SERC DISCUSSION PAPER 117

New Road Infrastructure: the Effects on Firms

Stephen Gibbons (SERC, Department of Geography & Environment, London School of Economics and Political Science)
Teemu Lyytikäinen (SERC, Government Institute for Economic Research, VATT)
Henry Overman (SERC, Department of Geography & Environment, London School of Economics and Political Science)
Rosa Sanchis-Guarner (SERC, The Grantham Research Institute on Climate Change and the Environment, London School of Economics and Political Science)

September 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.

New Road Infrastructure: the Effects on Firms

Stephen Gibbons*
Teemu Lyytikäinen**
Henry Overman*
Rosa Sanchis-Guarner***

September 2012

* Spatial Economics Research Centre (SERC) and Department of Geography and Environment, London School of Economics and Political Science
** Spatial Economics Research Centre (SERC), London School of Economics and Political Science and Government Institute for Economic Research (VATT)
***Spatial Economics Research Centre and The Grantham Research Institute on Climate Change and the Environment, London School of Economics and Political Science

Acknowledgements
This work is based on data from ARD, BSD, and NSPD, produced by the Office for National Statistics (ONS) and supplied by the Secure Data Service at the UK Data Archive. 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 or the Secure Data Service at the UK Data Archive in relation to the interpretation or analysis of the data. This work uses research datasets which may not exactly reproduce National Statistics aggregates. All the outputs have been granted final clearance by the staff of the SDS–UKDA. Additionally, any interpretations or opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the Department for Transport who also contributed to the study. This research has been presented at the Urban Economic Association Meetings (Denver, November 2010), at the SERC Work in progress seminar (December 2010), at the Royal Economic Society Conference (April 2011), at the LSE Labour Economics Workshop (November 2011), at the Regional and Urban Economics Seminar, Paris School of Economics (June 2012) and at the Commissariat Général au Développement Durable, Paris (June 2012). We thank the participants for their comments. All errors and opinions are our own. The authors acknowledge ESRC for funding (Grant ES/J007382/1: Transport Investments and Spatial Economic Performance).

Teemu Lyytikäinen wishes to thank the Yrjö Jahnsson Foundation for financial support and Rosa Sanchis-Guarner acknowledges funding support from the Bank of Spain.

Abstract

This paper estimates the impact of road improvements on firm employment and productivity using plant level longitudinal data for Britain. Exposure to transport improvements is measured by changes in employment accessibility along the road network. These changes are constructed using data on employment for small geographical units, details of the main road network and of road construction schemes carried out between 1998 and 2007. We deal with the central problem of endogenous scheme placement by using changes due to new road links and exploiting the spatial detail in our data to focus on accessibility changes close to new schemes. We find substantial effects on employment and numbers of plants for small-scale geographical areas (electoral wards), but no employment response at plant level. This suggests that road construction affects firm entry and exit, but not the employment of existing firms. We also find effects on labour productivity and wages at the firm level, although these results are less robust.

JEL Classifications: D24, O18, R12

Keywords: Productivity, employment, accessibility, transport

1 Introduction

Road networks dominate transport infrastructure in most countries. In the UK in 2008, 91% of passenger transport and around 67% of goods transport was by road. For transport within the European Union, in 2009, the corresponding figures were 92% and 47% and for the US in 2007, 88% and 40-48%.[2] Clearly road transport delivers economic benefits, and transport *improvements* are frequently proposed as a strategy for economic growth, integration and local economic development (e.g. European Commission 2006, World Bank, 2008). Transport improvements decrease transportation costs, improve access to markets and labour, foster economic integration, stimulate competition, generate agglomeration economies and a number of other 'wider' economic benefits. But for economies with well-developed transport networks, little is known about the extent of the gains that result from additions to the existing network. Direct evidence of the causal effects of such improvements on economic activity using ex-post evaluation of improvements is rare.

This paper provides such evidence by investigating the causal impact of road improvements on firms using administrative data on businesses in Great Britain from 1997-2008. We capture exposure to road improvements using an index of changing employment accessibility. For a given location, this index measures the amount of employment reachable per unit of travel time along the major road network. Improvements to the network affect optimal travel times between locations, and thus directly affect employment accessibility. To construct this index, we use a new dataset of road construction schemes carried out in Great Britain between 1998 and 2007. We combine this with road network data and use network

[2] Transport Statistics Great Britain 2010, EU Transport in Figures Statistical Pocketbook 2011, US Bureau of Transportation Statistics National Transport Statistics 2012. All figures based on passenger and tonne km

analysis tools to calculate year-by-year minimum travel times between electoral wards. This travel time information, combined with data on plant location and employment, allows us to calculate year-by-year employment accessibility.

The accessibility index offers advantages over alternative simple indicators such as distance to new roads, because it varies continuously over space and changes over time in ways that depend on a firm's site in relation to both the new *and* existing parts of the transport network. Therefore changes in the index are plausibly unrelated to location-specific characteristics that jointly influence new road placement and firm productivity. This allows us to use a panel data fixed-effects design that addresses concerns that road placement may be correlated with unobserved, time invariant firm or site characteristics.

Exploiting the geographical detail in our data – our geographical units of observation are circa 10,300 electoral wards [3] - we then identify effects using only areas that are close to road improvements. Because the transport schemes considered in this paper are improvements to the major highway system intended to improve performance on the long-distance network they were not specifically targeted at local economic development. Therefore, focussing on variation in the accessibility change within 10-30 km buffer zones close to the new roads mitigates potential biases caused by endogenous scheme planning decisions since the variation within these zones is incidental to the main aims of the projects. That is, this allows us to address endogeneity problems arising if time varying unobserved location specific factors affect scheme placement. Finally, we instrument for the local accessibility change holding employment constant and using changes stemming *only* from road construction.

[3] The average area of British wards is 21 square kilometres and the average population estimate in 2001 (for England and Wales) is around 6,000 people.

This further controls for the possibility that time varying unobserved factors may drive both employment change and scheme location. We provide more detail below. We find strong evidence road improvements affect local area level employment and number of businesses. However, we find no employment effects at plant level, implying that the local employment changes come about through firm entry and exit. Conversely, we find evidence of positive impacts on labour productivity, value-added and on wages at plant, but not area level. The rest of the paper is structured as follows. Section 2 reviews the related theoretical and empirical literature. Section 3 presents the empirical methodology and explains the construction of the accessibility, productivity and employment measures. Section 4 describes the data used, and Section 5 discusses the empirical results. Finally, in Section 6 we conclude.

2 Theoretical background and existing evidence

Theoretically, reduced transport costs and improved connectivity offer various direct benefits to firms. [4] Benefits to logistics, business travel and internal organisation may improve productivity. Improvements might affect input choices, if transportation services are a factor of production (Holl, 2006). Input mix may also change if relative prices of other inputs are affected by falling transport costs. For example, wages could rise if productivity effects are capitalised into wages, or could fall if they are set along the supply curve as a function of commuting costs (Gibbons and Machin, 2006). Land prices and commercial rents could also change in response to changed location-specific benefits, in ways that depend on the elasticity of supply of commercial space. These changes in wages and rents then change

[4] Our theoretical discussion draws mainly on Gibbons and Overman (2009) who provide an extensive analysis on the potential productivity and scale effects of transport infrastructure.

the use of labour and other production inputs. There may be additional scale effects if cost reductions feed through into lower output prices and higher demand (for example by increasing market area, as suggested by Lahr et al., 2005). Again these benefits may be (partly) offset by capitalisation in commercial rents. These effects combine to determine changes in employment and observed labour productivity.

In addition to these direct effects the literature considers a number of 'wider economic benefits' of transport improvements that involve total factor productivity effects arising from agglomeration economies (Graham, 2007). These agglomeration externalities have origins in sharing of resources, matching of workers to firms, and learning by information exchange (Duranton and Puga, 2004). Although usually associated with spatial concentration (e.g. cities or industrial clusters), these affects can just as well be attributed to a contraction in travel times between firms and workers (sometimes referred to as a change in effective density). Agglomeration benefits are traditionally assumed to act like a production function shifter increasing the amount produced with given inputs (Gibbons and Overman, 2009).

Transport improvements can also influence the spatial distribution of firms through selection and sorting effects (Baldwin and Okubo, 2006). Better transport may encourage start-ups and survival by lowering costs, or allowing firms to take advantage of increasing returns to scale or agglomeration economies. Conversely, improvements can force the exit of low-productivity firms previously protected from competition by transport cost barriers, leading to long run gains in aggregate productivity (Melitz, 2003) and benefits for consumers. Although some authors have explicitly included the role of transportation into spatial economics analysis (for example Combes and Lafoucarde, 2001; Puga, 2002; Behrens et al, 2004; Venables, 2007), the theoretical link between transportation infrastructure and spatial economic outcomes remains uncertain.

Theory alone provides little definitive guidance on whether to measure the effects on firms, through prices, output, or inputs, or what response to expect on any of these dimensions, particularly when production is characterised by increasing returns to scale and imperfectly competitive markets. The theoretical predictions on the net effects of transportation improvements on area level outcomes are similarly varied. Traditional appraisal methods used for assessing the ex-ante benefits of improvements have set many of these issues to one side, by assuming a world of perfect competition in which all the economic benefits of transport improvements are manifest in travel time savings and induced demand (Small, 2007). However, more recent work, both academic (Gibbons and Machin 2006, Gibbons and Overman 2009, Venables 2007) and policy-related (Eddington 2006), argues that this may not be a complete picture.In short, given the unclear net theoretical predictions, the size and direction of the effects of transport policy on economic outcomes is mainly an empirical question (Gibbons and Machin, 2006).

Unfortunately, the number of empirical studies that have tried to quantitatively assess these effects is still limited. Most of the empirical evidence of the effects of transport and infrastructure investment on economic outcomes has been provided at the macro-level (for a review see Straub, 2011). These studies generally estimate a nationally or regionally aggregated Cobb-Douglas production function where infrastructure expenditure or roads are treated as a factor of production (Garcia-Mila et al., 1996). Unfortunately the literature struggles to address concerns about the endogeneity of transport investment and generally ignores network issues. Results, for a variety of outcomes are generally mixed. [5]

Some recent papers have used careful identification strategies to estimate the causal effect of roads on a number of outcomes in the US: urban growth (Duranton and Turner, 2012), road traffic (Duranton and Turner, 2011), trade patterns (Duranton et al, 2012), sub-urbanisation (Baum-Snow, 2007), commuting patterns (Baum-Snow, 2010) or demand for skills (Michaels, 2008). These papers usually capture the effect of transport using connectivity to the network or some measure of the spatial density of the network. Identification comes from using historical networks or strategic plans as instruments, assuming that the original plans are exogenous to current economic shocks (e.g. Duranton and Turner, 2011) or that some places are incidental beneficiaries from strategic highway plans (Michaels, 2008). Other papers have used similar ideas in the EU, e.g. railways in Berlin (Ahlfeldt and Wendland, 2011) and in developing countries, e.g. highways in China (Faber, 2012) and railroads in colonial India (Donaldson, 2010).

Few papers study the effect of increased accessibility on firm outcomes, with most focus to-date on the positive effects on firm relocation (Coughlin and Segev, 2000; Holl, 2004a and 2004c) or birth (Holl, 2004b, Melo et al, 2010). Li and Li (2010) use the construction of the Chinese highway system to suggest that improvements reduce inventories (which should improve productivity). Related research looks at the effect of agglomeration economies and market-access on firm productivity (for example Ciccone and Hall, 1996; Henderson, 2003;

[5] Outcomes considered include aggregate productivity (Aschauer, 1989; Holtz-Eakin, 1994; Fernald, 1999), earnings (Chandra and Thompson, 2000) and employment (Jiwattanakulpaisarn et al., 2009). Some papers have tried to estimate spillover effects on neighbouring regions (Boarnet, 1998; Moreno and Lopez-Bazo, 2007).

