Road transport: the effects on firms

Stephen Gibbons[ab]

Teemu Lyytikäinen[ac]

Henry Overman[a]

Rosa Sanchis-Guarner[ad]

PRELIMINARY DRAFT

June 2012

[a] London School of Economics & Spatial Economics Research Centre, London, United Kingdom.

[b] Corresponding author: Stephen Gibbons (email: S.Gibbons@lse.ac.uk).

[c] The author wants to thank the Yrjö Jahnsson Foundation for financial support.

[d] The author acknowledges funding support from the Bank of Spain.

Abstract: This study estimates the impact of road transport infrastructure improvements on firms. Firms' exposure to transport improvements is measured by changes in employment accessibility (or effective density) along the road network. We estimate the effect of these changes on employment and productivity using plant level micro longitudinal data on firms in Britain, linked by detailed geographical location (10,500 wards) to the British road network and major improvements in it between 1998 and 2008. Estimates are based on an instrumental variables strategy that exploits localised accessibility changes due to road schemes. We find a substantial response in employment and numbers of plants at a small-scale geographical level of aggregation (electoral wards), but no response in employment at plant level, suggesting that road improvements encourage firm entry or discourage exit but do not effect existing firms. We also find effects on labour productivity and wages at the firm level, although the results are more mixed and sensitive to specification.

Key words: Productivity, employment, accessibility, transport

JEL codes: D24, O18, R12

This study has been presented at the Urban Economic Association Meetings (Denver, November 2010) and at the SERC Work in progress seminar in December 2010. We thanks the participants for their comments. All the errors and opinions of this paper remain our own.

1. Introduction

The road network is a hugely important part of infrastructure in most countries. In the UK, for example, more than 90% of passenger transport and around 65% of goods transport is done by road.[1] Understanding the relationship between transport improvements and economic outcomes is essential to the design of transport policy, and given the importance of road transportation for the movements of people and goods, the evaluation of the impact of road investments is also important for economic policy as a whole. The role of transportation in the spatial distribution of the economic activity and economic performance has become of increased interest to researchers in the last years. Decreasing transport costs is considered to be a central driver of economic integration and the emergence of agglomeration externalities, but solid empirical evidence on the channels through which these effects operate is still needed.

In this paper, we investigate the causal impact of road improvements on employment and productivity using micro data for Great Britain between 1998-2008. We use road network data and data on major road improvements to construct a measure of accessibility to employment through the road network in these years. We test the impact of changes in accessibility on firm level outcomes (employment and productivity) and on aggregate outcomes (ward employment, number of plants and aggregated productivity) using a novel instrumental variables strategy that exploits small scale spatial variation in the way road improvements change accessibility to employment.

The theoretical predictions on the net effects of transportation improvements on firms and local outcomes are ambiguous. Better transport infrastructure brings places and people closer together. This has two effects on the actual size of the markets. Firstly, for a given location of firms and workers, effective density increases, as it becomes easier to reach other locations using the improved transportation network. Secondly, new infrastructure increases the attractiveness of locations, which may boost spatial concentration if firms and workers relocate. These effects may reinforce each other and create positive agglomeration spillovers (Ottaviano, 2008). On the other side, improved access to

[1] Transport Statistics Great Britain.

markets also strengthens competition, forcing the exit of the less productive firms and thus increasing aggregate productivity (Melitz, 2003). Finally, firms use transport services as a production input, so changes in the supply and relative prices of transport affect the input mix used by the firms and their demand of other inputs, for example labour.

Gibbons and Overman (2009) provide an extensive discussion of the potential productivity and scale effects of transport infrastructure. At the firm level, transport improvements could affect the performance of firms. On the one hand, they may improve the logistics and the internal organisation of the firms, and can change the optimal input mix choice. Transportation services are used as production input and, if there is a substitution effect between inventories, labour and transport services, the demand and input mix will be affected (Holl, 2006). Input prices could decrease because of reduced transport costs or increased competition between the suppliers. Wages could also change if productivity effects are capitalised into wages or if wages are set as a function of commuting costs, which are affected by the transport network (Gibbons and Machin, 2006). Therefore, firms might change the demand of inputs, and depending on the internal returns to scale, this would affect its final output. If output increases with respect to inputs more than proportionally (due to increasing returns to scale), the output/input ratio will change, but this would be a scale effect and total factor productivity would be unaffected. Furthermore, better accessibility to consumers increases customer base and allows firms to expand production and exploit economies of scale.

In addition to the potential scale effects discussed above, firms total factor productivity could be affected by the *wider economic benefits of transport* (Graham, 2007). These refer to agglomeration externalities (sharing, matching and learning – Duranton and Puga, 2004), which can be internal to the industry (localisation economies) or just a consequence of an increase in the size of the markets (urbanisation economies). Firms benefit from the presence of other firms nearby (in the same or different sector of production) and from the increased proximity to suppliers which arises from the improvements on the transport network. Agglomeration benefits act like a production function shifter, i.e. for a given amount of inputs, the firm is able to produce more.

At an aggregate level, employment and productivity may be affected through firms entering and exiting the ward. Entry and exit rates could be affected as firms relocate to better benefit from scale effects and agglomeration externalities. This process may reinforce the scope for scale and agglomeration benefits further. For example, aggregate employment could increase because firms input demands increase (scale effects) and because the new firms move into the area. On the other hand, reduced transport costs can also increase spatial competition. If only the "fittest" firms survive (Melitz, 2003), then the number of firms and aggregate productivity level will change.

Given that the theoretical predictions are ambiguous, the effect of transport improvements on the firm level and aggregate level outcomes remains an empirical question. Although authors have explicitly included the role of transportation into spatial economics analysis (Combes and Lafoucarde, 2001; Puga, 2002; Behrens et at, 2004; Venables, 2007), there is still need to empirically establish the causal link between transportation infrastructure in spatial economic performance. 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). This literature has focused the impacts of investment in roads and public infrastructure on several economic outcomes, such as aggregate productivity, growth or employment, finding mixed results (Gramlich, 1994; Martin and Rogers, 1995; Boarnet, 1998; Chandra and Thompson, 2000; Jiwattanakulpaisarn et al, 2010).

Some recent papers have estimated, using careful identification strategies, the effect of roads on other outcomes in the US: urban sprawl (Burchfield et al, 2006), urban growth (Duranton and Turner, 2011), road traffic (Duranton and Turner, 2010), sub-urbanisation (Baum-Snow, 2007), commuting patterns (Baum-Snow, 2010) or demand for skills (Michaels, 2008). In these papers the effect of transport is usually capture by connectivity to the network (either connected or not) or by some measure of the density of the network within some geographic boundaries and the focus is on correct identification of (long-run) effects. Other papers (Faber, 2009; Donaldson, 2010) have focused on developing countries (highways in China and railroad in colonial India) to study the effect of the reduction of transport costs due to transport networks development on trade integration and the consequent economic development.

Only a handful of papers have studied the effect of increased accessibility on firms' outcomes, and they have mostly focused on the analysis of firm relocation (Coughlin and Segev, 2000; Holl, 2004a and 2004c) or firm birth (Holl, 2004b, Melo et al, 2010), all finding positive relationships between the presence of roads and firms' relocation and creation. Holl (2011) studies the relationship between market access and firm productivity when market access changes due to road investments and changes in population. She exploits data for a panel of firms during a period of intense road construction in Spain. When using plant fixed-effects the estimates are imprecise, so she relies GMM techniques in order to overcome endogeneity problems, with which she finds positive significant effects of markets access on productivity. Li and Li (2010) use the construction of the Chinese highways system to evaluate the impact of improved transport infrastructure on the amount of inventories held by firms, arguing that the reduced inventories due to road construction improve efficiency and aggregate productivity.

In this study we take advantage of the availability of micro data on firms and of road network data at a very detailed geographical level to overcome endogeneity problems. We construct a data set of road improvements carried out in Great Britain in different points in time during the period 1998-2008. We combine this data with road network data and use GIS network analysis tools to calculate minimum travel times by road between any location for every year in our sample. We use this information, jointly with information with the exact location of plant employment, to calculate accessibility to employment in different years. Accessibility to employment measures the amount of employment which is reachable using the road network from a given location, inversely weighed by the travel time to reach the other locations. One advantage of using this measure is that it is not constrained to artificial geographical boundaries like measures based on the density of roads. Moreover, it allows us to use variation due to transport improvements which affect optimal travel times, even if the location was previously connected to the network. This measure is, thus, very appropriate in a setting like Great Britain where road density was already high at the beginning of our period of study.

The geographical unit we use is the electoral ward. Wards are quite small and there are over 10,500 in Great Britain. The fine geographical detail of our data allows us also to implement a careful identification strategy. Having such a large number of units implies that we have variation in accessibility changes due to road improvements even when we focus on areas close to where the road improvement took place.

There are three endogeneity concerns which may challenge the validity of the estimated causal relationships. Firstly, cross-sectional estimates of the effect of accessibility on economic outcomes will be biased if the model does not capture underlying factors (such as place specific productive advantages) that affects both effective density and economic outcomes. We use a fixed-effects estimation method to address this problem. In the fixed effects framework, changes in accessibility can arise because of road improvements and because employment relocates. The second endogeneity issue is that accessibility changes due to relocation of employment may be partly driven by the outcome variable studied or be correlated with the same unobserved shocks. We address this source of bias, we constrict an instrument which uses accessibility changes stemming only from the transport improvements. Finally, in order to reduce the possible bias caused by the endogeneity of the placement of the transport investments, we will focus of firms which are located within 20 kilometres of road schemes. This way, we compare firms and locations which are close to the improvements and we exploit the fact that the impact of the improvements varies considerably even within the distance band. It is quite unlikely that the improvements are aimed at specific firms or wards within those narrowly defined distance bands.

