UNIVERSITY OF LEEDS

A Study of the value of local bus services to society

Task III Technical Report: Econometric Analysis

Daniel Johnson
Institute for Transport Studies
October 2016

ITS

1 SOCIAL IMPACTS OF BUS SERVICES

Local bus services have a strong social dimension, impacting on individual and community wellbeing. This value is typically not captured in standard economic appraisals, in which the social impacts of policy and investment are narrowly restricted to consideration on the distributional consequences.

The work reported on here aims to quantify the strength of the relationship between the quality of local bus services and a range of social outcomes.

The work is part of a larger project building on recent work by Greener Journeys on the economic and environmental value of local bus services[1]. The project seeks to articulate and quantify the important social impacts of local bus services, which when combined with economic and environmental impacts will help show the **true value** of local bus services. The project develops an overall picture of how better local bus services can be linked to better social outcomes.

2 INTRODUCTION

The first order effects of better bus services emerge through improvements in travel times, reliability, comfort or fares. The research question explored in this work is: Is there a link between bus accessibility and social outcomes?

This report details the results of econometric models we have applied to this task to identify whether there is a systematic variation in social outcomes at the local level with the quality of the bus network.

In this work we have circumvented many of the problems associated with data availability and data quality by adopting a 'reduced form' analysis that considers the strength of the relationship between measures of public transport connectivity and social impacts as measured by the Department for Communities and Local Government's Indices of Multiple Deprivation.

We analyse a cross-sectional dataset of bus accessibility indicators, social deprivation indicators and socio demographic information to examine effect of differences in public transport (primarily bus) journey times on areas' labour market outcomes.

The analysis uses data on bus accessibility indicators and the latest Index of Multiple Deprivation (IMD)[2] indicators for 2015. We control for differences in other characteristics between areas (population density, industrial structure, car availability, etc) using information from the census.

[1] http://www.greenerjourneys.com/publication/buses-economy-ii/
[2] https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015

Section 3 discusses the data used in this work, Section 4 the methodology used, Section 5 the results and Section 6 summarises.

3 DATA

3.1 Indices of Multiple Deprivation

The Index of Multiple Deprivation 2015 is the official measure of relative deprivation for small areas (or neighbourhoods)[3] in England each with a population of around 1,500 people. The Index of Multiple Deprivation ranks every small area in England from 1 (most deprived area) to 32,844 (least deprived area).

The underlying metrics reflect a mix of economic, environmental and social indicators. They are based on 38 separate measures, organised across seven distinct 'domains' of deprivation which are combined to calculate an overall score between 0 and 100 known as the Index of Multiple Deprivation (IMD). Table 11 provides details of the measures contributing to IMD scores.

- The Indices of Deprivation can be used for identifying areas with high levels of overall deprivation to look at areas with specific issues, such as health, that may not be considered as deprived on the overall index.
- The Indices are central to the evidence base for regeneration policy in England and help target limited resources appropriately.
- Key users of the Indices are local authorities where the Indices are used to identify the local areas with the greatest level of need for support or intervention
- They can be used to make spatial comparisons between areas but also to examine how deprivation of particular areas evolve over time.

The headline Index of Multiple Deprivation is part of the Indices of Deprivation and it is the most widely used of these indices. It combines information from seven separate 'domains' measuring different aspects of deprivation to produce an overall relative measure of deprivation.

The 7 IMD categories are:

- Income – the proportion of the population in an area experiencing deprivation related to low income
- Employment – measures involuntary exclusion of the working age population from the labour market
- Health and disability - measures premature death and the impairment of quality of life by poor health

[3] The small areas used are called Lower-layer Super Output Areas, of which there are 32,844 in England. Differing in spatial scale they are designed with an average of 1,500 residents each and are a standard way of dividing up the country as used and reported in the Census.

- Education Skills and Training - measures the extent of deprivation in terms of education, skills and training in an area
- Barriers to housing and services - measures the physical and financial accessibility of housing and key local services (including GPs and convenience/supermarket stores)
- Living environment deprivation - measures the quality of individuals' immediate surroundings both within and outside the home
- Crime - measures the rate of recorded crime in an area for four major crime types The full set of IMD use 38 separate indicators, across these seven distinct domains of deprivation.

The Index units in themselves are largely not meaningful – they are based on separate measures then ranked. Each domain ranking is transposed and represented as ordinal scale (0-100).