Graham 2007, Martin et al., 2011, Holl 2011). None of these studies successfully exploits changes in the transport network to identify agglomeration effects, or the effects of transport more generally. For example, Holl (2011) uses a small panel of firms during a period of intense road construction in Spain to find positive effects of market access on productivity, but relies on lags of endogenous variables as instruments (GMM) rather than using the transport changes in order to mitigate endogeneity problems. Our paper is, to our knowledge, the first to fully exploit localised changes in road network accessibility to identify the causal effects of transport improvements on production-related variables using firm level micro data. The detail in our transport network and firm data allows us to differentiate the impacts of the accessibility changes on employment and production from more general local effects and trends.

3 Empirical methods

As outlined above, we are interested in the effect of road improvements on employment and productivity. We measure the intensity of exposure to road improvements using changes in the accessibility to employment (or other measures of economic mass) along the road network. We adopt two approaches. The first one estimates aggregate effects for small spatial units (electoral wards). These estimates take into account firm exit, entry and relocation. In the second approach, we estimate firm level effects for existing firms. Both approaches use the same general estimation strategy applied to a panel of units (wards or firms) observed for up to 11 years, during 1998-2008. The data sources are described in Section 4 below.

3.1 General empirical set up

The underlying empirical model for our analysis is:

$$lny_{jt} = \beta lnA_{jt} + \mu_j + \tau_t + \varepsilon_{jt} \qquad (1)$$

Here y_{jt} is the outcome variable for unit j in year t, A_{jt} is a measure of employment accessibility along the road network (see Section 3.2 for details). Parameter β is the effect (elasticity) of accessibility on the outcome variable. [6] The unobserved component μ_j is a time invariant unit specific effect (for either wards or plants). Year effects τ_t represent general factors that influence all locations in a given year (e.g. macro shocks). The empirical specification also includes interaction of time effects with unit-specific control variables, supressed here for simplicity. Finally, ε_{jt} is an error term representing other unobservables.

We are interested in the estimation of parameter β, interpreted as the causal effect of accessibility on the outcome of interest. OLS estimates of equation (1) that ignore the time invariant component (μ_j), are biased if unobserved area effects are correlated with accessibility – if for example, and as seems likely, better transport connections and higher employment density have developed in places with productive advantages. [7] The first step to eliminating these biases is to eliminate fixed-over-time panel unit effects μ_j by demeaning the data within units (wards or plants), using the within transformation:

$$lny_{jt} - \overline{lny}_j = \beta\left(lnA_{jt} - \overline{lnA}_j\right) + \tilde{\tau}_t + \tilde{\varepsilon}_{jt} \qquad (2)$$

Here $\overline{lnx}_j$ denotes unit averages of lnx_{jt} over the time period (1998-2008), $\tilde{\tau_t} = \tau_t - \overline{\tau}$ and $\tilde{\varepsilon}_{jt} = \varepsilon_{jt} - \overline{\varepsilon_j}$. This formulation allows evaluation of the effects of transport policy on firms

[6] As explained in detail in section 4 below, accessibility is measured using the road network at the beginning of year *t*, and employment at the end of year *t-1*, and is linked to ward and firm level outcomes at the end of year *t*.

[7] As explained below in section 4.2, plant identifiers are location-specific, so plant fixed-effects are equivalent to including ward fixed-effect on the individual plant regressions.

because road construction generates changes in A_{jt} over time. Note, that the within-groups transformation is preferable to first differencing (which would also eliminate the fixed effects) when firm responses to accessibility change take time, because first difference estimation only uses changes occurring over a single year interval.

Within unit variation in accessibility over time can occur through: a) changes in the spatial distribution of employment, or b) because of changes in the transport network. But changes in accessibility through channel a) may be directly affected by the outcome variable or correlated with unobserved shocks in the error term of (2), leading to biased estimation. To address this issue, we instrument accessibility lnA_{jt} with $ln\widehat{A_{jt}}$. $\widehat{A_{jt}}$ is constructed only to pick up changes in accessibility due to new links in the transport network (i.e. channel b). We calculate the instrument using actual network changes combined with the *pre-improvement* spatial distribution of employment (for 1997, the first year in our sample). We then estimate equation (2) by two-stage least squares using this as an instrument for actual changes in accessibility. Details on the construction of A and $\hat{A}$ are provided in Section 3.2 below.

The instrumental variable (IV) estimates from (2) produce biased estimates if areas with different trends systematically experience different accessibility changes due to road improvements. That is, if $\tilde{\varepsilon}_{jt}$ in Equation (2) is correlated with the instrument ($ln\widehat{A_{jt}} - \overline{\ln \widehat{A_{jt}}}$). Such correlation occurs, for example, if transport investment is endogenous to employment and productivity trends in targeted locations, i.e. if the decision to improve the network is partly driven by unobserved production-related trends.

We address this potential bias by focusing the analysis on places and firms that are relatively close to transport improvements (within 20 km in our main results). In this way we compare places that differ incrementally in terms of accessibility changes experienced as a result of improvements. We argue that close to transport schemes, these differences are an incidental by-product of the scheme rather than its intended outcome (Michaels (2008) makes a similar argument for the US highway system). The main changes in travel times and accessibility occur close to the end points of improvements, although they are typically intended to improve the flow of traffic between areas *further away* from the improvement (Department of Transport 1997, Department of Transport 2009). There are also often long delays (10 years or more) between commissioning and opening of schemes, which weakens any link between local productivity trends and the decisions over where to site these projects.

We also drop wards (or firms located in them) which are crossed by or very close to improvements (within 1 km of any scheme) for three reasons. Firstly this further reduces bias due to targeting of wards as a result of endogenous routing of schemes e.g. if the route is chosen on the basis of low land prices (which may reflect low location productivity). Secondly, close to schemes production may be temporarily or permanently disrupted by construction works. Thirdly exclusion of the closest wards reduces error from inaccurate travel time calculations close to schemes, resulting from the map generalisation of our road network (see section 4.3.2).

To further reduce correlation of the instrument with underlying trends, we can control for differential trends in the vicinity of schemes. A set of 31 scheme dummies interacted with a linear time-trend in equation (2) control for different growth trends around the schemes. In some specifications we also control for linear time trends interacted with the straight line distance to closest scheme, a dummy indicating road opening year (or after) and salient electoral ward characteristics taken from 2001 Census data (unemployment rate, average age of population, proportion of population aged 16-74 with higher education and proportion of population living on social housing).

There are some specific points to consider when estimating equation (2) for plant level data. Firstly note that the plant identifiers in our data are location specific (changing if a plant moves to a different location). Thus, in the within-plant analysis changes in accessibility ($A_{jt}-\bar{A}_j$) are not related to relocation of plant j, but only to changes occurring at a fixed plant location, due either to transport improvements or changes of employment at other firms (and only the former when instrumenting with $\widehat{A_{jt}}$ as defined in Section 3.2). When instrumenting, estimation is only feasible using plants that exist, and appear in the data, both before and after the opening of the schemes that are used as the source of identifying variation. This means these plant level regressions do not capture changes in employment or productivity associated with the opening of new plants. In addition there are sample selection issues if firm re-location decisions in response to transport schemes are driven by unobserved characteristics that also affect outcomes (with a similar comment applying to the frequency with which firms appear multiple times in the data). These caveats aside, IV estimation of β from within plant changes give the micro-level impacts of improvements on firms, which are one component of the area level effects, and interesting in their own right.

Note, that the estimation strategy at both ward and plant level ignores whether or not specific firms or their employees and customers in fact use the network improvements. The effects are thus analogous to "intention to treat" estimates in the programme evaluation literature, and are the expected changes for firms or areas exposed to the 'treatment' (change in employment accessibility by road).

3.2 *Defining the accessibility index A and instrument Â*

Our aim is to estimate the causal effect of road improvements on economic outcomes using within unit changes over time. Therefore, a central challenge is to find a measure that captures changes in road infrastructure. Measures used in the literature include connectivity to the network (Faber, 2012), kilometres of roads within a given area (Melo et al., 2010; Duranton & Turner, 2011), distance to closest highway (Baum-Snow, 2007), number of rays crossing a given area (Baum-Snow, 2007, 2010), presence of highways in a given location in a particular year (Chandra & Thompson, 2000; Michaels, 2008), "lowest-cost route effective distance" (Donaldson, 2010) or even the amount of public expenditure on road infrastructure (Fernald, 1999). In our context, measures based only on connectivity or changes in the number of road kilometres are unlikely to capture effects of interest because the major road network is already very dense (49,816 km long in 1998) and does not expand much during our study period (increasing by 0.87% to 50,250 km by the end of 2008).

Instead, we use a measure of accessibility to employment along the road network. Accessibility to employment captures 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. A key advantage of this accessibility index over alternative indicators is that it varies continuously over space in ways that are partly unrelated to distance to improvements. This helps identify the effects of accessibility separately from the specific (dis)advantages of sites chosen for improvements. It also means we can potentially observe accessibility for all firms, irrespective of whether they are close to the site of the road improvement (though clearly firms closest to the improvements are more likely to use these new links, and hence tend to be the most exposed).

The accessibility index we use is identical in structure to market potential measures (e.g. Harris 1954) used in economic geography and agglomeration economies literatures, and to the accessibility indices used in the transport literature (e.g. El-Geneidy and Levinson 2006, Vickerman et al 1999), and is sometimes referred to as 'effective density' (Graham 2007). Consider a measure of economic activity, such as employment l. For a firm in location j at time t, an employment accessibility index A_{jt} is the weighted sum of employment in all destinations k that can be reached from origin j by incurring a transport cost c_{jkt} along some specified route between j and k (e.g. straight line or minimum cost route along a transport network). That is:

$$A_{jt} = \sum_{k \neq j} a(c_{jkt}) l_{kt} \qquad (3)$$

where $a(.)$ is a decreasing function of the cost of reaching destination k from origin j discussed further below. We use electoral wards as the spatial units, travel time along the major road network as the cost measure and employment as the measure of economic mass. Note that changes in A_{jt} are partly driven by changes in employment in destinations k. We can calculate accessibility, fixing employment at its initial (1997) level to give:

$$\hat{A}_{jt} = \sum_{k \neq j} a(c_{jkt}) l_{k1997} \qquad (4)$$

to ensure that changes in the index over time occur only as a result of changes in costs c_{jkt}, not in employment. As discussed above we can use $\hat{A}_{jt}$ to instrument for A_{jt}.

In Equations (3) and (4), $a(.)$ is a decreasing function of the cost of reaching destination k from origin j. Potential weighting schemes are discussed in Graham, Gibbons and Martin (2009). In line with common practice, we use the simple inverse cost weighting scheme ($a(c_{jkt}) = c_{jkt}^{-1}$) in which the cost is the "optimal" travel time between locations j and k at time t. Details on the construction of the origin–destination (O–D) matrix of optimal travel times are provided below. When constructing the matrix we apply a 75-minute drive time limit which facilitates computation but has negligible effects on the accessibility index because above this limit wards have negligible weights. We exclude location j from its own accessibility index to avoid mechanical endogeneity problems due to own outcomes appearing on the right hand side as a component of accessibility (although our IV strategy also solves this problem).

3.3 Accessibility changes arising from transport improvements

We calculate A_{jt} and $\hat{A}_{jt}$ using costs c_{jkt} in (3) and (4) based on routing along the transport network (see section 4.3 for more details). Transport improvements change the structure of the network and this changes costs between (some) origins and destinations. 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 will have a first order effect on the costs of the least-cost route between j and k if:

a) the least-cost route between j and k passes along the link p-q both pre and post-improvement, so the improvement reduces the cost of the journey along p-q and hence reduces c_{jkt}.

b) the least-cost route between *j* and *k* does not us link *p-q* pre-improvement, but switches to use the link *p-q* post-improvement because of the reduction in costs; again this reduces c_{jkt}.