We estimate the effect of road transport improvements at a small geographically aggregated scale (electoral ward) and at the plant level. The geographically aggregated regressions allow for firm entry and exit while the plant level regressions compare firms in locations with different accessibility changes. From these results, we find strong evidence of effects from road transport improvements on local area level employment and on the number of plants. However, no employment effects are found at plant level, implying that the employment changes at ward level come about through firm entry and exit. Conversely, at plant level, we do find evidence of positive impacts on labour productivity, value-added and on wages (labour costs per worker) which we do not find at the geographically aggregate level. A potential explanation for this is that it is smaller firms which experience productivity impacts, but these small firms contribute relatively little to local aggregate productivity.

The rest of the paper is structured as follows. Section 2 presents the empirical methodology and explains the construction of the accessibility, productivity and employment measures. Section 3 describes the data used, and in Section 4 the empirical results are discussed. Finally, in Section 5 we conclude.

2. Empirical methods

As outlined above, our empirical analysis is concerned with the estimation of the effect of road transport improvements on employment and productivity. The intensity of exposure to road transport improvements is measured in our empirical work by changes in the accessibility to employment (or other measures of economic mass) along the road network. We adopt two approaches. The first is to estimate aggregate effects for small spatial units (electoral wards). These estimates take into account firm exit, entry and geographical relocation. In the second approach, we estimate firm level effects for existing firms. Both approaches use the same general structure in the estimation, which is based on a panel of units (wards or plants) observed over a period of up to 11 years, from 1998-2008. The data sources are described in Section 3 below.

2.1. General empirical set up

The underlying empirical model for all our analyses is:

$$y_{jt} = \beta A_{jt} + \mu_j + \tau_t + \varepsilon_{jt} \qquad (1)$$

Here y_{jt} is the outcome variable in panel unit j in year t, A is a measure of employment accessibility along the transport network (see Section 2.2 for details). In general, y_{jt} and A_{jt} are in natural logarithms. Parameter β is the effect (elasticity) of accessibility via the road network on the outcome variable. The unobserved component μ is a time invariant unit specific factor. Year effects τ represent general changes that influence all locations in the study area in a given year (e.g. macro shocks). 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 an accessibility change on the outcome of interest. We use OLS as our first method of estimation. Traditional estimates of the effects of accessibility are often based on ordinary least squares (OLS) estimates of models like (1). The OLS estimates 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 evolved in places with productive advantages.

A first step to eliminating these biases is to eliminate fixed-over-time ward effects μ by time-demeaning the data within wards (so called within transformation). The within transformation is obtained by first averaging the equation (1) over time and then subtracting the ward averages from equation (1).

$$y_{jt} - \bar{y}_j = \beta\left(A_{jt} - \bar{A}_j\right) + \tau_t + u_{jt} \qquad (2)$$

Here $\bar{y}_j$ and $\bar{A}_j$ denote ward averages. This way the ward fixed effects disappear and the estimates are robust to time-invariant ward heterogeneity that can be arbitrarily correlated with accessibility. This formulation is a starting point for evaluating the effects of transport policy on firms, because transport improvements generate changes in A over time.

We prefer the within transformation to first differencing (which would also eliminate the panel unit fixed effects) because first differencing assumes instantaneous responses to accessibility changes while within estimation allows for a more flexible time pattern for the effect. For example, if we have ten years of data and accessibility in a ward changes in the fifth year, the within estimator will be based on the comparison of the value of y in years 1-5 to year 6-10 whereas the first differenced estimator compares changes between years 5 and 6 to changes in other years. If the response to the accessibility change takes longer than a year, the first differenced estimates will be biased downwards. The within estimator is better suited for capturing slow and gradual changes.

In general, the variation in accessibility over time within wards could come through changes in the spatial distribution of employment, or because of changes in the transport network. Changes in accessibility due to the relocation of employment across space may be directly affected by the outcome variable or correlated with unobserved shocks in the error term of (2), which may lead to bias in the estimation. To address the issue of endogenous determination of accessibility, we instrument accessibility index A (that varies within wards due to changes in both employment and transport network) with a measure, denoted by $\hat{A}$, which only picks up changes in the transport network. We calculate the accessibility based on the pre-improvement spatial distribution of employment (year 1997). We estimate (2) by two-stage least squares using this as an instrument for actual changes in accessibility. The construction of A and the instrument $\hat{A}$ are described in Section 3.3 below.

The instrumental variable (IV) estimates from (2) will produce biased estimates of the productivity effects of transport improvements, if areas with increasing or declining employment trends are those that experience the greatest accessibility changes due to road improvements. This implies that u_{jt} in Equation (2) is correlated with the instrument. The usual reason to suspect this kind of problem is the possibility that transport policy is endogenous to the employment and productivity trends in the targeted locations, i.e. the decision to improve the transport network might be partly driven by productivity trends.

We address this potential source of bias by focusing the empirical analysis on places and firms that are close to the transport improvement sites. In the results section below we present estimates for samples within 20km of the sites of improvement (10km and 30km used in robustness checks). In this way we are comparing closely neighbouring places that differ incrementally in terms of the changes in accessibility they experience as a result of the road network improvements. We assume that these differences in changes in accessibility close to transport schemes are an incidental by-product of the scheme rather than its intended outcome. The main changes in mean travel times and employment accessibility occur close to the end points of new road schemes, although they are typically intended to improve the flow of traffic between cities or areas further away from the improvement.[2] There are also often long delays between commissioning and opening of road schemes, which will weaken any link between pre-existing local productivity trends and the decisions over where to site these projects.

In order to ensure that the instrument is uncorrelated with the underlying trends, we further control for differential trends in the vicinity of road schemes in our final specification. This is done by including a set of scheme dummies (31 schemes) interacted with year in equation (2). Lastly, we test for the robustness of the results to the inclusion of salient ward characteristics (straight line distance to closest road scheme employment rate, average age, proportion of population aged 16-74 with higher education and proportion of population living on social housing). The area characteristics are measured in one year and do not change over time in our data but we interact them with time, which implies that the trend is allowed to differ by ward characteristics. We acknowledge that this robustness check is very demanding and may pick up some of the effect of road improvements, especially if the impact of road schemes is gradual in nature.

There are some specific points to consider regarding estimation of (2) on plant level data. Firstly it worth noting that the plant identifiers in our data are location specific (they change if a plant closes down and moves to a different location). Thus, in our within-plant analysis (where the panel units in (2) are plants) the changes in accessibility $(A_{jt} - \bar{A}_j)$ are not related to relocations of plant *j*, but only changes in accessibility occurring at a fixed plant location, due either to restructuring of employment in other firms and/or transport improvements. As discussed above, we instrument accessibility index *A* with the accessibility index $\hat{A}$ defined in Section $$, which only picks up changes in the transport network. Estimation of (2) using within-plant changes in a panel of plants is only feasible using plants that exist, and appear in the data, both before and after the opening of the transport schemes that are used as the source of identifying variation in accessibility. This introduces sample selection issues. Firstly, firms that stay in the location of the transport scheme are likely to be those that can benefit

[2] See Department for Transport - A new deal for trunk roads in England: Understanding the new approach (1997) and The Highways Agency Framework Document (July 2009)

most from it. Secondly, the method does not capture changes in employment or productivity associated with the opening of new plants. In addition, there will be sampling-related reasons why some firms appear in our data in multiple years whilst others do not. These caveats aside, the IV estimation of β from the changes within plants over time provides guidance to the micro-level impacts of transport improvements for firms, which is one of the components of the aggregate ward level effects, and interesting in its own right.

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

2.2. Defining the accessibility index A and instrument $\hat{A}$

The accessibility index we used to measure used is identical in structure to market potential measures used in economic geography (e.g. Harris 1954), and to the accessibility indices used more generally in the transport literature (e.g. Ahmed et al 2006, Vickerman et al 1999). This index is a measure of the economic mass accessible to a firm in a particular location, given the local transport network. The index is similar to market potential measures used in the agglomeration economies literature.

Consider a measure of economic activity or other variable of interest, such as employment l. For a firm in an origin location j at time t, an employment accessibility index A_{jt} is a (log) 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 distance, minimum cost route along a transport network). That is, the index has the structure:

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

We use wards as the spatial units, travel time along the primary road network as the cost measure and employment as the measure of economic mass. Note, that changes in *A* are partly driven by changes in employment in destinations *k*. This may lead to endogeneity problems in the estimation of the effect of accessibility, if the employment changes near the origin are causally linked with employment changes in the origin or driven by the same unobserved factors. In the empirical work below, we instrument A_{jt} with accessibility calculated fixing employment at the level at the first in year in our data sample (1997). We calculate accessibility with initial employment as

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

Fixing employment to the 1997 level ensures that changes in the accessibility index (8) over time occur only as a result of changes in the costs c_{jkt} (e.g. travel time) and not changes in employment.