The transformation used is as follows. For any LSOA, denote its rank on the domain, scaled to the range [0,1], by R (with R=1/N for the least deprived, and R=N/N, i.e. R=1, for the most deprived, where N=the number of LSOAs in England).The transformed domain, E, is

$$E = -23*\ln\{1 - R*[1 - \exp(-100/23)]\}$$

where *ln* denotes natural logarithm and *exp* the exponential or antilog transformation.

- 'High scores' (and low rankings) represent deprived areas, low scores (and high rankings) less deprived
- Overall IMD (again 0-100) derived from combining weighted domain scores.
- We removed access barriers from domain 6 and recalculated IMD as this includes accessibility measures.

3.2 Weightings and recalculated score

The original IMD is a weighted sum of the individual domain scores using the following weightings as reported in the first column of Table 1. The measure of deprivation used in the analysis was slightly adjusted from the published figures. As we are focusing on the link between transport accessibility and deprivation and for statistical purposes, we reworked the IMD score to only include six out of the seven original domains by removing the component of the barriers domain that related to access to services. This way we were able to isolate deprivation and accessibility separately and analyse their relationship. This reworking is shown in Table 1.

Table 1 - IMD Domain weighting

Domain	*Original Weighting*	*Re-worked weighting without services*
Income	22.5	23.6
Employment	22.5	23.6
Education, Skills and Training	13.5	14.2
Health Deprivation and Disability Rank	13.5	14.2
Crime	9.3	9.8
Barriers to Housing and services	9.3	4.9
Living Environment	9.3	9.8
Total	**100.0**	**100.0**

It is important to note that these statistics are a measure of relative economic, social and environmental deprivation, not affluence, and to recognise that not every person in a highly deprived area will themselves be deprived. Likewise, there will be some deprived people living in the least deprived areas.

3.3 Accessibility Indices

For measures of accessibility, we used DfT derived accessibility indicators (DfT, 2012) for journey times to key services identified in the 2003 Social Exclusion Unit report 'Making the Connections [4] along with access to town centres. These measures aim to provide a consistent dataset of accessibility of services at the lower super output area (LSOA)[5] within England. The aim of these figures is to help local authorities develop their evidence base for their accessibility strategies, and to support the DfT accessibility indicator on households with good transport access to key services or work.

The data consists of theoretical journey times calculated by modelling journeys between known sets of origins and destinations, using information on the road network, traffic speeds and public transport timetables.

Given correlation between alternative measures of access we select four measures including access to employment centres, GPs, Hospitals and town centres, in areas where bus services are the dominant form of public transport.

The accessibility statistics are constructed for walking/public transport and car.The public transport/walking variable primarily captures bus travel times as the main public transport (PT) mode but also includes rail. The travel time indicators measure the time taken for users

[4] Making the Connections: Final Report on Transport and Social Exclusion, Social Exclusion Unit, 2003.

[5] Lower layer super output areas (LSOAs) are used for the collection and publication of small area statistics, and have a minimum size of 1,000 residents and 400 households, but average 1,500 residents.

to reach the nearest employment centre/GP/Hospital/Town Centre by mode of transport (public transport/walking, cycle and car)[6].

