have estimated, using careful identification strategies, the effect of roads on other economic outcomes using data for the USA. Using the 1947 planned Interstate Highway System as an exogenous source of variation, Baum-Snow (2007b) studies the effect of highways on the process of suburbanisation of American cities since the 50s, and on the changes in commuting patterns since the 60s (Baum-Snow, 2010). Michaels (2008) uses a similar source of exogenous variation to estimate the effect of reduced trade barriers on the demand for skills. Hymel (2009) uses a cross section of US metropolitan areas to assess the impact of traffic congestion on aggregate employment growth. Duranton & Turner (2011a) examine the effect of road construction on vehicle-kilometres traveled (VKT) in US cities. These same authors, using a similar identification strategy, also estimate the effects of highway construction on urban growth (Duranton & Turner, 2011b) and on trade (Duranton et al., 2011). Finally, other papers (Donaldson, 2010; Faber, 2012) have focused on developing countries (highways in China and railroads in colonial India) to study the effect of the reduction of transport costs, due to transport network development, on trade integration and consequent economic development.

A large fraction of the evidence of the effect of transportation on labour market outcomes has been provided by the empirical testing of the spatial mismatch hypothesis. The related literature has focused on explaining, both from an empirical (Rogers, 1997) and a theoretical perspective (Gobillon et al., 2007; Gautier & Zenou, 2010), the relationship between the poor access to jobs of disadvantaged groups (like low-skilled, ethnic minorities or females) and their poor labour market outcomes (mainly unemployment or inactivity).

Most of theoretical and the empirical evidence on the spatial mismatch hypothesis has been developed in the US context. In the last decades many low-skill jobs have been relocated to the city suburbs due to suburbanisation. Poorer individuals and ethnic minorities remained residentially segregated in city centres, due to mobility constraints and cheaper house prices. If these city centres are poorly connected to the suburbs (due to the lack of good public transport links) or if these groups have worse access to private transportation (like cars), then their performance in the labour market is endangered due to poor access to potential vacancies. There exists abundant empirical evidence on the spatial mismatch hypothesis but mainly for the US (for a review see Inhanfeldt, 2006). Recent evidence using experimental data is provided by Phillips (2012). Some recent papers have studied some aspects of the spatial mismatch hypothesis using data for France (Détang-Dessendre & Gaigné, 2009; Gobillon et al., 2011), for the UK (Patacchini & Zenou, 2005) or for Spain (Matas et al., 2010). Aslund et al. (2010) study the impact of job proximity on individual employment and earnings using Swedish data and a refugee dispersal policy to obtain exogenous variation in individual locations.

Another channel through which transport policy affects labour market outcomes is by changing commuting costs (which can be proxied by commuting time or commuting distance). A reduction in commuting time and costs associated with transport improvements enables people to increase the scale of their job search and could also encourage potential workers to participate in the labour market (Vickerman, 2002; Jiwattanakulpaisarn et al., 2009).

Manning & Petrongolo (2011) use British data to study how "local" labour markets are. They find that workers tend to look for and to accept jobs which are relatively close to them, so labour markets are very local. They also find that workers are discouraged from applying to jobs in areas where there is a lot of competition for jobs (tight labour markets). Their research is related to the present paper as I indirectly investigate the effect that changes in commuting distance (which diminish effective distance to jobs) has on individual wages and hours. Reduced commuting time due to road construction could have an effect on the actual size of the local labour markets and could also affect the number of competitors for a give search area. Similarly to the authors, I also use small British geography (wards) to approximate "continuous" space.

There are a few papers that have tested the impact of commuting distance on individual labour market outcomes. van Ommeren & Gutiérrez-i-Puigarnau (2011) and Gutiérrez-i-Puigarnau & van Ommeren (2010) look at the effect of a reduction in commuting distance induced by a establishment-relocation (within the same firm) on hours worked and absenteeism (productivity). Mulalic et al. (2010) use changes in distance commuted caused by relocation of firms to study the effects of commuting on wages. Search costs can also lead to thin local labour markets which may explain the positive gradients between wages and commuting times (Manning, 2006).

Finally my research is also related to the effect of agglomeration economies and proximity to markets on nominal wages (Combes et al., 2008; Mion & Naticchioni, 2009; Hering & Poncet, 2010). Workers which are located in denser areas are able to capitalise positive agglomeration externalities on wages. The identification of the effects of agglomeration economies is not straightforward (Combes et al., 2011). Combes et al. (2010) point out two major problems with the identification of the effects of agglomeration on wages. Denser or better connected places might attract more workers and this in turn increases agglomeration (which they refer to as the "endogenous quantity problem"). They suggest the use of instrumental variables to deal with this problem. Also, more able workers might sort into dense areas in order to take advantage of the agglomeration externalities (which they refer to as the "endogenous quality problem"). To deal with this second problem, the authors suggest using individual panel data to control for unobservable individual and locational characteristics. In this paper I use a similar strategy, but I focus on the effect on agglomeration on labour market outcomes through changes on road accessibility.

Due to the multiplicity of channels through which transport policy and decreased commuting times can affect labour market outcomes and the non-conclusive empirical evidence, the estimation of robust causal effects is of considerable interest. The aim of the present paper is to provide such evidence.

4. Measuring road construction

4.1. Accessibility measures

The aim of this paper is to estimate the causal effect of road construction on economic outcomes. Therefore, the first challenge is to find a measure that captures changes in the road network. A variety of measures have been used in the literature [11]. As the current road network in Great Britain

[11] Some measures are connectivity to the network (Faber, 2012), kilometres of roads within a given area (Melo et al., 2010; Duranton & Turner, 2011b), distance to closest highway (Baum-Snow, 2007a), number of rays crossing a given area (Baum-Snow, 2007a, 2010), presence of highways in a given location in a particular year (Chandra & Thompson, 2000; Michaels,

is very dense and the length of the new links constructed is relatively short, to use measures like changes in connectivity or density of roads to capture the effects of road construction would not be appropriate for the UK context. Instead, I use a measure of accessibility to employment "through the road network" (or effective-density) from each location at every point in time.

Accessibility to employment measures the amount of employment which is reachable using the road network from a given location, inversely weighed by the travel time to reach these other locations. One advantage of using this measure is that it is not constrained to artificial geographical boundaries like some of the alternative measures. Moreover, it allows us to use variation due to road construction which affects optimal travel times, even if the location was previously connected to the network. Additionally, it captures the effects of transport improvements over the whole geography. This measure is, thus, appropriate in a setting where road density was already high at the beginning of the period of study.

Formally, accessibility to employment A_{rt} from a given location r at time t is defined as:

$$A_{rt} = \sum_{j \neq r}^{R} \left[a\left(c_{rjt}\right) * econ\_size_{jt} \right] \tag{1}$$

where $a(.)$ is the transport cost function, c_{rjt} are the transport costs between locations r and j at time t and $econ\_size_{jt}$ measures the economic size of the location at time t. Finally, t corresponds to years 2002 to 2008 [12].

This index is a measure of the economic mass accessible to a firm or a worker in a particular location, given the local transport network. At a given origin location r at time t, accessibility A_{rt} is a weighted sum of economic mass in all destinations j that can be reached from origin r by incurring a transport cost c_{rjt} along some specified route between r and j (for example straight line distance or minimum cost route along a transport network – measured in travel time or in distance). The function $a(.)$ determines how the weights enter in the calculation of A_{rt}. I use electoral wards (as defined in 1998) as the geographical unit. The economic size of locations is measured using total ward employment and travel time between wards is used as measure of transport costs (further details on the construction of ward-to-ward travel times and ward employment is provided in Section A.1.). Assuming inverse cost weights [13] and a cost decay equal to one [14], the final expression of accessibility to employment becomes:

$$A_{rt} = \sum_{j \neq r}^{R} \left[\left(1/travel\_time_{rjt}\right) * employment_{jt} \right] \tag{2}$$

2008), "lowest-cost route effective distance" (Donaldson, 2010) or amount of public expenditure on road infrastructure (Fernald, 1999).

[12] As explained in Section A.1.2., road networks and travel times are dated at the beginning of the year. As the ward employment data is dated at the end of the year, in practice I combine travel time data at the beginning of year t (for example 2002) with employment data at the end of year $t-1$ (end of 2001). This implies that in the empirical specifications of sections 5.1. and 6.1., accessibility measures are in fact lagged one year with respect to the outcome variables.

[13] In the definition of A_{rt} the value of the weight $a(.)$ attached to any destination r is a decreasing function of the cost of reaching destination j from origin r. α is the cost decay. Potential weighting schemes include: "cumulative opportunities" weights $a\left(c_{rjt}\right) = 1$ if j is within a specified distance of r, zero otherwise; "exponential weights" $a\left(c_{rjt}\right) = \exp\left(\alpha c_{rjt}\right)$; "logistic weights" $a\left(c_{rjt}\right) = \left[1 + \exp\left(-\alpha c_{rjt}\right)\right]^{-1}$ or "inverse cost weights" $a\left(c_{rjt}\right) = c_{rjt}^{-\alpha}$. See Graham et al. (2009) for further discussion of these indices.

[14] Graham et al. (2009), using the inverse cost weighting scheme, estimate the parameter of distance decay functions for several sectors using similar British data to ours and find values between 1.8 and 1, depending on the sector. The estimated alpha for the whole economy is around 1.6. In practice, I check the robustness of the results to different distance decays.

I exclude location r from the calculation of the accessibility measures to lessen the potential reverse causality problem. This problem arises because the dependent variables include labour market outcomes the individuals located in the ward and it is likely that these outcomes and the economic size of a given location are be jointly determined. As explained below, I use an instrument to address this issue, so the inclusion of the own economic-size in the calculation of A_{rt} would not be an issue.

The calculation of equation (2) requires the construction of an origin–destination (O–D) matrix whose components are travel times between the locations. When computing the O–D matrix I apply a limit of 75 minutes drive time (1.25 hours). This limit facilitates O–D matrix computation but hardly affects the value of the accessibility index because wards beyond 75 minutes have negligible weights in the calculation of A_{rt}. Moreover, as shown in Table A.8, more than 99% of commutes in the Great Britain are below 90 minutes.

The accessibility index as defined in (2) is similar in structure to market potential measures used in economic geography (e.g. Harris, 1954; Krugman, 1991), and to the accessibility indices used more generally in the transport literature (e.g. Vickerman et al., 1999; El-Geneidy & Levinson, 2006). This measure is similar to those usually used in empirical tests of the spatial mismatch hypothesis (for example Rogers, 1997) and closely related to the market-access measure used in tests of the effect of agglomeration economies on wages (Mion & Naticchioni, 2009; Hering & Poncet, 2010) or on productivity (Ciccone & Hall, 1996; Martin et al., 2011) [15].

Accessibility A_{rt} changes in a given origin r are driven both by changes in travel times between wards (stemming from road construction) and by changes in the employment of origin ward r and in the different wards j around r. This may lead to endogeneity problems in the estimation of the effect of accessibility, if the employment changes near the origin are causally linked with changes in the economic outcomes in the origin or driven by the same unobserved factors. Moreover, as I focus on the effect of road construction, the examination of A_{rt} would be limited as it would be impossible to differentiate between the changes driven by employment relocation and the changes driven by changes in travel time between locations. It is then useful to construct an alternative accessibility measure $\hat{A}_{rt}$ that focuses on the changes in accessibility that stem only from road construction:

$$\hat{A}_{rt} = \sum_{j \neq r}^{R} \left[(1/travel\_time_{rjt}) * employment_{jt_0} \right] \tag{3}$$

where t_0 is some fixed period of time at the beginning of the period of analysis. In the empirical analysis below t_0 corresponds to (end of) 2001. Fixing employment to its t_0 level ensures that changes in the accessibility index (3) over time occur only as a result of changes in the costs c_{rjt} (e.g. travel time) and not as a results of changes in wards employment. In the empirical work in the next sections, I instrument A_{rt} with $\hat{A}_{rt}$ in order to only use the variation in accessibility which stems from the road construction.

The empirical methodology uses the changes in the accessibility index at each location r to estimate the extent to which individuals in location r are "potentially" affected. In this sense, the measure

[15] The definition of accessibility used is different from other measures of accessibility used in the literature. Other authors, for example Ihlanfeldt & Sjoquist (1990); Raphael (1998); Ihlanfeldt (2002); Osland & Thorsen (2008); Détang-Dessendre & Gaigné (2009); Ahlfeldt (2011b) or Gobillon et al. (2011) have used indices of accessibility based on mean commuting time of employed workers within some given area, number of available jobs within a given distance band or similar gravity-weighted measures to A_{rt} but using different costs functions or distance decays, calculating skill-specific indices, or using employment growth as the measure of economic size of locations. Examples of authors who have use the same definition of accessibility or "market-access" measures are Graham (2007); Ahlfeldt (2011a); Holl (2012) and Donaldson & Hornbeck (2012).

(changes in accessibility) captures the "intention-to-treat". In fact, it is mainly through road construction and other road improvements that transport policy can affect travel times and alleviate congestion. This strengthens the validity of the approach.

Section A.1. in the appendix contains full details on the construction of the ward employment and the travel time between locations used in the calculation of the accessibility measures.

4.2. Descriptive statistics

Table 1 summarises the changes in the log of accessibility between 2002 and 2008, i.e. the growth in accessibility between the first and last year of the sample. I calculated accessibility in three different ways. The table shows the growth for these three accessibility indices for all the wards (10,540) and for wards which are situated within 10, 20 and 30 kilometres of the road schemes carried out during the period of analysis. In Figure A.1

In the upper panel (PANEL A) the accessibility measure is calculated fixing employment at 2001 levels ($\hat{A}_{rt}$), so that all the variation in accessibility comes from changes in the road network (due to road construction). Instead, the lower panel (PANEL C) shows the summary statistics when I calculate the accessibility allowing both employment and road infrastructure to vary (A_{rt}) in every year. The middle panel (PANEL B) shows the summary statistics when I calculate the accessibility fixing travel times at the beginning of the period (2002) and combining these weights with annual ward employment. In this case the variation in accessibility comes from relocation of employment across space. This latter measure is similar to the standard "market-access" measures calculated using geodesic distance.

	Wards	**Mean**	**Standard deviation**	**90^{th} percentile**	**Maximum value**	**Proportion of zeroes**
PANEL A: *2001 employment and time-varying travel times*						
All	10,540	0.30%	1.63%	0.57%	52.37%	42.88%
10 kms	862	2.44%	5.14%	5.93%	52.37%	9.86%
20 kms	2,453	1.12%	3.23%	2.40%	52.37%	11.29%
30 kms	3,888	0.79%	2.62%	1.65%	52.37%	10.65%
PANEL B: *Time-varying employment and 2002 travel times*						
All	10,540	13.43%	5.27%	18.33%	80.23%	0.00%
10 kms	862	12.37%	3.80%	15.61%	32.18%	0.00%
20 kms	2,453	12.57%	4.09%	15.94%	56.83%	0.00%
30 kms	3,888	12.57%	3.81%	16.00%	56.83%	0.00%
PANEL C: *Time-varying employment and time-varying travel times*						
All	10,540	13.74%	5.47%	18.70%	80.23%	0.00%
10 kms	862	14.84%	6.45%	19.85%	72.91%	0.00%
20 kms	2,453	13.71%	5.22%	17.59%	72.91%	0.00%
30 kms	3,888	13.37%	4.63%	17.12%	72.91%	0.00%

Source: Department of Transport, BSD and own authors' calculations.

Table 1: Summary statistics change log of accessibility between 2002 and 2008

The upper panel of Table 1 shows that employment accessibility induced by road construction was on average small, only 0.3% (which is very similar to the decrease in travel times observed in Table A.1 in the appendix). However, average accessibility change increases significantly when we focus on wards closer to projects. For example, within the 10-kilometre distance band the mean change is 2.44% and the 90^{th} percentile is 5.93%. As we expand the sample away from the schemes changes in accessibility tend to fall. Within 20 kilometres, mean accessibility change is 1.12% and the 90^{th} percentile is 2.4%. Within 30 kilometres of the schemes mean accessibility change is 0.79% and the 90^{th} percentile is 1.65%. It is worth noting that the standard variation of the changes is relatively large, suggesting a substantial amount of spatial variation of changes in accessibility. The lower panel of Table 1 shows that accessibility employment dynamics are a more important driver of variation in effective density than road construction. This is confirmed if we observe the changes in the middle panel, which focus on the variation on accessibility driven by employment relocation. Nevertheless, variation due to road construction is non-negligible relative to overall changes.

Table 1 illustrates the magnitude of total changes in accessibility between the first and the last years of the period of analysis. In fact, I use annual variation in accessibility in the estimation of the empirical results. For this reason, Table A.9 in the appendix provides summary statistics on the level of log accessibility in every year. The mean and standard deviation are very similar across the three panels.

Figures 1(a) and 1(b) illustrate the spatial relationship between the location of road schemes and resulting accessibility increases. Figure 1(b) shows the changes in log accessibility between 2002 and 2008 which stem only from road construction ($\hat{A}_{rt}$). The biggest changes in accessibility are around the schemes plotted in Figure 1(a), but there is substantial spatial variation across the country.

The amount of spatial variation is more evident in Figure 1(c). It represents the same values as Figure 1(b) but focuses on the Manchester-Bradford-Leeds area. This figure shows the small scale of the spatial data and also the substantial amount of variation in accessibility values close to the new links. It also allows us to illustrate the identification strategy (which is explained in Section 5.1.). The location of the new links constructed between 2002 and 2007 are indicated by bold white lines, with the name of the road labeled. The dark grey lines are ward boundaries. The thick black lines delimits the wards which are included within 30 kilometres of the road projects within the area. The map illustrates that the effects of road construction on accessibility vary considerably across wards in the vicinity of the same scheme. In the identification strategy I argue that these differences in accessibility changes across wards are coincidental and can be treated as exogenous, especially when controlling for differential time trends near different schemes.