In our empirical work we use these first order effects. We ignore second order effects arising when the least cost route between *j* and *k* does not use link *p-q* pre- or post-improvement but journeys between other origin and destination pairs switch to use link *p-q*, reducing congestion on the links in the network used by the routing between *j* and *k*. We have to ignore these second order effects because our network data does not allow us to observe travel time changes induced by changes in congestion resulting from improvements.[8]

4 Data sources and setup

4.1 Geographical units

Our analysis is based on plant level micro data. We have detailed information on the location of plants (postcodes) and can link this data geographically at various levels using the Office for National Statistics (ONS) National Statistic Postcode Directory. Postcode units correspond to a small number of addresses (around 14) or a single large delivery point such as a medium sized plant. As discussed in Section 3, for parts of our analysis we work with aggregates for circa 10,300 electoral wards. Ward boundaries are set so as to include roughly the same number of electorate. We use wards as defined in 1998, the first year of our study. This unit is very small, especially in dense areas.

[8] As discussed further below we deliberately exclude congestion when calculating the first order effect of a new link because this is possibly endogenous to changes occurring in nearby wards. This is likely to be much less of an issue for second order effects so in principle we might like to allow for these effects if data was available.

To construct the ward level characteristics used as control variables we use the Census 2001 from CASWEB. We calculate the share of population aged 15-64 with higher education, mean age of population, share of population living on social housing and the rate of unemployment. We also use straight line distance to the nearest improvement (undertaken at any point during our study period) and a dummy which indicates if the scheme is open, both calculated using GIS and the dataset of transport improvements described in 4.3. All these characteristics, as they are fixed over time, are interacted with a linear time trend in the estimation of the results.

4.2 Firm data

The data source for the aggregate analysis of employment and plant counts (number of establishments) at ward level is the Office for National Statistics (ONS) Business Structure Database (BSD) accessed through the UK's Secure Data Service (SDS). This is used both to construct dependent variables and for calculating the accessibility index. We use data from 1997 to 2008. The BSD is maintained by the ONS and contains a yearly updated register of the universe of businesses in the UK covering about 98% of business activity (by turnover).

The smallest unit of observation is the establishment or plant ("local unit"), but there is also information of the firm to which the plant belongs ("reporting unit") and the enterprise and enterprise group of the firms. The dataset provides detailed information on location (postcode), sector of production (up to 5 digit SIC) and employment of plants. It allows us to calculate employment and number of establishments at any geographical level aggregating up from postcodes. However, individual establishment identifiers are not always stable over long periods of time, which makes calculations of entry and exit of plants problematic.

For the productivity regressions, and for plant level employment regressions, we use the ONS Annual Respondents Database (ARD). The ARD holds responses to the Annual Business Inquiry (ABI) completed by a stratified random sample of units, extracted from the BSD (see Criscuolo, Haskel and Martin, 2003). The ABI is a comprehensive business survey covering balance-sheet information including gross value added, wages, intermediate inputs, employment, industry, and investment. We use the EU KLEMS Deflators (base 1995) to express the balance-sheet data in real terms. Although the ARD only contains a sample of small businesses, it is a census of large businesses so contains information for firms accounting for a large fraction of the employment (for example 90% of UK manufacturing employment). Historically, the ARD covered manufacturing, and is only reliable for service industries from 1997 onwards, which is why we limit the analysis to the post-1997 period.

We use the ARD to calculate plant (labour) productivity, output and labour costs. We cannot calculate total factor productivity for the whole of our study period because imputed estimates of the capital stock are only currently available up to 2004. Instead we look only at value-added (defined as price deflated revenue minus materials inputs), value-added per worker (labour productivity), gross revenue, and labour costs divided by employees (average wages).

As noted by Criscuolo et al (2003), a number of issues arise when deciding the level of aggregation at which to work. ARD reports information for both “local units” (LU) and “reporting units” (RU). Balance-sheet data is available at the RU level, while location and employment is available at the LU level. Employment can be used at LU level, since reporting units with multiple plants report on each local unit. For value-added per worker, the mean value-added per worker at RU (firm) level can be assigned to each LU (plant) assuming all plants are equally productive (see e.g. Criscuolo et al, 2012). Clearly this assumption need not hold. Note also that there is a risk of attenuating the estimated impacts of transport improvements on labour productivity, if not all plants are affected by the improvement, because the productivity gains of the affected plants will be combined with the productivity changes in unaffected plants when calculating firm level value-added per worker. Allocating value-added and revenue across plants from firm level data is similarly problematic. In short, given the structure of the data, strong assumptions are needed in order to calculate output related measures at the plant level.

4.3 *Road network data and origin-destination matrix construction*

The accessibility indices A_{jt} and $\hat{A}_{jt}$ are calculated using the ward employment data from BSD and the ward-to-ward origin-destination (O-D) travel times in each year. To calculate travel times we apply an optimal routing algorithm to a GIS-network that captures the location and travel speeds of the road links that exist in each year between 1998 and 2008. These GIS-networks are constructed by combining a dataset on new road schemes with two GIS-networks (for years 2003 and 2008) provided by the Department for Transport (DfT).

4.3.1 The road schemes

Information on completed road schemes for the British major roads network comes from information provided by the DfT and other sources including The Highways Agency, the Motorway Archive, Transport Scotland and Wikipedia. We analyse the effect of improvements carried out on major roads, which cover trunk roads, principal roads (class A) and motorways. Even if these roads only represent 13% of total road network length, they correspond to 65% of driven kilometres (Transport Statistics Great Britain 2010). We focus on major roads for two reasons. The first one is data availability: detailed data on road projects is only available for major schemes. The second reason is that transport policy is aimed at improving economic integration and reducing congestion of wide areas (Department of Transport, 2009), so we can expect the most substantial investments to be carried out on these types of roads.

These projects are diverse covering construction of new junctions, dualling, widening, upgrades and construction of new roads. We focus on new construction which we define to include new routes (where no direct link was previously available), faster routes (where a new road 'parallels' an existing minor link) and upgrades (where improvements, such as adding new lanes upgrade an existing minor link to be a major link). These improvements are the ones we can expect to have a substantial effect on travel times between wards and therefore provide a significant source of variation in optimal travel times.

From data collected on 75 projects completed between 1998 and 2007, restricting attention to new road construction leaves us with 31 road schemes, which are listed in table A1 in the Appendix. Some projects are small e.g. the A5 Nesscliffe bypass, costing £20.5 million. The largest project is the 6-lane, M6 Toll motorway bypassing Birmingham, which involved £0.9 billion of private investment. The total length of new links in our network between 1998 and 2008 is around 318 km. Total improvements represent 0.64% of the network length, with 43.6% corresponding to new roads and the remainder to faster routes or upgrades. Note that the lengths are measured from our simplified network data, and are less than the real length on the ground. The total change in major roads (motorways plus A-roads) reported in Transport Statistics Great Britain 2010 is 430 km, representing a 0.86% change. Figure A1 in the Appendix displays the location of these projects and the major road network at the end of our period of analysis (2008). Projects are scattered all over Britain.

4.3.2 The calculation of optimal travel times

Information on new road links is used to modify two GIS-networks, provided by DfT for years 2003 and 2008. The 2008 GIS-network contains all the major road links existing at the beginning of 2008. We use road construction as the source of variation in travel times. We geo-locate all the road links belonging to each of the 31 schemes listed in table A1 and match them to the 2008 road network. Starting from the 2008 network, in every year we remove the new links opened in that year in order to reconstruct the network as it was at the beginning of each year of the period 1998-2008.[9] In order to construct A_{jt} and $\hat{A}_{jt}$ we need to calculate the cost of travelling along each of the approximately 17,000 links in the network (c_{jkt}) which we assume are fully captured by travel times. A full measure of transport costs would require additional information on other characteristics of infrastructure (e.g. reliability), vehicle and energy use, as well as labor, insurance, tax and other charges (such as tolls). However, as demonstrated by Combes and Lafourcade (2005), using detailed French data, most of the spatial variation in transport costs is driven by infrastructure improvements. It seems likely that this will be the case in the UK too so, in practice, the major omission relates to issues such as reliability. For some projects (e.g. high speed rail) such reliability effects may be important. But for road projects changing travel times would be expected to be the most important factor affecting overall transport costs and so our data should provide a good proxy for changes to total costs.

[9] Some of the road schemes are bypasses around smaller settlements. Typically, bypasses replace a primary road through the settlement with the latter subsequently downgraded so the initial road is not present in our 2008 primary network. The same problem can apply to upgrades. Deleting links, therefore, creates artificial breaks in the network, when it comes to bypasses and upgrades. To correct for this we keep the bypasses in the network in years pre-opening but assume (consistent with available scheme evaluations) that travel time opening year was twice the post opening travel time.

In order to calculate travel times, we use data on traffic speeds from the 2003 generalised primary road GIS-network provided by DfT. We use journey times in the non-busy direction averaged over all time periods between Monday-Friday 08:00 and 18:00. We 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 (although this makes little difference in practice). Due to data availability, we use journey times in 2003 (based on 2003 traffic flows) for the whole period 1998-2008.

For links opened after 2003 we use estimated journey times from a regression model of link speed on various link characteristics (including traffic flows). The regression predicts speeds on the 2003 links reasonably well (R-squared = 0.76). We then use the regression results and link characteristics to predict travel times for links opened after 2003 for which no speed data is available. For some of the links, the prediction exceeded travel time implied by the speed limit. We replaced predicted speed with the speed limit for these links. Because we use traffic flows as one of the link characteristics that predict speeds we would reduce endogeneity problems if we could use traffic flows at the beginning of our study period. Given the only data are available for flows (for all links on the network) are for 2003 and 2008 we are forced to compromise. For the links existing in 2003 we use the modelled speeds provided by the DfT, which are calculated using 2003 traffic flows. For the links built between 2003 and 2008 we have to use 2008 flows. More detail on the construction of the O-D travel time matrix is provided in the Appendix.

It should be noted that the network is highly generalised. Journeys via the minor road network are not modelled nor are forbidden turns and one way systems. All link intersections are treated as junctions. When computing the O-D matrix we also apply a limit of 75 minute drive time. Changes in accessibility must therefore be regarded as approximate. This measurement error means our estimates of the effect of accessibility could be downward biased as a result of attenuation. To partially address this concern, we cross checked a sample of times and accessibility measures against estimates derived from Google maps, using the STATA 'travel time' module (Ozimek and Miles 2011). The cross sectional correlations in the journey times are high (in the order of 0.6-0.8), and the correlations in the accessibility indices (using address counts rather than employment) are even higher (0.8-0.95. However, the correlation for travel times is weaker for shorter journeys, presumably because shorter trips that do not use our generalised network are poorly approximated by our O-D calculation. For this reason, and because locations immediately proximate to new schemes may be adversely affected by the scheme (e.g. loss of premises, and environmental impacts), we drop wards and plants within 1 km of the road schemes in our analysis. As discussed above, this also helps to further mitigate concerns about the targeting of specific wards as a result of endogenous routing of schemes.

5 Results

5.1 Descriptive statistics

Table 1 summarises changes in (log) accessibility from 1998 to 2008 (reported as approximate % changes). In the upper panel we calculate overall accessibility allowing both employment and road infrastructure to vary (A). In the lower panel the accessibility measure is calculated fixing employment at 1997 level ($\hat{A}$), so that all the variation in accessibility comes from changes in the road network. Comparing the upper and lower panels shows that most of the variation in accessibility over time comes from changes in the spatial distribution of employment. The average accessibility change due to major road schemes was on average only 0.34%. However, this increases substantially for wards closer to improvements, and it is this variation which provides the basis for our IV strategy. For example, for wards within 10 km of a scheme the mean change is 1.2% and the 90th percentile is 3.2%. Within 20 km, which we use in our base specification, mean accessibility change is 1.2% and 90th percentile is 2.0%. Within 30 km these figures are 0.95% and 1.7%, respectively.