3.4 Town/city classification

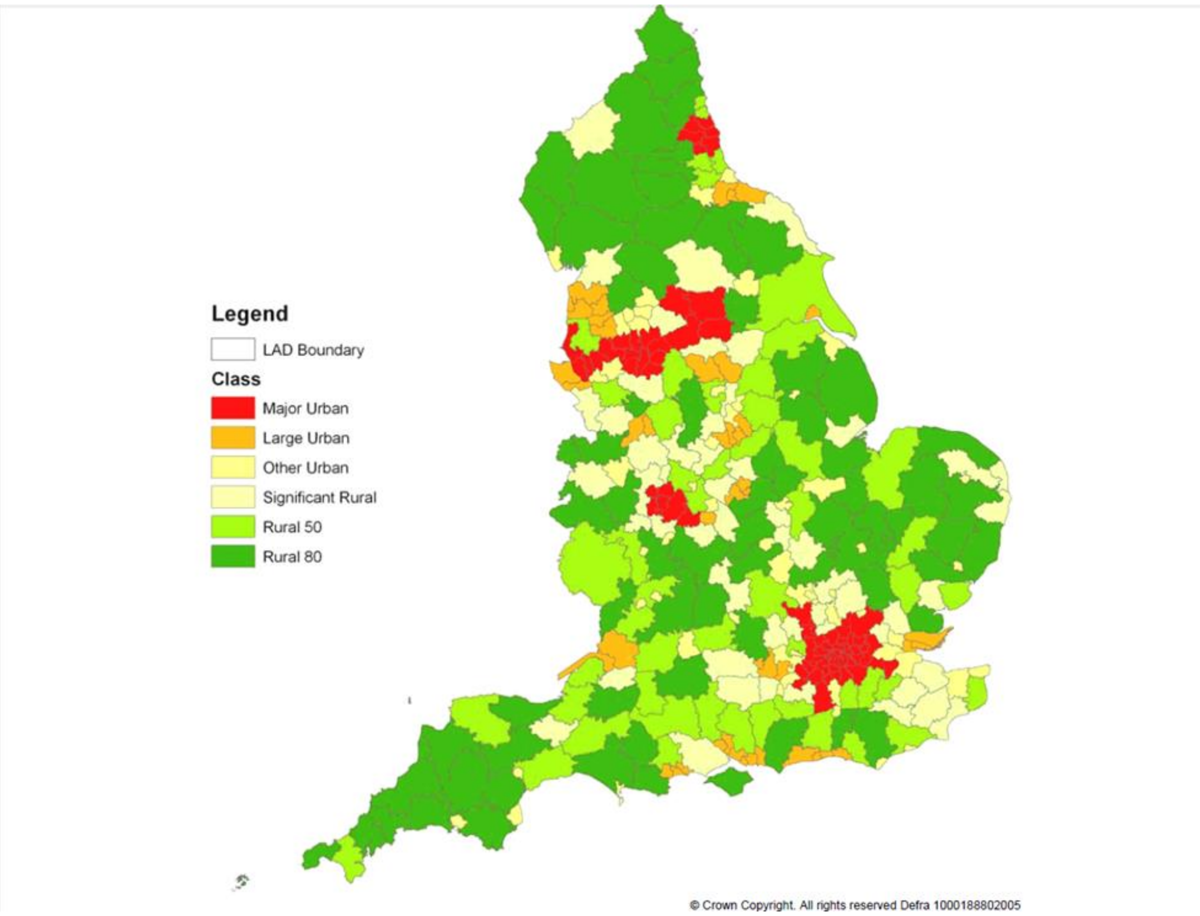
The aim of this work is to identify the link between bus accessibility and deprivation scores. However the accessibility figures only provide us with public transport accessibility which may incorporate other modes such as light rail, we focused our analysis on the subset of LSOAs which fell within the category of towns or cities outside of PTE areas. These figures can then more reliably be associated with bus accessibility. We followed the DEFRA rural/urban classification typology as detailed in Table 2 below and the large and other urban areas that we focus on appear as the yellow and orange areas in the map shown in Figure 3-1.

Table 2: Rural and Urban Stratification

DEFRA Classification	DEFRA Definition	Final Stratification
Major Urban	100k people or 50 percent of their population in an urban area with a population of more than 750,000	**Split into Dense Urban (ie Metropolitan) and London**
Large Urban	50k people *or* 50 percent of their population in one of 17 urban areas with a population between 250,000 and 750,000	**Towns/ Cities**
Other Urban	<37,000 people *or* less than 26 percent of their population in rural settlements and larger market towns	
Significant Rural	More than 37,000 people *and* more than 26 percent of their population in rural settlements and larger market towns	**Rural**
Rural-50	districts with at least 50 percent but less than 80 percent of their population in rural settlements and larger market towns	
Rural-80	districts with at least 80 percent of their population in rural settlements and larger market towns; there are 73 districts in this group	

[6] The calculation of these travel times is rather complex but described in https://www.gov.uk/government/publications/transport-connectivity-and-accessibility-of-key-services-statistics-guidance

Figure 3-1: Map of Urban/Rural Classifications of LAD areas



3.5 Other co-variates

We use the following additional covariates in our model, identified for each LSOA from the 2011 Census data[7].

Population density. By including this measure, we attempt to control for impacts on social outcomes that arise through differences in density. More dense areas may have better access to services and amenities than less dense areas.

