Spatial Economics Research Centre

SERC DISCUSSION PAPER **214**

New Road Infrastructure: The Effects on Firms

Stephen Gibbons (LSE , SERC and What Works Centre)
Teemu Lyytikäinen (VATT Institute of Economic Research)
Henry Overman (LSE, SERC and What Works Centre)
Rosa Sanchis-Guarner (Imperial and SERC/CEP, LSE)

May 2017

This discussion paper is a substantially revised version of SERCDP0117 and a slightly revised version of Centre for Economic Policy Research discussion paper 11239

This work is part of the research programme of the Urban Research Programme of the Centre for Economic Performance funded by a grant from the Economic and Social Research Council (ESRC). The views expressed are those of the authors and do not represent the views of the ESRC.

New Road Infrastructure: The Effects on Firms

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

May 2017

(This discussion paper is a substantially revised version of SERCDP0117 and a slightly revised version of Centre for Economic Policy Research discussion paper 11239)

* London School of Economics, SERC/CEP and What Works Centre
** VATT Institute for Economic Research
*** Imperial College Business School and SERC/CEP London School of Economics

The 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), at the University of Birmingham (October 2012) and at the UK Department for Transport (March 2013). We thank the participants for their comments. All errors and opinions are our own. This work was supported by the Economic and Social Research Council grant numbers ES/J007382/1 ES/J021342/1 and ES/G005966/1).

Abstract

This paper estimates the impact of new road infrastructure on employment and labour productivity using plant level longitudinal data for Britain. Exposure to transport improvements is measured through changes in accessibility, calculated at a detailed geographical scale from changes in minimum journey times along the road network. These changes are induced by the construction of new road link schemes. We deal with the potential endogeneity of scheme location by identifying the effects of changes in accessibility from variation across small-scale geographical areas close to the scheme. We find substantial positive effects on area level employment and number of plants. In contrast, for existing firms we find negative effects on employment coupled with increases in output per worker and wages. A plausible interpretation is that new transport infrastructure attracts transport intensive firms to an area, but with some cost to employment in existing businesses.

Keywords: productivity, employment, accessibility, transport
JEL Classifications: D24; O18; R12

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, which may foster economic integration, stimulate competition, generate agglomeration economies and a number of other 'wider' economic benefits.

A number of recent studies have looked at the impact of transport infrastructure networks on the spatial economy,[3] usually in an economic-historical context (Donaldson and Hornbeck, 2016; Duranton and Turner, 2012; Michaels, 2008; Baum-Snow, 2007) or developing country context (Ghani et al., 2016; Faber, 2014; Baum-Snow et al., 2016a). But for economies with well-developed transport networks, little is known about the effects at the micro level that result from additions to the network. Direct evidence of the causal effects of such improvements using ex-post evaluation of road network improvements is rare.

This paper provides such evidence by investigating the causal impact of road improvements on employment and productivity related firm outcomes, using administrative data on all businesses in Great Britain from 1997-2008. We measure exposure to road improvements using changes in a

[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.

[3] See Redding and Turner (2015) for a recent review of the literature.

continuous index of accessibility, calculated at a small geographical scale - electoral ward (there are 10,300 wards in Britain with an average land are of 24km^2 and population of 6000). These accessibility changes are based on optimal travel times, calculated from analysis of potential routes along the major road network. For a given location, this index measures the accessibility of potential destinations along the major road network. To construct this index, we use a bespoke dataset of road construction schemes carried out in Great Britain between 1998 and 2007, combined with road network data.

The principal challenge to estimation of the causal impact of road network changes on firm outcomes is that roads may be purposefully built to meet demand in places where productivity is growing, or to try to stimulate growth in places where productivity is falling. To address this problem of endogenous scheme location, we exploit the geographical detail in our data and identify the effects using only over-time variation in accessibility for wards that are very close to new schemes (within 10-30 km). This variation is incidental to the policy aims of the transport schemes, which are additions to the major road system aimed at improving network performance rather than local economic development. Balancing tests (discussed below) and an annex looking at the decision making process for a number of schemes (available on request), provide evidence to support this assumption.

Our study is unique in using this variation in accessibility changes close to specific road transport schemes to identify the effects of transport improvements. Restricting our attention to areas close to schemes results in little loss of generality relative to comparison across the whole country, because most of the variation in accessibility generated by relatively small scale road transport improvements occurs in areas close to new schemes. Our aim then is to infer the more general effects of changes in accessibility from these transport-induced changes occurring at a small geographical level.

Our continuous, network-based accessibility index is a crucial component in this research design, because it varies in complex ways that do not depend solely on a firm's distance to the transport improvement scheme. Instead, it varies across space, and changes over time, in ways that depend on a firm's position in relation to both the new *and* existing parts of the transport network, and whether new road links affect the optimal travel paths between the firm and destinations on the network. For example, two firms A and B sited at equal distance to a new road scheme will experience very different accessibility changes, if the new road becomes a link in many of the least-cost network paths from A to other destinations, whereas the new road is rarely if ever a link in the least-cost network path from B to other destinations. Therefore changes in the accessibility index are plausibly unrelated to location-specific characteristics that jointly influence new road placement and firm productivity and employment. To support this argument, we present tests that show there is no systematic relationship between accessibility changes and pre-trends in area characteristics. Other studies have used changes in similar, network based accessibility indices (an earlier version of our work Gibbons et al, 2010, Donaldson and Hornbeck, 2016; Holl, 2012; Baum-Snow et al 2016b) although, to our knowledge, no paper exploits localised variation in accessibility across areas that are close to road projects,[4] nor worked with such spatially refined data.

Our key finding is that road improvements increase ward-level employment and the number of businesses. A 1% increase in accessibility leads to a 0.3-0.4% increase in plants and employment. However, we find small negative employment effects at plant level, implying that the local

[4] Indeed, Donaldson and Hornbeck (2016) control for local changes in accessibility.

employment changes come about through firm entry and exit. Conversely, we find evidence of positive impacts on labour productivity (specifically on gross output and wage bill per worker).

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, Section 5 discusses the empirical results and Section 6 concludes.

2 Theoretical background and existing evidence

Theoretically, reduced transport costs and improved connectivity offer various direct benefits to firms. [5] Changes to logistics, business travel and internal organisation may improve productivity. If transportation services are a factor of production, reductions in transport costs will affect input choices. 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. 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). 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' that involve total factor productivity effects arising from agglomeration economies (Graham, 2007). These agglomeration externalities have origins in sharing of resources, matching of workers

[5] Our theoretical discussion draws mainly on Gibbons and Overman (2009) who provide further analysis.

to firms, and learning by information exchange (Duranton and Puga, 2004). Although usually associated with spatial concentration (e.g. cities or industrial clusters), these effects can just as well be attributed to lower travel times between locations (sometimes referred to as '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 market access, selection and sorting effects (Baldwin and Okubo, 2006). Better transport may encourage start-ups and survival by lowering costs, increasing returns to scale or agglomeration economies. Conversely, improvements can force the exit of low-productivity firms previously protected from competition (Melitz, 2003).

Due to the multiplicity of potential effects, theory 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. The theoretical predictions on the net effects of transportation improvements on area level outcomes are similarly varied. Traditional ex-ante appraisal of improvements have set many of these issues aside, by assuming a world of perfect competition in which all the economic benefits of transport improvements are captured by travel time savings and induced demand (Small, 2007). However, more recent studies (Eddington, 2006; Gibbons and Machin, 2006; Venables, 2007; Gibbons and Overman, 2009; Laird, Overman and Venables 2014) argue that this may not be a complete picture. In short, given the unclear theoretical predictions, the size and direction of the effects of transport policy on economic outcomes is mainly an empirical question.[6]

[6] A number of papers adopt a more structural approach to restrict the possible outcomes. See, for example, Donaldson and Hornbeck (2016) and Donaldson (2014). In contrast to these papers we use a reduced form approach paying particular attention to the issues of identification. Redding and Turner (2015) discuss both approaches in their recent survey.

A number of studies have tried to estimate the effects of transport improvements on the economy, but relatively few have looked at the impacts on firms at a micro or spatially disaggregated scale. Most of the empirical evidence considers the macro or regional level (for a review see Straub, 2011). These studies generally estimate aggregated production functions where infrastructure expenditure or roads are treated as a factor of production (García -Mila et al., 1996). Results, for a variety of outcomes are mixed.[7] Unfortunately, this literature struggles to address endogeneity concerns – that is, the fundamental problem that transport policy improvements are not randomly allocated, but are spatially targeted to meet specific economic and travel-related demands.

Recent papers have tackled this problem of endogenous transport improvements using various identification strategies. These include: a) using historical transport plans as instruments, under the assumption that the original plans are unrelated to current economic conditions; b) using physical geography as an instrument, under the assumption that physical geography is unrelated to current economic conditions; or c) assuming that some places are incidental beneficiaries of transport links, e.g. those located between big cities. Some papers use combinations of these ideas.

A number of papers in the US have used instruments derived from historical transport plans (or older routes) to look at various outcomes such as: urban growth (Duranton and Turner, 2012), road traffic (Duranton and Turner, 2011), trade patterns (Duranton et al., 2014), sub-urbanisation (García-Lopez et al., 2015; Baum-Snow, 2007), commuting patterns (Baum-Snow, 2010), 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. Baum-Snow et al. (2016) use a similar idea for China. Other studies use the second strategy we outline above. For example,

[7] Outcomes 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).

Faber (2014) uses optimal transport routes derived from physical geography, while Banerjee et al. (2012), Michaels (2008), Jedwab and Moradi (2016) use straight line paths between cities as instruments. Finally, several papers use the third strategy and claim that treatment of locations between cities or other network nodes is incidental to the aims of the transport policy, and therefore exogenous[8] (e.g. Chandra and Thompson, 2000; Holl, 2004a; Melo et al., 2010; Ahlfeldt and Feddersen, 2010; Ghani et al., 2016).