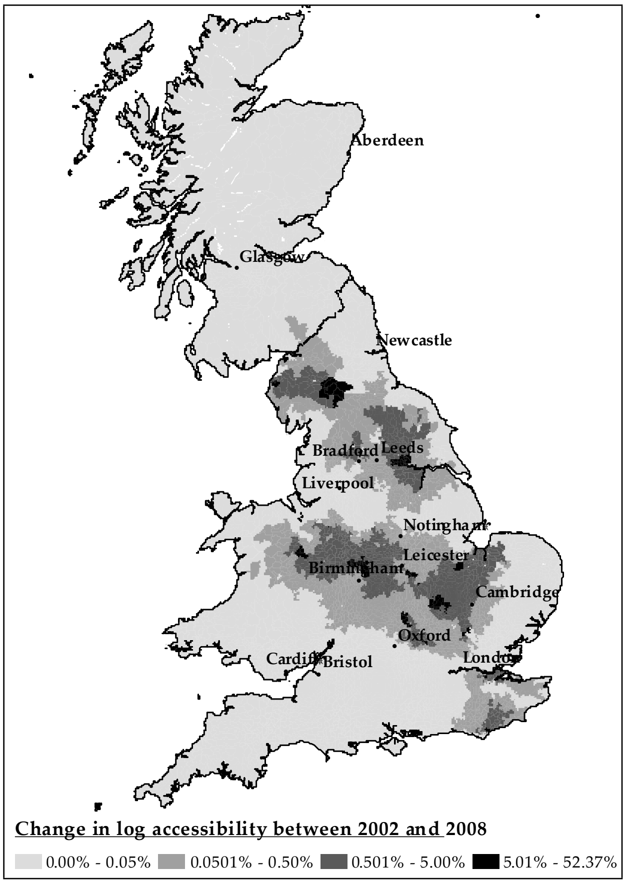
5. Empirical results: individual wages and hours worked

5.1. Empirical specification and identification strategy

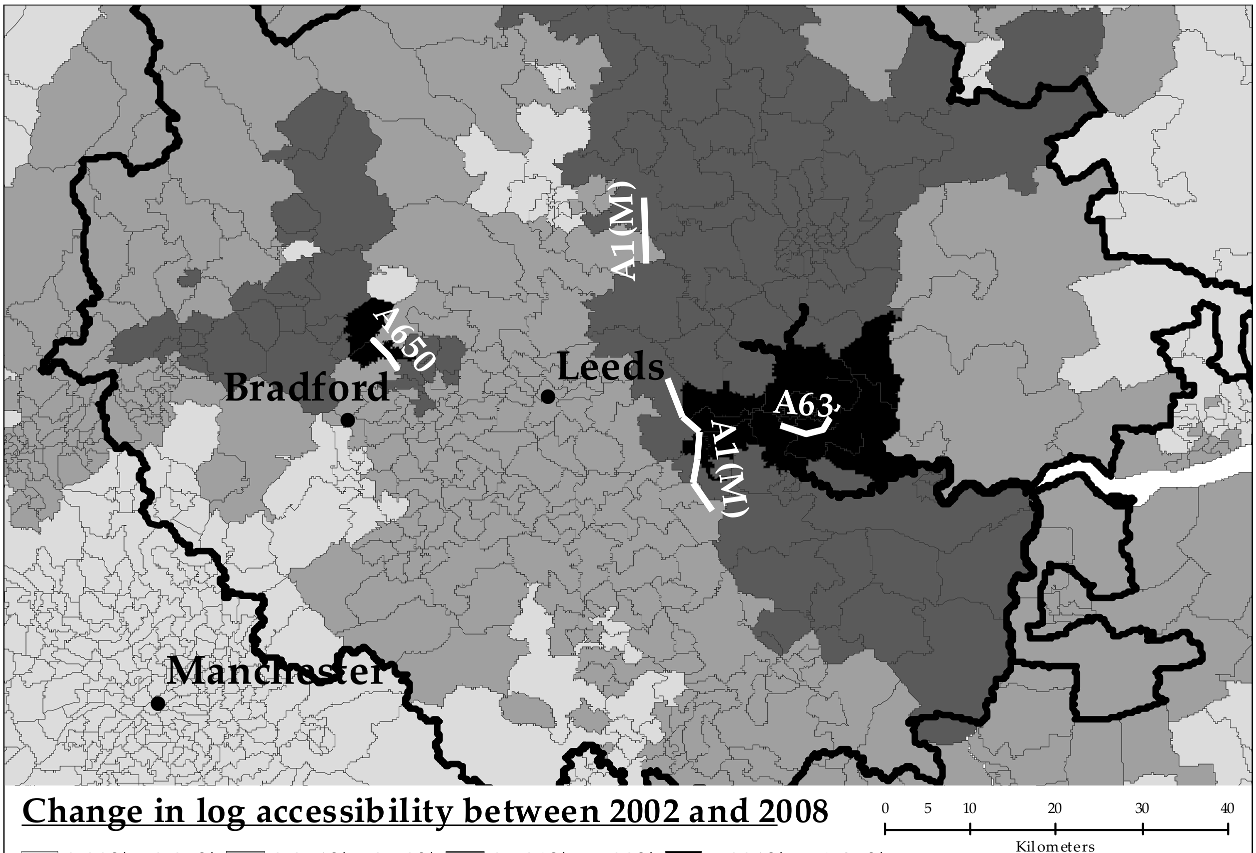
The Annual Survey of Hours and Earnings (ASHE) is used to obtain results on individual earnings and hours worked (more details in Section 5.2.). I use a panel of workers to estimate the effect of accessibility to employment, both from work and from home, on individual labour market outcomes. The main regressions use data for years 2002 to 2008. Three outcomes are investigated: weekly wages, weekly hours worked and hourly earnings. Both at basic and at total (which includes overtime) outcomes are analysed.

A worker i living in ward h and working in ward w at time t has labour market outcome y_{ihwt}. Given that accessibility is measured at the ward level, I ignore changes of home or job within the

(a) Major road network 2008 and new roads projects



(b) Change in log accessibility (2001 employment)



(c) Change in accessibility in Manchester-Leeds area (2001 employment)

Figure 1: Location of projects and induced changes in log accessibility 2002–2008

wards [16]. At each point in time, workers live and work in specific wards. Over time, workers location can change, if they change jobs, change home or change both. Initially, I use individual variation in accessibility and labour market outcomes which could also driven by spatial relocations (more on this below). The relationship estimated is:

$$y_{ihwt} = \beta_1 A_{iht} + \beta_2 A_{iwt} + \beta_3 c_{ihwt} + \theta X_{it} + \delta Z_{ht} + \lambda W_{wt} + \mu_i + \xi_t + \varepsilon_{ihwt} \tag{4}$$

where y_{ihwt} is the individual labour market outcome of worker i at time t (wages or hours), A_{iht} is accessibility to employment from home ward h at time t, A_{iwt} is accessibility to employment from work ward w at time t, c_{ihwt} denotes the commuting costs between home and work at time t, X_{it} is a vector of personal and job characteristics, Z_{ht} is a matrix of home ward characteristics and W_{wt} is a matrix of work ward characteristics. ξ_t are year fixed effects that control for common shocks affecting all wards in a given year and ε_{ihwt} is the idiosyncratic error. Both the labour market outcomes and the accessibility indices are transformed to natural logarithms so we can interpret the coefficients as elasticities. As explained above, accessibility and commuting time is indeed lagged one year with respect to the outcome variable, as this is measured at the end of year t and accessibility and travel time at the beginning of it.

The dataset used does not provide information on travel time or distance travelled to workplace. I use the road networks between 2002 and 2008 (created as explained in the appendix) to calculate optimal travel time between ward of home and ward of work along the road network. We do not have any information on the travel mode of the workers. It could be the case that they are not commuting by road or not commuting using the optimal route predicted by the GIS software. However, using optimal travel time to proxy for commuting costs has the advantage of getting rid of potential measurement error on self-reported travel time [17]. In practice, I estimate a reduced-form effect of commuting costs, in which the measure of commuting cost is the optimal travel time through the road network. Moreover, as this travel time measure changes over time due to the road construction, it does not drop out when I include individual-home ward-work ward fixed-effects below, as opposed to straight geodesic distance (used for example by Graham & Melo, 2009, to approximate commuting costs).

I include this information in the estimated specification (4) in order to capture the effect of commuting costs on labour market outcomes. This way I can estimate the effect of accessibility from home and work conditional on commuting costs. This helps to interpret the results given the numerous theoretical channels through which transport policy can affect labour market outcomes, as discussed above. For example, some of the effects of transport policy on wages could be due to employers compensating workers for longer commutes and some could come through increased spatial competition or agglomeration externalities. By controlling for commuting costs in equation (4) we can be more certain that the effects of accessibility are not due to compensation for longer commutes as I am explicitly controlling for that. Moreover, the interpretation of the coefficient β_3 also informs us about the relationship between commuting costs and labour market outcomes.

[16] In fact, most of workers change ward of work when they change job (enterprise). In the robustness checks I control for job and home changes within wards and the results remain unchanged.

[17] Which is common, as noted for example by Gutiérrez-i-Puigarnau & van Ommeren (2010). Indeed, when tabulating travel times reported in LFS responses are disproportionably cumulated in values such 5, 10, 15, 20 and 30 minutes, likely due to rounding-up.

I am interested in parameters β_1 and β_2. We may be worried that there are (time invariant) unobservable individual characteristics that affect both individual labour market outcomes and accessibility indices (in levels) at the same time, and that are not included in X_{it}. For example, more able individuals may live or work in areas where accessibility and wages are higher. I include worker fixed effects μ_i to control for this (which in practice is equivalent to estimate the demeaned model). By introducing the individual fixed effects I estimate coefficients β_1 and β_2 using the variation of accessibility over time with respect to the average individual accessibility. At each point in time the worker may hold different jobs in different locations (w) and live in different places (h). Therefore the level of employment accessibility to which the worker is exposed at home and at work varies over time for three reasons: when home-work location changes, when employment changes or when road construction (that changes travel times) take place.

If workers are sorting spatially in order to take advantage of the changes in accessibility we would not be able to identify the separate effects on labour market outcomes which stem from (endogenous) sorting and those which are due to changes in accessibility for a given location (externalities or spatial competition). Sorting could be an outcome of accessibility or could be due to other unobservable reasons correlated with the labour market outcomes. For example, if workers with higher ability move to areas where accessibility and wages are growing faster, then the correlation between the changes in accessibility and the (demeaned) error term could be different from zero. The same could occur if more able workers choose jobs in areas where wages and accessibility are growing procyclically. For these reasons, even after controlling for unobservable time invariant characteristics of the individuals, there would still be reasons to think the estimates of β_1 and β_2 could be biased.

To investigate this issue I define individual-home ward-work ward specific fixed effects, μ_{ihw}. Replacing the individual fixed effects in (4) with μ_{ihw} and we obtain:

$$y_{ihwt} = \beta_1 A_{iht} + \beta_2 A_{iwt} + \beta_3 c_{ihwt} + \theta X_{it} + \delta Z_{ht} + \lambda W_{wt} + \mu_{ihw} + \xi_t + \varepsilon_{ihwt} \tag{5}$$

I can rewrite equation (5) in demeaned terms by subtracting the individual-fixed location means across time (focusing only on the accessibility measures):

$$(y_{ihwt} - \bar{y}_{i.}) = \beta_1 (A_{iht} - \bar{A}_{i.}) + \beta_2 (A_{iwt} - \bar{A}_{i.}) + (\varepsilon_{ihwt} - \bar{\varepsilon}_{i.}) \tag{6}$$

By estimating (5) we are identifying the effects of accessibility on individual labour market outcomes exploiting the changes of accessibility over time for an individual while keeping their home and work locations constant.

To implement this, in the panel, each individual is allocated a different fixed-effect depending in the pair work-home location. Therefore, there are two types of individuals: those that never move home-work location pair (at least while observed in the data) and those that eventually change ward of work, ward of residence or both. Individuals in this last group may be spatially relocating (sorting) due to changes in accessibility, so in contrast to (4), (5) does not use this variation for the estimation of the effects of accessibility [18]. That is, when including μ_{ihw}, coefficients β_1 and β_2 are not capturing the effect of endogenously determined relocations after the changes in accessibility take place (that could be cause by increased accessibility or other unobserved reasons). The comparison of the estimates

[18] In the robustness checks I obtain the results focusing on the group of non-movers and the results remain unchanged.

obtained by estimating (4) and (5) thus inform us about how accessibility affects residential and work location sorting.

However, there are still other possible sources of bias in the estimation of the effect of accessibility. Even when keeping the home-work location fixed, accessibility changes around the individual means due to changes in employment (numerator) and in travel times due to road construction (denominator). If changes in labour outcomes, for example wages, increase accessibility by means of attracting workers to wards around which the worker lives or works, there would be reverse causality problem which would challenge the validity of the estimates. There could also exist unobservable trends which affect both the location of employment and the labour market outcomes which could bias the estimates.

To overcome this, as before, I instrument accessibility A_{rt} using $\hat{A}_{rt}$, the accessibility measure that keeps employment fixed before the period of analysis (end of 2001). By doing this, we only use the changes in accessibility that affect optimal travel times (due to road construction) and not from employment growth. The instrument is strong by construction, because both the instrument and the instrumented variable use the variation which comes from road construction. For the instrument to be valid, we need the instrument to be uncorrelated with unobservable shocks that affect changes in employment.

This relates to the issue of the potential endogeneity of the placement of the new road links (which affect travel times used in the calculation of $\hat{A}_{rt}$). It could be argued that the instrument is not valid if transport projects (the construction of new links) are aimed at areas which are experiencing unobservable shock which are correlated with individual labour market outcomes. Transport investments may be taking place in areas in which workers would have done better anyway. Nevertheless, improving economic outcomes is not one of the key objectives of the road projects carried out by the Government [19]. Transport projects are generally aimed at a higher spatial scale and designed to improve safety or reduce congestion within a wider area [20].

However it could still be the case that new road links placements are endogenous to unobservable trends in labour market outcomes. To correctly identify the effects I define three distance bands which indicate if the ward is at 10, 20 or 30 kilometres from any scheme undertaken during the study period. Individuals placed within a given distance band are more likely to be exposed to similar shocks and road projects are quite unlikely to be aimed at specific individuals within these narrowly defined distance bands. This implies that changes in accessibility, conditional on controls, can be considered exogenous to individual workers located close to the new links. Moreover, even if the placement of the new roads would be endogenous to the individual labour market outcomes, its exact placing within the road network, and thus how the accessibility of the wards is affected by the new link, is likely to be exogenous to individual workers. Within these bands we compare workers who are close to projects at some point in time but that vary in the intensity and the timing of the treatment (change in accessibility), because this depends on the characteristics of the new roads and their location within the network. This is the variation from which we identify the effects. The main results are obtained using the 30 kilometres band from both home and work wards. The home-location and work-location bands can overlap or not, depending how separated are the locations of work and residence. Figure

[19] More details on British Transport Policy context are provided in Sanchis-Guarner (2012).

[20] For example, the objectives of scheme "A6 Clapham Bypass" in table A.2 were to "improve road safety", "relieve congestion" and "provide the opportunity for environmental improvement in Clapham by removing through traffic". See *http://www.highways.gov.uk/roads/projects/6006.aspx* for more information and the evaluation report.

A.1 in the appendix displays the location of wards located within these distance bands. I exclude the individuals which work and live within 1 kilometre of the projects.

In order to ensure that the instrument is uncorrelated with the underlying trends, I further control for differential trends in the vicinity of road schemes in the final specification. This is done by including a set of scheme dummies (31 schemes) interacted with year in equation (5). In some specifications I also control for the distance to the closest scheme within the distance bands (trends and levels). I also control for differential trends of the wards (home and work) based on 2001 characteristics. I used CENSUS data (provided by CASWEB) at the ward level to calculate the share of population aged 15-64 with higher education, mean age of population, share of population living on social housing, the rate of unemployment, proportion of workers commuting using motor vehicles and the average distance traveled to work. I also calculated a residential density measure, using address counts data in 2001 from the National Statistics Postcode Directory (NSPD) and the area of the wards in square kilometres, obtained from EDINA-UKBORDERS. I interact these characteristics with a linear trend. These trends control for differential growth in labour market outcomes, e.g. wages, depending on the level of these ward characteristics in 2001, before the period of analysis which starts in 2002. In the estimation of (4) I also introduce these ward characteristics in levels.

I furthermore control for individual personal and job characteristics. Some of these characteristics, for example full time status or occupation, could be regarded as "bad controls" because they could be outcomes of the transport policy. To help address this issue, I define the level of the characteristics at the beginning of the period (the first time the individual is observed within each of the two panel definitions) and I interact that level with a time-trend. By doing this I control for differential trends in the evolution of the labour outcome depending on the initial level of the job and personal characteristics. I use occupation trends (9 categories defined below), age group (10 year groups, from 16 to 65), full-time status and gender trends.

Finally, I implement 2-way clustering of the standard errors to correct for arbitrary within-group correlation of the individual shocks in two distinct non-nested categories. There are 4 dimensions in the data: individuals, time, ward of home and ward of work. In the case of the estimation of (4) the categories are year and individual: the resulting standard errors are robust to arbitrary within-panel autocorrelation (clustering on panel id) and to arbitrary contemporaneous cross-panel correlation (clustering on time). Ideally I would have clustered at the level of the "treatment" (accessibility changes), which are the wards of location. But these categories are nested because individuals change location over time. In the case of the estimation of (5) I can cluster the standard errors at the home-ward and work-ward level, because now these categories are not-nested across time, i.e. individuals keep the same home-work pair over time in the panels as defined in (5). By doing this, I allow the errors of the workers to be correlated at the treatment level (wards). In addition, the standard errors are also robust to arbitrary heteroskedasticity.

5.2. The Annual Survey of Hours and Earnings

The ASHE is an annual survey of the earnings of employees in Great Britain, which from 2004 replaced the New Earnings Survey (NES). Its primary purpose is to obtain information about the levels, distribution and make-up of earnings, and for the collective agreements that cover them. It is designed to represent all categories of employees in businesses of all kinds and sizes. The questionnaire is directed to the employer, who completes it on the basis of payroll records for the employee. The earnings, hours of work and other information relate to a specified week in April of each year.

ASHE is based on a survey of a 1% sample of employees on the Inland Revenue PAYE register (Pay As You Earn). The information is provided by the employer. The sample consists of employees whose National Insurance numbers end with two specific digits. It covers approximately 160,000 individuals a year. The survey is designed as a panel of workers, in which the same workers are observed for multiple years. The sample is replenished as workers leave the PAYE system (e.g. to self employment, retirement, overseas or death) and new workers enter it (e.g. from school, self-employment, immigration).

ASHE contains information on the make-up of weekly earnings and hours worked (basic, total and overtime), occupation (using Standard Occupational Classification – SOC), industrial sector (using Standard Industry Classification – SIC), collective agreement status, whether the job is private or public sector, age, gender, postcode of workplace, and from 2002, postcode of residence. I use the information on the postcodes to allocate accessibility from home and from work wards using the National Statistics Postcode Directory (NSPD).

In order to clean the data I drop the 0.5% top and bottom extreme values of the labour market outcome variables (wages and hours) and of the commuting times. I define total pay consistently over the whole period 2002–2008 [21]. I also removed observations with negative values of the variables and individuals which show inconsistency in their age or gender over time. I only keep main jobs for those individuals that have more than one job in the same year. Finally, I drop the individuals for which earnings were affected by absence (loss of pay) and those paid at trainee/junior rates. In the main results, I focus on employees working in the private sector. I believe that the flexibility of wages and hours would be greater in the private than in the public sector, because the latter might be more regulated or constraint to specific types of jobs. That said, in the robustness check I also include public sector workers and the results remain unchanged.

ASHE provides information on the occupation level of the individuals using SOC 2000 codes from 2002. I define broader occupation codes using the first digit of the code [22]. I also define five broader industrial categories based on 2-digit codes from the SIC 2003 classification [23].