Maps 1 and 2 illustrate the spatial relationship between road schemes and the resulting accessibility increases.[10] The left panel of Map 1 shows new roads and major improvements and the right panel shows how the road schemes improved accessibility in the surrounding areas. Map 2 focuses on the Manchester-Leeds area in order to illustrate the identification strategy. The thin light white lines show the primary road network in 2008. New links between 1998 and 2007 are indicated by bold lines. The dark grey lines are ward boundaries. Clearly the effect of improvements on accessibility varies considerably across wards in the vicinity of the same improvement. As detailed above, we argue that differences in accessibility changes across wards near to schemes are coincidental and can be treated as exogenous, especially when controlling for differential time trends near different schemes.

As a partial test of this assumption that restricting our samples to within narrow distance bands reduces the problems associated with the targeting of road transport improvements to areas with higher/lower productivity or employment, we carried out a range of balancing regressions which show that, within our preferred 1-20km distance band, distance to a road transport scheme is largely uncorrelated with the initial ward conditions. These results are

[10] Note that the figures in the maps differ slightly from those in Table 1, because the maps do not exclude wards within 1km of the new roads.

shown in Appendix Table A2. These results are coefficients from regressions of ward level characteristics in the initial period (employment in 1997, GVA per worker in 1998, and so on as listed in the table headings) on distance to the nearest road scheme, with scheme dummy variables, on the sample of wards within the 1-20km band. As can be seen, in most cases there is a near-zero and insignificant coefficient in all these regressions, suggesting no relationship between employment, plants, population (measured by residential addresses), gross value-added or gross value added per worker, labour costs per worker or gross output per worker. We do find marginally significant coefficients for the gross output (revenue) and total labour costs, although as we show later, there is no substantive difference between the results on the effects of the transport schemes on these different measures of economic output and productivity, so we have no reason to suspect that these significant coefficients indicate an important failure of our identifying assumptions.

Appendix Table A3 provides further descriptive statistics for the accessibility index, and levels of employment and numbers of plants and total employment in the wards in our estimation samples.

5.2 Ward-level employment and plant count regressions

The first results, presented in Table 2, are for ward level regressions of log employment on log accessibility. Employment is from BSD data as described in Section 4.2. The main results in columns (1)-(5) use data on wards between 1 km and 20 km of road schemes for various model specifications. Columns (6) and (7) present the same specification as column (5), but applied to samples within 30 km and 10 km of schemes, respectively. Standard errors and F-stats are 'clustered' at ward level, to allow for arbitrary intra-ward correlation over time.

Alternative higher-level clustering schemes (e.g. based on Local Authorities) that allow for a degree of inter-ward error correlation give similar results – see Section 5.3 below.

The first specification in column (1) is a simple OLS regression which ignores all endogeneity issues. Coefficients are positive and significant showing that wards with higher employment tend to have better accessibility, which is probably unsurprising, since high employment wards tend to be near other high employment wards. In the second column, we add ward fixed effects to control for time-invariant ward specific factors.[11] The point estimate on accessibility reduces substantially, but still remains large, positive and significant at the 1% level. Column (3) uses accessibility index $\hat{A}$ as an instrument. Recall $\hat{A}$ keeps ward employment fixed at the 1997 level and so accessibility only varies due to road improvements. The first stage F statistics indicate that the instrument is very strong so not subject to weak instrument problems (Staiger and Stock, 1997). This is unsurprising given the mechanical relationship between $\hat{A}$ to and A. The point estimate in the IV specification is close to that in column (2), although less precisely estimated. The remaining columns of Table 2 add more control variables. In column (4), a dummy indicating the nearest scheme to the ward is interacted with a time trend to allow for changes in employment and accessibility that are common to all wards within a 20 km radius of a given scheme. Column (5) goes further and introduces a time trend interacted with distance to the scheme, and interacted with a dummy for the scheme being opened to allow for other common time patterns that might cause accessibility and ward level employment to move together over time. In both cases, the point estimate remains large, the magnitude increases and

[11] As discussed in section 3.1 we use the within group transformation to eliminate ward fixed effects. Results when first differencing (which rely on year-on-year changes for identification) are generally insignificant in the more demanding specifications.

significance improves relative to the basic IV estimate in column (3). In robustness results not reported, we included an interacted time trend with a set of census variables for each ward, to allow for time patterns related to the underlying demographics. In this overly saturated model the estimate becomes only marginally significant, although still large in magnitude which we interpret as evidence that the earlier results are not driven by unobserved common time trends.

Expanding the area considered in column (6) leaves the results unchanged, while reducing the area to within 10 km of schemes, in column (7), leads to smaller, less precise estimates. This might be surprising, given the largest changes in accessibility appear in these areas (see Maps 1 and 2). However, as pointed out in Section 4.3, some of these accessibility changes close to schemes are imprecisely measured due to the generalisation of our road network which lacks minor road detail. In keeping with this explanation, additional results (not reported) show that exclusion of wards within 2, 3 or 4 kilometres of the new links improves the precision of the estimates when using the 10 km band.

The headline story from these results is that accessibility changes induced by road improvements drive up local employment, with an elasticity of around 0.25-0.35. These estimates appear quite large, but remember that the actual changes in accessibility induced by the transport schemes in this study are small (see Table 1). On average, within 20 km the mean change in accessibility was only 0.83%, so the induced change in ward employment from the average scheme would be only around 0.25%.

The panels in Table 3 look at results by broad industrial sector.[12] The structure is otherwise identical to Table 2. The results suggest that most of the action on employment comes from producer services, land transport and 'other' sectors (a residual category that includes the primary and public sectors). Both the land transport and producer services effects are consistent with a story in which road transport has lowered transport costs for intermediates and business travel and stimulated employment in the logistics sector, though we cannot go further in pinning down the precise mechanisms empirically. The elasticity in the transport and producer services sectors is as high as 1 in some specifications. Once again, expanding to a 30 km distance band leaves results unchanged, while reducing the area to within 10 km leads to smaller, less precise estimates.

Moving now from employment to the number of plants, Table 4 and Table 5 present results analogous to those in Table 2 and Table 3, but for plant counts (number of establishments) from the BSD at ward level. Plant counts are potentially more reliable than employment measures, so we might expect the effects here to be more precisely measured. This turns out to be the case. The general picture from Table 4 is very similar to that for employment, although the results are even more stable and significant across different specifications, including within the 10 km band. Evidently, it is likely that the employment changes are at least partly driven by increases in the number of plants in wards experiencing transport improvements. Splitting the results by sector yields similar findings, with the notable exception that, in addition to effects on the land transport and producer services sector, we

[12] We use the 1992 Standard Industrial Classification (SIC) at 2 digits to define the 6 wide industrial categories. Manufacturing includes sector codes 15 to 37; construction includes sector codes 40, 41 and 45; consumer services includes sector codes 50 to 59; producer services includes sector codes 65 to 74, and land transport includes sector code 60. Other includes the rest of sectors, including primary activity, public sector, rest of transport and other sectors.

now detect strong impacts on the number of plants (i.e. firm entry and exit) in the manufacturing sector. The elasticity of plant numbers in the manufacturing sector is in the order of 0.4-1.

5.3 *Robustness checks: alternative accessibility indices, distance bands, spatial autocorrelation*

Table 6 presents the results of regressions (identical to column 5 of Table 2 and Table 4) but using alternative measures of accessibility as proxies for the intensity of exposure to the transport improvements. Column (1) uses residential population accessibility, computed by replacing ward level employment with counts of post office residential delivery addresses taken from the ONS National Statistics Postcode Directory when constructing A and $\hat{A}$. The coefficients here are slightly higher than when using employment accessibility, but of a similar magnitude given the standard errors. Column (2) uses an index of accessibility to plants, constructed in the same way from the BSD plant counts. These are nearly identical to those obtained using employment based accessibility indices.

Columns (3)-(5) experiment with alternative distance weighting schemes. The coefficients are somewhat higher when penalising distance less heavily in an inverse distance weighting scheme (column 3) and lower when penalising distance more heavily (column 4). Switching to an exponential distance weighting function (column 5) generates a bigger coefficient. The differences in the scale of these parameter estimates are largely explained by the change in the variance of the accessibility measures under these alternative weighting schemes. If we standardise the effects (divide by the standard deviation of the accessibility variables) we find a much more stable pattern. In sum, there is no evidence here that the results are substantively sensitive to changes in the definition of the accessibility index.

It is useful to explore, in more detail, at what distances these transport impacts occur. We have shown that the employment effects are strong within 20 km and 30 km, but appear weaker close to the schemes within 10 km. On the other hand, the plant count effects are strong at all distances. Table 7 explores this further using the standard IV specification of column 5 of Table 2 and Table 4. Column (1) simply repeats the results for 1 to 10 km, then columns (2) and (3) show results for bands from 10 to 20 km and 20 to 30 km. Looking at the employment results, it becomes clear that effects from 10 to 20 km dominate. Employment effects are still large, but imprecisely measured in the outer ring beyond 20 km.[13] For plant counts, the effects are again strong and significant in all distance rings, and increase in magnitude as we move outwards. These results suggest there is no strong evidence for the effects being driven predominantly through displacement from outer to inner wards nearer the scheme. Of course we cannot determine from this analysis (or probably any other) whether the employment and plant count effects come about through displacement of activity from low to high accessibility-change wards, or whether the gains are truly 'additional'.

One concern about the fact that we are using data on closely spaced wards and firms is that the unobservables in our regression models are spatially autocorrelated, leading to biased standard errors and incorrect inferences. Given we include ward fixed effects, scheme specific trends, and distance to scheme trends as control variables, this problem is probably not as important as it might at first seem. We need only be concerned about spatial autocorrelation in the deviations around these fixed effects and trends, not the simple cross

[13] Note, that it is not possible to add up the coefficients across these rings to arrive at the baseline estimates within 10-30km, because of two stage least squares procedure.

sectional patterns. Nevertheless, we carried out some direct tests by computing the Moran's I statistics for the residuals from our regressions, as reported in Appendix Table A4. The Moran's I statistics show no evidence of spatial autocorrelation in the residuals with tiny values (less than 0.01), even if these are occasionally statistically significant. In Appendix Table A5, we also report our main results with standard errors clustered at a larger geographical level, the Census district in which the ward or plant is located. If spatial autocorrelation was an issue we would expect to see a marked difference between the standard errors at these different levels of clustering, but comparison of the results in Appendix Table A5 with the main results in Table 2 to Table 5 shows that there is very little difference.

5.4 Plant-level employment regressions

The main results in the preceding tables suggest that increased accessibility leads to increased ward employment and number of plants, at least for some sectors. These findings could be driven by existing firms increasing employment, or by new firms entering. The plant count results show that firm entry appears to contribute to employment changes, but we can explore the issue further by looking at within-plant changes. [14] To do this, we estimate the effect of log accessibility on log plant employment using data from ARD for 1998-2008.

Table 8 presents the key results (the sectoral breakdown is in the Appendix table A6) using a similar structure to earlier tables. Additional control variables in these plant level regressions are sector-year dummies (using the 6 broad sectors used for the sector-specific

[14] BSD data on entry and exit rates prove to be too noisy to get precise estimates but the data proved too noisy to get reasonably precise estimates if used directly.

results above) and a dummy and a trend specific to single plant firms ('singletons'). As before, standard errors and F-statistics are clustered at ward (i.e. treatment) level.

The coefficients in Table 8, even in the simple OLS regression, start off small and rapidly become near-zero, insignificant and even negative as we add in additional controls. Note that the negative and marginally significant coefficient in the 10 km band is consistent with the pattern observed in Table 2 and Table 4. With ward level employment constant, but the number of plants up, employment within plants must shrink. This could arise due to firm restructuring or it may simply be a statistical artefact given that we have considered coefficients from many different specifications. More generally, it is clear that the overall area employment effects in Table 2 are due to new firms entering rather than within-plant changes. In short, transport improvements have very little effect on existing firms' decisions to expand or contract employment, but do have a sizeable impact on firm entry.