Studies of firm-related outcomes from road improvements are relatively rare. For India, Ghani et al. (2016) study a major national highway improvement programme in India and find that districts within 10km of the non-nodal sections of highways saw increases in entry rates and productivity in the manufacturing sector, compared with districts further away. Their identifying assumption is that the location of highway links between cities is exogenous, and they test this assumption by comparing with planned highways that were not constructed. Holl (2012) links a panel of firms to road network-based market potential indices for municipalities in Spain, and finds negative impacts of market potential on value-added in firm fixed effects specifications. She recovers positive effects by applying System GMM with instruments based on lags of the control variables, historical instruments and geology. Holl (2016) using Spanish data and historical roads as instruments, finds that access to highways is associated with higher firm level total factor productivity. The positive effect persists even when controlling for local employment density (to disentangle the direct effect from potential agglomeration effects). The identifying assumption in this approach is that the instruments affect firms only through their impact on later road network. Other papers have looked at firm relocation and entry (Coughlin and Segev, 2000; Holl, 2004a and 2004c) or birth (Holl, 2004b; Melo et al., 2010). Li and Li (2013) is distinctive in its focus on changes

[8] This is what Redding and Turner (2015) call the Inconsequential Units Approach.

in firm input inventories associated with expansion of the Chinese road network, using output inventories as a proxy for time varying confounding factors.

The identification strategy in our paper differs substantively from these previous approaches. We avoid instrumental variables approaches based on historical networks or plans (or any other historical variables) because: a) they would not be very relevant for the relatively small additions to the road network that we study; and b) we prefer not to rely on the excludability assumption that historical transport-related variables influence current economic performance only through changes to the current road network. Instead, we address the concern that transport schemes are endogenously targeted by estimating from treated places only; we do not use non-targeted places as a comparison group. Estimation is based on variation in the intensity of treatment within targeted locations, where treatment intensity is the magnitude of the change in accessibility induced by the new road scheme. Our paper is, to our knowledge, the first to exploit localised changes in road network accessibility in this way to identify the causal effects of transport improvements on production-related variables, using firm level micro data. The research design is discussed in detail in the next section.

3 Empirical methods

We measure the intensity of a firm's exposure to road improvements using an index of changes in 'accessibility', derived from changes in minimum travel times along the road network to potential destinations. These changes occur when new road links are constructed or existing links improved. We carry out our analysis at firm level and for small spatial units - electoral wards. The ward-level analysis allows us to capture effects working both through changes within firms, and through entry, exit and relocation of firms. The firm-level analysis captures within-firm changes only. Both approaches use the same general estimation strategy, applied to a panel of units (wards or firms) observed for up to 11 years, during the period 1998-2008. Data sources are described in Section 4.

3.1 General empirical set up

Our aim is to estimate the expected change in employment or productivity-related outcomes caused by a change in minimum travel times along the road network, induced by infrastructure improvements. We start from the basic regression equation:

$$lny_{jt} = \beta lnA_{jt} + \{u_j + \tau_t + \varepsilon_{jt}\} \quad (1)$$

Here y_{jt} is the outcome variable for unit *j* in year *t*, A_{jt} is a measure of travel time-based accessibility along the road network from origin unit *j* at time *t*. Unobservable factors include unit-specific time-invariant components (u_j) , year-specific unit-invariant components (τ_t) and year-by-unit varying (ε_{jt}) components.

The accessibility index at *j* is a proximity-weighted sum of economic activity at destinations *k*, where the proximity of *k* to *j* is a decreasing function of minimum journey times along the network, $a(time_{jkt})$:[9]

$$A_{jt} = \sum_{k \neq j} a(time_{jkt}) w_{k0} \quad (2)$$

The weights w_{k0} depend on the level of activity at destinations *k* in some base period. For the main part of our analysis the destination weights are 1997 employment, which precedes the first period in the estimation sample and the function $a(time_{jkt})$ is defined by a simple inverse time decay. Minimum journey times along the major road network are imputed from GIS network analysis. In

[9] Travel times are a proxy for travel costs. A generalised measure of transport costs would require additional information on other characteristics of infrastructure (e.g. reliability), vehicle and energy use, as well as labour, 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 time savings through infrastructure improvements.

existing literature, this index has been variously called an index of accessibility (e.g. El-Geneidy and Levinson 2006, Vickerman et al 1999), population or market potential (e.g. Harris 1954), effective density (Graham 2007), or market access (e.g. Donaldson and Hornbeck, 2016). The interpretation of this index as measuring market access has been shown to be theoretically grounded in trade theory (Donaldson and Hornbeck, 2016; Baum-Snow et al, 2016b). We use the generic word 'accessibility' throughout since we make no judgement on whether the effects work through access to markets or access to something else. We show results using alternative weights (e.g. employment, population etc.) and distance decay relationships, but the different indices are extremely highly correlated. Therefore it does not make sense in our context to place a theoretical interpretation on the accessibility index based on the choice of destination weights or functional form.

The parameter of interest is β in (1), interpreted as the causal effect of accessibility on economic outcomes. OLS estimates of β are very likely to be biased, because accessibility is non-random across space and time, and so is correlated with unobserved $\{u_j + \tau_t + \varepsilon_{jt}\}$. OLS regression compares units with high accessibility and units with low accessibility that are non-comparable on many unobserved dimensions. Part of this correlation occurs through unit specific fixed-over-time components u_j. In particular: a) faster transport connections may have been built to link more productive places; b) dense places may be more productive, and origins and destinations *j* and *k* are by definition closer together and network travel distances shorter in denser places, implying greater accessibility; c) the weights w_{k0}, if based on measures of economic activity will be endogenous if the outcome in *j* and in connected destinations *k,* are affected simultaneously by unobserved common productivity advantages.

A standard first step to eliminating the endogeneity induced by these fixed-over-time components of A_{jt} is to control for unit j fixed effects in (1), using standard within-groups regression (along with dummies to estimate general time effects $\tilde{\tau}_t$)

$$\widetilde{lny}_{jt} = \beta \widetilde{lnA}_{jt} + \tilde{\tau}_t + \tilde{\varepsilon}_{jt} \qquad (3)$$

Here the notation indicates $\widetilde{lnA}_{jt} = lnA_{jt} - \overline{lnA}_j$ and estimation of β comes from the within-unit changes in lnA_{jt}.[10] Given the structure of A_{jt} in Equation (2) and the way we construct A_{jt} (as described below in Section 4.3), these changes occur only through changes in minimum travel times between j and destinations k along the road network, caused by new or improved road infrastructure (where these changes are weighted by destination employment in the base period).

Now the concern is that changes in infrastructure incorporated in $\widetilde{lnA}_{jt}$ are correlated with changes in the time varying unobservables for unit j ($\tilde{\varepsilon}_{jt}$) if, for example, new road infrastructure is targeted at places that are experiencing better or worse than average productivity trends. To deal with this problem, we exploit the fact that in our data changes to minimum travel times are the result of a number of discrete road transport improvement schemes put in place in Britain over our study period (31 schemes over the period 1998-2008). It is primarily the location of these schemes within Britain that is potentially endogenous, due to policy targeting, not the accessibility changes occurring for units that happen to be close to these schemes. We can therefore control for endogenous scheme placement by controlling for geographical fixed and time varying effects related to scheme location.

[10] 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.

We do this in two ways. Firstly we restrict our sample to units *j* within a given distance buffer *b* of the nearest transport scheme (20 km in our main results). In this case, identification comes from comparison of units experiencing larger accessibility changes with units experiencing smaller accessibility changes, amongst the sub-sample of units that are all in close proximity to the road schemes that open over our study period. Secondly, we control for differential trends for each scheme within this sub-sample, by interacting nearest-scheme dummies with linear time trends.

Our identifying assumption is therefore that the variation in transport-induced, within-unit changes in accessibility is as good as random, when we compare units within a given radius of a particular road transport scheme. There are good arguments to support this assumption. The road schemes in our analysis are generally bypasses and motorway extensions that were intended to improve traffic flows between origins and destinations that are remote to the sites of the schemes (Department for Transport, 1997; Department for Transport, 2009). The variation in the changes in accessibility close to a scheme, while large relative to the changes elsewhere, are therefore an incidental by product of the scheme rather than its intended outcome. Related arguments have been made in other papers, e.g. Michaels (2008) argues that counties in intermediate rural locations between cities may be incidental beneficiaries of highways built between them. However in our case we are *not* using this argument to claim that the scheme location is exogenous, but that the incidental treatment means that the variation in accessibility changes amongst units in close proximity to each scheme location is exogenous. Appendix A provides a diagrammatic illustration of this point and, when we come to the empirical results later, Figure 3 shows an example drawn from our data. Our study is unique in using this variation in accessibility changes within distance buffers of specific road transport schemes to identify the effects of transport improvements.[11]

There are remaining concerns if the schemes are sited in such a way that their precise position and routing, and the accessibility changes they introduce, are correlated with very localised differential productivity trends. For instance, the route for a bypass may be chosen on the basis of low land prices, which in turn could indicate low potential productivity growth. Or, production close to schemes may be temporarily or permanently disrupted by construction works. We take a number of steps to mitigate these problems. Firstly we drop units which are crossed by, or very close to, improvements (within 1 km of any scheme). This also gets rid of rather mechanical sources of impact, such as service stations and fast food outlets built to serve road traffic.[12] Secondly we augment our regressions with further controls for differential trends over time and space within each nearest-scheme group. In the results, we show specifications with linear time trends interacted with: the straight line distance to nearest scheme; a dummy indicating observations in the periods after opening of the nearest scheme; and the level of accessibility at the beginning of our study period in 1997. We also experimented with trends interacted with 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) and discuss results for these in the text.