The great advantage of this data is the good quality of the earnings and hours information and the detailed information on the geographical location of both the workplace and the place of residence. Furthermore, its panel structure allows us to control for unobservable time invariant characteristics of the workers which might be correlated with the variable of interest. However, the survey contains information only on workers who are employed, so I am unable to observe unemployment spells. It also contains no information on household characteristics (for example housing tenure status, civil status, number of children) which might be relevant in shaping the response of labour market outcomes to transport policy. For these reasons, in Section 6. I use an alternative dataset, the LFS, to study other labour market outcomes for which information is not provided in ASHE.

Table A.10 displays the number of observations by year and gender. I restrict the sample to those

[21] Different stratifications of ASHE define gross pay differently. I use the most recent definition: *total (gross) pay=basic pay + incentive pay + shift and premium payments + overtime pay + other pay*.

[22] They correspond to: 1 Managers and Senior Officials; 2 Professional Occupations; 3 Associate Professional and Technical Occupations; 4 Administrative and Secretarial Occupations; 5 Skilled Trades Occupations; 6 Personal Service Occupations; 7 Sales and Customer Service Occupations; 8 Process, Plant and Machine Operatives; 9 Elementary Occupations.

[23] They correspond to: 1 Manufacturing, 2 Construction, 3 Consumer services, 4 Producer services and 5 Other sectors.

individuals for which the home-ward and the work-ward are situated within 30 kilometres of any new link constructed during the period 2002–2008. I exclude those which are located (live and work) too close to the projects (1 km). There are around 240,000 observations, even after restricting the sample to locations close to the road projects. As expected, there are more male workers than female workers. Table A.11 displays the number of observations by year and industrial sector. I observe that most of the workers are concentrated in the service sectors and a large fraction in the manufacturing sector. Between 2002 and 2008 there is a decrease on the importance of the manufacturing sector while the proportion of workers employed in the producer services sector increases over time. Finally, Table A.12 displays the number of individuals existing in each of the panel definitions. If individuals are allowed to change location over time, μ_i, they appear an average of 3 times over the 7-year panel, and there are almost 80,000 individuals. When I define the fixed-effects as an individual-home-work combination, μ_{ihw}, there are over 114,000 fixed-effects, which appear an average of 2.09 times in the panel.

5.3. Results

5.3.1. Individual fixed effects

This section presents the results for the estimates of equation (4). The main results of the tables (columns 1-7) are obtained using the 30 kilometre band both from work-ward and home-ward. Columns 8 and 9 replicate the results of column 7 for the 20 and 10 kilometre distance bands. All the specifications include sic-year dummies (5 wide sector definition explained above) to control for year-specific shocks in each sector. These dummies also control for nation-wide changes in economic conditions, for example CPI levels. The standard errors are robust to heteroskedasticity and clustered at the individual and year levels.

Results both for basic outcomes (Table 2) and for total outcomes (Table 3) are reported. The difference between basic and total outcomes are incentive and overtime pay (in the case of wages) and over time (in the case of hours worked). The results are presented for weekly nominal wages, weekly hours worked and weekly hourly earnings. The relationship between the three outcomes is as follows:

$$wage_{ihwt} = \frac{wage_{ihwt}}{hours_{ihwt}} * hours_{ihwt} = hourly\_wage_{ihwt} * hours_{ihwt} \tag{7}$$

In both tables PANEL A displays the results for log weekly wages, PANEL B the results for log hourly earnings and PANEL C the results for log hours worked. Given the log-log specification, the parameters can be interpreted as elasticities. The rows at the bottom of the tables provide details on the estimated model.

Tables 2 and 3 have the same structure. Only the coefficients on log accessibility from work, log accessibility from home and log of travel time between home and work are displayed. The coefficients on the dummies and controls are not provided in order to improve the clarity in the exposition of the results.

Column 1 shows the results obtained by OLS, where only sic-by-year dummies are included. The coefficients are estimated precisely due to the large number of observations (above 200,000). In all cases I find a positive and significant relationship between accessibility from work and outcomes, negative relationships between the outcomes and accessibility from home and positive coefficients

PANEL A: **Log of basic weekly pay**									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log of accessibility	0.233***	0.072***	0.074***	0.072***	0.066***	0.045***	0.037***	0.014	0.015
from work ward	[0.007]	[0.008]	[0.008]	[0.009]	[0.010]	[0.010]	[0.010]	[0.013]	[0.033]
Log of accessibility	-0.087***	0.042***	0.042***	0.030***	0.031***	0.027***	0.021***	0.005	-0.031
from home ward	[0.006]	[0.009]	[0.009]	[0.009]	[0.009]	[0.010]	[0.008]	[0.010]	[0.037]
Log of travel time	0.111***	0.029***	0.029***	0.029***	0.029***	0.028***	0.024***	0.025***	0.019***
between work and home	[0.002]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.002]	[0.003]	[0.005]
Observations	245,883	221,066	221,066	221,066	221,066	221,066	216,324	123,468	30,392
Kleibergen-Paap F-stat			184,524	280,260	282,681	221,022	231,113	94,225	18,883
Individual no of clusters	81,330	56,513	56,513	56,513	56,513	56,513	55,723	32,086	8,077
Year no of clusters	7	7	7	7	7	7	7	7	7
PANEL B: **Log of basic hourly earnings**									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log of accessibility	0.212***	0.047***	0.048***	0.044***	0.041***	0.023***	0.021***	0.013	0.011
from work ward	[0.005]	[0.004]	[0.004]	[0.004]	[0.004]	[0.006]	[0.006]	[0.012]	[0.016]
Log of accessibility	-0.073***	0.032***	0.032***	0.024***	0.024***	0.024***	0.022***	0.011	0.022*
from home ward	[0.005]	[0.005]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]	[0.007]	[0.012]
Log of travel time	0.072***	0.015***	0.015***	0.014***	0.014***	0.014***	0.012***	0.012***	0.006**
between work and home	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.001]	[0.001]	[0.002]	[0.003]
Observations	244,166	219,347	219,347	219,347	219,347	219,347	214,628	122,513	30,132
Kleibergen-Paap F-stat			178,198	278,318	283,270	223,385	234,337	91,067	18,564
Individual no of clusters	80,984	56,165	56,165	56,165	56,165	56,165	55,378	31,889	8,023
Year no of clusters	7	7	7	7	7	7	7	7	7
PANEL C: **Log of basic weekly hours worked**									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log of accessibility	0.023***	0.024***	0.025***	0.027***	0.024***	0.019**	0.013*	-0.001	0.005
from work ward	[0.003]	[0.005]	[0.005]	[0.007]	[0.008]	[0.007]	[0.007]	[0.012]	[0.032]
Log of accessibility	-0.014***	0.01	0.01	0.006	0.007	0.002	-0.001	-0.005	-0.048
from home ward	[0.003]	[0.007]	[0.007]	[0.007]	[0.007]	[0.007]	[0.006]	[0.010]	[0.031]
Log of travel time	0.040***	0.015***	0.015***	0.015***	0.015***	0.015***	0.012***	0.014***	0.012***
between work and home	[0.002]	[0.001]	[0.001]	[0.002]	[0.002]	[0.002]	[0.001]	[0.002]	[0.003]
Observations	245,774	220,932	220,932	220,932	220,932	220,932	216,181	122,903	30,194
Kleibergen-Paap F-stat			173,610	275,394	281,027	221,303	232,776	92,211	18,604
Individual no of clusters	81,363	56,521	56,521	56,521	56,521	56,521	55,729	31,973	8,037
Year no of clusters	7	7	7	7	7	7	7	7	7
Distance band	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms	20 kms	10 kms
Individual fixed effects		YES	YES	YES	YES	YES	YES	YES	YES
Instrumented			YES	YES	YES	YES	YES	YES	YES
Scheme dummies and trends				YES	YES	YES	YES	YES	YES
Distance to schemes dummies and trends					YES	YES	YES	YES	YES
Ward 2001 characteristic and trends						YES	YES	YES	YES
Personal and job characteristic trends							YES	YES	YES

Notes: Personal and job characteristics include log of firm size (employment) and female, initial 1-dig occupation, initial age group, and full-time dummies interacted with a linear trend. Ward 2001 attributes include mean age of the population, proportion of household living in social housing, proportion of household with higher qualifications, unemployment rate, address density, average distance traveled to place of work and proportion of employees going to work by motor vehicles. All specification include six-year dummies (5 categories). 2-way clustering (individual and years). * p<0.1, ** p<0.05, *** p<0.01. **Source:** ONS, DfT, CASWEB and author's own calculations.

Table 2: Individual fixed effects results, basic outcomes – 2002–2008

PANEL A: Log of total (gross) weekly pay									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log of accessibility	0.223***	0.069***	0.071***	0.068***	0.063***	0.040***	0.032***	0.016	-0.001
from work ward	[0.006]	[0.008]	[0.008]	[0.010]	[0.010]	[0.011]	[0.011]	[0.013]	[0.032]
Log of accessibility	-0.084***	0.041***	0.043***	0.029***	0.030***	0.025**	0.019**	0.003	-0.041
from home ward	[0.006]	[0.009]	[0.009]	[0.009]	[0.009]	[0.010]	[0.008]	[0.010]	[0.026]
Log of travel time	0.112***	0.027***	0.027***	0.027***	0.027***	0.026***	0.022***	0.022***	0.016***
between work and home	[0.002]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.005]
Observations	245,968	221,171	221,171	221,171	221,171	221,171	216,427	123,482	30,404
Kleibergen-Paap F-stat			182,145	278,097	280,805	219,965	230,144	94,319	18,815
Individual no of clusters	81,350	56,553	56,553	56,553	56,553	56,553	55,764	32,096	8,083
Year no of clusters	7	7	7	7	7	7	7	7	7
PANEL B: Log of total hourly earnings									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log of accessibility	0.206***	0.043***	0.043***	0.037***	0.033***	0.023**	0.022*	0.029	-0.025
from work ward	[0.006]	[0.007]	[0.007]	[0.009]	[0.010]	[0.011]	[0.012]	[0.027]	[0.026]
Log of accessibility	-0.071***	0.060***	0.063***	0.061***	0.062***	0.062***	0.061**	0.092*	0.000
from home ward	[0.005]	[0.017]	[0.018]	[0.021]	[0.021]	[0.024]	[0.025]	[0.047]	[0.027]
Log of travel time	0.072***	0.010***	0.010***	0.010***	0.010***	0.010***	0.008**	0.006	0.003
between work and home	[0.002]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.004]	[0.004]
Observations	244,642	219,842	219,842	219,842	219,842	219,842	215,112	122,770	30,206
Kleibergen-Paap F-stat			181,165	276,945	281,034	218,998	230,035	92,354	18,674
Individual no of clusters	81,084	56,284	56,284	56,284	56,284	56,284	55,495	31,954	8,042
Year no of clusters	7	7	7	7	7	7	7	7	7
PANEL C: Log of total weekly hours worked									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log of accessibility	0.019***	0.033***	0.035***	0.032***	0.031***	0.019	0.014	-0.009	0.042
from work ward	[0.003]	[0.009]	[0.009]	[0.011]	[0.011]	[0.014]	[0.015]	[0.030]	[0.036]
Log of accessibility	-0.015***	-0.021	-0.022	-0.04	-0.04	-0.048	-0.053*	-0.088*	-0.04
from home ward	[0.004]	[0.022]	[0.022]	[0.028]	[0.028]	[0.031]	[0.030]	[0.047]	[0.025]
Log of travel time	0.041***	0.018***	0.018***	0.018***	0.018***	0.017***	0.015***	0.021***	0.014***
between work and home	[0.002]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.004]	[0.005]
Observations	246,234	221,388	221,388	221,388	221,388	221,388	216,629	123,183	30,267
Kleibergen-Paap F-stat			179,590	277,381	281,622	218,453	229,866	93,280	18,741
Individual no of clusters	81,466	56,620	56,620	56,620	56,620	56,620	55,826	32,039	8,054
Year no of clusters	7	7	7	7	7	7	7	7	7
Distance band	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms	20 kms	10 kms
Individual fixed effects		YES	YES	YES	YES	YES	YES	YES	YES
Instrumented			YES	YES	YES	YES	YES	YES	YES
Scheme dummies and trends				YES	YES	YES	YES	YES	YES
Distance to schemes dummies and trends					YES	YES	YES	YES	YES
Ward 2001 characteristic and trends						YES	YES	YES	YES
Personal and job characteristic trends							YES	YES	YES

Notes: Personal and job characteristics include log of firm size (employment) and female, initial 1-dig occupation, initial age group, and full-time dummies interacted with a linear trend. Ward 2001 attributes include mean age of the population, proportion of household living in social housing, proportion of household with higher qualifications, unemployment rate, address density, average distance traveled to place of work and proportion of employees going to work by motor vehicles. All specification include sector-year dummies (5 categories). 2-way clustering (individual and years). * p<0.1, ** p<0.05, *** p<0.01. **Source:** ONS, DfT, CASWEB and author's own calculations.

Table 3: Individual fixed effects results, total outcomes – 2002–2008

on travel time. The positive effect of accessibility from work on wages and hours could be due to agglomeration externalities or to the fact that professionals, who earn more and work full time, are concentrated in work locations where accessibility is higher. At the same time, workers in low paid jobs such as basic services could also be living in these locations and this could explain the negative coefficient of accessibility from home on earnings and hours. Finally, these results suggest that longer commutes are capitalised into higher wages and that workers who work longer hours are those who also commute longer. This explanation is in line with Manning (2003), who suggests that part-time workers do not find it worthwhile to commute for long as the relative time spent in commuting versus working is smaller than for full-time workers.

Column 2 adds individual fixed effects, which control for individual heterogeneity. Thus, the fixed-effects control for unobservable time-invariant characteristics of workers which might be correlated with labour market outcomes and accessibility. The coefficient for accessibility from workplace is reduced substantially, and that of accessibility from home becomes positive. This suggest that individuals which have unobservable characteristics correlated negatively with wages and hours are located in wards in which accessibility from home is higher. This is in line with the previous argument that less skilled workers live in denser areas in which they can access more jobs easily, as predicted for example by the spatial mismatch hypothesis. When I introduce the fixed effects, the coefficient of accessibility from home on hours worked becomes very small and insignificant, both for basic and total hours (PANEL C of both tables).

Column 3 shows the results when instrumenting A_{rt} using $\hat{A}_{iht}$. The instrument is very strong and the results remain very similar to those of column 2. Column 4 adds scheme dummies and trends, both from work and home. Column 5 controls for the distance to the closest scheme from both home and work, both in levels and in trends. Column 6 adds work-ward and home-ward 2001 characteristics and trends. The results remain very similar across specifications. As an aside, I find some evidence that workers in bigger firms earn more and work fewer hours.

Column 7 presents results from the most demanding specification. It includes fixed effects and all the controls, including personal and job characteristics both in levels and in trends (defined as initial characteristics, as explained above). For weekly and hourly earning I find positive and significant elasticities, both for basic and total outcomes. I find a weak effect of accessibility from work on basic hours worked and no effect on total hours. I find a marginally significant negative effect of accessibility from home on total hours worked. The estimates of the effect of accessibility from work are similar in magnitude to the estimates of the effect of market potential on wages provided for example by Combes et al. (2008) or Mion & Naticchioni (2009).

In columns 8 and 9 I reduce the sample to individuals located in the 20 and 10 kilometres bands around the wards of workplace and residence. Most of the effects of accessibility become insignificant, but the positive effect of commuting time on wages and hours remains significant. In fact, when I cluster the errors only at the individual level (results available on request) the coefficients of accessibility from work are significant, so the lack of significance could be due to the arbitrary contemporaneous cross-panel correlation, which inflates the standard errors.

As discussed above, these results could be driven by endogenous relocation of the individuals if they are spatially sorting in order to benefit from the accessibility changes or in order to reduce their commuting time. In the next section I investigate if using the variation of accessibility for individuals while they keep their location fixed has any impact on the estimates of the effects.

5.3.2. Individual-work-home fixed effects

This section presents the results of the estimation of equation (5), i.e. including individual-work-home fixed effects μ_{ihw}. The results are displayed in Table 4 for the basic outcomes and in Table 5 for the total outcomes. As before, I report the results for weekly pay (PANEL A), hourly earnings (PANEL B) and hours worked (PANEL C). The specifications are the same as in Section 5.3.1.. The rows at the bottom of the tables provide details on the estimated model, from OLS to the model with all the controls.

Column 1 displays the results obtained by OLS which reproduce those of column 1 of tables 2 and 3. Column 2 includes the individual-work-home fixed-effects, so I only use the variation in accessibility and travel time for individuals across time, for a given home-work location pair. The standard errors are bigger than those of the previous section, as we would expect due to the smaller amount of variation within the panels when using individual-home-work fixed-effects. After introducing the fixed-effects, the coefficient of log accessibility from home becomes insignificant and quite imprecisely estimated. The coefficient of log accessibility from work remains positive and significant for basic weekly wages and basic hourly earnings, but becomes insignificant in the rest of the panels. The coefficient of log travel time is only significantly different from zero in PANEL B of Table 4, and remains significant across specifications.

Column 3 instruments the accessibility indices with the measures which only use the variation stemming from road construction. The coefficient of the effect of log accessibility from work on wages and on hourly earnings increases substantially as compared to those of column 2. For accessibility from home the coefficients remain insignificant as in the rest of the specifications.

The difference in size of the coefficients obtained in column 2 and those obtained in column 3 is quite substantial. Specifically, the IV estimates of column 3 are one order of magnitude larger than the FE ones. It is possible that part of the difference is due to attenuation bias caused by measurement error in the accessibility indices. As explained above in Section A.1.2., the calculation of travel times used in the computation of A_{rt} is approximate. This could be introducing some measurement error causing an attenuation bias that the instrument could be helping to reduce. It is nevertheless more likely that most of the difference in the size of both coefficients is due to the existence of unobservable trends which are correlated both with accessibility from work and with individuals labour market outcomes. If employment is concentrating in areas in which wages are growing slower, which would be sensible because is cheaper to hire employees, the fixed-effects estimator would be downward biased.