In Equations (7) and (8) the value of the weight *a*(.) attached to any destination k is a decreasing function of the cost of reaching destination *k* from origin *j*. Potential weighting schemes include: 'cumulative opportunities' weights $a(c_{jkt}) = 1$ if *k* is within a specified distance of *j*, zero otherwise; exponential weights $a(c_{jkt}) = \exp(\alpha c_{jkt})$; logistic weights $a(c_{jkt}) = (1+\exp(-\alpha c_{jkt}))^{-1}$ or inverse cost weights $a(c_{jkt}) = c_{jkt}^{-\alpha}$. See Graham, Gibbons and Martin (2009) for further discussion of these indices. 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 estimated travel time, although we present tests of robustness to alternative specifications.

2.3. Accessibility changes arising from transport improvements

Accessibility indicies A_{jt} and $\hat{A}_{jt}$ can be applied to the study of employment and productivity effects arising from accessibility by road when the costs c_{jkt} in (7) and (8) are calculated using routing along the transport network. This works because transport improvements change the structure of costs c_{jkt} along the transport network and the structure of costs along routes from *j* to potential destinations *k*. 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* in both the pre and post-improvement periods, such that the transport improvement reduces the cost of the journey along *p-q* and brings employment at destination *k* 'closer' to origin *j* in cost terms.
b) the least-cost route between *j* and *k* bypasses link *p-q* in the pre-improvement period, but switches to use the link *p-q* in the post-improvement period because of the reduction in costs; again this brings employment at destination *k* 'closer' to origin *j* in cost terms

There are also 'second order' effects arising when:

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

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

Changes in cost of all these types imply changes in the accessibility indices (i.e. a change in effective density). The amount of change in the accessibility index at a location *j* depends on the likelihood that a route between *j* and *k* uses the improved link *p-q*, and on economic mass in *k*. The idea in our method is to use the changes in the accessibility index at each location *j* to estimate the extent to which firms in location j are "potentially" affected. In turn, this change in potential accessibility enters into our regressions, as described in Sections 3.1 and 3.2 above.

3. Data setup and sample

3.4. Geographical units

All our analysis is based on micro data (plant level) sources. We have detailed geographical information on the location of the plants (postcodes) and can link data geographically at various levels using the ONS National Statistic Postcode Direcctory. A UK postcode unit corresponds to a small number of addresses (around 14) or a single large delivery such as a medium sized plant. As discussed in Section 2, for parts of the empirical analysis we work with aggregates at 'electoral ward' level. There are around 10,500 electoral wards represented in our data. Ward boundaries are set so as to include roughly the same size of the electorate. We use the ward as defined in 1998, the first year of our study. This unit is very small, especially in dense areas. For example, the City of London (which is a single local authority) contains approximately 25 wards. The advantage of using wards as the geographical units is that they are very small spatial units, which allows us to identify phenomena that would be unobservable at a higher geographical level. The detailed spatial scale is a crucial for our identification strategy that uses spatial variation in the accessibility increases in the vicinity of road schemes.

To construct the ward level background characteristics used as control variables we use information from the Census 2001 provided by 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 in time during our period of study) calculated using GIS and the dataset of transport improvements described in 3.2.

3.1. Firm data

The data source for the aggregate analysis of employment and plant counts at ward level is the Business Structure Database (BSD) for the period 1998-2008. Counts of employment and plants at ward level are used in the estimation of equation (2), both in construction of the dependent variables for the aggregate employment and plant count regressions, and for calculating the accessibility measures described in Section 2.2 above. The BSD is maintained by the UK Office of National Statistics (ONS) and contains a yearly updated register of the universe of businesses in the United Kingdom. It covers about 98% of business activity (by turnover) in Great Britain. The smallest unit of observation is the establishment or plant ("local unit"), but there is also information 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 the location (postcode), the sector of production (up to 5 digits) and employment of the 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 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 Annual Respondents Database (ARD). The ARD holds responses to the Annual Business Inquiry (ABI). The ABI is a stratified random sample, extracted from the BSD[3]. The ABI is a comprehensive business survey covering balance-sheet information like gross value added, gross output, wages, intermediate inputs, employment, industry, and investment. We use and the EU KLEMS Deflators (base 1995) to express the firm balance-sheet data in real terms. Although the ARD only contains a sample of small businesses, being a census of large businesses it contains information of firms which cover a large fraction of the employment (for example 90% of UK manufacturing employment). We use the balance-sheet data from ARD in the estimates of plant productivity, output and labour costs. Imputed estimates of the capital stock are available, which allow us, in principle to look at total factor productivity effects in firm level production functions. However, the data is only currently available up to 2004, and we not use it in the current version of this paper.[4] Instead we look only at value-added

3 For details see Criscuolo, Haskel and Martin (2003).

4 The capital stock variable is built from the gross investment flows using a perpetual inventory method and allowing for differential depreciation rates across the three main asset classes (equipment, structures and

(defined as price deflated revenue minus materials inputs), value-added per worker as an indicator of labour productivity, gross revenue, and labour costs divided by employees to measure 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). Balanced-sheet data is available at the RU level, while location and employment is available at the LU level. Questions related to employment can be investigated at LU level, since reporting units with several plants report on several local units that may be located in different wards[5]. Similarly, 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. Note though that there is a risk of attenuating the estimated impacts of transport improvements on labour productivity, if not all plants are affected by the transport improvement, because the productivity gains to the affected plants will be combined with the (potentially zero) productivity changes in unaffected plants when calculating the firm level value-added per worker. Allocating value-added and revenue across plants from firm level data is similarly problematic. Productivity and technology of production might vary across local units, across reporting units or indeed within local or reporting units. Hence, strong assumptions are needed in order to calculate output at the plant level (we would need to apportion the RU balance-sheet information across the LU which belong to it, for example based on their share of employment).

3.2. Road network data and origin-destination matrix construction

Once we have ward (by year) level employment aggregates, the second component in the employment accessibility index is an origin-destination (O-D) matrix containing the costs c_{jkt} (journey time) between each origin and destination. This matrix is required for different years.

We use data on traffic speeds on a generalised primary road GIS network for Great Britain provided by the Department for Transportation (DfT). The network includes all major roads that, according to

---

vehicles), available at the ONS-Business Data Linkage laboratory and constructed by the Centre for Research into Business Activity (CeRIBA).

5 As explained before, we define a plant within a ward as the presence of a firm within that ward, regardless of the number of sites the firm has in there. This is due to lack of precision of the local unit identifiers across time.

the DfT, cover roughly 65% of vehicle kilometres. Traffic speeds are modelled from traffic flow census data using the FORGE component of the National Transport Model (NTM). We construct the road network for years 1998 – 2008 (beginning of the year) by using the 2008 network provided by the DfT and information on major road improvement projects in 1998 – 2008 provided by the Highways Agency. We identify 31 major projects which added 320 kilometres of new roads to the network (total length in 2008 was roughly 50,000km).[6]

Using this generalised traffic network, we use the network analysis algorithms in ESRI ArcGIS to compute least-cost (minimum journey time) routes between each origin ward j and destination ward k in years 1998 - 2008. Therefore, the accessibility indices A and $\hat{A}$ are calculated using the ward employment data from BSD, and the ward-to-ward O-D travel times in years 1998 – 2008 using equations (7) and (8).

The road network consists of roughly 17,000 road links annually. We start from the 2008 network and construct the networks for earlier years 1998-2007 by deleting links from the network based on their opening years which we obtain by matching the links with the Highways Agency road improvement data.

We use journey times, obtained from the NTM, in the non-busy direction averaged over all time periods between Monday-Friday 08:00 and 18:00. We use journey times in 2003 for the whole period 1998 - 2008. For links opened after 2003 we use estimated journey times from a regression model using a dataset of links in the 2003 network. We regressed link speeds from the NTM on speed limit dummies, traffic flow, traffic flow squared, road category dummies (six categories) and local

6 The total length of the road schemes in our data corresponds closely to the official statistics by the Department of Transport (311 km of new main roads between 1999 and 2009).

authority dummies. The regression predicts speeds from the NTM reasonably well (R-squared = 0.76). We then used the results to predict travel times for links opened after 2003 for which no NTM speed 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.

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

When computing the O-D matrix we apply a limit of 75 minute drive time. This limit facilitates O-D matrix computation but does not affect the value of accessibility index because wards beyond 75 minutes would have negligible weights in the calculation of A.

It should be noted that the network is highly generalised. Journeys via the minor road network are not modelled. Forbidden turns and one way systems are not modelled. All link intersections are treated as junctions. The changes in accessibility must therefore be regarded as approximate.

We have cross checked the times and subsets of the accessibility measures against estimates derived from Google maps, using the stata 'travel time' module (citation). 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 between the Google maps travel time and ours becomes weaker as we consider shorter journeys, presumably because shorter trips that do not include sections of our generalised network are poorly approximated by our O-D calculation. This has implications for our empirical work, since it implies that the estimated accessibility changes in the vicinity of the transport improvements may be quite noisy. For this reason, and because locations immediately proximate to new road schemes may be adversely affected by the scheme (e.g. loss of premises, and environmental impacts), we drop wards and plants within 1km of the road schemes in our analysis.