---

[11] Even if related to the Inconsequential Units Approach discussed in Turner and Redding (2014) our approach uses a different identification strategy to estimate the causal impact of road network additions. First, we exclude all locations directly affected by the new link including both the end nodes and all the wards in between, rather than just excluding end nodes. Second, we use changes in employment accessibility induced by reductions in travel time between locations rather than a simpler measure of connectivity to the network. Identification comes from exploiting variation across locations close to new road links. For a given location, the level of treatment is determined by the timing, the characteristics and the location of the new road link within the whole road network.

[12] It also helps mitigate attenuation bias that could arise due to errors in travel time calculations close to schemes, resulting from the map generalisation of our road network.

In order to further check for potential targeting of scheme routes in relation to local economic outcomes, we carried out an extensive review of the literature associated with the planning and evaluation of a selection of road schemes covered in our study. This review was based on the Post Opening Project Evaluation (POPE) reports from Highways England[13] and various other web sources, and is available on request in an annex. In this annex, we outline the decision making process and time line for a number of specific schemes (these were a randomly chosen selection covering small, medium and large projects). The overriding message is that route selection is driven by technical and cost considerations, not small scale variation in local economic conditions. Second, any local influence is concerned with design features that have nothing to do with travel times (e.g. tree planting, sound mitigation measures, etc.) and are not driven by local economic considerations. Third, although the set of schemes that gets funded is potentially endogenous, there is a very long lag between the decision to go ahead with a scheme and completion. For example, in the schemes we consider, the gaps are 10, 14, 17 and 8 years respectively.

The main regression equations (1), (3) and their extensions that include scheme fixed effects and other control variables, are estimated on ward-by-year-level or plant-by-year-level data. There are some specific points to consider when estimating the fixed effects regression Equation (3) for plant data.[14] Firstly note that the plant identifiers are location specific (changing if a plant moves to a different location). Thus, in the within-plant analysis, changes in accessibility are not caused by relocation of plant *j*, but only by changes in the transport network for a fixed plant location. Therefore estimation requires plants to appear in the data both before and after the opening of the

[13] The POPE reports are available from Highways England on their web site: http://www.highways.gov.uk/publications?publicationType=report-2&tag=pope

[14] 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.

transport 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 potential sample selection issues, if these firms that stay in response to transport improvements differ on unobserved dimensions from those that relocate (with a similar comment applying to the frequency with which firms appear in the same location multiple times in the data). These caveats aside, 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, as well as 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 road transport accessibility).

3.2 Justification for using accessibility to measure exposure to transport improvements

The accessibility index A_{jt} in Equation (2) measures treatment exposure to transport infrastructure improvements. There are of course other ways of measuring this exposure. Alternatives used in the literature include whether an area is crossed by a highway (e.g. Chandra & Thompson, 2000; Faber, 2014; Michaels, 2008), kilometres of roads within a given area (Melo et al., 2010; Duranton & Turner, 2011), distance to closest highway (Baum-Snow, 2007; Ghani et al. 2016), number of radial roads from a city centre crossing a given area (Baum-Snow, 2007, 2010; Baum-Snow et al., 2016), or the amount of public expenditure in an area on road infrastructure (Fernald, 1999). In our context, the major road network is already very developed and dense (49,816 km long in 1998, in a land area of about 230,000 km^2) and does not expand much during our study period (increasing by

0.87% to 50,250 km by the end of 2008). We are also using small geographical areas and plant level data. This means that measures based only on whether a unit is crossed by a road, or changes in the number of road kilometres, or other simple indicators are unlikely to exhibit much variation (or are meaningless, in the case of plant level data). Proximity to roads is one viable alternative indicator. However, when new road scheme location is endogenous, distance to a new road is a poor indicator of network exposure, because it is infeasible to separate the influence of new transport infrastructure, from the influence of the place in which the new transport infrastructure is located. In fact, as discussed in Section 3.1, we view distance to the road schemes as a potentially important control variable, not an index of treatment.

Using an accessibility index (Equation (2)) has the key advantage that it varies continuously over space in ways that are partly unrelated to distance to improvements.[15] This helps identify the effects of transport improvements separately from the specific advantages or disadvantages of sites chosen for improvements. It also means we can potentially observe the degree of treatment 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). Other studies have employed similar indicators for this purpose, e.g. Graham (2007) uses cross sectional variation (but not changes) in accessibility. Holl (2012 and 2016) and Donaldson and Hornbeck (2016) like us, use changes in accessibility (or 'market potential'/'market access') as a treatment variable, but do not exploit the variation in accessibility within localised areas for identification, which is our key contribution.

[15] Donaldson and Hornbeck (2016) also claim that a continuous market access index of this type avoids the problems of spillovers in the effect of treatment that are inherent in designs with discrete, neighbouring, treatment and control areas. With an accessibility/market access index, all areas are treated to a greater or lesser degree. This is arguable, since spillovers between neighbouring areas with big accessibility gains to neighbouring areas with smaller accessibility gains will also lead to biased estimates of the accessibility impact.

4 Data sources and setup

4.1 Geographical units and area controls

Our analysis is based on plant level micro data. We have detailed information on the location of plants (postcodes, equivalent to around 17 houses or a medium sized plant) and can link this data geographically at various levels using the Office for National Statistics (ONS) National Statistics Postcode Directory. For most of our analysis we work with aggregates for approximately 10,300 electoral wards. Wards are defined to have roughly the same number of electorate and are geographically small in dense areas. We use wards as defined in 1998.

To construct ward level control variables we use the GB Census 2001 to calculate the share of population aged 15-64 with higher education, mean age of population, share of population living in social housing and the rate of unemployment. We also calculate straight line distances from each ward to the nearest scheme (undertaken at any point during our study period) using GIS and the dataset of transport schemes described in 4.3.

4.2 Firm data

Data for the analysis of employment and plant counts (number of establishments) at ward level, and for analysis of plant level employment, is from the Office for National Statistics (ONS) Business Structure Database (BSD)[16] accessed through the UK's Secure Data Service. We use data from 1998 to 2008 to construct dependent variables and data from 1997 to construct the accessibility index. The BSD contains a yearly updated register of the universe of businesses in the UK covering about 98% of business activity (by turnover). For consistency with our productivity data – described below – we do not use data for years past 2008.

[16] Office for National Statistics, *Business Structure Database, 1997-2014: 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

The smallest unit of observation is the establishment or plant ('local unit', LU), but there is also information on the firm to which the plant belongs ('reporting unit', RU). The dataset provides detailed information on location (postcode), sector of production (up to 5 digit SIC) and employment in plants. We can calculate employment and number of establishments at any geographical level aggregating up from postcodes.

For the productivity regressions, we use the ONS Annual Respondents Database (ARD).[17] The ARD holds responses to the Annual Business Inquiry (ABI) completed by a stratified random sample of units, extracted from the BSD (see Criscuolo et al, 2003). The survey covers balance-sheet information including gross output, value added, wages, intermediate inputs, employment, industry, and investment for both manufacturing and services. As no reliable yearly firm capital data exists for our period of analysis, we cannot report results for TFP. Instead, we use per-worker values of the balance sheet variables (turnover, output, value added and labour costs) as measures of labour productivity. We use the EU KLEMS Deflators (base 1995) to deflate the balance-sheet data. 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 employment (for example 90% of UK manufacturing employment). The Annual Business Inquiry ended in 2008 and was replaced by a different survey, resulting in a discontinuity in the series. Hence we do not look at effects past 2008.

Like the BSD, ARD reports information for LUs and RUs. Balance-sheet data is available at the RU level, location and employment at the LU level. Decisions have to be made to assign balance sheet data down to LU level. For output and other economic variables expressed in per-worker values,

[17] Office for National Statistics, *Annual Respondents Database, 1973-2008: 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

we assign the mean for a RU (firm) to each LU (plant). Total output and other economic variables are apportioned to LUs in proportion to the plant employment. Clearly this assumes that per-worker productivity is equal in all plants within a firm (see e.g. Criscuolo et al, 2012). There is therefore a risk of attenuating the estimated impacts of transport improvements, if not all plants within a firm are affected, because the productivity changes of the affected plants will be combined with changes in unaffected plants when calculating firm level value-added per worker. Results are robust to looking at single plant firms (singletons), for which these allocation issues are not relevant.

4.3 Road data and origin-destination matrix construction

4.3.1 The road schemes and road network data

Information on completed road schemes for the British major roads network comes from information provided by the Department for Transport (DfT) and other sources including The Highways Agency, the Motorway Archive, Transport Scotland, Wikipedia and other web based sources. We consider improvements carried out on trunk roads, principal roads (class A) and motorways. These roads represent only 13% of total road network length, but 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 these are the schemes we expect to have a substantial effect on travel times between wards.

An initial pool of 75 schemes from 1998 to 2007 covered construction of new junctions, dualing, 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 that we can identify in the data, and which usually have a substantial effect on travel times between wards. Restricting attention to new road construction leaves us with 31 road schemes, which are listed in Table A1 in the Appendix Tables and Figures. 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 January 1998 and January 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 1 shows the location of the schemes we consider. Projects are scattered all over Britain. Figure A1 in the Appendix Tables and Figures shows the complete major road network in 2008.

This information on new road links is combined with a snapshot, GIS road network for 2008 provided by the DfT. This DfT road network is generalised and covers only major roads. Combining the road scheme and network data sources allows us to reconstruct the major road network, complete with location, length and travel speeds of the road links, for each year between 1998 and 2008. We start with the 2008 network and locate all the road links belonging to each of the 31 schemes described above and listed in Table A1. By working backwards in time and deleting the new links opened in every year, we reconstruct the network as it was at the beginning of each year back to 1998.

