UNIVERSITY OF LEEDS

INSTITUTE FOR TRANSPORT STUDIES

ITS


# Land Value and Transport (Phase 2): Modelling and Appraisal


Final report

August 2019

Research funded by:

TRANSPORT FOR THE NORTH

West Yorkshire Combined Authority

EPSRC Engineering and Physical Sciences Research Council

# Contents

Version 1.1

Authors: J. Nellthorp, M. Ojeda Cabral, D. Johnson, C. Leahy and L. Jiang.

Please cite this work as:

Nellthorp, J., Ojeda Cabral, M., Johnson, D., Leahy, C. and Jiang, L. (2019*). Land Value and Transport (Phase 2): Modelling and Appraisal.* Final Report to TfN, WYCA and EPSRC. Leeds: Institute for Transport Studies, University of Leeds. https://doi.org/10.5518/100/18

# Glossary

| | |
|---|---|
| AADF | Annual Average Daily Flow – traffic |
| AURN | Automatic Urban and Rural Network of air pollution monitoring sites |
| BCR | Benefit:Cost Ratio |
| BRS | Business Rates Supplement |
| CadnaA | Noise modelling software |
| CBA | Cost benefit analysis |
| CBD | Central Business District – of a city |
| CGT | Capital Gains Tax |
| dB | Decibels – unit of noise |
| DfT | Department for Transport |
| DCLG | Department for Communities and Local Government |
| Defra | Department for Environment, Food and Rural Affairs |
| DfE | Department for Education |
| DfT | Department for Transport |
| DLR | Docklands Light Railway, London |
| EA | Environment Agency |
| END | Environmental Noise Directive |
| EO | Employment Opportunities |
| EP | Employment Potentiality |
| EPSRC | Engineering and Physical Sciences Research Council |
| *ex ante* | before implementation, analysis based on forecast data |
| *ex post* | after implementation, analysis taking advantage of available out-turn data |
| FE | Further Education |
| GC / GJC | Generalised Journey Cost – composite measure of journey cost including non-money items |
| GIS | Geographical Information Systems |
| GJT | Generalised Journey Time – composite, weighted measure of journey time used in PDFH and rail appraisal |
| GM | Greater Manchester |
| GP | General Practitioner – family doctor |
| GVA | Gross Value Added |
| GWR | Geographically Weighted Regression |
| HE | Higher Education |
| HP | Hedonic Pricing |
| HSL | Health and Safety Laboratory |
| IMD | Index of Multiple Deprivation |

| | |
|---|---|
| ITS | Institute for Transport Studies, University of Leeds |
| IVT | In-Vehicle Time |
| JLE | Jubilee Line Extension, London |
| JT | Journey Time |
| KS2 | Key Stage 2 (ages 7-11 years) |
| KS4 | Key Stage 4 (ages 14-16 years) |
| LA | Local Authority |
| LSOA | Lower Layer Super Output Area |
| LUTI | Land Use and Transport Interaction (model) |
| LVC | Land Value Capture |
| LVU | Land value uplift |
| MHCLG | Ministry of Housing, Communities and Local Government |
| MSOA | Middle Layer Super Output Area |
| NECA | North East Combined Authority |
| NELUM | Northern Economy & Land Use Model |
| NTEM | National Trip End Model |
| NIC | National Infrastructure Commission |
| NNDR | National Non-Domestic Rates |
| NoRMS | Northern Rail Modelling System |
| $NO_2$ | Nitrogen dioxide – air pollutant |
| NOx | Nitrous oxides – air pollutants |
| NPD | National Population Database (HSL) |
| NPR | Northern Powerhouse Rail |
| NPV | Net Present Value |
| OA | Output Area |
| OBC | Outline Business Case |
| OLS | Ordinary Least Squares estimation |
| ONS | Office of National Statistics |
| OS | Ordnance Survey |
| PAYE | Pay As You Earn income tax collection |
| PCM | Pollution Climate Models (Defra) |
| PDFH | Passenger Demand Forecasting Handbook (RDG, 2018) |
| $PM_{2.5}$ | Particulates smaller than 2.5 micrometres in diameter – air pollutant |
| PT | Public Transport |
| PVB | Present Value of Benefits |
| PVC | Present Value of Costs |
| SAR | Spatial Autoregressive model |

| | |
|---|---|
| SDLT | Stamp Duty Land Tax |
| SEM | Spatial Error Model |
| SO | Skills/Study Opportunities |
| SOBC | Strategic Outline Business Case |
| STATS19 | Accident reporting form used by the Police |
| TAG | DfT's online Transport Analysis Guidance |
| TfL | Transport for London |
| TfN | Transport for the North |
| TfGM | Transport for Greater Manchester |
| TRACC | Transport accessibility modelling tool by Basemap Ltd |
| TTW | Travel to work |
| VOA | Valuation Office Agency |
| WebTAG | DfT's online Transport Analysis Guidance (up to 2018) – now 'TAG' |
| WTP | Willingness to Pay |
| WYCA | West Yorkshire Combined Authority |

# Highlights

- The Report focuses on understanding and quantifying the relationships between transport and property values in the North of England.
- Three innovative models were built using available datasets – these are:
  - a *cross-sectional model of residential property values* across the whole of the North of England, using hedonic pricing methods and including both accessibility and place quality;
  - a *quasi-experimental model of residential property values,* focusing on the impact of Manchester Metrolink extensions between 1995 and 2018, using panel data (cross-sectional and time series);
  - a *cross-sectional model of commercial property values*, which is a prototype model based on a smaller dataset, covering Leeds and Bradford initially.
- Accessibility to employment by rail, car and walk – as the 'main mode' of travel – are each found to be highly significant in the residential property market. For every additional 10,000 jobs that are accessible from the home location, there is a property price premium of +0.16% for rail and +0.19% for car. The effect of walk accessibility to employment is even more powerful than for rail or car, and this is seen in city centre price premiums – although the total number of walk-accessible jobs tends to be lower than for the faster modes.
- Place quality also plays a highly significant role in residential property prices. Important place quality variables include accessibility to local centres, greenspace and play opportunities, community noise levels and air quality. As with previous research studies, local schools play an important role.
- Property type and floorspace have an influence as expected, and there is premium on new homes (+18%). Neighbourhood average income is included as a control variable, and the finding of a premium associated with this is both: (i) consistent with economic theory; and (ii) part of the explanation for price effects of 'gentrification' where that occurs.
- The average uplift due to new Manchester Metrolink stations was found to be +6.3% for the 0-1km catchment. This varied across Metrolink lines, e.g. +10.5% in the higher-income South Manchester catchment; lower or insignificant in East Manchester/Rochdale Line catchments. The Airport Line was exceptional (+20.6%), with employment centres at both ends and the opportunities of international connectivity.
- The study provides a first estimate of the potential uplift due to Northern Powerhouse Rail (NPR) service improvements. The maximum effect (on a very localised basis) was estimated at +9.3%, however this may be increased by dynamic effects through the economy and property market. Estimates of the total potential uplift are provided.
- The ability to use Land Value Capture (LVC) to recycle some of this property value gain into funding for infrastructure projects is considered: as with Crossrail and Crossrail 2, there are some serious limitations to the set of LVC tools available in England, relating to existing properties in particular (less so for new development).
- The study provides some insight into regional differences in uplifts in the North versus London and the South East. Uplifts are expected to be lower in the North at the present time – the models show how this is related to job density, scale of accessibility improvement, and income.
- Commercial property premiums are related to both station proximity and walk accessibility to other employment.
- Many model variants and sub-variants were tested: across these models, the parameter estimates and signs are remarkably stable and most are highly significant. This is encouraging for the robustness of the findings. When compared with previous well-regarded (though smaller scale) modelling, the results are very comparable (e.g. Nationwide, 2014/19). Internal consistency checks between the cross-sectional and quasi-experimental models also provide encouraging findings.

# Land Value and Transport (Phase 2): Modelling and Appraisal

## Final Report

## *Summary*


- This Report covers the findings of the Phase 2 study carried out by the Institute for Transport Studies (ITS), University of Leeds and sponsored by Transport for the North (TfN), West Yorkshire Combined Authority (WYCA) and the Engineering and Physical Sciences Research Council (EPSRC). Members of the study's Advisory Panel provided generous comments on an earlier draft, and those comments have been taken on board in this final version.
- The focus of the Report is on understanding and quantifying the relationships between transport and property values in the North of England.
- The main achievements described in this Report are the following:
  - The development of working models of three types:
    1. a Cross-Sectional model of Residential Property values;
    2. a Cross-Sectional model of Commercial Property values; and
    3. a Time Series model of Residential Property values.
  - The *Cross-Sectional Residential Property Model* makes use of a theoretical framework developed in this study, and builds on previous leading studies in the international literature (e.g. Mulley et al., 2018; He et al., 2018; Ahlfeldt, 2013). It covers the whole of the North of England: population 15 million; housing stock 7.1 million units; economy £350billion per annum Gross Value Added (GVA). The results address both the value of accessibility – by different modes of transport – and the value of place quality.
  - The *Commercial Model* is a new prototype model based on a small dataset, however additional data will allow it to be estimated more robustly. It opens up a set of questions about the linkages between commercial property and the wider economy, and offers some preliminary findings.
  - The *Time Series Residential Model* focuses on Manchester Metrolink – a case study where there have been significant improvements in accessibility by public transport over the last 25 years. The model takes a quasi-experimental approach to measure the impact of the intervention. The results are available line-by-line, for 0-1km and 1-2km station catchments. The pattern of the results can be understood with reference to the cross-sectional model and dynamic factors, and the implications are carefully considered.
  - The models have been used to explore the potential land value uplift due to Northern Powerhouse Rail (NPR) and some other hypothetical policy tests.
  - The implications for *appraisal, potential land value capture* in the North and the *distributional impacts* of rail investment are considered, using the quantitative results.
- The key points of the method and the main findings of the modelling work can be summarised as follows:

## *Cross Sectional Residential Model*

- A series of hedonic pricing (HP) models were developed, using a dataset of 160,563 property sales transactions in the year 2016 in the North of England, specifically the Transport for the North (TfN) area. The model form and explanatory variables are based in economic theory (see Section 2).
- The variables include: accessibility to employment (by mode of transport); place quality (defined quite widely); building attributes; income; and supply constraints in the local housing market.
- In these models, income is allowed to interact with the other variables, with significant results. This shows that the % premium for rail accessibility is greatest in high income areas – reasons for this and the implications of it are discussed (Sections 4.1.6 and 5.4). By contrast, the % premium for walk accessibility is uniform across high and low income areas – similarly car accessibility, where a minimal and not significant income effect was found.
- The models are spatially detailed, which has made it possible to estimate accessibility to employment parameters for rail, walk and car modes separately. Initially the model was estimated using MSOA-level[1] data for rail accessibility and some other key variables – there are roughly 3,400 homes per MSOA in the study area. The final models are based on rail accessibility at the much finer Output Areas (OA) level at the origin (there are around 130 homes in an OA) and at an intermediate LSOA level[2] at the destinations where employment is located. Going to an even finer level of spatial detail would require moving to high performance computing; for present purposes we think the existing model has sufficient detail.
- The *accessibility to employment* variable measures the number of jobs that can be 'seen' from a specific home location. It takes into account the number of jobs in each locality across the North[3], and the ease (or difficulty) of reaching them. Jobs that are remote from the home location are discounted by a 'deterrence function' based on generalised journey time (GJT) – this is estimated using census data on home and work locations in the North. The TfN NoRMS model is used to estimate GJTs across the rail network[4], including in-vehicle time, interchange, delay and crowding. A customised access and egress function was used to estimate travel times to/from stations.
- Figure ES1 shows how rail accessibility to employment varies across the TfN area.

### *Accessibility findings*

- Accessibility to employment across the North was shown to be a key factor in residential property prices.
- *Rail accessibility to jobs* has a significant positive impact on property values, i.e. properties with better rail access to jobs are worth more than those with worse

[1] MSOA = Middle Layer Super Output Area

[2] LSOA = Lower Layer Super Output Area – LSOAs are typically five times larger than an OA, and five times smaller than an MSOA.

[3] Job locations include the TfN area plus an external 'buffer' of 50km to the south, since residents will be able to 'see' jobs outside the study area.

[4] for car, GJT is based on Trafficmaster data, and for walk, GJT is based on walk time data generated by TRACC.

accessibility, holding everything else constant. Table ES1 indicates the strength of this effect.

Figure ES1: Rail accessibility to employment in the TfN area (2015)

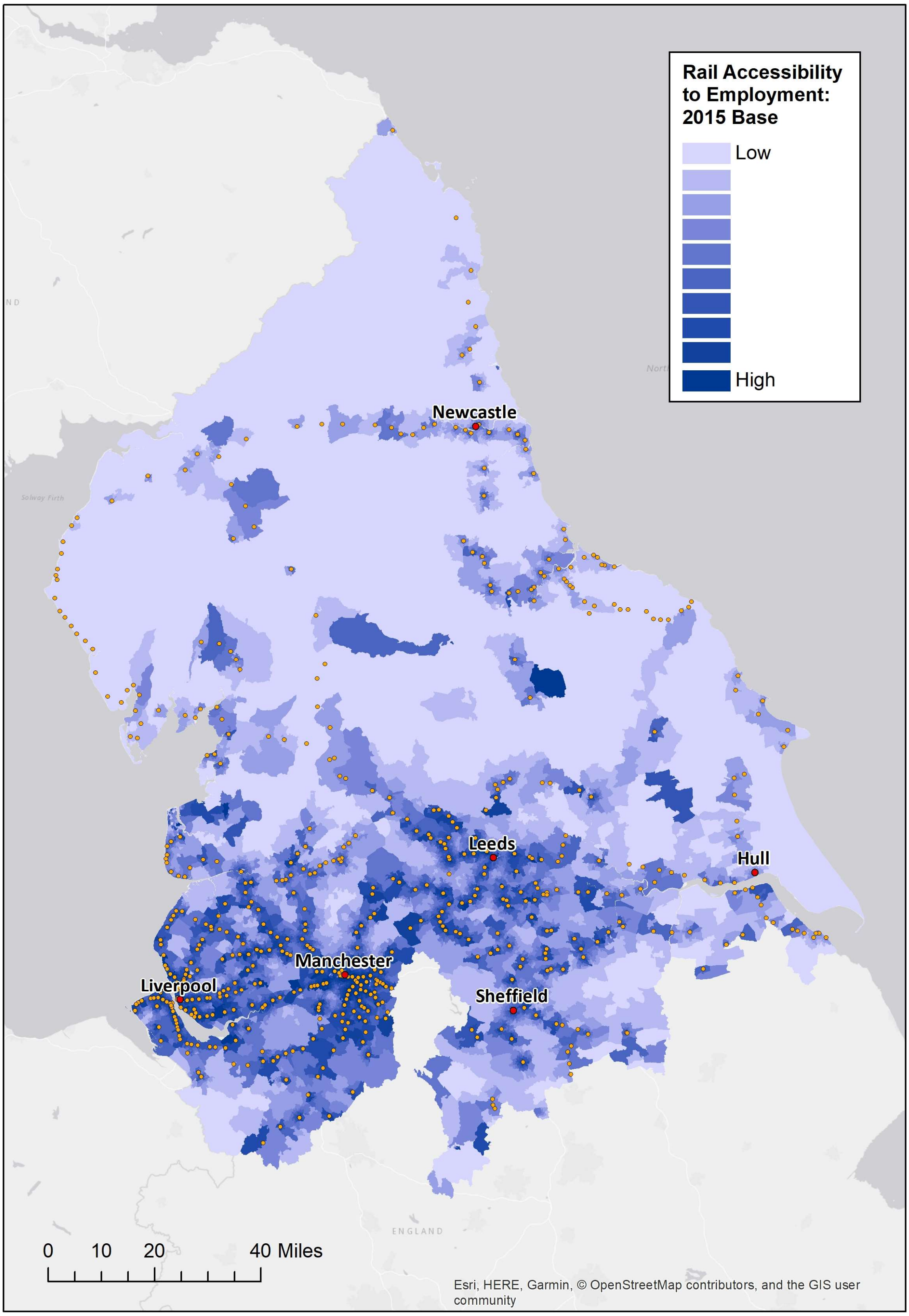


Table ES1: House price premium for rail accessibility to employment in the TfN area

| *% price premium per 100,000 jobs (GJT-discounted)* | *Worst-to-best rail accessibility in the North:* | |
|---|---|---|
| | *% price premium* | *difference in jobs accessible (worst-to-best)* |
| 1.6% | +14.3% | 917,000 |

- The interpretation is: rail accessibility to an additional 100,000 'GJT-discounted' jobs is worth a 1.6% premium on the price of a property, on average in the TfN area. The premium for rail accessibility between areas with the worst and best accessibility in the sample is 14.3%. This is based on the observed difference between worst and best, equal to approximately 917,000 GJT-discounted jobs.
- Once income interaction is included, the price impact of *rail accessibility* is found to vary between 1.1% per 100,000 jobs at the 10th percentile of household income, up to 2.1% at the 90th percentile. This implies that the property value premium can be expected to be very different between low and high income areas. If, hypothetically, the worst and best rail accessibility occurred in low income areas (at the 10th percentile of the income distribution) the model predicts that the price premium would be only 9.8%, whist in high income areas (the 90th income percentile) the premium would be 19.0% (Table ES2). See the Time Series model for further *ex post* evidence on this income effect.

Table ES2: Variation of the house price premium for rail accessibility with income

| *Comparison* | *% price premium* |
|---|---|
| Worst to best rail accessibility* | 14.3% |
| ... if low income LSOA (10th percentile) | 9.8% |
| ... if high income LSOA (90th percentile) | 19.0% |

Note: based on Model 14a with income interaction; * at the median LSOA income.

- For commuter rail stations, these results imply a pattern of rail accessibility premiums radiating out from the station (see Table ES3) which are familiar from previous research[5]. These premiums show the value – as seen in the housing market – of having a local rail station at different distances from the home.
- The larger premiums are found closer to the station, but in the new model there is evidence of a small premium even beyond 1500m from the station (using 2km as the base). The model captures noise and crime around stations to some extent[6] through the place quality variables, so although properties 250m from the station have the highest rail accessibility premium, they may not have the highest modelled prices – it also depends on their environmental and other characteristics.
- In the new model, the premium varies:
  - from station to station, based on its exact rail accessibility characteristics – i.e. how much connectivity it provides to what size clusters of employment;
  - also with the income profile of the catchment, because of the income interaction term.

[5] for a review, see Mohammad et al. (2013) or Debrezion et al. (2007).

[6] the main limitations here are that: only road noise is captured; and crime is at LSOA level.

Therefore two commuter rail stations were selected to illustrate – both in Greater Manchester, one in an area 9% above and one 9% below Greater Manchester average household income (Table ES3).

Table ES3: House price premiums for rail accessibility, applied to commuter rail stations in Greater Manchester (and compared with Nationwide, 2014/19)

| *Commuter Rail station at distance from home* | *ITS (2019) Cross-Sectional Model* | | *For comparison:* | |
|---|---|---|---|---|
| | *Station 1 (Higher Income)* | *Station 2 (Lower Income)* | *Nationwide (2014)* | *Nationwide (2019)* |
| 250m | +8.5% | +7.1% | | |
| 500m | +6.5% | +5.3% | +4.6% | +7.8% |
| 1000m | +3.3% | +2.7% | +2.0% | +3.3% |
| 1500m | +1.1% | +0.9% | 0 {BASE} | 0 {BASE} |
| 2000m | 0 {BASE} | 0 {BASE} | | |

Note: Stations 1&2 apply the Cross-Sectional model results to two illustrative Greater Manchester stations (East Didsbury and Moston stations 9km/6.5km and from central Manchester respectively); Nationwide (2014/19) give an *average* premium for Greater Manchester National Rail and Metrolink stations, also based on a cross-sectional model.

- Comparing with the previous cross-sectional models in Greater Manchester by Nationwide (2014/19), there is a slightly different spatial pattern. This may be because place quality variables (e.g. noise and crime) are included alongside rail accessibility in the new model, or due to other differences in modelling detail, or the differences in the stations sampled (as well as the different decision about where to set the base). In other respects, the results are encouragingly comparable. Changes in rail accessibility (including specific changes in the number of Metrolink destinations and frequency) could help to explain the changes between 2014 and 2019.
- *Walk accessibility to jobs* is also positive and highly significant. The premium for this mode appears to be substantial (Table ES4): the parameter implies that the price premium is 36.1% for each additional 100,000 jobs within reach by walking – but this is moderated by the relatively short distances that can be covered by walking. The premium going from worst to best walk accessibility in the TfN area is 51.2% (based on the sample difference of approximately 142,000 between worst and best). This would help to explain the premium for city centre properties, where job density is greatest.

Table ES4: House price premiums for accessibility to employment in the North of England

| *Mode* | *% price premium per 100,000 jobs (GJT-discounted)* | *Worst-to-best accessibility in the North:* | |
|---|---|---|---|
| | | *% price premium* | *difference in jobs accessible (worst-to-best)* |
| Rail | 1.6% | +14.3% | 917,000 |
| Walk | 36.1% | +51.2% | 142,000 |
| Car | 1.9% | +17.3% | 895,000 |

- *Car accessibility to jobs* completes the set of employment accessibility impacts and is also positive and significant. The average price premium for jobs accessible by car is slightly higher than for rail, with a 1.9% increase associated with an increase by 100,000 jobs. The difference in the indicator between worst and best is approximately 895,000, which is associated with a premium of 17.3%. The model parameters for rail, walk and car accessibility were all notably stable across different model variants tested.
- More than 99% of homes in the TfN area are within 1km of a bus stop, therefore bus accessibility is essentially part of the 'base' in this model. There is some evidence that remoteness from a bus stop (e.g. in rural areas) is negative for property values, however in most areas greater proximity to a bus stop does not seem to have a measurable value in the housing market and may even be negative – this mirrors the findings of some previous studies. It would be desirable to test bus accessibility further, using different model specifications and looking for any negative externalities, to understand this issue better.
- There is some evidence that proximity to centres of *skills training and education* is valuable alongside accessibility to jobs, although this is measured very simply in the model by number of resident students per LSOA, a full accessibility measure being too closely correlated to the 'accessibility to jobs'.
- Further details of the accessibility results are in Sections 4.1.1, 4.1.6 and 4.1.7.

***Place quality findings***

- Initially the work focused on identifying the place quality variables for which data could be obtained, and then testing these in the HP regression model. After mixed initial results, the approach to place quality was reconsidered carefully and a more structured approach taken, using the same type of theoretical approach for access to local facilities as for accessibility to jobs (above), though with different parameters based on the available evidence.
- The set of place quality variables showing a significant influence on property values includes:
    - *School quality* – both primary and secondary, Ofsted ratings of the five nearest schools;
    - *Accessibility to parks and gardens,* and to *playing fields and play spaces* (includes playgrounds – and is labelled 'Playgrounds' in the tables);
    - *Accessibility to local centres* – modelling identified that 'town centre' locations in DfT statistics do not capture all substantial local centres, and are based on 15 year old research for ODPM. The presence of a bank was found to be a better proxy for local centres, for the 2016 model. Going forward another proxy may be required, as the number of bank branches continues to decline.
    - *Crime* has a negative impact as expected, measured at LSOA level;
    - *Air quality* has an impact on price – lower $PM_{2.5}$ concentrations are associated with higher prices, all else equal, however the data could be much better – spatial resolution is only a 1km*1km grid;
    - *Road noise exposure* is more spatially detailed (10m*10m grid) and has a significant negative impact on price;
    - *Landfill sites* – increasing straight line distance from any landfill has a positive effect.

- Some place quality variables are found to interact with income. E.g. the influence of school ratings is greater in high income areas; accessibility to local centres (for shopping, etc) is valued very highly in high income areas, but local centres do not seem to be valued positively in the lowest income areas – perhaps reflecting a lack of discretionary spending power to afford 'lifestyle' facilities in these areas, or some negative externalities associated with local centres.
- An obvious question is: how large or small are these place quality effects relative to the effects of improving accessibility? This depends very much on the scale of the intervention and its effectiveness at a particular location. Table ES5 indicates that some place quality effects are potentially large: in certain specific, although not very common situations, they could be as powerful as the accessibility to employment effect, viewed from a specific residential location. These are best viewed as 'limiting cases' though – the effects will usually be smaller than this.

Table ES5: Relative magnitude of place quality and accessibility effects – illustrative examples of potential 'maximal' interventions

| *Place quality changes* | *Change in Value, %* |
|---|---|
| Local facilities vs. none (e.g. 'centre' planned into a new residential area, if no alternatives nearby) | +7.6% |
| Parks & playgrounds (good provision vs. completely absent) | +6.8% |
| Noise reduction 73→53 dB (e.g. trunk road placed in cut-and-cover tunnel) | +8.1% |
| *Example accessibility change* | |
| New suburban rail station (500m from neighbourhood) | +6.1% |

- Further details of the results for place quality are in Sections 4.1.2, 4.1.6 and 4.1.7.

***Building attributes***

- Floor area, in m$^2$, has a positive and significant effect on prices as expected. The effect is not exactly linear: larger homes are worth slightly less per m$^2$, all else equal (the same is found for commercial property). There are premiums on detached houses (+21% vs semi-detached) and new build homes (+18%), while terraces (-20%) and flats (-29%) are worth less than semi-detached houses – these results are in line with previous studies. Higher proportions of privately rented and owner occupied housing in the locality tend to increase values, the former indicating demand for housing in that location as a financial asset. See Section 4.1.3.

***Supply constraints in local housing markets***

- Recent growth in housing supply at Local Authority level in the previous five years was used as a proxy for a weaker housing supply constraint. This was associated with a dampening effect on property prices in 2016, as theory would suggest. See Section 4.1.4.

***Income (socio-economic variable)***

- The model finds that properties in areas where there is a higher average income will be worth more, holding all else equal (Sections 4.1.5 and 4.1.6). There is both a linear effect, +5.9% per £1,000[7] of equivalised[8] annual income, and a negative squared term. Therefore the premium varies from +7.3% per extra £1,000 of equivalised annual income at the 10th percentile of the income distribution, to +4.5% at the 90th percentile.

- This initially seems a large effect: bear in mind, however, that a difference of £10,000 in income between two LSOAs more-or-less corresponds to the difference between the 10th percentile (£22,000) and the 90th percentile (£33,600). This is a substantial effect, but not a surprising one in view of previous HP models.
- Interaction effects of income with rail accessibility, local school performance and accessibility to local centres have already been noted.

***Robustness***

- 14 main model variants, and many more sub-variants, were tested: across these models, the signs and magnitudes of the parameters are generally remarkably stable.
- Parameters are on the whole highly significant, and the large dataset helps to support a model with a large number of significant parameters.
- The behaviour of the model as variables are added or removed is quite intuitive, and the final series of tests described in Section 4.1.6 of the report lead us to Model 14a as a good representative model, summarising the work done so far (see Table 4.6).

***Models with Spatial Dummies, and Spatial Regression Models***

- In view of their roles in the literature, some alternative models were tested adding spatial fixed effects dummies or using spatial regression, to explore how the results changed, and whether model fit to the data would further improve (it was already good, Adjusted $R^2$=0.73, in the HP models).
- A model was estimated including 75 dummy variables representing the Local Authority (LA) areas in the North. This did produce a slight increase in Adjusted $R^2$ (to 0.77); at the same time, the car and rail accessibility variables became wrong sign (car) and insignificant (rail). The 75 LA dummies appeared to be picking up the wider-scale variation in accessibility instead of the accessibility variables doing so. It was interesting that *within* LA areas, the walk accessibility variable remained positive and significant – so local-scale accessibility differences remain and are captured.
- A Spatial Error Model (SEM) and a Spatial Autoregressive (SAR) model were also estimated, having tested for and found evidence of spatial autocorrelation. These models again showed rail accessibility becoming slightly less significant: this could be of concern when looking at the impact of rail accessibility to jobs on house prices – however, given the clustering of homes with high rail accessibility in close proximity to stations, the SEM/SAR models are likely attributing some of this high accessibility to the error as spatial correlation. It is worth noting that the parameter estimate on rail access to jobs remained relatively stable.
- There appear to be some parallels between the LA dummies model and the SEM & SAR models in terms of their effect on the main variables of policy interest.

[7] This coefficient varies across models, broadly in the range 4-6%.

[8] Household income divided by a weighted sum of household members (ONS, 2015).

Overall, these models added a little in terms of model fit, at the expense of reducing the model's ability to explain how accessibility influences house prices – as such their usefulness for exploring transport policy changes seemed less than the main models.

- A Geographically Weighted Regression (GWR) model was also tested but did not give defensible results – this is believed to be a consequence of the large size of the dataset and number of parameters. A peer reviewer commented that these models are generally difficult to estimate on large datasets and therefore would require substantial aggregation of this model in order to run (e.g. to LA level – which would defeat the purpose to some extent). The income interaction terms and spatially varying parameters in this model make the need for GWR less pressing in this case.
- Further details of these results are in Section 4.1.6.

### *Cross Sectional Commercial Model*

- As a starting point, we hypothesised that the value of commercial property is driven in part by its *accessibility for employees* and potential employees. This is similar to the accessibility from homes to jobs in the Residential model, but measured at the destination zone rather than the origin (home) zone. An increase in accessibility for employees should lead to an increase in labour supply at the work location – potentially advantageous for businesses who choose to locate there.
- We also hypothesised that the value of commercial property is driven in part by its *accessibility to other businesses*. This is a reflection of economic mass – as in agglomeration. The literature shows that there is a positive relationship between economic mass (or effective density) and productivity (e.g. Venables et al., 2014; Graham, 2007) and we might expect productivity helps drive the achievable rent at a particular location.
- A prototype model has been built using COSTAR data, for Leeds and Bradford local authority areas, purchased by WYCA and shared with ITS under the licence terms. This data usefully contains a quality variable for office space (1-5 star rating), but there were limitations on the amount of data that could be used (368 observations for the year 2016).
- Several different formulations have been tested, and most of the variables remain significant, with relatively stable coefficients across the different models – which is encouraging for the robustness of the model despite the small sample size. Adjusted $R^2$ is 0.70. The *accessibility* variables pose a number of challenges, including:
  - ideally we would include Walk and Rail as separate modes, for each of the two types of accessibility (above) – correlation makes this difficult in practice, and full solutions require a larger dataset;
  - for accessibility to businesses (& other workers) we would want to use evidence-based deterrence functions, similar to those used in the Residential model, and the decay parameters provided by Graham et al (2010) for DfT provide a starting point for this. However, the -1.8 decay parameter for Consumer and Producer Services found by Graham et al (2010), and even the default value -1.0 recommended for TAG (formerly WebTAG), lead to very steep deterrence functions when applied to inter-city travel. A person at office location *j* can only 'see' 10% of the jobs at another office location which is 3 mins walk away using the -1.8 decay

parameter (or 10 mins away using the -1.0 decay parameter). Given the time taken to reach a station, board a train and travel one stop, very few workers in other localities or cities would be reachable by rail within the time implied by the deterrence function. We need to establish whether this is a problem or not, and this suggests further investigation is needed.

- The Initial commercial model was estimated using the TAG decay parameter (-1.0), applied to effective density. Subsequent models used the Walk Travel-to-Work deterrence function (from the Residential model) instead (Figure 3.10 gives a comparison).
- The main findings relating to the variables in the commercial models are:
  - Effective density has a significant positive effect on commercial rents (per square foot) – whether the TAG decay parameter or the Walk accessibility (TTW) deterrence function are used;
  - Distance to the nearest rail station is significant (lower rent at greater distances, as expected) – it is not yet clear to what extent this is due to access to work or agglomeration/business clustering effects;
  - Attempting to separate proximity to a station (in Leeds or Bradford) from walk accessibility to other jobs, an average 12% premium is found for commercial property walkable to the station[9].
  - Quality matters – 1 or 2 star office rents are ~30% below rents for 3 star offices, and 4 or 5 start office rents are ~55% above 3 star offices;
  - Industrial rents are ~70% lower than 3 star Office rents, per sq ft, all else equal – consistent with wider evidence;
  - Rent per sq ft decreases slightly as the amount of space rented increases; and
  - Income (at LSOA level) serves as a proxy for place quality in this model, assuming that certain local facilities and amenity are influenced by household incomes in the locality.
- A larger dataset such as the VOA business ratings data would allow the model to be extended across the whole TfN area, and properly consider the role of rail connectivity between cities and other centres. It could also allow more of the available variables (e.g. on place quality) to be included. This is probably the most pressing need for improvement.
- Further details of the results can be found in Section 4.2.

## *Time Series Residential Model – Manchester Metrolink*

- A *case study* approach was chosen, since rail accessibility changes have been relatively limited across the TfN area as a whole over the last 20-25 years.
- Greater Manchester was selected as it has seen significant improvements in its rail-based mass transit network during the period for which detailed property price data is available (1995-2018), including the opening of Metrolink lines and stations.

[9] The literature indicates that the commercial rent premium is usually highest within 400m of a station (e.g. Debrezion et al., 2007). Whilst we find the expected relationship with distance from a station, we also find some effect out to 1500m, which requires further investigation in future work.

- Data was gathered with the assistance of TfGM.
- This model shares the hedonic price (HP) basis of the cross-sectional model, but uses a panel dataset (mixed time series-cross sectional data) and a quasi-experimental approach to identify the impact of new Metrolink stations – i.e. a 'treatment' vs 'control' approach to capture the value of the intervention.
- The model is built at LSOA level (~700 households per LSOA). Observations for 37 of the 1,673 LSOAs were dropped where there was more than one missing value for a particular year, leaving a useable sample of 24 years * 1,636 LSOAs = 39,264 observations.
- The model shows that for past changes in accessibility due to Metrolink network expansion, we observe a positive and statistically significant uplift of 6.3% from becoming in close proximity to a Metrolink station (within 1km). These uplifts are for average sold prices in each LSOA, having controlled for average house price increases across the Greater Manchester area, property type mix, and unobserved time-invariant LSOA-level characteristics ('fixed effects').
- A further model breaks the effects down by Line (Table ES6). There are significant uplifts for stations on the Airport Line (+20.6%) and the South Manchester Line to East Didsbury (+10.5%). Other lines do not show a significant uplift at 95% confidence, although there is evidence of an uplift from the East Manchester Line (+7.5%) at a lower confidence level.
- The pattern of uplift is similar in an outer ring 1km-2km from the new Metrolink stations; the uplift is smaller than in the 0-1km catchment, as expected (Table 4.17).

Table ES6: Property value uplifts within 1km of new Metrolink stations, overall and by Line

| | *Uplift within 1km of station* | *T-Ratio* |
|---|---|---|
| New Metrolink stations (1999-2017) | +6.3% | 3.94 |
| *Line* | *Uplift within 1km of station* | *T-Ratio* |
| Airport | +20.6% | 9.31 |
| South Manchester | +10.5% | 4.12 |
| East Manchester | +7.5% | 1.59 |
| Eccles | -3.4% | -0.56 |
| Rochdale | -1.1% | -0.54 |

- Although the Airport line passes through a number of deprived LSOAs[10] around Wythenshawe, it is distinctive in having a major concentration of employment at each end (Manchester City Centre and the Airport), which in the cross-sectional model would increase 'accessibility to employment' from intermediate stations. The South Manchester line also passes through some of the more affluent areas in the south of Manchester, in terms of household income.
- Parts of the Rochdale, East Manchester and Eccles lines serve areas of higher deprivation and relatively low income. It is probable that low underlying levels of property demand constrained the growth in prices in these places. The cross-

[10] in terms of the 2015 Index of Multiple Deprivation (IMD) (Smith et al., 2015).

sectional modelling would suggest that uplifts due to rail accessibility improvements would be lowest in areas with low LSOA-level average incomes (Tables ES2 and 4.5).

- The Rochdale line replaced an existing rail line so potentially the step change in accessibility was smaller. The Eccles line is a complex case with urban regeneration taking place, Media City opening and a substantial increase in housing supply in parts of the route. The result in Table ES6 is for the Eccles Line as a whole: further disaggregation of the line into sections might be needed to be able to say more. Also the result above is specifically for the effect of opening Metrolink stations – taking better account of the other factors may be important in this case.
- Using a different method, a simple comparison between the Metrolink catchment annual price growth and the Greater Manchester area annual price growth, the Eccles Line does appear to have a small positive effect (8.6% per annum vs 6.5% per annum, from year of opening to 2018), however this does not feed through into the quasi-experimental model, i.e. it is not statistically significant and may be caused by non-Metrolink factors.
- The quasi-experimental method using treatment and control areas helps to address the concerns about causality, which cannot easily be demonstrated in the cross-sectional model. The model is still relatively new, with scope for further improvement.
- Further details can be found in Section 4.3.

### *Model Interpretation*

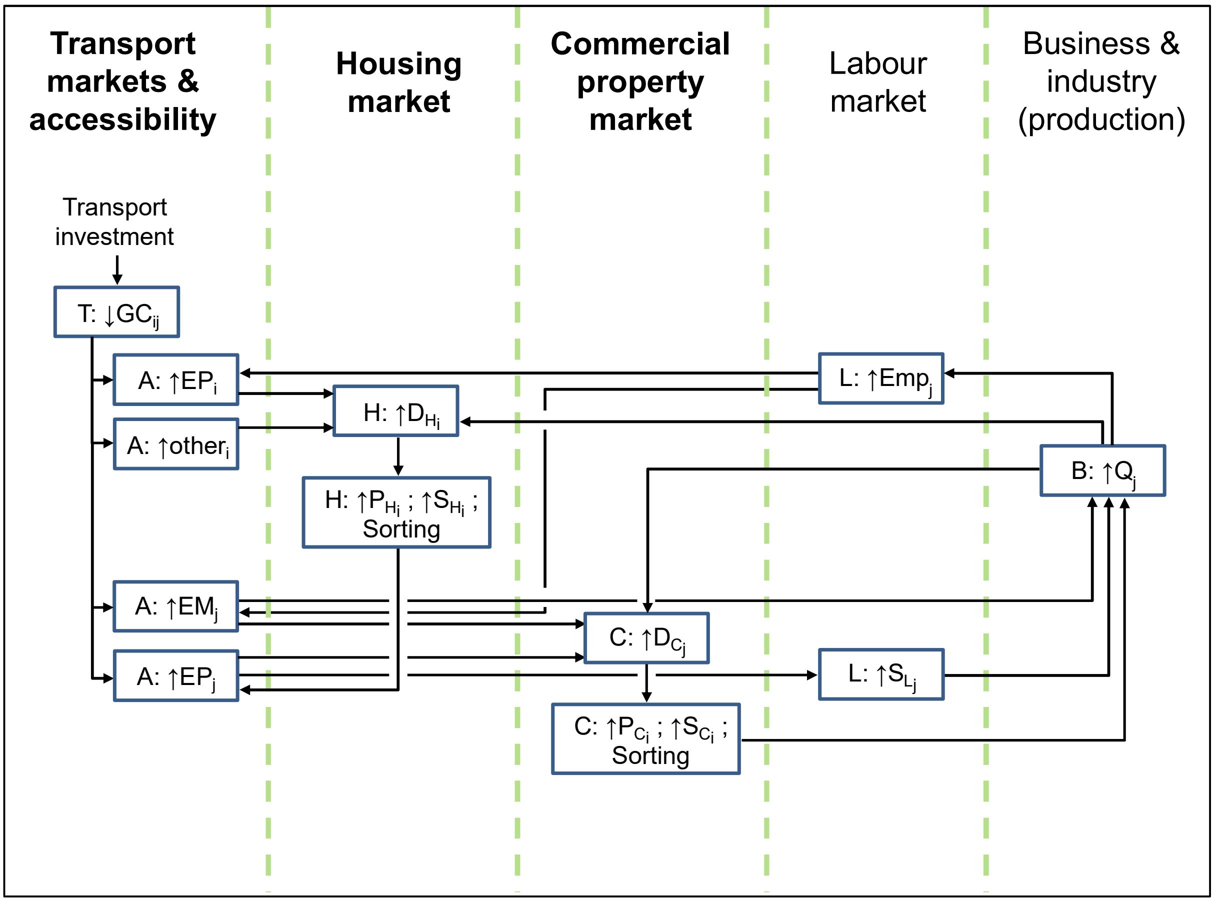
- The results of the Cross Sectional Residential Model are best seen as an indication of the amenity value of better accessibility to and from the home location. This is the value perceived by households, and expressed through willingness-to-pay (WTP) in the property market. It is – in a sense – the bedrock for all subsequent changes which 'market dynamics' bring about. The model controls for the effects of income, so the parameter estimates for accessibility always capture the value of access to more and better opportunities at a given initial income level. The results provide a starting point for estimating the total property value impact of accessibility changes. They are applied to Northern Powerhouse Rail (NPR) in the next section of this report (Section 5.2).
- There may be further effects, although these are more difficult to predict using available models. Property market sorting behaviour (Tiebout, 1956; Kuminoff et al., 2013) may lead to higher-income individuals moving into the areas made more accessible. The models in this study do not predict this, but do indicate that if it occurred this could lead to substantial localised increases in house prices: e.g. the price premium between inhabitants in the 5$^{th}$ and 6$^{th}$ income deciles is +7.7% (implied by the cross-sectional model[11]). It should be remembered that property market sorting activity involves households changing places, so in some other locations there would be a downward effect on household incomes and hence property prices (due to the income parameter).
- Housing supply may respond to the localised increase in demand – subject to planning restrictions. If it does so, there may be some dilution of the price impact,

[11] £1,300 per annum increase in income (equivalised) * 5.93% per £1,000.

however new development can also reduce the disamenity caused by derelict brownfield sites, and can rebalance the housing stock towards types that are locally scarce. As the cross-sectional model shows, the presence of new build housing units itself creates an uplift in average value, all else equal (since the value of each new housing unit is 18% greater than an equivalent existing unit).

- There may also be effects transmitted through the wider economy (Figure ES2). Labour supply improvements may lead to production increases and improvements in the public finances (direct tax revenue net of benefit payments). Agglomeration effects may increase productivity and output. These effects would likely feed back into the regional or sub-regional demand for housing: they are not modelled as part of this study, however they could be modelled by bringing together the models developed here with other model types.

Figure ES2: Linkages between transport, property markets and the wider economy




Note: GC = generalised cost of travel from home at i to work at j; D = demand; S = supply; EP = employment potentiality (accessibility to jobs); EM = economic mass. See Section 5.1 for details.

- Many of these further effects are likely to increase property prices near to a new transport service, over and above the initial uplift from the amenity value of the accessibility change. However, there will be exceptions: e.g. the effects of increased housing supply; and sorting effects between different areas with an improved transport service (e.g. demand may gravitate to areas with both transport improvements and other attractive features).

- Finally, there is the issue of *hope value* and the speculative aspect of the property market. To some extent demand is driven by investment considerations (even when the purchaser is an owner occupier) and hence the future potential of the local market comes into play. Since future potential is by its nature uncertain, there is the potential for inflows and outflows of capital to/from local markets based on changing expectations of future prospects and potential capital gains from investment now.

- The quasi-experimental model for Manchester Metrolink (the Time Series model) is a different type of model in that it includes all the following responses – provided they have had time to occur between the stations opening and the end of the dataset in 2018:
  - the initial uplift from the amenity value of improved accessibility (to employment and other opportunities);
  - any property market sorting behaviour, and development of new housing units;
  - any effects transmitted via the labour market or productivity impacts; and
  - any speculative responses associated with hope value in the local market.
- This allows for a comparison between the cross-sectional and quasi-experimental model results. First, the average uplift from new Metrolink stations, 6.3%, is well within the range for Best-to-Worst rail accessibility differences in the North, 14.3% (in the cross-sectional model) (Table ES7). In other words, as expected there is a significant uplift but still potential for further value gain from rail accessibility improvements in these places, e.g. if they were – hypothetically – connected directly by regional express rail to many more jobs.

Table ES7: Comparison of Metrolink uplift and Best-Worst rail accessibility premium (quasi-experimental vs cross-sectional models)

| | | | % |
|---|---|---|---|
| Uplift: | | | |
| | Metrolink Average | | 6.3 |
| | Airport Line | | 21 |
| | South Manchester Line | | 10 |
| | (East Manchester Line) | | (7.5) |
| | Others | | insig. |
| Premium: | | | |
| | Rail accessibility to jobs: Worst to Best (TfN area) | | 14.3 |
| | … at 90$^{th}$ percentile income | | 19 |
| | … at 10$^{th}$ percentile income | | 9.8 |
| | NPR increase in premium: | max. | 9.3* |
| | … based on Model 14a | | |

Note: Metrolink uplifts in 0-1km catchment; *at Output Area (OA) level.

- Secondly, the largest uplift for Metrolink (the Airport Line) slightly exceeds the range of the cross-sectional model, suggesting that some of the further effects in Figure ES2 and discussed above, have come into play. Meanwhile the uplifts for some Metrolink lines, e.g. the Rochdale line, are not significant, suggesting that the lower incomes and possibly lower accessibility gains in this corridor can neutralise the Metrolink uplift in specific local circumstances.

- Thirdly, work was undertaken to compare the rail accessibility premium at a South Manchester commuter rail station (in the cross-sectional model) with the Metrolink uplifts in the same area (South Manchester Line). Whilst this is not an exact comparison, because the form of rail accessibility is slightly different, the results are interesting (Table ES8). They indicate again that the uplift can (locally) exceed the initial accessibility premium, as a result of these further dynamic effects. This is true in both the 0-1km and 1-2km buffer zones, although the difference is particularly pronounced in the 1-2km buffer around the station. This area has slightly above average income (compared with Greater Manchester or compared with the TfN area as a whole).

Table ES8: Comparison of the rail accessibility premium and Metrolink uplift in South Manchester

<table>
<tr><th rowspan="2"></th><th rowspan="2">Premium (Cross-Sectional Model)</th><th colspan="2">Uplift (Quasi-Experimental Model)</th></tr>
<tr><th>0-1km</th><th>1-2km</th></tr>
<tr><td>250m from Commuter Rail station</td><td>+8.5%</td><td rowspan="3">+10.6%</td><td rowspan="3"></td></tr>
<tr><td>500m from Commuter Rail station</td><td>+6.5%</td></tr>
<tr><td>1000m from Commuter Rail station</td><td>+3.3%</td></tr>
<tr><td>1500m from Commuter Rail station</td><td>+1.1%</td><td rowspan="2"></td><td rowspan="2">+7.4%</td></tr>
<tr><td>2000m from Commuter Rail station</td><td>0 {BASE}</td></tr>
</table>

- It is worth noting that the uplift compares the Metrolink 0-1km and 1-2km buffer zones (the ‘treatment’ areas) with the rest of Greater Manchester (the ‘control’). If higher-income households are being attracted into the Metrolink buffer zone from the control area, there may therefore be some relative[12] reduction in values in the ‘rest of Greater Manchester’.
- In summary, the cross-sectional model provides an initial estimate of the value of rail accessibility, seen through the eyes of existing residents[13]. The outturn property uplift may be either higher or lower than the cross-sectional model suggests. There are a number of mechanisms through which further effects could occur: one is the relocation of households into or out of the locality, and the effect of their income on property prices. Other mechanisms were mentioned above and are discussed in Section 5.1.

### *Northern Powerhouse Rail*

- Northern Powerhouse Rail (NPR) is a major strategic rail programme, designed to transform connectivity between the key economic centres of the North. The

[12] Not necessarily observed as absolute price reductions, since they are likely to be spread across a wide area (therefore diluted), and may well be outstripped by general house price inflation.