4. Results

4.1. Descriptive statistics on accessibility

Table 1 summarises the changes in the Log accessibility from 1998 to 2008 (reported as approximate % changes). In the upper 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. In the lower panel we calculate the accessibility allowing both the employment and the road infrastructure to vary (A).

The upper panel of Table 1 shows that employment accessibility change induced by road improvements over this period was on average small, only 0.34%. However, this increases substantially when we focus on wards closer to improvements. For example, for the changes within the 10 km distance band the mean is 1.2% and the 90th percentile is 3.2%. As we expand the sample away from the sites of the schemes, changes in accessibility tend to fall. Within 20km, which we use in our base specification, mean accessibility change is 1.2% and 90th percentile is 2.0%. Within 30km of the schemes mean accessibility change is 0.95% and 90th percentile is 1.7%. The lower panels of Table 1 show what happens when we calculate he accessibility indices, allowing for the fact that the spatial distribution of employment changes over the period (but potentially endogenously to our outcomes of interest). It is evident that employment changes are a more important driver of variation in accessibility to employment than road improvements. Nevertheless, variation due to road improvements is non-negligible relative to overall changes, suggesting that the instrumental variable strategy outlined in Sections 3 is feasible.

Maps 1 and 2 illustrates the spatial relationship between road schemes and resulting accessibility increases. The left panel of Map 1 shows the locations of new roads and major improvements and the right panel shows that 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 grey lines show the primary road network in 2008. New links and significant improvements between 1998 and 2007 are indicated by bold lines. The black lines are ward boundaries. The map illustrates that the effect of road improvements on accessibility vary considerably across wards in the vicinity of the same improvement. We argue that differences in accessibility changes across neighbouring wards are coincidental and can be treated as exogenous, especially when controlling for differential time trends near different schemes.

Table 2 summarises the number of plants and total employment in the wards within 10-30km of the new road schemes.

4.2. Ward-level employment and plant count regressions

The first regression results, presented in Table 3, are ward level regressions of log of employment and log of the number of plants on log accessibility. Employment is based on BSD data as described in Section 3.1. The main results in columns (1)-(6) use data on wards within 20km of road schemes and contain various different model specifications. Columns (7) and (8) present the same specification as column (5), but applied to samples within smaller (within 10km of the schemes) and larger (within 30km of the schemes) samples. The tables show the coefficient on employment accessibility, its standard error, the number of observations used, and the first stage F-statistics for the IV specifications. Standard errors and F-stats are 'clustered' at ward level, to allow for arbitrary intra-ward correlation over time.

The first specification in column (1) is a simple OLS regression which neglects all endogeneity issues. The OLS coefficients are positive and significant showing that wards with higher employment tend to have better accessibility, which is probably unsurprising, since high employment wards will tend to be collocated with other high employment wards. In the second column, as a first step towards obtaining causal estimates, we add ward fixed effects to control for time-invariant ward specific factors, so the estimates are now based off changes in accessibility over time. The point estimate on accessibility is reduced substantially, but still remains large, positive and significant at the 1% level. Column (3) introduces the instrumental variable strategy, using accessibility index $\hat{A}$ as an instrument. Recall this 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, i.e. 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 3 add in more control variables. In column (4), a dummy for 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 the 20km radius of a given scheme. Column (5) goes further and interacts a time trend with distance to the scheme, and a time trend interacted with a dummy for year that the scheme opened to allow for other common time patterns that might cause accessibility and ward level employment to move together over time. Column (6) adds in an interaction of a time trend with a set of census variables for each ward, to allow for time patterns related to the underlying demographics. In all cases, the point estimate remains large, and the magnitude increases and significance improves relative to the basic IV estimate in column (3). In the saturated model in column (6), the estimate again becomes only marginally significant, although large in magnitude. Expanding the area considered in column (7) leaves the results unchanged, although reducing the area to within 10km of schemes leads to smaller and imprecise estimates. This might appear surprising, given of the largest changes in accessibility appear in these areas (see Maps 1 and 2). However, as pointed out in Section 3.2, some of these local accessibility changes may be imprecisely measured due to the generalisation of our road network which lacks minor road detail (e.g. a ward may appear closely linked to a new road because it is a short straight line distance away from one of the link nodes, even though there is no direct minor road link). The estimates may therefore be downward biased in close proximity to the schemes.

The headline story from this table suggests that accessibility changes induced by road transport 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 20km 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%.

Table 4 extends this analysis to look at the contributions by broad industrial sector, presented in separate panels moving down the table. The structure is otherwise identical to Table 3. The impression from this table is that most of the action on employment is not coming from manufacturing, construction or consumer services, but is coming from changes in the producer services, land transport and 'other' sectors. 'Other' here is a residual category that includes the primary sector and public sector. Both the land transport and producer services effects are consistent with a story in which road transport has lowered transport costs for intermediates, business travel and stimulated employment in the logistics sector, though we cannot go further with this analysis in pinning down the precise mechanisms empirically. The elasticity in the transport and producer services sectors is as high as 1 in some specifications.

Moving now from employment to the number of plants (business premises), Table 5 and Table 6 present results analogous to those in Table 3 and Table 4, but for plant counts from the BSD at ward level. Plant counts are potentially more reliable than employment measures, and less vulnerable to error in data collection, so we might expect the effects here to be more precisely measured. This indeed turns out to be the case. The general picture from Table 5 is very similar to that for employment, although the results are even more stable and significant across different specifications, even within the 10km band. Evidently, the employment changes are very likely 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 we 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.

4.3. Robustness checks: alternative accessibility indices and distance bands

Table 7 presents the results of regressions, identical to column 5 of Table 3 and Table 5, but using alternative measures of accessibility as proxies for the intensity of exposure to the transport improvements. Column (1) uses a 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 results here are slightly higher than when using employment accessibility, but of a similar order given the standard errors. Column (2) uses an index of accessibility to plants, constructed in the same way from the BSD plant. 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 generates a bigger number. The differences in the scale of these parameter estimates relates largely to the change in the variance of the accessibility measures under these alternative weighting schemes. If we were to standardise the effects (divide by the standard deviation of he accessibility variables) we would 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 at what distances these transport impacts are occurring in more detail. We have shown that the employment effects are strong within 20km and 30km, but appear weaker close to the schemes within 10km. On the other hand, the plant count effects are strong at all distances. One concern might be that we are observing displacement of firms and workers from more distance zones to zones closer to the schemes. Table 8 explores this further using the standard IV specification of column 5 of Table 3 and Table 5. It repeats the results for the 1-10km, then columns (2) and (3) show the outer bands from 10km-20km and 20km-30km. Looking at the employment results, it again becomes clear that effects beyond 10km dominate. The more important point to take away, is that the effect is large and significant even within a 10km width ring beyond 10km, implying that if displacement is occurring, it is not obviously occurring towards (or away) from wards sited close to the new schemes. The employment impacts show up in response to accessibility changes within this 10km width ring. Employment effects are still large, but imprecisely measured in the outer ring beyond 20km, implying that most of the action in our results is located within the 10km-20km band.[7] For plant counts, the effects are again strong and significant in all distance rings, and increase in magnitude as we move outwards. Again, there is no strong evidence for the effects being driven predominantly through displacement from outer to inner wards. 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'.

4.4. Plant-level employment regressions

The main results in the preceding tables suggest that increased accessibility leads to increased employment and number of plants, at least for some sectors. These findings could mean that existing firms are increasing their employment, and/or new firms are entering. We have already shown that firm entry appears to be a main contributor (through the count of plants) to employment changes, but we can explore the issue further by looking at within-plant changes. [8] Table 9 presents the key results (the sectoral breakdown is in the Appendix) using a similar structure to the earlier tables on wards. Additional control variables in these plant level regressions are a sector specific time trend (using the 6 broad sectors used for the sector-specific results above) and a time trend specific to single plant firms ('singletons'). As before, standard errors and F-statistics are clustered at ward level.

[7] 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 squared procedure.

[8] We made an attempt to study entry and exit rates with the BSD butt the data proved too noisy to get reasonably precise estimates.

The coefficients in Table 9, 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 30km band stacks up with the pattern observed in Table 3 and Table 5. The implication is that overall within the 30km radius the number of plants have grown, but employment within plants has shrunk in response to transport improvements. Hence the aggregate employment effects within the 30km radius are weak in Table 3 (column 8). The microeconomic explanation for employment shrinkage within firms remains unclear, but could arise due to firm restructuring. More generally, it is clear that the overall aggregate employment effects in Table 3 are not related to within-plant changes, but due to new firms entering. Evidently, transport improvements have very little effect on the existing firms decisions to expand or contract employment, but do have a sizeable impact on firm entry.

4.5. Productivity and other production related outcomes

Although we find no response of existing firms on the employment margin, existing firms may experience productivity gains due to lower transport costs and workforce reorganisation, which result in increased output, potentially leading to higher wages. The results in this section explore these possibilities directly by looking at value-added per worker, gross-value added, real revenue, and wages. Recall from Section 3.1 that the information on output and value added is only available at the higher RU (firm level) so approximations must be made in allocating these to plant level when there are multiple plants per firm. The key results for all sectors pooled together are in Table 10. We restrict attention to the 20km radius. The specifications are otherwise similar to Table 9.

The headline story here is that we observe positive effects from transport accessibility on all outcomes and the majority are 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 labour productivity responds to transport induced accessibility changes with an elasticity of 0.5. Similar figures emerge for gross value added and revenue (as expected, given we observed 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.