4.3.2 Origin-destination travel times and accessibility index construction

The essential first ingredient in our accessibility index $A_{jt} = \sum_{k \neq j} a(time_{jkt}) w_{k0}$ from Equation (2) is the bilateral, unit-to-unit road travel times ($time_{jkt}$). In our set up, the geographical units j and k are electoral wards and we calculate the ward-to-ward, origin-destination (O-D) minimum travel times at the beginning of each year using our GIS road network. These O-D travel times are derived using the ArcGIS Network Analyst optimal routing algorithm, with analysis of travel times along the approximately 17,000 road links. The simplified DfT road network structure means that the O-D travel time calculation assumes all intersections between links are junctions and ignores minor roads, forbidden turns and one way systems. When computing the O-D matrix we apply a limit of 75 minute total drive time. This limit is innocuous, given that our subsequent analysis focusses on the impacts of local differences in accessibility among wards close to the new road schemes, and differences in accessibility of remote destinations amongst relatively close neighbouring wards will be very small.

Unfortunately our 2008 DfT road network has no information on link journey times, so the journey time on individual links in the network – which form an input into the O-D travel time matrix calculation - comes from another edition of the DfT GIS road network from 2003. For links opened after 2003, we impute link journey times. The journey time calculation and imputation methods are discussed in Appendix B. Note that the set-up of the network and method of calculation of the ward-ward minimum travel times means that all changes in minimum travel times from year to year are due to the construction of new road links.

We calculate $time_{jkt}$ in (2) using minimum journey times routing along the transport network in each year. Transport improvements change the structure of the network and this changes journey times between some origin units j and destinations k. This in turn changes the accessibility index.

There are three potential channels for these changes in journey times. Firstly, a transport improvement that involves a journey time reduction on a road link *p-q* will have a first order effect on the time between *j* and *k* if the quickest route between *j* and *k* passes along the link *p-q* both pre and post-improvement. Secondly, the quickest route between *j* and *k* may not use link *p-q* pre-improvement, but switches to use the link *p-q* post-improvement because of the reduction in journey time. Lastly, second order effects arise when the quickest route between *j* and *k* does not use link *p-q* pre- or post-improvement, but other traffic switches to use link *p-q*, which reduces congestion on the quickest route between *j* and *k.* In our empirical work we exploit only the first two of these channels. We ignore second order effects because our network data does not allow us to observe travel time changes induced by changes in congestion resulting from improvements.

With the time-varying ward-to-ward O-D minimum journey time matrix in hand, we need to specify the proximity function $a(time_{jkt})$. This is a decreasing function of the minimum ward-to-ward travel time and, in line with common practice, we use a simple inverse-time weighting scheme for our main specifications ($a(time_{jkt}) = time_{jkt}^{-1}$). We show alternatives in robustness checks.

Finally, when constructing the accessibility index for our main regressions, we use workplace-based employment in destination wards (from the BSD) as weights (w_{k0}), measured in 1997 before the start of our estimation sample. We show results too for alternative residential population weights, and constant weights. Note, when we aggregate the components up to form the accessibility index, we exclude location *j* from its own accessibility index, to mitigate potential endogeneity problems.

5 Results

5.1 Descriptive statistics

Table 1 summarises the changes in (log) accessibility over the 1998 to 2008 period. The index is derived from the road network infrastructure changes and uses 1997 destination ward employment weights (calculated from our BSD data). The average accessibility change over the whole of Britain was only 0.34%, with a standard deviation of 1.22% and a 90th percentile of 0.79%. However, accessibility changes are bigger for wards closer to new road schemes, because it is in these wards that the new road links make the most difference to the minimum journey times to all potential destinations. It is the variation within these narrower distance bands from which we identify the effects in our regression analysis. Within 10 km of a scheme the mean change is 1.18% and the 90th percentile is 3.16%. Within 20 km, which we use in our base specification, mean accessibility change is 0.83% and 90th percentile is 1.97%. Within 30 km these figures are 0.66% and 1.71%, respectively.

As discussed in Section 1, alternative destination weights such as employment and population and different distance decay functions can be used in constructing the index. It is tempting to try to use these different indices to measure access to specific economic inputs or markets. However, the similarity in the spatial distribution of population, employment and other economic variables means that, in practice, indices computed using different weights are very highly correlated. Appendix Table A2, shows the summary statistics for the 1998-2008 changes in various alternative indices, and their correlation between the 1998-2008 changes in our preferred index using 1997 destination ward employment weights (for the 1-20km band). The means and standard deviations for indices with different destination are nearly identical, and differences due to alternative weighting schemes arise simply through a matter of scale (higher weights imply aggregation of employment over a shorter range). Evidently, all the correlations are above 0.9, and most are above 0.98 and it will be infeasible to identify their separate contributions in a regression analysis.

Figure 1, Figure 2 and Figure 3 illustrate the spatial relationship between road schemes and the accessibility increases they cause. Figure 1 shows new roads and major improvements. Figure 2 shows accessibility improvements between 1998 and 2008 for our main estimation sample (wards within 1-20 km of new links). Figure 3 zooms in on wards in the vicinity of the M6 Toll opened in 2003 to the north east of Birmingham. The original M6 motorway is a major route between London and the North West of England and goes through the centre of Birmingham, the old route involving a series of viaducts and interchanges which are subject to heavy congestion. The M6 Toll is indicated by the bold line and was constructed as an alternative to the main M6 motorway to relieve congestion for traffic travelling to and from the south and north of England. The shaded areas in the figure illustrate the employment accessibility changes in wards in our main estimation sample within close proximity to the road scheme, between 1 to 20km. Clearly the effect of the scheme on accessibility varies considerably across wards in the vicinity of the same improvement. As detailed in Section 3.1, our identifying assumption is that this variation in accessibility changes across wards within narrow distance buffers around these links are incidental to the policy aims of the road scheme, and can be treated as exogenous (especially conditional on additional scheme-specific fixed effects, and other local time trends). Our preferred distance buffer in the regression analysis is 1-20 km which, as can be seen from the map scale is a small distance compared to typical road links in our data.

One concern from Figure 3 might be that changes in accessibility induced by transport improvements are correlated with pre-treatment trends in area characteristics. Table 2 provides a direct test of this, showing results from the regression of 1981-1991 changes in area characteristics on the 1998-2008 log accessibility changes, conditional on our baseline controls: 1998 accessibility, distance to nearest scheme and scheme specific dummies. We do not have information on the characteristics of firms prior to our sample period, hence we rely on residential population demographic characteristics from the 1981 and 1991 population Census. These regressions provide balancing tests. If our identification strategy is valid, we should not see impacts from transport improvements on changes prior to the improvements occurring. Results show that this is the case. Of the 24 coefficients reported in Table 2, none are significant at conventional levels, suggesting that accessibility and pre-transport-improvement trends in local characteristics are uncorrelated. Overall these findings suggest that, within our preferred 1-20km distance band, changes in accessibility are uncorrelated with pre-treatment trends.

For reference, Appendix Table A3 provides further descriptive statistics for the number of plants and employment by sector in the wards in our estimation samples.

5.2 Ward-level employment and plant count regressions

Results, presented in Table 3, show the relationship between accessibility and employment (top panel) and numbers of plants (lower panel) at ward level. Each coefficient relates to a separate regression. Employment and plant counts come from BSD. The accessibility index uses base-period employment as destination weights. The main results in columns 1-3 use data on wards within 1-20 km of road schemes for various model specifications. Columns 4 and 5 present the same specification as column 3, but applied to samples within 1-30 km and 1-10 km of schemes, respectively. Standard errors are 'clustered' at ward level, to allow for arbitrary intra-ward correlation over time. Alternative higher-level clustering schemes (e.g. based on Census District) 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 the basic ward fixed effects regression of Equation 3 and shows the relationship between road network-induced changes in accessibility and changes in employment or number of plants inside the 1-20km band. The point estimate is large, highly statistically significant for plant counts, but only weakly significant for employment. A 1% improvement in accessibility increases the number of plants by 0.4% and employment by 0.3%. The greater precision (smaller standard errors) in the plant-count regressions is to be expected, given that employment will be subject to greater survey measurement error. In column 2 we add nearest-scheme specific trends to the ward fixed effects regression (i.e. it includes interactions between nearest-scheme dummies and a linear time trend). Column 3 introduces a time trend interacted with distance to the scheme, with a dummy for years after scheme opening, and with accessibility in 1997, all to allow for differences in trends across space close to the schemes, and general post-operation changes. Introducing these additional control variables improves the statistical significance (all significant at the 5% level or better) and shifts the point estimates around slightly, but not by much relative to standard errors. The effects on employment are generally slightly larger than the effects on number of plants, although again the differences are not large relative to standard errors. Results do not change substantively, for specifications (not reported) where 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.

Expanding the sample cut-off distance around the schemes in column 4 leaves the results largely unchanged. Reducing the distance to within 10 km of schemes in column 5 leads to smaller coefficients, and for employment, results are insignificant (although again less than one standard error below the largest point estimates in previous columns). The weaker results in the 10km band might seem surprising given that the variation in accessibility is biggest closes to the schemes (Table 1). However, close to schemes, a greater proportion of this variation is due to measurement error in the travel time calculations, since we do not have detail on minor roads journeys from wards to the new major roads (which is an additional reason for excluding wards within 1 km, see Section 3.1). Sample sizes are also smaller.

The headline story from these results is that accessibility changes induced by road improvements drive up local employment and the number of local plants, with an elasticity of around 0.3-0.4. These estimates appear quite large, but remember that the changes in accessibility are small (see Table 1). On average, within 20 km the mean change in accessibility was only 0.83%, implying an increase in employment of 0.25% and an increase in plants of 0.33%.