5.5 *Productivity and other production related outcomes*

Although we find no response for existing firms on the employment margin, these firms may experience productivity gains if lower transport costs allow reorganisation which results in increased output. We explore this possibility directly by looking at value-added per worker, gross-value added, real revenue, and wages using data from ARD during years 1998-2007. Specifications are identical to those in Table 8 although ARD data is available for one less year. Recall from Section 4.2 that output and value added are only available at the higher RU level so approximations must be made to allocate these to plant level for multi-plant firms. The key results for all sectors pooled together are in Table 9. We restrict attention to the 20 km radius, excluding plants located within 1 km of the schemes.

The headline story here is that we observe positive effects of accessibility on all outcomes with the majority of coefficients significant. The effects become bigger and more significant as we introduce more control variables. The preferred specifications at the right of the table, suggest that the elasticity of labour productivity with respect to transport induced accessibility changes is 0.5. The figures are similar for gross value added and revenue (consistent with the finding of no significant employment effects at plant level). Average wages (total wage bill per worker) increase too, implying that at least part of the productivity increase is paid out in worker wages. Again, it's important to remember that the average mean accessibility change within 20 km is 0.83%, so these coefficients imply induced productivity effects ranging from 0.2% for wages, to 0.4% for GVA per worker. Sector specific results, presented in Appendix Table A7, are more mixed. No clear pattern emerges here, because many of the estimates are very imprecise, although nearly all are positive and of similar orders of magnitude. In terms of significance, the consumer services sector results are strongest. Together with the sectoral results for employment, a picture emerges in which transport improvements induce entry of firms in most sectors apart from consumer services and construction, but with no employment effects for existing firms. In contrast, existing consumer services firms experience the most significant output, labour productivity and wage increases.

Additional analysis for these productivity-related outcomes aggregated to ward level, suggests that the productivity effects are not strongly evident at the area level (see Table 10). Although there are positive value-added effects, consistent with the earlier results, these are statistically insignificant. We observe no effects on labour productivity at this level (similar results obtain if we re-estimate the plant level regressions with employment weights). One way of reading these findings, in the light of earlier results, is that the inflow of new firms and increases in employment swamp the plant level productivity effects leaving area labour productivity relatively unchanged. In addition, if the productivity gains are concentrated on small and medium sized existing firms, we would expect to see weaker results at the aggregate level. This is because in the micro data, the sample is dominated by many smaller firms, whereas when aggregated taking into account employment shares, the effects in the larger firms play a stronger role. It turns out that the productivity effects observed in Table 9 are indeed concentrated in the small to medium sized firms with 6-50 employees, as shown in Appendix Table A8. These smaller firms, whilst making up nearly 60% of the firms, contribute less than 30% to total employment.

6 Conclusions

This paper uses unique data and innovative methods to assess the productivity and employment effects from transport improvements at a very detailed geographic scale. We measure the intensity of exposure to improvements using changes in employment accessibility constructed at a micro-geographic scale (electoral ward level). These accessibility indices are similar to those increasingly used in transport project appraisal. These are constructed using GIS network analysis of data on the major road network in Britain, and changes that occurred to it between 1998 and 2007, coupled to data from the administrative register of employment and businesses in Britain. We use a panel data, instrumental variables strategy applied only to wards relatively close to schemes to address the likely endogeneity of changes in employment accessibility. Identification comes from change in accessibility due to transport improvements. We argue that methods using cross-sectional variation in accessibility, effective density or other forms of 'agglomeration' or market potential indices are biased by the endogenous placement of transport infrastructure and the spatial restructuring of employment, so do not have a causal interpretation.

Our estimates of the benefits from transport improvements relate to those impacts that can be detected through changes in employment accessibility. These should incorporate agglomeration effects, and any direct effects related to transport cost savings that are correlated with the accessibility changes. By design, the effects we detect are fairly local to the sites of improvement, because any road link improvements feature more often in the optimal routes from nearby locations. Overall, we find strong effects from transport improvements on area employment and plant counts. A 10% improvement in accessibility leads to about a 3% increase in the number of businesses and employment, up to 30 km from the site of the improvement. The estimates range between zero and 10% according to sector and specification. The employment increases appear to come about through firm entry, rather than increases in the size of existing firms. We do, however, find evidence for increases in labour productivity, output and wages amongst existing firms, although these are not so evident at area level.

These effects appear substantial, when roughly translated into the expected increase in GDP (value-added) as a result of the public investment in new roads. The average effect of all the major new road schemes in Britain between 1998 and 2007 was to raise mean accessibility at ward level by 0.34% (Table 1). This implies a 0.012% increase in total employment from a year's investment in major road transport network improvements (using the elasticity of 0.36 from Table 2). Although a very small effect, if extrapolated to the whole workforce (roughly 30 million in 2008), the implied increase in total employment is 3600. It is difficult to assess the contribution to the economy from this increase in employment without information on the contribution to output from the marginal worker employed as a result of the improvements. An upper bound might be based on the average gross value-added per worker of £41000,[15] giving a value to one year of transport investment of around £148 million per year, or £4.2 billion in present value terms (using a 3.5% discount rate). A lower bound can be calculated come from the minimum wage (£6.00 per hour) suggesting benefits of £36 million per year (assuming a 35 hour week, 48 weeks per year), or £1 billion in present value terms. For comparison, expenditure on major road infrastructure investment of these types in 2007/8 was £1.8 billion (National Transportation Statistics 2010). A full cost-benefit analysis, which would require us to take in to account the opportunity cost of the resources (e.g. capital and labour) used in production as well as any productivity effects, environmental and other benefits, is beyond the scope of this paper.,.

Our results add substantially to existing evidence on the effects of transport policy on area and firm level economic outcomes. Our analysis highlights the importance of addressing endogeneity issues in a convincing way. We argue that utilizing small scale spatial variation in the impact of transport improvements on accessibility offers a promising quasi-experimental setting, even though data requirements are high. We provide evidence both at the area and firm level, allowing us to investigate the micro channels driving the area level results. Furthermore, we test the effects on a variety of outcomes (employment, number of plants, labour productivity, gross output and average wages), which sheds light on the impacts for economic outcomes for which existing evidence is scarce.

[15] Own calculation based on ONS 2011

7 References

Aschauer, D. A. (1989). “Is public expenditure productive?” Journal of Monetary Economics, 23(2): 177–200.

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

Baldwin R. E. and T. Okubo (2006). “Heterogeneous firms, agglomeration and economic geography: spatial selection and sorting”, Journal of Economic Geography 6(3): 323-346.

Baum-Snow N. (2007). “Did highways cause suburbanization?”, Quarterly Journal of Economics 122(2): 775–805.

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., and Tabuchi, T. (2004). “Testing the home market effect in a multi-country world: The theory”, CEPR Discussion Paper No. 4468.

Boarnet, M. (1998). “Spillovers and the Locational Effects of Public Infrastructure”, Journal of Regional Science, 38(3): 381 – 400.

Chandra, A and 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, July.

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

Criscuolo, C., J. Haskel, and R. Martin, “Building the Evidence Base for Productivity Policy Using Business Data Linking,” Economic Trends 600 (2003), 39–51.

Criscuolo, C., R. Martin, H. G. Overman and J. VanReenen (2012), “The Causal Effects of An Industrial Policy”, NBER Discussion Paper 17842.

Combes, P.P and Lafourcade, M. (2001), "Transport Cost Decline and Regional Inequalities: Evidence from France", CEPR Discussion Papers 2894.

Combes, P.P and Lafourcade, M. (2005), " Transport Costs: Measures, Determinants, and Regional Policy Implications for France ", Journal of Economic Geography 5(3): 319-349.

Coughlin C.C. and Segev E., (2000). "Location Determinants of New Foreign-Owned Manufacturing Plants", Journal of Regional Science, Blackwell Publishing, 40(2): 323-351.

Department for Transport (1997) “A new deal for trunk roads in England: Understanding the new approach” Department of Transport, London

Department of Transport (2009) Highways Agency Framework Document, Department of Transport, London

Donaldson, D. (2010). “Railroads of the Raj: Estimating the Impact of Transportation Infrastructure”, National Bureau of Economic Research, NBER Working Paper No. 16487.

Duranton, G. and Puga D. (2004). "Micro-foundations of Urban Agglomeration Economies" in J. F. Thisse and J. V. Henderson (eds.) Handbook of Urban and Regional Economics, Volume 4. North Holland.

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

Duranton, G. and Turner M.A., (2012). "Urban Growth and Transportation", forthcoming Review of Economic Studies.

Duranton, G., Morrow P and Turner M.A., (2012)."Roads and Trade: Evidence from the US", mimeo University of Toronto.

Eddington, R (2006) The case for action: Sir Rod Eddington's advice to Government, HMSO London

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.

European Commission (2006) Road Transport Policy, Open Roads Across Europe, European Commission, Energy and Transport DG, B-1049 Brussels

EU Transport in Figures Statistical Pocketbook (2011), Transport Statistics Publication, European Commission, Publications Office of the European Union, Luxembourg.

Faber, B. (2012), "Market Integration and the Periphery: The Unintended Consequences of New Highways in a Developing Country", Mimeo, 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.

Garcia-Mila, 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.

Gibbons, S. and 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.

Gibbons, S. and Overman, H. G., (2009). “Productivity in transport evaluation studies”, Report for Department of Transport, London.

Graham D. J. (2007). “Agglomeration economies and transport investments”, OECD/ITF Joint Transport Research Centre Discussion Papers 2007/11, OECD, International Transport Forum.

Graham, D. J., Gibbons, S. and Martin, R. (2009). “Transport investments and the distance decay of agglomeration benefits”, draft report for the Department of Transport.

Harris, C. (1954).”Market as a Factor in the Localization of Industry in the United States”, Annals of the Association of American Geographers, 44(4) 315-348.

Henderson, J. V. (2003). “Marshall’s scale economies.” Journal of Urban Economics, 53(1): 1–28.

Holl, A. (2004a). “Manufacturing location and impacts of road transport infrastructure: Empirical evidence from Spain”, Regional Science and Urban Economics 34(3): 341-363.

Holl, A. (2004b). "Transport infrastructure, agglomeration economies, and firm birth". Empirical evidence from Portugal', Journal of Regional Science, 44(4): 693-712.

Holl, A. (2004c). "Start-Ups and Relocations: Manufacturing Plant Location in Portugal", Papers in Regional Science, 83(4): 649-668.

Holl, A. (2006). "A review of the firm-level role of transport infrastructure with implications for transport project evaluation", Journal of Planning Literature 21(1): 3-14.

Holl, A. (2011). "Market potential and firm-level productivity in Spain", Journal of Economic Geography, in press.

Holtz-Eakin, D. (1994). "Public-sector capital and the productivity puzzle". The Review of Economics and Statistics, 76(1): 12–21.

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.

Lahr, M. L., Duran, R., & Varughese, A. (2005). Estimating the Impact of Highways on Average Travel Velocities and Market Size. EconWPA Urban/Regional 04003009, Center for Urban Policy Research.

Li, Z. and Li, H. (2010). " Road investment and inventory reduction: evidence from a large developing", country. Mimeo, Hong-Kong University.

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

Melitz M. (2003). “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity”, Econometrica 71: 1695-1725.

Melo, P. and Graham, D.J. and Noland, R. B., (2010), “Firm Formation and Transport Infrastructure: A Study of Portugal”, Transportation Research Record: Journal of the Transportation Research Board, 2163: 133—143.

Moreno, R. & Lopez-Bazo, E. (2007). “Returns to local and transport infrastructure under regional spillovers.” International Regional Science Review, 30(1): 47.

Michaels, G (2008) “The effect of trade on the demand for skill: evidence from the Interstate Highway System”. Review of economics and Statistics, 90(4): pp. 683-701.