Even after adding all the trends and controls in column 7, I find positive and highly significant effects of accessibility from work on basic weekly wages. In contrast, I find very weak effects on basic hours worked; column 7 of PANEL C of Table 4 displays a positive coefficient of the effect of accessibility from work on hours, but it is only significant at 10%. I find positive and significant effects of accessibility from work on basic hourly pay. Given that basic hourly wages are defined as weekly basic pay over weekly basic hours and basic hours worked seem not to be affected by changes in accessibility, the results on basic hourly earnings could be driven by increases in the "numerator" of the ratio. These results are robust when I restrict the sample to individuals located in the 20-kilometre distance bands, and the coefficient remains significant only for basic hourly earnings when I narrow the sample to the 10-kilometre band. Note however that the number of observations used in column

PANEL A: **Log of basic weekly pay**									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log of accessibility	0.233***	0.069*	0.314***	0.314***	0.342***	0.317***	0.328***	0.317***	0.151
from work ward	[0.013]	[0.037]	[0.116]	[0.112]	[0.112]	[0.113]	[0.116]	[0.122]	[0.152]
Log of accessibility	-0.087***	-0.007	-0.108	-0.107	-0.107	-0.076	-0.025	-0.064	0.09
from home ward	[0.013]	[0.032]	[0.091]	[0.094]	[0.097]	[0.095]	[0.092]	[0.095]	[0.136]
Log of travel time	0.111***	0.032	0.042	0.041	0.039	0.034	0.051	0.057*	0.079**
between work and home	[0.003]	[0.044]	[0.038]	[0.038]	[0.038]	[0.038]	[0.033]	[0.033]	[0.035]
Observations	245,883	183,091	183,091	183,091	183,091	183,091	180,105	105,168	26,935
Kleibergen-Paap F-stat			1,028	984	1,230	1,100	1,042	1,513	1,683
Work-ward no of clusters	3,794	3,557	3,557	3,557	3,557	3,557	3,551	2,226	755
Home-ward no of clusters	3,877	3,825	3,825	3,825	3,825	3,825	3,824	2,417	835
PANEL B: **Log of basic hourly earnings**									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log of accessibility	0.212***	0.049*	0.235***	0.233***	0.244***	0.224***	0.228***	0.232***	0.164**
from work ward	[0.012]	[0.030]	[0.059]	[0.058]	[0.058]	[0.060]	[0.056]	[0.058]	[0.073]
Log of accessibility	-0.073***	0.000	-0.058	-0.058	-0.065	-0.052	-0.02	-0.038	0.049
from home ward	[0.011]	[0.023]	[0.064]	[0.064]	[0.065]	[0.064]	[0.063]	[0.065]	[0.080]
Log of travel time	0.072***	0.040*	0.049**	0.049**	0.049**	0.049**	0.061***	0.063***	0.067***
between work and home	[0.002]	[0.021]	[0.023]	[0.022]	[0.022]	[0.023]	[0.019]	[0.019]	[0.016]
Observations	244,166	181,664	181,664	181,664	181,664	181,664	178,691	104,364	26,713
Kleibergen-Paap F-stat			1,035	988	1,236	1,104	1,047	1,526	1,745
Work-ward no of clusters	3,794	3,554	3,554	3,554	3,554	3,554	3,548	2,223	752
Home-ward no of clusters	3,877	3,824	3,824	3,824	3,824	3,554	3,823	2,417	835
PANEL C: **Log of basic weekly hours worked**									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log of accessibility	0.023***	0.02	0.099	0.098	0.114*	0.111	0.121*	0.104	0.013
from work ward	[0.004]	[0.029]	[0.070]	[0.068]	[0.068]	[0.067]	[0.065]	[0.069]	[0.080]
Log of accessibility	-0.014***	0.01	-0.013	-0.012	-0.005	0.015	0.034	0.014	0.083
from home ward	[0.005]	[0.027]	[0.052]	[0.052]	[0.054]	[0.053]	[0.050]	[0.051]	[0.077]
Log of travel time	0.040***	-0.009	-0.005	-0.005	-0.007	-0.012	-0.008	-0.003	0.014
between work and home	[0.001]	[0.032]	[0.029]	[0.029]	[0.029]	[0.029]	[0.027]	[0.027]	[0.032]
Observations	245,774	183,020	183,020	183,020	183,020	183,020	180,030	104,716	26,768
Kleibergen-Paap F-stat			1,023	979	1,223	1,092	1,035	1,527	1,749
Work-ward no of clusters	3,795	3,554	3,554	3,554	3,554	3,554	3,548	2,223	752
Home-ward no of clusters	3,878	3,827	3,827	3,827	3,827	3,827	3,826	2,417	835
Distance band	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms	20 kms	10 kms
Individual-work-home fixed effects		YES	YES	YES	YES	YES	YES	YES	YES
Instrumented			YES	YES	YES	YES	YES	YES	YES
Scheme trends				YES	YES	YES	YES	YES	YES
Distance to schemes trends					YES	YES	YES	YES	YES
Ward 2001 characteristic trends						YES	YES	YES	YES
Personal and job characteristic trends							YES	YES	YES

Notes: Personal and job characteristics include log of firm size (employment) and female, initial 1-dig occupation, initial age group, and full-time dummies interacted with a linear trend. Ward 2001 attributes include mean age of the population, proportion of household living in social housing, proportion of household with higher qualifications, unemployment rate, address density, average distance traveled to place of work and proportion of employees going to work by motor vehicles. All specification include sector-year dummies (5 categories). 2-way clustering (work-ward and home-ward). * p<0.1, ** p<0.05, *** p<0.01. **Source:** ONS, DfT, CASWEB and author's own calculations.

Table 4: Individual-home-work fixed effects results, total outcomes – 2002–2008

PANEL A: Log of total (gross) weekly pay

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log of accessibility	0.223***	0.061	0.240**	0.263**	0.285***	0.264**	0.302***	0.300**	0.124
from work ward	[0.013]	[0.039]	[0.110]	[0.105]	[0.103]	[0.106]	[0.111]	[0.120]	[0.144]
Log of accessibility	-0.084***	-0.035	-0.085	-0.059	-0.049	-0.025	-0.007	-0.047	0.054
from home ward	[0.013]	[0.035]	[0.082]	[0.085]	[0.090]	[0.087]	[0.086]	[0.090]	[0.117]
Log of travel time	0.112***	0.021	0.031	0.029	0.027	0.023	0.023	0.025	0.035
between work and home	[0.003]	[0.036]	[0.031]	[0.031]	[0.031]	[0.031]	[0.030]	[0.030]	[0.034]
Observations	245,968	183,188	183,188	183,188	183,188	183,188	180,200	105,199	26,947
Kleibergen-Paap F-stat			1,031	985	1,229	1,099	1,041	1,513	1,681
Work-ward no of clusters	3,794	3,557	3,557	3,557	3,557	3,557	3,551	2,226	755
Home-ward no of clusters	3,877	3,826	3,826	3,826	3,826	3,826	3,825	2,417	835

PANEL B: Log of total hourly earnings

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log of accessibility	0.206***	0.018	-0.021	-0.016	-0.014	-0.043	-0.037	-0.062	-0.252
from work ward	[0.012]	[0.047]	[0.139]	[0.137]	[0.138]	[0.139]	[0.143]	[0.153]	[0.214]
Log of accessibility	-0.071***	-0.009	0.154	0.157	0.171	0.19	0.215	0.239	0.445
from home ward	[0.011]	[0.047]	[0.191]	[0.191]	[0.192]	[0.192]	[0.198]	[0.209]	[0.310]
Log of travel time	0.072***	-0.025	-0.009	-0.012	-0.013	-0.015	-0.012	-0.006	0.061***
between work and home	[0.003]	[0.057]	[0.055]	[0.055]	[0.055]	[0.055]	[0.056]	[0.056]	[0.023]
Observations	244,642	182,082	182,082	182,082	182,082	182,082	179,101	104,593	26,775
Kleibergen-Paap F-stat			1,041	991	1,237	1,105	1,046	1,523	1,732
Work-ward no of clusters	3,794	3,554	3,554	3,554	3,554	3,554	3,548	2,223	752
Home-ward no of clusters	3,877	3,825	3,825	3,825	3,825	3,825	3,824	2,417	835

PANEL C: Log of total weekly hours worked

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log of accessibility	0.019***	0.043	0.281*	0.314**	0.341**	0.338**	0.375**	0.388**	0.423*
from work ward	[0.005]	[0.054]	[0.148]	[0.147]	[0.147]	[0.149]	[0.156]	[0.166]	[0.242]
Log of accessibility	-0.015***	-0.056	-0.223	-0.194	-0.201	-0.195	-0.204	-0.256	-0.385
from home ward	[0.005]	[0.063]	[0.217]	[0.217]	[0.218]	[0.218]	[0.224]	[0.234]	[0.361]
Log of travel time	0.041***	0.037	0.037	0.036	0.035	0.033	0.028	0.023	-0.04
between work and home	[0.002]	[0.070]	[0.062]	[0.062]	[0.062]	[0.062]	[0.062]	[0.061]	[0.036]
Observations	246,234	183,388	183,388	183,388	183,388	183,388	180,392	104,946	26,828
Kleibergen-Paap F-stat			1,028	982	1,225	1,094	1,036	1,525	1,739
Work-ward no of clusters	3,795	3,554	3,554	3,554	3,554	3,554	3,548	2,223	752
Home-ward no of clusters	3,878	3,827	3,827	3,827	3,827	3,827	3,826	2,417	835

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Distance band	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms	20 kms	10 kms
Individual-work-home fixed effects		YES	YES	YES	YES	YES	YES	YES	YES
Instrumented			YES	YES	YES	YES	YES	YES	YES
Scheme trends				YES	YES	YES	YES	YES	YES
Distance to schemes trends					YES	YES	YES	YES	YES
Ward 2001 characteristic trends						YES	YES	YES	YES
Personal and job characteristic trends							YES	YES	YES

Notes: Personal and job characteristics include log of firm size (employment) and female, initial 1-dig occupation, initial age group, and full-time dummies interacted with a linear trend. Ward 2001 attributes include mean age of the population, proportion of household living in social housing, proportion of household with higher qualifications, unemployment rate, address density, average distance traveled to place of work and proportion of employees going to work by motor vehicles. All specification include sector-year dummies (5 categories). 2-way clustering (work-ward and home-ward). * $p<0.1$, ** $p<0.05$, *** $p<0.01$. **Source:** ONS, DfT, CASWEB and author's own calculations.

Table 5: Individual-home-work fixed effects results, total outcomes – 2002–2008

9 is drastically reduced.

In the case of total labour market outcomes, reported in Table 5, the picture is slightly different. I find positive and significant effects of accessibility from work on weekly total wage. The coefficients are similar to those of PANEL A of Table 4. This would suggest that most of the effect on total pay can be attributable to the effect on basic pay. However, for total hours worked (PANEL C) I also find significant and positive effects of accessibility from work, while I did not find any effects on basic hours worked. This would suggest that the hours adjustment is working through overtime or through changes from part time to full time status. The estimates are very robust to the inclusion of controls and to the narrowing of the distance bands. As both wages and hours are adjusting, I do not find any effect on hourly earnings (PANEL B).

5.3.3. Interpretation of the results

Table 6 summarises the main results to facilitate the discussion of the interpretation of the size and sign of the coefficients. It reproduces some of the results of tables 2–3 and tables 4–5. For the 30-kilometre distance band Table 6 displays the coefficients obtained using OLS and using the two sorts of fixed-effects (individual and individual-work-home), with and without controls. Specification in column 7 shows the main findings. The results are very robust. A large set of robustness test are provided and discussed in the appendix in Section A.2..

PANEL A: Log of basic weekly pay							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log of accessibility	0.233***	0.072***	0.074***	0.037***	0.069*	0.314***	0.328***
from work ward	[0.007]	[0.008]	[0.008]	[0.010]	[0.037]	[0.116]	[0.116]
Log of accessibility	-0.087***	0.042***	0.042***	0.021***	-0.007	-0.108	-0.025
from home ward	[0.006]	[0.009]	[0.009]	[0.008]	[0.032]	[0.091]	[0.092]
Observations	245,883	221,066	221,066	216,324	183,091	183,091	180,105
Kleibergen-Paap F-stat			184,524	231,113		1,028	1,042
PANEL B: Log of total weekly hours worked							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log of accessibility	0.019***	0.033***	0.035***	0.014	0.043	0.281*	0.375**
from work ward	[0.003]	[0.009]	[0.009]	[0.015]	[0.054]	[0.148]	[0.156]
Log of accessibility	-0.015***	-0.021	-0.022	-0.053*	-0.056	-0.223	-0.204
from home ward	[0.004]	[0.022]	[0.022]	[0.030]	[0.063]	[0.217]	[0.224]
Observations	246,234	221,388	221,388	216,629	183,388	183,388	180,392
Kleibergen-Paap F-stat			179,590	229,866		1,028	1,036
Distance band	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms
Fixed effects		Indiv	Indiv	Indiv	Ind-w-h	Ind-w-h	Ind-w-h
Instrumented			YES	YES		YES	YES
All controls and trends				YES			YES

Notes: Personal and job characteristics include log of firm size (employment) and female, initial 1-dig occupation, initial age group, and full-time dummies interacted with a linear trend. Ward 2001 attributes include mean age of the population, proportion of household living in social housing, proportion of household with higher qualifications, unemployment rate, address density, average distance traveled to place of work and proportion of employees going to work by motor vehicles. All specification include sector-year dummies (5 categories). 2-way clustering (work-ward and home-ward). * p<0.1, ** p<0.05, *** p<0.01. **Source:** ONS, CASWEB, DfT and author's own calculations.

Table 6: Summary of main results – weekly basic wages and total hours worked

Two main conclusions can be drawn from the results of sections 5.3.1. and 5.3.2.. The first one is that labour market outcomes are affected by accessibility from work, but only the outcomes which are flexible, i.e. wages and overtime, while basic hours do not adjust. If accessibility from work is capitalising into higher wages, for example because of agglomeration externalities, conditional on commuting costs, workers could find it more worthwhile to work more hours [24]. If the number of basic work hours per week is fixed by contract, one way they can benefit from this wage increase is by working overtime. Another possible explanation is that workers are switching from part time to full time jobs, but I do not directly test this in the data.

The second one is that, for the correct identification and interpretation of the effects, it is key to account for workers spatial sorting. Comparison of the coefficients obtained in tables 2–3 and tables 4–5 informs us about the effect that controlling for sorting has on the estimates.

In tables 2–3, when I use the variation of accessibility which stems from workers relocating across wards, I find positive effects of accessibility both from work and from home. However, once I control for sorting, the results of tables 4–5 suggest that that improving access to jobs from home does not have any impact on labour market outcomes of employed workers whereas accessibility from work has substantial effects on wages and hours. To obtain these results I exploit the variation in accessibility over time for workers who remain in the same work-ward and live in the same home-ward. This finding suggests that the effects of accessibility from home on workers are partially driven by residential sorting. Workers might be moving into places in which accessibility is growing, because better connections makes these residential locations more desirable or because they can afford better housing in places that are further away from job centres and that are now more accessible.

Once individuals have chosen their residential location, access to more jobs from home does not seem to have an effect on their wages or hours, conditional on commuting time. The theoretical channel through which better accessibility from home would have an effect on wages and hours once controlling for sorting is unclear. The spatial mismatch hypothesis predicts positive effects of better access to jobs from home on labour market outcomes, especially on the probability of becoming employed or on the length of the unemployment spells. However, from the results obtained using the sample, in which all the individuals are already working, I do not find any evidence in favour of the spatial mismatch hypothesis. This could be due to households having already optimised their residential location in order to have access to better jobs and shorter commutes so further increases in accessibility from home do not have any impact on their wages or hours worked.

The positive effect of accessibility from work on wages and hours, given that tables 4–5 control for commuting travel time, could be working through spatial competition or agglomeration externalities. If firms in which the workers are employed can access a larger pool of workers, then employees might behave more competitively and work harder (specially high skilled workers, as suggested by Rosenthal & Strange, 2008). Also, workers could become more productive in these denser and better connected areas, as has been suggested and empirically verified in the agglomeration economies literature (see Duranton & Puga, 2004, for a review) or in the new economic geography literature (Krugman & Venables, 1995; Redding & Venables, 2004, for example).

In the context of the results, within the different mechanisms which affect industrial concentration and agglomeration [25], labour market pooling is possibly the most relevant channel through which agglomeration could be impacting wages and hours (for example Rosenthal & Strange, 2001, find that

[24] Unless they have a strong preference for leisure, which the results do not suggest.

[25] Although the different mechanisms are empirically hard to distinguish due to the "Marshallian equivalence", as noted by Duranton & Puga (2004) and Overman & Puga (2010).

this mechanism is the most relevant in determining industrial concentration). Even if I do not explicitly use a standard measure of labour market pooling which captures "the use of similar workers within an industry" (see Rosenthal & Strange, 2001; Overman & Puga, 2010; Jofré-Monseny et al., 2011, for examples), better accessibility to employment could potentially be improving the quality of job-workers matching (see Andersson et al., 2007, for some empirical evidence) or the search behaviour of the individuals (as suggested by Di Addario, 2011).

Concerning the size of the coefficients, the coefficient of the effect of accessibility from work on weekly wages (column 7 of PANEL A–Table 6) is larger (0.328) than previous estimates of the effect of Harris-type "market potential" measures on nominal wages. For example, the elasticities provided by Mion & Naticchioni (2005); Fingleton (2006); Combes et al. (2008); Amiti & Cameron (2007) and Graham & Melo (2009) are between 0.02 and 0.2. [26]

First of all, the measure of accessibility I use is not directly comparable to standard "market access" measures used in the literature. The papers mentioned above use a general market potential definition (Harris, 1954), which is similar in structure to ours but uses geodesic time invariant distance as the measure of proximity between locations. In their case, the identified effect of market access on wages stems from changes in the spatial distribution of employment/income over time, because distances between locations are kept fixed. Instead, I focus on a different channel of variation in the accessibility indices, which is changes in travel times (proximity) due to road construction. In the main results of Table 6, I instrument A_{rt} with $\hat{A}_{rt}$. In this way, I use a different source of variation for identification for the effects of accessibility from the authors above. Proximity between wards changes due to changes in optimal travel times between locations induced by road construction.

Spatial sorting could be explaining part of the difference in size between my estimates and the coefficients found in previous evidence. The identification strategy I follow substantially helps to reduce the bias caused by the spatial sorting of workers. As explained above, I control for sorting at the individual level (both residential and job sorting) by using the individual-home-work fixed effects. This way, for the identification of the effects I exploit the changes in accessibility over time for an individual while staying a given location combination. Moreover, as outlined above, the instrumental variables strategy helps to eliminate the bias induced by spatial relocation of workers across space which might be driven by the changes in accessibility. In other words, we tackle the sorting of workers at the ward level [27].

In fact in my results, the estimates that overlook sorting at the individual level and the endogenous spatial relocation of ward employment (columns 2 and 5 in Table 6), are similar in magnitude to previous estimates of the literature. Thus, the estimated elasticities of the effect of accessibility from work on weekly basic and total wages in PANEL A of tables 2–3 are 0.037 and 0.032 (in column 7, the preferred specification). Furthermore, even after controlling for individual sorting, when we do not

[26] Other authors, like Redding & Venables (2004); Hanson (2005); Head & Mayer (2006) or Hering & Poncet (2010) have used a measure of market access derived from a NEG model. Their estimate of the elasticity of nominal wages/GDP per capita with respect of market access are between 0.1–0.3.

[27] In their investigation of the determinants of individual wages using a large sample of French workers, Combes et al. (2008) find that differences in the skill composition of the labour force account for 40 to 50% of aggregate spatial wage disparities. They conclude that workers with better labour market characteristics tend to agglomerate in the larger, denser and more skilled local labour market. The intuition is that sorting of workers across space has an effect on individual wages because more skilled workers sort in specific areas. My approach is different as I explicitly control for the residential and job sorting of the individuals. Furthermore, I focus on accessibility changes stemming from road construction and do not use the variation coming from spatial changes in employment. This should help to reduce sorting issues in the same line as these authors.

control for spatial relocation of employment, e.g. in column 5 of Table 6, the estimated elasticity is 0.069, which is again in the same order of magnitude than previous findings.