Table 4 shows results by broad industrial sector for specification 3 in Table 3.[18] The results suggest that most of the action on employment and plants (in terms of the size of the effects) comes from producer services, land transport and 'other' sectors (a residual category that includes the primary and public sectors). The number of plants in the manufacturing sector also responds strongly, although this does not show up in the employment figures (presumably implying that new plants in the manufacturing sector are small). In additional results not reported, we find, as before, that expanding to a 30 km distance band leaves results unchanged, while reducing the area to within 10 km leads to smaller, less precise estimates. Both the land transport and producer services effects are consistent with a story in which road improvements lower transport costs for intermediate inputs and business travel and stimulate employment in the logistics sector. Additional results which break down the transport sector effects (not reported) confirm that the strongest positive

[18] 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 and energy 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.

effects come for employment and plants within the land transport sector, and specifically within the road freight and cargo handling sectors.

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

Table 5 presents the results of regressions (specification as for column 3 of Table 3) using alternative definitions of the accessibility index in Equation 2. Columns 1-5 use alternative destination weights (w_{k0}), instead of the 1997 BSD ward employment used in the previous tables: counts of post office residential delivery addresses (taken from the ONS National Statistics Postcode Directory) in column 1; ward population from the 1991 GB Census in column 2; ward plant counts from the BSD in column 3; ward level (residence-based) employment from the 1991 census in column 4; and no weights in column 5, so the accessibility index is simply the sum of the inverse minimum travel time to all potential destinations (within 75 mins). It is evident from the similarity in all these results that changing the definition of accessibility is immaterial, which is not surprising given the high correlations shown in Appendix Table A2. We have also estimated specifications using destination ward BSD employment in the current year as weights, instead of 1997 employment weights. When instrumenting this accessibility index with the version based on 1997 employment weights, results are essentially unchanged from the reduced form versions reported so far.

Columns 6-8 change the proximity function $a(time_{jkt})$ in Equation (2) from $(time_{jkt})^{-1}$. Columns 6 and 7 use alternative parameters in place of -1 in the time-decay function. Column 8 switches to an exponential weighting function. The coefficients are statistically significant regardless of the time function used, although the coefficients change quite dramatically, reflecting differences in the scale of the accessibility index as the weighting scheme changes. An important lesson here is that elasticity estimates based on one form of accessibility index clearly cannot be used to make predictions about the effects of changes in accessibility based on alternative functional forms. However, if we standardise the effects (multiply by the standard deviation of the accessibility variables reported in Appendix Table A2) we find a more stable pattern of effect sizes. For example, for the accessibility index using inverse time weighting (with a parameter of -1) the standardised effect on employment for a one standard deviation increase in accessibility is 0.365 x 1.97 = 0.72% (from Table 1 and column 3, Table 3). For the exponential time decay function in column 8 of Table 5, the standardised effect size is 2.13 x 0.46 = 0.98%.

Other robustness checks included estimating the regressions using samples within distance bands e.g. 11-20, 21-30km, so we are comparing firms that are at similar distances to the road schemes. We find strong effects within these bands, indicating that our results are not driven simply by firm relocation to sites close to schemes. We also ran regressions with differential time trends according to initial level of employment or number of plants (by interacting initial levels of the dependent variable with time trends), to allow for potential mean reversion, but find the results substantively unchanged.

In sum, there is no evidence that the results are substantively sensitive to changes in the definition of the accessibility index, or to other changes in specification.

One concern when using data on closely spaced wards and firms is that the unobservables in our regression models may be spatially autocorrelated, leading to biased standard errors and incorrect inference. Given we include ward fixed effects, scheme specific trends, and distance to scheme trends as control variables, this problem is likely 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 sectional patterns. Nevertheless, direct tests of the residuals from our regressions using Moran's I statistics take tiny values (less than 0.01) showing no evidence of important spatial autocorrelation in the residuals. In addition, re-running our main regressions with standard errors clustered at a larger geographical level, the Census district, makes 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. To do this, we estimate the effect of log accessibility on log plant employment using plant level data from BSD for 1998-2008.

Table 6 presents the key results using a similar structure to Table 3. Additional control variables in these plant level regressions are industry-year dummies (using the 6 broad sectors used for the sector-specific results above). As before, standard errors are clustered at ward (i.e. treatment) level.

The baseline plant fixed effects estimate in column 1 indicates plant size reductions in response to changes in road transport accessibility, although the coefficient is not statistically significant. Adding in further controls for local time trends in columns 2 and 3 increases the coefficient slightly (in absolute terms) and improves the statistical significance. The general picture remains the same when we change the sample to plants within a narrower area (1-10 km distance) or wider area (1-30 km distance) in columns 4 and 5. We also introduced interactions of linear time trends with GB Census characteristics as described at the end of Section 3.1, but this made little difference. A 1% increase in accessibility is associated with a 0.06-0.09% reduction in employees. Sector specific results (not reported) do not offer any strong insights, although interestingly there are (insignificant) positive employment effects for firms in the transport sector.

The evidence overall suggests that incumbent plants exposed to accessibility changes as a result of transport improvements are, on average, reducing employment. Read in conjunction with Table 3, the clear implication is that transport improvements boost local employment through a strong net gain in the number of plants and associated employment, while existing plants cut back (marginally) on employment. These employment cuts within incumbent plants could be due to increases in the price of labour relative to other inputs, causing substitution from labour to those inputs, or reductions in scale due to local factor price increases induced by demand from new plants entering the locality. The productivity results in section 5.5 shed more light on this question.

5.5 Productivity and other production related outcomes

Although we find a negative response for existing firms on the employment margin, these firms may experience productivity gains if lower transport costs allow reorganisation resulting in increased output per worker. We explore this directly by using various output and input-related balance sheet variables in the Annual Respondents Database data as dependent variables (described in Section 4.2). Specifications are similar to those in Table 6 supplemented to include a dummy and a trend specific to single plant firms or 'singletons' (available for the ARD, but not BSD). The regressions are weighted using ward-by-year specific employment weights derived from the BSD, in order to make the ARD sample more representative of the spatial employment structure in the BSD population. The key results for all sectors pooled together are in Table 7. We restrict attention to the 1-20 km distance band.

The headline story from the coefficients in the top panel is that there is very little impact on plant total output, total labour costs or the wage bill. There is an increase in inputs of goods, services, materials and road transport services, which in turn is associated with a reduction in value-added (which is gross output minus purchases of materials, goods and services). However, none of the coefficients is precisely estimated and we cannot rule out zero impact on any of these outputs and inputs. Once we switch to per-worker values in the second panel, as indicators of labour productivity, we find some stronger positive effects. Total labour costs and wage bill per worker (i.e. mean wages) increase, output per worker increases, as does the amount of non-labour inputs used per worker. The increases in the economic variables measured in per-worker terms are qualitatively consistent with the reductions in employment shown in the BSD data in Table 6. The overall picture from Table 6 and Table 7 is one in which output remains constant, but worker productivity is increased, with corresponding increases in wages and substitution from labour to non-labour inputs. Note however, that we can detect no labour productivity increases measured in terms of value-added per worker, because increases in output per worker are accompanied by increases in goods, services and materials purchases. Again, it is important to remember that the average mean accessibility change within 1-20 km is 0.83%, so these coefficients imply induced output per worker and wage effects of around 0.25% as a result of these schemes. Sector specific results (not reported) suggest that the strongest effects are in the manufacturing and consumer services sectors, which were some of the least responsive sectors in terms of aggregate employment in Table 4, although the picture is generally quite mixed and the effects imprecisely measured. The effects are also concentrated in larger firms with more than 10 employees.

Additional analysis for these productivity-related outcomes aggregated to ward level in the first panel of Table 8, suggests that output and input costs effects are also large at the ward level, consistent with the increases in the number of plants documented in Table 3. However, these effects are often too imprecisely measured to be informative, with the exception of expenditures on input goods, services, material and transport services which are highly responsive. A 1% increase in ward accessibility is associated with a 2-2.75% increase in total expenditure on non-labour inputs and transport services. There are also increases in the per-worker mean outputs and input costs, which are consistent with the output per-worker increases in incumbent firms (Table 7).

Taken 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/energy. At the same time, employment reductions occur in existing firms.

6 Conclusions

This paper estimates the impacts of recent improvements to the road network in Britain on a range of firm productivity-related and employment-related outcomes, using micro data at a very detailed geographic scale. Our results contribute to the evidence on the effects of transport on area and firm level economic outcomes, and provide unique evidence on the effects of relatively incremental changes in the network that are relevant to policy in developed economies.

We measure road transport access with a continuous index of accessibility based on minimum journey times, imputed from GIS network analysis. Our data-intensive research design uses policy evaluation methods applied to rich panel data. 'Treatment' as a result of road improvements is captured by changes in this accessibility index over time, in response to 31 new road link schemes over the 1998-2008 period. We identify the causal effects of changes in accessibility, from variation in this treatment amongst firms close to new road schemes, which mitigates biases arising from the location of schemes being potentially endogenous due to policy targeting. Places closest to the schemes are also those that experience the biggest changes in accessibility, because it is routes from these places that are most likely to make use of the new road links. Focussing on places close to schemes therefore makes the best use of the variation in accessibility generated by road network changes.

From our ward-level regressions we find strong evidence that road infrastructure improvement schemes increase the number of firms and employment in places that gain through better access to and along the road network. A 1% improvement in accessibility leads to about a 0.3-0.4% increase in the number of businesses and employment. The estimates range between zero and 1% according to sector and specification. Evidence from our plant level estimates suggests that, at the same time, incumbent firms shed workers, so employment gains must come about through firm entry. We detect output per worker and wage increases for incumbent plants: these plants maintain output whilst cutting workers, as they substitute to goods and services inputs. One theoretical story that is consistent with our findings (there may be others) is that accessibility improvements attract firms that benefit the most from transport accessibility, bidding up local wages relative to other input prices and transport costs. In response, incumbent firms (those that do not exit the area) substitute in-house labour with purchases of goods and services inputs. The sectoral picture is less clear, but reveals aggregate employment effects dominating in the producer services, transport and administrative sectors.