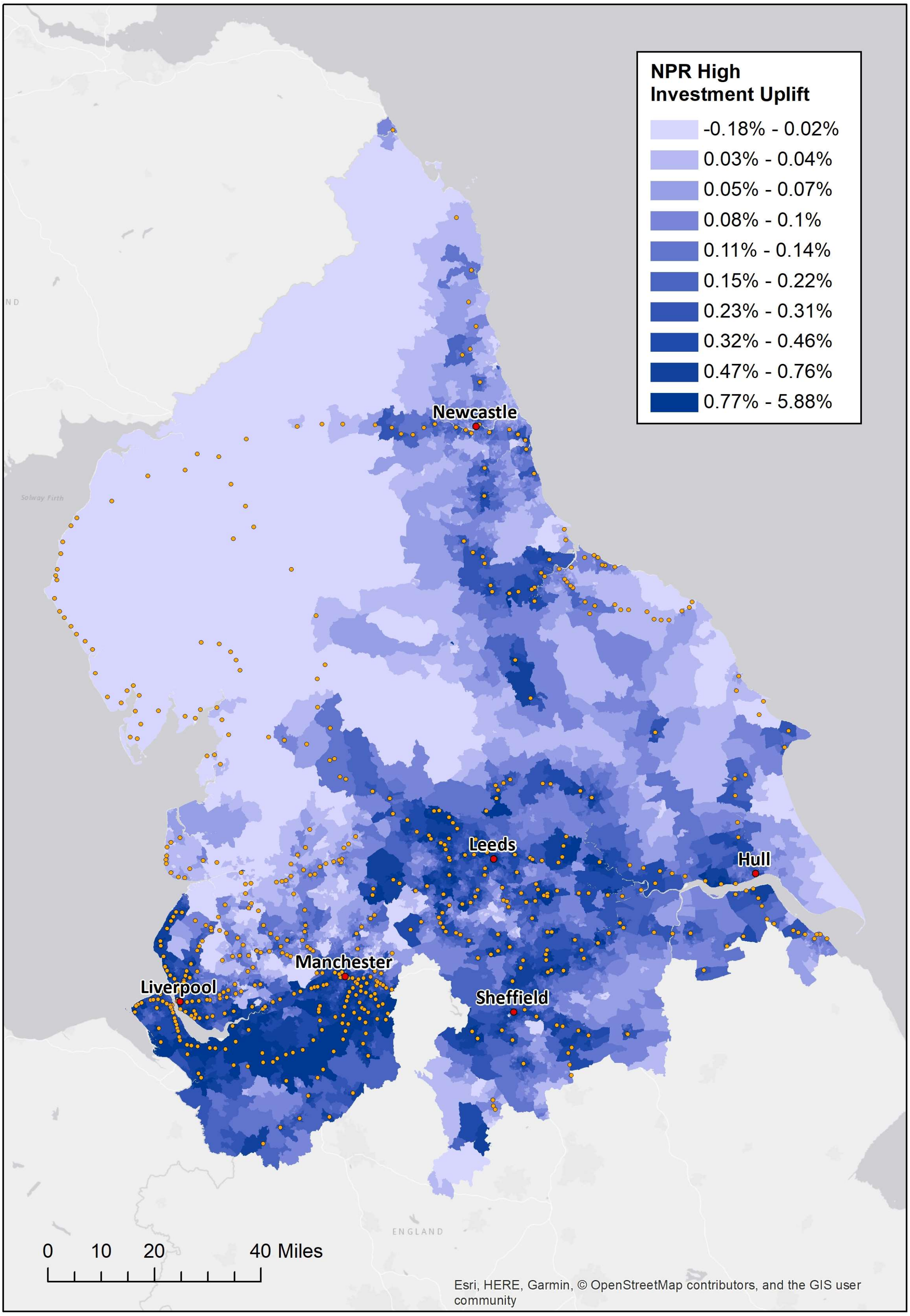
[13] And focusing on the amenity value of rail access to employment and other opportunities, rather than the ‘hope value’ in a rising asset market.

programme promises radical changes in service patterns and target journey times[14].

- An initial Strategic Outline Business Case (SOBC) for NPR was finalised by TfN and DfT in early 2019 and endorsed by TfN partners in February 2019. Based on this SOBC, the Department for Transport (DfT) agreed that NPR should continue to be developed as a scheme through 2019-2020. The initial SOBC made a broad strategic and economic case for the NPR network overall but left a number of potential network concepts on the table. As a complement to the transport and wider economic modelling carried out by TfN and DfT, the analysis presented here was included in the Strategic Case of the initial SOBC.
- The aims of the (upcoming) finalised SOBC are to narrow this down to a single network concept and optimise and refine the Strategic and Economic Cases. TfN is in the process of considering how further analysis of Land Value Uplift can form part of this case.
- As part of this work supporting the SOBC, policy scenarios for NPR were tested against a Do-Minimum.
- The parameters from the Cross-Sectional Residential model were used to investigate the potential change in the pattern of property prices across the TfN area, based on the following different scenarios:
  - an NPR High Investment scenario;
  - an NPR Medium Investment scenario;
  - a Do-Minimum scenario.
- Each scenario was defined by MSOA-to-MSOA generalised journey times (GJT) across the rail network, based on outputs from TfN's Northern Rail Modelling System (NoRMS) as the basis for the changes in rail accessibility. These GJT measures included access and egress time, in-vehicle time, crowding penalties, interchange penalties and a measure of delay time – a relatively comprehensive GJT metric.
- The Rail accessibility to employment measure in the property value model responded to these changes in GJT. The pattern of Rail accessibility to employment in the Do-Minimum is shown in Figure ES1 – the NPR improvements are all relative to this.
- Using the cross-sectional model (Model 14a), improvements in rail accessibility to employment in the NPR High Investment scenario could potentially produce uplifts in residential property values of up to 9.3% for Output Areas (OAs) – the areas with changes approaching this magnitude are generally close to stations, are very well-connected by rail to employment across the North and gain maximum benefit from NPR. This does not take account of any subsequent changes in the locations of households or businesses or the locations of investment (e.g. in residential or commercial property development) in response to the changes in the pattern of accessibility created by NPR. Figure ES3 maps these potential uplifts at LSOA level – since LSOA are larger than OAs the effect is diluted slightly and the largest impact at LSOA level is 5.88%.

[14] https://transportforthenorth.com/northern-powerhouse-rail/

Figure ES3: Potential residential property price changes implied by the NPR High Investment scenario (LSOA level)



- An Excel-based tool was developed to measure the total value accumulated in the housing market as a result of a particular policy test.
- Using the coefficients from the cross-sectional model, the total expected uplift or increase in value in the residential property market was estimated (Table ES9). The results are presented in three ways:
  - a snapshot for a single year assuming instantaneous change (based on 2017 property values);
  - a cumulative estimate over 8 years from NPR opening in 2033/34, to 2040/1, at the current nominal prices at that time (assuming the time profile of land value uplift is at a steady compound rate over 8 years from opening, based on comparisons with other recent projects where TfL have gathered this data); and
  - a Net Present Value (NPV) – using 2018/19 as the base year for discounting and 2010/11 as the price base year, as requested by TfN (otherwise the usual TAG price inflation and discounting rules apply).

Table ES9 Total potential value uplift due to NPR rail accessibility improvements, TfN Area (Residential)

| | *NPR High Investment* | *NPR Medium Investment* |
|---|---|---|
| Single year – instantaneous change (2017) | £2.9 billion | £2.5 billion |
| Cumulative over 8 years from opening in 2033/34-2040/1<br>(at nominal prices)<br>(NPV) | <br><br>£6.2 billion<br>£1.9 billion | <br><br>£5.3 billion<br>£1.6 billion |

- These results may be conservative in their estimation of potential uplifts, given the findings from the latest Cross-Sectional Modelling, which have identified:
  - a slightly higher rail accessibility parameter ($1.56*10^{-7}$ compared with $1.37*10^{-7}$) ;
  - positive income interaction which would increase uplift in areas with higher average household income (see Table 4.7 – Model 14a).
- In principle, it would be possible to compare these total uplifts with the transport user benefits measured in the CBA, and with the forecast user benefits. As the modelling tools used for the initial SOBC are undergoing further development, the results were not considered sufficiently finalised to share with this project for publication. However, TfN and DfT do intend to make this comparison in future. For completeness, this comparison would ideally include an estimate of commercial property value uplift, which has not been included in the analysis presented here.
- TfL has carefully studied Land Value Capture mechanisms and concluded that mechanisms used for Crossrail 1 capture only a small share of the total potential uplift (8-10%), however up to 30-40% is possible if more ambitious mechanisms, addressing existing properties in particular, are developed and used. The potential

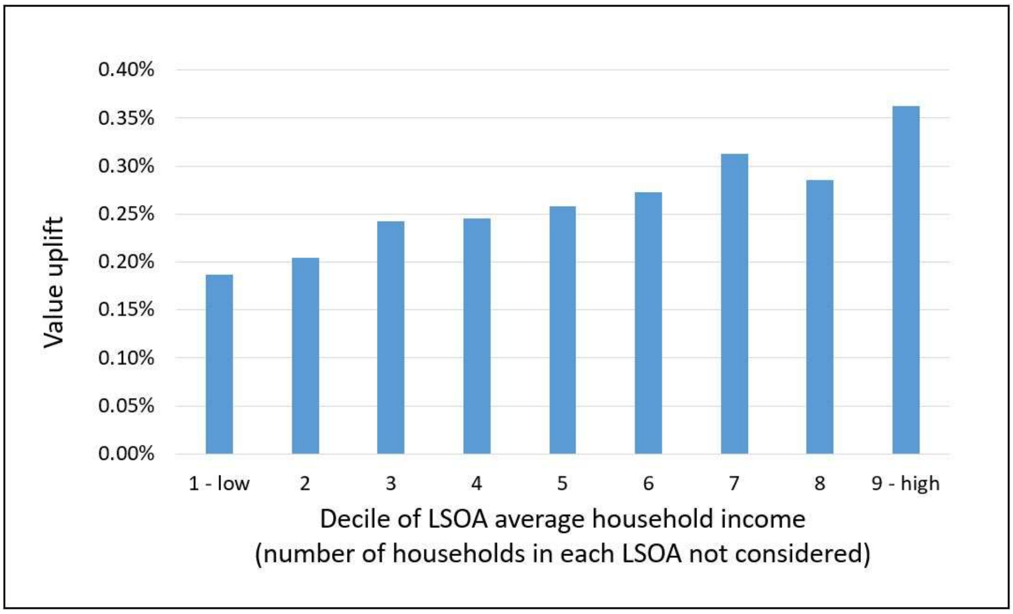
for revenue generation from the large total uplifts above needs to be seen in this light.

- There is a potential second-round effect: if NPR changes the spatial pattern of employment or the level of employment in city centres, this should lead to further land value gains (and this potentially has implications for housing and spatial planning in conjunction with NPR). We have explored this using a policy test, as follows.
- A further policy test was carried out separate from NPR, to understand the impact of increasing employment in economic centres. This could be due to, e.g., a change in urban planning policy to allow increased density, substantial redevelopment of brownfield/underutilised sites, or as a second-round effect of other policies which make central locations more attractive: such as investment in the rail network. The test modelled was a 10,000 increase in city centre employment in two central LSOAs in Leeds. This could be expected to produce a 1-3% increase in city centre residential property values, and up to 1% in the area outside that, to about a 12km radius in Leeds. Therefore if NPR changes the spatial pattern of employment or the level of employment in city centres, this should lead to further land value gains.

## *Distributional Impacts*

- The analysis undertaken so far allows us to extract some preliminary results on the distributional effect of the property value uplift from changes to the network. These results are based purely on the rail accessibility effect – including income interaction, without any of the further effects shown in Figure ES2. Figure ES4 reports the estimated average uplift by income decile, based on an LSOA level analysis. Decile 10 is missing because some of the data in this decile requires further cleaning to be used for this type of analysis; the other deciles are believed to be unaffected. The results suggest that higher income households will be affected more strongly by the property value uplift effects of a project to improve existing rail services.

Figure ES4: Potential property value uplift by average income decile (LSOA level)



- Figure ES5 provides a map of the predicted property value uplift with household incomes marked, for a selected urban area (Leeds). This highlights how uplift may extend to higher-income areas even some distance away from the rail stations with improved connectivity, whilst the LSOAs closest to the stations tend to experience uplift even where income is lower.

- Since we also have tenure data at LSOA level, it is possible to infer how the property value uplift affects areas with more or fewer renters. Figure ES6 shows that in the City Centre and denser suburbs near the Universities, there are larger uplifts associated with high shares of renting. In the outer suburbs, uplifts appear to impact on a larger share of owner-occupiers.

- In combination, these results make it possible to infer the pattern of impact on owners and renters at different income levels, for whom the welfare impact of the change in property values will differ. Owner-occupiers, landlords and renters will experience these impacts in different ways (there are wealth and income effects to be accounted for). Existing owners, first-time buyers, and downsizers will also experience the impact differently. An age breakdown of the impact is also possible.

Figure ES5: Pattern of potential property value uplift by income at LSOA level, Leeds

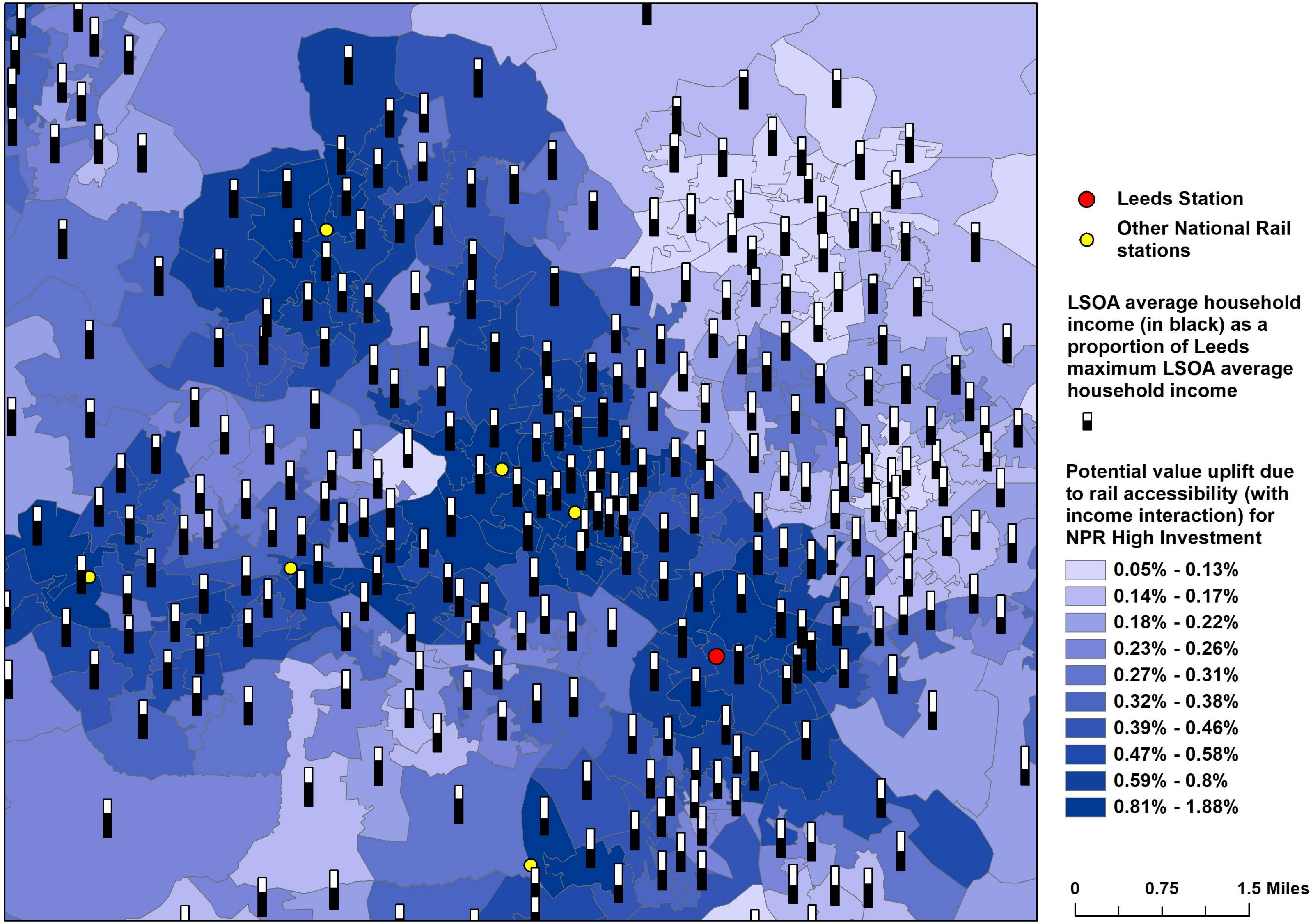
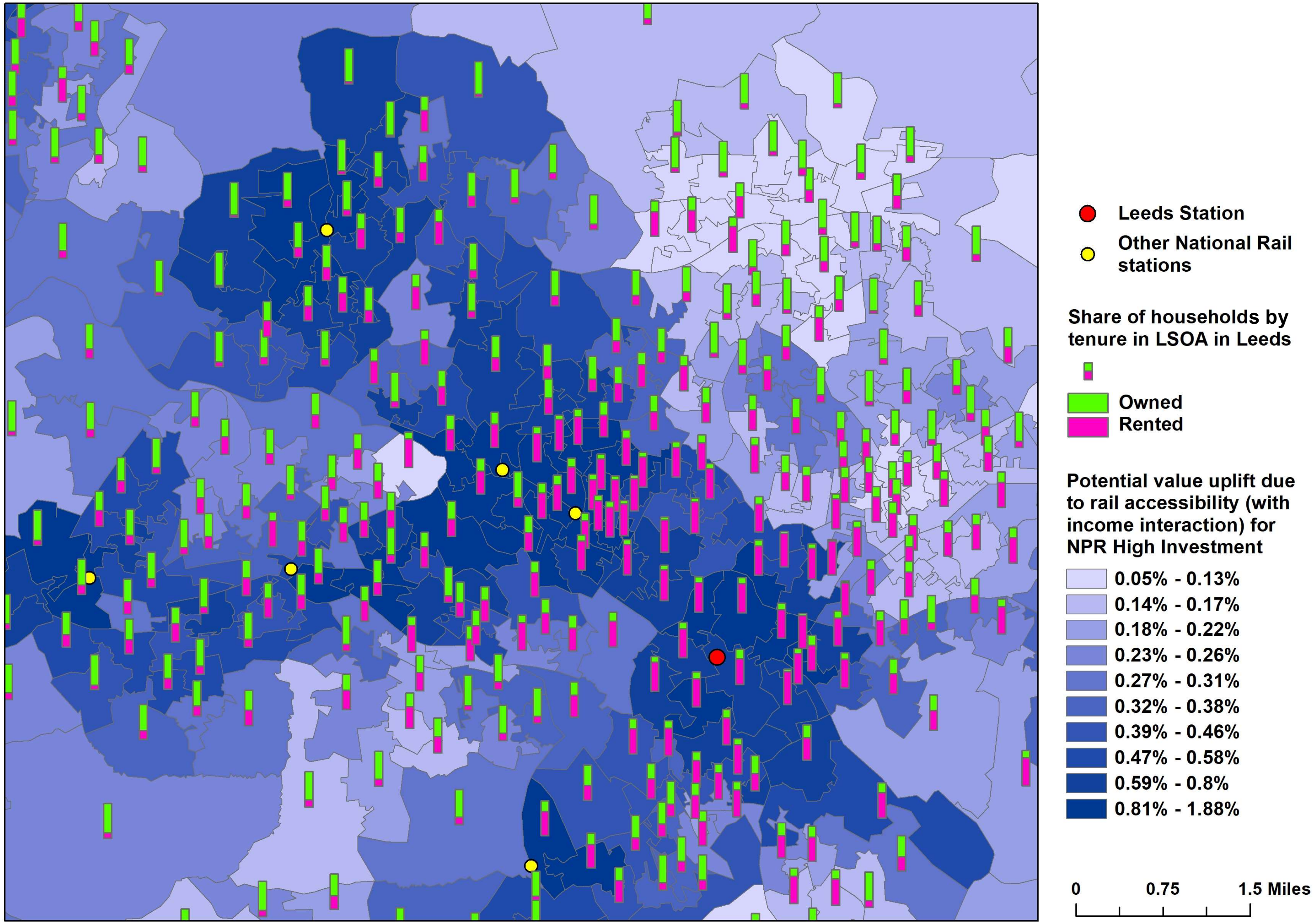


Figure ES6: Pattern of property value uplift by tenure (share of owned/rented), Leeds



## *Integration with Appraisal*

- Finally, major transport projects in England are required to have a Business Case (pre-implementation). At the heart of the Business Case is an Economic Case which is an assessment of the welfare (wellbeing) implications of the project, alongside a Strategic Case and a Financial Case – which together form three parts of the 'Five-Case Business Case' (DfT, 2013/2018). The role of the Business Case is to ensure that decision-makers are well-informed about the consequences of the proposed project and any alternative options within or around it. This evidence-based approach has stood the transport sector in good stead in Public Spending Rounds, and is part of the current practice in this sector internationally, not only in the UK.
- Table ES10 summarises this study's findings on how appraisal may need to evolve in future in order to integrate evidence about land value change within the Transport Business Case. Implementation would be subject to further methodological development. The main points relate to:
    - the need to represent both gains and losses to different groups in the appraisal – this is good practice anyway, and is reinforced in DfT's recent guidance on investment strategy (DfT, 2017a);
    - the need to see transport policy together with housing and economic development policy, not in isolation from it (again this is in line with DfT's stance on strategy – DfT, 2017a);
    - the potential role of land value capture (LVC); and
    - the issues that exist in measuring the welfare changes to people in the study area.

Table ES10: The roles of land value change in the Transport Business Case

| | *Type of analysis* | *Implications (from the study)* |
|---|---|---|
| Economic Case | Welfare economics, Social CBA | NPV=PVB-PVC<br>• Who benefits, who loses<br>• Dependent development<br>• Wider labour & productivity impacts<br>• LVC contributes to ↓C |
| Financial Case | Financial flows, Funding and financing arrangements, Financial sustainability check | • LVC contribution to Costs<br>• Borrowing against future LVC |
| Strategic Case | Policy fit, Objective achievement | • Policies relating to housing affordability<br>• Policies relating to land use and economic development |

CBA=Cost Benefit Analysis; NPV=Net Present Value; PVB=Present Value of Benefits; PVC= Present Value of Costs

- When land values change, e.g. as a result of a transport investment, it is expected that there will be both benefits and disbenefits as shown in Table ES11. The final pattern of benefits and disbenefits will depend on the interplay of the effects in Figure ES2, involving households, businesses, the land and property markets and the labour market.

Table ES11: Initial benefits and disbenefits of land value change

| *Impact group* | *Impact* | *Potential magnitude (Northern context)* |
|---|---|---|
| Owner occupier residents | Near to the new facility: land value uplift<br>• **capital value gain (windfall) (+)**<br>• potential Council Tax and insurance cost increases (−) | Moderate: cross-sectional and quasi-experimental models point to approx. range 0 to +20% (for interventions considered) |
| | Distant from the new facility: land value change of undetermined sign | Small: effect spread over a wide area |
| Renting residents | Near to the new facility:<br>• **rent ↑ (−)**<br>• potential Council Tax and insurance cost increases | Related to property value uplift. |
| Investors (residential and commercial) | Near to the new facility: land value uplift<br>• **capital value gain (windfall) (+)**<br>• **rental income ↑ (+)** | Moderate: cross-sectional and quasi-experimental models point to approx. range 0 to +20% (for interventions considered) |
| Future movers (in) | • prices, rents, SDLT on purchase and insurance costs increased (increased cost of living) (−) | Related to property value uplift. |
| Business occupiers | • **rent ↑ (−)**<br>• capital gain (if owners) (+)<br>• NNDR & insurance costs (−) | Magnitude uncertain: some evidence of rail accessibility premium approx. 12%; uplift may exceed this locally. |
| Local Authority | Potential ↑ in Council Tax and NNDR revenue (net across areas) (+) | Total limited by infrequent revaluations and low % rate. |
| Central Government | Potential ↑ SDLT and CGT revenue (net across areas) (+) | Total limited by applicability to movers/sellers but not to existing property, and by low % rate and exemptions. |

SDLT = Stamp Duty Land Tax; NNDR – National Non-Domestic Rates (revenues shared 50% locally, 50% pooled nationally); CGT = Capital Gains Tax.

- The wellbeing impact of these changes for residents will depend on their magnitude and the circumstances of the people impacted, including their household income (HM Treasury, 2018, outlines current official thinking on distributional weighting). Table ES10 highlights the potential complexity of these calculations. Relocation of households triggered by these changes might imply some further losses from the dislocation of an undesired move – potentially to a less well-connected location.
- In the case of *dependent development*, where transport improvements unlock development on a particular site or sites, the net value of the additional development may be counted as a measure of benefit under certain conditions (TAG Unit A2.1, DfT, 2018; DfT 2017b). However, due to the uncertainty

surrounding such estimates, these may only be included as 'indicative monetised impacts' which are not shown to the decision maker as part of the 'Initial Benefit:Cost Ratio (BCR)' or the 'Adjusted BCR'. Instead they should be presented as separate information[15]. Nevertheless, the dependent development (induced investment) case points towards the important role of development as a both a market response and policy response to rising accessibility premiums when transport investment occurs.

- Seen at a regional level, dependent development represents the relaxation of the supply constraint on housing (in particular) and also commercial property through transport investment. Barker (2003) highlights that there may be a welfare loss from an inefficiently-low level of housing supply, and by implication a welfare gain from a strategy to help to ease these constraints. Targeted transport investments as part of an overall strategy should serve to ease these constraints. A key concern for economists should be to establish the extent of any externalities or market failures in the land and property market, so that the welfare implications can be quantified.
- In some cities, including Manchester (Manchester City Council, 2018) and Helsinki, for example, there are explicit policies on housing affordability, which emphasise the role that housing costs play in achieving the cities' strategic economic and social objectives (Table ES12). For example, policy may target housing costs that are competitive with other cities (nationally/globally) to help attract mobile workers, and housing affordability may be part of achieving inclusive growth/social cohesion. These policies highlight that while accessibility improvement has economic benefits, the land value uplift it creates can be a variable to be managed rather than maximised. Increasing housing supply can be a part of that management – including policies for 'transit oriented development'.

Table ES12: Manchester Housing Affordability Strategy

| *Manchester Strategy outcomes* | *Summary of the contribution to the strategy* |
|---|---|
| A thriving and sustainable city: supporting a diverse and distinctive economy that creates jobs and opportunities | Ensure Manchester has the right mix of housing that is affordable across a range of tenure and income levels to support a functioning Manchester and sub regional economy. |
| A highly skilled city: world class and home grown talent sustaining the city's economic success. | The new and existing homes will be well connected to employment opportunities and schools. |
| A progressive and equitable city: making a positive contribution by unlocking the potential of our communities | Increasing the supply of good quality affordable homes for sale and rent will provide the opportunity for Manchester residents to raise their individual and collective aspirations. |

Source: extract from Manchester City Council (2018).

- As in Figure ES2, the outcome from a transport investment will be the result of interactions between several linked markets, including labour, land and property, and business and industry. The outcome *may be* the reshaping of the economic geography, with relocation of businesses to more productive locations, households to locations where they can access more productive jobs, benefits through labour supply and agglomeration, etc. All of this is subject to economic analysis and

[15] From a financial perspective such dependent development may be a source of revenue through Land Value Capture mechanisms (e.g. in Crossrail 2 and the East Leeds Orbital Road – TfL,2017; Leeds City Council, 2017) and hence reduce the Cost to the Broad Transport Budget (C) in the Economic Case.

modelling that goes beyond the scope of this study. The contribution of this study is to better understand the specific relationships between accessibility, place quality and value as measured in the property market – and as such to make a contribution to that wider modelling effort.

## Conclusions

- Finally, the report reflects on regional differences in uplifts between the North and the South-East of England. There is some evidence of smaller uplifts in the North although this is far from universal – the models help to explain the factors driving uplift, and indicate how uplifts will vary from project to project, as well as from region to region.
- The study of the Jubilee Line Extension and the Docklands Light Railway extension to Lewisham by Gibbons and Machin (2005) provides a very useful comparator for the quasi-experimental model of Manchester Metrolink in this study[16]. Gibbons and Machin found that the average uplift in the 0-2km catchment was 9.3%. For comparison, the uplift in the same catchment on the South Manchester Metrolink Line is 8.2%. This is the most affluent of the Metrolink corridors – uplifts are generally smaller elsewhere, pointing to the role of income in determining the size of the rail accessibility premium. The modelling suggests job density and the size of the accessibility increase are also factors (the frequency of Manchester Metrolink is typically 9-10 per hour, compared with 22-23 per hour on the JLE). The Metrolink Aiport Line is an exception – with a relatively high uplift of 12.3% in the 0-2km catchment – probably due to the concentration of employment at both ends of the line and the opportunity provided by international connectivity, rather than to income alone[17].
- The parameters estimated in this study – specifically in the cross-sectional model – were estimated on the whole TfN area, and should be widely applicable across the North of England as a measure of the underlying premium for accessibility to employment and other factors.
- The spatial pattern of land value uplift around new rail stations suggests that a mixed set of Land Value Capture tools would be required for rail projects, as it was for Crossrail or Barking Riverside, for example. Strong localised uplifts within 1km of a station will be accompanied by a wider but weaker impact pattern. Localised (and flexible) Section 106 agreements and land assembly strategies are well suited to capture localised gains from new development, whilst CIL and Business Rate Supplements/Retention can be applied in a more general way across a wider area. The challenge of capturing a significant share of value gains to existing residential properties remains an issue under English legislation – e.g. Council Tax Precepts are too limited and do not allow for the kind of fine spatial differentiation that these models show will be needed if the LVC instrument is to be related to the property value gain.
- Land value uplift may give rise to a complex pattern of gains and losses in a welfare analysis – not only spatially but by income, tenure, age and other factors. The evidence points to network effects being a helpful phenomenon in spreading the benefits from transport investment more widely, so that more of the population become ‘gainers’. A complete analysis of the impacts of a project will include the wider economic impacts via productivity and wages, for example, and the dynamic

[16] Gibbons and Machin (2005) also used a quasi-experimental method.

[17] Average income is markedly lower in the Airport corridor than the South Manchester corridor (Table 4.15).

effects through business or residential relocation. Thus land value uplift effects are only part of the picture – even from an individual's perspective.

- The key areas identified as evidence gaps where further investigation would be valuable include: dynamics of the system (Figure ES2) including land and property values; wide-area estimation of a commercial model; incorporation of additional built environment/urban realm variables; and potentially disaggregation of employment by sector/category to explore the role of skills matching in the regional housing and labour markets.

# 1. Introduction

*Background and Aims of the Research*

The aims of the study are both scientific and applied:

- The scoping study* found that there are important gaps in the understanding of, and ability to quantify, property value responses to transport infrastructure investment. There is therefore an opportunity for a quantitative model representing the complex relationships between transport/accessibility and property values across a specific study area.
- Moreover, there is a need to define more clearly:
    - how evidence on property value responses should be integrated with appraisal;
    - how Land Value Capture proposals can be informed by modelling evidence in the property market;
    - the distributional impacts of land value change; and
    - what can be inferred about the optimal timing and phasing of investment, from an understanding of the property market response.
- The study is designed with the specific application to Northern Powerhouse Rail (NPR) in mind. The NPR project aims to improve connectivity across the North of England in support of economic growth. Currently at Strategic Outline Business Case (SOBC) stage, NPR will be moving at the end of 2018 to the Outline Business Case (OBC) stage.

> * The 'scoping study' referred to here is the work on *'Transport and Land Value Uplift: Evidence and Implications for Appraisal, Modelling and Strategy*', also referred to as the 'Phase 1 study', which was sponsored by the West Yorkshire Transport Research Innovation Fund, and which identified research needs, with participation by DfT, DCLG, NIC, TfN, TfGM, WYCA, TfL, Regional Authorities in the North, property professionals and others. See Nellthorp et al. (2016) in the References list.

The central aim can be stated very simply: to build the evidence base, using available data and newly-developed models, on *the relationship between rail accessibility and property prices*, using the North of England as the study area.

The intention is to provide authorities at national & regional level with quantitative evidence which will help inform Business Cases for rail improvement. At the same time, the intention from an academic perspective is to take advantage of increasingly available data, and learn from recent modelling experience, to go beyond previous models of transport-property value interaction, addressing various limitations and issues which have arisen.

This work is not being developed in isolation. It fits with other modelling work being undertaken for NPR, including Land Use and Transport Interaction modelling ('NELUM') and rail demand modelling ('NoRMS'). The focus here is specifically on the link between accessibility and property prices, and the wider implications for the Business Case.

*Reporting Stages*

The purpose of this Final Phase 2 Report is to provide a complete set of key findings from the study.

Previous reporting stages were:

- Advisory Panel Discussion Paper (21st June 2018);
- Interim Report: Modelling (31st August 2018);
- Input to the NPR Strategic Outline Business Case (31st October 2018); and
- Final Report Draft to Advisory Panel (21st January 2019).

*Sequence of the work*

The overall sequence of the work is shown in Table 1.1. Out-turn was slightly later than plan for milestones A-C due to the time taken to bring the team together at the start; the project caught up with the planned timescale from August 2018 onwards.

Table 1.1: Project Timescale

| Month | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | Writing-up |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Jan 2018 | | Mar 2018 | | May 2018 | | July 2018 | | Sep 2018 | | Nov 2018 | | Jan-Aug 2019 |
| Initial Development | | | | | | | | | | | | | |
| Interim Development | | | | | | | | | | | | | |
| Final Development | | | | | | | | | | | | | |
| NPR Results | | | | | | | | | | | | | |
| NPR SOBC | | | | | | | | | | | | | |
| Final Outputs | | | | | | | | | | | | | |
| Milestone | | | | A | B | | C | D | | E | | F | G |
| Advisory Panel: | | | | | | | | | | | | | |
| Report (comments) | | | | * | | | * | | | | | | * |
| Meeting | | | | | * | | | | * | | | | * |

The Initial Development stage of the work covered all workstreams: (i) Modelling – the main workstream; (ii) Integration with appraisal; (iii) Land Value Capture; (iv) Distributional impacts; (v) Temporal dimension.

Modelling work was then prioritised from June-September, using the full study team.

In October, input was provided for the NPR SOBC which included:

- modelling results;
- policy tests based on those results;

- aggregation of value changes across the housing stock in the TfN area;
- projection of value changes over the appraisal period;
- mapping the spatial pattern of value changes;
- interpretation of the findings.

The Final Outputs stage of the Phase 2 Study included further modelling work on the three main model types, leading up to December 2018. The writing-up stage has included further minor model extensions and optimisation, and the writing-up process itself.

*Sponsorship and Acknowledgements*

This study was sponsored by Transport for the North (TfN), West Yorkshire Combined Authority (WYCA) and the Engineering and Physical Sciences Research Council (EPSRC). The authors are grateful to Jack Snape (TfN) and Patrick Bowes (WYCA) for their careful guidance and support throughout the work. Members of the study's Advisory Panel contributed to three very useful meetings and provided generous comments on drafts – those comments have been taken into account in preparing this final report, but responsibility for the findings and the views expressed remains with the authors. The Advisory Panel and client leads* were (with some changes over time): Tom Bousfield, NIC; Iven Stead, DfT; Henry Kelly, DfT; Josh Nava, DfT; Alistair Baldwin, NECA; Ian Palmer, TfGM; Olaolu Adeboye, TfGM; Ian Raymond, Merseytravel; Damian Walne, HS2; Professor Gordon Mitchell, University of Leeds; Professor Corinne Mulley, University of Sydney Business School; Tim Foster, TfN; Jack Snape, TfN*; Andrea Barry, TfN*; Patrick Bowes, WYCA*. The Study Team gratefully acknowledges the vital support received from these people and organisations, especially from TfN and WYCA.

*Structure of the Report*

The main Sections of this report cover the following aspects of the study:

- the Theory underlying the models (Section 2);
- Modelling Strategy and Data (Section 3) – which describes the empirical research strategy used to test the theory set out in Section 2;
- Results (Section 4) – presents and discusses the results for each of the three model types developed within the study, i.e.:
    - cross-sectional residential property models;
    - cross-sectional commercial property models;
    - time series (or 'panel data') residential property models;
- Other Workstreams (Section 5) – presents the findings on integration with appraisal, land value capture, distributional impacts and the temporal dimension;
- Conclusions (Section 6) – reflects on the findings and includes further research needs identified during the course of the work.

# 2. Land and Property Markets and Transport – Theoretical Background

The scoping study (see Nellthorp *et al.*, 2016, the Phase 1 study) identified significant gaps in the understanding of property value changes due to transport infrastructure investment. Property markets are complex, and transport is just one element of many that can influence prices. One of the needs identified by the Phase 1 study was to develop a theoretical framework capable of bringing these elements together, in order to provide a basis for the empirical modelling. The purpose of this section is to report on this work.

This section describes an economic theoretical framework developed to support the analysis of property prices in order to understand price sub-components or 'hedonic prices'. The practice of hedonic price modelling, including the applications to property markets, has traditionally been built upon the theoretical foundation laid out by Rosen (1974) to analyse implicit prices within differentiated products or composite goods. More recent developments include the work by Rouwendal (1992 and 1998) and Rouwendal and van der Straaten (2008). The framework below also largely follows Rosen (1974), but is set out specifically for the context of property markets, as in Rouwendal's work.

## 2.1 Property market and price determination

The fundamental presumption is that we are dealing with differentiated products (housing units; properties), which are fully described by a vector *z* of objectively measured characteristics (Rosen, 1974):

$$z = (z_1, z_2, \dots, z_n) \quad (1)$$

Where $z_n$ indicates the amount of the $n$th characteristic embodied in each property. The term 'objectively' refers to the condition that the characteristics present in each property is read in the same way by all agents (e.g. floorspace, semi-detached, two bedrooms, performance rating of the nearest primary school, distance walk to the city centre); each consumer can of course value them differently.

The central 'hedonic' assumption means that a property is valued for its utility-bearing characteristics. Hedonic prices are then defined as the implicit prices of those characteristics. Observed property prices and amount of characteristics $(z_1, z_2, \dots, z_n)$ act as channels through which implicit or hedonic prices for $z_1, z_2, \dots, z_n$ are revealed to economic agents. The (set of) implicit prices guides both consumer and producer choices in the characteristics space, i.e. agents buy and sell packages of characteristics (Rosen, 1974).

The price *p* of a property can then be defined as the price of a bundle or package of characteristics *z*:

$$p(z) = p(z_1, z_2, \dots, z_n) \quad (2)$$

This function $p(z)$ is identical to the set of hedonic prices. Assuming competition in the market and that a competitive equilibrium is reached, the price of a property $p(z)$ and hence the implicit prices of characteristics $z$ are determined by market clearing conditions following buyers' and sellers' maximising behaviour. Each agent individually has a negligible impact on prices and take these as given when making decisions. Distributions of consumer preferences (in the demand side of the market) and supply conditions together determine $p(z)$. Empirically, the implicit prices are estimated by a regression analysis where the property price is regressed on property characteristics (Rosen, 1974).

### 2.2 Theory of demand

A competitive property market leaves price determination to be an outcome of both, jointly, individuals' choices based on their preferences and sellers' choices based their assessment of market value and their costs. The interest here is understanding the demand function that underlies the hedonic prices, noting that, in practice, prices reveal simultaneously supply and demand structures (Rosen, 1974; Rouwendal, 1992). A clear theoretical understanding of underlying consumer decision-making process is fundamental to guide both the empirical estimation and the use of hedonic price estimates for policy-making, forecasting or appraisal purposes.

Consider households' preferences can be understood within a neo-classical utilitarian microeconomic framework in a conventional way. Households seek to maximise their utility subject to a budget constraint. Households' utility (U) is conventionally a function of housing consumption (H) and other non-housing goods (X). Both H and X can be equally defined as a housing composite good and a non-housing composite good respectively.

$$U = U(H, X) \tag{3}$$

Let us further assume that the utility that a household derive from a composite housing good or house (H) is a function of the set of property characteristics (*z*), such that $H = H(z)$.

$$U = U(H, X) = U(z, X) \tag{4}$$

The budget constraint can be written as:

$$M = p(z) + p_x X \tag{5}$$

where $M$ is household income, $p(z)$ is the price of the property (defined by the vector of *i* characteristics *z*) such that $p(z) = p(z_1, z_2, \dots, z_n) = \sum p_i(z_i)$, and $p_x$ is the price of the composite non-housing commodity X.

We have assumed that a household only consumes one unit of housing, but the model can easily be expanded to accommodate multiple units as shown by Rosen (1974). This practical simplification allows the model to resemble most housing consumption choices, where a household has previously decided to buy only one unit. Thus the choice households make, and also that of interest here, is a choice among different bundles of characteristics. Utility maximization can consequently be framed in the space of characteristics.

There are obvious issues in framing the problem in terms of choice of characteristics, particularly that properties are indivisible, and an assumption of divisibility may seem to be needed (Rouwendal, 1992). However, in line with a similar argument made by Rosen (1974, p. 36-37), we can argue that this assumption is feasible "if it is assumed that a sufficiently large number of differentiated products are available so that choice among various combinations of *z* is continuous for all practical purposes". This practical assumption also enables us to formulate the problem such that it is closely analogous to standard consumer theory.

The Lagrangian from maximising equations (3) (or 4) subject to (5) is:

$$L = U(z, X) - \lambda(p(z) + p_x X - M) \tag{6}$$

Noting that $p(z) = \sum p_i(z_i)$, the first order conditions are:

$$\frac{\partial u}{\partial z_n} = \lambda \frac{\partial p(z)}{\partial z_n} \quad \forall\, n = 1 \dots N \tag{7}$$

The first order conditions are analogous to those of a standard consumer theory problem. It can easily be seen that conditions in (7) state that the marginal price of the $n$th characteristic should be equal to the ratio of the marginal utility of the $n$th characteristic and the marginal utility of income. Optimality is achieved by the purchase of the house that offers the desired combination of characteristics (Rosen, 1974).

Econometric estimation can reveal the implicit prices $\partial p(z)/\partial z_n$ of each $n$th characteristic, which can be interpreted as the marginal rate of substitution between the $n$th characteristic and household income.

To add detail to the analysis and discussion, the generic set of characteristics *z* can be decomposed into several subsets of characteristics that share common features, including: the accessibility to jobs and skills from the property, which we refer to as 'accessibility' (*A*); the characteristics of the property surroundings such as environmental quality, local amenities and safety which can be referred to as place quality characteristics (Q); and the physical building and plot characteristics (*B*). We also recognise that the utility of living in a particular neighbourhood may be related to the socio-economic characteristics of the neighbourhood (*G*) in some way. Additionally we will allow for the supply-demand balance to differ across local areas, represented by *S* in equation (8), which is discussed further in Section 2.7. The utility function from (3) and (4) can therefore be further defined as:

$$U = U(z, X) = U(A, Q, B, G, S, X) \tag{8}$$

In the next sections, we analyse further theoretical details in relation to each relevant subset of characteristics, namely *A, Q, B, G* and *S*, where each set can comprise multiple characteristics. Specific characteristics will also be discussed where relevant.

### 2.3 Accessibility (A)

The first two categories – accessibility and place quality – may be classed together as the 'external characteristics' of the property, in the sense that their level of provision is subject to decisions by parties outside the control of the household. These characteristics are of particular relevance for policy.

*Accessibility* is a fundamental determinant of property prices, and accessibility to employment seems to be a particularly important aspect of this (see the review by Nellthorp *et al.*, 2016, in the Phase 1 study). We now need to consider more deeply what types of accessibility are relevant. The Northern Powerhouse Independent Economic Review (SQW & CE, 2016) found that to transform the Northern economy and start closing the 25% gap in GVA per capita between the North and the Rest of England, it is essential to increase access to economic opportunities including employment and skills training. There is a large body of literature on types of accessibility: we do not attempt to cover this fully in this report (e.g. see Geurs and van Wee, 2004; van Wee, 2016), instead in this section the broad theoretical concept is outlined. This will be expanded in more detail in the section covering empirical modelling strategy (Section 3).

Individuals' preferences for good accessibility can be related to two factors: i) the inconvenience of travelling to other locations *j* of interest, which in turn can be represented by the concept of Generalized Journey Cost, *GJC*; and ii) the opportunities that accessible locations $j$ can provide, e.g. employment opportunities but also education and any other opportunities (e.g. see Ahlfeldt, 2013). We distinguish between employment opportunities (*EO*), skills/study opportunities (*SO*) and other opportunities (OO) that are somewhat accessible from a given property *i*, and define the *Accessibility* $(A_i)$ level of property *i* as follows:

$$A_i = f(GJC_{ijm}, EO_j, SO_j, OO_j) \quad (9)$$

where $GJC_{ijm}$ represents the generalized journey costs between the property *i* and the available set of destinations $j$ by transport modes *m;* $EO_j$ is a metric of the employment opportunities available at the set of destinations $j$ (e.g. number of jobs/workers in each area *j*), and $SO_j$ is a metric of the study opportunities available at the set of destinations $j$ (e.g. student positions in each area *j*). $OO_j$ represents any other opportunities that are relevant for people and is included here only for completeness and to illustrate the link between Accessibility and Place Quality variables (e.g. green spaces, local amenities, etc.), i.e. there can be an element of 'accessibility value' for place quality variables which will be reflected in the modelling (Section 3.4). Under the heading 'Accessibility', however, we will focus on and refer mainly to 'accessibility to employment' and also – in a more limited way – 'accessibility to skills'.

The $GJC_{ijm}$ element of accessibility can be associated with various transport modes (e.g. rail, car, walk, other public transport). The accessibility of a property to jobs and to skills/study opportunities will be a function of the different modes available, where some modes will be substitutes for each other, and others will be complementary to some extent – e.g. walk access to rail. Theoretically, the degree of substitutability or complementarity should influence the preferred model specification. Modes with worse (higher) $GJC_{ij}$ should contribute less than others, although the presence of multiple modes in a corridor increases its resilience and offers residents more choices (and option value), which should increase property values as well. Overall, for modelling purposes, this means that there may be interrelationships between the $GJC_{ij}$ by different modes, and the formulation of $A_i$ is not straightforward.

Defining accessibility as in equation (9) is intuitive. For instance, focusing on accessibility to jobs, it follows that the utility that a property provides to an individual can increase thanks to two distinct factors: i) shorter *GJC* to some employment locations $j$ ; and/or ii) a greater number of jobs at some employment locations *j*. An implication is that policies that increase the pool of job opportunities can increase individuals' utility – and hence property prices – even if travel times do not change. Of course the question of additionality also applies to 'increasing the pool of opportunities', i.e. whether new jobs are additional or displaced from somewhere else will need to be assessed.

The above is a generic definition of accessibility that encompasses a range of specific formulations for modelling purposes (e.g. Ahlfeldt, 2013; Adair et al., 2000). Handy and Niemeier (1997) outline the three key formulations: Cumulative Opportunity, Gravity, and Logsum. Gravity-type specifications are very popular and varied within the literature. For instance, Ahlfeldt (2013) used an exponential function that included specific definitions of $EO_j$ and $GJC_{ij}$ of the fastest mode (he used Generalized Journey Time, $GJT_{ij}$ and the monetary component of travel cost was not included in the calculations) within a gravity-model type formulation. Using only the fastest mode to define accessibility is practical as it avoids correlation issues, across access to (potentially the same pool of) jobs by different modes. It implicitly assumes complete substitutability, however, and assigns no value to modes which are available but are slower. The logsum measure uses a 'composite *GJC*' across multiple modes available to define accessibility (e.g. Ben-Akiva and Lerman, 1985; Simmonds, 2004). This formulation also relies on measuring the relative merits of each mode, but is demanding in data, needing utility-like parameters from mode choice models to underpin calculations. It is also not clear the extent to which mode complementarity and 'option values' of multiple modes is recognised. Finally, others (e.g. Mulley, 2014) use separate accessibility indicators by different modes. This approach does recognise the added value that multiple modes can have, although is limited in its account of substitutability. Correlation issues are common and it is somewhat expected that the data will reveal which mode plays which role in property prices. Other formulations are possible (see the review studies cited above).

In summary, there are pros and cons with all existing approaches to measuring and valuing accessibility, from a theoretical perspective. The choice of approach will be driven by a combination of factors, including these theoretical considerations, data availability, empirical performance and usability of outcomes for policy-relevant purposes.

### 2.4 Place quality characteristics (Q)

Like accessibility, place quality comprises attributes that can be seen as external to the property. There is a myriad of characteristics that can be included in this category, ranging from neighbourhood environmental and design characteristics, amenities within close reach, quality of key services like schools, and crime levels. In fact, from a technical point of view, accessibility to opportunities (as in the previous Section) is also a 'place quality' feature, but we choose to keep it separate as it is the key focus of this study. The following list mentions some of the potential place quality attributes, but the list is not exhaustive:

- School quality;
- Neighbourhood safety (crime levels);
- Peace and quiet (noise levels);
- Air quality;
- Parks / squares / other green or open spaces;
- Shops and other retail facilities;
- Street design (pavements, street trees, etc.);
- Architectural beauty / historic character;
- Industrial estates nearby;
- Landfill sites nearby.

Arguably, all of the above characteristics could be objectively defined in line with Rosen (1974)'s definition: i.e. the *reading* of a characteristic (e.g. school quality, safety, peace and quiet…) is identical for all individuals. Then, of course, valuation can vary.