If we compare the coefficients in column 5 to those in columns 6 and 7 of Table 6, once controlling for individuals sorting, when I instrument for spatial relocation of employment the estimated coefficients increase substantially. As explained above in Section 5.3.2., the downward bias in the fixed-effect estimates of column 5 could be due to employment concentrating around areas in which wages are growing slower. Conditional on other wards characteristics, firms would prefer to employ workers in wards in which the price of labour is growing slower. The instrument could also be reducing attenuation bias caused by measurement error in accessibility [28]. Compared to column 3, attenuation bias could be amplified in column 5 as the number individual-work-home fixed effects is much larger than before (see Table A.12). Besides, the precision of the estimates of columns 6 and 7 is lower because the amount of variation in accessibility for individuals that keep their home-work location pair fixed is quite small. Compared to columns 3 and 5, when I control for individual sorting and instrument in columns 6 and 7, the standard errors increase substantially, and it does so the confidence interval of the estimates. As a matter of fact, the confidence intervals of the coefficients of column 5 and 6 overlap and are not statistically different at 5% confidence level.

Finally, Table 7 provides some interpretation of the size of the estimates within the context of the analysis. It reports the mean change of log accessibility $\hat{A}_{rt}$ between 2002 and 2008 (growth rate). I focus on $\hat{A}_{rt}$ because the analysis focuses on changes in accessibility stemming from road construction. In the top panel it displays the growth in accessibility for wards situated within 10-20-30 kilometres of any road project (All projects) and in wards within the same distance bands from three specific projects undertaken in 2000, 2003 and 2006 the length of which is provided too. In the first place, from this table we see that the changes in accessibility around a project undertaken before the period of analysis (2000) is very small but not zero. This illustrates the fact that changes in accessibility, even when using only the variation due to road construction, affect the whole geography. A change in the network impacts travel times between wards depending on the relative position of the wards within the network and with respect to the rest of the road projects. For the schemes carried out in 2003 and 2006 the changes in accessibility are more substantial.

Additionally, this Table helps us to evaluate the size of the coefficients taking into account the average growth in accessibility within the context of the empirical exercise. I focus on the results obtained for basic weekly pay (which are very similar to the results for gross pay) and total weekly hours worked. The benchmark estimated elasticities are 0.328 for basic weekly pay and 0.375 for total hours, as presented in column 7 of Table 6. The average growth in accessibility for Great Britain between 2002 and 2008 (approximated by the change in logs), as displayed in Table 1, is 0.3%. The product of the elasticity and the growth in accessibility informs us of the relative growth in weekly wages and hours which can be attributed to the changes in accessibility (due to road construction). For the whole Great Britain, the estimates predict a growth in weekly wages between 2002–2008 due to changes in accessibility of 0.098% , and in total hours worked of 0.113%.

When we focus in wards closer to the projects larger impacts are found, as the growth in accessibility is bigger. The bottom panel of Table 7 provides the predicted effects on the 10-20-30 kilometres distance bands. For wards within 30 kilometres of projects, approximately 0.26% of the growth of weekly nominal wages and almost 0.3% of the growth in hours worked in Great Britain during

[28] Which is possible as annual changes in ward employment calculated using the BSD are likely to be slightly noisy.

	DISTANCE BAND		
PANEL A: Change in log accessibility $\hat{A}_{rt}$ between 2002–2008	*10 kms*	*20 kms*	*30 kms*
All projects	2.440%	1.120%	0.790%
2000 - M66 Denton - Middleton (15.3 kms)	0.011%	0.011%	0.030%
2003 - A5 Nesscliffe Bypass (21.48 kms)	3.622%	2.026%	1.292%
2006 - A1(M) Ferrybridge to Hook Moor (19.2 kms)	3.131%	2.108%	1.272%
PANEL B: Effect on basic weekly wages (coefficient=0.328)	*10 kms*	*20 kms*	*30 kms*
All projects	0.800%	0.367%	0.259%
2000 - M66 Denton - Middleton (15.3 kms)	0.004%	0.003%	0.010%
2003 - A5 Nesscliffe Bypass (21.48 kms)	1.188%	0.665%	0.424%
2006 - A1(M) Ferrybridge to Hook Moor (19.2 kms)	1.027%	0.691%	0.417%
PANEL C: Effect on total weekly hours (coefficient=0.375)	*10 kms*	*20 kms*	*30 kms*
All projects	0.915%	0.420%	0.296%
2000 - M66 Denton - Middleton (15.3 kms)	0.004%	0.004%	0.011%
2003 - A5 Nesscliffe Bypass (21.48 kms)	1.358%	0.760%	0.484%
2006 - A1(M) Ferrybridge to Hook Moor (19.2 kms)	1.174%	0.791%	0.477%

Source: ONS, DfT and author's own calculations.

Table 7: Effect of accessibility growth on wages and hours – 2002–2008

the period can be attributed to the changes in accessibility that stem from road construction (right column). These impacts increase the closer we get to the projects (10 and 20 kilometres distance bands). Around specific projects these effects are larger. For example in within 10 kilometres of the "A5 Nesscliffe Bypass" project, 1.2% of the growth in wages and 1.36% of the growth in total hours are attributable to changes in accessibility.

6. Further results: employment probability and commuting time

The results in Section 5. inform us about the effect of accessibility on labour market outcomes of workers which are already employed. But changes in accessibility from work and home wards could also affect other outcomes. This section provides evidence on the effect of changes in accessibility on the probability of being employed versus being unemployed or out of work. I also investigate if the effects of accessibility on labour market outcomes are working through the channel of reducing commuting costs, measured by travel time to workplace in the data. These results also inform us about the channels through which accessibility from work is affecting wages and hours.

6.1. Empirical specification and identification strategy

For the analysis of the effect of accessibility on other labour market outcomes I use data from the Labour Force Survey - LFS during the period 2002–2008. Due to the setup of this dataset, I use a slightly different identification strategy. I test the effect of accessibility on two outcomes: on employment status (two different dummy variables, which take the value 1 if the individual is employed and 0 if he is either unemployed or out of work) and on commuting time (which is only available in some quarters).

LFS provides quarterly observations for individuals for which there is information on their personal characteristics (gender, age, ethnicity, marital status, household composition, tenancy status, level of education, etc.) and their job characteristics (industrial sector, occupation, full time/part time status, public/private sector, main/secondary job, hours worked, income, etc.). Information on the home location of households at the ward level is provided. However, there is limited information on the location of jobs. When the individual is working, we know if the individual works and lives within the same local authority. For these reasons, I can only test the effect of accessibility from home. There is also information on their travel time to workplace or commuting time [29].

Even if the individual data has quarterly frequency, the accessibility measures only change on a yearly basis. Due to the design of the survey, we cannot define individual fixed effects [30]. Instead I can define so called "pseudo-fixed effects" ($\tilde{\mu}_i$) and estimate a pseudo-panel. This strategy is commonly used in survey data which has the same structure as the LFS, i.e. repeated cross-sections (see for example Nickell et al., 2002; Warunsiri & McNown, 2010; Dearden et al., 2011). We allocate a different fixed-effect to individuals with a combination of specific set of characteristics. I exploit the variation of different individuals around the mean of a "representative" individual, defined by the set of characteristics. Depending on how many different characteristics used to define these dummies, the estimation strategy becomes more demanding [31]. I define the pseudo fixed effects using gender, 10-year age group and ethnicity (the base category is white, male, aged 16-25).

I estimate the following specification:

$$y_{iht} = \beta_0 + \beta_1 A_{iht} + \theta X_{it} + \delta Z_{jt} + \lambda W_{ht} + \tilde{\mu}_i + \xi_t + \xi_m + \varepsilon_{iht} \tag{8}$$

where y_{iht} is a given labour market outcome for individual i, living in ward h at time t; A_{iht} is accessibility to employment from home, X_{it} is a matrix of individual characteristics (not included in the pseudo fixed-effects), Z_{jt} is a matrix of job characteristics (when employed), W_{ht} is a matrix of home-ward characteristics, $\tilde{\mu}_i$ denotes the pseudo fixed-effects as defined above, ξ_t and ξ_m are year and monthly dummies which control for common annual shocks and monthly seasonality and finally ε_{iht} denotes the idiosyncratic error term.

As for the results on wages and hours worked, I address the identification of coefficient β by estimating a reduced-form specification which uses $\hat{A}_{iht}$ (2001 employment) or by instrumenting A_{iht} with $\hat{A}_{iht}$. To tackle the endogeneity of the placement of the new links I only use information on individuals living within 30 kilometres from a project opened between 2002 and 2007.

6.2. The Labour Force Survey

I use data from the Labour Force Survey (LFS), which is a quarterly sample survey of households living at private addresses in Great Britain. Its purpose is to provide information on the UK labour market that can then be used to develop, manage, evaluate and report on labour market policies. The survey seeks information on respondents' personal circumstances and their labour market status during a specific reference period, normally a period of one week or four weeks (depending on the question) immediately prior to the interview. Data is available quarterly from 1992. The survey is based on household residing at a given address, not on specific individuals. It contains interviews

[29] And on the transport mode used to go to work, but only for very few observations.

[30] In some cases the same individual is surveyed twice in the same address in quarters 1 and 5 so for a sub-sample we can indeed construct a yearly individual panel, with only two observations per individual.

[31] We could for example define an individual-home ward pseudo fixed-effect, $\tilde{\mu}_{ih}$ which would exploit the variation for the individuals defined by the gender-age-ethnicity within a given location. For this we need enough observations within each ward for each category defined by the pseudo-fixed effect. Given the small spatial unit of analysis we are using this is not feasible.

of all members of the household in five consecutive quarters. The same address can host different individuals across the five quarters of interviews. Each quarter a new wave enters the survey and an old wave leaves. Table A.13 displays the number of observations by year and quarter.

The smallest geographical unit available for the location of the household is ward (as defined in 1998). The location of workplace is only reported at the regional level (for around 20 regions). However the LFS indicates if the worker works and lives in the same Local Authority District (of which there are 354 in Great Britain). The survey provides information on personal and household characteristics, level of education, sector, occupation, hours worked, travel to work and employment status, although information on earnings is not as good and reliable as that contained in ASHE [32]. From this LFS we can then recover information on unemployment or inactivity status. It also has information on the household composition and detailed information on personal characteristics. The main disadvantage of the survey is that it is not a panel. However, as discussed above, we can construct a panel of repeated cross sections for a given quarter and use the abundant individual information to construct a pseudo-panel and to control for characteristics like age and gender, occupation group, industry, etc.

6.3. Results

6.3.1. Travel time to workplace

Table 8 shows the results for the effect of log accessibility from home on travel time to workplace. I use data between 2002 and 2008. In columns 1 to 8 the results are obtained using a "reduced-form" estimation, i.e. using $\hat{A}_{iht}$ as the main dependent variable. Columns 9 and 10 use A_{iht} and in columns 11 and 12 I use individual fixed-effects instead of pseudo fixed-effects. The number of observations is much smaller than in section 6.3.2. because the information on travel to work is only available in the LFS 4^{th} quarter (Autumn).

The results are obtained using a pool of individuals (repeated cross-section) and include year, monthly dummies and pseudo-fixed effects. The standard errors are robust to heteroskedasticity and clustered at the ward level. Column 1 shows the results obtained by OLS. The parameter suggests a positive association between travel time and accessibility. This could be just because in denser areas, where accessibility is higher, workers travel longer for jobs because of traffic congestions. Column 2 introduces personal characteristics: student status, skill level, housing tenure, marital status, household composition, number of children [33]. Column 3 adds the job characteristics: occupation (1 digit), full time dummy, second job dummy, work from home dummy, public sector worker dummy and industrial sector dummies (2 digits). Columns 2 and 3 also show positive and significant effects of accessibility on travel time. However in these specifications we do not take into account that individuals might be changing jobs when accessibility from home is increased because in relative terms it becomes worthwhile traveling further.

Column 4 controls unobservables at the area level by including district dummies (Electoral wards group into districts. These correspond to Local/Unitary Authorities and there are around 350 in Great

[32] The response rate for the earning and hours questions is quite low and the data on earnings and hours is self-reported, and thus less reliable than employer provided data.

[33] The baseline category is 1st quarter 2002, white-male-16-25, without qualifications, household owns the house, single or living with other people, 1 adult with no children, managers and senior officials, working in sector 01:agriculture,hunting,etc.

Log of travel time to workplace												
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Log of accessibility	0.095***	0.090***	0.017***	-0.078***	-0.048***	-0.052***	-0.094	-0.418			0.112	
(2001 employment)	[0.007]	[0.006]	[0.005]	[0.012]	[0.012]	[0.012]	[0.353]	[0.355]			[0.784]	
Log of accessibility									-0.189	-0.027		0.106
(timevar employment)									[0.134]	[0.310]		[0.741]
Live and work in the			-0.698***	-0.766***	-0.765***	-0.765***	-0.776***	-0.776***	-0.776***	-0.776***	-0.579***	-0.579***
same LAD			[0.009]	[0.008]	[0.008]	[0.008]	[0.009]	[0.009]	[0.009]	[0.009]	[0.045]	[0.045]
Observations	81,114	81,114	81,114	81,114	81,114	81,114	81,114	81,114	81,114	81,050	15,734	15,734
Adjusted R2	0.0298	0.06	0.28	0.312	0.313	0.314	0.287	0.287	0.287	0.231		
F-stat	44.1	66.9		144	139	133				172	82.3	82.3
No of clusters	4,066	4,066	4,066	4,066	4,066	4,066	4,066	4,066	4,066	4,003	2,664	2,664
Kleibergen-Paap F-stat										3,301		652
Distance band	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms
Year & month dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Personal characteristics		YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Job characteristics			YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
District dummies				YES	YES	YES						
Ward attributes/trends					YES	YES		YES	YES	YES		
Scheme dummies/trends						YES		YES	YES	YES	YES	YES
Ward fixed-effects							YES	YES	YES	YES		
Instrumented										YES	YES	YES
Individual fixed-effects											YES	YES

Notes: Clustered (ward) s.e. All specifications include year, monthly dummies & pseudo-fe (age, gender, white). Distance band 30 kms. Personal characteristics include student status, skill level, housing tenure, marital status, household composition, number of children. Job characteristics include occupation (1 digit), full time dummy, second job dummy, work from home dummy, public sector worker dummy and industrial sector dummies (2 digits). Ward 2001 attributes include distance to closest transport scheme, mean age of the population, proportion of household living in social housing, proportion of household with higher qualifications, average distance traveled to place of work, proportion of households with cars or vans, proportion of employees going to work by motor vehicles, area. * $p<0.05$, ** $p<0.01$, *** $p<0.001$. **Source:** ONS, DfT, CASWEB and author's own calculations.

Table 8: Travel time to workplace results 2002–2008

Britain) and the coefficient becomes negative and significant. The district dummies control by time-invariant unobservable characteristics of the districts which might be correlated with accessibility and commuting time at the same time. These could include aspects related to road infrastructure like the district initial endowment of roads, its quality or its level of congestion. Column 5 and 6 add ward 2001 attributes (in levels and in trends) and scheme dummies and trends. Column 6 controls for district fixed-effects, ward attributes and scheme dummies: it provides substantial evidence of a negative relationship between accessibility and commuting time. In columns 7 and 8, when I introduce the ward fixed-effects, the estimates become very imprecise, possibly due to the lack of variation across pseudo-fixed effects within each ward.

One might think that the lack of significant results in column 8 is due to lack of variation in the instrument $\hat{A}_{iht}$. To test this, column 9 uses the accessibility measure with time-varying employment, A_{iht}, but the estimates are still insignificant. In column 10 I use instrumental variables, and the results are even less significant (smaller parameter and higher standard error). In columns 11 and 13 we use a sub-sample of observations, those for which we can observe the individual in two consecutive years. The coefficients increase substantially but also do the standard errors, and the effect of home accessibility on travel time to workplace is estimated very imprecisely.

These results suggest a negative partial correlation between employment accessibility from home and commuting time, but given the characteristics of our data, more demanding causal parameters (from column 7) are imprecisely estimated. This implies that cannot rule out a zero causal effect of accessibility on commuting time. However, specification in column 6 is quite demanding and still yields a negative and significant coefficient, providing quite substantial evidence that increased accessibility from home reduces travel time to workplace.

6.3.2. Employment status

This section provides estimates of the effect of accessibility on the probability of being employed versus being out of work (Table 8) and versus being unemployed (Table 9). This results are interesting for two reasons. First of all they are a more direct test of the spatial mismatch hypothesis discussed in Section 2.. Secondly, the estimation of the effect of one variable on individual wages is always subject to problems of sample selection bias. As discussed by Heckman (1979), we only observe wages of workers which are employed, and those might be positively selected into employment due to unobserved factors. If we find no effect of accessibility on the probability of being employed this would, at least partially, reduce the concern of the estimates in Section 5.3.2. being biased due to sample selection.

As in the previous section, columns 1 to 7 estimate a reduced-form specification. Columns 8 and 9 use A_{iht} and columns 10 and 11 use individual fixed-effects. I estimate a linear probability model (LPM). It is worth noting that the standard errors in Table 9 are more precise than before because the number of observations is larger as I use individuals sampled in all quarters.

Column 1 shows a negative significant association between accessibility and the probability of being employed. This could be because individuals with specific characteristics which make them "less employable" locate in places where accessibility is higher. Column 2 controls for personal characteristics. The effects in both panels become weaker and insignificant. Column 3 adds district dummies. The coefficients of accessibility in both panels (for employment status when 0 is out of work) be-

PANEL A: Employment status (1 employed, 0 out of work)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Log of accessibility	-0.007***	0.002	0.003	0.003	0.004	-0.131	-0.028			0.078	
(2001 employment)	[0.002]	[0.001]	[0.003]	[0.004]	[0.004]	[0.087]	[0.095]			[0.104]	
Log of accessibility								-0.024	-0.126**		0.075
(timevar employment)								[0.036]	[0.056]		[0.100]
Observations	836,364	836,364	836,364	836,364	836,364	836,364	836,364	836,364	836,361	186,338	186,338
Adjusted R2	0.0836	0.23	0.233	0.234	0.234	0.205	0.205	0.205	0.140		
F-stat	358	825	790	744	600	775	545	545	1017	21.9	22
No of clusters	4,110	4,110	4,110	4,110	4,110	4,110	4,110	4,110	4,112	4,088	4,088
Kleibergen-Paap F-stat									19,242		3,623

PANEL B: Employment status (1 employed, 0 unemployed)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Log of accessibility	-0.004***	0.000	0.000	0.001	0.001	-0.025	0.017			0.034	
(2001 employment)	[0.001]	[0.001]	[0.002]	[0.002]	[0.002]	[0.036]	[0.037]			[0.052]	
Log of accessibility								-0.037**	0.003		0.032
(timevar employment)								[0.019]	[0.023]		[0.050]
Observations	656,960	656,960	656,960	656,960	656,960	656,960	656,960	656,960	656,957	140,962	140,962
Adjusted R2	0.0369	0.0808	0.0825	0.083	0.0831	0.0678	0.0682	0.0682	0.0334		
F-stat	87	105	103	94.8	77.4	97.6	69.2	69.2	130	1.28	1.28
No of clusters	4,110	4,110	4,110	4,110	4,110	4,110	4,110	4,110	4,111	4,075	4,075
Kleibergen-Paap F-stat									20,487		2,660

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Distance band	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms	30 kms
Year & month dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Personal characteristics		YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
District dummies			YES	YES	YES						
Ward attributes/trends				YES	YES		YES	YES	YES		
Scheme dummies/trends					YES		YES	YES	YES	YES	YES
Ward fixed-effects						YES	YES	YES	YES		
Instrumented									YES	YES	YES
Individual fixed-effects										YES	YES

Notes: Clustered (ward) s.e. All specifications include year, monthly dummies & pseudo-fe (age, gender, white). Distance band 30 kms. Personal characteristics include student status, skill level, housing tenure, marital status, household composition, number of children. Ward 2001 attributes include distance to closest transport scheme, mean age of the population, proportion of household living in social housing, proportion of household with higher qualifications, average distance traveled to place of work, proportion of households with cars or vans, proportion of employees going to work by motor vehicles, area. * $p<0.05$, ** $p<0.01$, *** $p<0.001$. **Source:** ONS, DfT, CASWEB and author's own calculations.