Sector specific results for these economic variables are more mixed, but are presented in the Appendix for reference. No clear pattern emerges here, because many of the estimates are very imprecise, although nearly all are positive and of similar order of magnitude. In terms of significance, the consumer services sector results are strongest. Put together with the sectoral results on employment, a picture emerges in which transport improvements are inducing entry of firms in most sectors apart from consumer services and construction, with no employment effects on existing firms. Existing consumer services firms experience output increases, labour productivity increases and pay higher wages.

Additional analysis on these productivity-related outcomes aggregated to ward level, suggests that the productivity effects are not strongly evident at this more aggregated level (see Table 11). 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. One way of reading these findings, in the light of earlier results, is that we are seeing productivity and output gains in smaller, consumer services firms, which do not translate into strong output gains at the aggregate level, moreover, coupled with the inflow of firms and increase in employment, aggregate labour productivity remains relatively unchanged.

5. 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 transport improvements using changes in employment accessibility constructed at a micro-geographic scale (electoral ward level). These accessibility indices are similar to those often used in the process of 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 to address the likely endogeneity of changes in employment accessibility, by relying solely on transport improvements for identification in a panel data fixed effects setting. We argue that methods using cross-sectional variation in accessibility, effective density or other forms of ‘agglomeration’ or market potential index are biased by the spatial restructuring of employment, and do not have a causal interpretation.

Our estimates of the benefits from road transport improvements relate to those impacts that can be detected through changes in employment accessibility (i.e. changes in connectivity to economic mass). These benefits 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 than they do in the optimal routes from more remote locations. Overall, we find strong effects from transport improvements of this type on local aggregated employment and plant counts. A 10% improvement in transport induced accessibility leads to about a 3% increase in the number of businesses and employment, up to 30km 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 labour productivity improvements, more output and higher wages amongst existing firms, although these are not so evident at a more aggregated level.

Although these effects are large, when translated into the expected benefits from specific schemes, they are not so impressive. The average effect of all the road schemes in Britain between 1998 and 2007 was to raise mean accessibility at ward level by 0.34%. This implies, at best, something like 0.16% on total employmemt and a similar boost to the productivity of existing firms.

6. References

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, American Economic Association, vol. 100(2), pages 378-82, May.

Behrens, K., Lamorgese, A., Ottaviano, G.I.P., and Tabuchi, T. (2004). “Testing the home market effect in a multi-country world: The theory”, Centre for Economic Policy Research, Discussion Paper No. 4468.

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

Burchfield M., Overman H.G., Puga D. and Turner M. A., (2006). "Causes of Sprawl: A Portrait from Space," The Quarterly Journal of Economics, MIT Press, vol. 121(2), pages 587-633, May.

Chandra, A and Thompson, E, (2000). "Does public infrastructure affect economic activity?: Evidence from the rural interstate highway system", Regional Science and Urban Economics, Elsevier, vol. 30(4), pages 457-490, July.

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

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

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

Donaldson, D. (2010). “Railroads of the Raj: Estimating the Impact of Transportation Infrastructure”, mimio, 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," Forthcoming, American Economic Review.

Duranton, G. and Turner M.A., (2010). “"Urban Growth and Transportation”, mimeo, University of Toronto. December.

Faber, B. (2009), “Market Integration and the Periphery: The Unintended Consequences of New Highways in a Developing Country”, Mimeo, London School of Economics.

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.

Gramlich, E. M., (1994). "Infrastructure Investment: A Review Essay," Journal of Economic Literature, American Economic Association, vol. 32(3), pages 1176-96, September.

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

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.

Jiwattanakulpaisarn, P, Noland, R B. and Graham, D J., (2010). "Causal linkages between highways and sector-level employment," Transportation Research Part A: Policy and Practice, Elsevier, vol. 44(4), pages 265-280, May.

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

Martin, P and Rogers, CA, (1995). "Industrial location and public infrastructure," Journal of International Economics, Elsevier, vol. 39(3-4), pages 335-351, November.

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, volume 2163, number 1, pages 133—143.

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.

Puga D., (2002). "European regional policies in light of recent location theories," Journal of Economic Geography, Oxford University Press, vol. 2(4), pages 373-406, October.

Ottaviano I. P., G., (2008). "Infrastructure and economic geography: An overview of theory and evidence," EIB Papers 6/2008, European Investment Bank, Economic and Financial Studies

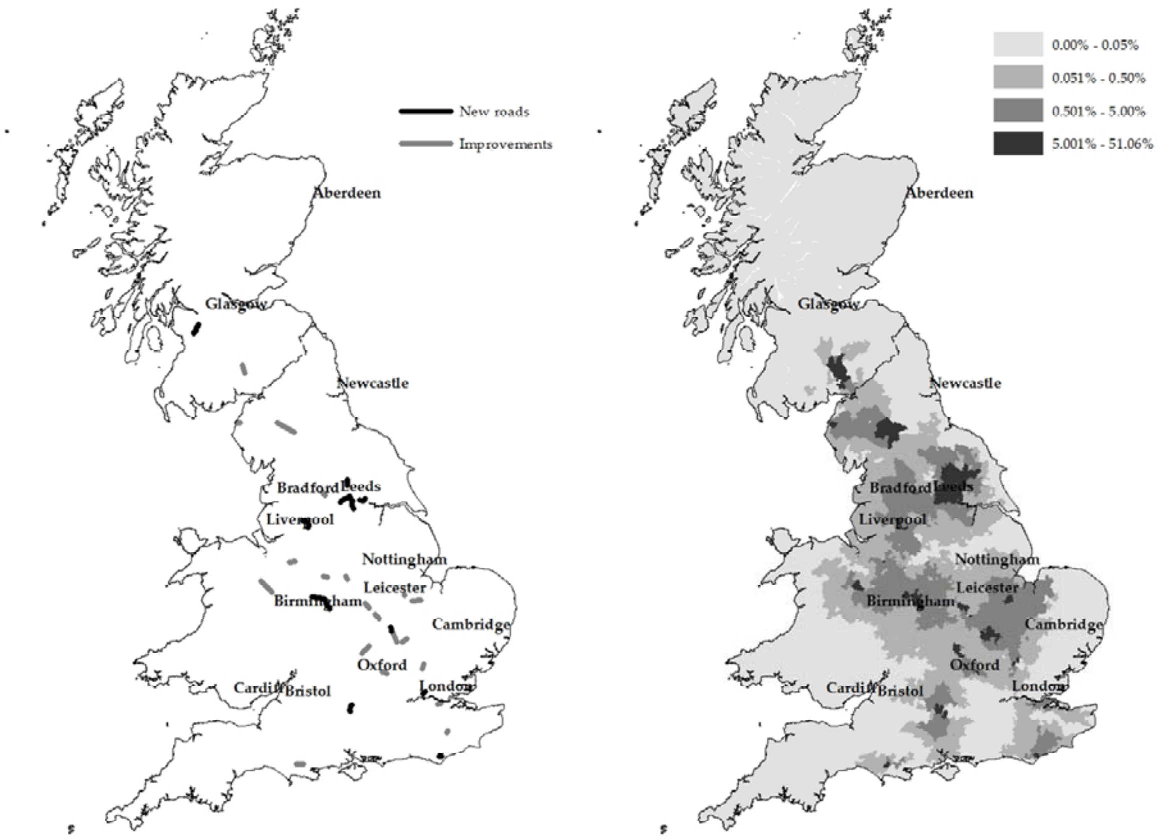
Staiger, Douglas and James Stock. 1997. "Instrumental Variables Regression with Weak Instruments," Econometrica, 65(3), 557-586.