Theoretically, each characteristic above – and any other belonging to the set 'place quality' $Q_i$ – can have a direct bearing on households' utility and consequently have an implicit or hedonic price which contributes to the overall property price. Also in theory, if there is sufficient variability in the bundles of characteristics (which make up a property) that are for sale, then the hedonic price of each characteristic should be discoverable through data analysis. This is more likely to be the case the larger the study area is, and we believe the Transport for the North area is sufficiently large and diverse to offer a wide range of levels of most characteristics.

However, even if there is sufficient variability in each of the characteristics within the property market under consideration, it is known that certain combinations of characteristics are more common than others – which translates into endogeneity and correlation issues that have been widely acknowledged in the literature (e.g. Gibbons and Machin, 2008; Ahlfeldt, 2013). For instance, areas of environmental or architectural beauty and good accessibility also tend to have good school quality and lower crime rates. This kind of pattern is associated with households self-sorting (e.g. Koster et al., 2016). The observation that certain (non-random) patterns exist in the formation of area characteristics means that clear theoretical reasoning is necessary to understand the underlying process. A conceptual framework can then inform the modelling of hedonic prices, especially when the formation and observation of those prices is complex (e.g. Bayer et al., 2007). Section 2.6 discusses some of these underlying mechanisms that lead to the observation of specific patterns and correlation among certain characteristics.

## 2.5 Physical building & plot characteristics (B)

This category of characteristics refer to the physical attributes of a property. They are generally features that do not vary with external effects (e.g. public sector interventions or neighbourhood sorting). An extensive list of common attributes would include:

- Type of property (e.g. detached, semi-detached, terraced or flat[18]);
- Number of bedrooms;
- Number of bathrooms;
- Floor space (in squared metres);
- Plot size;
- Ceiling height;
- Garden feature;
- Balcony;
- Garage;
- Drive;
- Energy efficiency;
- Property age;
- Property condition;
- Property design / character.

All of these characteristics are potentially observable and measurable (perhaps with the exception of condition, design and character), although data is not always available or associated with the key variable – sale price (*P*). Improvements/increases in these features (with the exception of property age) tend to be associated with higher property prices.

This group of variables is in some ways the most straightforward and also the least interesting one from a transport policy perspective. Some of them are, however, important price explanatory factors and therefore a comprehensive modelling exercise should attempt to account for them to the extent that is feasible.

[18] in the case of flats, also floor/height above ground (e.g. Higgins et al., 2019)

## 2.6 The interplay of property characteristics and other variables

This section considers the interplay of variables within $z$ , which can require special attention. Readers will have noticed that some of the variables that influence property prices are not, strictly speaking, characteristics of the property or the local area, but are characteristics of the households or people in the neighbourhood (the $G$ variables being good examples). Section 2.6.1 examines the role of household income and preferences, which may translate into endogeneity and correlation problems that need careful treatment. Section 2.6.2 addresses the heterogeneity of prices that may result. Section 2.6.3 briefly considers simpler cases of correlation between variables within $z$.

### *2.6.1 Income and sorting of households*

Existing studies tend to encounter the complexity of modelling hedonic prices when they find endogeneity and correlation problems emerging. Sometimes these issues can be overcome through very tightly focused case studies, e.g. the study of one specific feature (e.g. the value of historic amenities or the value of school quality) in one particular city. This allows the collection of data with a high level of granularity and implementation of very specific methodologies (e.g. Ahlfeldt and Holman, 2018; Koster et al., 2016). However, as opposed to other studies, in our project the area of interest is a large polycentric economic region, the North of England. Our research questions revolve around the problem of accessibility in this region, and the goal is to obtain results and reach conclusions that are widely applicable to this context – thus it is more difficult to resort to a detailed analysis in a small self-contained area. In this context, it is even more important to have a clear conceptual understanding of the underlying mechanisms that lead to specific patterns in the available sets of characteristics and thus also play a role in property price determination and hedonic prices.

Two features of the property market that allow some property characteristics to be more commonly observed jointly with others is that property can be privately bought and sold, and that property markets are free markets. Imagine instead the property market being fully under the control of a government body which allocates properties to households through a random allocation process and where households cannot negotiate changes with each other. In that scenario, households would not be able to self-sort into their preferred locations. Then, property and place characteristics would be randomly spread into a much greater mix of bundles than we currently observe. One would then expect no correlation among attributes other than for those that are inherently related (e.g. air quality and green spaces).

In Western economies, housing markets are liberalised and households self-sort based on their income and preferences. Kuminoff et al. (2013) offers an extensive account of the economics of sorting. In particular, the paper develops a theoretical framework for equilibrium sorting in housing markets, based on the same standard microeconomic theory discussed in the previous sections. One of the key conclusions of the framework, consistent with Tiebout's (1956) seminal work, is that "conditional on preferences, wealthier households always choose to live in communities with more public goods (…) and, conditional on income, households with stronger preferences choose communities with more public goods" (Kuminoff et al., 2013, p.14). This implies a positive link between households' income and public goods, which is observed empirically in the stratification of communities (and 'gentrification' as the socio-economic composition of areas changes).

### *2.6.2 Heterogeneity in prices*

Varying income and preferences leads to heterogeneous hedonic prices, i.e. there may not be a single price for each characteristic (Rouwendal, 1998; Koster et al., 2016; Gibbons and Machin, 2008). The theory described by equations (1) to (8) above can be used to explain heterogeneity in hedonic prices. Rearranging equation (7) we have:

$$\frac{\partial p(z)}{\partial z_n} = \frac{\frac{\partial u}{\partial z_n}}{\lambda} \quad \forall\, n = 1 \dots N \tag{10}$$

The first order condition of the consumer behaviour problem indicated that implicit prices $\partial p(z)/\partial z_n$ of each $n$th characteristic are equal to the ratio of the marginal utility of the $n$th characteristic and the marginal utility of income ($\lambda$). Both marginal utilities can vary by household leading to a range of hedonic prices, based on households' characteristics (including income). For instance, it is commonly argued – and empirically proven - that marginal utility of income is lower for richer households, which would translate into higher valuation of a $n$th characteristic ($\partial p(z)/\partial z_n$). Equation (10), as well as the notation in the wider theoretical framework, could be rewritten to make heterogeneity among *k* households explicit as follows:

$$\frac{\partial p_k(z)}{\partial z_n} = \frac{\frac{\partial u_k}{\partial z_n}}{\lambda_k} \quad \forall\, n = 1 \dots N; k = 1 \dots K \tag{11}$$

where now preferences are explicitly different across different households *k* leading to different marginal utilities of the $n$th characteristic and of income. The consequence is a hedonic price that can vary across *k* households. Thus, in estimation, it is an option to consider a distribution of values instead of a unique value. In practice, some hedonic price studies have estimated heterogeneous values for characteristics and found them to increase with income, e.g. Koster et al. (2016) on historic amenities and Rouwendal and van Der Straaten (2008) for open spaces. But income is not the only variable that could explain a distribution of values; other socio-demographic variables, albeit related like education, could also lead to varying preferences and hedonic prices. Similarly, Rouwendal (1998) points out that the hedonic prices of a characteristic may depend on the region of the household. But it may be argued that differences across regions are closely related to sorting and thus to differences in socio-demographics like income or education. Various *spatial regression models* exist, which in principle are capable of representing such spatial patterns. Spatial Error Models (SEM) and Spatial Autoregressive (SAR) models are able to account for patterns of spatial correlation in the error term, whereas Geographically Weighted Regression (GWR) can be used to represent variation in parameter estimates across geographic space (LeSage and Pace, 2009; Fotheringham et al., 2002). These models do however risk giving up the ability to explain the variation using policy-relevant variables (see Section 4.1.6).

Finally, linking this to the previous section, a consequence of acknowledging heterogeneity is that income can have a different role in the model, other than as an explanatory variable in its own right. Income can also be used as a modifier on some of the other hedonic variables, capturing the distribution of values with income levels. Such an approach would help to exploit its explanatory power in a more theoretically sound manner, potentially avoiding some of the conceptual, endogeneity and correlation issues described earlier.

### *2.6.3 Other causes of correlation between characteristics*

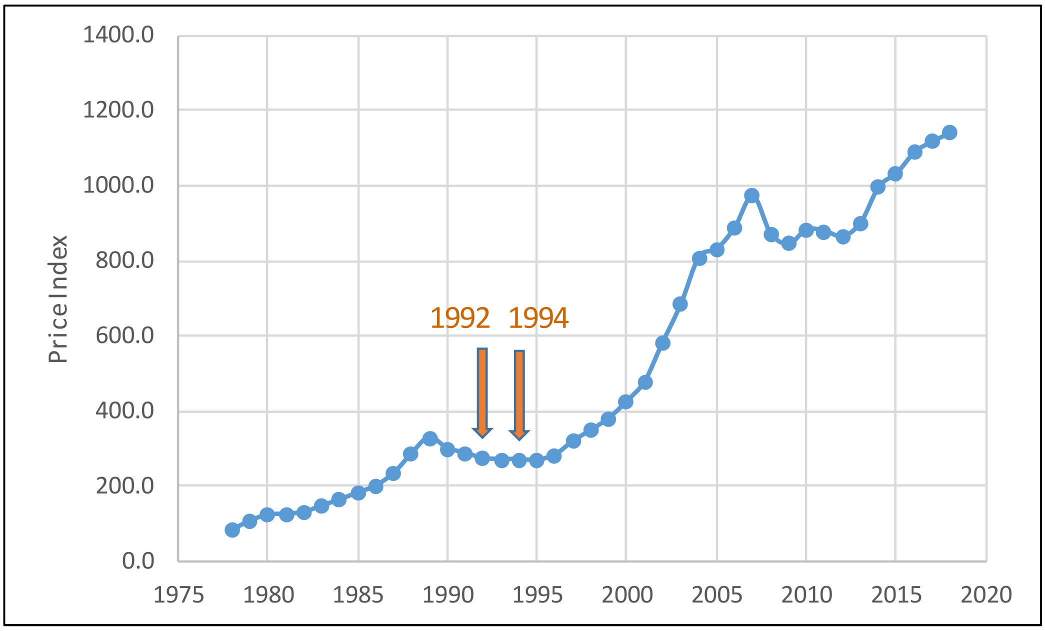
Sometimes correlations between characteristics may not be related to sorting behaviour and the influence of incomes, but instead due to simpler underlying causes. For example, both air quality and noise measures may be related to and even derived from the same traffic flow variable(s). In that case air quality and noise may – depending on the functions involved – be highly correlated with each other and with traffic itself. Statistical problems caused by this include instability of the coefficients in the model, with potentially the wrong sign on one or more of the correlated variables. Potential solutions may be to adopt a composite variable that captures the various elements, or to delete the less significant variables and accept that the remaining variables are representing a wider set of effects.

## 2.7 Supply-demand balance and macroeconomic factors (S)

In this section we discuss the arguments for a variable representing the *supply-demand balance* in the property market as an explanatory variable for house prices, and the potential indicators that might be used.

Property markets are observed to be cyclical. In the Phase 1 study, it was noted that the property market impact of rail openings can be reduced or suppressed entirely when the service opens during a property market recession – two examples being Manchester Metrolink (1992) and Sheffield Supertram (1994). Conversely, the market has cyclical peaks on a national level, such as the pre-crisis peak in 2007 and the 'Lawson boom' peak in 1989 (see Figure 2.1). These aspects could be described as macroeconomic, influenced by macroeconomic policy (e.g. interest rates, fiscal policy) and by market conditions and the business cycle. They may also be influenced by net migration and foreign investment into the UK economy or housing market. Whilst an advanced model of housing market dynamics is not in the scope of this Phase of research, this suggests it is important that the models we build are based on a long-term view of property values, and that the data used to estimate cross-sectional models is from a year in which the national market is not experiencing a marked cyclical low (*or* an exceptional cyclical high).

Figure 2.1: UK House Price Index 1978-2018 (1978=100)




Source: Nationwide House Price Index (Nationwide, 2018)

Property markets are also regionally differentiated, with some regional or sub-regional markets being identifiably 'hot' whilst others are 'cool'. In order to build a model for the TfN area, there was a desire to represent – in a way which could be quantified and captured in the model – this regional differentiation. This led to consideration of ways in which potential demand-supply disequilibrium at a regional level could be analysed, initially from a theoretical point of view.

An underlying principle in the study of the housing market is that prices are determined by the interaction between the demand and supply of properties (or equivalently 'bundles of characteristics' *z* in the hedonic model). The observed prices can be assumed to be the outcome of a market equilibrium, as discussed at the start of the theory section. At the equilibrium, the demand for housing ($D_h$) can be assumed to be a function of income *M* and the price level *p(z)*:

$$D_h = f(M, p(z)) \tag{12}$$

On the supply side of the market, the production of housing units $S_h$ is determined by the costs of provision *C* (land, material, labour, finance) and by the current real house prices *p(z)* (Kenny, 1999; Tsai, 2018).

$$S_h = f(C, p(z)) \tag{13}$$

As Tse et al. (1999, p.629) describe it, "If the number of households demanding housing units exceeds the number of units available at some moment, competition among the demanders will drive up prices until equilibrium is attained. The change in house prices will stabilize when demand equals supply. However, such adjustments do not occur instantaneously". It is the difficulty the market has in adjusting quickly (from the supply side) that can generate imbalances in the short term that will be reflected in the observed prices. The question is then: to what extent is a disequilibrium related to the characteristics of properties (e.g. A, Q and B)? So far, the theoretical framework that underpins hedonic price modelling (e.g. Rosen, 1974) explicitly assumes that markets are at equilibrium.

Let us consider three ways in which the ratio of demand to supply could be increased and prices be driven up:

a. Improvements in the demand factors, namely A, Q or B above, would lead to a higher demand and hence a higher price of housing, all other things equal. This is the foundation of the hedonic price model.
b. Restrictions (or reductions) in the supply of housing in the area.
c. Increases in demand due to a net increase in the population interested in buying (e.g. new generations of individuals joining the market, net migration, foreign investment).

While the impact 'a)' should be captured in an HP model through the demand variables – for example an improvement in rail accessibility – impacts 'b)' and 'c)' are somewhat different effects on the supply-demand balance and their addition to an HP model separately might be considered. However, there may be reservations about this: even for option c) it can be argued that if the additional population prefers 'area 1' to 'area 2', leading to an excess of demand in area 1, that would not be random – i.e. a pure population effect – and would still be the result of the relative characteristics of the two areas, i.e. a demand effect that should be picked up through the hedonic prices of the area characteristics. Only a universal demand increase across all regions would be entirely in the category 'c)'.

In conclusion, the demand-supply ratio is an interesting variable that surely impacts upon property prices, but it requires careful consideration whether to include it as an additional factor in an HP model. Theoretically, particularly in a cross-sectional model, we are inclined to let any demand-driven effects be an inherent part of the estimated hedonic prices. However, it is worth exploring the extent to which supply constraints can influence property prices. For this study, we will collect data to generate a supply constraint indicator and will test its inclusion in our preferred model. Data on demand-supply ratio indicators will also be explored, but for the theoretical reasons outlined above will not be tested further. The following sections describe potential indicators that may be used.

### *2.7.1 Demand-Supply indicators*

There are a number of alternatives to define a Demand-Supply (*DS*) ratio for use in an HP model. Coppola et al. (2013) define *DS* as a the ratio between residential demand and supply, where:

$D_i$ = number of residents in zone $i$

$O_i$ = number of habitable square meters in zone $i$

This is the only HP study identified in our literature review that includes a *DS* ratio. This formulation picks up overcrowding, since areas with more people per $m^2$ would be assigned a higher *DS* ratio, and vice versa.

Other indicators of market conditions either control for overcrowding or use other market information that avoids the issue. For instance, Demand can be defined using transaction information and Supply using availability of houses for sale (e.g. Tsai, 2018; Tse et al., 1999). One way of defining *D* and *S* in this fashion could be as follows:

$D_i$ = number of transactions of housing units in zone $i$ and period $t$

$S_i$ = number of housing units for sale in zone $i$ and period $t$

Alternatively, if data on housing units for sale was difficult to obtain, there may be variables that serve as proxy. For example, the average time on market for a property can be seen as a proxy for the ratio *transactions/units for sale* (inversely related to the DS ratio, with properties in low demand (relatively high supply) remaining longer on the market).

### *2.7.2 Supply Constraint indicators*

Using supply constraint indicators means that demand-related information is left completely to the existing elements of the Hedonic Price model. Thus, a supply constraint measure will aim at picking up any property price premium related to a shortage of supply. Supply constraints can be measured in different ways. For instance, one could use information on planning restrictions on each area; however this information is rarely available and is difficult to generalise into a common metric. On the other hand, proxies for supply constraint can also be derived from information on the dwelling stock in each area. Information on dwelling stock is more easily available and can be shaped in various ways to generate a proxy. In particular, the change in dwelling stock (from a previous time period) can reflect the extent to which supply is constrained more in some areas than others. Thus, a generic metric for Supply Constraint (SC) is a follows:

$SC_i$ = % change in dwelling stock in period $t$

where the change in period t can be related to the previous year or another earlier year (e.g. 3 or 5 years ago). Since supply constraints may have a more or less immediate impact, but also a medium to long-term impact, it is possible to test changes relative to various lags. This is possible also within the cross-sectional model: the purpose of this variable is to act as a proxy of the supply constraint.

### 2.8 Additional Theory – Commercial Model

The value of commercial property can be understood within a microeconomic theoretical framework analogous to the one used for the residential property market. Thus, the 'hedonic' assumption means that a property is valued for its utility-bearing characteristics, and hedonic prices are then defined as the implicit prices of those characteristics. There are of course differences in the commercial property market:

i. The buyers (renters) are businesses rather than households.
ii. The utility-bearing characteristics of the property are different as businesses seek to maximise not their 'residential-related utility' but their 'commercial-related utility', driven by (we assume) profit maximisation.

Just as it happens in the residential market, observed property prices and amount of characteristics $(z_1, z_2, \ldots, z_n)$ act as channels through which implicit or hedonic prices for $z_1, z_2, \ldots, z_n$ are revealed to economic agents. Recalling again the theory, the (set of) implicit prices guides both consumer and producer choices in the characteristics space, i.e. agents buy and sell packages of characteristics (Rosen, 1974). The price of a property can then be defined just like in equation (1), as the price of a bundle or package of characteristics *z,* giving the hedonic prices $p(z)$:

$$p(z) = p(z_1, z_2, \ldots, z_n) \qquad (14)$$

In the residential market we decomposed the generic set of characteristics *z* associated with a property into a combination of certain subsets of characteristics that share common features. In the commercial market, the same approach is taken but the contents of the subsets vary as follows. There are three types of Commercial property, each with distinct Accessibility requirements (Table 2.1).

Table 2.1: Commercial property types and forms of accessibility

| *Commercial Property Type* | *Accessibility Requirements* |
|---|---|
| Office | Staff (commuting)<br>Business-to-Business (personal travel)<br>Deliveries (goods) and Servicing (staff) |
| Industrial | Staff (commuting)<br>Goods in (heavy/light)<br>Goods out (heavy/light)<br>Business-to-Business (personal travel)<br>Servicing (staff) |
| Retail | Staff (commuting)<br>Customers<br>Deliveries (goods) and Servicing (staff) |

- Accessibility (A)

  'Accessibility' needs are somewhat different in the commercial property sector from the residential sector. Access requirements vary across property types as above:

  - Accessibility to the labour market (i.e. to residential properties) is common to all. This is the mirror image of the 'accessibility to jobs' in the residential model. Commercial properties with better access to the labour market will be more valuable, all other things equal. The accessibility function will be summed across all *i* instead of all *j*. Deterrence functions estimated for the residential market may be applied here (trip purpose is commuting). This aspect of accessibility can be defined as 'Accessibility to Labour Market' for commercial property 'j', as follows:

    $$A_{LMj} = f(GJC_{jim}, LM_i) \tag{15}$$

    Where $GJC_{jim}$ is the Generalized Journey Cost between commercial property at location 'j' and households (labour market) in locations 'i' by mode 'm'; $LM_i$ is a measure of the labour market at location 'i', which would be measured ideally using active population (alternatively, if disaggregated data is not available, population or number of households could be used).

  - Accessibility to other businesses & workers – particularly for Office property. Office properties near other businesses are expected to have a price premium due to all the benefits of clustering (agglomeration effects). The externalities of being located near other business (e.g. due to knowledge transfers, spill overs, etc) are valuable. The impact of travel time here is different to the impact in the 'access to labour', as it is not commuting time that matters: thus, deterrence functions are likely to differ too. It is not certain what the deterrence function should look like; walk is potentially the most important mode given the rather steep decay functions found in the literature (Graham et al., 2010). This aspect of accessibility can be defined as 'Accessibility to Other Firms' for commercial property 'j', as follows:

    $$A_{OFj} = f(GJC_{jjm}, OF_j) \tag{16}$$

    Where $GJC_{jjm}$ is the Generalized Journey Cost between commercial property at location 'j' and other firms in locations 'j' by mode 'm'; $OF_i$ is a measure of the amount of other firms' activity at location 'j', which could be approximated using number of jobs but also other measures of the economic activity in the area 'j'.

  - Accessibility to customers (retail property). Market demand in the area: a big driver of prices in the commercial sector is the ability of the location to attract customers to the premises, which may be a function of the density of retail floorspace and variety/quality in the local area. For instance, high streets or shopping areas/centres with many people walking every day would have a higher premium. Since customers will very often travel from home or work, it may be difficult to separate this effect from the accessibility to other firms ($A_{OFj}$) and households ($A_{LMj}$) and in practice they are likely to be highly correlated.

- Place quality characteristics (Q)

  Place quality comprises attributes that can be seen as external to the property, including environmental and design characteristics of the urban area and amenities within close reach. Just as with the residential property market, from a technical point of view, accessibility to opportunities could be regarded a 'place quality' feature, but again we choose to keep it separate as it is the key focus of this study. The following list mentions some of the potential place quality attributes that are applicable to the commercial property market but different from those of the residential market:

- Urban realm: in simple terms, a more pleasant area where the premises are located. Urban realm can be made up of multiple attributes (e.g. street trees, green spaces, open spaces, street design and building character). Areas with better quality urban realm are more likely to attract workers to a firm and more customers. While urban realm was also, in general terms, a feature of the residential property market, the details on what is valuable might be different.
- Many of the factors identified for the residential market would not apply (e.g. school quality, playgrounds) or would do so in a different way (e.g. road noise, air quality).

- Physical building and plot characteristics (B)
  - Size ($m^2$);
  - Quality of the premises;
  - Use classification of the building: generally in one of the three categories above, i.e. offices, industrial premises or retail. This represents three 'separate' sub-markets with Commercial property;
- Socio-economic characteristics of the area (G) – the relevance of this may come through retail spend, or the types of labour available in the walking catchment, or indirectly as a proxy for place quality (this is in fact how it is interpreted below).
- Supply constraints (S) – in theory, analogous to their role in the Residential market.

Overall, the utility function is defined for commercial premises as (X is again the non-property composite good in the business's utility):

$$U = U(z, X) = U(A, Q, B, G, S, X) \tag{17}$$

# 3. Modelling Strategy and Data

Having developed the theoretical background in Section 2, this Section describes the empirical research strategy that has been used to test the theory. Inter-related strands of modelling have been undertaken using different model types, to explore different aspects of the subject.

## 3.1 Overall strategy

The primary objectives of the empirical work are:

1. to test whether rail accessibility has a significant effect on property prices in the Transport for the North area;
2. if it is significant, then to quantify the relationship between rail accessibility and property prices – both for residential property prices and, if possible, commercial property prices (data permitting);
3. also to quantify the relationships with other variables which might be significant, giving an understanding of the influence of rail accessibility in the context of other factors.

### *3.1.1 Model types*

The chosen model types for this work are:

- first and foremost, a *hedonic pricing (HP)* model based on the theory set out in Sections 2.1-2.7, providing a cross-sectional analysis of the residential property market;
- second, another HP model, focusing on the commercial property market (using the theoretical starting point in Section 2.8);
- thirdly, using *quasi-experimental methods* in a selected case study area, where there has been a substantial change in rail accessibility over time since 1995[19] – for example, the area covered by the Metrolink system;
- finally, as sensitivity tests to the main HP model, spatial regression models, which could take the form of:
    - a Spatial Error Model (SEM) or Spatial Autoregressive (SAR) model, capable of accounting for spatial correlation; or
    - a Geographically Weighted Regression (GWR) model, which has the potential to estimate spatially varying parameters.

  These have the potential to improve model fit to the data, although not (as it turns out) to improve upon the policy relevance of the HP model.

The adoption of *HP* as the main modelling approach reflects the conclusions on modelling options in the Phase 1 study. The main challenges found with some previous HP models are:

i. that a poorly-specified HP model may contain many highly correlated variables, leading to unstable coefficients and potential biases in the values emerging – our work over the last 2 years has been focused particularly on improving property market HP models from this perspective, and the models emerging from this study show an encouraging degree of robustness to changes in specification;

---

[19] 1995 is the start year for the Land Registry price data.

ii. that (like any model) they rely on good data, and a particular issue in previous studies, including published studies, has been the measurement of accessibility. In some cases 'distance to the CBD' or 'journey time to the CBD' has been used, which is problematic in a multi-modal transport system, or where quality is not uniform. Our approach is therefore to use high quality data on accessibility, via the main competing modes. The assumption of a single CBD is also unrealistic in a polycentric region like the North, where the purpose of the investment is partly to open up the economies of the adjoining city regions, encouraging cross-commuting and business-to-business interaction. Our approach is to measure accessibility to jobs and economic opportunities across the study area, not limited to the nearest CBD.

*Spatial regression models* are distinct from OLS in their treatment of spatial correlation and some allow relationships to vary across localities. Spatial correlation is a feature of property prices across LSOAs and MSOAs[20]. Studies including Mulley et al (2018), Dziauddin et al (2015) and Northall (2014) use spatial regression models, and such models may offer an improved model fit compared with more conventional HP model forms, however policy interpretation may be more difficult. This study explored SEM, SAR and GWR models, noting issues where they arose. See Section 4.1.6.

*Quasi experimental models*. Finally, these are relevant because they offer a complementary approach using panel data (time series and cross-sectional). By focusing on treatment and control areas, and investigating property prices before and after a transport intervention using – for example – a differences-in-differences method, these models allow us to focus on price changes in a particular case study area, and are useful in that they provide a way of addressing causality (see Section 3.9 for further description). The challenge is partly in identifying comparable treatment and control areas and partly in generalising the results from one specific station or corridor to other contexts. Any downward impact on land values at the 'control' sites (due to reduced *relative* accessibility) may be missed if the only comparison is between the 'treatment' and the 'control' sites. Also, compared with the above HP and spatial modelling approaches, these models leave a great deal of area-wide data 'on the table'. On balance, these models are best used alongside the cross-sectional models to gain a more complete picture of 'impact' and to address questions of causality[21].

We have also looked at fuzzy logic, random forest algorithm and simulation techniques. These lack a clear economic interpretation at present, although that could potentially be addressed in the longer term. We consider that these are beyond scope for now, but have potential for future research.

Our approach to model types is a fairly flexible one: we are open-minded about the exact form of the final model. We will be guided by: model performance; fit with theory; and ability to provide the types of output needed to inform the Business Case. These are our main criteria.

---

[20] Lower Layer Super Output Areas and Middle Layer Super Output Areas

[21] We also considered instrumental variable approaches. Whilst regression (or propensity matching) may adjust for observed differences between treated and untreated groups, there may be unobserved characteristics which determine treatment, i.e. differences in accessibility between otherwise similar areas. Addressing this concern could either involve the use of an instrument which can explain accessibility but is exogenous to land value or some discontinuity in treatment. Regarding the former, in the literature instruments have included long lags of population or historic transport plans (eg Ciccone and Hall, 1996) to explain current day density (or in our case accessibility), with the logic that these are exogenous to current day economic factors that might drive accessibility. Discontinuities in treatment include some random element for treatment selection such as natural disaster closing a rail line (Tyndall, 2017) or intercity schemes leading to better intra-city connections (Gibbons et al., 2012).

### *3.1.2 Model structure*

*Model structure* is key. The updated approach to HP models, building on lessons learned from previous research, relies on careful selection of model variables, driven more strongly by economic theory.

- *Income* is often included in HP models as a 'socio-economic' variable alongside others, such as unemployment rate; proportion of professionals; % with a degree-level qualification; etc (e.g. see Powe et al., 1995). We have found that these socio-economic variables are highly correlated with each other, and also correlated with some neighbourhood variables such as school quality and crime/safety. It is therefore unhelpful to include a large set of socio-economic variables in the model (to do so causes bias and instability in the coefficient estimates, particularly the 'G' and 'Q' variables). Our starting point for this study is to use the minimum number of socio-economic variables. We begin with just *neighbourhood average household income* (at LSOA level) in this role, and we note that Cortright (2009) successfully used a very similar variable – "median household income for the census block group[22] in which each house was located" – in a relevant US-based property market HP study.

  In economic theory, income affects the budget constraint (the individual or household's 'ability to pay') and therefore potentially influences all the hedonic prices emerging from the model (see Section 2.6.2). We allowed for this by testing income interaction terms throughout the hedonic function. The approach has parallels with the approach to income in the recent UK value of time study (Batley et al., 2017) and more broadly with state-of-the-art theory and practice in the field of valuation of non-market goods.

*Accessibility* is a fundamental determinant of property prices – this is clear from the review work undertaken in Phase 1 (Nellthorp et al., 2016). In particular, accessibility to employment drives property prices. The Northern Powerhouse Independent Economic Review (SQW & CE, 2016) found that to transform the Northern economy and start closing the 25% gap in GVA per capita between the North and the Rest of England, it is necessary to increase access to economic opportunities including employment and skills training.

*Place quality* is often broken down into neighbourhood/amenity/environment categories in HP models. We aim to bring these together into a single set of carefully selected variables which describe – as best we can with the available data – the attributes of the local area which the evidence indicates are valued in the property market. This is likely to include: schools; local facilities; greenspace; urban realm; and noise & air pollution – among other potential variables.

We have tested whether these variables play a role in the form of, for example:

- an indicator of quality, e.g. the level of noise exposure or air quality;
- a simple dummy variable based on an understanding of catchment areas;

[22] similar to an LSOA in England: US Census Block Group = 600 to 3,000 people; LSOA = 1,000 to 3,000 people. Cortright interpreted the variable as "a proxy for perceived differences in neighborhood quality and to reflect the external effects associated with the income level of one's neighbors. Neighborhood income levels are frequently associated with crime rates and school quality ... (High-income neighborhoods tend to have better local schools, neighborhoods with lower incomes tend to have higher crime rates)". We have found that with good enough data it is possible to separate some these effects, and they are found in the 'Q' or Place Quality category, nevertheless income remains a powerful socio-economic variable in the model.

– a distance weighted function so that the benefit of an amenity is greater if it is closer to the home location (e.g. WalkScore – explained in the article by Cortright (2009)).

We aimed to take account of place quality at stations, since place-making at major stations is a feature of NPR. For properties located near the station, this would feed into the model through the place quality variables; for journeys passing through the station as an interchange (between trains or modes) this could be incorporated in the journey experience element of generalised cost & accessibility. Ultimately the evidence for this was not sufficient: the types of improvements proposed in NPR go beyond what is covered by our dataset, and an alternative research approach would be more fruitful for this question.

*Model Variables*

Overall, the variables in the model include the following categories:

- P, Price variables (see Section 3.2);
- A, Accessibility variables (Section 3.3);
- Q, Place quality variables: neighbourhood/amenity/environment variables (Section 3.4);
- B, Building and plot characteristics (Section 3.5);
- G, Socio-economic variables, e.g. income (Section 3.6);
- S, Supply-demand balance variables (Section 3.7).

### *3.1.3 Modelling process*

The modelling process included:

- an initial model containing a basic set of variables across the categories above, pragmatically using the data we had gathered at the spatial levels initially available;
- testing alternative functional forms – in the HP model, we expected the semi-log form to perform best, based on both our own previous experience and the findings in the wider literature;
- variable addition and deletion with a view to developing a model covering all the variables that we wish to test for a significant relationship with property prices;
- adding data where necessary, for example to increase the amount of spatial detail;
- monitoring the goodness-of-fit (adjusted $R^2$), t-ratios and signs of coefficients, the correlation matrix, and other statistics – aiming to develop a model which meets our main criteria above (end of Section 3.1.1);
- the same broad process has been followed for each of the model types being tested. Table 3.1 shows the sequence of models produced. The results are presented in the next section of the report (Section 4).

Table 3.1: Modelling sequence

| *Date* | *Cross-Sectional Residential* | *Time Series Residential* | *Cross-Sectional Commercial* |
|---|---|---|---|
| 24 July | Model 1 | | |
| 17 Sept | Model 6 | | |
| 31 Oct | SOBC Model | Initial Model | |
| 30 Nov | Post-SOBC Model | ↓ | Initial Model |
| 21 Dec | Model 14 / 14a | Improved Model | Models 2-5 |

Through this process, we have explored the potential not only to optimise the set of variables included, their specifications, and the overall functional form, but also to optimise the spatial detail in the data and gradually to converge on a set of models which shed light on the questions about transport and land value posed at the start (Section 3.1).

### *3.1.4 Spatial scale*

In this report, reference will be made to the spatial scale or level of spatial detail in various datasets. This is relevant because some characteristics are very localised – e.g. noise, as shown in the lower-right panel of Figure 3.1 – whilst some data may be available only at a more spatially averaged level, with less spatial detail. For example, many of the variables are only available for Middle Layer Super Output Areas (MSOA) or Lower Layer Super Output Areas (LSOA) (see Figure 3.1). Obviously, there is a risk that property characteristics which can vary a lot on a detailed spatial level are represented by more averaged, and therefore less accurate, datasets. How much this matters for the estimation of a hedonic model in the property market is an empirical question, on which we have had to focus a considerable amount of effort.

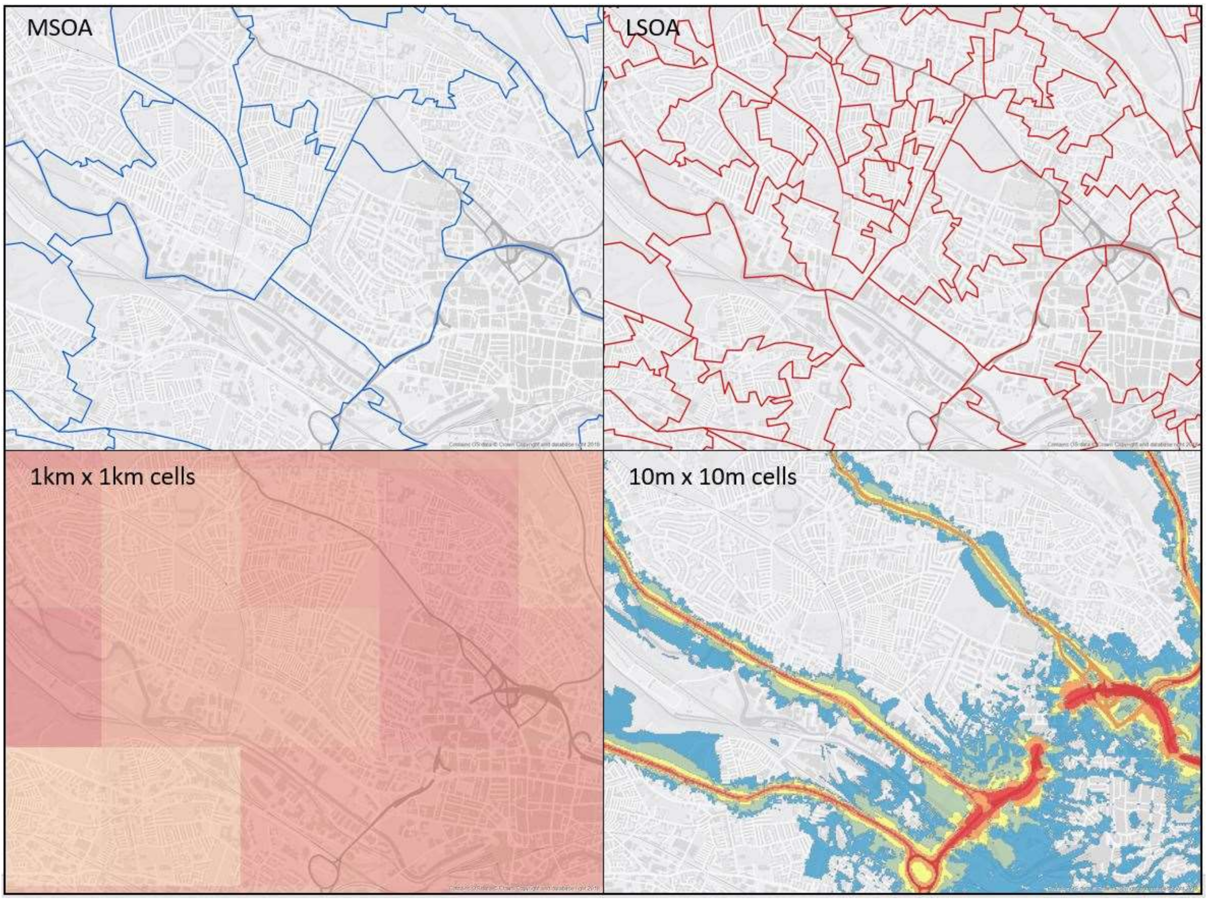
The level of spatial detail also, obviously, has implications for the size of datasets and the computing power required to estimate models. Table 3.2 summarises the number of spatial units in the TfN area.

Table 3.2: Number of units at different spatial scales in the TfN Area

| *Spatial Unit* | *Number in TfN Area* |
|---|---|
| MSOA (or Zones, as used in NoRMS model) | 2,016 |
| LSOA | 9,764 |
| OA | 50,984 |
| Postcode | 624,880 |
| Street Address | 6,900,000 (approximately)* |

Note: *based on number of dwellings, MHCLG Table 100 (https://www.gov.uk/government/statistical-data-sets/live-tables-on-dwelling-stock-including-vacants)

Figure 3.1: Local geographies and spatial scales – MSOA; LSOA; 1km square cells (Defra PCM model of air quality); and 10m square cells (END Noise data)




### 3.2 Property price

The dependent variable is property price. Data on property prices was gathered as follows:

- Residential property: Land Registry Price Paid data provides a comprehensive dataset of transaction prices, for the period 1995-2018, by street address, which allows for detailed matching to other data;
- Commercial property: The data is more restrictive – commercially available data is primarily for asking rents (rents paid are not always available) but includes quality ratings (which have a substantial impact on rents). Meanwhile official VOA data are harder to obtain (data have been requested) and lack a quality rating, but are for rents paid. As a result, the prototype commercial model is based on a relatively small sample size.

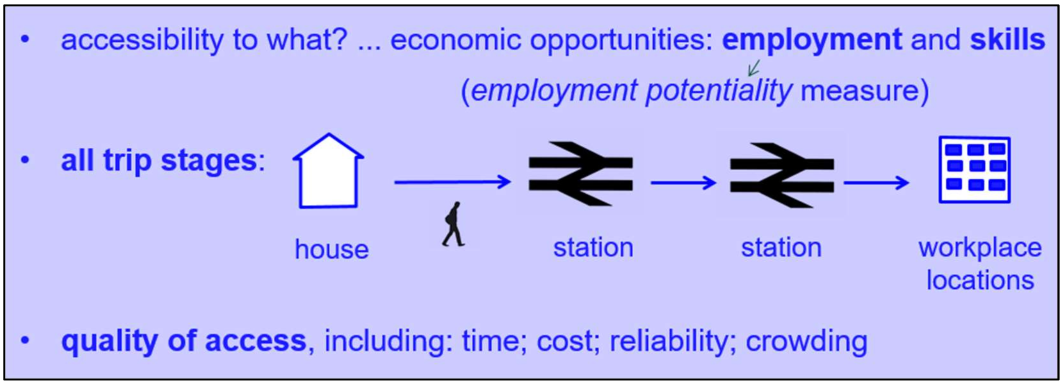
After cleaning the data, the numbers of observations used in each model are: in the cross-sectional residential model, 160,563 individual properties (all for the year 2016); in the time-series residential model, 39,264 LSOA average prices (1995-2018); while the prototype commercial model uses just 294 observations and its results should therefore be treated as indicative pending expansion of the sample.

### 3.3 Accessibility

As described in the theory section (Section 2), our starting hypothesis is that accessibility to opportunities adds to the value of a property, in particular accessibility to jobs and skills (i.e. study and training opportunities). Access to other opportunities is treated as part of 'Place quality' in Section 3.4. In the model, we have therefore focused on:

- representing accessibility to employment using policy-relevant variables that take account of:
    - employment locations across the North – we have used an employment potentiality measure, further developing the measure used by Ahlfeldt (2013), to provide a composite measure of accessibility from residential locations to employment locations across the study area;
    - the main components of accessibility from a residential location to employment opportunities, including travel time, cost and other generalised cost elements (crowding and reliability) (see Box 1):
        - by main mode (rail; car; other PT (bus); walk);
        - taking note of access/egress needs – e.g. note that access from the home location to the rail station is potentially an important differentiator, and will be modelled specifically;
        - for properties located in city centres and near to other employment centres, walk is a viable main mode which may influence property prices (there was some evidence of this in previous prototype models, e.g. Xue, 2016);
- exploring whether access to skills training locations including Higher and Further Education and other training centres influences property value.

Box 1: Scope of 'accessibility'




Note that in practice, there have been barriers to measuring money cost for these journeys – specifically for rail – and as a result the models rely on generalised journey time (GJT) comprising the other components of 'quality of access' in Box 1. A distance-based estimate of rail fares was considered, however this would add little to the usefulness of the results. One future enhancement of the model could be to develop money cost estimates for each mode. For now, the parameter estimates for rail, car and walk in the models reported from this study, take account of the costs of these modes implicitly rather than explicitly.

### *3.3.1 The mode-based model*

A gravity model-based accessibility indicator has been calculated for three key modes: Rail; Car; and Walk. (Further tests relating to the bus mode are described in Section 4.1.7). In each case, the starting point for calculating accessibility is the general formula given in section 2.3:

$$A_i = f(GJC_{ijm}, EO_j, SO_j) \tag{18}$$

where the accessibility at location *i* ($A_i$) is a function of generalised journey cost (or time) between *i* and locations *j* by mode *m,* as well as the employment and skills opportunities at locations *j.*

In each case a separate dataset is used for GJC (or in fact GJT). Initially these datasets were at MSOA level – work to increase the level of spatial detail was undertaken over the course of August-November 2018, such that for Rail the final level of detail was OA at the property *i,* to LSOA for destinations *j* :

- Rail – zone-to-zone [23] journey time (JT) outputs from TfN's NoRMS model, covering: access/egress time; initial wait time; in-vehicle time; interchange time; delay time; and crowding penalty. Weights have been applied to elements of the JT in accordance with the latest version of the rail industry Passenger Demand Forecasting Handbook (RDG, 2018). Access and egress times have been recalculated so that they relate as closely as possible to the location of a particular property (instead of relying on the MSOA centroids).
- Car – peak travel times, taking observed traffic speeds into account – these speeds are based on TrafficMaster data – at MSOA to MSOA level which should be adequate for Car;
- Walk – although initially at MSOA level, the additional work has been completed to produce LSOA-to-LSOA walk times, assuming an average speed of 4.8km/h. Testing showed that the new Walk Access (LSOA) variable hardly changes the model outcomes in general compared with the earlier model (Model 6), however the Rail Access variable has now a significantly higher coefficient. This may be related to a reduction in correlation between the Walk and Rail variables, which is observed when the spatial scale is changed. Model fit remains practically the same.

Employment opportunities in each LSOA, $E_j$ , have been calculated, using the HSL National Population Database to provide the quantities and locations of employment (the HSL dataset is at LSOA level so it was aggregated for the initial models at MSOA level). The study area for the purpose of calculating employment opportunities is defined as the TfN area plus a 50km buffer outside the TfN boundary, where residents of the TfN area are assumed to be able to consider employment opportunities.

A number of approaches exist for assessing the access to employment potential from each location (for an overview see Handy and Niemeier, 1997). In a polycentric area such as the North of England, a gravity-based model will allow a resident at origin *i* to potentially access the employment in any destination zone, subject to their travel budget constraint. Broadly speaking, employment that is easily accessible to an origin will be more highly weighted than employment that is less accessible. This calculation over all destinations for each origin is

[23] zones in the NoRMS model do not correspond exactly to MSOAs but in most cases are very similar to MSOAs.

called the *employment potentiality*, *EP*. In our case, the function is defined to be general across modes:

$$EP_{im} = \sum_j E_j e^{\alpha_m GJT_{ijm}} \tag{19}$$

where:

$EP_{im}$ is the employment potentiality for origin *i* and mode *m* (origins are within the TfN area);

$E_j$ is the number of jobs observed at destination *j* ;

$\alpha_m$ is a mode-specific parameter, referred to as the *deterrence parameter*, which defines how employment is discounted as a function of GJT between *i* and *j* ;

$GJT_{ijm}$ is the generalised journey time between origin and destination using mode *m.*

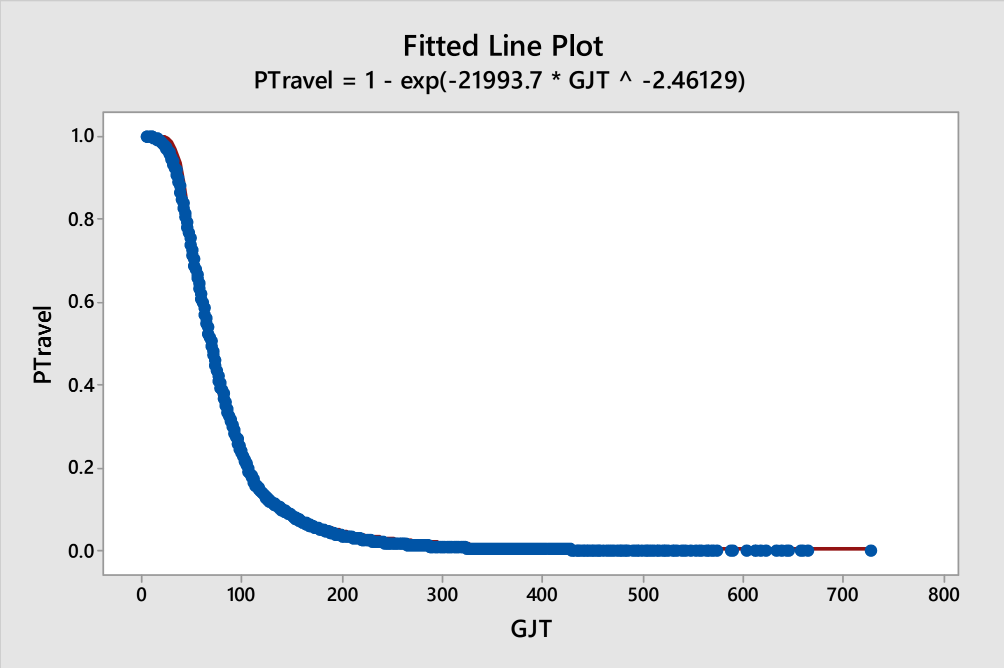
Previous research has estimated deterrence parameters for a single mode as part of the regression estimation (e.g. Ahlfeldt, 2013). However in this study, the deterrence parameter is estimated separately for each mode based upon zone-to-zone GJT combined with observed travel demand for that mode from the 2011 census (Table WU03EW). This results in a 'willingness to travel' function for each mode, which can be used in the employment potentiality function for that mode. The shape of the curve in Figure 3.2 below is similar to the negative exponential (as is commonly used), but has been calibrated to a two parameter function, so that equation 19 becomes:

$$EP_{im} = \sum_j \left( E_j \left( 1 - e^{\alpha_{1m} GJT_{ijm}{}^{\alpha_{2m}}} \right) \right) \tag{20}$$

Note the addition of the mode specific parameter $\alpha_{2m}$.