Our evidence does not shed light on whether these effects arise because new roads improve access to output markets, intermediate inputs or workers, or just reduce travel times in general, and we treat our accessibility index simply as an indicator of policy treatment. This accessibility index is identical in structure to those used previously in transport project appraisal and in the trade and spatial economic literature, where claims are sometimes made about the ability to distinguish market access effects from other changes. However, we show that accessibility indices constructed to measure access to destination employment, residential population, or simply the number of destinations are all highly correlated and yield similar results.

In common with all empirical work that estimates causal effects from statistical comparisons across time and place, it is impossible to know for sure whether these employment increases are additional to the economy as a whole. Our design ensures that the effects we observe are not estimated from simple displacement to areas near new road schemes from areas elsewhere in Britain, because we only estimate from variation within areas close to new road schemes. However, it is fundamentally impossible to rule out more subtle effects where firms relocate precisely in response to the accessibility changes *within* the vicinity of the road schemes. Even if the local employment gains we estimate are not all additional, they are still important in terms of understanding the expected impact of transport projects on local development – a fundamental concern for many policy makers. If we were to assume that the local employment gains are additional, they appear substantial when roughly translated into the expected increase in GDP as a result of the public investment in new roads (see Appendix C). An upper bound on the estimate is around £4.2 billion in present value terms. A lower bound is around £1 billion. 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.

References

Ahlfeldt, G. M. and Feddersen, A. (2010) "From periphery to core: economic adjustments to high speed rail". London School of Economics & University of Hamburg. LSE Research Online.

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

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.

Banerjee A. & Esther Duflo & Nancy Qian, (2012). "On the Road: Access to Transportation Infrastructure and Economic Growth in China," NBER Working Papers 17897, National Bureau of Economic Research, Inc.

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.

Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., & Zhang, Q. (2016a). "Roads, railroads and decentralization of Chinese cities". The Review of Economics and Statistics.

Baum-Snow, N., Henderson, J. V., Turner, M. A., Zhang, Q., & Brandt, L. (2016b). "Highways, Market Access and Urban Growth in China," SERC Discussion Papers 0200, Spatial Economics Research Centre, LSE.

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.

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.

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

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

Department for Transport (1997) "A new deal for trunk roads in England: Understanding the new approach" Department for Transport, London.

Department for Transport (2009) Highways Agency Framework Document, Department for Transport, London.

Donaldson, D. (2014) "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure", American Economic Review, forthcoming.

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

Duranton, G. 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", Review of Economic Studies 79 (4): 1407-1440.

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

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.

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

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

Faber, B. (2014), “Trade Integration, Market Size, and Industrialization: Evidence from China’s National Trunk Highway System”, Review of Economic Studies, 81(3): 1046–1070.

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

García-López, M.-Á., Holl, A. & Viladecans-Marsal, E., (2015). "Suburbanization and highways in Spain when the Romans and the Bourbons still shape its cities," Journal of Urban Economics, 85: 52-67.

García -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.

Ghani, E., Goswami, A. G., & Kerr, W. R. (2016). “Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing”. The Economic Journal, 126(591): 317-357.

Gibbons, S. Lyytikainen, T., R. Sanchis-Guarner, H.G. Overman, J. Laird (2010) Evaluating the productivity impacts of road transport schemes, Report for Department for Transport, August https://www.researchgate.net/publication/48910180_Evaluating_the_productivity_impacts_of_road_transport_schemes_report_on_pilot_study_findings

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 for Transport.

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

Graham D. J. (2007). 'Agglomeration, productivity and transport investment', Journal of Transport Economics and Policy, 41(3) 317-343.

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. (2012). "Market potential and firm-level productivity in Spain", Journal of Economic Geography, 12 (6): 1191-1215.

Holl, A. (2016) "Highways and productivity in manufacturing firms", Journal of Urban Economics, 93: 131-151.

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

Jedwab R. and Alexander Moradi (2016). The Permanent Effects of Transportation Revolutions in Poor Countries: Evidence from Africa”, Review of Economics and Statistics, 98(2), 268-284.

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.

Vanables, A.J., J. Laird and H.G. Overman (2014). Transport investment and economic performance: Implications for Project Appraisal, Department for Transport, London

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, Han & Li, Zhigang, (2013). "Road investments and inventory reduction: Firm level evidence from China," Journal of Urban Economics, Elsevier, vol. 76(C), pages 43-52.

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.

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.

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

Redding, S. J., Turner, M. A. (2015). “Transportation Costs and the Spatial Organization of Economic Activity.” in Gilles Duranton, Vernon Henderson, and William Strange eds. Handbook of Urban and Regional Economics, Vol. 5: Elsevier.

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

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.

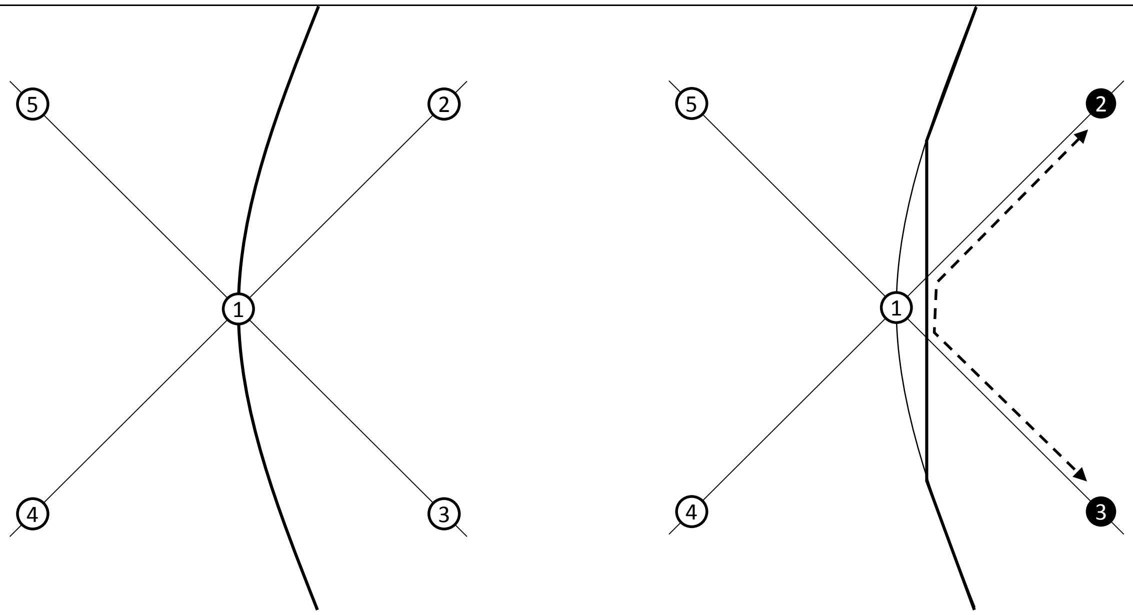
U.S. Department for Transportation (2012), Research and Innovative Technology Administration Bureau of Transportation Statistics, National Transportation Statistics.

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

Vickerman R., 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 A: Illustration of source of identifying variation in local accessibility



5 locations, trunk road and connecting roads pre-improvement

5 locations, trunk road and connecting roads post-bypass construction. Shortened link between locations 2 and 3 generates accessibility changes in 2 and 3. Other link times between locations and accessibility in 1-5 unchanged.

The figure above illustrates how local improvements to the strategic road network can induce accessibility changes that are plausibly unrelated to the road's intended purpose. The circles represent centres of employment and the lines connecting roads. The thick line through 1 represents a trunk road which is improved by a bypass in the second period. Note that as a result of this change, the only improvement in journey time on local connections between these employment centres is on journeys between 2 and 3. This improves accessibility in locations 2 and 3, but nowhere else in this locality. Our identifying assumption is that the choice of route for the bypass in period 2 is related only to technical considerations (e.g. route length, feasibility and cost) not factors affecting future employment growth in locations 2 and 3. Our search of the literature related to the planning and evaluation of benefits of the schemes in our study area and period suggests that this is invariably the case (annex discussing examples is available on request) and the balancing tests provided in the main text provide empirical support for this assumption.

Appendix B: Further details on construction of road travel times

The construction of the road network in 1998-2008 uses information on new road links which is used to modify a GIS road network provided by DfT for the year 2008. The 2008 GIS-network contains all the major road links existing in the beginning of 2008, and includes information on several characteristics of the road links but lacks information on travel times. Travel times are available for the 2003 network, and we use them for the whole period.

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 count point code (a link identifier). 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. For genuinely new road links, i.e. roads for which we do not have an alternative minor road flowing in parallel, we simply remove these links as we move backwards to recursively reconstruct the network. Other projects involve 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 typically correspond to bypasses which relieve traffic congestion from villages and usually run in parallel to an existing alternative minor road. These 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 these 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 optimal travel times between locations, we use data on travel times from the 2003 generalised primary road GIS-network provided by DfT. These travel times are available for each link of the network and modelled from traffic flow census data using the Road Capacity and Costs Model (FORGE) component of the National Transport Model (NTM).[19]

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. 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. This has the additional advantage that the variation in travel costs over time stems only from new additions to the network.

Links constructed or upgraded after 2003 do not appear in the 2003 network and thus for these we do not have information on travel time from FORGE. The model that we use to estimate journey times for links opened 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. The regression predicts speeds from the FORGE reasonably well (R-squared = 0.76). We then use the regression results and link characteristics in the 2008 network to

[19] 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.