Office for National Statistics, *Annual Respondents Database, 1973-2009: Secure Data Service Access* (computer file). Colchester, Essex: UK Data Archive (distributor), June 2012. SN: 6644, http://dx.doi.org/10.5255/UKDA-SN-6644-1

Office for National Statistics, *Business Structure Database, 1997-2010: Secure Data Service Access* (computer file). *2nd Edition.* Colchester, Essex: UK Data Archive (distributor), November 2011. SN: 6697, http://dx.doi.org/10.5255/UKDA-SN-6697-2

National Statistics Postcode Directory (2010), ONS Geography, Office for National Statistics.

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

Ozimek Adam and Daniel Miles (2011) “Stata utilities for geocoding and generating travel time and travel distance information”, Stata Journal 11(1) 106-119

Small, K. (2007) The Economics of Urban Transportation, Routledge, New York and London

Staiger, Douglas and James Stock. 1997. "Instrumental Variables Regression with Weak Instruments," Econometrica, 65(3): 557-586.

Straub S., (2011). "Infrastructure and Development: A Critical Appraisal of the Macro-level Literature," The Journal of Development Studies, Taylor and Francis Journals, 47(5): 683-708.

Transport Statistics Great Britain (2010), Transport Statistics Great Britain, 25th November 2010, Department for Transport.

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

U.S. Department of Transportation (2012), Research and Innovative Technology Administration Bureau of Transportation Statistics, National Transportation Statistics.

Vickerman R., Spiekermann K, and Wegener M. (1999). "Accessibility and Economic Development in Europe", Regional Studies 3(1): 1-15

World Bank (2008). "World Bank Transport Business Strategy 2008-2012," Washington DC.

Appendix: Further details on construction of road travel times

Information on new road links is used to modify two GIS-networks, provided by DfT for years 2003 and 2008. The 2008 GIS-network contains all the major road links existing at the beginning of 2008. 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 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. 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.

We use road construction as the source of variation in travel times over time. We geo-locate all the road links belonging to each of the 31 schemes listed in table A1 and we match them to the 2008 road network based on their CP code. Starting from the 2008 network, in every year we remove new links opened in that year to reconstruct the network as it was at the beginning of each year of the period 1998-2008. The exact treatment of new links depends on the type of link. Projects fall into two categories. Type 1 (new routes) corresponds to "genuinely" new roads, i.e. roads for which we do not have an alternative minor road flowing in parallel. As we move backwards to recursively reconstruct the network, we simply remove these links. Type 2 projects (faster routes) correspond to either (a) roads for which there was an alternative route before, but the road was a minor road (not existing in the major road network) or (b) an upgrade (which involves improvement and the construction of new lanes) so the road becomes part of the major road network. These mainly correspond to bypasses which relieve traffic congestion from villages and usually flow in parallel to an existing alternative minor road. Type 2 schemes usually involve the downgrading of an existing link (so the old link is not present in our 2008 primary road network). This causes an artificial break in the primary road network when we delete these links based on their opening years. To resolve this problem we keep type 2 projects in the network in pre-opening years and assume that travel time was twice that post-opening. Scheme evaluation reports support the assumption of significantly longer travel time.

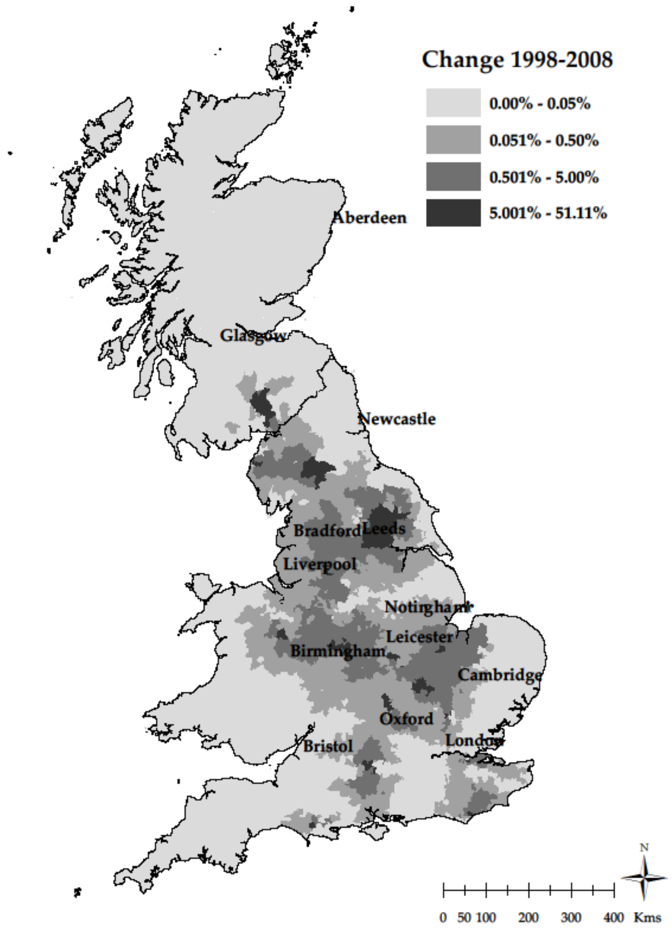
In order to calculate travel times, we use data on traffic speeds from the 2003 generalised primary road GIS-network provided by DfT. Traffic speeds are modelled from traffic flow census data using the Road Capacity and Costs Model (FORGE) component of the National Transport Model (NTM).[16] As discussed in the text, we use journey times, obtained from FORGE, in the non-busy direction averaged over all time periods between Monday-Friday 08:00 and 18:00. Due to data availability, we use journey times in 2003 (based on 2003 traffic flows) for the whole period 1998-2008.

[16] 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. 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. As explained in the DfT documentation: "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 links 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 we use in the calculation of travel times between wards.

The model that we use to estimate journey times after 2003 regresses 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. As we discuss in the text, the regression predicts speeds from the FORGE reasonably well (R-squared = 0.76). We then used the regression results to predict travel times for links opened after 2003 for which no FORGE speed is available.

8 Figures

Map 1. Road improvements and accessibility changes from 1998 to 2008



Map 2. Changes in accessibility due to road improvements from 1998 to 2008 in the Manchester-Leeds area

9 Tables

Table 1. Change in log accessibility 1998-2008

	Wards	Mean	Std. Dev	90th percentile	Max	Proportion of zeroes
Time-varying employment and time-varying travel times						
All	10318	7.51%	8.06%	13.88%	137.07%	0.00%
1-10 km	1514	6.39%	6.03%	12.04%	52.10%	0.00%
1-20 km	3487	6.32%	5.75%	11.68%	67.09%	0.00%
1-30 km	4903	6.37%	5.57%	11.75%	67.09%	0.00%
1997 employment and time-varying travel times						
All	10318	0.34%	1.22%	0.79%	31.37%	32.52%
1-10 km	1514	1.18%	2.45%	3.16%	31.37%	5.28%
1-20 km	3487	0.83%	1.97%	1.91%	31.37%	6.05%
1-30 km	4903	0.66%	1.71%	1.57%	31.37%	6.00%

Notes: Table provides summary stats for wards located more than 1km from any road construction site over the period of analysis. Source: Own calculations using BSD.

Table 2: Ward level employment regressions: all sectors

ALL SECTORS	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log accessibility	0.429***	0.252***	0.275*	0.363**	0.361**	0.355**	0.199
	(0.020)	(0.060)	(0.162)	(0.176)	(0.178)	(0.177)	(0.207)
Observations	38247	38247	38247	38247	38247	53823	16566
Wards	3477	3477	3477	3477	3477	4893	1506
Distance band	1-20 km	1-20 km	1-20 km	1-20 km	1-20 km	1-30 km	1-10 km
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ward FE		Yes	Yes	Yes	Yes	Yes	Yes
IV			Yes	Yes	Yes	Yes	Yes
IV first stage F-stat			2667	4958	4651	4653	4522
Scheme trends				Yes	Yes	Yes	Yes
Controls					Yes	Yes	Yes

Notes: Standard errors in brackets (clustered at the ward level). IV first stage is the first stage Kleibergen-Paap F-stat for the IV specifications. *, **, *** indicate significant at the 10%, 5% and 1% levels, respectively. Controls include a linear trend interacted with the distance to closest improvement and a trend interacted with year of opening. Reduced form estimate in preferred specification (5) is 0.399*** (0.198).

Table 3: Ward level employment regressions: by sector

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
MANUFACTURING							
Log accessibility	0.252***	0.526***	0.411	0.09	-0.003	0.226	0.033
	(0.034)	(0.201)	(0.501)	(0.519)	(0.526)	(0.512)	(0.600)
CONSTRUCTION							
Log accessibility	0.239***	0.188	0.239	0.068	0.221	0.1	0.397
	(0.022)	(0.163)	(0.336)	(0.342)	(0.348)	(0.341)	(0.430)
CONSUMER SERV							
Log accessibility	0.448***	0.107	0.068	-0.199	-0.276	-0.315	-0.435
	(0.023)	(0.086)	(0.253)	(0.281)	(0.288)	(0.282)	(0.352)
PRODUCER SERV							
Log accessibility	0.749***	0.649***	1.646***	0.995***	0.878**	1.014***	0.898**
	(0.027)	(0.126)	(0.388)	(0.371)	(0.376)	(0.377)	(0.453)
LAND TRANSPORT							
Log accessibility	0.198***	0.637**	1.285**	1.177*	1.06	1.078*	0.061
	(0.028)	(0.286)	(0.599)	(0.640)	(0.647)	(0.642)	(0.753)
OTHER							
Log accessibility	0.344***	0.095	0.482**	0.570**	0.617***	0.653***	0.334
	(0.020)	(0.081)	(0.219)	(0.229)	(0.232)	(0.233)	(0.287)
Observations	38246	38246	38246	38246	38246	53820	16566
Ward clusters	3477	3477	3477	3477	3477	4893	1506
Distance band	1-20 km	1-20 km	1-20 km	1-20 km	1-20 km	1-30 km	1-10 km
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ward FE		Yes	Yes	Yes	Yes	Yes	Yes
IV			Yes	Yes	Yes	Yes	Yes
IV First stage F-stat			2667	4956	4650	4652	4518
Scheme trends				Yes	Yes	Yes	Yes
Controls					Yes	Yes	Yes

Notes: Standard errors in brackets (clustered at the ward level). IV first stage is the first stage Kleibergen-Paap F-stat for the IV specifications. *, **, *** indicate significant at the 10%, 5% and 1% levels, respectively. Controls include a linear trend interacted with the distance to closest improvement and a trend interacted with year of opening. Observations refers to maximum number of ward x year cells. Numbers vary slightly by sector.

Table 4. Ward level plant count regressions: all sectors

ALL SECTORS	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log accessibility	0.332***	0.105***	0.379***	0.283***	0.288***	0.335***	0.262***
	(0.015)	(0.034)	(0.093)	(0.088)	(0.088)	(0.089)	(0.096)
Observations	38269	38269	38269	38269	38269	53834	16577
Wards	3479	3479	3479	3479	3479	4894	1507
Distance band	1-20 km	1-20 km	1-20 km	1-20 km	1-20 km	1-30 km	1-10 km
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ward FE		Yes	Yes	Yes	Yes	Yes	Yes
IV			Yes	Yes	Yes	Yes	Yes
IV First stage F-stat			2667	4956	4650	4652	4518
Scheme trends				Yes	Yes	Yes	Yes
Controls					Yes	Yes	Yes

Notes: Standard errors in brackets (clustered at the ward level). IV first stage is the first stage Kleibergen-Paap F-stat for the IV specifications. *, **, *** indicate significant at the 10%, 5% and 1% levels, respectively. Controls include a linear trend interacted with the distance to closest improvement and a trend interacted with year of opening. Reduced form estimate in preferred specification (5) is 0.318*** (0.098). First stage F stat >3000.