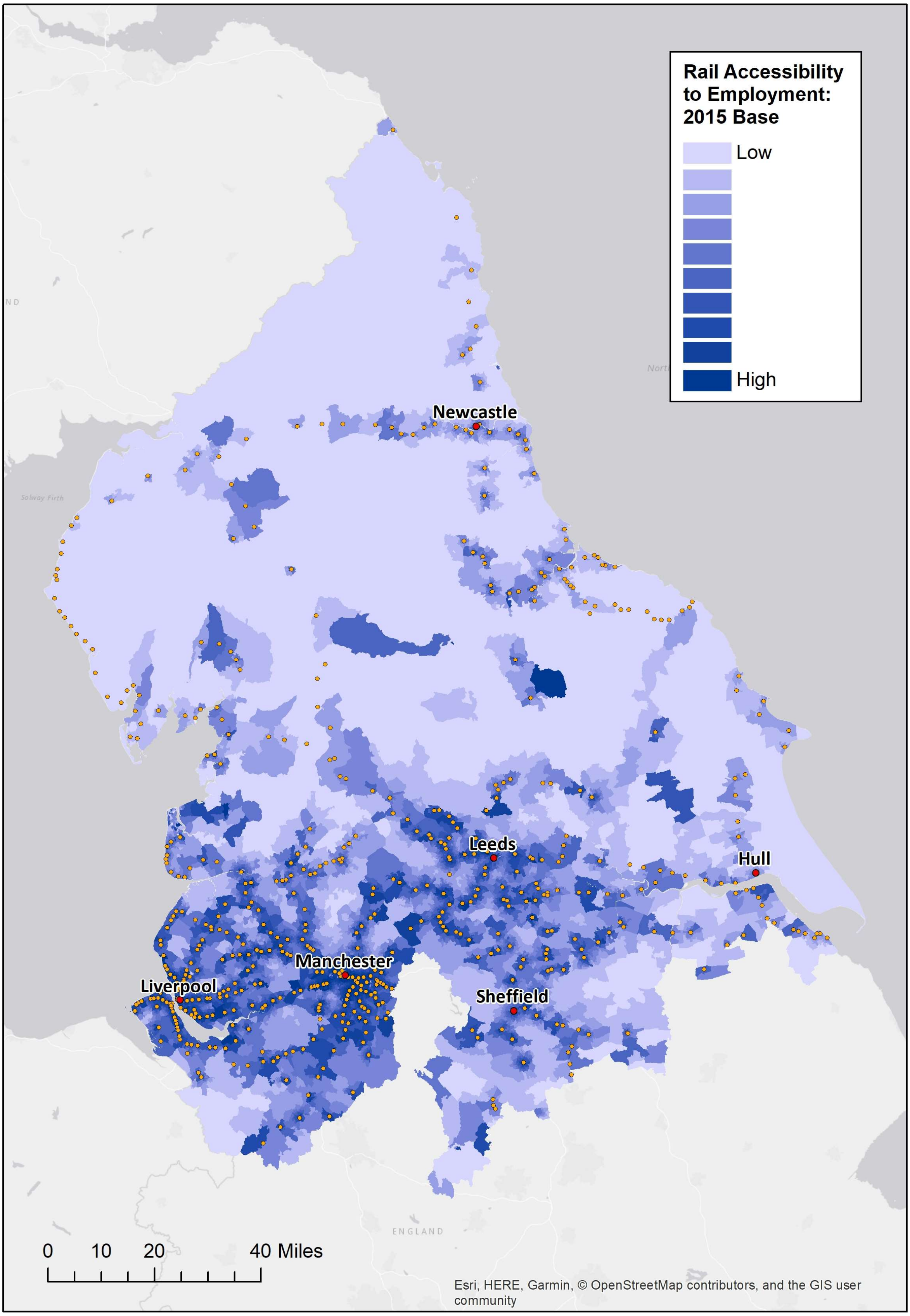
An advantage of this approach is that it is calibrated on observed data and therefore offers greater accuracy, compared to using parameter values from the literature. The result of applying this function to all origin zones, together with NoRMS-based GJT and HSL employment data is shown in Figure 3.3 (which is drawn at MSOA level).

Figure 3.2: Deterrence function for Rail accessibility to employment based on Census data




Turning to Skills, an attempt was made to estimate a model where Access to Higher Education (HE) opportunities was included alongside Access to Employment, however this was unsuccessful. Strong correlation between the accessibility measures for access to HE and access to employment made it impossible to include both in the model with correct signs and satisfactory t-ratios (the underlying correlation between the GJT and deterrence functions is perfect of course). Instead, it was found to be possible to estimate models which included the number of Students resident in the same LSOA as the property *i* (counted as term time resident students), which we interpret as a rough proxy for accessibility to study opportunities, the number of students being related to the number of student places available in a location.

Figure 3.3: Rail accessibility to employment in the TfN area, based on NoRMS, HSL and Census data (2015 Base scenario)



## 3.4 Place quality

This is not the main focus of the model. Its inclusion is partly (i) to ensure that key variables which certainly vary across the study area and may have changed over time, and may therefore help to explain the pattern of property prices, are addressed where appropriate, and partly (ii) to help enable any place quality improvements as part of the NPR project – e.g. at/around stations – to be evaluated, dependent on the model outcomes (which variables are significant and whether they correspond with the variables which would be influenced by the NPR on/off-station place quality work).

Through this Study, including the modelling results, the Advisory Panel discussions and follow-up meetings, it has become clear that spatial scale is key. Across the main place quality aspects we have been focusing-on – including air quality, noise, greenspace, crime level and school quality, etc. – there are substantial street-by-street variations. Streets may be as little as 20-30m apart, whilst the data used in previous HP models is sometimes at 350m Hex cell level or even 1km squares. We have made the best use we can of the data that is available. There is considerable scope for improvement here, requiring fresh modelling or measurement approaches to generate the data that is required.

### *3.4.1 Air quality*

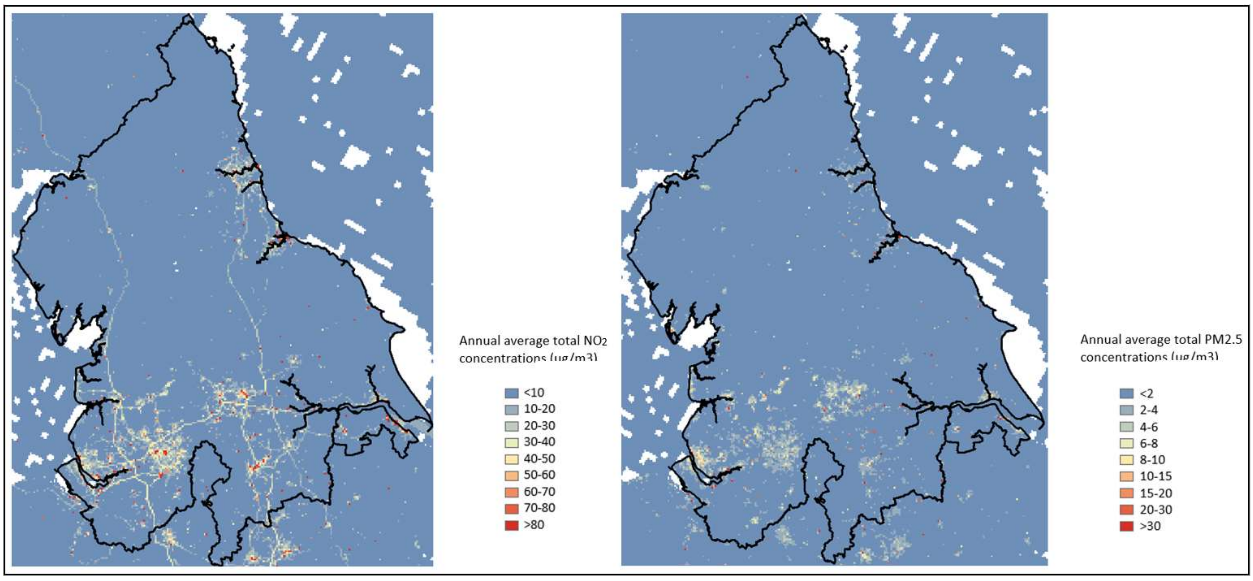
Outputs of Defra's Pollution Climate Mapping (PCM) were used for the calculation of air quality for each property (Figure 3.4). PCM is a collection of models for per pollutant designed to fulfil part of the UK's EU Directive (2008/50/EC) requirements to report on the concentrations of particular atmospheric pollutants. It provides outputs on a 1 km x 1 km grid of background conditions plus around 9,000 representative road side values. UK wide maps are available for $NO_x$ (as $NO_2$), $PM_{10}$ and $PM_{2.5}$ for the current reference year 2015, in terms of annual average concentrations. Detailed modelling methodology can be found in Ricardo Energy & Environment (2016).

Values of annual average concentrations of $NO_2$, $PM_{10}$ and $PM_{2.5}$ were assigned to each property by overlaying the concentration maps onto location points of properties in GIS. For this model, centroids of post code units where the properties were located were used as property location points.

Limitations of the Defra PCM data are that its spatial resolution is very low at 1 km x 1 km, and it is only available for 2015. Archived air quality datasets are based on past year emissions factors and used different modelling methodologies, which makes comparisons over time problematic.

One reason for caution with air quality as an explanatory variable in an HP model of property prices, is the difficulty that people have detecting some pollutants – particularly NOx and ultrafine particles (e.g. see Paas et al., 2016), although other pollutants such as larger particulates (PM10s) do seem to be perceptible (Nikolopolou et al., 2011).

Figure 3.4: Outputs of Defra's Pollution Climate Mapping (PCM) for $NO_2$ and $PM_{2.5}$




### *3.4.2 Noise*

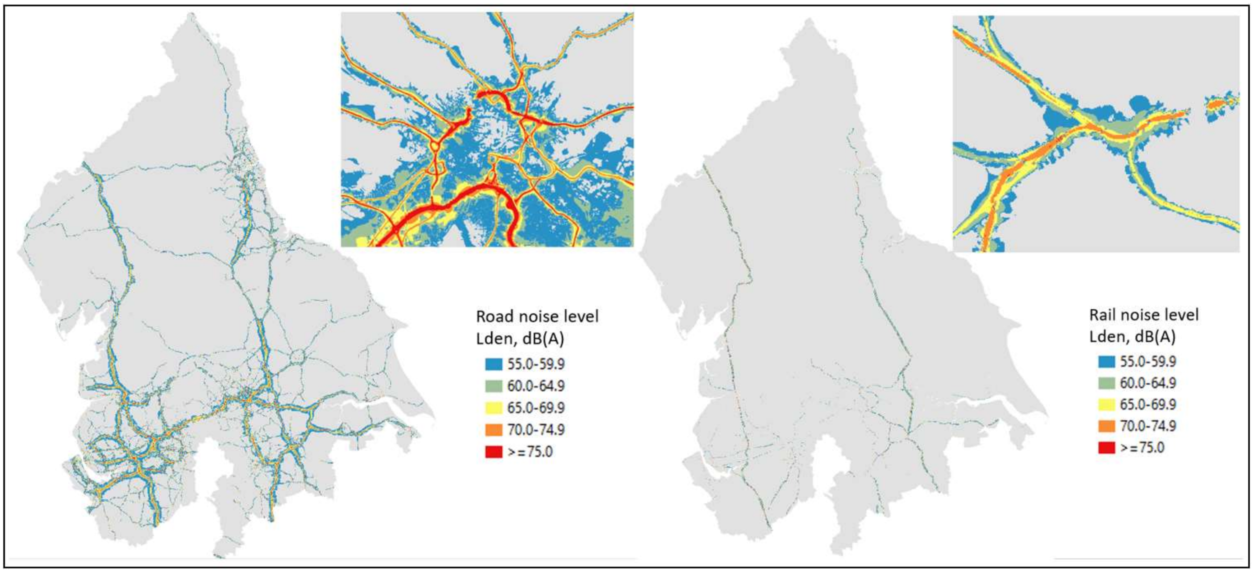
There is widespread evidence of noise being perceived by residents and influencing property prices (e.g. literature reviewed in Nellthorp et al., 2007). Potentially, noise is influential alongside air quality – particularly if the spatial pattern of impacts differs (at the spatial scale covered by the data). It is important to capture the benefits of proximity to local facilities such as shops and services separately (see Section 3.4.5), otherwise there is risk that the noise variable will pick up proximity to economic activities proxied to some extent by traffic and other noise sources.

For this model, outputs of Defra's strategic noise mapping for the year 2012 were used for the calculation of noise exposure at each property (Figure 3.5). The strategic noise mapping was developed as part of implementing the Environmental Noise Directive, and estimated noise from major road and rail sources across England in 2012, with noise from major airports mapped separately by relevant airport operators. For road and rail noises, the maps were published at 10 m x 10 m spatial resolution for 3 different noise indicators: annual average $L_{DEN}$, $L_{Aeq,16hr}$ and $L_{night}$. Detailed modelling methodology can be found in Defra (2015).

Values of annual average Lden of road and rail noises were assigned to each property by overlaying the noise maps onto location points of properties in GIS. Centroids of post code units where the properties were located were used as property location points.

Defra strategic noise maps have a high spatial resolution of 10m x 10m. However, they are only available for 2012, only cover major (and some minor) roads and railways, and only map noise levels above 55dB(A). Better noise maps (or noise levels at given receiver points) can be produced using commercial noise modelling software. However, like dispersion modelling for air quality, the noise modelling approach requires high quality traffic data, which can be a problem for a large study area considering data availability and computational cost.

Figure 3.5: Outputs of Defra strategic noise mapping 2012 for road and rail noise in $L_{den}$ (showing Leeds detail)




### *3.4.3 Greenspace*

Greenspace data was obtained from the MasterMap Greenspace Layer, which covers various types of greenspace including public parks and gardens, private gardens, playing fields, play spaces, allotments, religious grounds, outdoor sports facilities, woodlands, open natural and semi-natural lands, etc. Types of greenspaces considered in model developing are categorised and listed below, with examples shown in Figure 3.6:

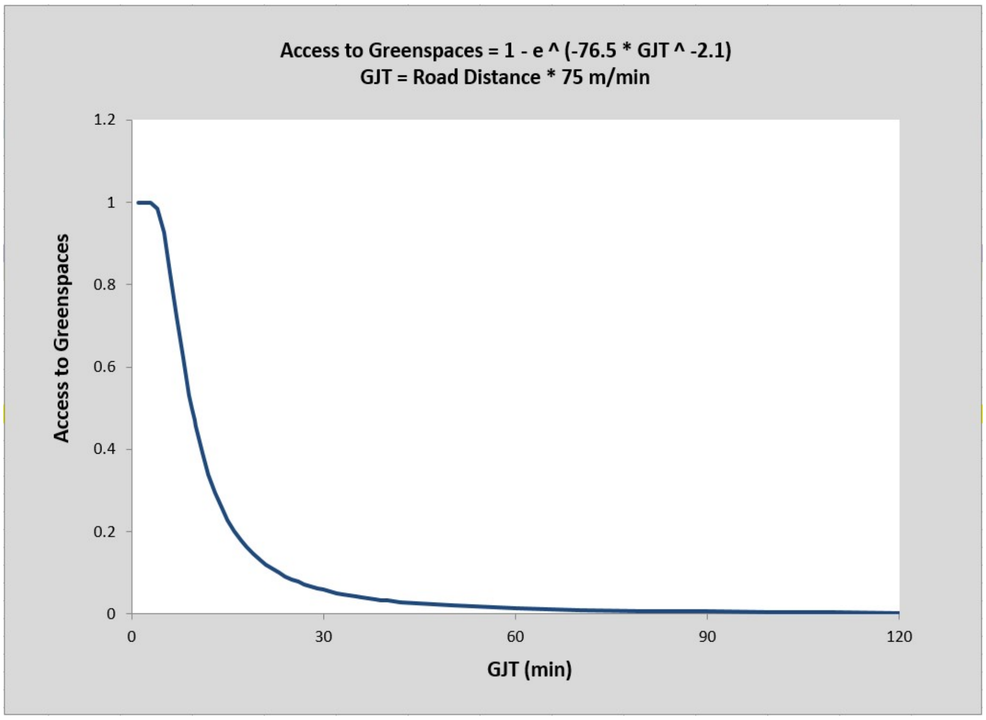
- Parks and gardens – including public parks and gardens;
- Playing fields and play spaces – including playing fields, play spaces, bowling greens and other maintained open fields for general activities;
- Open natural lands – including woodlands, open natural and semi-natural lands.

Other greenspaces such as allotments, school grounds, institutional grounds and religious grounds were not included, as they would be correlated with proximities to certain facilities and services, and many of them are not open to or used by the general public.

Greenspace variables initially calculated include areas of greenspaces of each category within ¼ mile and 1 mile radiuses, areas of and road distances to the nearest 3 parks/gardens and the nearest 3 playing fields/spaces, which are typically used in literature (Greater London Authority, 2010; Kong et al., 2007; Panduro and Veie, 2013; Talen, 1997). We further applied a deterrence function (Figure 3.7) to the distance-based variables, to account for the non-linear relationships found between property price and proximity to public services and facilities (Cortright, 2009; Debrezion et al., 2007). This deterrence function was based on the same functional form as the ones for Accessibility to Employment, but calibrated to an ‘access to local facilities’ context using DfT journey time statistics and other data.

Figure 3.6: Examples of greenspace types considered in model development – top left: park; bottom left: woodland; top right: playing field; bottom right: play space

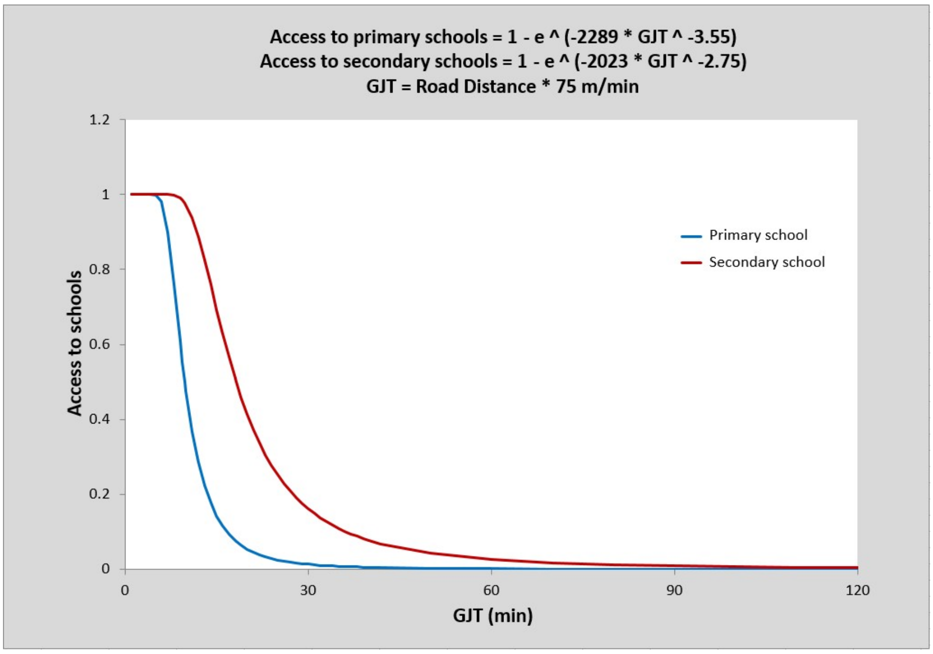

Figure 3.7: Deterrence function for access to greenspaces



### *3.4.4 Schools*

School variables consist of two types: (i) pupil performance by LSOA of residency; or (ii) quality of schools near to property *i*. For pupil performance, variables used were the average point scores of Key Stage 2 (primary) and Key Stage 4 (secondary) of the pupils residing in each LSOA in the 2013/14 academic year, published by the Department for Education. For quality of nearby schools, variables considered include Ofsted ratings of and road distances to the nearest 5 primary schools and the nearest 5 secondary schools, road distances to the nearest Ofsted outstanding primary school and to the nearest Ofsted outstanding secondary school. We further applied deterrence functions (Figure 3.8) to the distance-based variables to account for possible non-linear relationships between distance to schools and property price. Again these were based on the Accessibility deterrence functions, but recalibrated using DfT statistics on journey times to Primary and Secondary schools in the North.

Figure 3.8: Deterrence functions for access to schools

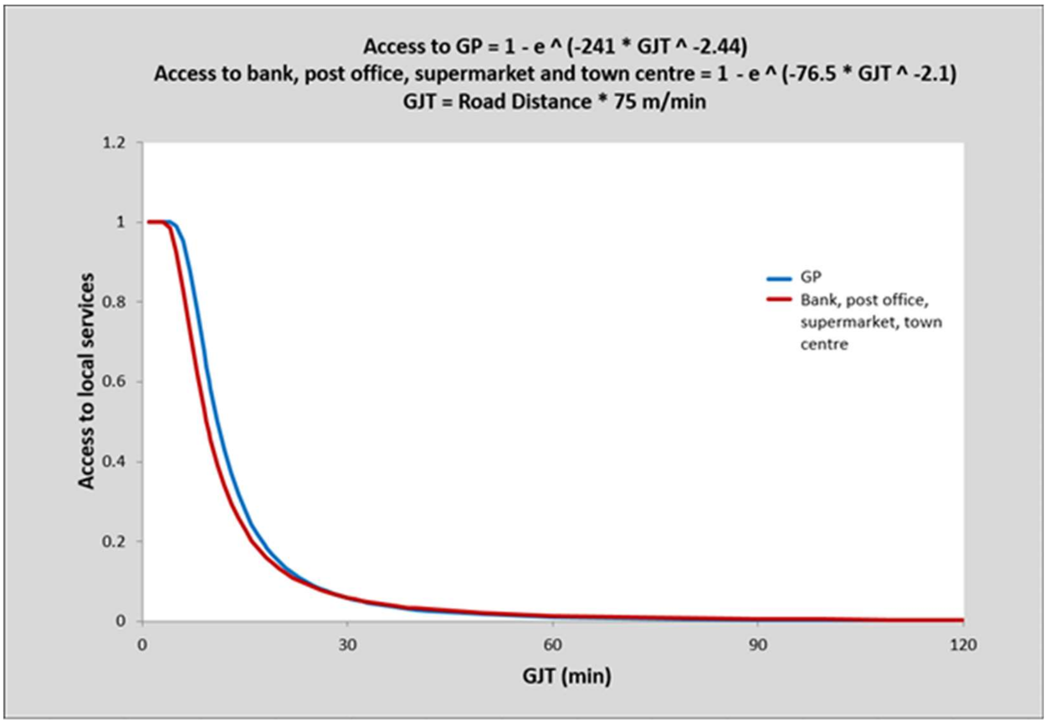



### *3.4.5 Local services and town centres*

Accessibility to other local services and town centres was used as a further set of descriptors of place quality, which measure the provision of amenities and key services in the local area, and are different from accessibility to employment opportunities and skills as described in Section 3.3. For local services, the variables used include road distances (for walking, cycling or driving) to the nearest supermarket, GP, bank and post office, with deterrence functions applied (Figure 3.9).

For town centres, travel time to the nearest town centre by walk and/or public transport for each LSOA obtained from DfT's journey time statistics was used, and also with deterrence functions applied (Figure 3.9). The central focal point for the town centre mapped to the nearest road was used as the location of the town centre for travel time calculation. Definition and calculation of town centre boundaries can be found in Office of the Deputy Prime Minister (2004). Briefly, the boundaries were calculated based on three components: Economy, Diversity, and Property. The Economy component contained numbers of jobs in commercial offices, retailing, entertainment, public administration, *etc.* as positive indicators, and numbers of jobs in manufacturing, warehousing, utilities, *etc.* as negative indicators, in each postcode unit. The Diversity component was calculated as the total number of different employment types as positive indicators in each postcode unit. The Property component was captured by mapping the density of retail and office floor space in each postcode unit. Note that these 2004 boundaries, calculated using data collected before 2004, will become increasingly inaccurate over time as different town centres expand and contract. Also note that this definition of town centres is heavily reliant on employment – we return to this issue in the Results section (4.1.6).

Figure 3.9: Deterrence functions for access to supermarkets, GPs, banks, post offices and town centres




### *3.4.6 Landfill sites*

Euclidean distances to and areas of the nearest three landfill sites were included as place quality variables, to account for the disbenefit of landfill site proximity. The data was obtained from the Environment Agency.

### *3.4.7 Students and tourists*

These variables were added although they did not appear in the initial list (Section 2.4). Apart from the potential role of students in representing Accessibility to Skills (Section 3.3.1 above), numbers of students and tourists can be indicators of some distinctive aspects of a place, such as vibrancy, attractiveness and diversity. The number of tourists may be a reflection of cultural and heritage quality – which are not reflected in any other variables included here. Both may also be linked to high demand for short-term and long-term accommodation. As input variables for the model, number of students in each LSOA was calculated as university term-time population minus non-term-time night time population, and number of tourists was estimated by annual visits to the LSOA by tourists. This data was from the HSL National Population Database.

### *3.4.8 Crimes and accidents*

The crime variable used in model development is the number of crimes in 2016 in each LSOA. Crimes include reported and recorded crimes, of which the data was taken from the Police dataset (data.police.uk). It does not reflect the number of actual crimes committed or those resulting in a conviction.

The accident variable used in model development is the number of road accidents occurring in 2016 in each LSOA. Accidents were recorded through the Police's STATS19 process, including all injury accidents (Killed, Seriously Injured, Slight), and excluding damage-only accidents and those not recorded using the STATS19 process. Multiple injuries occurring as part of the same accident are recorded as a single accident.

### *3.4.9 Other potential place quality variables not currently included*

The following variables could potentially be included however suitable datasets were not immediately available: airport noise; urban realm (tree canopy; road width; pavement width; severance); traffic – all roads.

## 3.5 Building & plot characteristics

The building characteristics come from two datasets: price paid data, provided by HM Land Registry; and energy certificate data provided by MHCLG. The data are matched based on street address and date of sale.

The price paid dataset denotes the type of property (Detached, Terrace, etc.) which is then coded into dummy variables in the modelling dataset.

The total floor area comes from the energy certificate dataset and is the total of all enclosed spaces within a property, measured to the internal face of the external walls, i.e. the gross floor area ($m^2$).

Habitable rooms, also from the energy certificate data, include any living room, sitting room, dining room, bedroom, study and similar.

## 3.6 Socio-economic variables / income *(G)*

Average income (equivalised, before housing costs) is calculated based upon small area estimates provided by the ONS for 2015/16. These have been further developed to LSOA area using the PAYE/benefit estimates, provided also by the ONS for the same year.

## 3.7 Supply-Demand balance variables *(S)*

Having identified this as a relevant issue across the North, we worked towards including variables relating to the *supply constraint* in local property markets. The following were considered:

a) residents or households vs housing stock (as described in Section 2.7.1);

b) rate of transactions, or transactions vs housing units for sale (as described in Section 2.7.1);

c) % of households on housing waiting lists;

d) property vacancy rates;

e) net additions to housing stock (as described in Section 2.7.2);

f) % changes in housing stock (as described in Section 2.7.2)

g) time on market or % of asking price achieved.

Pure supply side variables such as (e) and (f) have some theoretical appeal (as described in Section 2.7.2) and can be assembled in practice using ONS/MHCLG data at the LA level, which fits fairly well with local housing market areas. Other variables are harder to source on a consistent basis across the whole North, with the exception of (a) which touches on both supply and demand, and so risks picking up demand variations as well as supply constraints.

## 3.8 Commercial property model

As outlined in Section 3.2, price data for commercial properties is not as open as the residential data described previously, however it does exist on a number of databases. One such system is COSTAR[24], which collects and aggregates price data from commercial agents. Although large scale data retrieval is not possible from the system, a dataset covering the Leeds district for 2016 was made available. This dataset consists of asking rent per square foot by year for each commercial property advertised, which is used as the dependent variable in the hedonic model estimated in this report.

The model makes use of this bespoke commercial dataset, as well as other available datasets, including those used in the Residential model. The types of variables used (previously described in Section 2.8) are B (Building Characteristics), A (Accessibility) and Q (Place Quality). Further work beyond this report will aim to incorporate other variables.

### *3.8.1 Building (B) characteristics*

Section 2.8 set out the key building characteristics that would affect the lease price: Purpose, Quality and Size. For "Purpose of use", the COSTAR dataset includes an identifier to show whether a property is for office or industrial use. In the case of offices, a star rating is also available, indicating the quality of a premises[25]. Initial analysis shows that this is likely to influence asking rents, however there are very few high or low observations (those with one or five stars). Therefore categories have been created that combined one and two star together, and four and five star together. The COSTAR data also indicates the total size of the area for lease, which will also be included in model estimation.

### *3.8.2 Accessibility (A) of Commercial Premises*

Similar to the Residential model, accessibility of commercial premises is hypothesised to be a driver of prices. In the commercial model this would take two forms – accessibility to other economic activity (i.e. agglomeration effects), and accessibility to a workforce.

Initial analysis focussed on agglomeration using a measure of effective density; a form of Gravity Model discussed in TAG Unit 2.4 and defined as:

---

[24] See http://www.costar.co.uk for more information

[25] Further information on the rating system available: http://www.costar.co.uk/docs/librariesprovider5/knowledge-centre-documnets/ratingsytem.pdf

$$ED_J = \sum_{j \neq J} E_j \, GJT_{Jj}{}^{\alpha} \quad (21)$$

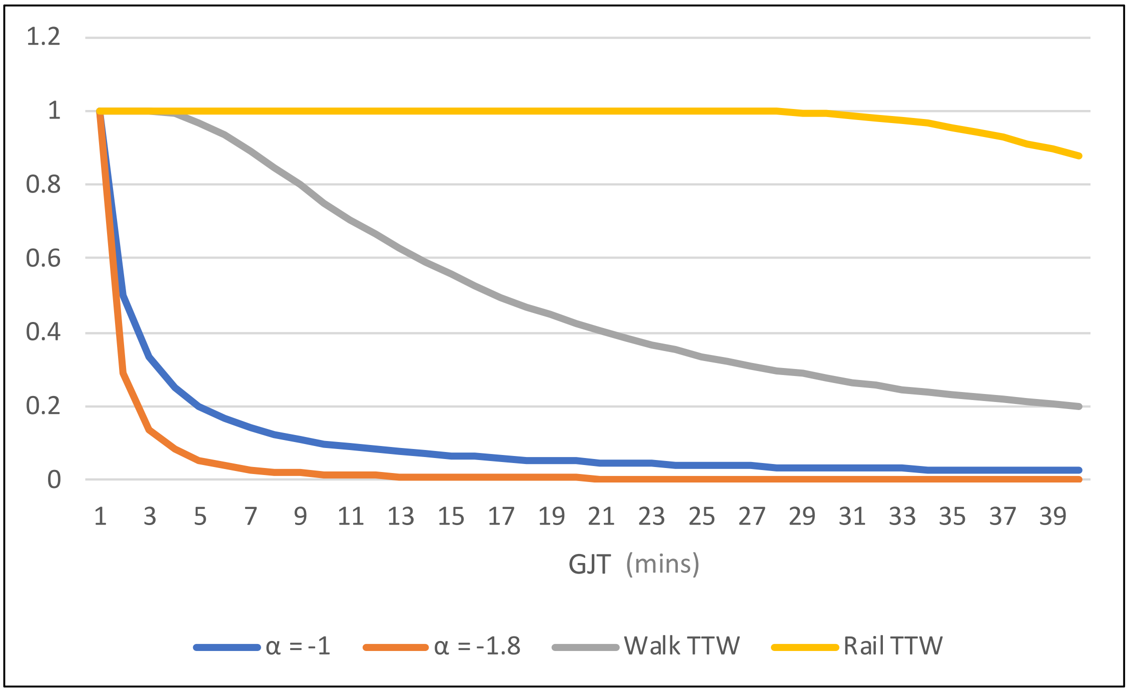
where $ED_J$ is the effective density at location *J*, $E_j$ is the employment at other locations *j*, $GJT_{Jj}$ is the generalised journey time between locations *J* and *j*, and *α* is the decay parameter, typically taking a value of between -1 and -2 (e.g. Graham et al, 2010, gives a value around -1.8 for Consumer and Producer Services). The deterrence function implied by this is steep and largely omits the effect of businesses more than a few minutes of GJT away (Figure 3.10). Conversely, it was found that walk and rail access to other employment sites, using the deterrence functions previously calibrated for travel-to-work in the Residential model (Section 3.3.1) allowed a wider and possibly more realistic range of accessibility (although this really requires more specific investigation), and so this was taken forward for model estimation (see Figure 3.10 for a comparison of the different deterrence functions).

The equation for the latter approach, based on employment potentiality in the residential model, is:

$$EP_J = \sum_{j \neq J} E_j \; 1 - e^{\alpha_{1m} GJT_{Jj}{}^{\alpha_{2m}}} \quad (22)$$

where $EP_J$ is the Employment Potentiality at location *J*, $E_j$ is the number of jobs at location *j*, and $\alpha_{1m}$ and $\alpha_{2m}$ are parameters which define the deterrence function for mode *m.*

Figure 3.10: Deterrence functions in agglomeration versus the access to employment model – comparison of alphas




Access to a workforce was calculated for rail and walk modes, again using previously calculated TTW deterrence curves. The approach uses a similar gravity model to that in equation (23), with the mass of the calculation being working age population in an MSOA:

$$PP_{jm} = \sum_{i} P_i \; 1 - e^{\alpha_{1m} GJT_{ji}{}^{\alpha_{2m}}} \quad (23)$$

where $PP_{jm}$ is the population potentiality at *j* for mode *m*, $P_i$ is the population at MSOA *i* and $\alpha_{1m}$ and $\alpha_{2m}$ are parameters which define the deterrence function for mode $m$.

For the rail and walk modes, both access to economic activity and access to population were calculated, separately, using the mode-specific deterrence function (there was no evidence found of how the function would vary for different journey purposes). This results in these two indicators being highly correlated for the rail mode; the empirical model would therefore be unlikely to include both for a single mode.

#### *3.8.3 Place Quality Variables (Q)*

In the cross-sectional residential model, income turns out to be a consistently positive and statistically significant variable (Section 4.1). In the case of the commercial model, income of the local resident population is used somewhat differently – as a proxy for the 'place quality' of the area in which the commercial property is located. The hypothesis is that commercial premises in higher income areas will tend to experience external benefits from the place quality attributes which tend to be associated with income of residents, e.g. the quality of local retail facilities, parks and urban realm. This in turn will allow the commercial property in these areas to command higher rents.

### 3.9 Time Series Residential model – Case Study approach

The cross-sectional hedonic model allows us to explain differences in house prices in different areas at a particular point in time. It is also important to consider temporal variation in property prices and how this is linked to changes in the explanatory variables, particularly the impact of transport investments. For this we can – in principle – use panel data (time series and cross sectional data, e.g. from 1995-2018) provided a sufficient dataset containing suitable variables can be compiled. Such an approach helps to address questions of causality: e.g. does the opening of a new or improved rail service increase property prices in a given area after vs before opening?

The panel data approach allows us to control for any impact of unobserved time invariant heterogeneity of areas (e.g. natural endowments of resources, historical employer links, topography, etc) on property prices – a particular shortcoming of cross-sectional models. It could also potentially be informative to examine the property market dynamics in this context, i.e. how long before or after a policy announcement we might expect to see a land value impact and how this plays out over time.

A panel-based analysis for the whole of the region would be too large a task to undertake with available resources. Therefore we agreed with TfN and WYCA on a case study approach, focusing on one intervention or package of interventions in a more limited spatial area. Analysis of this type has typically been structured by identifying 'treatment' areas subject to the transport improvement and 'control' areas which are otherwise similar but not subject to the transport improvement, with more robust approaches exploiting a random element of scheme selection.

A number of candidate case studies were considered (below):

- *Manchester Metrolink* – large scale investment in light rail which has been opened over a number of phases during the last 26 years. With the support and assistance of TfGM we have assembled a number of years' data across a range of variables, to complement the Land Registry price data.

- *Transpennine Express* – service frequency has been gradually during the period covered by the property market data. This provides an opportunity to study the impact of rail accessibility, as close as possible to the policy context of NPR. A consideration is that some of the variables may need to be synthesised – e.g. GJT from available evidence on past service patterns of Transpennine rail services.
- *Airedale and Wharfedale lines upgrade* – these were significant rolling stock based enhancements: both vehicle quality and capacity (not simultaneous).
- *Accrington-Burnley Manchester new services via Todmorden Curve* – new commuter rail services connecting towns/small cities with economic development needs and low property prices with the economic hub of Manchester city centre.
- *New stations* – other ‘new stations’ (1995-2018) exist across the TfN area network including Burkshaw Parkway, Kirkstall Forge and Apperley Bridge.

After examining data availability and giving the options full consideration, a choice was made in September 2018, Manchester Metrolink was selected for the case study. There have been surprisingly few major rail-based changes in accessibility across the region over the last 20-25 years, and the Metrolink network represents a relatively large scale change in accessibility affecting a relatively large population within the region – which should assist in identifying any property market impact. As well as providing initial data, TfGM have been supportive with discussions around scoping and potential modelling directions. The modelling approach has been to build a relatively simple panel regression model based on the difference between ‘treatment’ areas close to Metrolink stations and the rest of the GM area as ‘control’, but taking account of the multiple lines and Phases of Metrolink, testing different definitions of the ‘treatment’ area, and also controlling for the effect of multiple property types.

### 3.10 Data summary

Table 3.3 summarises the data contained in our dataset for modelling purposes.

Table 3.3: Data by category of variables

| ***P variables*** | ***Source; Features*** |
|---|---|
| Residential property price | Land Registry Price Paid data (street address) |
| Commercial property value | COSTAR data (VOA Rateable Value data – we are continuing to pursue detailed data through official channels) |
| ***A variables*** | ***Source; Features*** |
| Accessibility to employment * | HSL National Population Database (employment locations: LSOA level, by sector);<br>NoRMS (TfN Rail Model), 2015 for model calibration, 2033 for NPR scenario tests;<br>TRACC (initial data and walk access) |
| Accessibility to skills * | HE, FE and Training locations (skills locations: HSL National Population Database);<br>NoRMS (TfN Rail Model) and TRACC (initial data and walk access) |
| * by mode: rail (main mode); car (main mode); walk (main mode); other PT (main mode, e.g. bus). Generalised time-based measure including journey quality and time. | |

| *Q variables* | *Source; Features* |
|---|---|
| Greenspace - parks, public gardens, playing fields, play spaces, woodland, open semi-natural land | OS MasterMap Greenspace Layer, 2017 (location centroids, access points, area polygons); |
| Local schools | Ofsted school inspections, 2016/2017 (4-point school quality rating)<br>DfE Schools in England address data, 2017 (street addresses)<br>National Pupil Database KS2 and KS4 performance by pupil residency-based LSOAs, 2014 (KS2 and KS4 score statistics, LSOA) |
| Local facilities – supermarkets, post offices, banks, town centres, GPs | OS Points of Interest, 2016 (location points)<br>NHS Prescription Services GP Practices data, 2018 (street addresses)<br>DfT Journey time Statistics, 2016 |
| Students and tourists | HSL National Population Database, 2017 (population statistics, LSOA) |
| Safety (crime) | Police crime data, 2016 (crime statistics, LSOA)<br>IMD Crime domain data, 2015 (crime statistics, LSOA) |
| Road safety | STATS19 road safety data, 2016 (road traffic accident statistics, LSOA)<br>IMD Road traffic accident data, 2015 (road traffic accident statistics, LSOA) |
| Air pollution – $NO_2$, $PM_{2.5}$, $PM_{10}$ | UK National Atmospheric Emissions Inventory - Defra's Pollution Climate Mapping (PCM), 2015 (air pollution maps) |
| Noise pollution – road and rail noise | Defra strategic noise mapping, 2012 (noise maps) |
| Landfill proximity | Environment Agency Permitted waste sites – authorised landfill site boundaries, 2015 (location centroids, area polygons) |
| Greenspace - parks, public gardens, playing fields, play spaces, woodland, open semi-natural land | OS MasterMap Greenspace Layer, 2017; |
| Local schools | Ofsted school inspections, 2016/2017<br>DfE Schools in England address data, 2017<br>National Pupil Database KS2 and KS4 performance by pupil residency-based LSOAs, 2014 |
| Local facilities – supermarkets, post offices, banks, town centres, GPs | OS Points of Interest, 2016<br>NHS Prescription Services GP Practices data, 2018<br>DfT Journey time Statistics, 2016 |
| Students and tourists | HSL National Population Database, 2017 |
| Safety (crime) | Police crime data, 2016<br>IMD Crime domain data, 2015 |
| Road safety | STATS19 road safety data, 2016<br>IMD Road traffic accident data, 2015 |

| | |
|---|---|
| Air pollution – $NO_2$, $PM_{2.5}$, $PM_{10}$ | UK National Atmospheric Emissions Inventory - Defra's Pollution Climate Mapping (PCM), 2015 |
| Noise pollution – road and rail noise | Defra strategic noise mapping, 2012 |
| Landfill proximity | Environment Agency Permitted waste sites – authorised landfill site boundaries, 2015 |
| ***B variables*** | |
| Type | Detached/Semi-detached/Terraced/Flat, Land Registry |
| Tenure | Freehold/Leasehold, Land Registry |
| New build | Y/N, Land Registry |
| Floorspace | MHCLG Energy Performance of Buildings data |
| Number of habitable rooms | MHCLG Energy Performance of Buildings data |
| Ceiling height | MHCLG Energy Performance of Buildings data |
| ***G variables*** | ***Source; Features*** |
| Area average income per capita (equivalised) | MHCLG LSOA-level |
| ***S variables*** | ***Source; Features*** |
| Supply-demand variables | ONS/MHCLG; Dwelling stock (2011-16) at LA level |

The correlation matrix is shown in Appendix I.

# 4. Results

This section presents the results from the following models, as well as interpretation and the findings of sensitivity testing / additional modelling:

- cross-sectional residential models (Section 4.1);
- cross-sectional commercial models (Section 4.2);
- time series (panel data) residential models (Section 4.3).

### 4.1 Cross-Sectional Residential Model results

The final number of observations in the data is 160,563 properties, after a careful data cleaning process was undertaken to ensure all observations contain information on all relevant variables. All transactions took place in 2016 (according to the Land Registry 'Price Paid' dataset) within the boundaries of the TfN area.

The models estimated as part of an extensive specification search are all based on the conceptual model set out in Section 2 and refined in Section 3. The price of a property is understood as a function of hedonic prices (*p*) of the multiple attributes *z* (internal and external) that define it:

$$p = p(z_1, z_2, \dots, z_n) \tag{24}$$

and where the *z* attributes are grouped within an agreed set of categories based on the theoretical work, such that:

$$p = A, Q, B, G, S \tag{25}$$

where *p* = market price of a residential property;

*A* = accessibility to economic opportunities, using the employment potentiality metrics by different modes, from location *i*;

*Q* = place quality attributes of location *i*;

*B* = building and plot characteristics of the property;

*G* = socio-economic characteristics of households in the locality;

*S* = supply constraint variables.

During the model specification search, criteria for retention were:

- theoretical rationale – the model needs to have a theoretical foundation as described in Section 2, as well as good empirical performance – these considerations do generally reinforce each other when the modelling is done carefully;
- significance – t-ratios and significance levels of individual variables;
- impact on the statistical performance of the model as a whole, with particular attention to interactions with other variables where there is found to be a high level of correlation;
- goodness of fit – adjusted $R^2$ measure;
- expected sign – where variables were found to be 'wrong sign', i.e. the opposite of the expected sign, this was investigated.

The preferred model emerging from this sequence of work is a semi-logarithmic model of the following generic form. The time trend accounts for house price change over the course of a year:

$$LN(Price\ Paid) = \beta_0 + \beta_{A_i} Accessibility\ to\ Jobs_i + \beta_{Q_i} Place\ Quality_i + \beta_{B_i} Building + \beta_{S_i} Supply\ Constraint_i + \beta_{G_i} SocioEconomic_i + \beta_t Time\ trend \qquad (26)$$

Numerous variants of this model now exist. We will begin by reporting the version that has been used as the base version since early November when the NPR SOBC outputs were produced. We will refer to this as the *'Post-SOBC base model'*, Model 12 in the modelling sequence. This model has the same Rail accessibility coefficient as the 'SOBC model' (to 3 significant figures) so is consistent with the uplifts reported in the SOBC. It features just one minor change from the SOBC model: the additional squared term on floor area (see Section 4.1.3 below).

This model remains a good reflection of the cross-sectional models produced by the study overall. Since this model was estimated, a large number of sensitivity tests have been conducted – some of them producing useful additional findings – without fundamentally changing the nature of the model or the orders of magnitude of the most important coefficients. These additional models are discussed in Section 4.1.6.

The model parameters were estimated using ordinary least squares (OLS) and are reported on the next page in Table 4.1. The final column of Table 4.1 shows the implications for percentage change (%Δ) in house prices for each variable.

The following general observations can be made:

- The model is a fairly good fit to the data: Adjusted $R^2$ = 0.729 (this value is in line with existing similar models in the literature, normally achieving values between 0.5 and 0.7; for comparison, the global model of Mulley (2014) had an Adjusted $R^2$ value of 0.64).
- All reported explanatory variables are the expected sign and significant at the 99% level of confidence. Only a small number of variables have been omitted from the final model and none of them was expected to play a central role thus no issues identified – excluded variables are discussed in Section 4.1.7.
- Following an extensive model specification search, the outcomes are robust to different model specifications, including different sets of variables included and also different structures (e.g. linear model instead of semi-log). Multiple sensitivity tests were conducted to this end.
- Four main sets of model extensions have been explored:
    i. Income interactions;
    ii. A model including fixed effects dummies for Local Authorities (LA);
    iii. Spatial Error Model (SEM);
    iv. Different 'Town Centre' definition.

  The outcomes of these additional tests are reported in Section 4.1.6, and help to demonstrate that the main model presented below is robust.

Table 4.1: Cross-Sectional Residential Model results – Post-SOBC base model (Model 12)

| Variable | Units | Estimate | t-stat | %Δ House Price | |
|---|---|---|---|---|---|
| **Accessibility to Employmt.** | | | | | |
| Rail (P) | N of jobs (discounted by GJT) | 1.37E-07 | 15.42 | 0.14% | per 10k jobs |
| Walk (P) | N of jobs (discounted by GJT) | 3.18E-06 | 26.67 | 3.18% | per 10k jobs |
| Car (P) | N of jobs (discounted by GJT) | 1.86E-07 | 25.89 | 0.19% | per 10k jobs |
| **Place Quality** | | | | | |
| Primary School rate (Avg) | OFSTED; avg rating of 5 nearest schools | -0.09511 | -27.7 | -1.9% | per scale point in 1 school |
| Primary School rate (1) | OFSTED rating of nearest school | -0.00947 | -6.7 | -2.8% | per scale point |
| Primary School rate (5) | OFSTED rating of 5th nearest sch. | 0.00829 | 5.84 | -1.1% | per scale point |
| Secondary School rate (Avg) | OFSTED; avg rating of 5 nearest sch. | -0.12228 | -59.16 | -2.4% | per scale point in 1 school |
| Secondary School rate (1) | OFSTED rating of nearest school | -0.00300 | -3.1 | -2.7% | per scale point |
| Parks (P) | Potentiality [0,1] = f(walk time) | 0.03247 | 14.63 | 9.7% | worst to best |
| Playground (P) | Potentiality [0,1] = f(walk time) | 0.02234 | 9.14 | 2.2% | worst to best |
| Bank (P) | Potentiality [0,1] = f(walk time) | 0.01377 | 3.95 | 1.4% | worst to best |
| TownCentre (P) | Potentiality [0,1] = f(walk time) | 0.20814 | 36.32 | 20.8% | worst to best |
| Landfill (dist) | Metres (straight distance) | 0.00001 | 17.6 | 0.6% | per 1km |
| Students | Count term time resident students | 0.00013 | 11.03 | 1.3% | per 100 students |
| Tourists | Count annual tourists | 0.00000 | 14.89 | 2.1% | per 100,000 tourists p.a. |
| PM 2.5 | µg/m3 | -0.00869 | -11.77 | -0.9% | per µg/m3 |
| Crime | Annual recorded incidents | -0.00016 | -23.31 | -1.6% | per 100 incidents |
| Road noise (55 to 59 dB) | Dummy 1/0 (<55dB = BASE) | -0.02563 | -11.63 | -2.6% | relative to base |
| Road noise (60 to 64 dB) | Dummy 1/0 | -0.03909 | -11.45 | -3.9% | relative to base |
| Road noise (65 to 69 dB) | Dummy 1/0 | -0.04964 | -8.41 | -5.0% | relative to base |
| Road noise (70 to 75 dB) | Dummy 1/0 | -0.07483 | -5.18 | -7.5% | relative to base |
| **Building & Plot** | | | | | |
| Total Floor Area | m2 gross | 0.00854 | 114.08 | 6.5% | per 10m2 (at 90m2) |
| Total Floor Area^2 | m2 gross (sq) | -0.00001 | -41.63 | 4.5% | per 10m2 (at 180m2)* |
| New Build | Dummy 1/0 | 0.18106 | 79.06 | 18.1% | new-build premium |
| Property type (Semi=BASE) | | | | | |
| Terrace | Dummy 1/0 | -0.21427 | -119.3 | -21.4% | relative to Semi |
| Detached | Dummy 1/0 | 0.20573 | 110.97 | 20.6% | relative to Semi |
| Flat | Dummy 1/0 | -0.29864 | -78.12 | -29.9% | relative to Semi |
| % Owner occupied | %dwellings in LA(Social rent=BASE) | 0.00195 | 22.26 | 2.0% | per 10% relative to base |
| % Privately rented | %dwellings in LA | 0.00629 | 34.78 | 6.3% | per 10% relative to base |
| **Supply constraint** | | | | | |
| Dwelling Stock (5-year %Δ) | Dwelling stock %Δ(2016-2011); LA | | | | |
| More constrained areas | %Δ(2016-2011); below average | -2.69430 | -17.13 | -2.69% | per 1%Δ up to 2.5% |
| Less constrained areas | %Δ(2016-2011); above average | -1.35534 | -15.18 | -1.36% | per 1%Δ above 2.5% |
| **Socio Economic** | | | | | |
| LSOA Average Income | £ | 0.00004 | 165.7 | 4.11% | per £1,000 p.a. |
| (Constant) | | 10.16661 | 721.93 | | |
| Time trend | 11 dummies (one per month) | (not reported) | | | |
| **Adj $R^2$** | 0.729 | | | | |
| **Observations** | 160,513 | | | | |

*Notes: (P) = Potentiality indicator, using a deterrence function on journey time; (1) = nearest facility; (5) = 5th nearest facility; (Avg) = Average among 5 nearest facilities; (dist) = distance; (LA) = Local Authority level; *Illustrative example of the distribution of values*

In the following sub-sections, the findings for each group of variables are explained and discussed.