Population – to control for differences in social impacts based on population of an area.

Household Non car availability – this is a measure capturing the proportion of the population who do not have access to a vehicle in their households. Typically people without access to a car will not be able to access the same kinds of amenities and opportunities that those with a car can, effectively acting as a social barrier. However, arguably it is also a proxy for income so part of the dependent variable, however we will test for this in subsequent analysis.

[7] Downloaded from https://www.nomisweb.co.uk/

Industrial Structure: We capture industrial structure by setting a dummy value of 1 to the industrial grouping which has the highest relative concentration (ie the highest percentage uplift relative to the national average). These groupings are based on the 2007 Standard Industrial Classification classifications from the census which we have categorised as Manufacturing, Retail, Business services, Professional, Public and Service sectors[8]. These measures attempt to control for the structure of employment in each area. Some more deprived areas are associated with higher concentrations of employment in Manufacturing, Retail and Service sectors.

Local Authority Identifier. The nature of this analysis is that we cannot observe all the variables which determine social outcomes in a given area. These variables include, for example, natural resources (eg a coalfield), local geography (eg ports may improve employment prospects in coastal areas), the presence of large historical employers in an area. A Fixed Effects approach is a standard way of controlling for any of these time-invariant unobserved characteristics, in this case of an area through the estimation of a Local Authority District[9] (LAD) area constant. By modelling using fixed effects we are looking to explain variation in outcomes 'within' Local Authority Districts (LAD), as any underlying differences in outcomes 'between' LADs are controlled for by the dummy variables.

Underlying skills as captured by educational attainment forms part of the IMD measure itself so were omitted as co-variates.

3.6 Descriptive Statistics

Table 3: Descriptives for all areas and Towns and Cities

		All areas				Town and City			
Variable Name	*Description*	*Mean*	*Std. Dev*	*Min*	*Max*	*Mean*	*Std. Dev*	*Min*	*Max*
DIMD2	Adjusted Index Multiple Deprivation	21.7	16.1	0.3	95.2	21.0	16.2	0.4	93.9
DIMD2Rank	IMD Ranking	16423	9481	1	32844	16948	9659	2	32843
Empl_pttime	Time to nearest employment centre (mins)	8.7	4.5	5.0	77.7	8.0	2.9	5	31.3
Gp_pttime	Time to nearest GP	8.2	3.8	5.0	81.5	8.0	2.6	5	32.4
Hosp_pttime	Time to nearest hospital by PT	25.3	16.5	5.0	120.0	23.7	14.6	5	120.0
Town_pttime	Time to nearest town centre by PT	15.2	9.3	5.0	120.0	13.9	6.4	5	57.4
pop	Population	1614	301.3	983.0	8300	1594	299.3	983	8300.0
pop_dens	Population per hectare	42.6	42.3	0.0	684.7	38.2	28.8	0.4	337.8
NCA	Non Car Availability 2011	0.25	0.17	0.00	0.86	0.23	0.13	0.00	0.82
Observations			**32840**				**14514**		

[8] Manufacturing comprised SIC groups C (Manufacturing) and F (Construction); Retail is represented by group G (Wholesale and retail trade; repair of motor vehicles and motor cycles); Business services comprised SIC groups J (Information and communication) and K (Financial and insurance activities); Professional is represented by group M (Professional, scientific and technical activities). Public comprised SIC groups O (Public administration and defence; compulsory social security), P (Education). Q (Human health and social work activities)

[9] Includes Unitary Authorities too

Table 3 shows descriptive statistics for the key variables for all England and for Towns and Cities. It shows the adjusted index of deprivation varies from 0.3 to 95.2 (although there is a potential minimum of 0 and a maximum of 100). There is little difference in average IMD scores between the full and sub-sample. Journey times by public transport have a minimum of 5 minutes and a maximum of 120 minutes observed for access to hospitals and town centres, presumably in more remote rural areas.

Populations vary from 983 to 8300 with a mean of around 1600 for all areas and the sub-sample. Population densities are slightly lower for the sub-sample.

Levels of car availability are similar between the two samples and average around 0.25 with a maximum value of over 0.8 (80% of people with no car access).