Table 5: Ward level plant count regressions: by sector.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
MANUFACTURING							
Log accessibility	0.325***	0.358***	0.935***	0.580**	0.620**	0.675***	0.638**
	(0.019)	(0.097)	(0.236)	(0.247)	(0.248)	(0.245)	(0.292)
CONSTRUCTION							
Log accessibility	0.171***	0.188***	0.299*	0.131	0.211	0.258	0.01
	(0.016)	(0.072)	(0.174)	(0.181)	(0.184)	(0.185)	(0.213)
CONSUMER SERV							
Log accessibility	0.374***	0.069	0.076	-0.079	-0.088	-0.045	-0.029
	(0.017)	(0.052)	(0.149)	(0.144)	(0.148)	(0.145)	(0.174)
PRODUCER SERV							
Log accessibility	0.596***	0.252***	1.179***	0.682***	0.633***	0.670***	0.765***
	(0.020)	(0.088)	(0.198)	(0.183)	(0.183)	(0.183)	(0.222)
LAND TRANSPORT							
Log accessibility	0.080***	0.204*	0.874**	0.954**	0.929**	1.004***	0.529
	(0.014)	(0.109)	(0.342)	(0.371)	(0.373)	(0.368)	(0.438)
OTHER							
Log accessibility	0.134***	0.073*	0.306**	0.429***	0.433***	0.486***	0.454**
	(0.016)	(0.044)	(0.141)	(0.145)	(0.144)	(0.143)	(0.178)
Observations	38268	38268	38268	38268	38268	53831	16577
Wards	3479	3479	3479	3479	3479	4894	1507
Distance band	1-20 km	1-20 km	1-20 km	1-20 km	1-20 km	1-30 km	1-10 km
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ward FE		Yes	Yes	Yes	Yes	Yes	Yes
IV			Yes	Yes	Yes	Yes	Yes
IV First Stage F-Stat			2667	4956	4650	4650	4518
Scheme trends				Yes	Yes	Yes	Yes
Controls					Yes	Yes	Yes

Notes: Standard errors in brackets (clustered at the ward level). IV first stage is the first stage Kleibergen-Paap F-stat for the IV specifications. *, **, *** indicate significant at the 10%, 5% and 1% levels, respectively. Controls include a linear trend interacted with the distance to closest improvement and a trend interacted with year of opening. Observations refers to maximum number of ward x year cells. Numbers vary slightly by sector.

Table 6: Robustness of ward employment and local unit count results to distance decay.
20 km radius

ALL SECTORS	(1)	(2)	(3)	(4)	(5)
Economic size	Addresses	LU counts	Employment	Employment	Employment
Cost function	$Distance^{-1}$	$Distance^{-1}$	$Distance^{-0.5}$	$Distance^{-1.5}$	$e^{-0.2*Distance}$
Employment					
Log accessibility	0.564*	0.396*	0.711**	0.188	1.558**
	(0.305)	(0.215)	(0.322)	(0.132)	(0.692)
Plant count					
Log accessibility	0.475***	0.354***	0.534***	0.182***	1.157***
	(0.153)	(0.107)	(0.163)	(0.070)	(0.355)
Observations	38247	38247	38247	38247	38247
Wards	3477	3477	3477	3477	3477
Year FE	Yes	Yes	Yes	Yes	Yes
Distance band	1-20 km	1-20 km	1-20 km	1-20 km	1-20 km
Ward FE	Yes	Yes	Yes	Yes	Yes
IV	Yes	Yes	Yes	Yes	Yes
IV First stage F-stat	197	3762	3994	628	834
Scheme trends	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors in brackets (clustered at the ward level). IV first stage is the first stage Kleibergen-Paap F-stat for the IV specifications. *, **, *** indicate significant at the 10%, 5% and 1% levels, respectively. Controls include a linear trend interacted with the distance to closest improvement and a trend interacted with year of opening.

Table 7: Ward employment and plant count: within bands rings from scheme.
IV estimates

ALL SECTORS	(1)	(2)	(3)
Distance band	1-10 km	10-20 km	20-30 km
Employment			
Log accessibility	0.199	0.986***	1.14
	(0.207)	(0.326)	(1.043)
Plant count			
Log accessibility	0.262***	0.338**	1.040**
	(0.096)	(0.170)	(0.518)
Observations	16566	21681	15576
Wards	1506	1971	1416
Year FE	Yes	Yes	Yes
Ward FE	Yes	Yes	Yes
IV First stage F-stat	4522	1139	89
Scheme trends	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: Standard errors in brackets (clustered at the ward level). IV first stage is the first stage Kleibergen-Paap F-stat for the IV specifications. *, **, *** indicate significant at the 10%, 5% and 1% levels, respectively. Observations reported for employment regressions; number of observations differs slightly in plant count regressions. Controls include a linear trend interacted with the distance to closest improvement and a trend interacted with year of opening.

Table 8: Plant level employment: all sectors. 20 km radius

ALL SECTORS	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log accessibility	0.069***	0.068***	0.095	-0.009	-0.048	-0.062	-0.228*
	(0.015)	(0.026)	(0.094)	(0.088)	(0.089)	(0.090)	(0.118)
Observations	2065343	2064780	2064780	2064780	2064780	2737108	977378
Ward clusters	3487	3487	3487	3487	3487	4903	1514
Distance band	1-20 km	1-20 km	1-20 km	1-20 km	1-20 km	1-30 km	1-10 km
Year-Sic FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ward FE		Yes	Yes	Yes	Yes	Yes	Yes
IV			Yes	Yes	Yes	Yes	Yes
IV First Stage F-Stat			658	1087	1054	1114	1813
Scheme trends				Yes	Yes	Yes	Yes
Controls					Yes	Yes	Yes

Notes: Standard errors in brackets (clustered at the ward level). IV first stage is the first stage Kleibergen-Paap F-stat for the IV specifications. *, **, *** indicate significant at the 10%, 5% and 1% levels, respectively. Regressions include sic-year dummies, a dummy for singleton plants. The sample excludes the plants situated 1 km or closer to the improvements and the plants which employment is on the top 0.05%. Controls include a trend for the distance to the improvement and a trend for the year of opening of the closest improvements.

Table 9: Plant level economic outputs: all sectors. 20 km radius

ALL SECTORS	(1)	(2)	(3)	(4)	(5)
GVA per worker					
Log of accessibility	0.076***	0.088***	0.438***	0.515***	0.497***
	(0.011)	(0.033)	(0.127)	(0.133)	(0.133)
Total labour costs per worker					
	0.080***	0.031	0.244***	0.238***	0.242***
	(0.010)	(0.024)	(0.077)	(0.083)	(0.084)
Total revenue					
	0.178***	0.044	0.125	0.426**	0.380*
	(0.028)	(0.042)	(0.184)	(0.196)	(0.198)
Total value-added					
	0.189***	0.121**	0.426**	0.516**	0.460**
	(0.027)	(0.051)	(0.211)	(0.223)	(0.226)
Observations	824,980	687,877	687,877	687,877	687,877
Wards	3487	3473	3473	3473	3473
Distance band	1-20 km	1-20 km	1-20 km	1-20 km	1-20 km
Year-Sic FE	Yes	Yes	Yes	Yes	Yes
Ward FE		Yes	Yes	Yes	Yes
IV			Yes	Yes	Yes
IV First Stage F-stat			339	594	583
Scheme trends				Yes	Yes
Controls					Yes

Notes: Standard errors in brackets (clustered at the ward level). IV first stage is the first stage Kleibergen-Paap F-stat for the IV specifications. *, **, *** indicate significant at the 10%, 5% and 1% levels, respectively. Observations reports maximum number of plant x year observations. Controls include a linear trend interacted with the distance to closest improvement and a trend interacted with year of opening.

Table 10: Ward level economic outputs: all sectors. 20 km radius

ALL SECTORS	(1)	(2)	(3)	(4)	(5)
Real gross value added pw					
Log accessibility	0.062***	0.014	0.122	-0.093	0.023
	(0.011)	(0.128)	(0.437)	(0.475)	(0.470)
Real gross value added					
Log accessibility	0.398***	0.372	0.273	0.776	0.443
	(0.031)	(0.244)	(0.860)	(0.927)	(0.919)
Observations	34380	34378	34378	34378	34378
Wards	3487	3485	3485	3485	3485
Distance band	1-20 km	1-20 km	1-20 km	1-20 km	1-20 km
Year FE	Yes	Yes	Yes	Yes	Yes
Ward FE		Yes	Yes	Yes	Yes
IV			Yes	Yes	Yes
First stage F-stat			2582	4199	4014
Scheme trends				Yes	Yes
Controls					Yes

Notes: Standard errors in brackets (clustered at the ward level). IV first stage is the first stage Kleibergen-Paap F-stat for the IV specifications. *, **, *** indicate significant at the 10%, 5% and 1% levels, respectively. Controls include a linear trend interacted with the distance to closest improvement and a trend interacted with year of opening.

10 Appendix Tables

Table A1: Major road schemes in Britain 1998-2007

Opening year	Type	Road	Scheme	Length in km
1998	Faster route	A16	A16 Market Deeping/Deeping St James Bypass	1.6
1998	New route	A34	A34 Newbury Bypass	9.3
1998	Faster route	A50	A50/A564 Stoke - Derby Link (DBFO)	5.1
1999	New route	A12	A12 Hackney Wick - M11 Contracts I-IV	4.7
1999	Faster route	A35	A30/A35 Puddleton Bypass (DBFO)	9.3
1999	New route	M1	M1/M62 Link Roads	16
1999	Faster route	M74	A74(M). Paddy's Rickle - to St Ann's (J16)	11.6
2000	New route	M60	M66 Denton - Middleton Contract I	15.3
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	4.5
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
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	18.3
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 length of new links 1998-2007	318.03

Figure A1: Location of major road schemes in Britain 1998-2007 showing the overall road major network