### *4.1.1 Accessibility (A)*

We can infer that:

- *Rail accessibility to jobs* has a positive impact on price, i.e. properties with better rail access to jobs are worth more than those with poorer accessibility, all else equal. The accessibility index is measured in units of number of jobs accessible, discounted by a deterrence function based on generalised journey time (Figure 3.2), and thus the interpretation is straightforward: access to an additional 100,000 jobs is worth a 1.4% premium on the price of a property, on average in the TfN area. The premium for rail accessibility between areas with the worst and best accessibility in the sample is 12.6%. This is based on the observed difference between worst and best, equal to approximately 917,000 'GJT-discounted' jobs.
- *Walk accessibility to jobs* is also positive and highly significant. Furthermore the premium for this mode appears to be substantial: the parameter implies that the price uplift is 31.8% for 100,000 jobs within reach by walking; and the premium going from worst to best walk accessibility would be 45.3% (based on the sample difference of approximately 142,000 between worst and best).
- *Car accessibility to jobs* completes the set of employment accessibility impacts and is also positive and significant. The average price premium for jobs accessible by car is slightly higher than for rail, with a 1.9% increase associated to an increase in 100,000 jobs. The difference in the indicator between worst and best is approximately 895,000, which is associated with a premium of 16.6%.
- *Accessibility to skills / training* is reflected, relatively simply, by the 'resident students per LSOA' measure which is grouped with Place quality in Section 4.1.2 below.

Further interpretation can be given to these results using some scenario tests – please see Section 4.1.8 below – and then in their application to NPR (Section 5). The omission of bus accessibility should not be interpreted as suggesting that bus accessibility is not valuable in the residential property market – far from it. The best interpretation is probably that bus service provision is in fact widespread across the study area – we found that less than 1% of homes in the TfN area are more than 1km away from a bus stop – and therefore bus accessibility is part of the 'base' in this model. Our attempts to measure the value of being closer to a bus stop, or the disbenefit of being in that 1% of non-bus connected homes are described in Section 4.1.7 below.

### *4.1.2 Place quality (Q)*

- Explanatory variables for school quality retained in this model are based on Ofsted ratings, of which 1 means 'Outstanding' and 4 means 'Inadequate'. The average rating of the five nearest *primary schools* shows a negative parameter estimate, which means that as expected the better the rating of nearby schools, the higher the property prices, all else equal. The premium per scale point improvement in one school is 1.9% (this is calculated dividing by five the estimate of -0.09511, which applies to the average of the five schools). Beyond this average effect, the uplift is higher for the nearest school and lower for the 5$^{th}$ nearest: 2.8% and 1.1% respectively. These are captured by additional coefficients on those schools' ratings that act as modifiers of the 'average effect'.

- Similar impacts are found for the five nearest *secondary schools*. The premium per scale point improvement in one school is 2.4% on average, and slightly higher (2.7%) for the nearest secondary school. The aggregated premium can potentially be very high in a scenario of five high ratings vs. five low ratings.

- Access to *parks and gardens* is modelled as a function of walking time to the 3 nearest parks or gardens; and access to *playing fields and play spaces*, as a function of walking time to the nearest playing fields or play spaces. These two variables are significant and positive. The walking time is adjusted by a deterrence function to account for the non-linear relationship between distance to these amenities and property prices shown in literature (the details of the deterrence functions were provided in Section 3 with the description of the dataset). For playing fields and play spaces, the access scale is from 0 to 1 where 0 means lowest access and 1 means highest, as this includes only the nearest facility. For parks, the access scale covers the summation of the nearest 3 parks and thus can take a value between 0 and 3. The premium for access to parks and gardens going from locations of worst to best access would be 9.7%; the premium for 'best' access – relative to no access – to playing fields and play spaces is 2.2%. However, both variables only apply to properties located in LSOAs where the LSOA average income is equal or higher than the average income in the sample (approximately equal to 27,600). A significant positive coefficient was not found for lower income areas. This suggests that this effect is not homogeneous across neighbourhoods, which could be linked to quality and other characteristics (e.g. safety perception) of the green spaces, as well as to preference heterogeneity among different individuals.

- Access to *town centres* and access to *banks* are also modelled using a scale from 0 to 1, as functions of walking time to the nearest ones and adjusted by deterrence functions. Both variables have positive parameters estimated, which means properties closer to town centres and/or banks are more expensive, all else equal. The premium for access to town centres between locations with best versus the worst access to town centres is 20.8% (1.4% for access to banks).

- Distance to nearest *landfill site*, measured as straight line distance, has a positive impact on price, which means properties further away from landfill sites are more expensive, all else equal. The premium per 1km further is 0.6%.

- Number of *resident students* positively impacts on price, and the premium is substantial at 1.3% per 100 increase in term time number of resident students in the LSOA . (Note that this is a separate effect from the premium (below) relating to the % of properties being privately rented).

- Number of *tourists* also positively impacts on price, and the premium per 100,000 increase in annual tourist visits in the LSOA is 2.1%.

- The parameter estimate on *crime* is negative, suggesting that areas with higher levels of recorded crime incidents have lower house prices, all else equal. The premium per 100 increase in number of annual recorded incidents in LSOA is -1.6%.

- *Air quality* has a positive impact on price. The lower the $PM_{2.5}$ concentration, the higher the price, all else equal. The premium per μg/$m^3$ increase in annual average PM2.5 concentrations is -0.9%.

- *Road noise* shows a negative impact on price. The price reduction increases with noise levels, with a 7.5% reduction in the price of properties exposed to the largest annual average noise levels of between 70 and 75dB, compared to those that experience noise levels lower than 55 dB. The premium increases gradually: relative to the base level (under 55dB), the prices discounts are 2.6% at 55-59 dB, 3.9% at 60-64 dB and 5% at 65-69dB.

#### *4.1.3 Building/Plot characteristics (B)*

- The parameter estimate for *floor area* is positive and highly significant: as expected, larger properties are more expensive, all else equal. A second parameter on the squared term (floor area squared) is negative and highly significant, indicating a non-linear relationship by which there are diminishing returns to additional floorspace (as a % of price) – so the estimated premium per additional $m^2$ is smaller the larger the property. For instance, for a property of $90m^2$, the premium per additional $m^2$ is 0.65%; the premium drops to 0.45% for a property of $180m^2$. However, in absolute terms, the money value of the premium per $m^2$ may still be similar across the board, as larger properties are also more expensive..
- *Detached houses* are, on average, worth more than other housing types (20.6% more than a semi-detached house); semi-detached houses are the second most expensive, followed by terraced houses which are on average priced 21.4% lower than a semi-detached. Flats are worth less than houses, all else equal (-29.9%). The magnitude of these terms is in line with previous research.
- *New builds* have an 18.1% premium on average, again consistent with previous research.
- The proportion of *privately rented*, *socially rented* and *owner occupied* properties in the Local Authority is included in the model as a proxy for various factors not included elsewhere, in particular housing quality and financial value. Social housing is not always built to the same standards as housing units in the private market, and recent research has shown that the housing quality across an area may have negative externalities (Koster and van Ommeren, 2019). Additionally, areas with a strong rental market make property more of a financial asset, thus generating a financial value premium for properties in the area. The coefficients are in line with expectations and are significant: an additional 10% of properties being privately rented (relative to socially rented) carries a premium of 6.3%; an additional 10% of properties being owner occupied (relative to socially rented) carries a premium of 2%.

#### *4.1.4 Supply constraint (S)*

- The impact of the *supply constraint* is approximated by measuring increases in the dwelling stock in the Local Authority area, and assessing their price impact. The % increase in dwelling stock over the last five years is used – representing medium term changes in supply. Areas with a higher % increase in stock are expected to be less supply-constrained than areas with a low %, hence a downward influence on price.
- The effect of % increase in dwellings turns out to be negative and significant, i.e. higher increases in dwellings are associated with lower prices. Furthermore, further sensitivity tests indicated that the premium was non-linear and thus the final model includes two separate coefficients to estimate the impact on low-growth (i.e. more constrained) and high-growth (i.e. less constrained) areas. The average % increase in dwellings in the sample (approximately 2.5%) was used to generate the split:
    - for LAs with dwelling stock growing at less than 2.5% in the last 5 years, an additional 1% of dwelling stock is associated with a decrease in price of 2.7%;

  - for LAs with dwelling stock growing more than 2.5% in the last 5 years, an additional 1% of dwelling stock is 'only' associated with a decrease in price of 1.36%. Thus, the 'inflation control' benefits of additional dwelling stock are larger in areas characterised by low growth in recent years.

#### *4.1.5 Socio-economic variables (G)*

- The model finds that properties in areas where there is higher *average income* will also be worth more (4.11% higher per extra £1,000 of average annual income), all else being equal. This is a substantial effect, and not a surprising one in view of previous HP models. Note that in this version of the model, this effect is separate from the other variables (rail accessibility, local school performance, etc).
- Additional modelling suggests that the impact of income does not necessarily accrue solely to the overall price of a property but also to the individual sub-components (i.e. the hedonic prices of accessibility, school quality, etc) as proposed in the theory section (2.6). The results of these and other additional model tests can be found in the next section (4.1.6).

#### *4.1.6 Additional Tests: Income Interactions; Model with LA Dummies; Spatial Regression Models; and Town Centre definition*

This section presents the results of a series of model extensions and additional tests. Four main topics have been investigated:

i) Interaction of variables with LSOA income. In the Post-SOBC base model (Model 12), income affects the overall property price, but we have argued theoretically that income might also affect the different components of the price separately. As we shall see, it has been found that some hedonic prices vary with the area-level income, e.g. rail accessibility seems to have a higher price premium in higher income areas.

ii) Fixed effects dummies for local authorities. This model includes a dummy for each local authority to pick up any price premium inherent to each LA. The results are in line with expectations, and most hedonic prices remain surprisingly robust in relation to the base model, however there are some interesting implications for the value of accessibility at local versus regional level.

iii) Spatial Error Models (SEM), Spatial Autoregressive (SAR) models and Geographically Weighted Regression (GWR). The OLS model assumes spatial independence between observations, however tests for spatial independence indicate that the residuals from this model are not randomly distributed over space. SEM and SAR models are estimated to take account of similar characteristics of nearby data points. These models have a higher adjusted $R^2$ than the OLS model, but the impact of this model specification on the meaning of parameter estimates requires careful consideration. A GWR model was also estimated, however the parameter estimates are not plausible – the size of the model (both in terms of parameters and dataset) appear to be causing issues and further work is required before a GWR model can be reported with confidence.

iv) Different 'town centre' definition. The town centre accessibility measure has so far – including extensions i) to iii) above – been based on the definition of town centres used in the DfT Travel Time Statistics. However, problems with that definition include that it is strongly related to employment numbers, weakly related to the retail element of a town centre, and based on 2004 data. Thus some places that are known locally as town centres (e.g. Horsforth or Chapel Allerton in the case of Leeds) are not identified in the data. Thus, we test a model that removes this Town Centre indicator, and interprets the presence of any 'Bank' as an alternative indicator for town centre – this increases the number of local centres that are treated as Town Centres.

#### 4.1.6.1 *Model with income interactions*

In the theory and modelling strategy sections (2.6.2 and 3.1.2), it was indicated that the income variable would be interacted with other variables. This allows the hedonic prices to be functions of income, thus allowing the housing characteristics to be valued differently depending on people's income.

To this end, several model specifications have been explored. These include the testing of income interactions with multiple sub-groups of variables and within different model structures (linear, semi-log and log-log). The testing was performed both at an initial stage with initial models, and refined later on with the final models. Below we set out the specification of income included as part of the final model (the rest of the model is identical to the previous specification):

$$\ln(Price) = \dots + \beta_{inc} * INC + \beta_x * X + \beta_{x\_inc} * X * (INC - \overline{INC}) \qquad (27)$$

Where INC is income (at the LSOA level; as specified in the data section), $\overline{INC}$ is the average LSOA income level in the sample (equal to £27,678), *X* is any other variable in the model (e.g. accessibility, school quality) that is interacted with income; $\beta_{inc}$, $\beta_x$ and $\beta_{x\_inc}$ are parameters estimated in the model.

Under this specification, the hedonic price of a variable *X* is thus defined as a function of income:

$$\frac{\partial\, ln(Price)}{\partial X} = \beta_x + \beta_{x\_inc} * (INC - \overline{INC}) \qquad (28)$$

The value of *X* can then vary with the income level of different areas. Furthermore, using a mean-centred income metric (i.e. $* INC - \overline{INC}$) allows us to still interpret $\beta_x$ as the average % impact on price of a unit increase in variable X (thus, facilitating a direct comparison with the estimates from the main cross-sectional model reported earlier). Hence, the resulting hedonic price (or marginal value) of *X* using only $\beta_x$ corresponds with the value at the sample average income.

The model specification that is reported below (Table 4.2) includes the income interactions with the following variables: all accessibility to employment indicators; secondary school rating; and the town centre accessibility indicator. The rest of the model is identical to the Post-SOBC base model.

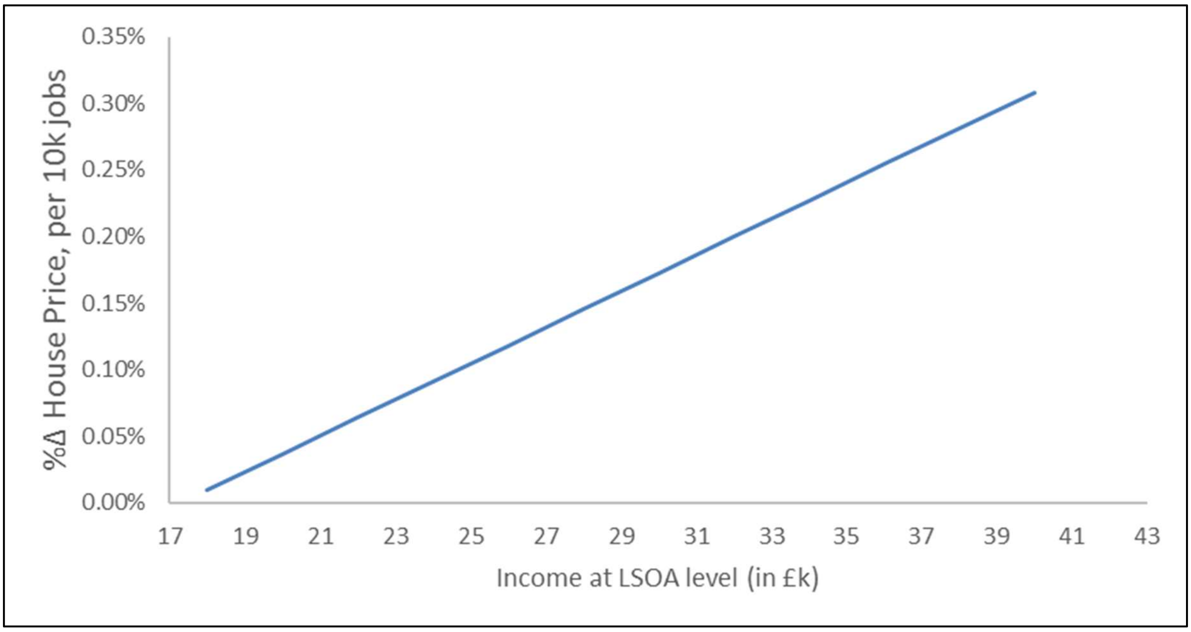
Table 4.2: Cross-Sectional Residential Models – with Income Interactions and LA Dummies

| | 12a. Income interactions | | | 12b. Local Authority dummies | | | |
|---|---|---|---|---|---|---|---|
| **Variable** | **Estimate** | **t-stat** | **%Δ Price** | **Estimate** | **t-stat** | **%Δ Price** | **Interpretation of %Δ** |
| **Accessibility to Employmt.** | | | | | | | |
| Rail (P) | 1.41E-07 | 15.95 | 0.14% | 1.31E-08 | 1.39 | 0.01% | per 10k jobs |
| Walk (P) | 3.24E-06 | 25.7 | 3.24% | 1.21E-06 | 8.9 | 1.21% | per 10k jobs |
| Car (P) | 1.81E-07 | 25.12 | 0.18% | -1.96E-07 | -13.33 | -0.20% | per 10k jobs |
| **Place Quality** | | | | | | | |
| Primary School rate (Avg) | -0.09413 | -27.48 | -1.9% | -0.10362 | -31.19 | -2.1% | per scale point (1 school) |
| Primary School rate (1) | -0.00987 | -6.99 | -2.9% | -0.01362 | -10.44 | -3.4% | per scale point |
| Primary School rate (5) | 0.00890 | 6.3 | -1.0% | 0.00900 | 6.81 | -1.2% | per scale point |
| Secondary Sch. rate (Avg) | -0.12245 | -59.07 | -2.4% | -0.06547 | -29.04 | -1.3% | per scale point (1 school) |
| Secondary School rate (1) | -0.00368 | -3.82 | -2.8% | -0.00898 | -9.85 | -2.2% | per scale point |
| Parks (P) | 0.01833 | 8.07 | 5.5% | 0.03574 | 16.91 | 10.7% | worst to best |
| Playground (P) | 0.02774 | 11.35 | 2.8% | -0.00719 | -3.16 | -0.7% | worst to best |
| Bank (P) | 0.01089 | 3.14 | 1.1% | 0.00787 | 2.42 | 0.8% | worst to best |
| TownCentre (P) | 0.23021 | 40.49 | 23.0% | 0.17202 | 31.31 | 17.2% | worst to best |
| Landfill (dist) | 6.72E-06 | 18.45 | 0.7% | 6.77E-06 | 17.8 | 0.7% | per 1km |
| Students | 0.00012 | 10.77 | 1.2% | 0.00014 | 10.81 | 1.4% | per 100 students |
| Tourists | 2.08E-07 | 15.21 | 2.1% | 1.23E-07 | 9.78 | 1.2% | per 100,000 tourists p.a. |
| PM 2.5 | -0.00841 | -11.92 | -0.8% | -0.00723 | -11.28 | -0.7% | per µg/m3 |
| Crime | -0.00015 | -21.87 | -1.5% | -0.00010 | -15.49 | -1.0% | per 100 incidents |
| Road noise (55 to 59 dB) | -0.02530 | -11.52 | -2.5% | -0.00500 | -2.36 | -0.5% | relative to base |
| Road noise (60 to 64 dB) | -0.03857 | -11.36 | -3.9% | -0.02334 | -7.17 | -2.3% | relative to base |
| Road noise (65 to 69 dB) | -0.04873 | -8.29 | -4.9% | -0.03291 | -5.9 | -3.3% | relative to base |
| Road noise (70 to 75 dB) | -0.07320 | -5.09 | -7.3% | -0.07345 | -5.36 | -7.3% | relative to base |
| **Building & Plot** | (not reported) | | | (not reported) | | | |
| **Supply constraint** | (not reported) | | | (not reported) | | | |
| **Socio Economic** | | | | | | | |
| LSOA Average Income | 3.57E-05 | 38.15 | 3.57% | 0.000038 | 161.59 | 3.80% | per £1,000 p.a. |
| *LSOA Income interactions* | | | | | | | (units as above, at two extreme points of the income distribution) |
| Income * Rail Jobs (P) | 1.36E-11 | 7.34 | 0.06% | | | | at 10% percentile of Inc |
| | | | 0.22% | | | | at 90% percentile of Inc |
| Income * Walk Jobs (P) | 1.38E-11 | 0.68 | 3.16% | | | | at 10% percentile of Inc |
| | | | 3.32% | | | | at 90% percentile of Inc |
| Income * Car Jobs (P) | 1.26E-12 | 0.86 | 0.17% | | | | at 10% percentile of Inc |
| | | | 0.19% | | | | at 90% percentile of Inc |
| Inc * Secondary rate (Avg) | -9.44E-07 | -2.52 | -2.34% | | | | at 10% percentile of Inc |
| | | | -2.56% | | | | at 90% percentile of Inc |
| Income * TownCentre (P) | 2.73E-05 | 23.22 | 7.49% | | | | at 10% percentile of Inc |
| | | | 39.27% | | | | at 90% percentile of Inc |
| **Local Authority Dummies** | | | | Reported in the Appendix | | | |
| (Constant) | 10.3082 | 347.45 | | | | | |
| Time trend | (not reported) | | | (not reported) | | | |
| **Adj R-squared** | 0.7307 | | | 0.7717 | | | |
| **Observations** | 160513 | | | 160513 | | | |

From the model in Table 4.2 with income interactions (Model 12a), it can be observed that:

- In comparison to the base model (Model 12), this model is a marginally better fit to the data (Adjusted $R^2$ = 0.7307 versus 0.7291).
- The model estimates (other than the new terms) are very similar, nearly identical, to those of the main model for all groups of variables (including Building and Plot characteristics and Supply Constraint variables, not reported in the table). Thus, all reported explanatory variables are again the expected sign and significant at the 99% level of confidence.
- The income interactions also take the expected sign and are highly significant in two cases: rail accessibility to jobs and town centre accessibility.
- The income interaction term with rail accessibility indicates that the value of improving rail access to jobs is always positive but is higher in higher income areas (and close to zero in the most deprived areas). The hedonic price at the average income in the sample, remains very close to that of the main model, estimated again at 0.14% increase in price for an additional 10,000 jobs.
- Additionally, the distribution of the hedonic price is now estimated as a function of income in the LSOA area of the property (see Figure 4.1). For instance, the premium for 10,000 jobs is +0.06% in an area of approximately £22,000 (the 10th percentile of the income distribution in the sample), and +0.22% in an area with approximately £33,600 (the 90th percentile of the income distribution in the sample).

Figure 4.1: Variation of the hedonic price premium on Rail accessibility to employment, by LSOA income




- The price premiums for Car access to jobs and Walk access to jobs do not vary significantly with income levels in the area.
- The value of town centre access also remains similar on average (23% versus 21% from the main model), but it now reflects also a large spread due to income in the area: i.e. being close to a town centre is more valuable the higher the income in the area.

*4.1.6.2 Model with Local Authority dummies (fixed effects)*

The second additional model is a Fixed Effects model which includes Local Authority dummies on top of the base model specification from the previous section:

$$\ln(Price) = \ldots + \beta_{LA_m} * LA_m \tag{29}$$

where $LA_m$ is a set of 76 dummy indicators (taking a value 0 or 1), corresponding to the 76 Local Authorities of the TfN area in the sample; $\beta_{LA_m}$ is a set of 75 parameters to be estimated, one for each LA excluding the base category. Liverpool was arbitrarily chosen as the base category (the choice of base category does not influence the results). Each $\beta_{LA_m}$ would pick up the price premium associated with that particular LA, in theory above and beyond the impacts of the observed variables estimated in the rest of the model.

Due to a collinearity problem, four variables of the model had to be dropped. These are the supply constraint indicators and the percentage of owner occupied and privately rented properties, all of which are measured at the LA level, thus causing an identification problem once the LA dummies are included.

From the model in Table 4.2 with Local Authority dummies, the following observations can be made:

- Model fit is better than in previous models (Adjusted R-squared = 0.7717), with most LA dummies' coefficient estimates being significant (see the Table in Appendix II).
- Most variables take the expected sign and are still highly significant, with a few exceptions: rail access (not significantly different from zero), car access to jobs (negative sign), local access to playgrounds (negative sign).
- Overall, considering that now 75 LA dummies are picking up much of the variation in price across properties, it is fair to say that the model is robust as the large majority of estimates remain significant and with only small changes in the order of magnitude in most cases.
- It comes to no surprise that the most affected estimates in this model, compared with the model without LA dummies, are the accessibility indicators. This is because these indicators measure rail and car accessibility to jobs, and variation in those indicators is larger across LAs than within LAs. Thus, there is confounding between the LA dummies and the rail and car accessibility indicators, with the LA dummies capturing part of the value of accessibility to jobs in different LAs.
- When the LA dummies control for differences across LAs, the remaining accessibility estimates are naturally picking up the remaining localised variation in accessibility within each LA:
  - Walk access to jobs remains positive and significant, although its value is now roughly a third of the value in the base model (1.21% per 10,000 jobs instead of 3.18%)
  - Rail access to jobs is not significant.
  - Car access to jobs is now negative and highly significant, possibly indicating some negative effects associated with car accessibility or roads in general at the local level that are not included anywhere else in the model. The outcomes suggest that at the local level (intra-LA), being accessible to jobs by car does

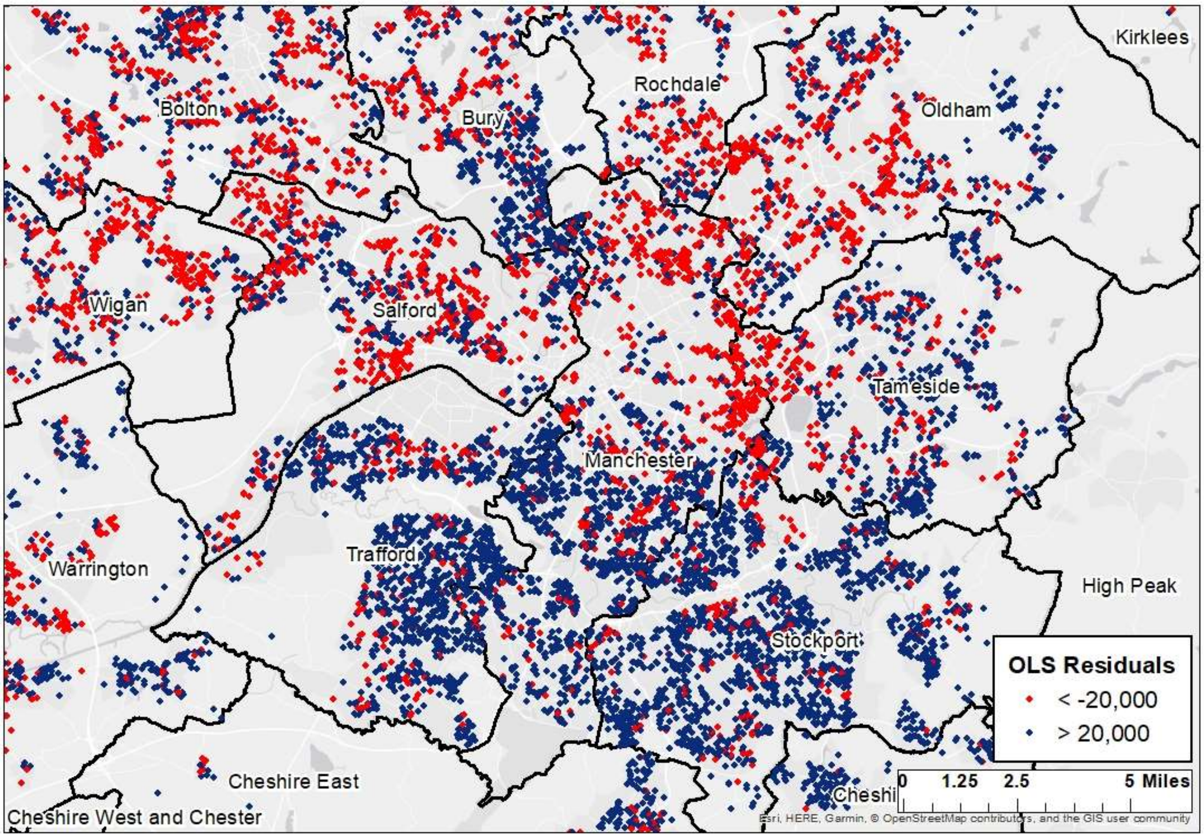
not carry a positive premium compared with other properties in that LA, but instead is associated with a disbenefit.

- Place quality coefficients are of the same order of magnitude as those in previous models; the same applies to Building and Plot characteristics (not reported). These variables vary much more at the local level and thus the introduction of LA dummies has less of an impact.

*4.1.6.3 Spatial Error Models (SEM), Spatial Autoregressive (SAR) Models and Geographically Weighted Regression (GWR)*

In order to formally test for spatial correlation in the OLS dataset, a spatial weights matrix was estimated based upon Thiessen polygons and Queen Contiguity i.e. the matrix indicated a “1” if data points were neighbours, otherwise “0”. Distance based bandwidths could not be calibrated in a dataset of this size (approximately 160,000 records dispersed across 40,000km$^2$). Using this matrix, a *Moran’s I* test revealed evidence of spatial autocorrelation. This phenomenon can be confirmed by spatial examination of the residuals from the OLS model. Figure 4.2 below is a plot of OLS residuals across the Greater Manchester area, showing (generally) negative residuals in the north and positive residuals in the south i.e. the errors are not spatially independent.

Figure 4.2: Map of residuals resulting from the global OLS model for the Manchester area




Further tests indicated that the Spatial Error Model (SEM) would be an appropriate model. This model takes the form:

$$\ln(Price) = \beta_0 + \beta X + \lambda W\mu + \upsilon \quad (30)$$

Where there are $K*1$ $\beta$ regression coefficients, $X$ is an $N*K$ matrix of observations, $W$ is the $N*N$ weights matrix, $\mu$ is the spatial error term, $\upsilon$ is the classical error term, and $\lambda$ is the spatial error coefficient. Note that if $\lambda=0$ then the regression becomes the global OLS regression as estimated earlier.

The SEM is compared to a reduced parameter OLS in Table 4.3.

Table 4.3: Cross-Sectional Residential Models – comparison of OLS and SEM

| | **12c. OLS Comparator** | | **13. Spatial Error Model** | | |
|---|---|---|---|---|---|
| **Variable** | **Est.** | **t-ratio** | **Est.** | **z-value** | **% difference** |
| **Accessibility to jobs** | | | | | |
| Rail Jobs (P) | 1.57E-07 | 18.13 | 3.97E-08 | 1.91 | -74.75% |
| Walk Jobs (P) | 3.24E-06 | 31.44 | 1.86E-06 | 6.60 | -42.68% |
| Car Jobs (P) | 1.89E-07 | 27.17 | 2.12E-07 | 10.58 | 11.73% |
| **Place Quality** | | | | | |
| Primary School rate (Avg) | -0.09413 | -28.20 | -0.11164 | -16.28 | 18.60% |
| Primary School rate (1) | -0.00923 | -6.58 | -0.01454 | -7.05 | 57.56% |
| Primary School rate (5) | 0.00785 | 5.54 | 0.01242 | 6.92 | 58.34% |
| Secondary School rate (Avg) | -0.11895 | -59.46 | -0.11661 | -22.76 | -1.96% |
| Secondary School rate (1) | -0.00348 | -3.65 | -0.01248 | -6.31 | 258.30% |
| Parks (P) | 0.03557 | 13.02 | 0.00519 | 1.52 | -85.41% |
| Bank (P) | 0.01623 | 5.10 | -0.04720 | -7.66 | -390.80% |
| TownCentre (P) | 0.21329 | 40.53 | 0.14627 | 15.02 | -31.42% |
| Landfill (dist) | 6.60E-06 | 20.00 | 7.96E-06 | 8.28 | 20.46% |
| Students | 0.00012 | 19.77 | 5.75E-05 | 6.50 | -52.84% |
| Tourists | 2.08E-07 | 18.72 | 7.70E-08 | 5.72 | -63.02% |
| PM 2.5 | -0.00853 | -36.80 | -0.00427 | -12.53 | -49.94% |
| Crime | -0.00016 | -35.45 | -4.99E-05 | -8.01 | -69.67% |
| Road noise (55 to 59 dB) | -0.02630 | -11.40 | -0.00160 | -0.65 | -93.91% |
| Road noise (60 to 64 dB) | -0.03982 | -11.54 | -0.01689 | -5.09 | -57.58% |
| Road noise (65 to 69 dB) | -0.04927 | -8.29 | -0.03604 | -6.91 | -26.86% |
| Road noise (70 to 75 dB) | -0.07649 | -5.68 | -0.09476 | -7.78 | 23.89% |
| **Building & Plot** | | | | | |
| Total Floor Area | 0.00851 | 131.62 | 0.00823 | 158.33 | -3.29% |
| New Build | 0.17873 | 71.92 | 0.17030 | 55.10 | -4.71% |
| **Property type (Semi=BASE)** | | | | | |
| Terrace | -0.21536 | -127.47 | -0.17188 | -115.93 | -20.19% |
| Detached | 0.20432 | 100.06 | 0.21863 | 128.10 | 7.00% |
| Flat | -0.29957 | -85.45 | -0.35136 | -116.78 | 17.29% |
| % Owner occupied | 0.00188 | 23.95 | 0.00277 | 11.77 | 47.29% |
| % Privately rented | 0.00622 | 36.72 | 0.00521 | 10.10 | -16.29% |
| **Socio Economic** | | | | | |
| LSOA Average Income | 4.32E-05 | 224.39 | 2.66E-05 | 88.57 | -38.44% |
| (Constant) | 10.0662 | 809.01 | 10.5625 | 362.78 | |
| R-squared | 0.73 | | 0.84 | | |
| AIC | 44328 | | -21966 | | |
| Observations | 160,563 | | 160,563 | | |
| Lambda | | | 0.77879 | 339.64 | |

Notes: certain variables were omitted from the OLS and SEM run: Playground, Floor Area$^2$ and Supply Constraint; (P) = Potentiality indicator, using a deterrence function on journey time.

Comparing the Spatial Error Model (SEM) with the OLS model in Table 4.3, it can be observed that:

- Model fit is improved: the pseudo $R^2$ of the SEM is 0.84, compared to the corresponding OLS $R^2$ of 0.73 (although note these $R^2$ measures have a different basis). To confirm, the Akaike Information Criterion (AIC) decreases substantially in the case of the SEM, indicating better fit to the data.
- The Lambda estimate is statistically significant, confirming that the model is capturing spatial autocorrelation errors.
- The variables Playground, Floor Area$^2$ and Supply Constraint were removed from the SEM estimation due to apparently causing issues with model estimation. Estimation of the OLS without these variables shows little overall difference to the base model in Table 4.1.
- All parameter estimates take the expected sign with the exception of Bank, which is now negative in the SEM. The impact of proximity to banks on house prices may now be being captured in the error term i.e. there being clusters of similarly priced houses close to banks. This may also account for the reduction in the Town Centre variable.
- The Rail Accessibility parameter decreases substantially in the SEM (approx. -75%) when compared to the corresponding parameter in the OLS model and is no longer significant at 5% (z-value 1.91). This is of concern when looking at the impact of rail access to jobs on house prices. Our analysis has shown that housing proximity to stations has an impact upon this variable. However as houses with high rail access values are clustered around rail stations, the SEM is likely attributing some of this high accessibility to the error as spatial correlation. This impact (likely to be affecting other variables, such as access to Banks and Schools) raises the question as to whether accounting for spatial autocorrelation is problematic in a policy context where we are attempting to isolate the effect of location specific characteristics on house prices.
- The walk parameter decreases by 50% although remains significant at 5%. This is also likely due to clustering of properties with high walk access to jobs (e.g. city centres).
- There appear to be some parallels between the impact of the LA fixed effects model (immediately above, in Section 4.1.6.2) and the impact of the SEM model reported here, in terms of their effect on the variables of policy interest – notably rail accessibility and walk accessibility.

Further tests were undertaken using an alternative Spatial Autoregressive (SAR) model in place of the SEM model. SAR models incorporate a ‘spatial lag’ term which is a weighted average of the prices of neighbouring properties. The findings were broadly in line with those for SEM, although the model fit was slightly lower and the rail parameter estimates slightly closer to OLS. Table 4.4 shows the indicators of model fit, which suggest that if model fit was the only criterion, SEM would outperform SAR and OLS.

Table 4.4: Spatial regression models: model fit and spatial parameters

| | OLS | SAR | SEM | OLS+income | SAR+income |
|---|---|---|---|---|---|
| Rho-squared | 0.7249 | 0.8115 | 0.8397 | 0.7322 | 0.8149 |
| AIC | 45913.2 | -6758.3 | -22057.9 | 41613.4 | -9884 |
| Wy | | 0.5066 (254.4) | | | 0.4989 (249.8) |
| Lambda | | | 0.7213 (324.2) | | |

Note: figures in parentheses are t-ratios for the spatial parameters; "+income" indicates income interaction terms are included.

Table 4.5 shows the implications for the rail, walk and car accessibility to employment parameters, when the model type is changed. The relative stability of these parameter estimates, in the face of fundamental changes to the model type, gives further encouragement on the robustness of the results.

Table 4.5: SAR and SEM model estimates compared with OLS, % change in value per 10,000 accessible jobs

| | OLS | | SAR | | SEM | |
|---|---|---|---|---|---|---|
| | Estimate | t-ratio | Estimate | t-ratio | Estimate | t-ratio |
| Accessibility to Employment by: | | | | | | |
| Rail | 0.19% | 22.1 | 0.17% | 11.3 | 0.13% | 6.6 |
| Walk | 3.28% | 31.2 | 5.52% | 31.3 | 2.45% | 9.9 |
| Car | 0.19% | 26.8 | 0.18% | 14.9 | 0.18% | 10.8 |

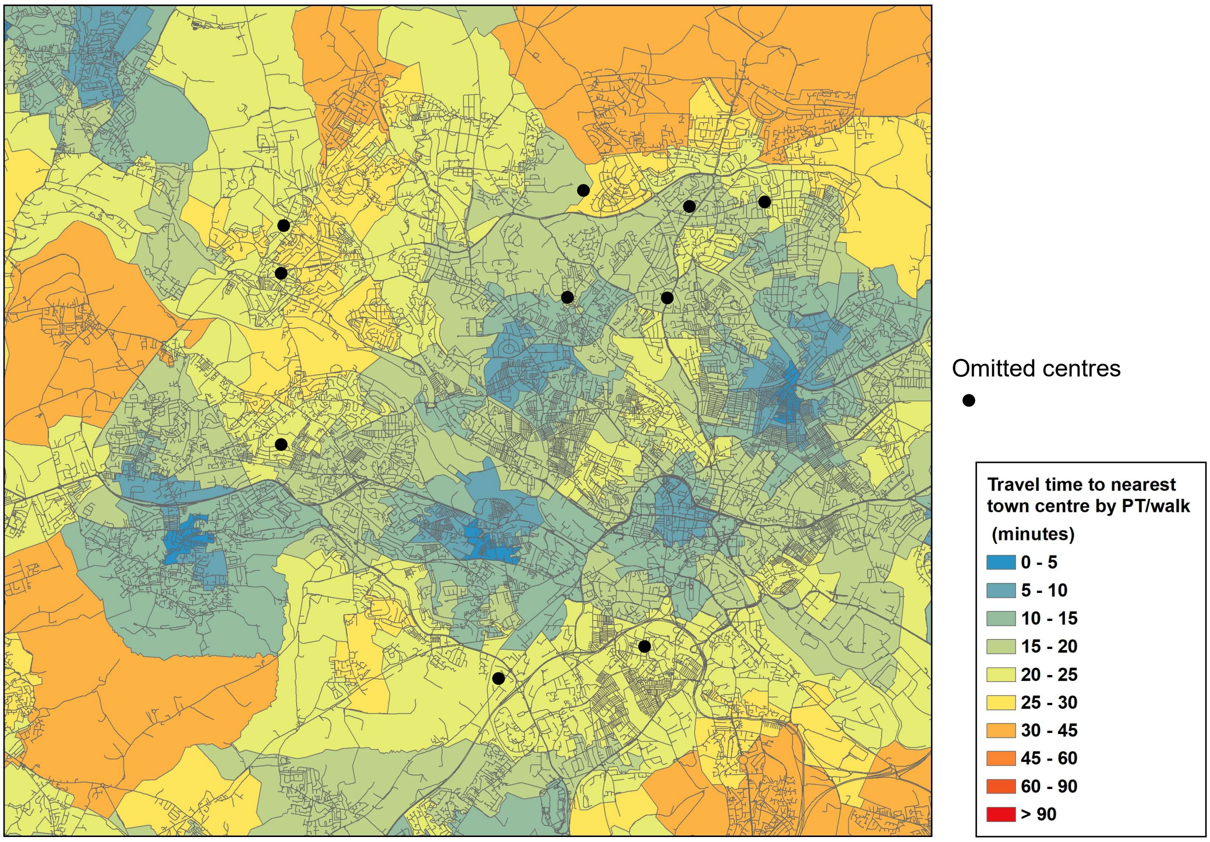
Finally, Geographically Weighted Regression (GWR) has the potential to estimate spatially varying parameters, which could account for correlations in errors over space, as well as revealing differences in the weight placed on attributes in each area. Although a GWR model was estimated in this study, with parameter estimates initially appearing sensible, it is not covered further in this report. This is because plots of the rail parameter estimates are inconsistent with expectations; the patterns of the spatially varying parameters are not plausible. Ultimately the model results could not be defended. The size of the model (both in terms of parameters and dataset) appears to be causing issues with estimation, and further work is required before a GWR model can be reported with confidence. A peer reviewer commented that these models are generally difficult to estimate on large datasets and therefore would require substantial aggregation of this model in order to run (e.g. to LA level – which would defeat the purpose to some extent). The income interaction terms and spatially varying parameters in this model make the need for GWR less pressing in this case.

#### *4.1.6.4 Models with a different town centre definition*

The Town Centre accessibility variable has so far been based on town centres as defined in the DfT Travel Time Statistics, a definition which is itself based on work by ODPM (2004, see Section 3.4.5 for details). However, given that employment is an element of that definition, an in-depth analysis of this variable and its impact in the models has been conducted. Key observations from our preliminary analysis on the town centre variable are:

- the Town Centres identified in the Town Centre accessibility variable are an incomplete set;
- using Leeds as an example (Figure 4.3 below), it can be seen that some centres are recognised – these are shaded in the darkest blue (Leeds City Centre, Headingley, Harehills, Armley, Pudsey and Yeadon) – while several others, including some substantial centres, are omitted (e.g. Horsforth, Chapel Allerton, Moortown, Roundhay, Meanwood, Beeston, Bramley, and some significant out-of-town centres such as Moor Allerton Centre and J1 Retail Park).

Figure 4.3: Town Centre definition and omitted centres (Leeds)




Reasons for the above observations are primarily:

- the ODPM Town Centres are based heavily on employment – which we already capture in other variables (i.e. our accessibility to jobs indicators for Walk, Rail and Car) – and commercial floorspace, whereas we are really only interested in the retail floorspace element of the ODPM definition;
- a variable which captures a wider range of local centres, including those which have a substantial retail/community offer, would be useful;
- also since they were defined in 2004 or earlier, it is questionable whether the ODPM Town Centres capture the current retail offer in places like Chapel Allerton and Horsforth, as opposed to the retail offer 15 or more years ago.

After considering options, two additional models were tested (see Table 4.6):

i) Model 14. A model that removed the original Town Centre variable from the base model (everything else remains identical). We also make the assumption that the 'Bank' variable is reinterpreted as a proxy for town centre, under the assumption that banks are typically located in town centres and a wider set of local centres meet this criterion. (Note that the original 'Town Centre' variable and 'Bank' are correlated but only at 0.5);

ii) Model 14a. As Model 14, but adding income interaction terms with key variables (similar to Model 12a), and taking the opportunity to improve upon the specification of Model 12a.

From Model 14, which excludes the original Town Centre variable, it can be observed that:

- In comparison to the base model in Table 4.1, the model fit is marginally worse (Adjusted $R^2 = 0.7264$). This is expected as we have only removed a previously highly significant coefficient – but we have done so purposely on theoretical/ definitional grounds.
- The majority of the model estimates remain largely unaffected when compared to Model 12 – showing once more the robustness of the results to changes in the set of explanatory variables.
- The rail access to jobs term is now slightly larger, with the effect per additional 10,000 jobs being 0.17% (compared to 0.14% in Model 12). Changes in the other accessibility terms are negligible on average.
- The estimate on Bank accessibility is 6 times larger than in Model 12, indicating a 6.6% premium for properties with Bank accessibility indicators equal to 1 (best; i.e. closest to a bank) relative to 0 (worst; furthest away from a bank). Since there is no strong argument to believe that banks – by themselves – are highly valuable amenities and drivers of property prices, we take this estimate to suggest that banks are indeed a useful proxy for town centres.
- Relative to the base model (Model 12), the price premium on town centre accessibility (now measured through banks and not via the former town centre definition) thus drops from roughly 20% to the aforementioned 6.6%.

Model 14a, which adds income interaction terms to the previous model, brings the following further insights:

- This model is a slightly better fit to the data, relative to the previous Model 14, relative to the base model (Model 12) and relative to the base model with income interactions (Model 12a). This model has an Adjusted $R^2 = 0.7308$.
- Most model estimates are very close to those of Model 14 for all groups of variables (including Building and Plot characteristics, Supply Constraint and Time trend variables, not reported in the table).
- The premium for Rail access to jobs is always positive and significant but significantly varies with income. For instance, while it is on average 0.16% per additional 10k jobs, it is estimated at 0.11% for areas with average income of £22,000, and at 0.21% for areas with average income equal to £33,600 per annum (the $10^{th}$ and the $90^{th}$ percentiles).