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

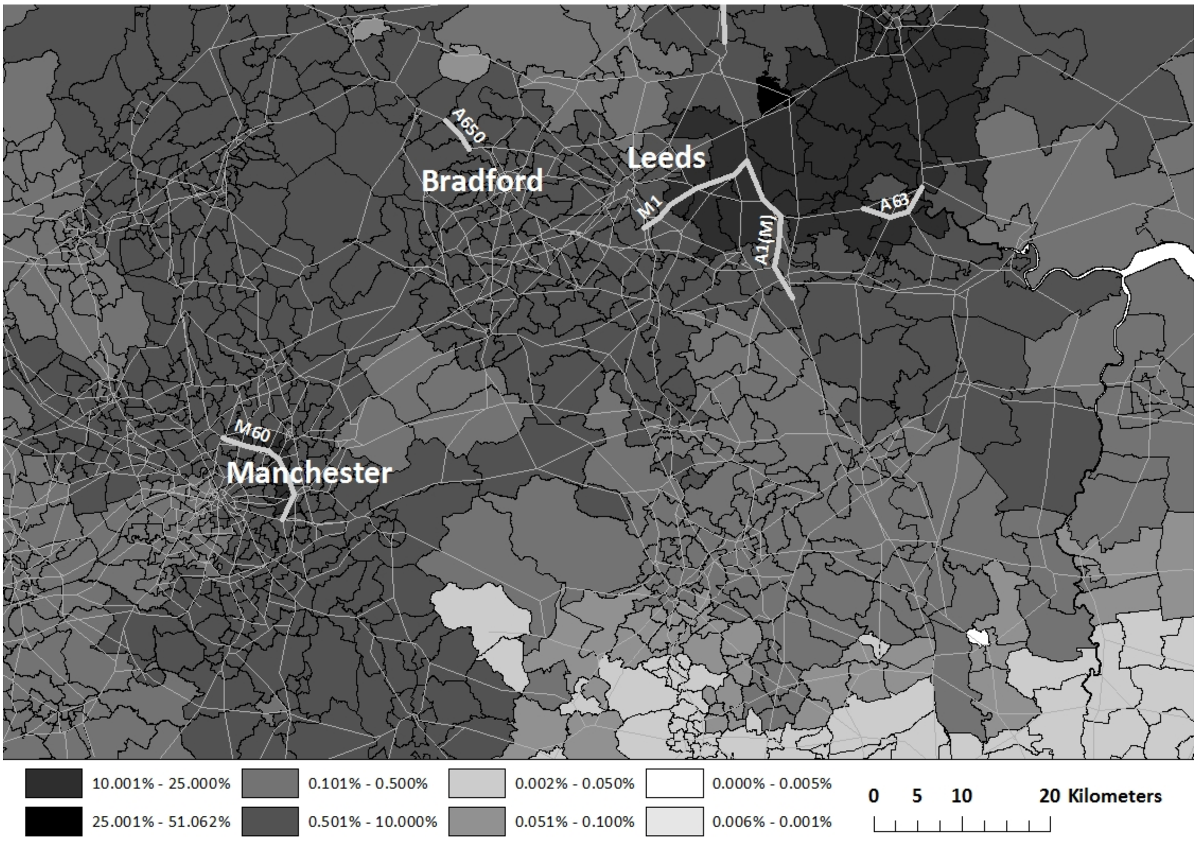
Venables, A. J. (2007) “Evaluating Urban Transport Improvements: Cost-Benefit Analysis in the Presence of Agglomeration and Income Taxation”, Journal of Transport Economics and Policy, Volume 41, Part 2, pp. 173–188

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

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 Manchester-Leeds area



7. Tables

Table 1. Long change in log accessibility between 1998 and 2008

	Wards	Mean	Std. Dev	90th percentile	Max	Proportion of zeroes
1997 employment and time-varying travel times						
All	10318	0.34%	1.22%	0.79%	31.37%	32.52%
10kms	1514	1.18%	2.45%	3.16%	31.37%	5.28%
20kms	3487	0.83%	1.97%	1.91%	31.37%	6.05%
30kms	4903	0.66%	1.71%	1.57%	31.37%	6.00%
Time-varying employment and 1998 travel times						
All	10318	7.17%	7.93%	13.55%	137.07%	0.00%
10kms	1514	5.20%	5.01%	10.21%	52.06%	0.00%
20kms	3487	5.48%	5.01%	10.47%	67.04%	0.00%
30kms	4903	5.70%	5.03%	10.83%	10.83%	0.00%
Time-varying employment and time-varying travel times						
All	10318	7.51%	8.06%	13.88%	137.07%	0.00%
10kms	1514	6.39%	6.03%	12.04%	52.10%	0.00%
20kms	3487	6.32%	5.75%	11.68%	67.09%	0.00%
30kms	4903	6.37%	5.57%	11.75%	67.09%	0.00%

Table 2: Employment and number of plants in wards (check numbers)

Distance band	10km		20km		30km	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Accessibility						
Ln(A), 1997 empl	15.44	0.86	15.41	0.82	15.35	0.84
Ln(A), time-varying empl	15.45	0.87	15.42	0.83	15.35	0.84
Employment						
ALL SECTORS	3730	9639	3155	7266	2879	6324
MANUFACTURING	472	942	435	897	409	853
CONSTRUCTION	163	298	151	306	143	287
CONSUMER SERVICES	796	2103	711	1615	657	1431
PRODUCER SERVICES	1034	4614	768	3295	674	2812
OTHER	1264	2976	1090	2355	996	2077
Local units						
ALL SECTORS	314	636	285	459	267	402
MANUFACTURING	25	38	22	34	21	30
CONSTRUCTION	24	18	24	17	23	18
CONSUMER SERVICES	85	142	77	112	72	100
PRODUCER SERVICES	104	353	91	244	84	215
OTHER	76	129	71	94	67	82

Table 3: Ward level employment regressions: all sectors

ALL SECTORS	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log accessibility	0.429***	0.252***	0.275*	0.363**	0.361**	0.314*	0.355**	0.199
	(0.020)	(0.060)	(0.162)	(0.176)	(0.178)	(0.177)	(0.177)	(0.207)
Observations	38247	38247	38247	38247	38247	38247	53823	16566
Wards	3477	3477	3477	3477	3477	3477	4893	1506
Distance band	1-20 kms	1-20 kms	1-20 kms	1-20 kms	1-20 kms	1-20 kms	1-30 kms	1-10 kms
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ward FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes
IV			Yes	Yes	Yes	Yes	Yes	Yes
IV first stage F-stat			2667	4958	4651	4507	4653	4522
Scheme trends				Yes	Yes	Yes	Yes	Yes
Controls					Yes	Yes	Yes	Yes
Census trends						Yes		

Notes: controls includes a linear trend interacted with the distance to closest improvement and a trend interacted with year of opening
Census trends includes 2001 census characteristics (mean age, unem rate, share of WAP with high degree, share of social housing) interacted with a linear trend
Reduced form estimate in preferred specification (5) is 0.399*** (0.198)

Table 4: Ward level employment regressions: by sector

	(1)	(2)	(3)	(4)	(5)	(6)	(8)	(9)
MANUFACTURING								
Log accessibility	0.252***	0.526***	0.411	0.09	-0.003	-0.155	0.226	0.033
	(0.034)	(0.201)	(0.501)	(0.519)	(0.526)	(0.522)	(0.512)	(0.600)
CONSTRUCTION								
Log accessibility	0.239***	0.188	0.239	0.068	0.221	-0.012	0.1	0.397
	(0.022)	(0.163)	(0.336)	(0.342)	(0.348)	(0.343)	(0.341)	(0.430)
CONSUMER SERV								
Log accessibility	0.448***	0.107	0.068	-0.199	-0.276	-0.335	-0.315	-0.435
	(0.023)	(0.086)	(0.253)	(0.281)	(0.288)	(0.291)	(0.282)	(0.352)
PRODUCER SERV								
Log accessibility	0.749***	0.649***	1.646***	0.995***	0.878**	0.792**	1.014***	0.898**
	(0.027)	(0.126)	(0.388)	(0.371)	(0.376)	(0.375)	(0.377)	(0.453)
LAND TRANSPORT								
Log accessibility	0.198***	0.637**	1.285**	1.177*	1.06	1.146*	1.078*	0.061
	(0.028)	(0.286)	(0.599)	(0.640)	(0.647)	(0.657)	(0.642)	(0.753)
OTHER								
Log accessibility	0.344***	0.095	0.482**	0.570**	0.617***	0.655***	0.653***	0.334
	(0.020)	(0.081)	(0.219)	(0.229)	(0.232)	(0.237)	(0.233)	(0.287)
Observations	38246	38246	38246	38246	38246	38246	53820	16566
N_clust	3477	3477	3477	3477	3477	3477	4893	1506
Distance band	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	30 kms	10 kms
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ward FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes
IV			Yes	Yes	Yes	Yes	Yes	Yes
IV First stage F-stat			2667	4956	4650	4508	4652	4518
Scheme trends				Yes	Yes	Yes	Yes	Yes
Controls					Yes	Yes	Yes	Yes
Census trends						Yes		

Notes: controls includes a linear trend interacted with the distance to closest improvement and a trend interacted with year of opening
Census trends includes 2001 census characteristics (mean age, unem rate, share of WAP with high degree, share of social housing) interacted with a linear trend
Observations refers to maximum number of ward x year cells. Numbers vary slightly by sector.

Table 5. Ward level plant count regressions: all sectors

ALL SECTORS	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log accessibility	0.332***	0.105***	0.379***	0.283***	0.288***	0.203**	0.335***	0.262***
	(0.015)	(0.034)	(0.093)	(0.088)	(0.088)	(0.086)	(0.089)	(0.096)
Observations	38269	38269	38269	38269	38269	38269	53834	16577
Wards	3479	3479	3479	3479	3479	3479	4894	1507
Distance band	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	30 kms	10 kms
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ward FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes
IV			Yes	Yes	Yes	Yes	Yes	Yes
IV First stage F-stat			2667	4956	4650	4508	4652	4518
Scheme trends				Yes	Yes	Yes	Yes	Yes
Controls					Yes	Yes	Yes	Yes
Census trends						Yes		

Notes: controls includes a linear trend interacted with the distance to closest improvement and a trend interacted with year of opening. Census trends includes 2001 census characteristics (mean age, unem rate, share of WAP with high degree, share of social housing) interacted with a linear trend. Reduced form estimate in preferred specification (5) is 0.318*** (0.098). First stage F stat >3000