4 METHODOLOGY

We estimate a cross sectional model of deprivation at the LSOA level utilising 14514 observations on town/city based LSOAs constructed matching the accessibility, deprivation and census data together at the appropriate level of aggregation. Our analysis considers the strength of the relationship between transport connectivity and levels of economic, social and environmental deprivation after taking account of other factors that influence levels of deprivation.

The cross sectional approach allows us to investigate the relationship between spatial differences in public transport accessibility and differences in social outcomes as captured by the index of multiple deprivation, controlling for other localised factors. In order to estimate this model we conduct fixed effects regression analysis by estimating a set of constants for each Local Authority District which capture area-wide unobserved characteristics influencing employment.

Our model formulates deprivation in LSOA *i* within LAD/UA *k* in the following way:

$$IMD_i = f(A_i, C_{LADk}, V_i)$$

where:

A_i represents the accessibility measures for area *i*; *Town_pttime; Empl_pttime; Hosp_pttime; GP_pttime* for public transport times to nearest town centre, employment centre, hospital and GP respectively.

V_i are variable factors such as population and car availability

C_{LADk} are constants capturing the impact of unobserved variables within LAD area *k*

We experimented with a variety of functional forms including:

Linear levels- the model is regressed using the variables at the level they are reported at

Log-levels – all non-dummy variables are logged. This functional form can be used to derive proportional responses (elasticities) for the impact of differences in travel time on deprivation.

Logisitic Model – here the dependent variable (the IMD) is transformed using the logistic distribution. Because the dependent variable is actually bounded to be between 0 and 100, it can be treated as a proportion. Predictions from OLS can produce values outside the range of possible outcomes. The logit transformation ensures that values fall strictly within the unit interval. The transformed response variable is

$Y^*=y/(1-y)= X\beta +u.$

We can then use a linear regression of *Y** as a function of linear or log-linear transformed explanatory variables

4.1 Endogeneity and Instrumental Variable Analysis

Another important issue within this analysis is to understand if any of our explanatory variables might be endogenous with the dependent variable. Different levels of deprivation may also have differential effects on Non Car Availabliity in that lower levels of deprivation may facilitate higher levels of car availability through an income effect,

We investigate this issue of endogeneity through the use of instrumental variable (IV) approaches to control for the endogeneity between deprivation and Non Car Availability and employment. Successful IV estimation here requires the identification of at least one variable (instrument) which influences car availability but is uncorrelated with the error term in the deprivation equation, i.e. is only correlated with deprivation through its impact on car availability. We use the 2 stage-least squares (2SLS) approach: in the first stage an instrumenting regression is estimated where the endogenous variable is modelled using the instrument(s) and all the explanatory variables from the second stage; in the second stage an instrumented regression is estimated using the predicted variable(s) from the first stage as one of the explanatory variables in place of its actual counterpart.

There are two requirements for a good instrumental variable. Firstly, the instrument must be highly correlated with the endogenous explanatory variable it is instrumenting (ie non car availability). Secondly, the instrument must have a very low correlation with the residual error from the second stage regression (on deprivation). These two requirements are referred to as instrument relevance and instrument exogeneity.

We cannot test this second requirement, instrument exogeneity, other than appealing to economic intuition. Given a selected instrument (or instruments) we can however test for instrument relevance by comparing the OLS and 2SLS estimates for employment to determine whether the differences are significant. If they differ this implies there may be some degree of endogeneity and IV is appropriate. This is known as the Wu-Hausman test.

Successful IV estimation often involves use of long time-lags of the instrumented variable as instruments in the (first stage) instrumenting regression. This is because it can be argued that these lagged values cannot have been influenced by the current level of the dependent variable (e.g. deprivation) and are thus exogenous. We were able to make use of this, as we were able to use information from the 1991 census on non car availaibility. If there is endogeneity and it is not controlled for then OLS estimates of the relationship may be biased, yielding parameters which do not accurately reflect the direction of causation.

5 RESULTS

5.1 Linear vs Log vs Logistic Models

The first 3 sets of 2 columns shown in Table 4 are the results of the estimation of the deprivation regression for the different functional forms, linear, log and logistic described above.

Initial modelling was based on all areas but subsequently it was decided to focus on Town and Cities as it was felt that the PT accessibility variables would reflect bus services rather than other forms of public transport which feature in larger conurbations and that more rural areas are predominantly served by private transport. The reported models are based on the 14514 LSOAs which fall into the category of towns and cities only.