Table 4.6: Cross-Sectional Residential Models – excluding original 'Town Centre' variable

| | **14. Town centre excluded** | | | **14a. As 14+Income interactions** | | | |
|---|---|---|---|---|---|---|---|
| **Variable** | **Estimate** | **t-stat** | **%Δ Price** | **Estimate** | **t-stat** | **%Δ Price** | **Interpretation of %Δ** |
| **Accessibility to Employmt.** | | | | | | | |
| Rail (P) | 1.66E-07 | 18.61 | 0.17% | 1.56E-07 | 17.93 | 0.16% | per 10k jobs |
| Walk (P) | 3.15E-06 | 26.3 | 3.15% | 3.61E-06 | 29.36 | 3.61% | per 10k jobs |
| Car (P) | 1.83E-07 | 25.4 | 0.18% | 1.93E-07 | 26.72 | 0.19% | per 10k jobs |
| **Place Quality** | | | | | | | |
| Primary School rate (Avg) | -0.099 | -28.75 | -2.0% | -0.10028 | -29.26 | -2.0% | per scale point (1 school) |
| Primary School rate (1) | -0.00871 | -6.13 | -2.9% | -0.00981 | -6.97 | -3.0% | per scale point |
| Primary School rate (5) | 0.007242 | 5.08 | -1.3% | 0.007387 | 5.24 | -1.3% | per scale point |
| Secondary Sch. rate (Avg) | -0.12196 | -58.65 | -2.4% | -0.12086 | -58.55 | -2.4% | per scale point (1 school) |
| Secondary School rate (1) | -0.00421 | -4.33 | -2.9% | -0.00587 | -6.09 | -3.0% | per scale point |
| Parks (P) | 0.041152 | 18.49 | 12.3% | 0.019767 | 8.9 | 5.9% | worst to best |
| Playground (P) | 0.018665 | 7.59 | 1.9% | 0.008668 | 3.52 | 0.9% | worst to best |
| Bank (P) – Town C. proxy | 0.066146 | 19.89 | 6.6% | 0.075899 | 23.35 | 7.6% | worst to best |
| TownCentre (P) | excluded | | | excluded | | | worst to best |
| Landfill (dist) | 6.37E-06 | 17.46 | 0.6% | 7.29E-06 | 20.16 | 0.7% | per 1km |
| Students | 0.00014 | 11.27 | 1.4% | 0.000136 | 10.68 | 1.4% | per 100 students |
| Tourists | 2.05E-07 | 14.77 | 2.1% | 1.97E-07 | 14.55 | 2.0% | per 100,000 tourists p.a. |
| PM 2.5 | -0.00824 | -11.96 | -0.8% | -0.00784 | -12.13 | -0.8% | per µg/m3 |
| Crime | -0.00013 | -20.25 | -1.3% | -9.7E-05 | -15.32 | -1.0% | per 100 incidents |
| Road noise (55 to 59 dB) | -0.02642 | -11.95 | -2.6% | -0.02795 | -12.79 | -2.8% | relative to base |
| Road noise (60 to 64 dB) | -0.0393 | -11.46 | -3.9% | -0.04137 | -12.2 | -4.1% | relative to base |
| Road noise (65 to 69 dB) | -0.05037 | -8.53 | -5.0% | -0.05315 | -9.02 | -5.3% | relative to base |
| Road noise (70 to 75 dB) | -0.07425 | -5.2 | -7.4% | -0.0749 | -5.26 | -7.5% | relative to base |
| **Building & Plot** | (not reported) | | | (not reported) | | | |
| **Supply constraint** | (not reported) | | | (not reported) | | | |
| **Socio Economic** | | | | | | | |
| LSOA Average Income | 4.08E-05 | 164.68 | 4.08% | 0.000125 | 48.33 | 5.93% | per £1,000 p.a. |
| *LSOA Income interactions* | | | | | | | (units as above, at two extreme points of the income distribution) |
| Income * Rail Jobs (P) | | | | 8.65E-12 | 4.74 | 0.11% | at 10% percentile of Inc |
| | | | | | | 0.21% | at 90% percentile of Inc |
| Income * Car Jobs (P) | | | | 1.75E-12 | 1.41 | 0.18% | at 10% percentile of Inc |
| | | | | | | 0.20% | at 90% percentile of Inc |
| Inc * Primary rate (Avg) | | | | -5.5E-06 | -9.65 | -1.38% | at 10% percentile of Inc |
| | | | | | | -2.66% | at 90% percentile of Inc |
| Inc * Secondary rate (Avg) | | | | -5.1E-06 | -13.18 | -1.84% | at 10% percentile of Inc |
| | | | | | | -3.02% | at 90% percentile of Inc |
| Income * Bank (P) | | | | 2.08E-05 | 31.48 | -4.24% | at 10% percentile of Inc |
| | | | | | | 19.97% | at 90% percentile of Inc |
| Income * Income (P) | | | | -1.2E-09 | -43.37 | 7.28% | at 10% percentile of Inc |
| | | | | | | 4.52% | at 90% percentile of Inc |
| Not reported variables: | Building & Plot, Supply constraint, Time trend and constant (very similar to previous models) | | | | | | |
| **Adj R-squared** | 0.7264 | | | 0.7308 | | | |
| **Observations** | 160513 | | | 160513 | | | |

- Walk accessibility is slightly larger than in previous models (3.61% versus 3.15% in Model 14), presumably due to a better specification of the income and town centre effects through their respective income interaction terms.
- The premium on 'town centre' accessibility measured using the Bank proxy varies greatly with the income of the area as expected, but the estimates are within a more reasonable range than in previous Model 12a.
- The income effect is modelled to be non-linear through a squared term on income which is also highly significant – this is an improvement relative to Model 12a and shows that the area income effect on property prices is diminishing, which is consistent with economic theory and expectations.
- This new model – thanks to the more refined specification of income – allows us to also observe a significant and expected income-varying premium for the quality of schools (both primary and secondary).
- Finally, it must be reiterated that parks and playgrounds are interacted – as in all previous models reported – with a dummy which takes value 1 if the area has an LSOA income above average. Thus, the estimates show a positive effect which is only applicable to those areas with above-average income. As mentioned in the discussion of the base model (Model 12), this was a pragmatic solution and shows that the impact of parks and playgrounds on house prices is not uniform across areas, and is likely to only be significant and positive in relatively better-off areas. This test was repeated under the model specifications chosen for Models 14 and 14a, with the same result: no positive premium for parks and playgrounds in relatively worse-off areas.

As a concluding remark, based on all the tests conducted, there are reasons to believe that the final model reported, Model 14a is superior to previous models and thus preferable for future policy tests. Nevertheless, it does not fundamentally change the findings on the impact of rail accessibility on property prices relative to the SOBC Model, rather it refines them, and shows how the uplift is likely to vary with income – this is a consequence of the inclusion of income interaction terms in the model. Table 4.7 summarises the findings.

Table 4.7: Impact of rail accessibility to jobs on house prices (point estimate and range)

| | Model 12<br>(Post-SOBC base model) | | | Model 14a<br>(Improved model) | | |
|---|---|---|---|---|---|---|
| | Estimate | t-ratio | %Δhouse price | Estimate | t-ratio | %Δhouse price |
| Rail accessibility | 1.37E-07 | 15.42 | 0.14% | 1.56E-07 | 17.93 | 0.16% |
| | | | | | | |
| *Income interactions:*<br>Income*Rail accessibility | | | | 8.65E-12 | 4.74 | 0.11% (10th %ile)<br>0.21% (90th %ile) |

### *4.1.7 Unused variables: Bus accessibility; Rail noise; NOx; Access to other amenities*

The base model presented in Table 4.1 (Model 12), and the additional models discussed in the preceding section (Section 4.1.6), do not use all of the variables gathered. In getting to this point, a handful of variables were found to be insignificant or wrong sign in all the alternative model specifications tested. These unused variables are discussed here, in order to give further context to the model selection and highlight some additional lessons learnt in the study.

#### *4.1.7.1 Bus accessibility*

Due to data limitations, it was not possible to compute a 'bus accessibility to jobs' metric analogous to the accessibility to jobs variables by rail, car and walk. A simpler approach to bus accessibility was tested instead, using the measured distance from a property to a bus stop as a proxy. The main expectation (from theory) was that greater distance would be associated with a lower property price; however, in theory it is possible that close proximity to bus stops may be associated with negative externalities as well (e.g. due to traffic or noise issues). Previous models (e.g. in London) have found (net) negative effects of being located closer to a bus stop.

Including the simple variable 'distance to bus stops' in the model revealed a significant impact on property prices: prices decreased the closer a property is to a bus stop. Alternative models were tested to explore this further, with mixed results. Firstly, dummy variables for being located within 400m and within 600m of a bus stop were tested, both revealing a negative impact on prices. Secondly, a non-linear specification was tested, including the squared term of the distance variable: the results were interesting and suggested that the negative effect of bus stop proximity decreases with distance and, beyond 1.0km, being closer to a bus stop actually has a positive effect – and this effect becomes large, e.g. +2% per 100m closer at 2km, suggesting there may be a negative value on remoteness. Finally, a 1km dummy was tested, but this was estimated with the wrong sign and only marginally significant (t-ratio=2.76).

Our interpretation is that there is some evidence that remoteness from a bus stop is negative for property values; this may apply mainly to rural areas. However the number of properties beyond 1km from a bus stop in the dataset (only 1,174 out of 160,513) is a tiny proportion of the whole sample and the estimate is not robust to alternative specifications. More generally, beyond the negative value of extreme cases of remoteness, with the current measurement of bus distance we are unable to detect the positive value of access to jobs by bus. A more sophisticated metric which includes number of jobs reachable and travel times, connected by a deterrence function, (i.e. analogous to those we have for the other modes) would be worth exploring in future expansions of the model beyond this stage of the project. For now, the tests were not satisfactory for bus access and it was omitted from the selected models reported earlier in Section 4.

#### *4.1.7.2 Rail noise*

This is a set of categorical indicators similar to the road noise indicators that are included in the models reported above. At lower noise levels, parameter estimates were insignificant. The parameter estimate on the higher noise band was positive, contrary to expectation. The interpretation of this result may be that high average noise levels are indicative of being close to rail stations, which, as outlined in the Phase 1 report of this project, will likely have a price premium for properties. This variable was consequently dropped from the modelling.

#### 4.1.7.3 Air pollution: Nitrous oxide (NOx)

- NOx – The parameter estimate on the NOx variable was negative as expected when NOx was the only air quality indicator in the model (increasing NOx levels reduces house prices, all else equal). However when the PM 2.5 indicator was included, the NOx variable became positive, against expectations. This may be explained by the correlation between NOx levels and PM 2.5. There is also a potential issue with spatial fidelity of the air quality data at present, which is being investigated. As the parameter estimate on PM 2.5 was consistently more significant in model estimates, it has been left in the reported model above.

#### 4.1.7.4 Access to other amenities

Accessibility to some local amenities, in particular GP surgeries, health services, food stores and Post Offices, by walking were found to be insignificant or the wrong signs when estimated in the models. These variables have been tested multiple times throughout the project, including a new test at the very end with the final set of reported models; the results were again unsatisfactory – insignificant and/or wrong sign – and the variables were not used in any of the selected models reported in this document.

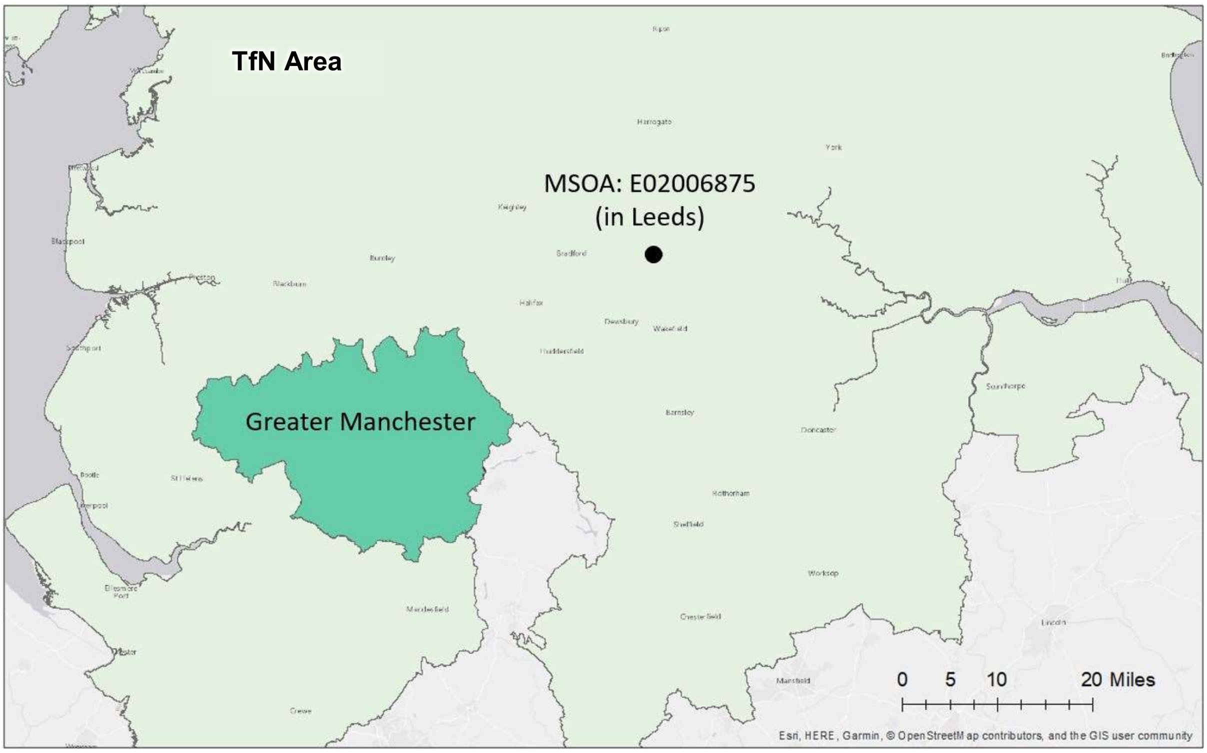
### 4.1.8 Model interpretation and application

Prior to applying the modelling results to the appraisal of Northern Powerhouse Rail (NPR), which is covered in Section 5 of this report, some preliminary work was done to show how the rail and walk accessibility terms in these hedonic pricing models can be interpreted, by considering a Do-Something scenario involving accessibility improvements, compared to a Do-Minimum scenario of no improvement.

The examples below are hypothetical, and the results are for a single MSOA in the centre of Leeds (E02006875) (see Figure 4.4 overleaf). In each scenario, the accessibility to employment variable (as defined in Sections 2.3 and 3.3.1) is recalculated and applied using the coefficient from the estimated hedonic model. All other variables are assumed to remain constant.

The hypothetical scenarios broadly cover two types of accessibility improvement. The first is a direct rail accessibility improvement, answering "how will reduced rail in-vehicle times to other areas impact upon property prices in the chosen origin MSOA"? Secondly, a scenario is tested where there is increased employment in the MSOA, and how that affects property prices through improved (local) walk accessibility. Obviously this latter scenario is not a direct impact of any specific transport intervention, but it may be an indirect impact of transport investment in the North or of spatial policy, for example.

Figure 4.4: Locations used in hypothetical scenarios




The scenarios tested are as follows:

- **Do-Minimum:** The employment potentiality value for E02006875 is calculated for walk and rail over the entire study area (including employment within a buffer of 50km outside the TfN area). This is the value that was calculated using the baseline NoRMS data and subsequently used as part of the hedonic model calibration.
- **Do-Something (1):** In this scenario, it is assumed that rail in-vehicle time between the origin MSOA and all MSOAs in Greater Manchester will reduce by 20 minutes.
- **Do-Something (2):** In this scenario it is assumed that rail in-vehicle time between the origin MSOA and all destination MSOAs in the whole study area (TfN area+50km) will reduce by 20 minutes, in cases where the Do-Minimum in-vehicle time is greater than 30 minutes.
- **Do-Something (3):** In this scenario, no rail changes are applied. Instead, the number of jobs within origin MSOA in Leeds City Centre increases by 10% on the baseline value (of 98,755). Other areas are unaffected. A new walk access potentiality measure is calculated for the Leeds City Centre origin and applied with the hedonic price model.

The results of these tests are shown in Table 4.8. Note that these results are for a single MSOA in Leeds City Centre. The property uplift is calculated using the following formula:

$$Uplift\ (\%) = \sum_m \beta_{EP,m}(EP_{DS,m} - EP_{DM,m}) \tag{31}$$

where

$\beta_{EP,m}$ is the parameter on employment potentiality (=accessibility to employment), for mode *m,* from Model 6 (the latest model at the time);

$EP_{DS,m}$ is the employment potentiality in the do-something scenario for mode *m ;*

$EP_{DM,m}$ is the employment potentiality in the do-minimum scenario for mode *m.*

Table 4.8: Results of initial scenario testing (MSOA: E02006875)

| | **Rail Accessibility (Employment Potentiality)** | **Walk Accessibility (Employment Potentiality)** | **Property Value uplift %** |
|---|---|---|---|
| Do-Minimum | 0.15319 | 0.01449 | - |
| Do-Something (1) | 0.15660 | 0.01449 | 0.26 |
| Do-Something (2) | 0.18420 | 0.01449 | 2.34 |
| Do-Something (3) | 0.15319 | 0.01547 | 4.39 |

The table above demonstrates the impact on residential property prices in the three Do-Something scenarios when compared to a Do-Minimum. In Scenario 1, a spatially limited improvement in rail in-vehicle time results in small property price uplifts being forecast. When the extent of the rail improvements is widened across the TfN area in Scenario 2, this uplift increases, reflecting the dispersed nature of employment locations across the North of England. Scenario 3 represents the case where new jobs appear in the origin zone. In this case, the local walk access to employment improves, and a property value uplift in the MSOA is observed.

For comparison, we also calculated the property value uplift using the rail accessibility parameter from an earlier model (Model 1). It can be seen in Table 4.9 that the specification of the accessibility terms in the model has a significant impact on the results. The work later on in the study shows that the rail accessibility parameter has become stable and this kind of sensitivity has – we believe – been left behind.

Table 4.9: Comparison of initial scenario testing (MSOA: E02006875) using Model 1 vs Model 6

| | **Rail Accessibility (Employment Potentiality)** | **Walk Accessibility (Employment Potentiality)** | **Property Value uplift %<br>in Model 1 (Model 6)** |
|---|---|---|---|
| Do-Minimum | 0.15319 | 0.01449 | - |
| Do-Something (1) | 0.15660 | 0.01449 | 0.55 (0.26) |
| Do-Something (2) | 0.18420 | 0.01449 | 5.01 (2.34) |
| Do-Something (3) | 0.15319 | 0.01547 | 4.39 |

### *4.1.9 Model comparison*

The research team was aware of previous models developed by Nationwide (2014/19) for their Transport Special publications, which included models for Greater Manchester. These were also cross-sectional models, although of a smaller study area, using a simpler 'distance to rail station' measure of rail accessibility, but well-regarded in the property sector.

In order to make comparisons with these Nationwide models, the ITS cross-sectional model of the North (Model 14a) was applied to the characteristics of two selected 'representative' stations in Greater Manchester, in order to derive a set of premiums by distance to the station. These premiums are affected by the degree of rail connectivity to jobs, hence by distance from the CBD in particular, and by income. Stations 1&2 in the Table represent: 1) a slightly more distant but slightly higher income are; and 2) a slightly closer but lower income area. The income differences are 9% above the Greater Manchester average for Station 1 and 9% below the Greater Manchester average for Station 2.

Table 4.10: House price premiums for rail accessibility, applied to commuter rail stations in Greater Manchester (and compared with Nationwide, 2014/19)

| *Commuter Rail station at distance from home* | *ITS (2019) Cross-Sectional Model* | | *For comparison:* | |
|---|---|---|---|---|
| | *Station 1 (Higher Income)* | *Station 2 (Lower Income)* | *Nationwide (2014)* | *Nationwide (2019)* |
| 250m | +8.5% | +7.1% | | |
| 500m | +6.5% | +5.3% | +4.6% | +7.8% |
| 1000m | +3.3% | +2.7% | +2.0% | +3.3% |
| 1500m | +1.1% | +0.9% | 0 {BASE} | 0 {BASE} |
| 2000m | 0 {BASE} | 0 {BASE} | | |

Note: Stations 1&2 apply the Cross-Sectional model results to two illustrative Greater Manchester stations (East Didsbury and Moston stations 9km/6.5km from central Manchester respectively); Nationwide (2014/19) give an *average* premium for Greater Manchester National Rail and Metrolink stations, also based on a cross-sectional model.

Overall the models are encouragingly consistent. The premiums at 500m, 1000m and 2500m from a station are similar in the latest models. There is a slightly different spatial pattern. Nationwide do not report uplifts at 250m: this may be because place quality variables (e.g. noise and crime) are included separately alongside rail accessibility in the ITS model, whereas the Nationwide model may be picking both up together in the area very near the station.

The ITS model also finds a small premium between 1500m and 2000m. Changes in rail accessibility (including specific changes in the number of Metrolink destinations and frequency) could help to explain the changes between 2014 and 2019.

## 4.2 Commercial Model results

This section presents the results of the hedonic price modelling in the commercial property sector. The additional theory for this sector was covered in Section 2.8, and the empirical research strategy in Section 3.8. This Report covers the initial model and several further models, gradually developing the approach.

### *4.2.1 Initial Model*

As a starting point, we hypothesised that the value of commercial property is driven in part by its accessibility for employees. This would be the mirror image of accessibility from a home at *i* to many jobs at locations *j.* It would be a many-to-one mapping instead of one-to-many.

We also hypothesised that the value of commercial property is driven in part by its accessibility to other businesses. This is a reflection of *effective density* – as in agglomeration*:* the literature shows that there is a positive relationship between effective density and productivity (e.g. Graham, 2007) and we might expect productivity helps drive the achievable rent at a particular location.

Other variables sought for the model include:

- Floor space, sq ft;
- Type: office/retail/industrial;
- Quality of the space – particularly for office space – described by a star rating on a 1..5 scale, see below;
- Other variables used in the Residential Model, e.g. income and certain place quality variables.

Two data sources have been pursued for the dependent variable, commercial property rents:

i. Valuation Office Agency (VOA) data – previously provided to IFS and requested TfN for use in the present context;
ii. COSTAR data – a privately-provided data source, used by WYCA and shared with ITS under the licence terms.

The COSTAR data has been the first to come through: it is available in limited quantities so we have focused on a smaller Case Study area initially – in this case Leeds (Bradford is added later – Section 4.2.3). The data relates to leases completed in the year 2016, of which there are 294 useable observations in the database, and covers: addresses; dates; use type; the star rating for office buildings; asking and achieved rents; and other costs such as service charges and business rates.

The Initial Model is very simple. The dependent variable used is: LN(Annual Rent per sq ft in 2016). Note that this is asking rent, as the data on rent achieved is incomplete. The first Accessibility variable tested was simple walk distance (via the street network) to the rail station, and shorter distances were found to be associated with higher rents. This was then replaced with the proposed Accessibility variable. Access to economic mass (i.e. jobs), or 'effective density', was calculated for the rail and walk modes, using a function very similar to the Accessibility variables in the Residential model. A decay parameter of -1 was used, as recommended by the DfT TAG Guidance. As perhaps might be expected, it was not possible to include both modes in the model due to correlation (although we are looking into a possible way to address this); in the model below the rail accessibility variable is used.

In the Initial Model (Table 4.11 below), significant variables are:

- Effective density, which is positive (as expected) and significant – underlying this variable is an assumption that the TAG -1.0 decay parameter applies;
- Type – 'NA STAR' indicates Industrial properties, which would be expected to be cheaper (so this variable is correct sign and significant) – for comparison Ibanez and Pennington-Cross (2013) indicate 77% cheaper than office rents[26], the same as the coefficient here;
- Quality 'star rating' with 3 star as the base – as expected, higher rated properties (4 and 5 star) are more expensive, 1 and 2 star are cheaper;
- Total area (sq ft) leased is negative and significant in this model i.e. larger premises are cheaper *per sq ft* (this variable is not significant in a linear version of the model);
- Income (at LSOA), which is positive and significant.

Table 4.11: Initial Commercial model results

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | 1.796 | .156 | | 11.547 | .000 |
| | Income | 1.373E-5 | .000 | .127 | 4.029 | .000 |
| | NA_Star | -.768 | .044 | -.654 | -17.362 | .000 |
| | @1_2_Star | -.312 | .051 | -.210 | -6.097 | .000 |
| | @4_5_Star | .556 | .067 | .293 | 8.283 | .000 |
| | Effective_Density_Rail | 6.136E-6 | .000 | .121 | 3.347 | .001 |
| | TotalSFLeased | -6.497E-6 | .000 | -.101 | -3.045 | .003 |

a. Dependent Variable: LN_AskingRentSFYr

The adjusted $R^2$ of the model is 0.72, indicating a fairly good fit to the data. There is certainly scope to refine the measurement of accessibility and treatment of modes – and to address accessibility from households which is not yet included, as well as other specific variables which theory/literature would suggest are relevant, some of which can be transferred from the Residential model.

### *4.2.2 Further Models*

Further work on the commercial model focused on developing the theoretical basis (Section 2.8) and the definition of the accessibility indicators(s) (see Section 3.8). Several aspects differentiate this line of work from the residential modelling, mainly:

- The context – due to current data availability – is not the whole of TfN, but instead only Leeds (Bradford data was added later – Section 4.2.3), i.e. the variability of the data is only at the local level and not regional level.

[26] although in a US context

- Only 294 observations are available, as opposed to over 160,000 for the cross-sectional residential model and over 30,000 for the time-series model in Greater Manchester.

The modelling therefore has limitations due to the small sample within a limited geographical area.

Following the theoretical work, a series of accessibility variables were developed, to reflect the different sources of 'value premium' that commercial properties may have. These cover:

- Walk access to other jobs (i.e. the value of locating near other firms & employees at their workplace);
- Walk access to population (i.e. the value of locating near workers and consumers at their home location);
- Rail access to other jobs (i.e. the value of having good rail access to other firms);
- Rail access to population (i.e. the value of having good rail access for workers and consumers)

In practice, however, there are a number of problems with this set of accessibility indicators which force us to make more limited selections of variables to include in the model:

- We have to decide what deterrence function(s) to use for 'Access to other jobs' (see Figure 3.10). We have discussed above (Section 3.8) whether the agglomeration-based functions are too steep (alpha = -1.8 or -1.0); if they are then we need to find an alternative.
- If we re-use the deterrence functions for 'Access to population' (shown as TTW in Figure 3.10) then we get a systematic relationship between 'Access to other jobs' and 'Access to population' for a given mode (e.g. rail), hence correlation, and therefore only one can be used for modelling purposes.
- Since rail is used extensively for commuting, an obvious choice is to use the 'rail access to population' deterrence function (Rail TTW in Figure 3.10). For walk, however, we expect that the stronger commercial property driver is the location near to other firms (i.e. walk access to jobs) – thus this is the selected variable for Walk access. Note, however, that models with all possible combinations have been tested, without providing grounds for altering these choices.
- Therefore, Walk access to jobs and Rail access to population are pre-selected as preferred variables.
- A further complication arises from the fact that these two variables are also highly correlated – see Figure 4.5. The reason for this is that it is common for many firms to desire a location near other firms but also near a rail station. Therefore, the outcome is that firms locate both near each other and near a station. This is observed in our sample (Leeds – and later Bradford) – the focus on only two cities limits the variability in the dataset in this respect.
- A solution to this problem has been to try instead using a simpler metric for rail access, independent of the number of jobs or population: distance to a train station.

The following tables (Tables 4.12 & 4.13) summarise the results of two models, which both include the Walk Access to jobs indicator and a metric of distance to the nearest rail station: Model C1 includes distance; and Model C2 includes a dummy for being located within 1,500m of a station. The reason for the latter is, again, a non-negligible correlation between Walk access to jobs and Distance to station – see Figure 4.6 – but also a low representation of properties very close to stations in the data (only 13 properties within 500m). Both C1 and C2 use the 'Walk TTW' deterrence function, though as noted above there is a case for using the agglomeration-based functions instead, and further testing is needed. The dependent variable in all models below is the log of the asking price variable, for which the dataset was complete.

Figure 4.5: Relationship between walk access to jobs and rail access to population (Leeds)

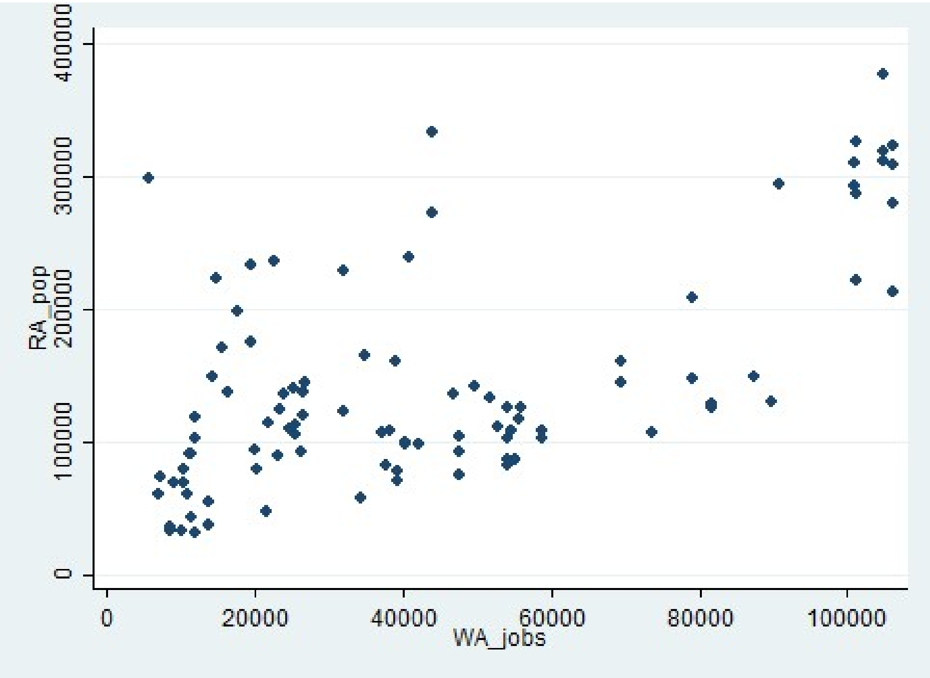



Figure 4.6: Relationship between distance to rail stations and walk access to jobs (Leeds)

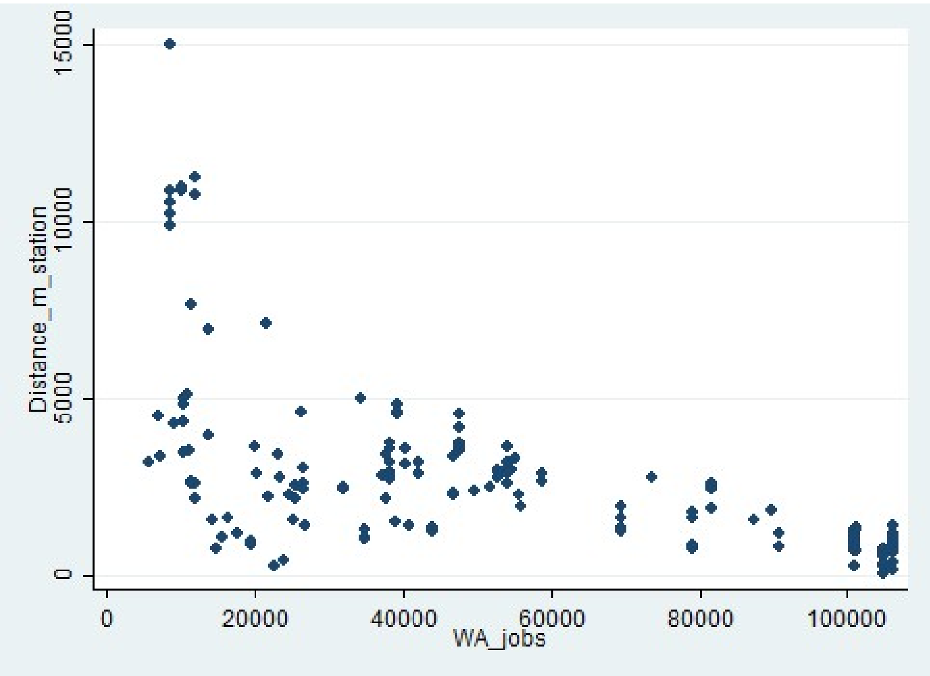


Table 4.12: Commercial results – Model C1 (Leeds)

| Variable | Units | Est. | t-ratio | %Δ Rent | |
|---|---|---|---|---|---|
| **Accessibility metrics** | | | | | |
| Walk Jobs (P) | N of jobs (discounted by GJT) | 1.72E-06 | 2.74 | 1.72% | per 10k jobs |
| Distance to rail station | Metres to nearest rail station | -1.38E-05 | -1.9 | -0.69% | per 500m |
| **Building characteristics** | | | | | |
| Dummy for Industrial | equal to 1 if industrial property | -0.74343 | -16.16 | -74% | relative to office space |
| Low quality space | equal to 1 if quality = 1, 2 stars | -0.29278 | -5.35 | -29% | if stars = 1 or 2 (relative to 3) |
| High quality space | equal to 1 if quality = 4, 5 stars | 0.544281 | 10.84 | 54% | if stars = 4 or 5 (relative to 3) |
| Floorspace | Squared foot leased | -6.5E-06 | -3.29 | -0.65% | per 1,000 sq. foot |
| **Place Quality** | | | | | |
| LSOA Average Income | £ (serves as a proxy for area quality) | 1.75E-05 | 5.12 | | per £1,000 p.a. |
| Constant | | 1.991044 | 17.95 | | |
| R-squared | | 0.7304 | | | |
| Observations | | 294 | | | |

Table 4.13: Commercial results – Model C2 (Leeds)

| Variable | Units | Est. | t-ratio | %Δ Rent | |
|---|---|---|---|---|---|
| **Accessibility metrics** | | | | | |
| Walk Jobs (P) | N of jobs (discounted by GJT) | 1.41E-06 | 1.84 | 1.41% | per 10k jobs |
| Within 1,500m of station | equal to 1 if within 1,500m of station | 9.70E-02 | 1.8 | 9.7% | if within 1,500m of station |
| **Building characteristics** | | | | | |
| Dummy for Industrial | equal to 1 if industrial property | -0.75847 | -17.16 | -76% | relative to office space |
| Low quality space | equal to 1 if quality = 1, 2 stars | -0.28411 | -5.17 | -28% | if stars = 1 or 2 (relative to 3) |
| High quality space | equal to 1 if quality = 4, 5 stars | 0.530117 | 10.99 | 53% | if stars = 4 or 5 (relative to 3) |
| Floorspace | Squared foot leased | -5.9E-06 | -3.05 | -0.59% | per 1,000 sq. foot |
| **Place Quality** | | | | | |
| LSOA Average Income | £ (serves as a proxy for area quality) | 1.36E-05 | 4.1 | | per £1,000 p.a. |
| Constant | | 2.040861 | 17.97 | | |
| R-squared | | 0.7312 | | | |
| Observations | | 294 | | | |

The main observations from the models reported above, as well as from all other model tests undertaken, are as follows:

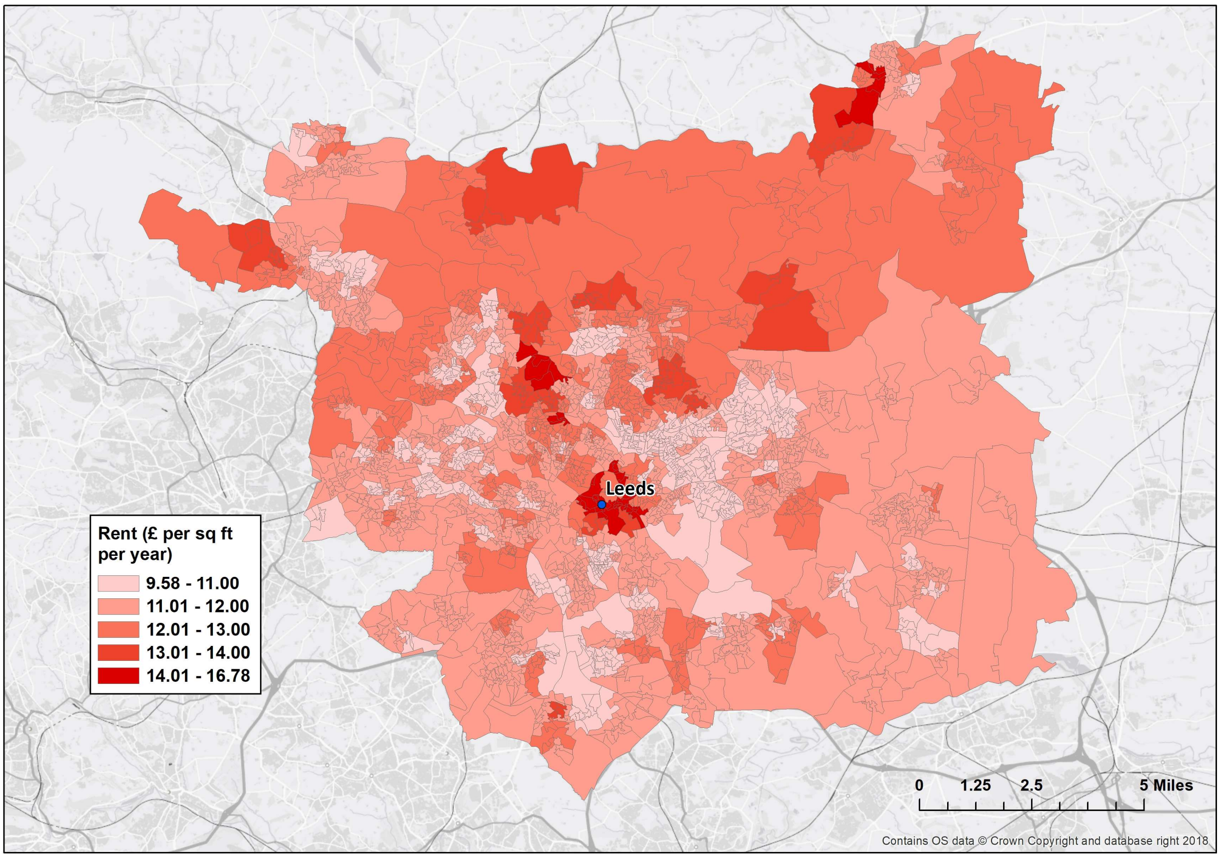
- There is evidence of a value premium for 'good' accessibility in the commercial sector, and this is associated with both being located near other firms (walk access) and near a train station (related to access to jobs and population by rail), however, so far the models are not able to fully unpack these in a robust way;
- The premium for accessibility to jobs appears to be approximately 1.4–1.7% per 10,000 accessible jobs (after applying the deterrence function) when using the TTW Walk deterrence function, or 6.1% per 10,000 accessible jobs when using the alpha= –1 decay parameter – this difference in the results seems consistent with the difference in the deterrence functions in Figure 3.10;
- The premium for commercial property located near a rail station appears to be in the region of 10% (within 1,500m of the station) – this should be treated with caution given the t-ratio of this variable is 1.8, and because most of the sample will be located within the catchment of Leeds Station, which is relatively well connected to other economic centres (stations with less extensive rail service may bring less of a premium);
- The estimates are from a model at a local level and need to be interpreted very carefully, thus, their direct use for policy in the context of the TfN area is not recommended at this stage;
- All other estimates in the model are fairly consistent across model specifications, suggesting:
    - higher premium for office spaces (relative to industrial);
    - the quality of the space drives property prices significantly;
    - leasing larger spaces is associated with reduction in the price per sq foot, as expected;
    - income in the area, interpreted as a proxy for the place quality of that area, has the expected positive effect on commercial prices.

The analysis of commercial property prices is still considered to be work in progress and the following avenues for improvement have been identified:

- More observations, both locally and at the regional level (from other LAs), should help to improve the analysis and statistical inferences;
- There is scope to improve existing knowledge on deterrence functions in the context of business trips – this would enhance the options to unpack the more varied sources of 'accessibility value' identified for the commercial sector (relative to the residential market) in the theory section;
- The current data does not include retail spaces – existing literature suggests that the patterns of accessibility value can be different for retail vs. offices; for completeness, having retail spaces in the data would be desirable.

Finally, the following map (Figure 4.7) shows the pattern of rents implied by the model, for mid-quality (3 star) office space in Leeds. Bear in mind that this is based on 2016 data, and that there is roughly a 35% discount for 3 star quality, compared with 4 or 5 star quality, according to the models estimated to date. The map highlights the rent gradient between central areas and most non-central locations. The influence of factors including income in driving rents outside the centre is noteworthy. As above, there is scope to improve the representation of accessibility (including rail accessibility) in this model, and to extend the coverage to the wider TfN area.

Figure 4.7: Modelled commercial office rents (Leeds)




### *4.2.3 Models including Bradford data*

With the addition of 77 observations for Bradford (in 2016), the COSTAR dataset increased to 371 useable observations. A series of models were run, and the results are summarised in Table 4.14.

All the parameters in these models appear to be the correct sign. Adjusted $R^2$ measures are encouraging for such a small model (sample N=371). The first two models use the Effective density formulation used in agglomeration analysis, but applied to jobs accessible via rail only. We find this variable is significant: despite the very steep decay function used, rail accessibility is significant.

Next we check whether the difference in accessibility between Leeds Station (many more services) and Bradford stations is significant. This Leeds dummy variable is insignificant: although initially surprising, this does suggest the rest of the model is adequately capturing the reasons for the differences Leeds vs Bradford – i.e. the Effective Density (Rail only) variable serves to explain the difference. The Leeds dummy variable continues to be insignificant through the remaining models.

Table 4.14: Commercial models – summary (Leeds and Bradford)

| *Accessibility & Agglomeration variables* | *Additional variables:* | *Coefficient, B* | *t ratio* | *Model fit, Adjusted $R^2$* |
|---|---|---|---|---|
| Effective Density (Rail only) | | 6.663E-06 | 3.582 | 0.691 |
| Effective Density (Rail only) | | 6.534E-06 | 3.497 | 0.691 |
| | Leeds dummy variable | 0.059 | 0.824 | |
| Walk accessibility to jobs | | 1.740E-06 | 2.683 | 0.703 |
| Distance to nearest station | | -1.758E-05 | -2.113 | |
| Walk accessibility to jobs | | 1.737E-06 | 2.635 | 0.702 |
| Distance to nearest station | | -1.757E-05 | -2.107 | |
| | Leeds dummy variable | 0.002 | 0.026 | |
| Walk accessibility to jobs | | 1.432E-06 | 1.999 | 0.704 |
| Station within 1500m | | 0.116 | 2.269 | |
| Walk accessibility to jobs | | 1.390E-06 | 1.890 | 0.703 |
| Station within 1500m | | 0.117 | 2.279 | |
| | Leeds dummy variable | 0.019 | 0.264 | |

The alternative formulations in the Table aim to bring in *both* accessibility to jobs within walking distance in the same city *and* the more distant rail-connected jobs in Manchester, York, Huddersfied, etc (and whichever of Bradford or Leeds the property is not located in);

- $R^2$ increases slightly – suggesting that having both walk accessibility and rail accessibility in the model is worthwhile;
- Both rail variables are very simple in these models;
- 'Station within 1500m' performs marginally better than 'Distance to Station' but the t-ratios are similar (magnitudes).

To round off this set of tests, both Rail and Walk were tested using the more sophisticated functions – i.e. Effective Density for Rail, and Walk accessibility to jobs – in the same model. As expected, parameters for both cannot be estimated concurrently, because of the correlation that comes from the 'walk to rail station' part of the rail journey (and which is highly weighted in GJT).

In conclusion:

- there is evidence here that commercial property values are sensitive to 'Accessibility to employment' – both by walk (locally) and by rail;
- we would really want a larger dataset to explore this further, and will focus attention on that going forward.

## 4.3 Time Series Residential Model results

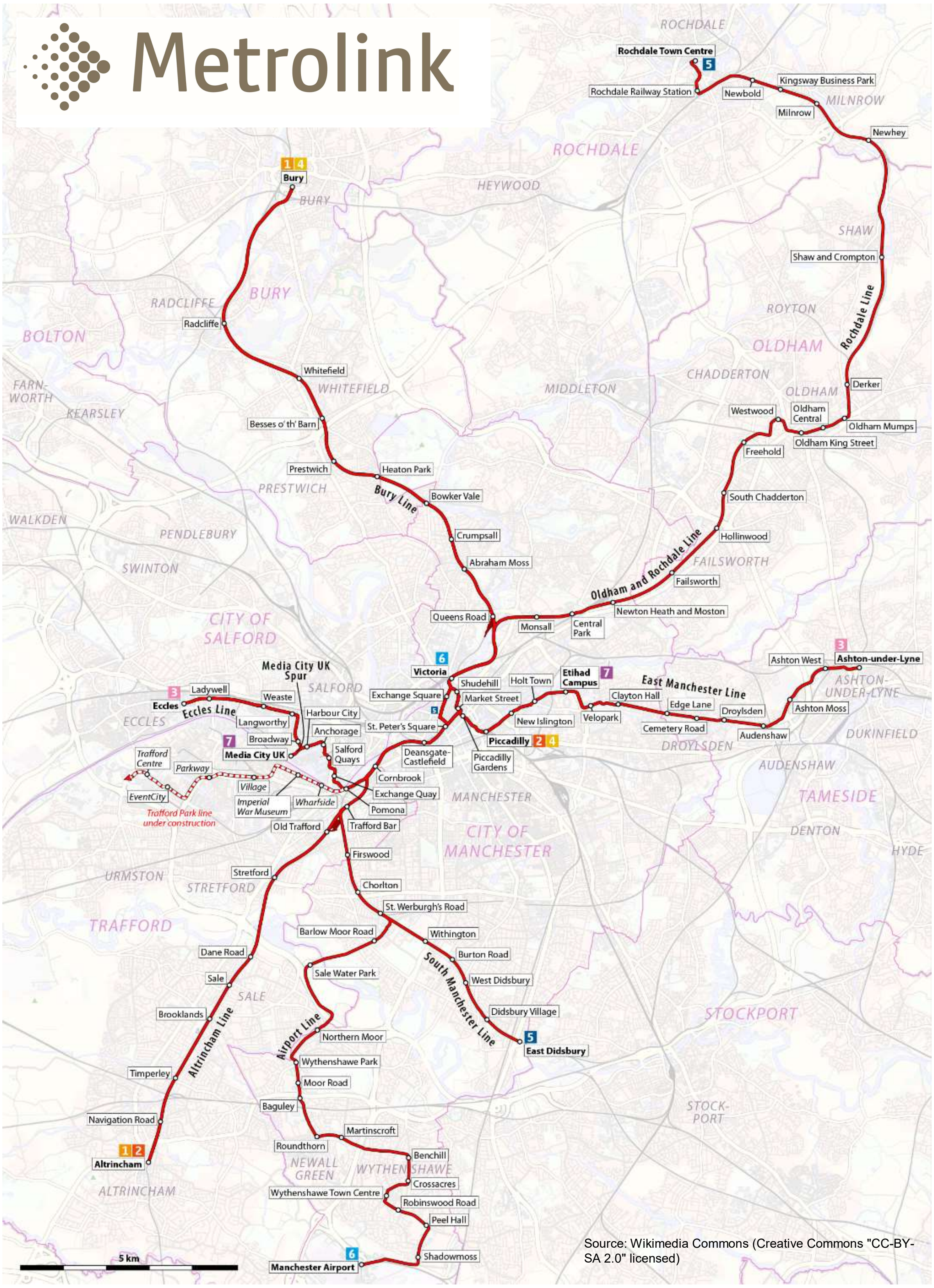
### *4.3.1 Definition of the 'treatment'*

Manchester Metrolink was chosen as the case study, and the geographic scope of the analysis is the Greater Manchester (GM) area. Metrolink currently has a network of 57 miles and 93 stops spread over seven lines radiating from Manchester City Centre (Bury, Rochdale, East Manchester, South Manchester, Airport, Altrincham, Eccles) (Figure 4.8). The 'treatment' in the model is based on the opening of these 93 stops at specific points in time. The Phasing of the Metrolink schemes has been:

- Phase 1: 20 tram stops linking Bury to Altrincham via on-street running in the city centre. Opened 1992.
- Phase 2: Line to Eccles. Opened 1999
- Phase 3a: This involved development of 3 lines: conversion of the Oldham loop heavy rail line from Victoria to Rochdale via Oldham completed in 2013; a South Manchester Line from Trafford Bar to Chorlton (opening 2011) and the East Manchester Line from Piccadilly to Droylsden opening 2013.
- Phase 3b: This involved further extension of the East Manchester Line to Ashton under Lyne and of the South Manchester Line to Didsbury and the Airport Line to Manchester Airport stemming off St Werburgh's Road. Phase 3b was completed in November 2014.

The Exchange Square tram stop, opened in 2017 as part of the Second City Crossing, was considered as part of the Rochdale line for the purposes of modelling.

Figure 4.8: Manchester Metrolink Network (2018) – mapping



Source: Wikimedia Commons (Creative Commons "CC-BY-SA 2.0" licensed)

### *4.3.2 Method*

Using the opening of these Metrolink stops to define the 'treatment', we look at two approaches to identifying a property price premium from being in close proximity to a Metrolink station.

1. A *comparative approach*. Here we compare the overall property uplifts for properties within 1km of stations along the different lines for the relevant operation periods, with the property price growth across the whole of the GM area for the same time periods.
2. *Fixed effects approach*. A panel dataset aggregated to LSOA level is used to estimate a fixed effects regression model, aiming to explain property price variations over time. As well as controlling for the property mix for transacted property in each LSOA we use a measure of 'treatment' as an explanatory variable. Here we employed two measures of 'treatment':
   I. An inner zone which is where LSOA centroids become less than 1km from the nearest Metrolink station (where they were more than 1km away in 1995).
   II. An outer zone which is where LSOA centroids are now between 1-2km from the nearest Metrolink station (where they were more than 2km away in 1995)

   The fixed effects estimation approach allows the model to control for unobserved time invariant heterogeneity in residential property prices at the LSOA level. Annual dummies are used to control for 'macro' level property price fluctuations. Average property characteristics are also controlled for.