Table 9: Employment status results 2002–2008

comes positive but are insignificant. Column 4 adds ward attributes which are obtained from the 2001 census. Column 5 adds scheme trends and the coefficients remain insignificant.

In columns6 to 9 I add ward fixed-effects to control for unobservable factors at the ward level which might be correlated with employment and accessibility. The coefficient becomes very imprecise and the standard errors increase considerably.Column 7 adds scheme trends and ward attributes trends, and the coefficients become even less precise. As for the results of the effect on travel time, it is very likely that the big standard errors of columns 6 and 7 are due to the lack of variation of the data once I only exploit the within ward variation, because there are not sufficient observations in each ward in each year and in each pseudo-fixed effect category to be able to identify the parameters.

Column 8 uses the accessibility measure with time-varying yearly employment A_{iht} in order to have more variation in the data. The results are more precise than in column 7 as expected, because accessibility changes much more from year to year when I use yearly employment. The coefficient in PANEL A (0 if out of work) remains insignificant but that in PANEL B (0 if unemployed), become significant at 5% confidence level. This could be do to endogenous relocation of employment across the space. To control for this, in column 9 I instrument A_{iht} with $\hat{A}_{iht}$. Now the situation of column 8 is reverse. PANEL A suggests that accessibility from home has a negative effect on the probability of being employed for active individuals, while PANEL B suggests no effect if the alternative is being out of work (unemployed but also inactive). This results is puzzling and contrary to previous evidence in favour of the spatial mismatch hypothesis. To test if this could be due to unobserve individual heterogeneity, columns 10 and 11 include individual fixed-effects, for the reduced-form and the IV estimation. The coefficients are precisely estimated, close to zero and insignificant. This suggests that in fact there is no effect of accessibility from work on the probability of being employed.

To summarise, the results in sections 6.3.1. and 6.3.2. provide some evidence that accessibility to employment from home reduces travel time to workplace but it does not have any effect on the probability of being employed. These results ignore any effect that accessibility might have on relocation because we do not observe individuals over time. These results are consistent with those found is section 5., in which we found that accessibility from home does not have an effect on labour market outcomes once I only use the variation for a given location.

7. Conclusion

In this paper I investigate the effect of changes in accessibility, induced by road construction, on individual labour market outcomes. I make use of rich individual datasets and small geographical scale to be able to infer causality on the estimates.

I provide evidence of the effect of accessibility both from workplace and residence on individual wages and hours worked. Controlling for commuting time allows us to learn about the potential theoretical channels through which accessibility might be impacting on labour market outcomes. The methodology used overcomes several endogeneity problems common in the identification of effects of access to jobs: unobserved individual heterogeneity, endogeneity of the residential and job location, reverse causation between accessibility and spatial distribution of local employment and endogenous placement of the new road links. After controlling for sorting, I find positive effects of accessibility from work on earnings and total hours. These results could be driven by agglomeration externalities which are capitalised into higher nominal wages and by increased spatial competition which might make employees work longer hours. Workers could also be responding to higher nominal wages by increasing overtime or switching from part time to full time jobs. These results are very robust to different specifications and across different groups of workers.

Two main conclusion can be drawn from these results. First, changes in accessibility impact workers by affecting "flexible" outcomes, i.e. wages and overtime. These results are consistent with what would be expected if basic hours are fixed by contract. If higher accessibility from workplace is capitalised into higher wages, workers could be increasing the number of hours work to benefit from the wage increase, but the margin in which they can adjust this is through over time. Secondly, my identification strategy allows us to separate the effect of spatial sorting from other channels through which accessibility might be impacting labour market outcomes. As discussed in the text, once I focus on the variation of accessibility for a given work-home location combination, there is no effect of accessibility from home on hours or wages. But the effect of accessibility from workplace is significant even once controlling for sorting, and the size of the coefficient increases substantially. This result stresses the importance of accounting for sorting to be able to correctly attribute the changes in earnings and hours to changes in accessibility. It also suggest that workers sort residential location in response to changes in accessibility.

I also investigate the effect that changes in accessibility might have on employment status and travel times. I find no effect of accessibility from home on the probability of being employed. This rules out the existence of sample selection bias on the estimation of the effect on wages and hours, at least selection into employment due to changes in accessibility. They also provide robust evidence which does not support the spatial mismatch hypothesis. I also find some negative correlations between accessibility from home and travel time to workplace, but these become imprecise when we become more demanding in terms of the identification.

The main limitation of the paper is the lack of substantial evidence on how accessibility affects residential and job spatial sorting. Using individual-home-work fixed-effects allows to focus on the variation stemming only from changes in accessibility for a given home-work location pair choice. Comparing the results using this approach and a standard individual fixed-effects estimation (which allows variation in accessibility to arise from job and home relocations) sheds some light on the effect of accessibility on residential location. The fact that the effect of home accessibility on hours and wages becomes insignificant when keeping location fixed suggest that individuals sort home location as a results of accessibility changes. But the causal relationship between accessibility and residential and job location is relevant on its own and requires further investigation.

The findings in this paper provide new evidence on the effect of agglomeration externalities and proximity to markets on individual wages and hours worked using a novel strategy that carefully tackles multiple endogeneity issues and a detailed dataset on road construction. I use changes in optimal travel times between location stemming from road construction as the source of changes in accessibility. Road construction can be directly used by policy makers to influence travel times and therefore affect effective density and proximity between locations. Transport policy is a substantial part of economic policy and the estimates of this paper help to shed light on the economic impacts that transport infrastructure investments can have on individual labour market outcomes.

References

Ahlfeldt, G. (2011a). If Alonso Was Right: Modeling Accessibility And Explaining The Residential Land Gradient. *Journal of Regional Science*, 51(2), 318–338.

Ahlfeldt, G. (2011b). If We Build, Will They Pay? Predicting Property Price Effects of Transport Innovations. *SERC Discussion Papers*, 0075.

Ahlfeldt, G. & Wendland, N. (2011). Fifty Years of Urban Accessibility: the Impact of the Urban Railway Network on the Land Gradient in Berlin 1890-1936. *Regional Science and Urban Economics*, 41(2), 77–88.

Amiti, M. & Cameron, L. (2007). Economic Geography and Wages. *The Review of Economics and Statistics*, 89(1), 15–29.

Andersson, F., Burgess, S., & Lane, J. I. (2007). Cities, Matching and the Productivity Gains of Agglomeration. *Journal of Urban Economics*, 61(1), 112–128.

Aslund, O., Östh, J., & Zenou, Y. (2010). How important is access to jobs? Old question–improved answer. *Journal of Economic Geography*, 10(3), 389–422.

Baum-Snow, N. (2007a). Did Highways Cause Suburbanization? *The Quarterly Journal of Economics*, 122(2), 775–805.

Baum-Snow, N. (2007b). Suburbanization and transportation in the monocentric model. *Journal of Urban Economics*, 62(3), 405–423.

Baum-Snow, N. (2010). Changes in Transportation Infrastructure and Commuting Patterns in US Metropolitan Areas, 1960-2000. *American Economic Review*, 100(2), 378–82.

Behrens, K., Lamorgese, A., Ottaviano, G. I. P., & Tabuchi, T. (2004). *Testing the Home Market Effect in a Multi-Country World: The Theory*. CEPR Discussion Papers 4468, Centre for Economic and Policy Research.

Chandra, A. & Thompson, E. (2000). Does public infrastructure affect economic activity?: Evidence from the rural interstate highway system. *Regional Science and Urban Economics*, 30(4), 457–490.

Ciccone, A. & Hall, R. E. (1996). Productivity and the density of economic activity. *American Economic Review*, 86(1), 54–70.

Combes, P.-P., Duranton, G., & Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, 63(2), 723–742.

Combes, P.-P., Duranton, G., & Gobillon, L. (2011). The identification of agglomeration economies. *Journal of Economic Geography*, 11(2), 253–266.

Combes, P.-P., Duranton, G., Gobillon, L., & Roux, S. (2010). Estimating agglomeration economies with history, geology, and worker effects. In *Agglomeration Economics*, NBER Chapters (pp. 15–65). National Bureau of Economic Research, Inc.

Combes, P.-P. & Lafourcade, M. (2001). *Transport Cost Decline and Regional Inequalities: Evidence from France*. CEPR Discussion Papers 2894, Centre for Economic and Policy Research.

Dearden, L., Fitzsimons, E., & Wyness, G. (2011). *The Impact of Tuition Fees and Support on University*. IFS Working Papers W11/17, Institute for Fiscal Studies.

Department of Transport (1998a). *A new Deal for Transport: Better for Everyone*. White paper - Department of Transport.

Department of Transport (1998b). *A New Deal for Trunk Roads in England: Understanding the New Approach*. Report - Department of Transport.

Détang-Dessendre, C. & Gaigné, C. (2009). Unemployment duration, city size, and the tightness of the labor market. *Regional Science and Urban Economics*, 39(3), 266–276.

Di Addario, S. (2011). Job search in thick markets. *Journal of Urban Economics*, 69(3), 303–318.

Donaldson, D. (2010). Railroads of the raj: Estimating the impact of transportation infrastructure. *NBER Working Papers*, 16487.

Donaldson, D. & Hornbeck, R. (2012). Railroads and American Economic Growth: A "Market Access" Approach. Working Paper, MIT - Massachusetts Institute of Technology.

Duranton, G., Morrow, P., & Turner, M. A. (2011). Roads and trade: Evidence from the US. Mimeo, University of Toronto.

Duranton, G. & Puga, D. (2004). Micro-foundations of urban agglomeration economies. In J. V. Henderson & J. F. Thisse (Eds.), *Handbook of Regional and Urban Economics*, volume 4 of *Handbook of Regional and Urban Economics* chapter 48, (pp. 2063–2117). Elsevier.

Duranton, G. & Turner, M. A. (2011a). The fundamental law of road congestion: Evidence from US cities. *American Economic Review*, 101(6), 2616–52.

Duranton, G. & Turner, M. A. (2011b). Urban growth and transportation. Mimeo, University of Toronto.

El-Geneidy, A. M. & Levinson, D. M. (2006). *Access to destinations: Development of accessibility measures*. Report, Research Services Section. Mn/DOT 2006-16, University of Minnesota. Dept. of Civil Engineering and Minnesota. Dept. of Transportation.

Faber, B. (2012). *Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System*. Unpublished, London School of Economics.

Fernald, J. G. (1999). Roads to prosperity? Assessing the link between public capital and productivity. *American Economic Review*, 89(3), 619–638.

Fingleton, B. (2006). The new economic geography versus urban economics: an evaluation using local wage rates in great britain. *Oxford Economic Papers*, 58(3), 501–530.

García-Milá, T., McGuire, T. J., & Porter, R. H. (1996). The effect of public capital in state-level production functions reconsidered. *The Review of Economics and Statistics*, 78(1), 177–80.

Gautier, P. A. & Zenou, Y. (2010). Car ownership and the labor market of ethnic minorities. *Journal of Urban Economics*, 67(3), 392–403.

Gibbons, S., Lyytikäinen, T., Overman, H., & Sanchis-Guarner, R. (2012). New road infrastructure: the effects on firms. *SERC Discussion Papers*, (0117).

Gibbons, S. & Machin, S. (2005). Valuing rail access using transport innovations. *Journal of Urban Economics*, 57(1), 148–169.

Gibbons, S. & Machin, S. (2006). *Transport and Labour Market Linkages: empirical evidence, implications for policy and scope for further research*. Background paper for the Eddington report to the Department of Transport.

Glaeser, E. L. & Kohlhase, J. (2003). Cities, regions and the decline of transport costs. *Papers in Regional Science*, 83(1), 197–228.

Gobillon, L., Magnac, T., & Selod, H. (2011). The effect of location on finding a job in the paris region. *Journal of Applied Econometrics*, 26(7), 1079–1112.

Gobillon, L., Selod, H., & Zenou, Y. (2007). The mechanisms of spatial mismatch. *Urban Studies*, 44(12), 2401–2427.

Graham, D. J. (2007). Variable returns to agglomeration and the effect of road traffic congestion. *Journal of Urban Economics*, (pp. 103–120).

Graham, D. J., Gibbons, S., & Martin, R. (2009). *Transport investments and the distance decay of agglomeration benefits*. Draft – Department of Transport.

Graham, D. J. & Melo, P. C. (2009). Agglomeration economies and labour productivity: Evidence from longitudinal worker data for gbs travel-to-work areas. *SERC Discussion Papers*, (0031).

Gramlich, E. M. (1994). Infrastructure investment: A review essay. *Journal of Economic Literature*, 32(3), 1176–96.

Gutiérrez-i-Puigarnau, E. & van Ommeren, J. N. (2010). Labour supply and commuting. *Journal of Urban Economics*, 68(1), 82–89.

Hanson, G. H. (2005). Market potential, increasing returns and geographic concentration. *Journal of International Economics*, 67(1), 1–24.

Harris, C. D. (1954). The Market as a Factor in the Localization of Industry in the United States. *Annals of the association of American geographers*, 44(4), 315–348.

Head, K. & Mayer, T. (2006). Regional wage and employment responses to market potential in the EU. *Regional Science and Urban Economics*, 36(5), 573–594.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1), 153–61.

Hering, L. & Poncet, S. (2010). Market access and individual wages: Evidence from China. *The Review of Economics and Statistics*, 92(1), 145–159.

Highways Agency (2009). *The Highways Agency: Framework Document*. Highways Agency - Department of Transport.

Holl, A. (2012). Market potential and firm-level productivity in Spain. *Journal of Economic Geography*, Forthcoming.

Hymel, K. (2009). Does traffic congestion reduce employment growth? *Journal of Urban Economics*, 65(2), 127–135.

Ihlanfeldt, K. R. (2002). Spatial mismatch in the labor market and racial differences in neighborhood crime. *Economics Letters*, 76(1), 73–76.

Ihlanfeldt, K. R. & Sjoquist, D. L. (1990). Job accessibility and racial differences in youth employment rates. *American Economic Review*, 80(1), 267–76.

Inhanfeldt, K. R. (2006). A primer on spatial mismatch within urban labor markets. In *A companion to urban economics*, Blackwell companions to contemporary economics (pp. 404 – 417). Wiley-Blackwell.

Jiwattanakulpaisarn, P., Noland, R. B., Graham, D. J., & Polak, J. W. (2009). Highway infrastructure and state-level employment: A causal spatial analysis. *Papers in Regional Science*, 88(1), 133–159.

Jofré-Monseny, J. (2009). The scope of agglomeration economies: Evidence from catalonia. *Papers in Regional Science*, 88(3), 575–590.

Jofré-Monseny, J., Marín-López, R., & Viladecans-Marsal, E. (2011). The mechanisms of agglomeration: Evidence from the effect of inter-industry relations on the location of new firms. *Journal of Urban Economics*, 70(2-3), 61–74.

Krugman, P. & Venables, A. J. (1995). Globalization and the inequality of nations. *The Quarterly Journal of Economics*, (pp. 857–880).

Krugman, P. R. (1991). Increasing returns and economic geography. *Journal of Political Economy*, 99(3), 483–99.

Manning, A. (2003). The real thin theory: monopsony in modern labour markets. *Labour Economics*, 10(2), 105–131.

Manning, A. (2006). A generalised model of monopsony. *Economic Journal*, 116(508), 84–100.

Manning, A. & Petrongolo, B. (2011). *How local are labor markets? Evidence from a spatial job search model*. CEPR Discussion Papers 8686, Centre for Economic and Policy Research.

Marshall, A. (1890). *Principles of Economics*. London: Macmillan.

Martin, P., Mayer, T., & Mayneris, F. (2011). Spatial concentration and plant-level productivity in france. *Journal of Urban Economics*, 69(2), 182–195.

Matas, A., Raymond, J.-L., & Roig, J.-L. (2010). Job accessibility and female employment probability: The cases of barcelona and madrid. *Urban Studies*, 47(4), 769–787.

Melo, P. C., Graham, D. J., & Noland, R. B. (2010). Impact of transport infrastructure on firm formation: Evidence from the portuguese municipalities. *Transportation Research Record*, 2163, 133–143.

Michaels, G. (2008). The effect of trade on the demand for skill: Evidence from the interstate highway system. *The Review of Economics and Statistics*, 90(4), 683–701.

Mion, G. & Naticchioni, P. (2005). *Urbanization Externalities, Market Potential and Spatial Sorting of Skills and Firms*. CEPR Discussion Papers 5172, Centre for Economic and Policy Research.

Mion, G. & Naticchioni, P. (2009). The spatial sorting and matching of skills and firms. *Canadian Journal of Economics*, 42(1), 28–55.

Mulalic, I., van Ommeren, J. N., & Pilegaard, N. (2010). Wages and commuting: Quasi-natural experiments' evidence from firms that relocate. *Tinbergen Institute Discussion Papers*, (10-093/3).

Nickell, S., Jones, P., & Quintini, G. (2002). A picture of job insecurity facing British men. *Economic Journal*, 112(476), 1–27.

Osland, L. & Thorsen, I. (2008). Effects on housing prices of urban attraction and labor-market accessibility. *Environment and Planning A*, 40(10), 2490–2509.

Overman, H. G. & Puga, D. (2010). Labor pooling as a source of agglomeration: An empirical investigation. In *Agglomeration Economics*, NBER Chapters (pp. 133–150). National Bureau of Economic Research, Inc.

Patacchini, E. & Zenou, Y. (2005). Spatial mismatch, transport mode and search decisions in england. *Journal of Urban Economics*, 58(1), 62–90.

Phillips, D. (2012). Getting to work: Experimental evidence on job search and transportation costs. Unpublished, Department of Economics, Hope College MI.

Puga, D. (2002). European regional policies in light of recent location theories. *Journal of Economic Geography*, 2(4), 373–406.

Raphael, S. (1998). The spatial mismatch hypothesis and black youth joblessness: Evidence from the san francisco bay area. *Journal of Urban Economics*, 43(1), 79–111.