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. 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. Moreover, journey times for the links may be imprecise. Changes in accessibility must therefore be regarded as approximate. This measurement error means our estimates of the effect of accessibility could be attenuated.

To partially address concerns about measurement error in the accessibility index, we cross checked a sample of times and accessibility measures against estimates derived from Google maps, using the STATA 'traveltime' 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.

Appendix C: Back of the envelope representation of potential GDP gains

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 0.013% increase in total employment from a year's investment in major road transport network improvements (using the elasticity of 0.37 from Table 3). 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,[20] 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.

[20] Own calculation based on ONS 2011

Figures

Figure 1. Location of road schemes opened 1998-2008

Figure 2. Changes in log accessibility from 1998 to 2008 – all wards

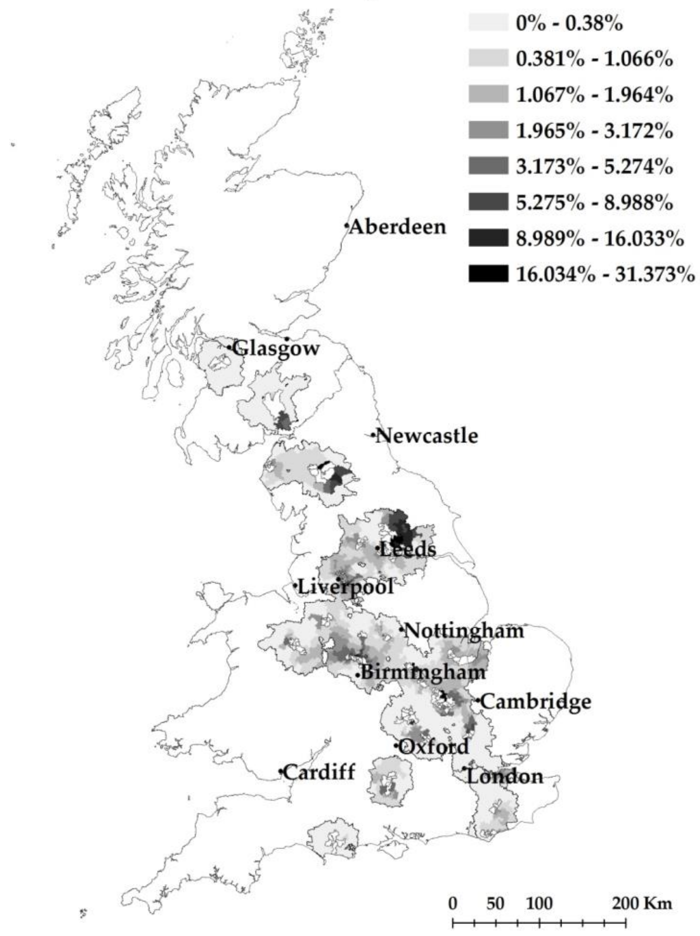
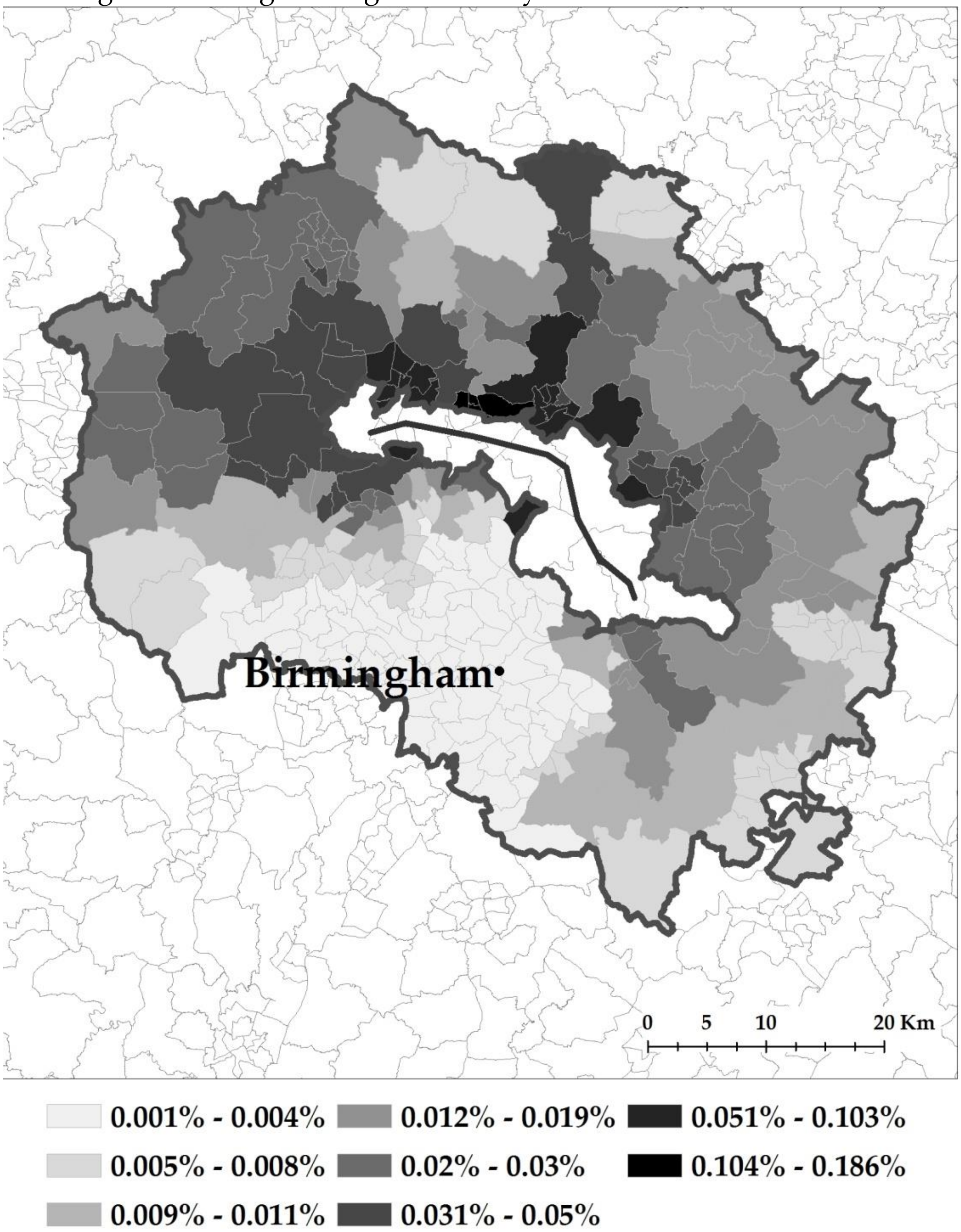


Figure 3: Changes in log accessibility from 1998 to 2008 – M6 Toll



Notes: Map shows changes in accessibility from 1998 to 2008 for wards within 1-20km from M6 Toll.

Tables

Table 1. Change in log accessibility 1998-2008 based on road network minimum travel time changes and BSD 1997 employment weights.

	Wards	Mean	Std. Dev	90th percentile	Max	Proportion of zeroes
1997 BSD employment weights						
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: Balancing regressions: Association of 1981-1991 ward pre-trends with 1998-2008 changes in accessibility, within 1-20km band

	(1)	(2)	(3)	(4)	(5)	(6)
	UK born	Age 17-29	Age 30-49	Age 50-64	Age 64+	Age 30-44 w. degrees
Log accessibility	0.0136	0.0412	0.00162	0.0394	0.00146	0.0192
	[0.0114]	[0.0296]	[0.0213]	[0.0338]	[0.0258]	[0.0212]
R-squared	0.273	0.044	0.080	0.129	0.144	0.105

	(7)	(8)	(9)	(10)	(11)	(12)
	Owner occupied	Council tenants	Econ. Active	Working	Retired	Students
Log accessibility	-0.0484	0.0732	0.0491	0.0388	-0.0174	0.0188
	[0.0769]	[0.0722]	[0.0428]	[0.0469]	[0.0311]	[0.0119]
R-squared	0.098	0.164	0.212	0.256	0.150	0.054

	(13)	(14)	(15)	(16)	(17)	(18)
	Manufact.	Energy (inc mining)	Distrib. & catering	Transport	Construct.	Other ind.
Log accessibility	0.184	-0.0591	-0.0768	0.0125	0.0525	0.0525
	[0.117]	[0.0540]	[0.0532]	[0.0317]	[0.0395]	[0.0395]
R-squared	0.176	0.290	0.051	0.050	0.043	0.043

	(19)	(20)	(21)	(22)	(23)	(24)
	Profession /mangers	Int. non-manual	Junior manual	Pers. serv. /semi sk.	Manual	Unskilled manual
Log accessibility	0.0183	-0.0349	-0.0389	0.0951*	-0.0114	-0.0216
	[0.0638]	[0.0429]	[0.0581]	[0.0565]	[0.0544]	[0.0268]
R-squared	0.070	0.057	0.078	0.037	0.057	0.049

Ward level regressions of 1981-1991 changes in Census residential population shares with characteristics described in column headings on 1998-2008 change in log accessibility. Sample restricted to wards within 1-20km of a road scheme. Table reports regression coefficients and robust standard errors (clustered at Ward level). *, **, *** indicate significant at the 10%, 5% and 1% levels, respectively. All regressions include scheme dummies, log accessibility 1998, distance to scheme. Obs. 3469.