Estimation in the fixed effects models is based on the following specification:

$$lnP_{it}=f(A_{it},B_{it},e_i) \qquad (32)$$

where:

- $lnP_{it}$ is the average price of properties in $LSOA_i$ at time $t$ ;
- $A_{it}$ is the (accessibility) treatment measure dummy for $LSOA_i$ at time $t$ ;
- $B_{it}$ captures the average building characteristics of properties in $LSOA_i$ at time $t$ ;
- $e_i$ captures the unobserved time invariant component of property prices in $LSOA_i$ .

Both the fixed effects and comparative approaches are weighted by the total number of property transactions in each LSOA.

The treatment measures were applied globally to all lines and then broken down separately as follows:

- Eccles line
- Airport Line
- East Manchester Line
- Rochdale Line
- South Manchester Line

It was impossible to identify the impact of Phase 1 in the fixed effects model as the time series commenced after its opening. Each LSOA within 1km of a Phase 1 station will be associated with a dummy $A_i$ which will be 1 for all years in the dataset – perfectly correlated with, and thus indistinguishable from, the fixed effect constant $e_i$ for that LSOA.

It is important to note that the uplifts between the two methods are not strictly comparable:

- The fixed effects uplifts are a treatment effect – i.e. differences in property prices we would expect to see in an otherwise identical LSOA area with and without a Metrolink station within 1km catchment. In other words these are the property uplifts attributable to the presence of a nearby Metrolink station.
- The comparative uplifts are based on the cumulative proportional uplifts observed on property prices for LSOAs within 1km of a station on a particular line over the period of operation as against the cumulative uplifts on properties observed over the whole of the Greater Manchester area for the same time period. These do not control for differences in types of properties over time or other time invariant factors which may have led to differential uplifts in areas.

### *4.3.3 Data*

To construct our panel dataset we first identified our dependent variable as sold prices on completed house purchases between 1995 and 2018, averaged (mean) at the LSOA level in Greater Manchester. Observations on 37 of the 1673 LSOAs were dropped where there were more than one missing value for a particular year. Where there was just one missing value we used the average value either side of that year. This left us with 24*1636=39,264 observations. Matching this structure we established a number of explanatory variables based on the mix of transacted property types (flat, semi, terrace or detached) in each LSOA and year, the 'treatment' measure and annual dummies to control for underlying property price changes in the GM area.

### *4.3.4 Results*

*Comparative Analysis*

Based on properties in LSOAs where we observe a change in proximity of the nearest Metrolink station to the LSOA centroid, we calculate two sets of effects. Firstly, through the comparative approach we see the property uplifts between the treatment areas and the overall GM area reported in Table 4.15, alongside sample size and income levels along each line.

Table 4.15: Average annual property value uplifts within 1 and 2km catchments of new Metrolink stations compared with overall GM area

| Line | Pha-se | Year open | 1km catchment area | | 2km catchment area | | Average Annual % Uplift across GM area during operation period to 2018 | Average equivalised income 2015-16, £ |
|---|---|---|---|---|---|---|---|---|
| | | | Average % Annual Uplift during oper'n period to 2018 | LSOAs in catch-ment | Average % Annual Uplift during operation period to 2018 | LSOAs in catch-ment | | |
| Altrincham | 1 | 1992 | 6.16 | 27 | 6.05 | 90 | 5.34 | 31,220 |
| Bury | 1 | 1992 | 5.62 | 33 | 5.42 | 81 | 5.34 | 24,310 |
| City | 1 | 1992 | 4.47 | 22 | 5.25 | 44 | 5.34 | 30,038 |
| Eccles | 2 | 1999 | 8.58 | 12 | 8.45 | 34 | 6.54 | 24,512 |
| Airport (Firswood- St. Werburgh's Rd) | 3a | 2011 | 5.37 | 14 | 5.76 | 27 | 3.94 | 28,244 |
| Airport (Withington onwards) | 3b | 2014 | 7.11 | 28 | 6.38 | 53 | 5.35 | 25,332 |
| East Manchester | 3a+b | 2013 | 4.85 | 28 | 6.30 | 63 | 5.39 | 24,062 |
| South Manchester | 3a+b | 2013 | 7.37 | 18 | 7.68 | 42 | 5.39 | 31,120 |
| Rochdale | 3a | 2012 | 2.39 | 56 | 3.35 | 161 | 4.58 | 22,400 |

The table shows the average property price growth in LSOA zones within 1km and 2km of new Metrolink stations for each line, compared to the overall GM trend during the operating period of the line.

- LSOAs within catchment of Phase 1 lines have annual growth in line with the overall GM average, although the Altrincham line appears slightly higher (just over 6% for both catchments compared to GM average of just over 5%)
- LSOAs within catchment of the Eccles line also have property price increases above the GM trend for this time period.
- Both Phases of the Airport line are associated with property uplifts, particularly on the Phase 3a development around Chorlton with property increases over 5% per year compared to the underlying trend of 4% in GM for the time period 2011-2018.
- The East Manchester line associated uplifts are similar to the GM average but interestingly growth is 1% higher per year compared to GM once the catchment area is broadened to 2km.
- Properties along the South Manchester line grow in value around 2% per year above the GM average over both catchment sizes. This is the largest uplift in growth and it is worth noting that LSOAs within 2km catchment of this line have the highest average income from all areas examined.

- The annual uplifts observed along the Rochdale line were considerably below overall trend for the 2012-2018 period although they become larger at the 2km catchment than for the 1km catchment. Catchment areas around this line have the lowest average income level.

*Fixed Effects Modelling*

Using a fixed effects approach we observe a statistically significant positive uplift (T-statistic of just above 2) on residential property prices of on average £3,338. This average is over the 24 years, so is perhaps more appropriate to look at percentage uplifts. We find again a positive and statistically significant uplift of 6.3% from being in close proximity to a Metrolink station. These uplifts are on average sold prices in each LSOA having controlled for average house price increases across the GM area, house type mix, and unobserved time invariant LSOA level characteristics.

Extending the fixed effects approach we find the following uplifts on property prices changing in proximity to the new lines from Phase 2 onwards (Table 4.16).

Table 4.16: Property value uplifts within 1km of new Metrolink stations

| Line | Uplift within 1km of station | T-Ratio |
|---|---|---|
| Eccles | -3.4% | -0.56 |
| Airport | 20.6% | 9.31 |
| East Manchester | 7.5% | 1.59 |
| Rochdale | -1.1% | -0.54 |
| South Manchester | 10.5% | 4.12 |

Using the 2km zones in conjunction with the 1km yields the results shown in Table 4.17. The full regression results are reported in Table 4.18.

Table 4.17: Property value uplifts within 2km of new Metrolink stations

| Line | Uplift within 1km of station | T-Ratio | Uplift between 1km and 2km from station | T-Ratio |
|---|---|---|---|---|
| Eccles | -3. 5% | -0.58 | -2.5% | -1.13 |
| Airport | 20.9% | 9.32 | 9.4% | 2.99 |
| East Manchester | 7.5% | 1.61 | 5.7% | 1.66 |
| Rochdale | -1.1% | -0.54 | -4.9% | -3.28 |
| South Manchester | 10.6% | 4.16 | 7.4% | 3.95 |

The Airport and South Manchester lines generate the highest significant uplifts on property prices for LSOAs within 1km around the new stations of around 21% and 11% respectively. There appears to be a distance decay effect with lower but still significantly positive values for the 2km ring. The impacts for the other lines are insignificant (barring the Rochdale line 2km ring).

Table 4.18: Fixed Effects Model – Full regression results

| | 1km zone model | | 1 and 1-2 km zone model | |
|---|---|---|---|---|
| logprice | Coef. | T-Stat | Coef. | T-Stat |
| _cons | 11.414 | 775.29 | 11.414 | 778.35 |
| T96 | -0.002 | -0.42 | -0.002 | -0.41 |
| T97 | 0.045 | 10.22 | 0.045 | 10.22 |
| T98 | 0.092 | 17.41 | 0.092 | 17.42 |
| T99 | 0.137 | 22.87 | 0.138 | 22.88 |
| T100 | 0.223 | 34.68 | 0.224 | 34.71 |
| T01 | 0.331 | 45.48 | 0.332 | 45.44 |
| T02 | 0.493 | 66.90 | 0.494 | 66.78 |
| T03 | 0.716 | 107.69 | 0.717 | 107.40 |
| T04 | 0.945 | 169.80 | 0.945 | 169.05 |
| T05 | 1.067 | 214.29 | 1.068 | 212.90 |
| T06 | 1.142 | 219.97 | 1.142 | 218.90 |
| T07 | 1.217 | 228.33 | 1.218 | 227.48 |
| T08 | 1.189 | 196.07 | 1.190 | 195.63 |
| T09 | 1.074 | 178.39 | 1.074 | 177.64 |
| T10 | 1.089 | 182.69 | 1.089 | 181.72 |
| T11 | 1.059 | 176.32 | 1.059 | 175.52 |
| T12 | 1.056 | 173.08 | 1.056 | 173.38 |
| T13 | 1.060 | 174.05 | 1.063 | 172.07 |
| T14 | 1.108 | 164.45 | 1.108 | 161.60 |
| T15 | 1.154 | 180.00 | 1.153 | 177.31 |
| T16 | 1.222 | 184.34 | 1.221 | 183.27 |
| T17 | 1.280 | 180.45 | 1.280 | 178.21 |
| T18 | 1.295 | 166.50 | 1.294 | 165.20 |
| flat | -0.824 | -27.39 | -0.827 | -27.45 |
| semi | -0.664 | -40.96 | -0.663 | -40.96 |
| terraced | -0.968 | -47.08 | -0.968 | -47.34 |
| Eccles_1k | -0.035 | -0.56 | -0.035 | -0.58 |
| Airport_1k | 0.188 | 9.31 | 0.190 | 9.32 |
| East_1k | 0.072 | 1.59 | 0.073 | 1.61 |
| South_1k | 0.100 | 4.12 | 0.101 | 4.16 |
| Rochdale_1k | -0.011 | -0.54 | -0.011 | -0.54 |
| Eccles_2k | | | -0.025 | -1.13 |
| Airport_2k | | | 0.090 | 2.99 |
| East_2k | | | 0.055 | 1.66 |
| South_2k | | | 0.072 | 3.95 |
| Rochdale_2k | | | -0.053 | -3.28 |
| Overall R-sq. | overall | 0.7023 | | 0.7039 |
| | obs. | 39264 | | |
| | groups | 1636 | | |

Parts of the East Manchester, Rochdale and Eccles lines serve areas of higher deprivation and relatively low income. It is possible that low underlying levels of property demand constrained the growth in prices. The cross-sectional modelling would suggest that uplifts due to rail accessibility improvements would be lowest in areas with low LSOA-level average incomes (Table 4.7).

It is worth noting the Rochdale line replaced an existing rail line so potentially the step change in accessibility was not so great. The South Manchester and (to a lesser extent) Airport lines pass through more affluent areas in South Manchester. Although the Airport line passes through parts of Wythenshawe that are relatively deprived (measured by IMD), the Airport line is distinctive in having a major concentration of employment at each end (Manchester City Centre and the Airport itself), which in the cross-sectional model would increase 'accessibility to employment' from intermediate stations.

#### *4.3.5 GIS Mapping*

Here we aim to examine the spatial distribution of changes in property prices alongside the Metrolink network. Figure 4.9 shows the Metrolink stations by line (with Media City subsumed into the Eccles line).

Next we divide the panel dataset into two periods. In Figure 4.10 we see the changes in average prices per LSOA between 1995 and 2011 (excluding stations opening in 2011). This period covers the Phase 1 and Phase 2 lines. In Figure 4.11 we plot the distribution of changes in average prices per LSOA between 2011 and 2018. This period covers the Phase 3 lines.

A general observation is that there are many LSOAs away from Metrolink routes where there is evidence of large property price increases, so Metrolink is not a necessary condition for strong property price growth. Equally, there are some LSOAs near to new Metrolink stations that are showing relatively low price growth, so neither is Metrolink a sufficient condition for strong property price growth. Nevertheless, there are some general patterns that can be observed.

- From Figure 4.10, the Altrincham line going South and to a lesser extent the Bury line going North do appear to pass through multiple areas with higher than average house price growth (with a higher concentration of relatively darkly shaded LSOAs). Phase 2 to Eccles shows relatively few LSOAs with high price growth, however much of the regeneration around MediaCity (and the new tram stop there) came right at the end of the period shown by the Figure (1995-2011), so the price growth shown in the Figure may not reflect its full impact. Property development around Salford Quays increased the supply of housing units during this period, so relaxation of the supply constraint should have helped dampen price growth to some extent.
- Figure 4.11 does seem to show a pattern of darker shaded areas along the Airport and South Manchester lines corresponding to higher price changes, consistent with the comparative analysis and fixed effects modelling. There are some exceptions, and close inspection suggested that a few of these match with concentrations of low income and social housing, where the cross-sectional model would suggest the potential for uplift is reduced. Overall the concentrations of deep red along much of the length of these lines suggests that conditions for uplift are generally met.
- By contrast, in Figure 4.11 the Rochdale line shows many stops surrounded by light shading, indicating low price growth and even price decreases, 2011-2018. This accords with findings of the comparative analysis and fixed effects modelling, and checking the IMD map reveals that many of these are LSOAs in the top deciles for deprivation.

Figure 4.9: Metrolink Stations by Line

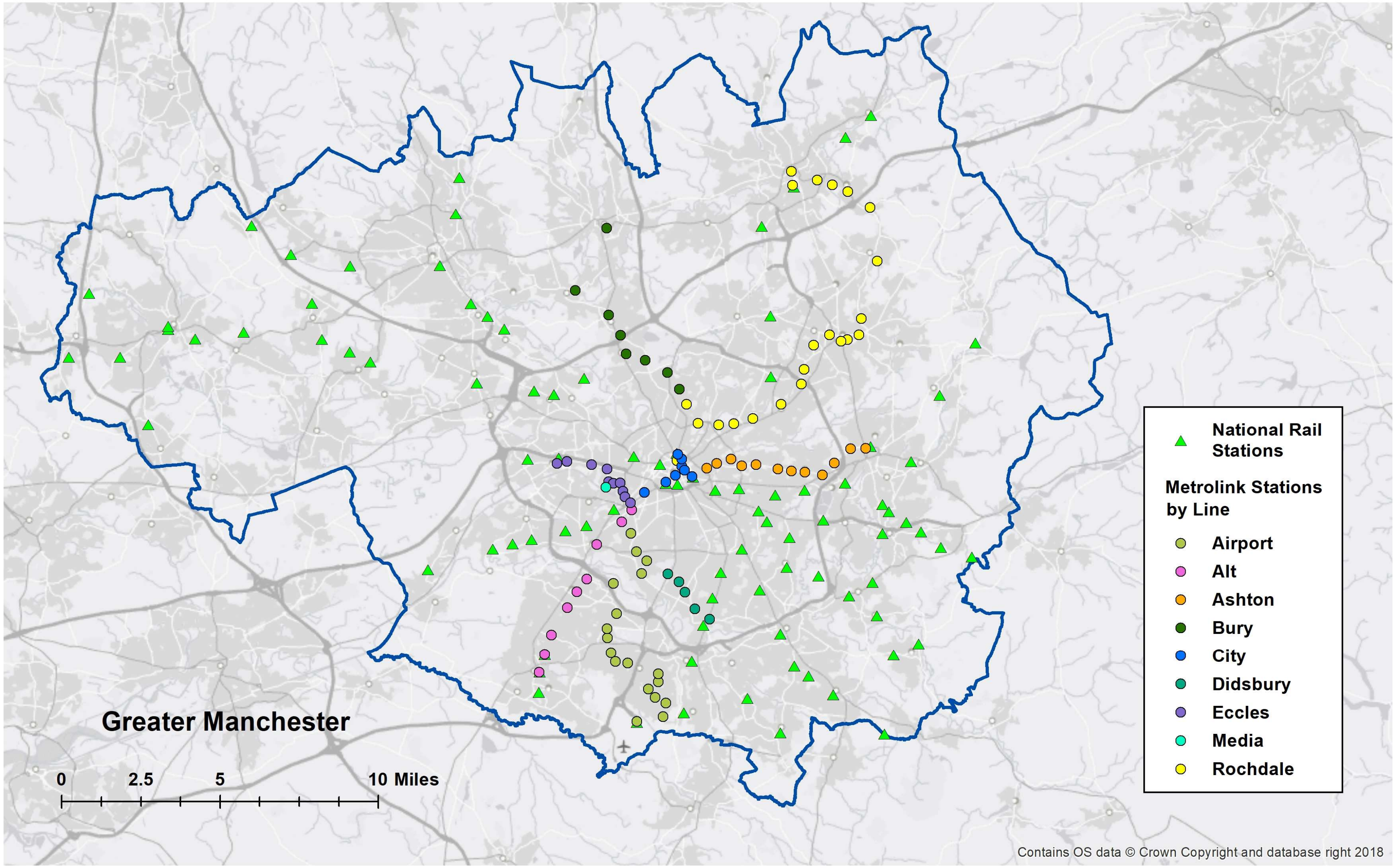


Figure 4.10: Spatial distribution of price changes 1995-2011

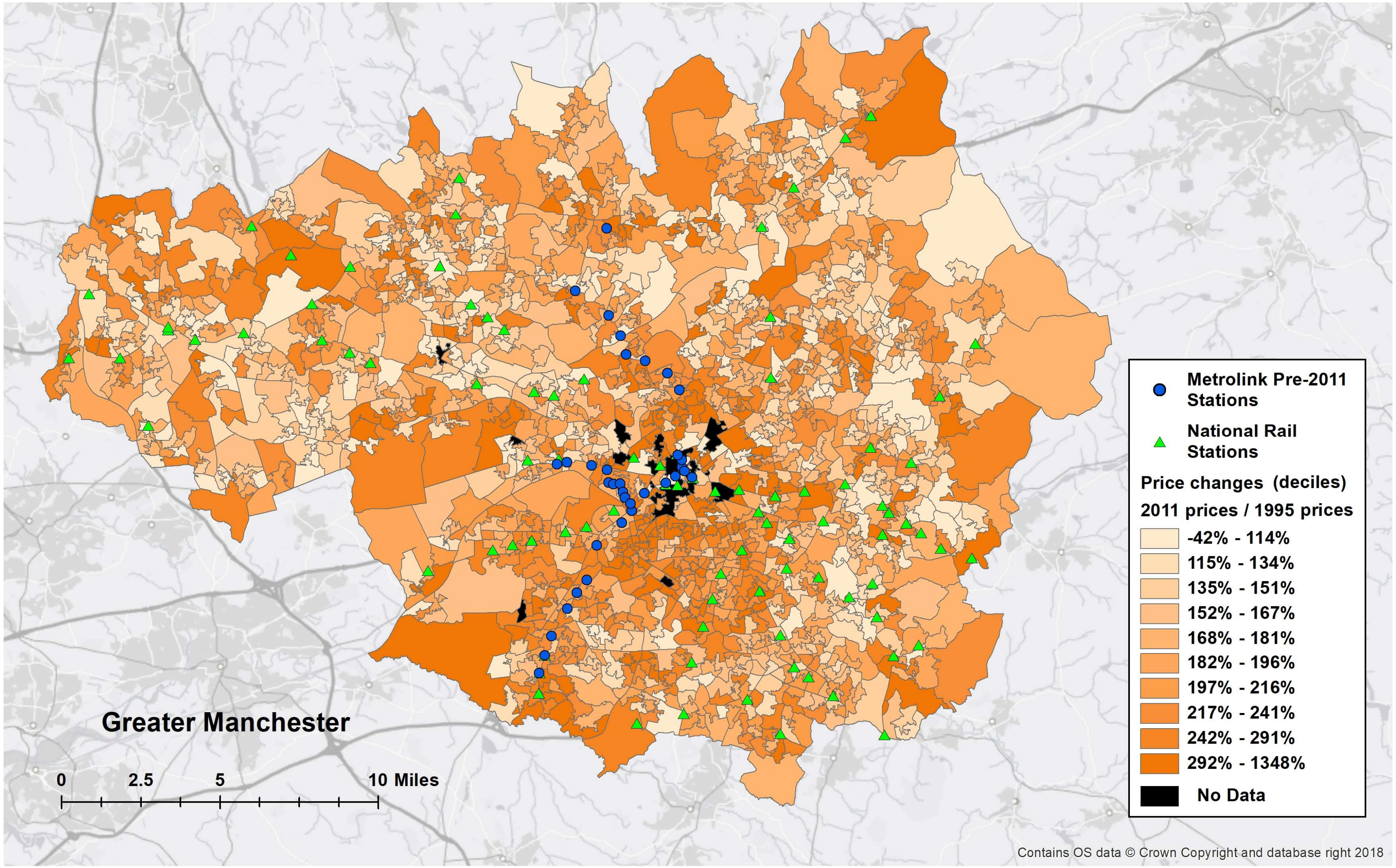
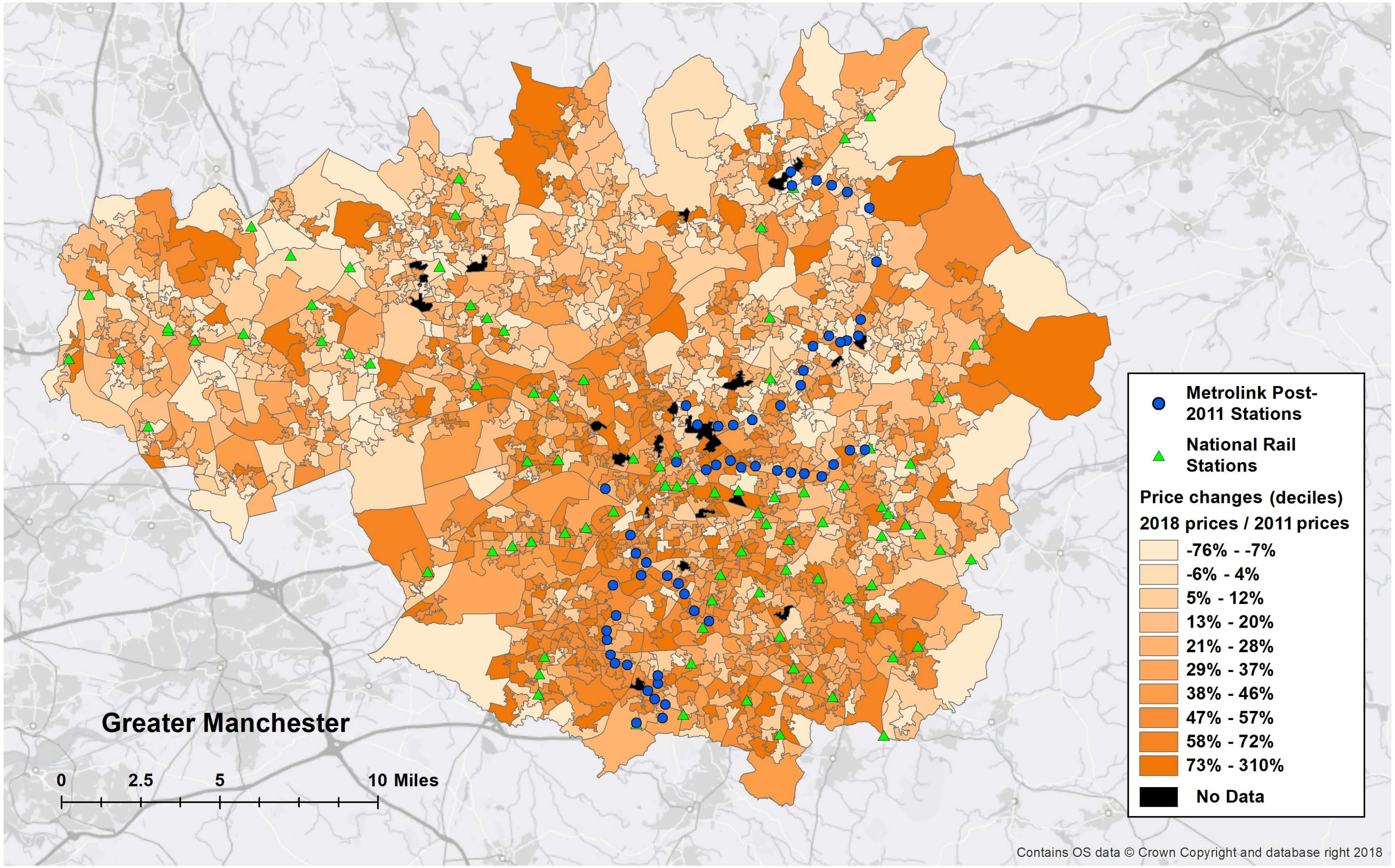


Figure 4.11: Spatial distribution of price changes 2011-2018



### *4.3.6 Summary*

Both the comparative analysis and fixed effects models indicate the highest uplifts and growth are found along the Airport and South Manchester lines and this was visible in the GIS mapping. There are also noticeable uplifts along the Phase 2 corridors in the comparative analysis but this is not reflected in the fixed effects modelling. This could be due to the small sample size of LSOAs affected by the Phase 2 line not yielding robust estimates. The more deprived areas of East Manchester and those on the Rochdale line seem to exhibit lower uplifts than elsewhere and growth in these areas could be constrained by the underlying property market and economic conditions.

# 5. Other Workstreams

This Section focuses on providing results from the application to Northern Powerhouse Rail (NPR) and further reflection on the study findings.

The evidence emerging from the modelling work in Section 4 goes beyond what is usually generated for a rail project Business Case. It is important to establish how it can inform the appraisal of a project such as NPR – or any other project that improves accessibility across a region – and how it relates to the other sources of evidence that are usually prepared for a Business Case. This study therefore included a workstream on 'Integration with Appraisal', which also touched on the potential for Land Value Capture in the North, the distributional impacts, and the implications for timing and phasing of investment.

This Section of the report begins by reflecting on the modelling results (5.1), then covers:

- Northern Powerhouse Rail (5.2);
- Implications for potential Land Value Capture (5.3);
- Distributional impacts (5.4); and
- Integration with appraisal (5.5).

### 5.1 Model interpretation

The results of the ***cross sectional residential model*** are best seen as an indication of the value of better accessibility to and from the home location. This is the value perceived by households, and expressed through willingness-to-pay (WTP) in the property market. It is – in a sense – the bedrock for all subsequent changes which 'market dynamics' bring about. The model controls for the effects of income, so the parameter estimates for accessibility always capture the value of access to more and better opportunities at a given initial income level. The results provide a starting point for estimating the total property value impact of accessibility changes. They are applied to Northern Powerhouse Rail (NPR) in the next section of this report (Section 5.2).

There may be further effects, although these are more difficult to predict using available models. Property market sorting behaviour (Tiebout, 1956; Kuminoff et al., 2013) may lead to higher-income individuals moving into the areas made more accessible. The models in this study do not predict this, but do indicate that if it occurred this could lead to substantial localised increases in house prices – due to not only the income interaction term but the pure income term in the model. For example, if the local average household income increased from the 5th to the 6$^{th}$ income decile, the price difference associated with this is +7.7% (implied by Model 14a[27]). It should be remembered that property market sorting activity involves households changing places, so in some other locations there would be a *downward effect* on household incomes and hence property prices (due to the income parameter).

Housing supply may respond to the localised increase in demand – subject to planning restrictions. If it does so, there may be some dilution of the price impact, however new development can also reduce the disamenity caused by derelict brownfield sites, and can rebalance the housing stock towards types that are locally scarce. As the cross-sectional model shows, the presence of new build housing units itself creates an uplift in average value,

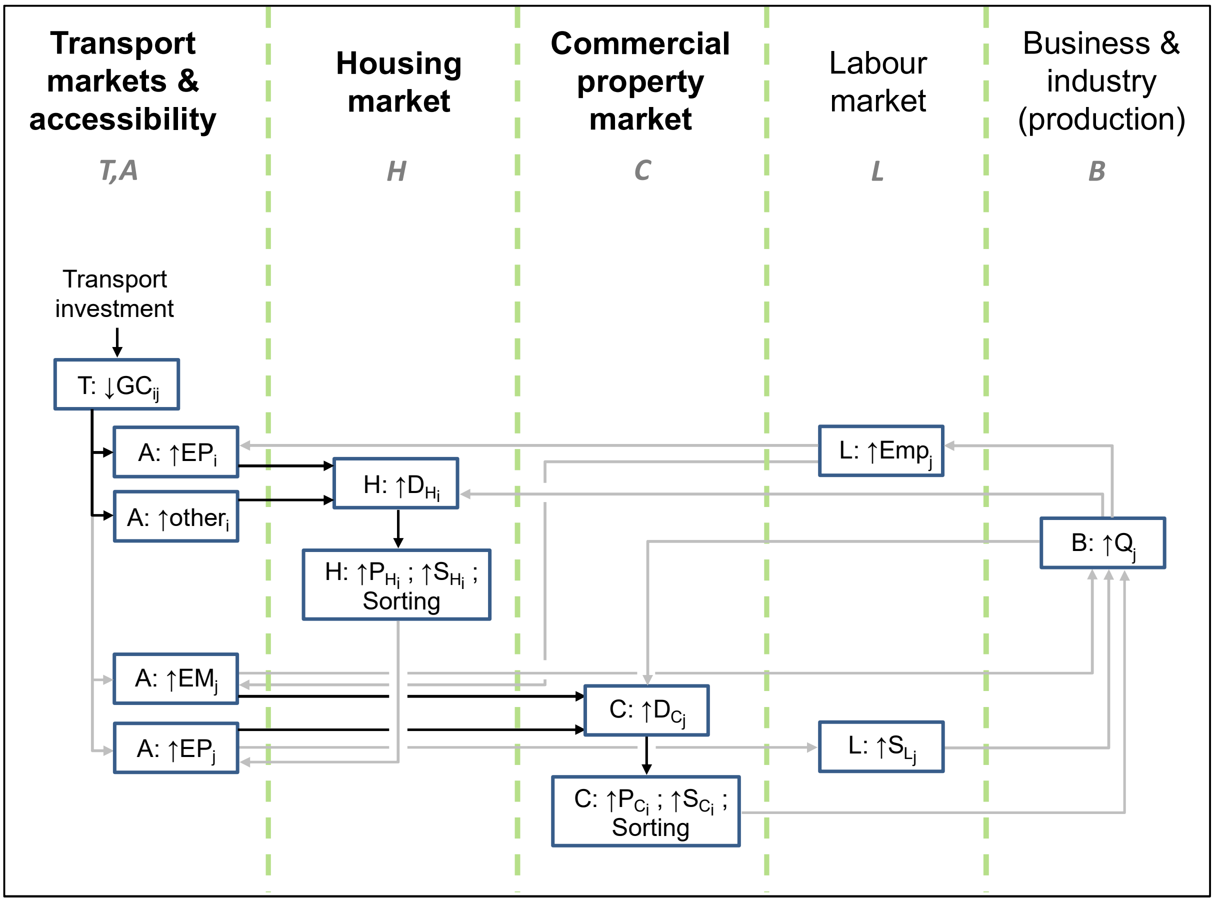
[27] £1,300 per annum increase in income (equivalised) * 5.93% per £1,000. See Table 4.6.

all else equal (since the value of each new housing unit is 18% greater than an equivalent existing unit).

There may also be effects transmitted through the wider economy. Labour supply improvements may lead to production increases and improvements in the public finances (direct tax revenue net of benefit payments). Agglomeration effects may increase productivity and output. These effects would likely feed back into the regional or sub-regional demand for housing: they are not modelled as part of this study, however they could be modelled by bringing together the models developed here with other model types.

These impact pathways are summarised in Figure 5.1. The initial impact through the rail accessibility premium is shown by the black arrows, and the further effects are shown by the grey arrows. It is worth recalling that even the initial rail accessibility effect on price is mediated by local area household income (see Figure 4.1) and area supply conditions (Section 4.1.4). For example, the model indicates that the initial price effect of a rail investment will be extremely small in local areas with household income below £20,000 (less than +0.05% per 10,000 additional jobs accessible). This may be due to a lack of job-skills matching with opportunities reachable by rail, and/or affordability of rail commuting relative to some other modes.

Figure 5.1: Linkages between transport, property markets and the wider economy




Note: GC = generalised cost of travel from home at i to work at j; D = demand; S = supply; EP = employment potentiality (accessibility to jobs); EM = economic mass.

Many of the further effects are likely to increase property prices near to a new transport service, over and above the initial uplift from the amenity value of the accessibility change.

However, there will be exceptions: e.g. the effects of increased housing supply; and sorting effects between different areas with an improved transport service (e.g. demand may gravitate to areas with both transport improvements and other attractive features).

Finally, there is the issue of hope value and the speculative aspect of the property market. To some extent property demand is driven by investment considerations (even when the purchaser is an owner occupier) and hence the future potential of the local market comes into play. Since future potential is by its nature uncertain, there is the possibility of inflows and outflows of capital to/from local markets based on changing expectations of future prospects and potential capital gains from investment now.

The ***quasi-experimental model*** for Manchester Metrolink (the Time Series model) is a different type of model in that it includes all the following responses – provided they have had time to occur between the stations opening and the end of the dataset in 2018:

- the initial uplift from the amenity value of improved accessibility (to employment and other opportunities);
- any property market sorting behaviour, and development of new housing units;
- any effects transmitted via the labour market or productivity impacts; and
- any speculative responses associated with hope value in the local market.

This allows for a comparison between the cross-sectional and quasi-experimental model results. First, the average uplift from new Metrolink stations, 6.3%, is well within the range for Best-to-Worst rail accessibility differences in the North, 14.3% (in the cross-sectional model) (Table 5.1). In other words, as expected there is a significant uplift but still potential for further value gain from rail accessibility improvements in these places, e.g. if they were – hypothetically – connected directly by regional express rail to many more jobs.

Table 5.1: Comparison of Metrolink uplift and Best-Worst rail accessibility premium (quasi-experimental vs cross-sectional models)

| | | | |
|---|---|---|---|
| Uplift: | | | **%** |
| | Metrolink Average | | 6.3 |
| | Airport Line | | 21 |
| | South Manchester Line | | 10 |
| | (East Manchester Line) | | (7.5) |
| | Others | | insig. |
| Premium: | | | |
| | Rail accessibility to jobs: Worst to Best (TfN area) | | 14.3 |
| | … at 90th percentile income | | 19 |
| | … at 10th percentile income | | 9.8 |
| | NPR increase in premium: | max. | 9.3* |
| | … based on Model 14a | | |

Note: estimation of NPR premium – see Section 5.2.

Secondly, the largest uplift for Metrolink slightly exceeds the range of the cross-sectional model (Airport Line), suggesting that some of the further effects in Figure 5.1 and discussed above, have come into play. Meanwhile the uplifts for some Metrolink lines, e.g. the Rochdale line, are not significant, suggesting that the lower incomes and possibly lower accessibility gains in this corridor (Metrolink replaced an existing rail service) can neutralise the Metrolink uplift in specific local circumstances.

Thirdly, work was undertaken to compare the rail accessibility premium at a South Manchester commuter rail station (in the cross-sectional model) with the Metrolink uplifts in the same area (South Manchester Line). Whilst this is not an exact comparison, because the form of rail accessibility is slightly different, the results are interesting (Table 5.2). They indicate again that the uplift can (locally) exceed the initial accessibility premium, as a result of these further dynamic effects. This is true in both the 0-1km and 1-2km buffer zones, although the difference is particularly pronounced in the 1-2km buffer around the station. This area has slightly above average income (compared with Greater Manchester or compared with the TfN area as a whole).

Table 5.2: Comparison of the rail accessibility premium and Metrolink uplift in South Manchester

<table>
<tr><th rowspan="2"></th><th rowspan="2">Premium (Cross-Sectional Model)</th><th colspan="2">Uplift (Quasi-Experimental Model)</th></tr>
<tr><th>0-1km</th><th>1-2km</th></tr>
<tr><td>250m from Commuter Rail station</td><td>+8.5%</td><td rowspan="3">+10.6%</td><td rowspan="3"></td></tr>
<tr><td>500m from Commuter Rail station</td><td>+6.5%</td></tr>
<tr><td>1000m from Commuter Rail station</td><td>+3.3%</td></tr>
<tr><td>1500m from Commuter Rail station</td><td>+1.0%</td><td rowspan="2"></td><td rowspan="2">+7.4%</td></tr>
<tr><td>2000m from Commuter Rail station</td><td>0 {BASE}</td></tr>
</table>

It is worth noting that the uplift compares the Metrolink 0-1km and 1-2km buffer zones with the rest of Greater Manchester. If higher-income households are attracted into the Metrolink buffer zone there may be some relative reduction in values in the 'rest of Greater Manchester'.

In summary, the cross-sectional model provides an initial estimate of the value of rail accessibility, seen through the eyes of existing residents. The outturn property uplift may be either higher or lower than the cross-sectional model suggests. There are a number of mechanisms through which further effects could occur – summarised in Figure 5.1 and discussed in this section. These include: property market sorting effects; the price impact of sorting (evident from the cross-sectional model); supply responses (for both residential and commercial property); and effects transmitted through the labour market, business and industry.

## 5.2 Northern Powerhouse Rail

### *5.2.1 Introduction*

Northern Powerhouse Rail (NPR) is a major strategic rail programme, designed to transform connectivity between the key economic centres of the North. The programme promises radical changes in service patterns and target journey times[28].

---

[28] https://transportforthenorth.com/northern-powerhouse-rail/

An initial Strategic Outline Business Case (SOBC) for NPR was finalised by TfN and DfT in early 2019 and endorsed by TfN partners in February 2019. Based on this SOBC, the Department for Transport (DfT) agreed that NPR should continue to be developed as a scheme through 2019-2020. The initial SOBC made a broad strategic and economic case for the NPR network overall but left a number of potential network concepts on the table.

As a complement to the transport and wider economic modelling carried out by TfN and DfT, the analysis presented in this section was included in the Strategic Case of the initial SOBC. This work was conducted in parallel with the main body of work on the NPR SOBC. The aims of the (upcoming) finalised SOBC are to narrow the option set down to a single network concept and optimise and refine the Strategic and Economic Cases. TfN is in the process of considering how further analysis of Land Value Uplift can form part of this case.

The evidence provided by this study in support of NPR is primarily: an understanding of the role of rail accessibility in property values in the North of England; and an assessment of the potential changes in values when NPR is implemented. The following text includes an overview of the modelling work, as well as the results from testing policy scenarios for NPR against a Do-Minimum for the year 2033.

### *5.2.2 Method*

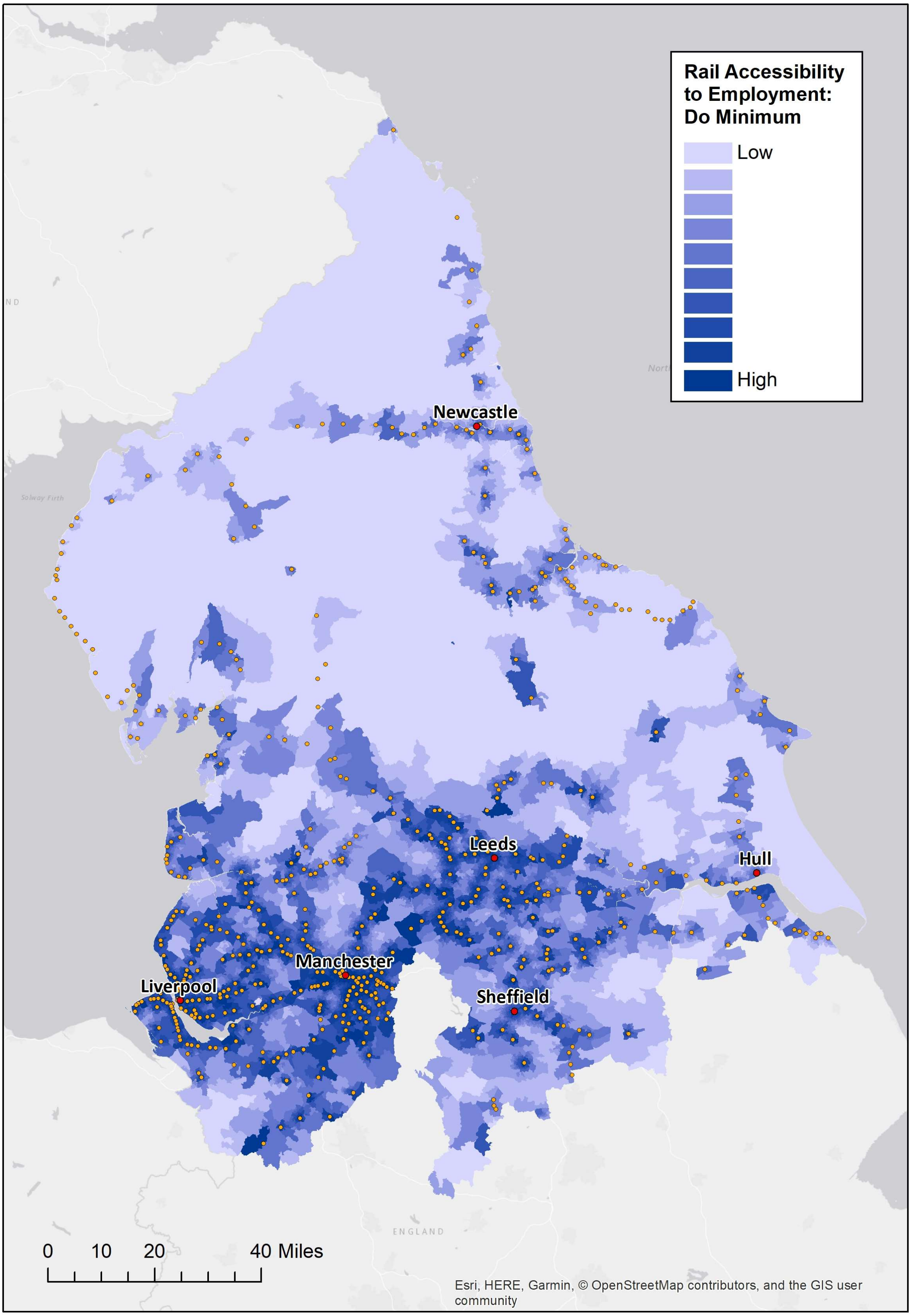
Modelling of the residential property market in the North of England, including the role of rail accessibility, has been undertaken using a set of purpose-built econometric models (Sections 2-4). Both cross-sectional and panel (time-series) models have been produced. These models are built using large datasets containing property characteristics, spatial economic characteristics including rail accessibility, and the Land Registry 'Price Paid' data as the dependent variable. The models combine geospatial analysis and econometrics, and are relatively sophisticated in their treatment of accessibility, using measures of generalised journey time (GJT) that are consistent with the NoRMS rail model used elsewhere in the NPR Strategic Outline Business Case.

This research has found that accessibility to economic opportunities – especially employment – is an important factor influencing property prices in the TfN area (Sections 4.1.1 and 4.1.6). The model takes account of multiple modes, but we begin by focusing on rail.

The number of jobs that a resident can 'see' via the rail network in the Do-Minimum scenario, varies depending on how easily they can access the rail system from their home location, and how well that station is connected to employment locations across the North. The worst rail-connected location in the North can only 'see' 2,994 jobs via rail – taking into account Census evidence on acceptable travel times to work. By contrast, the best rail-connected location can 'see' 917,948 jobs, or roughly 12% of the jobs available across the North.

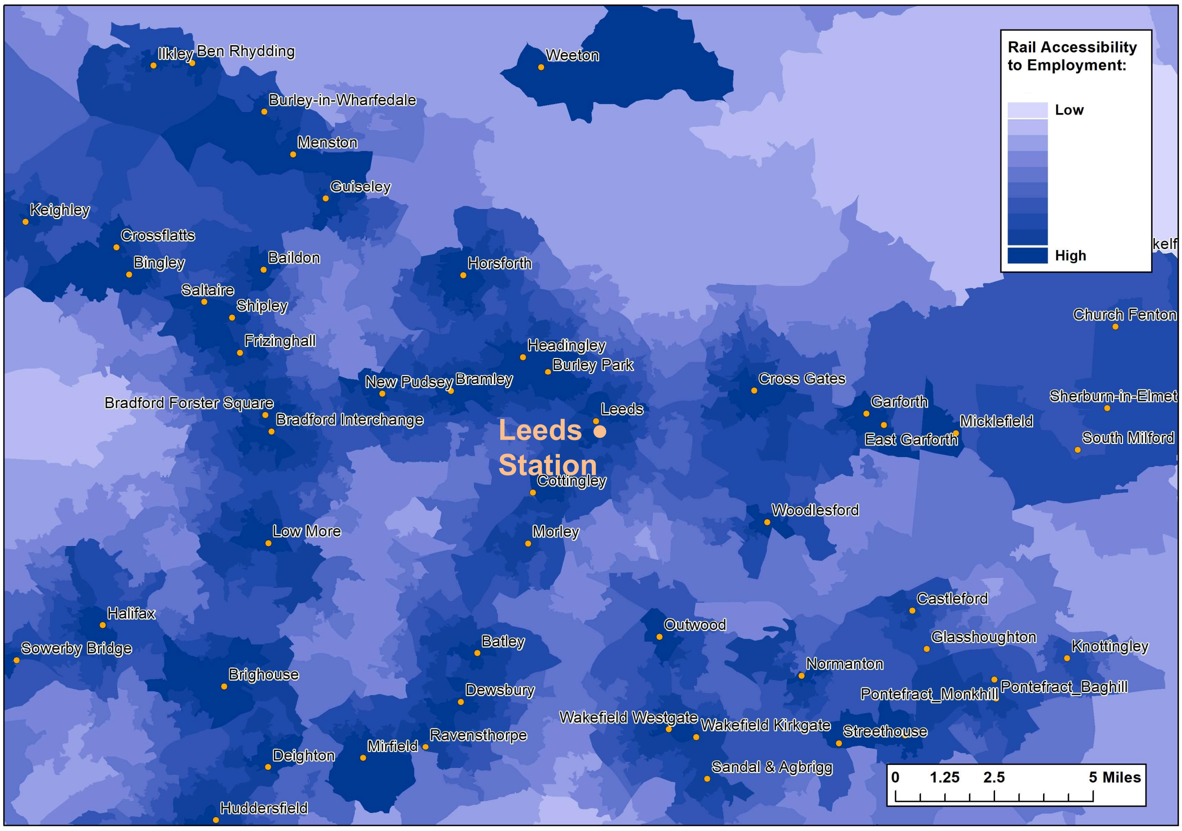
Figures 5.2 and 5.3 show this uneven distribution of rail accessibility to employment across the North. The first map shows the wider scale of the TfN area as a whole, whilst the second map focuses on an area of West Yorkshire 30 miles by 20 miles and in this case the stations are all labelled. Both maps show how the most rail accessible places to live (the darker blue areas) are clustered around the existing stations – but not all stations are equal: those with the best rail connectivity to the largest concentrations of employment are differentiated from stations with poor rail connectivity or a long distance from large numbers of jobs. For example, some of the stations on the Calderdale line between Manchester and Leeds score highly for rail accessibility to employment because they face towards both cities' labour markets and have rail journey times under an hour to both, with frequencies up to four trains per hour, whereas the Cumbrian Coast Line has journey times over 2 hours to Manchester and a basic service frequency of one train per hour.

Figure 5.2: TfN Area rail accessibility to employment, based on NoRMS, HSL and Census data (2033 Do-Minimum, stations shown)




Note: this scenario includes HS2 and other planned interventions up to 2033.

Figure 5.3: Rail accessibility to employment in an area of West Yorkshire (stations shown)




The models allow us to understand how residential property prices are affected by rail accessibility to employment. The current version of the cross-sectional model indicates a premium of 14.3% for properties in the most rail-accessible vs the least rail-accessible locations, controlling for all other factors. This is based on analysis of 2016 property price data, and the other matched datasets.

The parameters from the Cross-Sectional Residential model were used to investigate the potential change in the pattern of property prices across the TfN area, based on the following different scenarios:

- an NPR High Investment scenario;
- an NPR Medium Investment scenario;
- a Do-Minimum scenario.

Each scenario was defined by MSOA-to-MSOA generalised journey times (GJT) across the rail network, based on outputs from TfN's Northern Rail Modelling System (NoRMS) as the basis for the changes in rail accessibility. These GJT measures included in-vehicle time, crowding penalties, interchange penalties and a measure of delay time, and were supplemented with customised access and egress time estimates (Section 3.3) – a relatively comprehensive GJT metric.

The Rail accessibility to employment measure in the property value model responded to these changes in GJT. The pattern of Rail accessibility to employment in the Do-Minimum is shown in Figure ES1 – the NPR improvements are all relative to this.