Table 6: Ward level plant count regressions: by sector

	(1)	(2)	(3)	(4)	(5)	(6)	(8)	(9)
MANUFACTURING								
Log accessibility	0.325***	0.358***	0.935***	0.580**	0.620**	0.439*	0.675***	0.638**
	(0.019)	(0.097)	(0.236)	(0.247)	(0.248)	(0.243)	(0.245)	(0.292)
CONSTRUCTION								
Log accessibility	0.171***	0.188***	0.299*	0.131	0.211	0.07	0.258	0.01
	(0.016)	(0.072)	(0.174)	(0.181)	(0.184)	(0.177)	(0.185)	(0.213)
CONSUMER SERV								
Log accessibility	0.374***	0.069	0.076	-0.079	-0.088	-0.143	-0.045	-0.029
	(0.017)	(0.052)	(0.149)	(0.144)	(0.148)	(0.151)	(0.145)	(0.174)
PRODUCER SERV								
Log accessibility	0.596***	0.252***	1.179***	0.682***	0.633***	0.534***	0.670***	0.765***
	(0.020)	(0.088)	(0.198)	(0.183)	(0.183)	(0.183)	(0.183)	(0.222)
LAND TRANSPORT								
Log accessibility	0.080***	0.204*	0.874**	0.954**	0.929**	0.960**	1.004***	0.529
	(0.014)	(0.109)	(0.342)	(0.371)	(0.373)	(0.376)	(0.368)	(0.438)
OTHER								
Log accessibility	0.134***	0.073*	0.306**	0.429***	0.433***	0.440***	0.486***	0.454**
	(0.016)	(0.044)	(0.141)	(0.145)	(0.144)	(0.141)	(0.143)	(0.178)
Observations	38268	38268	38268	38268	38268	38268	53831	16577
Wards	3479	3479	3479	3479	3479	3479	4894	1507
Distance band	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	30 kms	10 kms
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ward FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes
IV			Yes	Yes	Yes	Yes	Yes	Yes
IV First Stage F-Stat			2667	4956	4650	4508	4650	4518
Scheme trends				Yes	Yes	Yes	Yes	Yes
Controls					Yes	Yes	Yes	Yes
Census trends						Yes		

Notes: controls includes a linear trend interacted with the distance to closest improvement and a trend interacted with year of opening. Census trends includes 2001 census characteristics (mean age, unem rate, share of WAP with high degree, share of social housing) interacted with a linear trend. Observations refers to maximum number of ward x year cells. Numbers vary slightly by sector.

Table 7: Robustness of ward employment and local unit count results to distance decay. 20km radius

ALL SECTORS	(1)	(2)	(3)	(4)	(5)
Economic size	Addresses	LU counts	Empl	Empl	Empl
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)
Plants					
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
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

Note: observations reported for employment regressions; number of obs differs slightly in plant count regressions

Table 8: Ward employment and plants: within distance bands from scheme. IV estimates

ALL SECTORS	(1)	(2)	(3)
Distance band	1-10k	10-20k	20-30k
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

Note: observations reported for employment regressions; number of obs differs slightly in plant count regressions

Table 9: Plant level employment: all sectors

ALL SECTORS	(1)	(2)	(3)	(4)	(5)	(6)	(8)	(9)
Log accessibility	0.069***	0.068***	0.095	-0.009	-0.048	-0.04	-0.062	-0.228*
	(0.015)	(0.026)	(0.094)	(0.088)	(0.089)	(0.088)	(0.090)	(0.118)
Observations	2065343	2064780	2064780	2064780	2064780	2064780	2737108	977378
N_clust	3487	3487	3487	3487	3487	3487	4903	1514
Distance band	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	30 kms	10 kms
Year-Sic FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ward FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes
IV			Yes	Yes	Yes	Yes	Yes	Yes
IV First Stage F-Stat			658	1087	1054	951	1114	1813
Scheme trends				Yes	Yes	Yes	Yes	Yes
Controls					Yes	Yes	Yes	Yes
Census trends						Yes		

Regressions include sic-year dummies, a dummy for singleton plants. The sample excludes the plants situated 1km 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 10: Plant level economic outputs: all sectors. 20km 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	824980	687877	687877	687877	687877
Wards	3487	3473	3473	3473	3473
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

Observations reports maximum number of plant x year observations

Table 11: Ward level economic outputs: all sectors. 20km radius

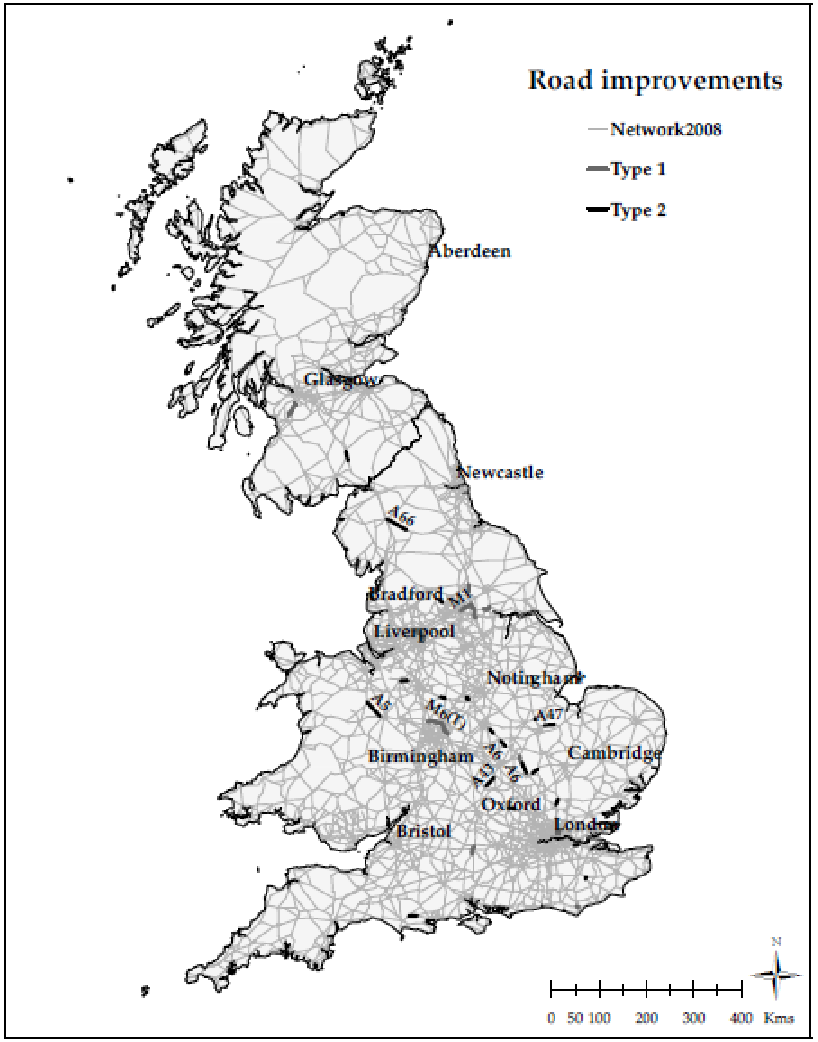
Real gross value added pw					
Log accessibility	0.062***	0.014	0.122	-0.093	0.023
	(0.011)	(0.128)	(0.437)	(0.475)	(0.470)
Real gross value added					
Log accessibility	0.398***	0.372	0.273	0.776	0.443
	(0.031)	(0.244)	(0.860)	(0.927)	(0.919)
Observations	34380	34378	34378	34378	34378
Wards	3487	3485	3485	3485	3485
Distance band	20 kms	20 kms	20 kms	20 kms	20 kms
Year FE	Yes	Yes	Yes	Yes	Yes
Ward FE		Yes	Yes	Yes	Yes
IV			Yes	Yes	Yes
widstat			2582	4199	4014
Scheme trends				Yes	Yes
Controls					Yes

8. Appendix Tables

Major road schemes in Britain 1998-2007

Opening year	Type	Road	Scheme	Length in kms
1998	Type 2	A16	A16 Market Deeping/Deeping St James Bypass	1.6
1998	Type 1	A34	A34 Newbury Bypass	9.3
1998	Type 2	A50	A50/A564 Stoke - Derby Link	5.1
1999	Type 1	A12	A12 Hackney Wick - M11 Contracts I-IV	4.7
1999	Type 2	A35	A30/A35 Puddleton Bypass	9.3
1999	Type 1	M1	M1/M62 Link Roads	16
1999	Type 2	M74	A74(M). Paddy's Rickle - to St Ann's (J16)	11.6
2000	Type 1	M60	M66 Denton - Middleton	15.3
2002	Type 1	A27	A27 Polegate Bypass	3.2
2002	Type 2	A43	A43 Silverstone Bypass	14.2
2002	Type 2	A6	A6 Clapham Bypass	14.57
2002	Type 2	A66	A66 Stainburn and Great Clifton Bypass	4.1
2003	Type 2	A41	A41 Aston Clinton Bypass	7.3
2003	Type 2	A5	A5 Nesscliffe Bypass	21.48
2003	Type 2	A500	A500 Basford, Hough, Shavington Bypass	7.7
2003	Type 2	A6	A6 Alvaston Improvement	4.7
2003	Type 2	A6	A6 Great Glen Bypass	6.8
2003	Type 2	A6	A6 Rothwell to Desborough Bypass	8.43
2003	Type 1	A6	A6 Rushden and Higham Ferrers Bypass	5.4
2003	Type 2	A650	A650 Bingley Relief Road	4.4
2003	Type 1	M6(T)	M6 Toll. Birmingham Northern Relief Road	29.7
2004	Type 2	A10	A10 Wadesmill to Colliers End Bypass	7
2004	Type 1	A63	A63 Selby Bypass	9.5
2005	Type 1	A1(M)	A1(M) Wetherby to Walshford	8.1
2005	Type 2	A21	A21 Lamberhurst Bypass	2.4
2005	Type 2	A47	A47 Thorney Bypass	10.7
2005	Type 1	M77	M77 Replaces A77 from Glasgow Road	18.25
2006	Type 1	A1(M)	A1(M) Ferrybridge to Hook Moor	19.2
2006	Type 2	A421	A421 Great Barford Bypass	7.6
2007	Type 2	A2	A2 / A282 Dartford Improvement	4.2
2007	Type 2	A66	A66 Temple Sowerby Bypass and Improvements at Winderwath	26.2
			TOTAL	318.03