Table 4: Model Results for Towns and Cities

	Linear Model		Log Model		Logistic Model		Log IV Model	
employment	Coeff	T-stat	Coeff	T-stat	Coeff	T-stat	Coeff	T-stat
pop	-0.00219	*-6.2*	-0.111	*-3.8*	-0.00016	*-6.8*	-0.113	*-5.1*
pop_dens	-0.0243	*-2.7*	-0.0187	*-3.7*	-0.00099	*-1.8*	-0.0289	*-6.9*
Town_pttime	0.239	*11.0*	0.169	*9.8*	0.0118	*7.5*	0.198	*20.6*
Empl_pttime	0.0845	*2.6*			-0.00471	*-2.1*		
Hosp_pttime	0.0334	*3.0*	0.0505	*3.5*	0.00182	*2.5*	0.0584	*8.6*
Gp_pttime	0.204	*6.1*	0.0332	*1.9*	0.00423	*1.5*	0.0587	*4.6*
reta_dum	0.662	*2.9*	0.0249	*2.1*	0.0322	*2.1*	0.0218	*2.3*
busi_dum	-1.849	*-5.6*	-0.139	*-7.5*	-0.231	*-10.8*	-0.124	*-9.1*
prof_dum	-2.760	*-5.6*	-0.204	*-8.4*	-0.291	*-9.4*	-0.193	*-12.8*
publ_dum	-2.193	*-6.3*	-0.132	*-7.4*	-0.180	*-7.9*	-0.122	*-10.0*
serv_dum	-1.269	*-3.3*	-0.0530	*-2.9*	-0.128	*-5.7*	-0.0588	*-5.6*
NCA	109.5	*64.6*	1.104	*87.8*	6.778	*62.7*	1.177	*133.5*
_cons	-4.837	*-4.3*	4.871	*21.1*	-2.962	*-37.0*	4.963	*28.4*
R-squared (within)	0.760		0.733		0.726		0.730	
R-squared (overall)	0.769		0.753		0.747		0.751	
Groups (FE Constants)	281							
Observations	14514							

The linear model results show a negative impact of population (*pop*) and population density (*pop_dens*) on deprivation which means denser and more populous areas are associated with less deprivation, all else equal.

We find positive and significant coefficients on all four public transport accessibility measures which implies that areas which are further from town, employment centres, hospitals and GPs by public transport (as measured in minutes by *Town_pttime, Empl_pttime, Hosp_pttime, Gp_pttime* respectively) are typically more deprived, all else equal.

A positive and highly significant coefficient on non car availability (*NCA*) suggests areas with lower levels of car availability are associated with areas of higher deprivation.

The following (mutually exclusive) dummies were used to represent industrial structure:

reta_dum is 1 for areas with relatively higher concentration of retail employment

busi_dum is 1 for areas with relatively higher concentration of business services employment,

prof_dum is 1 for areas with relatively higher concentration of professional service employment,

publ_dum is 1 for areas with relatively higher concentration of public sector employment,

serv_dum is 1 for areas with relatively higher concentration of service employment,

with manufacturing and construction the omitted base group.

The results show that, relative to manufacturing and construction, areas with a relatively high concentration of business, professional services, public sector or, to a lesser degree, service sector employment were associated with areas with lower levels of deprivation.

The linear model reported an overall R^2 value of 0.769 which represents a good fit and predicted only 2% of observations to be outside the permitted range of 0-100 for the deprivation score.

Results from the log model in the next two columns of Table 4 show a similar pattern overall with significant negative relationship between (log) population and (log) deprivation and positive between (log) deprivation and (log) non-car availability. However, in this case the coefficient on (log) public transport time to nearest employment site was negative but highly insignificant so dropped from the final reported regression model here. The impact of the log of other time measures were positive and significant.

The log model reported an overall R^2 value of 0.753 which represents a good fit and predicted only 1% of observations to be outside the permitted range of 0-100 for the deprivation score. The model fit is not comparable to that from the linear model as the dependent variable is on a different scale. However, an adjustment was made to rescale the models (based on the geometric mean) so as to be comparable and it was found that the log model had a better fit than the linear model[10].

The third set of results from the logistic model reported in Table 4 follow a similar pattern of sign and significance to the log model. Results from this model cannot be adjusted to be compared to the linear model and it is impossible to derive marginal effects (Marginal effects in logistic models depend on the level of other covariates whereas this is not an issue in linear models) or constant elasticities, so interpretation of the results is difficult. For these reasons it was decided not to proceed with this model.