Redding, S. J. & Venables, A. J. (2004). Economic geography and international inequality. *Journal of International Economics*, 62(1), 53–82.

Rogers, C. L. (1997). Job search and unemployment duration: Implications for the spatial mismatch hypothesis. *Journal of Urban Economics*, 42(1), 109–132.

Rosenthal, S. S. & Strange, W. C. (2001). The determinants of agglomeration. *Journal of Urban Economics*, 50(2), 191–229.

Rosenthal, S. S. & Strange, W. C. (2003). Geography, industrial organization, and agglomeration. *The Review of Economics and Statistics*, 85(2), 377–393.

Rosenthal, S. S. & Strange, W. C. (2004). Evidence on the nature and sources of agglomeration economies. In J. V. Henderson & J. F. Thisse (Eds.), *Handbook of Regional and Urban Economics*, volume 4 of *Handbook of Regional and Urban Economics* chapter 49, (pp. 2119–2171). Elsevier.

Rosenthal, S. S. & Strange, W. C. (2008). Agglomeration and hours worked. *The Review of Economics and Statistics*, 90(1), 105–118.

Sanchis-Guarner, R. (2012). *Essays on Urban and Spatial Economics*. PhD Theses. The London School of Economics and Political Science.

Straub, S. (2011). Infrastructure and development: A critical appraisal of the macro-level literature. *The Journal of Development Studies*, 47(5), 683–708.

van Ommeren, J. N. & Gutiérrez-i-Puigarnau, E. (2011). Are workers with a long commute less productive? An empirical analysis of absenteeism. *Regional Science and Urban Economics*, 41(1), 1–8.

Venables, A. J. (2007). Evaluating urban transport improvements: Cost-benefit analysis in the presence of agglomeration and income taxation. *Journal of Transport Economics and Policy*, 41(2), 173–188.

Vickerman, R. (2002). *Transport and Economic Development*. Transport and economic development round table 119, European Conference of Ministers of Transport, OECD.

Vickerman, R., Spiekermann, K., & Wegener, M. (1999). Accessibility and economic development in europe. *Regional Studies*, 33(1), 1–15.

Warunsiri, S. & McNown, R. (2010). The returns to education in Thailand: A pseudo-panel approach. *World Development*, 38(11), 1616–1625.

Zenou, Y. (2009). *Urban Labor Economics*. Cambridge University Press.

Appendix

A.1. Further details on data construction

The data described in this section follows closely Gibbons et al. (2012). When necessary, the data construction was adapted to the setup of the empirical analysis if this paper.

A.1.1. Economic size of the locations

The accessibility indices A_{rt} and $\hat{A}_{rt}$ of section 4.1. are calculated combining the economic size of locations and ward-to-ward travel times in years 2002–2008 using the expressions (2) and (3). I use total employment in the ward as measure of economic size of the locations. The employment figures at the ward level were obtained using data from the Business Structure Database (BSD). The BSD is a yearly snapshot of the Inter-Departmental Business Register (IDBR), which is accessible to researchers through the Secure Data Service (SDS) delivered by the UK Data Archive (UKDA). I use data from 2001 to 2008.

This dataset is maintained by the Office of National Statistics (ONS) and contains a yearly updated register of the universe of businesses in the United Kingdom. It is drawn from administrative registers. It covers about 98% of business activity (by turnover) in Great Britain.

The smallest unit of observation is the establishment or plant ("local unit"), but there is also information on the firm to which the plant belongs ("reporting unit") and the enterprise and enterprise group of the firm. The dataset provides detailed information on the location (postcode), the sector of production (up to 5 digits) and employment of the plant. This level of detail allows us to calculate employment at any geographical level aggregating up from postcodes.

Alternatively indices (2) and (3) can be defined using different measures of economic size. Specifically, I use number of residential addresses in the ward [34], obtained from the National Statistics Postcode Directory (NSPD) and the number of establishments at the ward level (local unit counts), calculated using the BSD. I use these alternative indices in the robustness checks carried out below in Section A.2..

A.1.2. The calculation of travel times

The second component in the accessibility index is then an origin-destination (O–D) matrix containing the costs c_{rjt} (journey time) between each origin and destination (the ward centroid). This matrix is required for all the years of the sample. To calculate travel times I construct GIS-networks for every year between 2002 and 2008 (dated at the beginning of the year). I do this by combining the dataset on road schemes described in Section A.1.3. and two GIS-networks (for years 2003 and 2008) for Great Britain provided by the Department of Transport (DfT).

The 2008 GIS-network contains all the major road links existing at the beginning of 2008. It includes all major roads that, according to the DfT, cover roughly 65% of vehicle kilometres. The network includes information on several characteristics of the road links: the count point code (CP) of each road section (which helps is to identify the links and refers to the point where the traffic

[34] We use address counts due to the lack of available data on population at the ward level on an annual basis.

is counted), the grid reference for the traffic count point, a unique reference for the local and national transport authorities which manage the link, the road to which the link belongs to (number and type), the maximum permitted speed, the total length of network road link in kilometres and the traffic total flows [35]. Total flow is defined as the "Annual Average Daily Flow" (AADF) and it is measured in terms of number of vehicles. It corresponds to the average over a full year of the number of vehicles passing a point in the road network each day.

I use road construction as the source of variation in travel times over time. I geo-locate all the road links belonging to each of the 23 schemes listed in Table A.2 and I match them to the 2008 road network based on their CP code. Starting from the 2008 network, in every year I remove the new links opened in that year in order to reconstruct the network as it was in years prior to 2008. I construct a network at the beginning of each year of the period 2002–2008. The road network consists of roughly 17,000 road links annually. These networks contain the links existing every year. In order to construct A_{rt} and $\hat{A}_{rt}$ I need to calculate the cost of crossing those links c_{rjt}, which I define based on travel times.

In order to calculate travel times, I use data on traffic speeds from the 2003 generalised primary road GIS-network provided by DfT. Traffic speeds are modeled from traffic flow census data using the Road Capacity and Costs Model (FORGE - Fitting On of Regional Growth and Elasticities) component of the National Transport Model (NTM). FORGE is the highway supply module of NTM [36]. The National Transport Model provides "a means of comparing the national consequences of alternative national transport policies or widely-applied local transport policies, against a range of background scenarios which take into account the major factors affecting future patterns of travel". It is used to produce forecasts on traffic flows in order to design transport policies [37]. The Road Capacity and Costs Model is one of the three sub-models included in the NTM and it corresponds to the highway supply module.

The Road Capacity and Costs Model (FORGE) is used to show the impact of road schemes and other road-based policies [38]. As explained in the documentation [39]: "The inputs to the Road Capacity and Costs Model are car traffic growth (based on growth in car driver trips) and growth in vehicle-miles from other vehicle types. This traffic growth is applied to a database of base year traffic levels to give future "demand" traffic flows. These are compared to the capacity on each link, and resulting traffic speeds are calculated from speed/flow relationships (which link traffic volumes, road capacity and speed) for each of 19 time periods through a typical week". One of the outputs of FORGE is therefore vehicle speeds by road type, and this is what I use in the calculation of travel times between wards.

I use journey times, obtained from FORGE, in the non-busy direction averaged over all time periods between Monday-Friday 08:00 and 18:00. I focus on non-busy travel directions because the busy travel directions are, in principle, more sensitive to changes in congestion induced by new travel links. In practice, there are only minor differences between the modeled journey times in the busy and non-busy directions, or between averages over the working day and full 24-hour periods (see Gibbons et al., 2012, for more details). Due to data limitations, I use journey times in 2003 for the

[35] More information in `http://www.dft.gov.uk/matrix/estimates.aspx`.
[36] See `http://www.rudi.net/files/FORGE.pdf` for more information.
[37] See `http://www2.dft.gov.uk/pgr/economics/ntm/` for more information.
[38] See `http://www.rudi.net/files/FORGE.pdf` for more information.
[39] See `http://www2.dft.gov.uk/pgr/economics/ntm/pdfntmoverview.pdf` for more information

Year	Number of observations	Mean	Standard deviation	1st percentile	10th percentile	90th percentile	99th percentile
2002	9,170,886	49.745	17.515	6.681	24.104	70.872	74.593
2003	9,170,886	49.735	17.512	6.679	24.099	70.859	74.591
2004	9,170,886	49.613	17.476	6.674	24.049	70.724	74.573
2005	9,170,886	49.597	17.471	6.672	24.044	70.704	74.571
2006	9,170,886	49.591	17.469	6.672	24.043	70.696	74.569
2007	9,170,886	49.575	17.463	6.672	24.037	70.675	74.566
2008	9,170,886	49.554	17.457	6.669	24.025	70.645	74.560
Growth 2002–2008		-0.38%	-0.33%	-0.18%	-0.33%	-0.32%	-0.04%

Notes: Times in minutes. Min=0 and Max=75 mins. Times measured at the beginning of the year. **Source:** Department of Transport and own authors' calculation.

Table A.1: Summary statistics O–D matrices of travel times – 2002–2008

whole period 2002–2008. These speeds are based on traffic flows for 2003, and I keep these constant for all years in the sample.

For links opened after 2003 I use estimated journey times from a regression model using a dataset of over 17,000 links in the 2003 network. I regressed link speeds from the 2003 FORGE network on speed limit dummies, traffic flows, traffic flows squared, road category dummies (six categories) and local authority dummies [40]. The regression predicts speeds from the FORGE reasonably well (R-squared = 0.76). I then used the results to predict travel times for links opened after 2003 for which no FORGE speed is available. To obtain these predictions I use the link characteristics and traffic flows.

I are interested in using the variation in accessibility which stems from road construction, which is how I capture changes transport investments. Therefore in the calculation of predicted traffic speeds I need to avoid using endogenous changes in traffic flows that could be induced by road construction. For the links existing in 2003 I use the FORGE modeled speeds provided by the Department of Transport, which are calculated using 2003 traffic flows. For the links after 2003 I use 2008 flows. These are available for all the links constructed after 2003 because they were provided in the 2008 GIS-network. Alternatively I could use annual traffic flows on the year of opening on the link [41]. However, in order to avoid endogenous traffic flow changes, I prefer to use traffic flows for the same year in all the new links, and the only year in which this is possible is 2008. For some of the links, the prediction exceeded travel time implied by the speed limit. I replaced predicted speed with the speed limit for these links.

Some of the road schemes in the Highways Agency data are bypasses around villages and small towns (Faster Routes in Table A.2). Typically, before the bypass was opened there was a primary road through the village or town but after the introduction of the bypass the old road was downgraded. The downgrading of the old road implies that it is not present in the 2008 primary road network. Hence, using the method of deleting links based on their opening years would create an artificial break in the primary road network, when it comes to bypasses. Therefore, I keep the bypasses in the network in the pre-opening years and assume that travel time on the bypass before opening year was twice the post opening travel time. Scheme evaluation reports available to us support the assumption of significantly longer travel time through the village/town before the bypass is opened.

After constructing the generalised traffic networks of each year, I use the network analysis al-

[40] The results of this regression are available on request.
[41] These are available in `http://www.dft.gov.uk/matrix/Search.aspx`.

gorithms in ESRI ArcGIS® to compute least-cost (minimum journey time) routes between each origin ward j and destination ward k in years 2002–2008. When computing the O-D matrix I apply a limit of 75 minute drive time. This limit facilitates O-D matrix computation but does not affect the value of accessibility index because wards beyond 75 minutes would have negligible weights in the calculation of A_{rt}.

Table A.1 contains summary statistics for the travel times between wards for year 2002–2008. The number of O–D crossings is over 9,000,000 in every year. Over the period, the average travel time between two wards decreased 0.38%.

It should be noted that the network is highly generalised. Journeys via the minor road network are not modeled. Forbidden turns and one way systems are not modeled. All link intersections are treated as junctions. The changes in accessibility must therefore be regarded as approximate. This might have implications in the estimation of the effect of accessibility on economic outcomes if it induces measurement error in the accessibility indices calculation. Hence, the estimates could be too small due to attenuation bias. In this case, I should interpret any effect as a lower bound of the "real" effect.

A.1.3. The road schemes

Opening year	Type	Road	Scheme	Length in kms
2002	New Route	A27	A27 Polegate Bypass	3.2
2002	Faster Route	A43	A43 Silverstone Bypass	14.2
2002	Faster Route	A6	A6 Clapham Bypass	14.6
2002	Faster Route	A66	A66 Stainburn and Great Clifton Bypass	4.1
2003	Faster Route	A41	A41 Aston Clinton Bypass	7.3
2003	Faster Route	A5	A5 Nesscliffe Bypass	21.5
2003	Faster Route	A500	A500 Basford, Hough, Shavington Bypass	7.7
2003	Faster Route	A6	A6 Alvaston Improvement	4.7
2003	Faster Route	A6	A6 Great Glen Bypass	6.8
2003	Faster Route	A6	A6 Rothwell to Desborough Bypass	8.4
2003	New Route	A6	A6 Rushden and Higham Ferrers Bypass	5.4
2003	Faster Route	A650	A650 Bingley Relief Road	4.4
2003	New Route	M6(T)	M6 Toll. Birmingham Northern Relief Road	29.7
2004	Faster Route	A10	A10 Wadesmill to Colliers End Bypass	7.0
2004	New Route	A63	A63 Selby Bypass	9.5
2005	New Route	A1(M)	A1(M) Wetherby to Walshford	8.1
2005	Faster Route	A21	A21 Lamberhurst Bypass	2.4
2005	Faster Route	A47	A47 Thorney Bypass	10.7
2005	New Route	M77	M77 Replaces A77 from Glasgow Road	18.2
2006	New Route	A1(M)	A1(M) Ferrybridge to Hook Moor	19.2
2006	Faster Route	A421	A421 Great Barford Bypass	7.6
2007	Faster Route	A2	A2 / A282 Dartford Improvement	4.2
2007	Faster Route	A66	A66 Temple Sowerby Bypass and Improvements at Winderwath	26.2
			TOTAL	**245.1**

Source: Own authors calculations using information from the Department for Transport, the Highways Agency, the Motorway Archive, Transport Scotland and Wikipedia.

Table A.2: Road schemes – 2002–2007

In order to calculate travel times between location in every point of time, I need to know which road links existed every year. I collected information on completed road projects for the British major roads network by combining information provided by the Department of Transport and other data sources[42]. I collected data on around 75 projects which were completed between 2002 and 2007 [43].

The nature of these projects is diverse. They cover construction of new junctions, dualling, widening, upgrades and construction of new roads. I focus on construction of new roads and, as our analysis covers the period 2002–2008, I retain 23 road schemes, which are listed in Table A.2. The length of these schemes is also provided in Table A.2. As clearly stated in the policy documentation (see for example Department of Transport, 1998b; Highways Agency, 2009, for more information), road interventions undertaken by the local and national authorities [44] aim to improve traffic flows and road security, and indirectly affect the environment and the economy by reducing traffic congestion.

	Length	**Percentage**
Major Roads Network in 2008	50,093.7	
All new links 2002–2007 (both types)	245.1 kms	0.49%
of which: New Routes	93.35 kms	38.1%
of which: Faster Routes	151.78 kms	61.9%

Notes: The right column reports the percentage of total new links over the whole network (2^{nd} row) and the percentage of the new links by type over total improvement length (3^{rd} and 4^{th} rows). **Source:** Own authors calculations using information from the Department for Transport, the Highways Agency, the Motorway Archive, Transport Scotland and Wikipedia.

Table A.3: Length of schemes – 2002–2007

I define two types of projects. Type 2 (Faster Routes) correspond to roads for which there was an alternative route before, but the road was a minor road (not existing in the major road network) and an upgrade (which involves improvement and the construction of new lanes) was carried out so the road becomes part of the major road network. They mainly correspond to bypasses which relieve traffic congestion from villages and usually flow in parallel to an existing alternative minor road. Type 1 (New Routes) corresponds to "genuinely" new roads, i.e. roads for which I do not have an alternative minor road flowing in parallel. In practice, as detailed in Section A.1.2., I defined these two types of improvements in order to be able to calculate travel times given the characteristics of the road network data that I use. They are not substantially different, both involve road construction, except for the treatment I give them in them in the calculation of the travel O–D matrices.

Figure 1(a) displays the location of these projects and the major road network at the end of the period of analysis (2008). Projects are scattered all over Britain. I focus on new construction because these improvements are the ones I can expect to have a substantial effect on travel times between wards. Table A.3 summarises the length of the links by type and displays the percentage of kilometres they represent with respect to the total network length. Total new links represent 0.49% of the

[42] Mainly The Highways Agency, the Motorway Archive, Transport Scotland and Wikipedia.

[43] More details in Gibbons et al. (2012). For details on the policy context see Sanchis-Guarner (2012).

[44] Due to rising concerns about the increase in traffic flows and the limited capacity of existing networks, in 1997 the Labour Government carried out a reform of the management of major roads. Major roads comprise trunk roads, motorways and principal roads (A roads). The aim of this reform was to "radically change transport policy" (Department of Transport, 1998a). This reform involved the transfer of parts of the English Trunk Network to local authorities, while the Department of Transport kept control of the most strategic roads (those connecting major population centers, ports and airports, key cross-border links and the Trans-European Road Network). The management of these roads was transferred to the Highways Agency in England (which is part of the Department of Transport), to Transport Scotland and to the Welsh Assembly Government.

network length, and the majority of them (61.9%) correspond to type 2 links, i.e. faster routes.

A.1.4. Accessibility changes arising from road construction

Accessibility indices A_{rt} and $\hat{A}_{rt}$ can be applied to study the economic effects arising from changing accessibility by road when the costs c_{rjt} in (2) and (3) are calculated using routing along the transport network. This works because road construction changes the structure of costs c_{rjt} along the transport network and the structure of costs along routes from r to potential destinations j. This in turn changes the accessibility index.

For example, consider a transport improvement that involves a journey time reduction on a road link between two nodes p and q. This scheme has a first order effect on the costs of the least-cost route between r and j if:

(a) the least-cost route between r and j passes along the link p–q in both the pre and post-improvement periods, such that the transport improvement reduces the cost of the journey along p–q and brings employment at destination j "closer" to origin r in cost terms.

(b) the least-cost route between r and j bypasses link p–q in the pre-improvement period, but switches to use the link p–q in the post-improvement period because of the reduction in costs; again this brings employment at destination j "closer" to origin r in cost terms.

There are also "second order" effects arising when:

(c) the least cost route between r and j bypasses link p–q in the pre-improvement and in the post-improvement periods. However, journeys between other origin and destination pairs have switched to using the link p–q, which reduces congestion on the alternative links in the network used by the routing between r and j; again this brings employment at destination j "closer" to origin r in cost terms.

In the empirical work below I rely only on the first order effects of type (a) and (b) arising from new transport infrastructure. I have to ignore second order effects of type (c) because, as explained in the previous section, the road transport network data does not allow us to observe changes in travel time induced by changes in congestion occurring as a result of road construction (I have no information on traffic flows observed prior to the construction).