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



Robustness to alternative accessibility definitions; sector results

	MANUFACTURING					*CONSTRUCTION*					*CONS SERVICES*				
Log accessibility	-0.171	-0.128	-0.188	0.17	-0.434	0.285	0.19	0.405	0.045	0.852	-0.516	-0.371	-0.464	-0.278	-0.949
	(0.847)	(0.600)	(1.015)	(0.367)	(2.191)	(0.579)	(0.418)	(0.627)	(0.269)	(1.348)	(0.463)	(0.334)	(0.512)	(0.215)	(1.092)
Observations	37625	37625	37625	37625	37625	38184	38184	38184	38184	38184	38246	38246	38246	38246	38246
widstat	177	3604	3975	618	831	197	3771	4010	632	837	197	3761	3993	628	834
N_clust	3471	3471	3471	3471	3471	3477	3477	3477	3477	3477	3477	3477	3477	3477	3477
	PROD SERVICES					*LAND TRANSPORT*					*OTHER*				
Log accessibility	1.385**	0.982**	1.532**	0.618**	3.189**	1.734	1.2	2.229**	0.49	4.877**	1.049***	0.796***	1.311***	0.319**	2.894***
	(0.639)	(0.447)	(0.695)	(0.270)	(1.494)	(1.113)	(0.796)	(1.131)	(0.520)	(2.407)	(0.396)	(0.278)	(0.447)	(0.149)	(0.969)
Observations	38219	38219	38219	38219	38219	35198	35198	35198	35198	35198	38246	38246	38246	38246	38246
widstat	197	3762	3985	628	833	175	3409	3556	621	737	197	3762	3994	628	834
N_clust	3477	3477	3477	3477	3477	3387	3387	3387	3387	3387	3477	3477	3477	3477	3477
Economic size	Address	Plants	Empl	Empl	Empl	Address	Plants	Empl	Empl	Empl	Address	Plants	Empl	Empl	Empl
Cost function	Inv 1	Inv 1	Inv 0.5	Inv 1.5	Inv exp0.2	Inv 1	Inv 1	Inv 0.5	Inv 1.5	Inv exp0.2	Inv 1	Inv 1	Inv 0.5	Inv 1.5	Inv exp0.2
Distance band	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms
IV	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE and controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robustness to alternative accessibility definitions; sector results

	MANUFACTURING			*CONSTRUCTION*			*CONSUMER SERV*		
Log accessibility	0.033	0.217	4.661*	0.397	0.509	1.616	-0.435	0.42	0.664
	(0.600)	(1.077)	(2.775)	(0.430)	(0.656)	(1.809)	(0.352)	(0.513)	(1.777)
Observations	16290	21335	15326	16517	21667	15537	16566	21680	15576
widstat	4438	1131	93.1	4542	1141	89.4	4522	1139	89
N_clust	1504	1967	1411	1506	1971	1416	1506	1971	1416
	PRODUCER SERV			*LAND TRANSPORT*			*OTHER*		
Log accessibility	0.898**	0.917	6.165**	0.061	2.577**	-1.526	0.334	1.481***	0.95
	(0.453)	(0.629)	(2.779)	(0.753)	(1.094)	(3.328)	(0.287)	(0.388)	(2.298)
Observations	16555	21664	15558	15352	19846	14189	16566	21680	15574
widstat	4520	1139	89.7	4171	884	91.8	4522	1139	89
N_clust	1506	1971	1416	1471	1916	1377	1506	1971	1416
Distance band	1-10k	10-20k	20-30k	1-10k	10-20k	20-30k	1-10k	10-20k	20-30k
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ward FE	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	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Employment and plant effects from 1998-2008 long differences in transport accessibility

ALL SECTORS	(1)	(2)	(3)	(4)	(5)
Employment					
ΔLog accessibility	0.643***	0.262**	0.301	0.273	0.217
	(0.068)	(0.110)	(0.184)	(0.270)	(0.271)
Plants					
ΔLog accessibility	0.403***	0.220***	0.444***	0.534***	0.525***
	(0.039)	(0.061)	(0.116)	(0.159)	(0.157)
Observations	3477	3477	3477	3477	3477
widstat			936	2319	2118
N_clust	3477	3477	3477	3477	3477
Distance band	20 kms	20 kms	20 kms	20 kms	20 kms
First-differences	Yes	Yes	Yes	Yes	Yes
Year dummies		Yes	Yes	Yes	Yes
IV			Yes	Yes	Yes
Scheme dummies				Yes	Yes
Controls (levels)					Yes

Plant level employment: sector 1-3

MANUFACTURING	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log of accessibility	-0.101***	0.141**	-0.138	0.002	-0.013	-0.152		-0.028	0.204
	(0.019)	(0.069)	(0.258)	(0.262)	(0.260)	(0.264)		(0.261)	(0.334)
Log of accessibility (1997)							-0.014		
							(0.282)		
Observations	220876	220338	220338	220338	220338	220338	220338	287754	105410
widstat			1190	2686	2530	2399		2433	1791
N_clust	3112	3104	3104	3104	3104	3104	3104	4355	1357
CONSTRUCTION									
Log of accessibility	0.069***	0.144	0.183	0.243	0.279	0.356		0.189	0.187
	(0.017)	(0.089)	(0.294)	(0.303)	(0.308)	(0.314)		(0.311)	(0.381)
Log of accessibility (1997)							0.302		
							(0.333)		
Observations	123752	123066	123066	123066	123066	123066	123066	168233	54222
widstat			1500	3914	3549	3390		3379	2567
N_clust	3319	3305	3305	3305	3305	3305	3305	4624	1424
CONSUMER SERV									
Log of accessibility	0.069***	0.068**	-0.033	-0.04	-0.064	-0.102		-0.067	-0.266*
	(0.015)	(0.030)	(0.108)	(0.110)	(0.111)	(0.113)		(0.112)	(0.158)
Log of accessibility (1997)							-0.071		
							(0.123)		
Observations	848761	847193	847193	847193	847193	847193	847193	1134518	388079
widstat			431	702	661	601		694	1642
N_clust	3477	3477	3477	3477	3477	3477	3477	4888	1512
Distance band	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	30 kms	10 kms
Year-Sic FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ward FE		Yes	Yes	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
Census trends						Yes			

ALL SECTORS regressions include sic-year dummies, a dummy for singleton plants. The sample excludes the plants situated 1km 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

Plant level employment: sector 4-6

PRODUCER SERV	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log of accessibility	0.132***	0.019	0.706**	0.221	0.087	0.29		0.108	-0.094
	(0.029)	(0.073)	(0.278)	(0.266)	(0.265)	(0.246)		(0.270)	(0.338)
Log of accessibility (1997)							0.097		
							(0.296)		
Observations	456766	455095	455095	455095	455095	455095	455095	591130	229588
widstat			444	636	677	608		701	687
N_clust	3429	3429	3429	3429	3429	3429	3429	4812	1492
LAND TRANSPORT									
Log of accessibility	0.040**	0.049	-0.129	-0.209	-0.24	-0.226		-0.282	-0.542**
	(0.019)	(0.045)	(0.170)	(0.177)	(0.177)	(0.181)		(0.180)	(0.245)
Log of accessibility (1997)							-0.265		
							(0.194)		
Observations	377565	374140	374140	374140	374140	374140	374140	495734	179617
widstat			755	1262	1227	1107		1317	1531
N_clust	3449	3442	3442	3442	3442	3442	3442	4836	1498
OTHER									
Log of accessibility	0.079**	0.012	-0.085	-0.283	-0.326	-0.445		-0.219	-0.315
	(0.032)	(0.163)	(0.584)	(0.587)	(0.590)	(0.580)		(0.554)	(0.581)
Log of accessibility (1997)							-0.343		
							(0.619)		
Observations	37623	37539	37539	37539	37539	37539	37539	49974	16959
widstat			1039	1827	1823	1733		1730	1436
N_clust	2220	2211	2211	2211	2211	2211	2211	3040	958
Distance band	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	30 kms	10 kms
Year-Sic FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ward FE		Yes	Yes	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
Census trends						Yes			

ALL SECTORS regressions include sic-year dummies, a dummy for singleton plants. The sample excludes the plants situated 1km 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

Plant level economic outputs: 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
widstat			731	2092	1884			599	921	871			255	433	416
N_clust	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
widstat			286	436	485			345	652	625			406	850	1212
N_clust	3415	3066	3066	3066	3066	3421	3201	3201	3201	3201	2118	1322	1322	1322	1322
Distance band	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms	20 kms
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