For reasons of fit, prediction and ease of interpretation we decided to proceed with the log model as the preferred model.

The adjusted IMD score model was preferred to the ranking on the basis that it recovered more robust coefficients and had fewer predicted values (half as many) outside the observed

[10] The re-scaled linear model had a residual sum of squares (RSS) of 2839 whereas the re-scaled log model had a RSS of 1992

data range. Also the rankings corresponded to the national picture rather than for just towns and cities.

5.2 Instrumental Variable Models

One concern is that the strong results for car availability may be due to the influence of deprivation on car availability rather than of car availability on deprivation, ie that car availability might be a measure itself of deprivation rather than a causal factor. To examine this we instrumented for non-car availability using historic measures of non-car availability from the 1991 UK census which could not be affected by current levels of deprivation.

The results of this approach are reported in the final 2 columns of Table 4. Casual inspection of the coefficients indicates this model performs very similarly to the non-instrumented log model reported in the second set of columns in Table 4.

Because of the two-stage nature of the estimation the standard errors and hence t-stats are not directly comparable with those from the 1-stage linear model.

The Wu-Hausman test score (with an associated p-value of 0.00) suggests we reject the exogeneity of car availability. Whilst we cannot verify that the use of lagged car availability is exogenous, it is intuitively appealing and a typical approach for this kind of analysis. It is for this reason we prefer this model.

IV estimation is inefficient, ie leads to lower t-statistics and less precise estimates than under OLS, so if OLS is appropriate it is preferable. If there is endogeneity and it is not controlled for then OLS estimates of the relationship may be biased, yielding parameters which do not accurately reflect the direction of causation.

5.3 Interpretation of results

The preferred final IV based log form of the model (the final set of columns in Table 4) provides us with a set of elasticities which show us the magnitude of the linkage between public transport accessibility and deprivation. The results of this model form provide us with coefficients which represent an 'elasticity of deprivation' to small differences in each explanatory variable. They suggest for example an elasticity of deprivation of 0.2 for public transport times to town centres – ie a 10% (positive) difference in travel times by bus to town centres between areas within a given local authority area would be associated with a (positive) difference in IMD scores of 2% (ie 2% more deprived) all else equal. If we compound the effect of the elasticities for travel to town, hospital and GPs, we have an elasticity of 0.34, ie a 10% difference in travel times by bus between areas within a given local authority area would be associated with a difference in IMD scores of 3.6% all else equal.

Although the econometric analysis alone does not prove causation, the interpretation of the results in the context of the conclusions from the literature review and stakeholder engagement from the main report do indicate that bus accessibility may have important economic, social and environmental impacts.

In order to understand the results further we looked at how these differences in IMD scores between areas with differing levels of accessibility break down in terms of the individual deprivation domains. These are not the entire set of indicators, but are some of the key ones and easy to interpret. For each area we calculate a ‘smoothed’ value of domain measures based on the average of the 10 areas either side of it in the IMD distribution. This effectively provides a picture of what areas typically look like in terms of their domain scores at this point in the IMD distribution.

Following the accessibility improvement our model predicts a new IMD value for each area. This puts the area in a different part of the distribution (ie it becomes ‘less’ deprived). We then recalculate the ‘smoothed’ value of domain measures in this new part of the distribution associated with town/city areas which have similar IMDs to the area in question following a 10% improvement in accessibility.

This enables the comparison of before and after domain scores for each area. We repeat this process for every area in our sample and take the average changes in the domain scores for areas in each decile of the IMD distribution, enabling us to compare the average values of individual domains. We use our model results to examine how improvements in accessibility, as modelled by a 10% improvement in public transport travel times, would be associated with changes the IMD scores on average in each of these deciles of the distribution.

As the analysis is carried out for each of the areas in the ten different deciles of the IMD distribution, this enables comparisons of IMD scores and associated underlying domains at these different parts of the distribution. For example we can see the changes in domain scores which are associated with improvements in accessibility for the 10% of most deprived areas and compare these with the changes in the 10% least deprived areas. The percentage changes in the key domain scores are shown in Table 5 below.