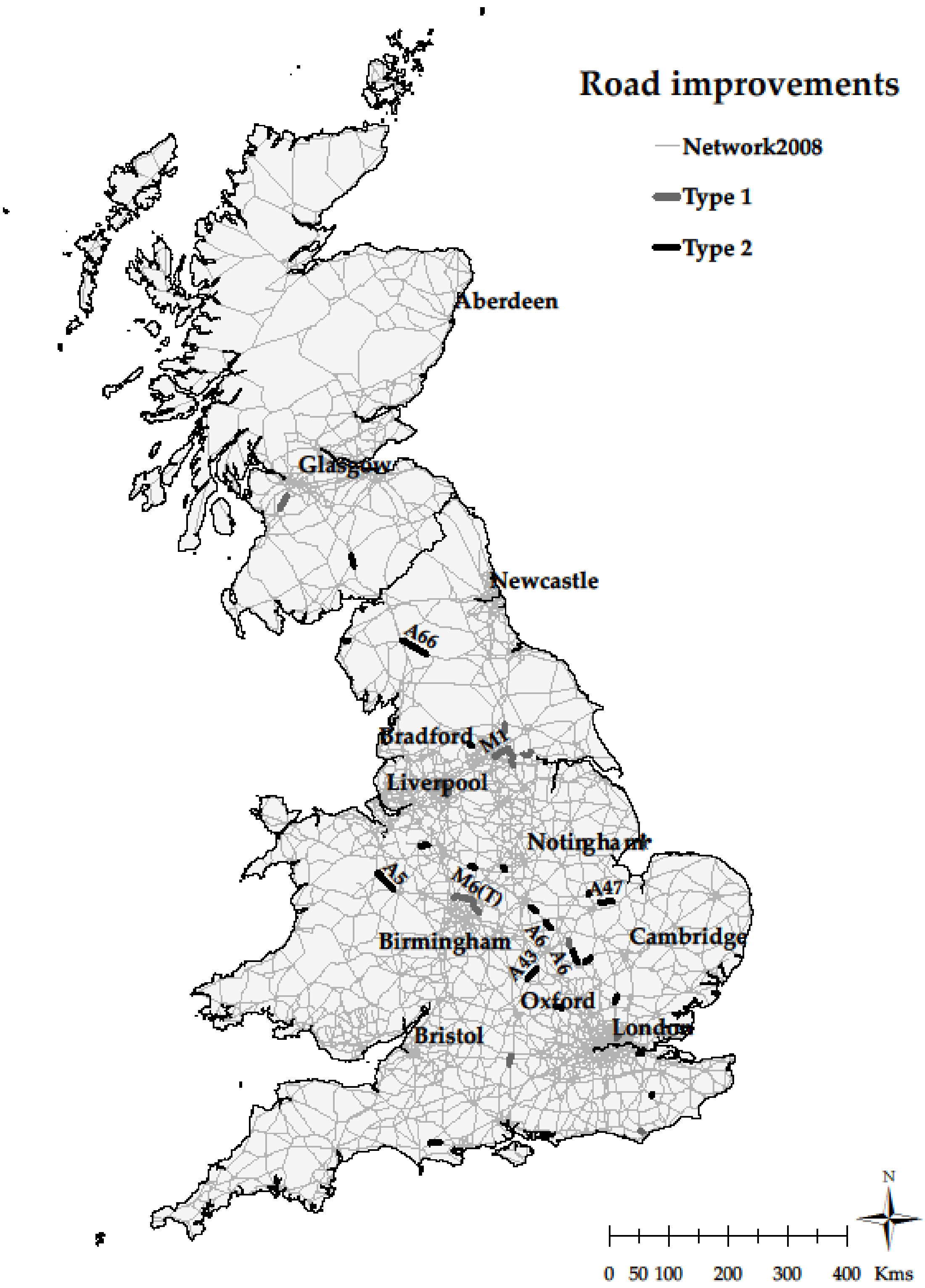


Table A2: Balancing tests, initial conditions versus distance within 1-20 band

Initial (1997)	Log employment	Log number of plants	Log no of residential addresses			
Distance to scheme	-0.001	0.001	0.003			
(km)	((0.003)	((0.002)	((0.002)			
Scheme dummies	Yes	Yes	Yes			
Observations	3487	3487	3487			
Ward clusters	3487	3487	3487			
R2	0.171	0.226	0.236			
Initial (1998)	Log GVA	Log GVA per worker	Log labour costs	Labour costs per worker	Log gross output	Log gross output per worker
Distance to scheme	0.009	-0.002	0.011*	0.000	0.012**	0.002
(km)	(0.006)	(0.002)	(0.006)	(0.002)	(0.006)	(0.002)
Scheme dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3428	3428	3418	3418	3419	3419
Ward clusters	3428	3428	3418	3418	3419	3419
R2	0.0525	0.0136	0.06	0.0206	0.0429	0.00535

Notes: Standard errors clustered at ward level. 1-20 km distance band.

Table A3: Employment and number of plants in wards

Distance band	10km (16,654 obs)		20km (38,357 obs)		30km (53,933 obs)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Accessibility						
Log accessibility, 1997 empl	15.434	0.893	15.390	0.845	15.329	0.855
Log accessibility, time-varying empl	15.421	0.878	15.379	0.836	15.321	0.848
Employment						
ALL SECTORS	3629.79	9244.31	3095.69	7112.44	2852.41	6295.96
MANUFACTURING	475.28	997.57	434.72	915.21	399.62	851.64
CONSTRUCTION	159.46	320.12	150.43	322.57	141.39	290.89
CONSUMER SERVICES	777.97	2009.28	699.03	1588.20	650.91	1426.67
PRODUCER SERVICES	991.72	4379.83	749.63	3197.79	668.04	2764.43
TRANSPORT	1148.79	2745.60	995.19	2196.40	932.26	2035.23
OTHER	76.57	261.11	66.70	224.66	60.19	198.61
Local units						
ALL SECTORS	311.12	607.02	280.40	446.33	264.09	395.72
MANUFACTURING	25.06	41.37	22.12	33.48	20.42	30.03
CONSTRUCTION	23.70	17.99	23.64	17.33	23.15	17.49
CONSUMER SERVICES	82.82	137.18	75.14	109.46	70.68	98.96
PRODUCER SERVICES	103.18	334.06	89.67	236.95	83.42	211.62
TRANSPORT	70.37	118.73	64.34	87.51	61.13	77.11
OTHER	5.99	24.18	5.50	16.31	5.29	14.02

Notes: Own calculation using BSD and optimal travel times calculated as described in the text

Table A4: Spatial autocorrelation testing for ward level regressions

	ALL	MANUF	CONS	C SERV	P SERV	TRANS	OTHER
Errors log employment							
Spatial lagged residuals	-0.002	0.001	0.002	0.001	-0.001	-0.003***	0.000
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)
Observations	3477	3477	3477	3477	3477	3477	3477
Errors log no of plants							
Spatial lagged residuals	-0.003	-0.006***	-0.001	0.000	-0.005*	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.001)	(0.002)
Observations	3479	3479	3479	3479	3479	3479	3479
Distance band	1-20kms	1-20kms	1-20kms	1-20kms	1-20kms	1-20kms	1-20kms

Notes: Table shows results from a regression of spatially weighted residuals on own residuals for wards within the 1-20km band. In order to allow a sufficient number of neighbours when constructing the spatially weighted residuals, we run regressions using the same specification as column 5 in tables 2-3 and 4-5, but for wards located in a 1-50kms distance band. We then calculate spatially weighted residuals using these results and the spatial weights as in the accessibility measures (inverse travel time with decay 1).

Table A5: Main ward results with district clustering level, employment and plant counts

	ALL	MANUF	CONS	C SERV	P SERV	TRANS	OTHER
Log employment							
Log of accessibility	0.361**	-0.003	0.221	-0.276	0.878***	1.06	0.617**
	(0.173)	(0.418)	(0.295)	(0.290)	(0.276)	(0.647)	(0.259)
Observations	38247	37625	38184	38246	38219	35198	38246
widstat	915	937	924	914	914	1092	915
N_clust	209	209	209	209	209	209	209
Log no plants							
Log of accessibility	0.288**	0.620**	0.211	-0.088	0.633***	0.929***	0.433*
	(0.112)	(0.251)	(0.225)	(0.128)	(0.178)	(0.276)	(0.234)
Observations	38269	37655	38209	38268	38242	35243	38268
widstat	912	934	922	912	911	1090	912
N_clust	209	209	209	209	209	209	209
Distance band	1-20kms	1-20kms	1-20kms	1-20kms	1-20kms	1-20kms	1-20kms
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ward FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IV	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Scheme trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table reproduces the results of column 5 in tables 2-3 and 4-5 but with standard errors clustered at district level.

Table A6: Plant level employment regression estimates, by sectors

	MANUFACTURING					*CONSTRUCTION*					*CONSUMER SERV*				
Log of accessibility	-0.101***	0.141**	-0.138	0.002	-0.013	0.069***	0.144	0.183	0.243	0.279	0.069***	0.068**	-0.033	-0.04	-0.064
	(0.019)	(0.069)	(0.258)	(0.262)	(0.260)	(0.017)	(0.089)	(0.294)	(0.303)	(0.308)	(0.015)	(0.030)	(0.108)	(0.110)	(0.111)
Observations	220876	220338	220338	220338	220338	123752	123066	123066	123066	123066	848761	847193	847193	847193	847193
First Stage F-stat			1190	2686	2530			1500	3914	3549			431	702	661
Ward clusters	3112	3104	3104	3104	3104	3319	3305	3305	3305	3305	3477	3477	3477	3477	3477
	PRODUCER SERV					*LAND TRANSPORT*					*OTHER*				
Log of accessibility	0.132***	0.019	0.706**	0.221	0.087	0.040**	0.049	-0.129	-0.209	-0.24	0.079**	0.012	-0.085	-0.283	-0.326
	(0.029)	(0.073)	(0.278)	(0.266)	(0.265)	(0.019)	(0.045)	(0.170)	(0.177)	(0.177)	(0.032)	(0.163)	(0.584)	(0.587)	(0.590)
Observations	456766	455095	455095	455095	455095	377565	374140	374140	374140	374140	37623	37539	37539	37539	37539
First Stage F-stat			444	636	677			755	1262	1227			1039	1827	1823
Ward clusters	3429	3429	3429	3429	3429	3449	3442	3442	3442	3442	2220	2211	2211	2211	2211
Distance band	1-20km	1-20km	1-20km	1-20km	1-20km	1-20km	1-20km	1-20km	1-20km	1-20km	1-20km	1-20km	1-20km	1-20km	1-20km
Year-Sic FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ward FE		Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes
IV			Yes	Yes	Yes			Yes	Yes	Yes			Yes	Yes	Yes
Scheme trends				Yes	Yes				Yes	Yes				Yes	Yes
Controls					Yes					Yes					Yes

Notes: As Table 8.

Table A7: Plant level economic output regression estimates: value-added per worker by sectors

	MANUFACTURING					*CONSTRUCTION*					*CONSUMER SERV*				
Log of accessibility	0.046***	0.437***	-0.151	-0.203	-0.247	0.047***	-0.297*	0.33	-0.925*	-0.748	0.053***	-0.005	0.726***	0.602***	0.583***
	(0.012)	(0.120)	(0.384)	(0.388)	(0.390)	(0.014)	(0.174)	(0.515)	(0.560)	(0.564)	(0.008)	(0.042)	(0.185)	(0.186)	(0.185)
Observations	72480	57903	57903	57903	57903	32564	19809	19809	19809	19809	397078	339615	339615	339615	339615
First Stage F-stat			731	2092	1884			599	921	871			255	433	416
Ward clusters	3015	2445	2445	2445	2445	3250	2046	2046	2046	2046	3475	3346	3346	3346	3346
	PRODUCER SERV					*LAND TRANSPORT*					*OTHER*				
Log of accessibility	0.125***	0.132	-0.55	0.794	0.717	0.089***	0.197***	0.196	0.371	0.372	0.064***	-0.112	1.592	0.505	1.131
	(0.023)	(0.121)	(0.381)	(0.512)	(0.511)	(0.015)	(0.073)	(0.278)	(0.272)	(0.273)	(0.014)	(0.271)	(1.457)	(1.520)	(1.577)
Observations	158563	119239	119239	119239	119239	149290	125037	125037	125037	125037	15005	12277	12277	12277	12277
First Stage F-stat			286	436	485			345	652	625			406	850	1212
Ward clusters	3415	3066	3066	3066	3066	3421	3201	3201	3201	3201	2118	1322	1322	1322	1322
Distance band	1-20km	1-20km	1-20km	1-20km	1-20km	1-20km	1-20km	1-20km	1-20km	1-20km	1-20km	1-20km	1-20km	1-20km	1-20km
Year-Sic FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ward FE		Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes
IV			Yes	Yes	Yes			Yes	Yes	Yes			Yes	Yes	Yes
Scheme trends				Yes	Yes				Yes	Yes				Yes	Yes
Controls					Yes					Yes					Yes

Notes: As Table 9.

Table A8: Firm GVA pw regressions by plant size (average) 6 categories

	Average Employment					
	1 to 5	6 to 10	11 to 25	26 to 50	51 to 150	Over 150
ALL SECTORS						
Log of accessibility	-0.192	1.069***	0.774***	0.752**	0.07	0.45
	(0.290)	(0.335)	(0.273)	(0.294)	(0.294)	(0.455)
Observations	129697	130145	165278	95727	98957	56793
First stage F-stat	590	438	680	320	548	856
Ward clusters	3283	3017	3255	2899	2767	2109
Distance band	1-20km	1-20km	1-20km	1-20km	1-20km	1-20km
Year-Sic FE	Yes	Yes	Yes	Yes	Yes	Yes
Ward FE	Yes	Yes	Yes	Yes	Yes	Yes
IV	Yes	Yes	Yes	Yes	Yes	Yes
Scheme trends	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: These results are obtained using the specification of column 5 in table 9 but restricting the sample to the plants in the different size categories (defined by average employment during the period).

BIS Department for Business Innovation & Skills

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Llywodraeth Cynulliad Cymru
Welsh Assembly Government

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.