Table 3: Effect of accessibility improvements on employment and number of plants: ward-by-year level regressions; all sectors

	(1)	(2)	(3)	(4)	(5)
Employment					
Log accessibility	0.319*	0.398**	0.465**	0.451**	0.250
	[0.186]	[0.195]	[0.197]	[0.195]	[0.221]
Plant counts					
Log accessibility	0.437***	0.313***	0.292***	0.366***	0.228**
	[0.107]	[0.098]	[0.098]	[0.098]	[0.103]
Observations	38357	38357	38357	53933	16654
Wards	3487	3487	3487	4903	1514
Distance band	1-20 km	1-20 km	1-20 km	1-30 km	1-10 km
Year FE	Yes	Yes	Yes	Yes	Yes
Ward FE	Yes	Yes	Yes	Yes	Yes
Scheme trends		Yes	Yes	Yes	Yes
Controls			Yes	Yes	Yes

Notes: Table reports coefficients from ward-level regression of log employment or log plant counts on accessibility. Each coefficient is from a separate regression. Standard errors in brackets (clustered at the ward level). *, **, *** indicate significant at the 10%, 5% and 1% levels, respectively. 'Scheme trends' are closest-scheme dummy variables interacted with a linear time trend. 'Controls' are a linear trend interacted with: distance to closest scheme, a dummy for years in which the scheme is open and the initial level of (log) accessibility in 1997.

Table 4: Effect of accessibility improvements on employment and number of plants: ward-by-year level regressions; by industry sector; 1-20km band

	(1)	(2)	(3)	(4)	(5)	(6)
	Manufacturing	Construction/ Energy	Consumer Services	Producer Services	Transport	Other
Employment						
Log accessibility	-0.037	0.175	-0.284	0.720*	1.085	0.867***
	[0.584]	[0.388]	[0.318]	[0.416]	[0.723]	[0.266]
Plant counts						
Log accessibility	0.640**	0.093	-0.127	0.432**	1.074***	0.572***
	[0.274]	[0.210]	[0.165]	[0.205]	[0.416]	[0.159]
Observations	37743	38297	38356	38330	35331	38356
Wards	3481	3487	3487	3487	3398	3487

Notes: Table reports coefficients from ward-level regression of log employment or log plant counts on accessibility. Each coefficient is from a separate regression. Standard errors in brackets (clustered at the ward level). *, **, *** indicate significant at the 10%, 5% and 1% levels, respectively. All specifications as column 3 in Table 3 and control for year dummies, ward fixed effects, and linear trends interacted with: closest scheme dummies, distance to closest scheme, a dummy for years in which the scheme is open, the initial level of (log) accessibility in 1997. Manufacturing includes SIC92 2 digit sectors 15 to 37, construction/energy includes SIC92 2 digit sectors 40, 41 and 45, consumer services includes SIC92 2 digit sectors 50 to 59, producer services includes SIC92 2 digit sectors 65 to 74, transport includes SIC92 2 digit sector 60 and others includes the remaining sectors (agriculture, financial sector and public administration).

Table 5: Robustness of ward employment and plant count results to alternative accessibility measures; 1-20 km band.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Destination weights	Address	1991 Pop.	Plants	1991 Empl.	None	BSD Empl.	BSD Empl.	BSD Empl.
Decay:	$Time^{-1}$	$Time^{-1}$	$Time^{-1}$	$Time^{-1}$	$Time^{-1}$	$Time^{-0.5}$	$Time^{-1.5}$	$e^{-0.2*}$ Time
Employment								
Log access.	0.474**	0.479**	0.457**	0.482**	0.450*	0.903**	0.258*	2.130**
	[0.210]	[0.210]	[0.225]	[0.211]	[0.249]	[0.361]	[0.151]	[0.836]
Plant count								
Log access.	0.314***	0.316***	0.343***	0.318***	0.369***	0.556***	0.180**	1.311***
	[0.105]	[0.105]	[0.115]	[0.105]	[0.129]	[0.181]	[0.080]	[0.422]

Notes: Table reports coefficients from ward-level regression of log employment or log plant counts on accessibility. Each coefficient is from a separate regression. Standard errors in brackets (clustered at the ward level). *, **, *** indicate significant at the 10%, 5% and 1% levels, respectively. All specifications as column 3 in Table 3 and control for year dummies, ward fixed effects, and linear trends interacted with: closest scheme dummies, distance to closest scheme, a dummy for years in which the scheme is open, the initial level of (log) accessibility in 1997. Observations 38357 for 3487 wards in all columns.

Table 6: Effect of accessibility improvements on plant employment: plant-by-year level regressions; all sectors; BSD dataset.

	(1)	(2)	(3)	(4)	(5)
Log accessibility	-0.062	-0.069*	-0.091**	-0.071*	-0.114*
	[0.044]	[0.042]	[0.043]	[0.043]	[0.059]
Observations	9763020	9763020	9763020	12952535	4669444
Wards	3487	3487	3487	4903	1514
Distance band	1-20 km	1-20 km	1-20 km	1-30 km	1-10 km
Year-Sic FE	Yes	Yes	Yes	Yes	Yes
Plant FE	Yes	Yes	Yes	Yes	Yes
Scheme trends		Yes	Yes	Yes	Yes
Controls			Yes	Yes	Yes

Notes: Table reports coefficients from plant-level regression of log plant employment on accessibility. Each column is from a separate regression. Standard errors in brackets (clustered at the ward level). *, **, *** indicate significant at the 10%, 5% and 1% levels, respectively. 'Scheme trends' are closest-scheme dummy variables interacted with a linear time trend. 'Controls' are a linear trend interacted with: distance to closest scheme, a dummy for years in which the scheme is open and the initial level of (log) accessibility in 1997.

Table 7: Effect of accessibility improvements on output, productivity and non-labour inputs: plant-by-year level regressions; all sectors; ARD database; 1-20 km band

	Total labour costs	Wage bill	Gross VA	Gross O/P	Goods & services purchases	Materials	Road transport purchases
Plant totals							
Log of accessibility	0.046	0.046	-0.118	-0.000	0.082	0.310	0.275
	[0.180]	[0.180]	[0.267]	[0.190]	[0.216]	[0.365]	[0.444]
Per worker							
Log of accessibility	0.284**	0.284**	0.133	0.340**	0.317*	0.500	0.423
	[0.126]	[0.126]	[0.232]	[0.133]	[0.166]	[0.340]	[0.413]
Observations	569145	569147	538800	456599	571133	523383	405649
Wards	3456	3456	3451	3436	3455	3452	3404

Notes: Table reports coefficients from plant-by-year-level regression of log plant employment on accessibility. Each column is from a separate regression. Standard errors in brackets (clustered at the ward level). *, **, *** indicate significant at the 10%, 5% and 1% levels, respectively. All outcome variables are in logs. All regressions include industry-by-year dummies, plant fixed effects, and linear trends interacted with: closest scheme dummies, distance to closest scheme, a dummy for years in which the scheme is open, the initial level of (log) accessibility in 1997, and a dummy indicating single plant firms. Observations reports maximum number of plant x year observations. Total labour costs = employment costs + national insurance. Gross value added = turnover adjusted for stocks, insurance claims, purchases of goods and services, minus taxes duties levies plus subsidies. Gross output = turnover plus increase in stocks and work in progress. Some values were interpolated due to missing data (we include a dummy variable to identify these observations). Regressions weighted by ward level BSD employment/ARD employment to make ARD representative of BSD spatial employment structure.

Table 8: Effect of accessibility improvements on output, productivity and non-labour inputs: all sectors: ward-by-year level regressions, all sectors; ARD database; 1-20 km radius

	Total labour costs	Wage bill	Gross VA	Gross O/P	Goods & services purchases	Materials	Road transport purchases
Plant totals							
Log of accessibility	0.787	0.787	0.968	1.180	2.017**	2.748**	2.518*
	[0.708]	[0.708]	[1.381]	[1.071]	[1.001]	[1.127]	[1.307]
Per worker							
Log of accessibility	0.509	0.509	0.437	1.240	1.431*	1.108	1.667*
	[0.358]	[0.358]	[0.818]	[0.781]	[0.849]	[0.854]	[0.936]
Observations	37205	37205	37220	36558	37412	37254	36505
Wards	3481	3481	3482	3480	3485	3482	3479

Notes: Table reports coefficients from ward-by-year-level regression of log plant employment on accessibility. Each column is from a separate regression. Standard errors in brackets (clustered at the ward level). IV first stage is the first stage F-stat for the IV specifications. *, **, *** indicate significant at the 10%, 5% and 1% levels, respectively. Observations reports maximum number of ward x year observations. Total labour costs = employment costs + national insurance. Gross value added = turnover adjusted for stocks, insurance claims, purchases of goods and services, minus taxes duties levies plus subsidies. Gross output = turnover plus increase in stocks and work in progress. Scheme trends are scheme dummy variables interacted with linear time trend. Controls are a linear trend interacted with the distance to closest improvement, interacted with a dummy which is one the year and after the scheme has been opened and interacted with the initial level of (log) accessibility in 1997.

Appendix Tables and Figures

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: Road major network in 2008

Table A2: Summary statistics and correlations of changes in alternative accessibility indices with changes in our preferred 1997 employment-based accessibility index. 1-20km band

Destination weight	*Decay function*	*Mean*	*s.d.*	*Correlation with 1997 employment accessibility index*
1997 employment	$time^{-1}$	0.83%	1.97%	1.000
1997 Addresses	$time^{-1}$	0.86%	1.88%	0.995
1991 census population	$time^{-1}$	0.87%	1.89%	0.994
1997 plants	$time^{-1}$	0.79%	1.79%	0.983
1991 census employment	$time^{-1}$	0.86%	1.88%	0.993
None	$time^{-1}$	0.83%	1.59%	0.903
1997 employment	$time^{-0.5}$	0.46%	1.08%	0.986
1997 employment	$time^{-1.5}$	0.96%	2.63%	0.947
1997 employment	exp(-0.2*time)	0.20%	0.46%	0.984

Note: Number of observations is 3,487.

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.
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/energy	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
Plant count						
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/energy	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.

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

Web: www.spatialeconomics.ac.uk