The table shows that a 10% improvement in local bus service accessibility in the most deprived 10% of town/city areas is predicted to change the IMD to reflect areas which have 2.7% fewer unemployed, 2.8% fewer individuals from benefit claiming households, 1.4% fewer ‘unskilled’ adults and 0.7% more young people staying on to post 16 education. The impacts are proportionally lower in the least deprived 10% of town/city areas, with the change in IMD associated with areas with 1.3% fewer unemployed, 1.6% fewer individuals from benefit claiming households, 0.7% fewer ‘unskilled’ adults and 0.3% more young people staying on to post 16 education. Entry to higher education changes very little.

Table 6 aggregates the results from each decile to impute the absolute change in numbers unemployed and in benefit claiming households associated with a 10% improvement in journey times. We see that for the 10% most deprived areas in England this improvement in IMD would be associated with areas with around 10,000 less unemployed, whilst in the least deprived areas this would be associated with areas with just under 600 fewer unemployed. In terms of individuals from benefit claiming households, for the 10% most deprived areas in England this improvement would be associated with areas with over 22,000 fewer claimants, whilst in the least deprived areas this would be associated with just over 1,000 fewer unemployed.

Table 5: Comparison of model results for indicators by deciles of the IMD distribution.

		% improvement in indicators associated with 10% improvement in accessibility				
Decile	**Description**	**Employment deprivation**	**Income deprivation**	**Post 16 education**	**Entry to higher education**	**Adult skills**
1st	**Most deprived**	2.7	2.8	0.7	0.1	1.4
2		3.3	3.1	1.0	0.1	1.3
3		2.6	2.7	0.4	0.2	1.4
4		2.8	3.3	1.8	0.2	1.6
5		2.8	3.1	0.1	0.1	1.1
6		2.8	3.5	0.6	0.1	0.8
7		2.3	2.5	0.7	0.1	1.4
8		2.3	2.6	0.2	0.1	0.5
9		2.0	2.6	0.5	0.2	1.3
10th	**Least deprived**	1.3	1.6	0.3	0.1	0.7
Average		**2.7**	**2.9**	**0.7**	**0.1**	**1.2**

Notes: The **Employment Deprivation** Domain measures the proportion of the working-age population in an area involuntarily excluded from the labour market.
The **income deprivation** index measures the proportion of the population experiencing deprivation relating to low income, including both those people that are out-of-work, and those that are in work but who have low earnings
The **Post 16 education** indicator measures the proportion of young people not staying on in school or non-advanced education above age 16
The **Entry to higher education** indicator measures the proportion of young people aged under 21 not entering higher education
The **adult skills** indicator is the proportion of working-age adults with no or low qualifications combined with he proportion of the working-age population who cannot speak English or cannot speak English 'well'.

Table 6: Comparison of changes in unemployed and income deprived individuals in towns/city areas by deciles.

Decile	**Description**	**Change in Unemployed**	**Change in Income deprived**	**Change in those with no adult skills**	**Population**
1st	**Most deprived**	-9909	-22647	7310	2,246,950
2		-8735	-17671	6359	2,325,528
3		-5472	-12110	6154	2,324,472
4		-4914	-11825	6071	2,316,807
5		-4040	-9115	4097	2,334,958
6		-3259	-8139	2474	2,309,631
7		-2305	-4590	4096	2,303,004
8		-1903	-3895	1263	2,310,787
9		-1292	-2966	3399	2,312,724
10th	**Least deprived**	-571	-1079	1246	1,983,367
Average		-3.0	-6.6		

6 SUMMARY

The work reported on here uses econometric models to analyse the linkage between deprivation as measured by the index of multiple deprivation and public transport accessibility, controlling for car availability, industrial structure and unobserved local authority area characteristics.

Our models appear plausible in terms of signs and magnitudes for all estimated coefficients.

Our findings show that, after allowing for other factors that influence deprivation, neighbourhoods with better bus services have lower levels of unemployment, lower levels of families on income support, lower numbers of unskilled workers, more young people staying on in education after the age of 16 and entering higher education and longer life expectancy amongst other things.

Our results would indicate that, all else equal, we find a statistical link between public transport accessibility (as measured by journey times) and deprivation, which suggests a worsening of public transport could have detrimental effects on areas. Whilst good local bus services are not an end in themselves they do enable individuals take employment, participate in education and take better care of themselves - activities which may otherwise be out of reach.