Changes in cost of all these types imply changes in the accessibility indices (i.e. a change in effective density). The amount of change in the accessibility index at a location j depends on the likelihood that a route between r and j uses the improved link p–q, and on economic mass in j.

A.2. Robustness of the results in Section 5.3.2.

This section tests the robustness of the findings of section 5.3.. I use the model of column 7 of table 6, which is the most demanding specification. I focus on the results on basic weekly wages (which are very similar to those for total weekly wages) and on total hours worked.

Columns 1 to 3 of Table A.4 use different measures of the economic size of the wards in the definition of accessibility indices. Column 1 replicates baseline results. Column 2 uses address counts as the measure of economic size [45] and column 3 uses number of plants (local unit counts). The instruments were calculated using 2001 employment, address counts and local unit counts respectively. The results are very similar to those of section 5.3.2.. In columns 4 to 6 I compute accessibility using employment as economic mass but I use different instruments. Column 4 calculates $\hat{A}_{iht}$ using 1997 employment, and columns 5 and 6 use 2001 address counts and local unit counts as economic mass. The instruments of columns 4 and 5 are weaker than that of column 1, but again, the results are very robust.

PANEL A: **Log of basic weekly pay**						
	(1)	(2)	(3)	(4)	(5)	(6)
Log of accessibility	0.328***	0.298***	0.312***	0.325***	0.321***	0.324***
from work ward	[0.116]	[0.112]	[0.116]	[0.116]	[0.119]	[0.118]
Log of accessibility	-0.025	-0.051	-0.048	-0.028	-0.031	-0.033
from home ward	[0.092]	[0.090]	[0.093]	[0.092]	[0.093]	[0.093]
Log of travel time	0.051	0.043	0.046	0.05	0.049	0.049
between work and home	[0.033]	[0.034]	[0.033]	[0.033]	[0.033]	[0.033]
Observations	180,105	180,105	180,105	180,105	180,105	180,105
Kleibergen-Paap F-stat	1,042	1,290	3,428	1,037	814	690
Work-ward no of clusters	3,551	3,551	3,551	3,551	3,551	3,551
Home-ward no of clusters	3,824	3,824	3,824	3,824	3,824	3,824
PANEL B: **Log of total weekly hours worked**						
	(1)	(2)	(3)	(4)	(5)	(6)
Log of accessibility	0.375**	0.364**	0.386**	0.372**	0.380**	0.396**
from work ward	[0.156]	[0.162]	[0.162]	[0.154]	[0.163]	[0.163]
Log of accessibility	-0.204	-0.232	-0.243	-0.204	-0.22	-0.23
from home ward	[0.224]	[0.231]	[0.240]	[0.221]	[0.239]	[0.240]
Log of travel time	0.028	0.025	0.024	0.028	0.027	0.027
between work and home	[0.062]	[0.063]	[0.061]	[0.062]	[0.061]	[0.061]
Observations	180,392	180,392	180,392	180,392	180,392	180,392
Kleibergen-Paap F-stat	1,036	1,301	3,401	1,032	809	684
Work-ward no of clusters	3,548	3,548	3,548	3,548	3,548	3,548
Home-ward no of clusters	3,826	3,826	3,826	3,826	3,826	3,826
Accessibility	Empl	Addrct	LUsct	Empl	Empl	Empl
Instrument	Empl01	Addrct01	LUsct01	Empl97	Addrct01	LUsct01

Notes: All specification include sector-year dummies (5 categories), μ_{ihw} fixed-effects, instrument and all the controls and trends. 2-way clustering (work-ward and home-ward). * p<0.1, ** p<0.05, *** p<0.01. **Source:** ONS, DfT, CASWEB and authors own calculations.

Table A.4: Robustness: different economic sizes and different instruments – 2002–2008

Table A.5 shows results using alternative cost functions in the calculation if the accessibility measure. Columns 1 and 4 and 2 and 5 use the "inverse cost weights" function ($a\left(c_{rjt}\right) = c_{rjt}^{-\alpha}$) but with distance decays $\alpha = 0.5$ (flatter) and $\alpha = 1.5$ (steeper). Columns 3 and 6 use the "exponential weights" function ($a\left(c_{rjt}\right) = \exp\left(\alpha c_{rjt}\right)$) with distance decay $\alpha = 0.2$. The elasticities are larger than before but still positive and largely significant.

Table A.6 tests the robustness of the results to the inclusion or exclusion of some variables. Column 1 drops commuting time and in column 2 introduces travel time in levels instead of in logs. The results remain very similar, the only different is that travel time is now significant and positive in column 2 contrary to insignificant in column 7 of Table 4. Column 3 adds a dummy which indicates

[45] Ideally I would have used population, but data on population at the ward level on a yearly basis is not available.

	PANEL A:Log of basic weekly pay			PANEL B:Log of total weekly hours worked		
	(1)	(2)	(3)	(4)	(5)	(6)
Log of accessibility	0.723***	0.176**	1.843***	0.874**	0.198**	2.202**
from work ward	[0.260]	[0.071]	[0.683]	[0.355]	[0.085]	[0.867]
Log of accessibility	-0.044	-0.013	-0.003	-0.433	-0.114	-0.936
from home ward	[0.197]	[0.052]	[0.465]	[0.470]	[0.138]	[1.123]
Log of travel time	0.052	0.049	0.061*	0.032	0.027	0.039
between work and home	[0.032]	[0.033]	[0.032]	[0.062]	[0.063]	[0.062]
Observations	180,105	180,105	180,105	180,392	180,392	180,392
Kleibergen-Paap F-stat	366	1914	58	362	1905	57
Work-ward no of clusters	3,551	3,551	3,551	3,548	3,548	3,548
Home-ward no of clusters	3,824	3,824	3,824	3,826	3,826	3,826
Cost function	Inverse	Inverse	Exponent	Inverse	Inverse	Exponent
Decay	0.5	1.5	0.2	0.5	1.5	0.2

Notes: All specification include sector-year dummies (5 categories), μ_{ihw} fixed-effects, instrument and all the controls and trends. 2-way clustering (work-ward and home-ward). * p<0.1, ** p<0.05, *** p<0.01. **Source:** ONS, DfT, CASWEB and authors own calculations.

Table A.5: Robustness: different costs functions and distance decays – 2002–2008

if the individual is changing jobs (that I identify using the enterprise reference number) within the ward of work. The estimated elasticities remain virtually unchanged. This is also the case in column 4 when I introduce a dummy if the worker changes home (identified using the postcode) within the home-ward. Finally, column 5 drops those individuals who live and work in the same ward. Once more, I obtain very similar coefficients to those of section 5.3.2..

In Table A.7 tests the robustness of the results to the inclusion or exclusion of some specific individual groups. Column 1 excludes all individuals whose wage setting is subject to collective agreement. The number of observations decreases substantially and the estimates become imprecise, and only weekly significant for total hours worked. Column 2 includes workers from the public sector, which are excluded in the main results. The results are very similar to before but the elasticity of accessibility from work is slightly larger. In column 3, when I obtain the results only using individuals working in the public sector, the coefficient of accessibility from work on wages and hours are both significant (but weaker) and positive, and the size is bigger than in the baseline results.

In columns 4 to 6 I test the robustness of the results to the exclusion of London. Column 4 excludes individuals working or living in London (inner and outer), in column 5 only those living in London and in column 6 only working in the capital. The results are very similar to the main estimates. Finally, column 7 restricts the estimation to those individuals which do not change home-ward location while observed in the panel, i.e. those for which μ_i is equal to μ_{ihw}. The coefficient of the effect of accessibility from work on wages is very similar to the baseline, but the effect on hours is larger. These individuals might not be moving work or home for some unobserved reasons, for example they are home owners or social renters, they have restricted mobility due to family reasons, etc. Therefore, if they cannot relocate to take advantage of the changes in accessibility, they might adjust by working longer hours or switching to full time employment. This could explain the larger estimated effect on hours worked.

PANEL A: **Log of basic weekly pay**					
	(1)	(2)	(3)	(4)	(5)
Log of accessibility	0.286***	0.316***	0.327***	0.328***	0.326***
from work ward	[0.110]	[0.116]	[0.116]	[0.116]	[0.119]
Log of accessibility	-0.04	-0.011	-0.024	-0.026	-0.019
from home ward	[0.097]	[0.094]	[0.092]	[0.092]	[0.093]
Log of travel time			0.05	0.051	0.076**
between work and home			[0.033]	[0.032]	[0.033]
Travel time		0.236**			
between work and home		[0.102]			
Individual changes			0.002		
job within work ward			[0.003]		
Individual changes				0.006*	
house within home ward				[0.003]	
Observations	185,818	180,105	180,105	180,105	158,609
Kleibergen-Paap F-stat	1,104	1,076	1,042	1,042	787
Work-ward no of clusters	3,556	3,551	3,551	3,551	3,353
Home-ward no of clusters	3,830	3,824	3,824	3,824	3,785
PANEL B: **Log of total weekly hours worked**					
	(1)	(2)	(3)	(4)	(5)
Log of accessibility	0.355**	0.373**	0.375**	0.375**	0.350**
from work ward	[0.140]	[0.153]	[0.156]	[0.156]	[0.147]
Log of accessibility	-0.209	-0.167	-0.204	-0.203	-0.21
from home ward	[0.229]	[0.228]	[0.224]	[0.224]	[0.233]
Log of travel time			0.028	0.028	0.072
between work and home			[0.062]	[0.062]	[0.064]
Travel time		0.284			
between work and home		[0.230]			
Individual changes			-0.001		
job within work-ward			[0.005]		
Individual changes				-0.001	
house within home-ward				[0.006]	
Observations	186,209	180,392	180,392	180,392	158,998
Kleibergen-Paap F-stat	1,099	1,070	1,036	1,035	778
Work-ward no of clusters	3,553	3,548	3,548	3,548	3,351
Home-ward no of clusters	3,834	3,826	3,826	3,826	3,790
Specification	No ltrvt	Level trvt	Chg jobs	Chg houses	Diff HW

Notes: All specification include sector-year dummies (5 categories), μ_{ihw} fixed-effects, instrument and all the controls and trends. 2-way clustering (work-ward and home-ward). * p<0.1, ** p<0.05, *** p<0.01. **Source:** ONS, DfT, CASWEB and authors own calculations.

Table A.6: Robustness: changes in the specification – 2002–2008

PANEL A: Log of basic weekly pay							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log of accessibility	0.256	0.337***	0.501*	0.301***	0.306***	0.308***	0.341***
from work ward	[0.211]	[0.110]	[0.297]	[0.108]	[0.108]	[0.108]	[0.117]
Log of accessibility	0.034	-0.051	-0.135	-0.059	-0.06	-0.058	0.068
from home ward	[0.143]	[0.083]	[0.150]	[0.082]	[0.082]	[0.081]	[0.105]
Log of travel time	0.111*	0.022	-0.042	0.022	0.022	0.023	0.035
between work and home	[0.058]	[0.021]	[0.028]	[0.021]	[0.021]	[0.021]	[0.031]
Observations	96,119	254,294	73,488	205,991	208,091	213,822	114,919
Kleibergen-Paap F-stat	1,132	952	352	1,259	1,253	1,271	987
Work-ward no of clusters	3,315	3,710	2,703	3,192	3,194	3,549	3,350
Home-ward no of clusters	3,745	3,847	3,511	3,333	3,629	3,337	3,724
PANEL B: Log of total weekly hours worked							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log of accessibility	0.261*	0.407***	0.645**	0.387***	0.389***	0.385***	0.698***
from work ward	[0.142]	[0.136]	[0.254]	[0.135]	[0.135]	[0.136]	[0.218]
Log of accessibility	0.064	-0.17	-0.029	-0.177	-0.177	-0.181	-0.237
from home ward	[0.119]	[0.179]	[0.138]	[0.177]	[0.177]	[0.176]	[0.282]
Log of travel time	0.019	0.008	-0.038	0.005	0.005	0.005	0.033
between work and home	[0.049]	[0.046]	[0.041]	[0.046]	[0.046]	[0.046]	[0.074]
Observations	96,485	254,619	73,537	205,655	207,768	213,739	115,172
Kleibergen-Paap F-stat	1,123	949	348	1,259	1,254	1,272	990
Work-ward no of clusters	3,310	3,707	2,696	3,190	3,192	3,546	3,346
Home-ward no of clusters	3,748	3,848	3,510	3,332	3,627	3,336	3,728
Specification	No coll	Priv-Publ	Public	No London	No hLond	No wLond	No move

Notes: All specification include sector-year dummies (5 categories), μ_{ihw} fixed-effects, instrument and all the controls and trends. 2-way clustering (work-ward and home-ward). * p<0.1, ** p<0.05, *** p<0.01. **Source:** ONS, DfT, CASWEB and authors own calculations.

Table A.7: Robustness: different groups – 2002–2008

A.3. Additional Tables and Figures

Commuting time	Travel mode				
	Rail/tube	Motor	Walk/Cycle	*Commuters*	*% times*
15 mins or less	0.25%	78.13%	21.63%	25,298	49.27%
16-30 mins	2.12%	87.65%	10.23%	15,666	30.51%
31-45 mins	8.07%	87.85%	4.08%	5,489	10.69%
45-60 mins	21.63%	76.71%	1.66%	3,315	6.46%
61-90 mins	33.74%	65.78%	0.48%	1,242	2.42%
91-120 mins	30.21%	69.79%	0.00%	288	0.56%
120 mins or more	31.11%	68.89%	0.00%	45	0.09%
% modes	4.04%	81.63%	14.33%	100%	51,343

Notes: Percentages with respect to commuters. **Source:** British Household Panel Survey, years 2002–2008.

Table A.8: Commuting times by travel mode – 2002–2008

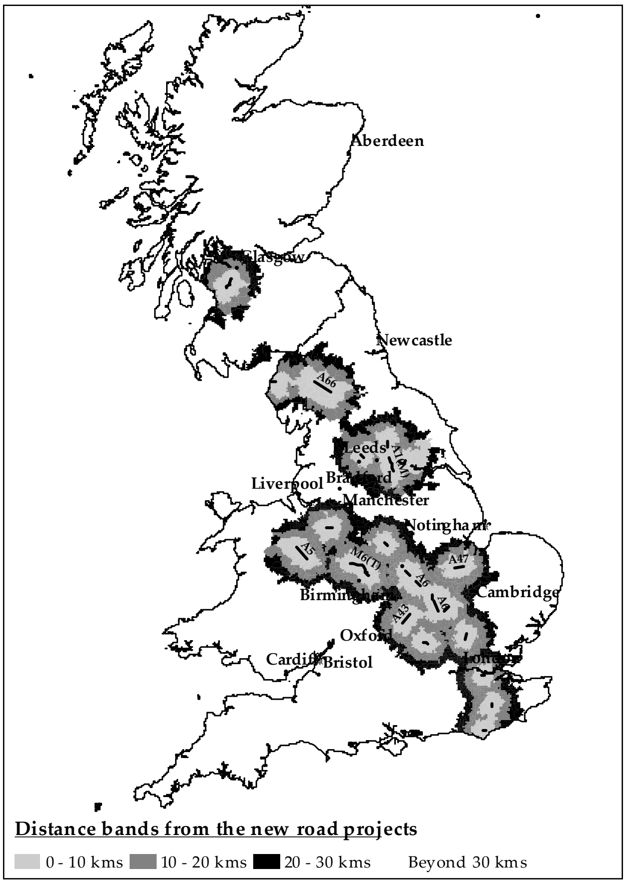


Figure A.1: Wards within 10-20-30 kms distance bands around the project sites - 2002–2008

	Wards	Mean	Standard deviation	90^{th} percentile	Maximum value
PANEL A: *2001 employment and time-varying travel times*					
All	73,780	14.710	1.170	19.320	19.320
10 kms	6,034	15.050	0.720	17.620	17.620
20 kms	17,171	15.190	0.710	18.170	18.170
30 kms	27,216	15.260	0.790	18.190	18.190
PANEL B: *Time-varying employment and 2001 travel times*					
All	73,780	14.770	1.170	16.060	19.500
10 kms	6,034	15.090	0.730	15.810	17.800
20 kms	17,171	15.240	0.720	16.000	18.300
30 kms	27,216	15.310	0.790	16.220	18.500
PANEL C: *Time-varying employment and time-varying travel times*					
All	73,780	14.760	1.170	16.060	19.500
10 kms	6,034	15.100	0.730	15.810	17.800
20 kms	17,171	15.240	0.720	16.000	18.300
30 kms	27,216	15.310	0.790	16.220	18.590

Source: Department of Transport, BSD and own authors' calculations.

Table A.9: Summary statistics annual log of accessibility – 2002–2008

YEAR	Males	Female	**Total**
2002	17,756	13,457	31,213
2003	19,451	15,355	34,806
2004	19,248	15,687	34,935
2005	20,100	17,012	37,112
2006	20,111	17,178	37,289
2007	16,846	14,945	31,791
2008	16,732	14,658	31,390
Total	130,244	108,292	238,536

Source: ONS. Observations within 30 kms of home and work wards.

Table A.10: Number of observations by year and gender in ASHE – 2002–2008

YEAR	MANUFACT.	CONSTRUCT.	CONSUMER SERVICES	PRODUCER SERVICES	OTHER	**ALL SECTORS**
2002	7,638	1,615	7,331	7,380	7,249	31,213
2003	7,871	1,747	8,579	8,582	8,027	34,806
2004	7,336	1,706	9,375	8,371	8,147	34,935
2005	7,445	1,692	10,300	8,911	8,764	37,112
2006	6,765	1,921	10,249	9,267	9,087	37,289
2007	4,806	1,800	8,671	9,392	7,122	31,791
2008	4,773	1,850	8,429	9,244	7,094	31,390
Total	46,634	12,331	62,934	61,147	55,490	238,536

Source: ASHE–ONS. Observations within 30 kms of home and work wards.

Table A.11: ASHE Number of observations by year and industrial sector – 2002–2008

OBSERVATIONS	**Individual home-work ward id**	**Individuals allowed to move wards**
Total	238,536	238,536
Number of individuals	114,370	79,286
Average number of T	2.09	3.01

Source: ASHE–ONS. Observations within 30 kms of home and work wards.

Table A.12: ASHE Number of individuals for each panel – 2002–2008

YEAR	QUARTER				
	1	2	3	4	**Total**
2002	27,362	25,280	22,153	19,504	94,299
2003	22,233	23,609	25,123	24,288	95,253
2004	22,071	21,283	20,465	18,568	82,387
2005	23,314	26,330	26,254	26,159	102,057
2006	26,300	26,385	26,244	26,153	105,082
2007	26,406	26,403	26,489	26,403	105,701
2008	26,453	26,049	25,520	26,062	104,084
Total	174,139	175,339	172,248	167,137	688,863

Table A.13: LFS number of observations – 2002–2008

Spatial Economics Research Centre (SERC)
London School of Economics
Houghton Street
London WC2A 2AE

Tel: 020 7852 3565
Fax: 020 7955 6848
Web: www.spatialeconomics.ac.uk

SERC is an independent research centre funded by the Economic and Social Research Council (ESRC), Department for Business Innovation and Skills (BIS) and the Welsh Assembly Government.