### *5.2.3 Results*

Using cross-sectional Model 14a, improvements in rail accessibility to employment in the NPR High Investment scenario could potentially produce uplifts in residential property values of up to 9.3% for Output Areas (OAs). The potential increases in property price are greatest in areas close to NPR stations, and around other stations with strong connections into NPR (e.g. from well-served National Rail suburban stations, and from areas with high frequency bus services connecting to NPR stations) – these are locations which experience the greatest increase in rail connectivity to employment across the North. This does not take account of any subsequent changes in the locations of households or businesses or the locations of investment (e.g. in residential or commercial property development) in response to the changes in the pattern of accessibility created by NPR.

Figure 5.4 maps these potential uplifts for the High Investment scenario – since LSOAs are larger than OAs the effect is diluted slightly and the largest impact at LSOA level is 5.88%. Figure 5.5 maps the potential uplifts for the Medium Investment scenario – at LSOA level the largest potential uplift is 5.38% in this scenario.

Figure 5.4: Potential residential property price changes implied by the NPR High Investment scenario (LSOA level)

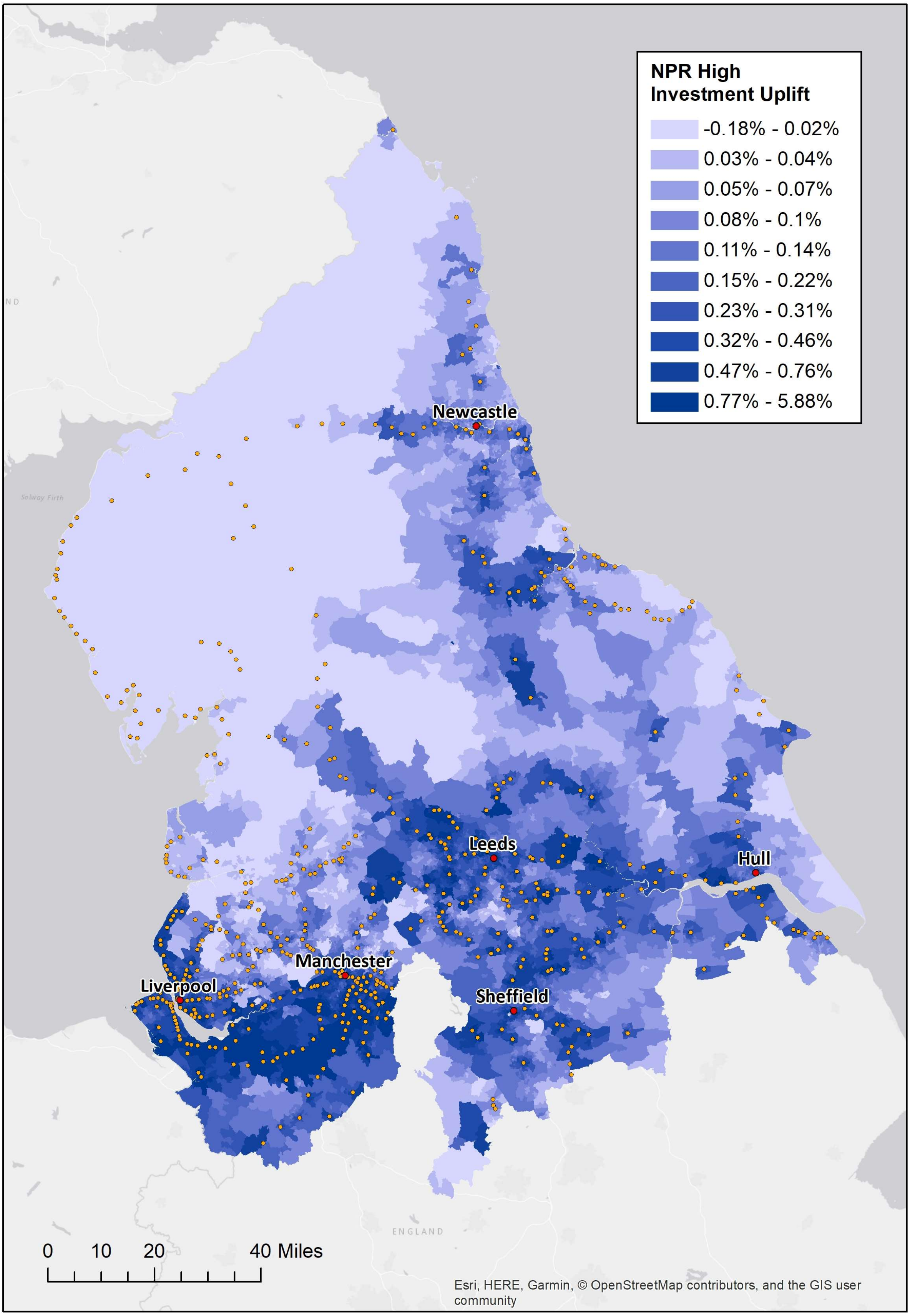
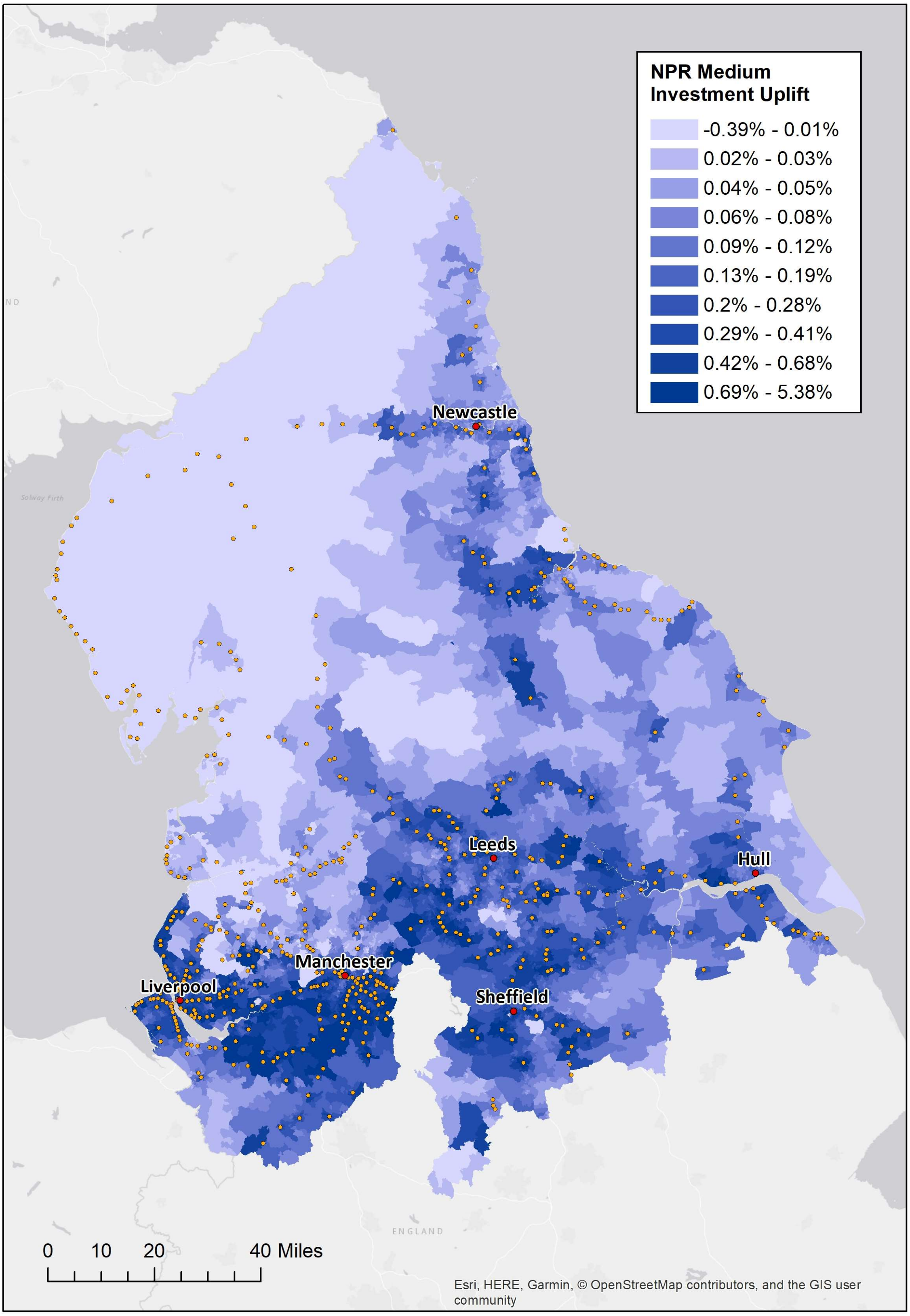


Figure 5.5: Residential property price changes implied by NPR Medium Investment scenario




Modelling has also been undertaken using a time series (panel data) modelling approach in a case study area (Section 4.3 above) – Greater Manchester was selected as it has seen significant improvements in its mass transit network over the period for which detailed property price data is available (1995 to 2018). Comparison with these results helps to get a sense of the robustness of the cross-sectional model. The time series model shows that for past changes in accessibility due to Metrolink network expansion, we observe a positive and statistically significant uplift of 6.3% from becoming in close proximity (1km) to a Metrolink station. For the Airport Line and South Manchester Line the uplifts are in the region of 20% and 10% respectively. These uplifts are on average sold prices in each LSOA, having controlled for average house price increases across the Greater Manchester area, property type mix, and unobserved time-invariant LSOA-level characteristics.

The average uplift found (6.3%) sits comfortably within the range of worst-to-best rail accessibility found using the cross-sectional model (12.5%), and is slightly less than the potential uplifts identified for NPR by Model 12. It is worth noting that:

i. Model 14a, the most recent cross-sectional model including income interaction terms, implies somewhat higher uplifts for NPR on average, and particularly higher uplifts in areas with higher average household income;

ii. Metrolink stations often, although not always, bring high frequency rail-type accessibility to locations where none was previously available, whereas NPR is in most locations a (major) improvement on already existing rail accessibility – this will affect the magnitude of the impacts;

iii. The South Manchester Line serves areas with higher-than average incomes.

iv. The Airport Line provides accessibility to both Manchester City Centre in one direction and Manchester Airport in the other direction, and in that respect is distinct from the other Metrolink lines.

An Excel-based tool was developed to measure the total value accumulated in the housing market as a result of a particular policy test. This Excel tool aggregates property value uplifts across the current and future dwelling stock of the TfN area, for use in the Business Case. Key evidence and assumptions underlying this included:

- Coefficients from the cross-sectional model (Model 12);
- Dwelling stock by LA area and tenure, 2017/8 (MHCLG data);
- Dwelling stock growth rates to 2051 and beyond (from NTEM);
- ONS House Price Statistics for Small Areas – at LA level (2017/18) – as a check on the values in the LVT model sample;
- Uplifts for social housing as well as privately owned stock, based on ratios from the literature;
- For NPR, the profile of uplift over 8 years from the opening year (inclusive) – based on comparisons with data on the time profile of uplift for other recent projects.

Table 5.3 shows the results for the total expected uplift or increase in value in the residential property market, showing how the results are sensitive to the NPR High Investment or Medium Investment scenarios. The results are presented in three ways:

- a snapshot for a single year assuming instantaneous change (based on 2017 property values);
- a cumulative estimate over 8 years from NPR opening in 2033/34, to 2040/1, at the current nominal prices at that time (assuming the time profile of land value uplift is at a steady compound rate over 8 years from opening, based on comparisons with other recent projects where TfL have gathered this data); and

- a Net Present Value (NPV) – using 2018/19 as the base year for discounting and 2010/11 as the price base year, as requested by TfN (otherwise the usual TAG price inflation and discounting rules apply).

Table 5.3: Total potential value uplift due to NPR rail accessibility improvements, TfN Area (Residential property)

| | NPR High Investment | NPR Medium Investment |
|---|---|---|
| Single year – instantaneous change (2017) | £2.9 billion | £2.5 billion |
| Cumulative over 8 years from opening in 2033/34-2040/1 | | |
| (at nominal prices) | £6.2 billion | £5.3 billion |
| (NPV) | £1.9 billion | £1.6 billion |

These results may be conservative in their estimation of potential uplifts, given the findings from the latest Cross-Sectional Modelling since the SOBC inputs were provided, which have identified:

- a slightly higher rail accessibility parameter ($1.56*10^{-7}$ compared with $1.37*10^{-7}$);
- positive income interaction which would increase uplift in areas with higher average household income (see Table 4.7 – Model 14a).

In principle, it would be possible to compare these total uplifts with the transport user benefits measured in the CBA, and with the forecast user benefits. As the modelling tools used for the initial SOBC are undergoing further development, the results were not considered sufficiently finalised to share with this project for publication. However, TfN and DfT do intend to make this comparison in future. For completeness, this comparison would ideally include an estimate of commercial property value uplift, which has not been included in the analysis presented here.

There are two points worth emphasising in summing-up the discussion on NPR:

- There are a number of potential further effects following the initial change in rail accessibility premiums at specific locations: for example, if NPR changes the spatial pattern of employment or the level of employment in city centres, this should lead to further land value gains (and this potentially has implications for housing and spatial planning in conjunction with NPR). We outlined these further effects in Section 5.1, and have explored them using one other simple policy test (below).
- Further improvements to the Land Value model *may* change the coefficients on Rail Accessibility, however we are monitoring this carefully. The most recent modelling work reported above (Section 4.1.6) has not changed the order of magnitude of the coefficient, and this stability is encouraging. Further, when an income interaction term is included in the model, we find that property prices are more sensitive to Rail Accessibility in higher income areas, and this may increase the uplift overall (see Table 4.5).

A further policy test was carried out separate from NPR, to understand the impact of increasing employment in economic centres. This could be due to, e.g., a change in urban planning policy to allow increased density, substantial redevelopment of brownfield/underutilised sites, or as a second-round effect of other policies which make central locations more attractive: such as investment in the rail network. The test modelled was a 10,000 increase in city centre employment in two central LSOAs in Leeds. This could be expected to produce a 1-3% increase in city centre residential property values, and up to 1% in the area outside that, to about a 12km radius in Leeds. Therefore if NPR changes the spatial pattern of employment or the level of employment in city centres, this should lead to further land value gains.

### 5.3 Implications for Potential Land Value Capture

The measured uplifts due to Manchester Metrolink and the potential uplifts due to NPR were discussed in Sections 5.1 and 5.2. This Section considers one potential use of these uplift estimates, which is to inform consideration of Land Value Capture (LVC) as a tool of infrastructure funding in the North of England. This Section is general in nature, intended to aid discussion rather than to support specific conclusions.

There has been a rise in interest in the use of LVC in the UK, and a corresponding rise in its actual use for transport projects. Crossrail and the proposed Crossrail 2 in London are probably the highest profile applications to date, although there are other interesting examples, e.g. Barking Riverside Overground Extension/Opportunity Area – also in London – and there is some relevant experience in the North. Before looking at applications, it is worth briefly introducing LVC and the mechanisms available.

The name *land value capture* suggests a focus on the underlying land rather than the assets built over that land – homes, commercial premises, etc[29]. It is worth noting that our models find a relationship with the value of the property as a whole. TfL (2017) cite a rule of thumb that land values make up one third of house prices[30]. The latest data from ONS show that since 1995, for the UK as a whole, this proportion has been higher: rising from 55% in 1995 to 72% in 2017. Therefore any LVC-type charge on residential property value is equivalent to a proportionate charge on land value (a factor of 100/72 = 1.39 on average).

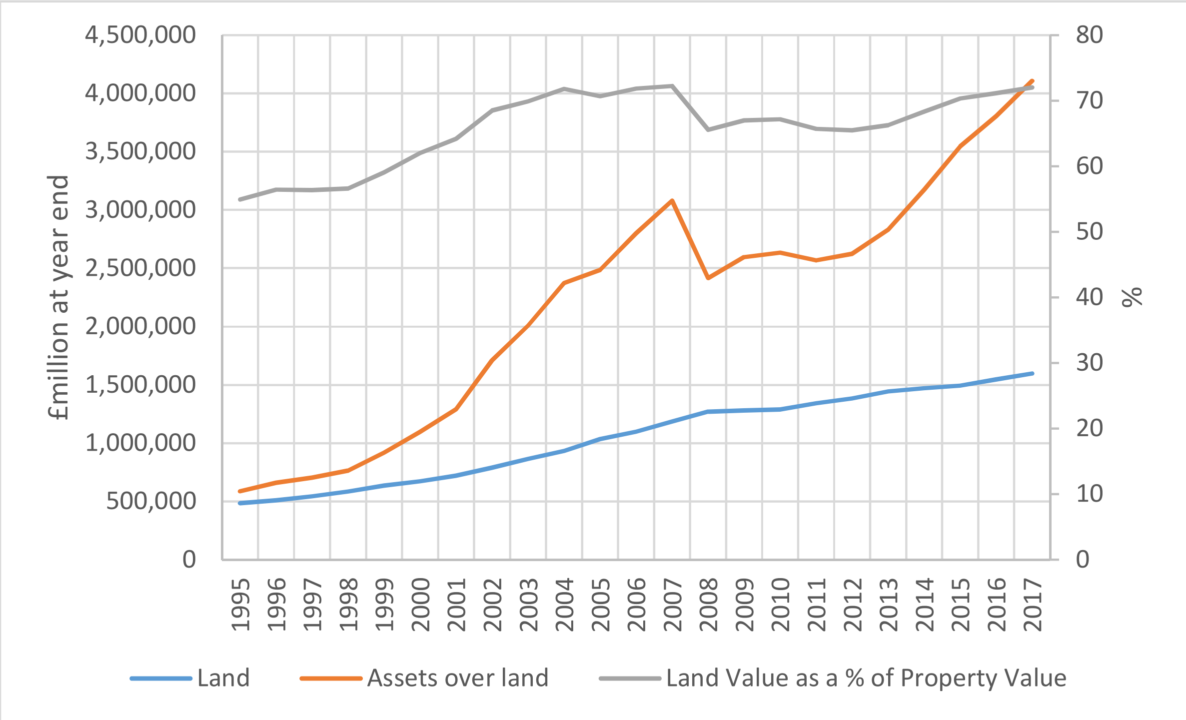
Data tends to be for property transactions rather than land only. Furthermore, the development status of land can change, particularly through:

- planning permissions to develop for housing, office or industrial use – permission can change the value of land by factors of more than one hundred, which puts the uplifts from transport infrastructure in perspective (see Table 5.4);
- intensification of development – e.g. by building higher or denser on the available land area;
- placemaking – where it is recognised that high quality places are not simply the most intensively developed, but those where the masterplan, the design and the mix of uses is most beneficial in a broader sense. E.g. in a densely-developed urban area, provision of a park could increase land and property values (the Cross-Sectional Residential Model gave an indication of the potential effect on property values – Section 4.1.2 and Summary).

[29] In their latest work, Salon et al. (2019) broaden the scope of LVC to include all types of '*location value capture'*, and include residents, businesses and property owners in scope. In effect, this is what the LVC methods in the UK already do.

[30] TfL (2017) p28, citing "Mayor of London (2011) Draft Charging Schedule for the MCIL".

Figure 5.6: Land value as a % of property value, for properties owned by households (UK)




Data source: ONS (2018). Note: 'assets over land' include dwellings (98% by value in 2017) and other buildings & structures over land.

The values in Table 5.4 are area-wide estimates published by MHCLG; but in each specific location a land value component is embedded in the prices of property transactions (e.g. in the dataset developed for this study). If a developer obtains permission, and then develops a mixed-use development, for example, on a site within walking distance of a new rail station, then the uplift in *land value* may be due to a combination of the improved accessibility and the permissions granted. This subject deserves further analysis in the context of multi-sectoral appraisals (and evaluations) involving – for example – transport, housing and commercial property development, but is beyond the scope of this current research project.

Table 5.4: Land values in selected locations, different uses (permission granted), £/hectare

| **Location** | **Residential** | **Office (edge of City Centre)** | **Office (out of town Business Park)** | **Industrial** | **Agricultural** |
|---|---|---|---|---|---|
| Carlisle | 520,000 | 865,000 | 370,000 | 370,000 | 26,000 |
| Leeds City Region | | 11,140,000 | 1,250,000 | 1,250,000 | 25,000 |
| (Leeds district) | 2,720,000 | | | | |
| Westminster | 113,300,000 | 573,000,000 | | | |
| (Central London - South) | | | | 6,175,000 | |
| (Outer London)[1] | | | 3,000,000 | 3,200,000 | 23,500 |

Data source: MHCLG (2018)

Table 5.5 gives an overview of the types of Land Value Capture mechanisms available in the UK including the TfN area. Some of these apply to existing property and some to new developments; some relate only to residential property, some only to commercial property (see the Table).

Whilst the number of different mechanisms is fairly large and growing, the professional consensus is that the set of mechanisms available in the UK is flawed and would benefit from reform (e.g. House of Commons, 2018; Murphy, 2018; TfL, 2017; SDG, 2016).

- Part of the problem is the low rates of tax/levy that can be applied to existing properties, e.g.:
    - Stamp Duty Land Tax would only capture 5% of a £50,000 uplift = £2,500, on a residential property whose starting value was £300,000; even in London, TfL (2017) found that only 3% on average of any uplift from transport investment could be captured through SDLT;
    - Council Tax receipts are very unresponsive to property values, because of the structure of the tax itself (which is fairly flat) and because properties are revalued infrequently (last in 1991); and
    - any ‘Council Tax Levy’ (or Precept) such as that used for the Olympics in London is usually small and is politically sensitive.
- There are also gaps in the tax base, e.g. Capital Gains Tax is in principle levied on increases in the capital value of assets including property, however a person’s main residence is exempt from the tax.
- Another aspect of the problem is the lack of local retention, for example SDLT receipts are collected and allocated by central government. The more recent developments in LVC are partly about ensuring that regional and local authorities can benefit from increases in their own economic growth and tax take, e.g. TIF/Business Rate Retention as applied in the City Deals in the North of England.
- The two main mechanisms for capturing value from *development* each have issues:
    - Community Infrastructure Levy (CIL) is rather inflexible as it must be set at a flat rate per unit of development and is difficult to reset when new transport projects are approved – although for major projects like Crossrail 1&2 this long term approach appeared to work, with a special Mayoral CIL being introduced for Crossrail 1 and then amended for Crossrail 1&2 (see Table 5.6);
    - Business Rate Supplement (BRS) provided the largest share of LVC funding for Crossrail 1 (it is expected to produce £6.6bn overall towards project costs in the range £16-18bn) and in this sense has been a successful and effective LVC mechanism – the main concern with it is its impact on businesses that rely on property, e.g. high street retail. In any context, including the North of England, an assessment of its impact would be important, pre-implementation. The funding package adopted for Crossrail emerged from consultations with business groups at an early stage, which helped to understand the level and pattern of contributions which would be acceptable.

Table 5.5: Land Value Capture mechanisms in use in the UK, and relevant characteristics

| *LVC mechanism* | *Characteristics and examples* | *Applicability* | | | |
|---|---|---|---|---|---|
| | | *Existing Properties* | *New Development* | *Residential Property* | *Commercial Property* |
| Business Rate Supplement | e.g. Crossrail (+2p/£ on rateable value of business premises) – retained for London | ✓ | ✓ | | ✓ |
| Tax Increment Financing (TIF) and Business Rate Retention | e.g. Battersea Northern Line Extension, Greater Manchester Earn Back Model, Newcastle and Gateshead City Deal – agreed to allocate (a share of) incremental revenue locally.<br>Enables borrowing against future ring-fenced 'tax increment' (such as business rates/property taxes – used in the US and Canada, as well as UK examples) | ✓ | ✓ | | ✓ |
| Stamp Duty Land Tax | (Low) Fixed rates across England, revenue retained by HM Treasury – paid on purchase of property | ✓ | ✓ | ✓ | |
| Council Tax Levy | Limited – must be ratified by referendum if over 2%, and politically sensitive (was used for Olympics and Greater Manchester Transport Package) | ✓ | ✓ | ✓ | |
| Community Infrastructure Levy (CIL) | Only applies to developments (any sector, anywhere within designated area). E.g. Crossrail 1&2, levied at different rates by London Borough. | | ✓ | ✓ | ✓ |
| Section 106 payments | Negotiable between authorities and developers – flexible; usually used for major projects and specific sites; e.g. used by Milton Keynes to set up a general Strategic Land and Infrastructure Contract, developers pay £18,500 per residential dwelling and £260,000 per hectare of commercial land developed | | ✓ | ✓ | ✓ |

Sources: House of Commons (2018); Murphy (2018); TfL (2017); SDG (2016).

Table 5.6: Mayoral CIL rates used to fund Crossrail

| Zone | Boroughs | £ per sq m | |
|---|---|---|---|
| | | MCIL1 (Apr 2012 – Mar 2019) Crossrail 1 | MCIL2 (Apr 2019 -) Crossrail 1&2 |
| Central London & Isle of Dogs – office/retail/hotel only | | | £185 office<br>£165 retail<br>£140 hotel |
| 1 | Camden, City of London, City of Westminster, Hammersmith and Fulham, Islington, Kensington and Chelsea, Richmond-upon-Thames, Wandsworth | £50 | £80 |
| 2 | Barnet, Brent, Bromley, Ealing, Greenwich, Hackney, Haringey, Harrow, Hillingdon, Hounslow, Kingston upon Thames, Lambeth, Lewisham, Merton, Redbridge, Southwark, Tower Hamlets<br>(in MCIL2: + Waltham Forest, London Legacy Development Corporation (LLDC), Old Oak and Park Royal Development Corporation (OPDC)<br>– Greenwich) | £35 | £60 |
| 3 | Barking and Dagenham, Bexley, Croydon, Enfield, Havering, Newham, Sutton, Waltham Forest<br>(in MCIL2: + Greenwich<br>– Waltham Forest) | £20 | £25 |

Source: based on TfL (2019a). The high rates for Central London and Isle of Dogs were formerly part of planning obligations/Section 106 under MCIL1.

Both in the UK and internationally, important issues with land value capture are *horizontal and vertical equity* – these are important not only from a distributional or welfare perspective (see the following two Sections) but in developing an acceptable LVC strategy which can be implemented (e.g. Mulley, 2017).

Horizontal equity refers to the spatial variation in payment compared with the spatial variation in benefits – including land value uplift. This report found that for NPR, for example, there is a small potential uplift across a wide area, and a larger potential uplift more concentrated in the local areas which gain the most accessibility (Figure 5.4). There are marked differences in uplift at OA level and LSOA level close to the stations where connectivity will increase, and areas which are very well connected to those stations by high frequency bus services (for example). This creates a challenge for the design of LVC: how to share the gains of uplift in a way which captures significant revenue for the authority, whilst recognising that some property owners stand to gain significantly more than others?

The practical solution in the case of Crossrail has been to apply CIL and BRS across a wide area at a low level – and then separately negotiate developer contributions (S106) for specific large development sites at or near stations. For the 'wide area' part of this, TfL (2017) reports that the Mayoral CIL for Crossrail equates to 0.48% to 1.13% of the price of an average house across different areas of London. There is some differentiation of CIL by area – see Table 5.6. This kind of spatial differentiation could be considered in the North. There is a precedent for agreement among authorities over LVC-type mechanisms and the allocation of revenues from them, in the form of the Greater Manchester Earn Back model, for example, and the other Northern City Deals (SDG, 2016).

In the case of Crossrail, the Section 106 developer contributions relate particularly to Canary Wharf Group (£150m) at the Crossrail Canary Wharf station, Berkeley Group at Woolwich station and other developments within the 'contribution areas' defined at the outset (including Central London, the Isle of Dogs and 1km radii around certain stations – these were later amalgamated into the MCIL2 charges shown in Table 5.6). Overall, the original funding package for Crossrail was as shown in Table 5.7. The annual stream of BRS payments (~£250m p.a.) has been used to allow GLA to borrow ~£3.5bn to finance the construction phase.

Table 5.7: Overall Crossrail funding package

| | |
|---|---|
| Mayor of London (TfL & GLA) underwritten: | |
| Transport for London (TfL) | £1.9bn |
| Crossrail Business Rate Supplement | £4.1bn |
| Mayoral Community Infrastructure Levy (CIL) | £0.3bn |
| Developer Contributions (S.106) | £0.3bn |
| Over-site Development Opportunities | £0.5bn |
| | |
| SubTOTAL | £7.1bn |
| | |
| DfT underwritten: | |
| UK Government Grant | £4.8bn |
| City of London contribution | £0.25bn |
| Heathrow Airport Ltd contribution | £0.07bn |
| Other | £0.16bn |
| | |
| SubTOTAL | £5.3bn |
| | |
| Network Rail-funded works (recouped through TAC) | £2.3bn |
| | |
| **TOTAL** | **£14.8bn** |

Sources: Buck (2017); TfL (2017).

In the case of the London Overground Extension to Barking Riverside – a regeneration area with permission for 10,800 new homes, a school, healthcare facilities and a district centre with commercial and leisure facilities – 63% of the project cost will be paid by developer contributions (£172m of £273m) and the remainder will be paid largely by TfL's Growth Fund, which is targeted at unlocking housing and employment growth (TfL, 2019b). The 'anchor' developer is Barking Riverside Ltd., a joint venture between the GLA (49%) and London and Quadrant New Homes (51%). As such the *private sector* developer contribution is 32%. This is a case where land has been assembled and the share of costs and rewards between the developer and the public authorities has been carefully negotiated, based on a prediction of the likely uplift and the implications for sale prices and development profitability.

Locations like Kings Cross St Pancras, and the planned HS2 stations in the larger northern cities, are likely to generate large uplifts because they concentrate the flow of large numbers of people including ‘high value/income’ users, through these locations, and make development nearby very attractive. They also contain (typically) improvements to place quality, urban realm and connecting services as well, strengthening the uplift. These locations require (and are receiving) special assessment and planning. Masterplanning of residential and commercial development is part of the ongoing activity, and as this report has found, planning for clusters of business and employment in the city is potentially a source of further uplift (Section 4.1.8).

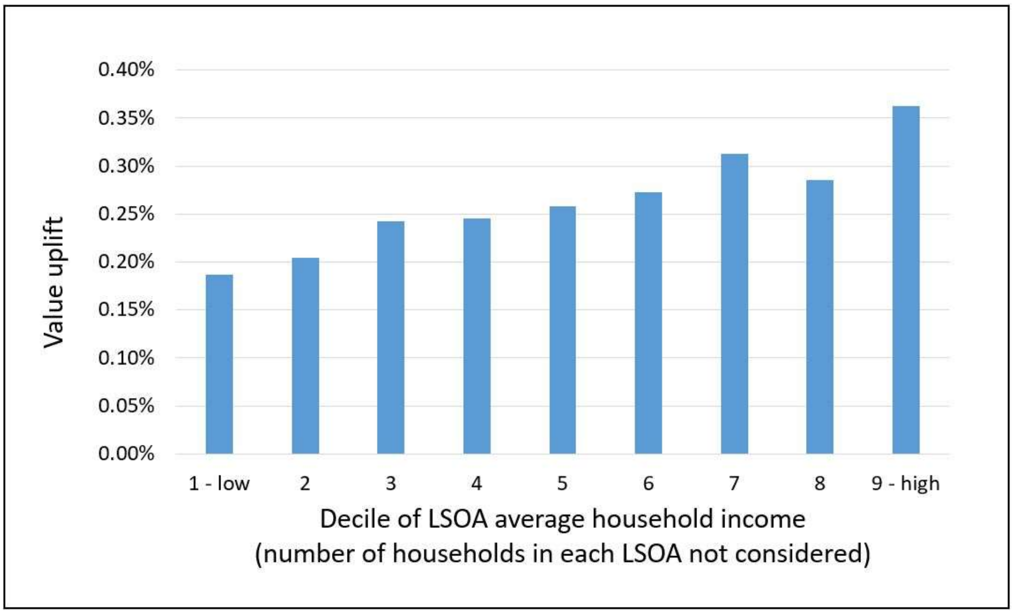
Finally, it is worth considering whether smaller areas could be identified for differentiated LVC rates, bearing in mind the strength of the localised effects identified by the cross-sectional and quasi-experimental models (at OA and LSOA level, and in 250/500/1000/2000m catchments around new stations). The models provide an evidence base for a localised rail accessibility premium in the housing market and also in the commercial property market. The Metrolink case study would suggest caution, because in the case of the Rochdale Line particularly there is little evidence of significant uplift, six years after opening, which highlights the role of local market conditions at least in the short term – until regional growth leads demand to rise in this corridor. If a more locally-differentiated LVC model was adopted, it would need to be designed to be responsive to such local conditions.

Table 5.3 showed the size of the uplift potential from NPR in the residential market, based on the early cross-sectional model (Model 12). It is not possible at this stage to determine an ideal LVC model for NPR or to predict with confidence what proportion of the potential value may be captured, however Crossrail 1 provides a useful benchmark. TfL found that applying LVC mechanisms to Crossrail 1 will result in it capturing about 8-10% of the total potential uplift (TfL, 2017). However up to 30-40% is possible if more ambitious mechanisms, addressing existing properties in particular, are developed and used – these are not available off-the-shelf and require work with Government to make them available.

### 5.4 Distributional Impacts

The analysis undertaken so far allows us to extract some preliminary results on the distributional effect of the property value uplift from changes to the network. These results are based purely on the rail accessibility effect – including income interaction, without any of the further effects shown in Figure 5.1. Figure 5.7 reports the estimated average uplift by income decile, based on an LSOA level analysis. Decile 10 is missing because some of the data in this decile requires further cleaning to be used for this type of analysis; the other deciles are believed to be unaffected. The results suggest that higher income households will be affected more strongly by the property value uplift effects of a project to improve existing rail services.

Figure 5.7: Potential property value uplift by average income decile (LSOA level)




Note: 10th decile omitted due to data issues.

Figure 5.8 provides a map of the predicted property value uplift with household incomes marked, for a selected urban area (Leeds). This highlights how uplift may extend to higher-income areas even some distance away from the rail stations with improved connectivity, whilst the LSOAs closest to the stations tend to experience uplift even where income is lower.

Since we also have tenure data at LSOA level, it is possible to infer how the property value uplift affects areas with more or fewer renters. Figure 5.9 shows that in the City Centre and denser suburbs near the Universities, there are larger uplifts associated with high shares of renting. In the outer suburbs, uplifts appear to impact on a larger share of owner-occupiers.

In combination, these results make it possible to infer the pattern of impact on owners and renters at different income levels, for whom the welfare impact of the change in property values will differ. Owner-occupiers, landlords and renters will experience these impacts in different ways (there are wealth and income effects to be accounted for). Existing owners, first-time buyers, and downsizers will also experience the impact differently.

An age breakdown of the impact is also possible – Figure 5.10 provides an example for the same Leeds Centre-North West area covered by the income and tenure maps. There is a bar on the map for each LSOA, and the black part of this bar indicates the average age in this LSOA as a proportion of the highest in Leeds district as a whole. An observation is that the most uplifted areas are neither uniformly older nor uniformly younger than the rest of the city – there is mixed pattern of impact by age. In the North Western suburbs (around Horsforth station) the impacted population is notably above average age, while in the City Centre and around Burley Park station the impacted population appears relatively young. The area around Bramley station illustrates how two LSOAs both close to an existing station can have entirely different population age characteristics.

Figure 5.8: Pattern of potential property value uplift by income at LSOA level, Leeds

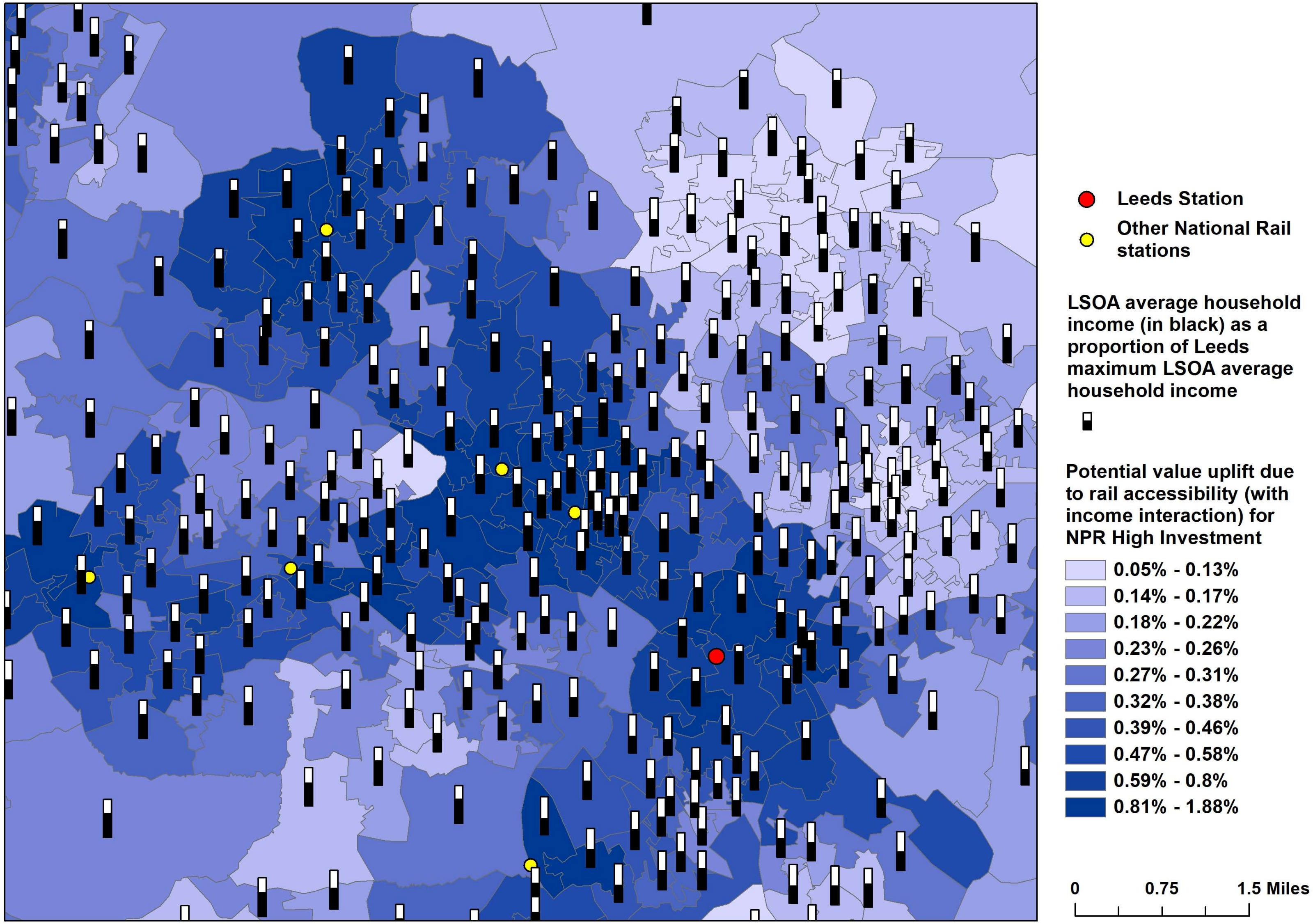


Figure 5.9: Pattern of property value uplift by tenure (share of owned/rented), Leeds

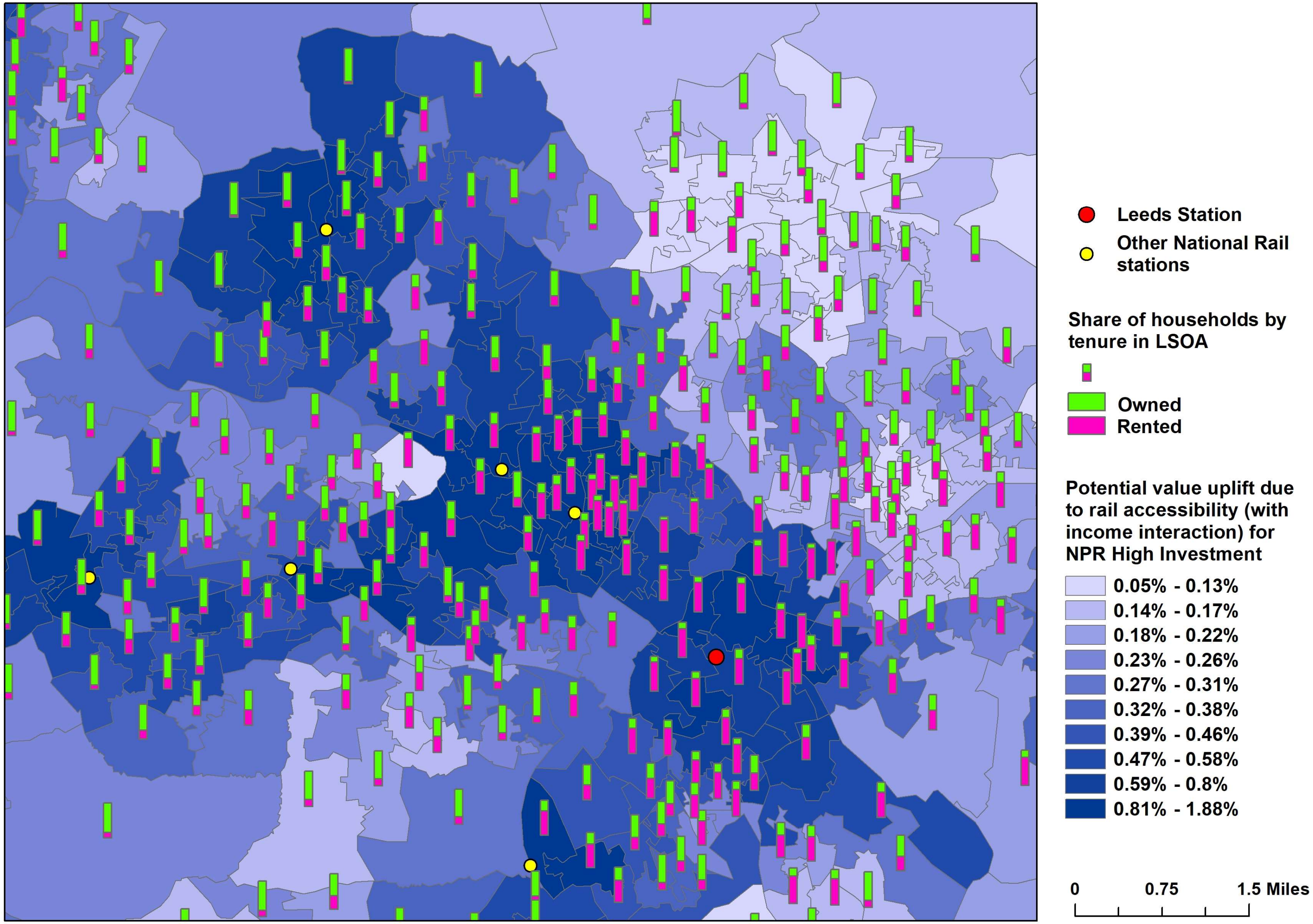
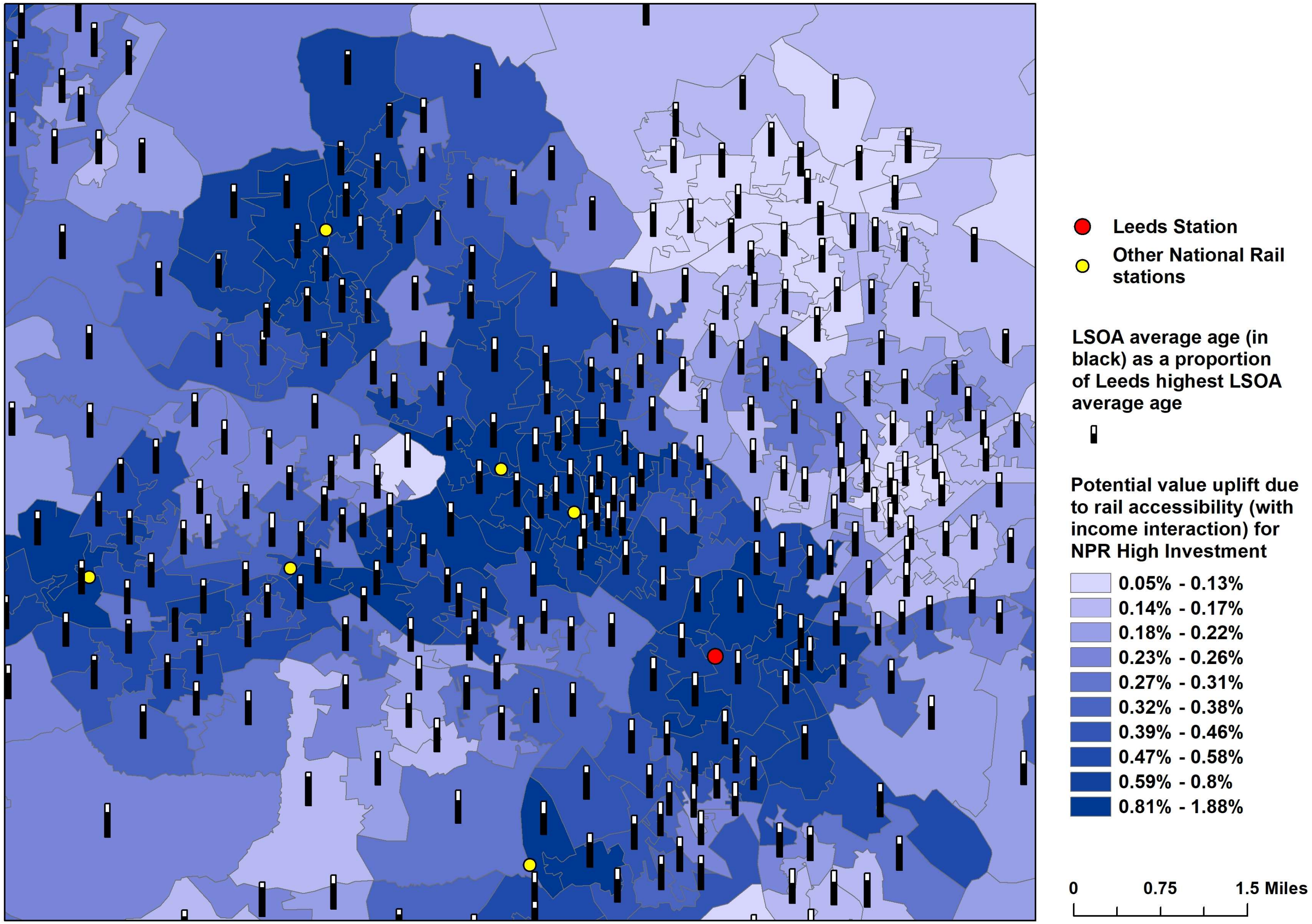


Figure 5.10: Pattern of property value uplift by average age in LSOA, Leeds



In conclusion, it is recommended that future research on the distributional impacts of transport projects & policies make use of this kind of detailed spatial data on impact and population characteristics, to understand who will benefit and who (if anyone) will lose from the intervention, and the distribution of the benefits (and any losses) across income groups, age groups and other relevant population characteristics. It is worth noting that the total impact of a project like Northern Powerhouse Rail will depend on the full set of effects that flow from the project, not only these house price and rental impacts – e.g. movement to more productive jobs across the regional economy, development around rail 'hubs', changes in labour force participation and participation in training. It will also depend on any accompanying measures that form part of the policy package, which could include changes in spatial planning around rail 'hubs' and – in the event that any Land Value Capture mechanisms are used – the details of how those impact on owner-occupiers, private and social renters, investors and businesses, for example. There will almost always be areas that are located away from a particular project and/or benefit less: distributional analysis of the area-wide strategy is a natural counterpart of project-level analysis.

### 5.5 Integration with appraisal

Finally, major transport projects in England are required to have a Business Case (pre-implementation). At the heart of the Business Case is an Economic Case which is an assessment of the welfare (wellbeing) implications of the project, alongside a Strategic Case and a Financial Case – which together form three parts of the 'Five-Case Business Case' (DfT, 2013/2018). The role of the Business Case is to ensure that decision-makers are well-informed about the consequences of the proposed project and any alternative options within or around it. This evidence-based approach has stood the transport sector in good stead in Public Spending Rounds, and is part of the current practice in this sector internationally, not only in the UK.

Table 5.8 summarises this study's findings on how appraisal may need to evolve in future in order to integrate evidence about land value change within the Transport Business Case. Implementation would be subject to further methodological development. The main points relate to:

- the need to represent both gains and losses to different groups in the appraisal – this is good practice anyway, and is reinforced in DfT's recent guidance on investment strategy (DfT, 2017a);
- the need to see transport policy together with housing and economic development policy, not in isolation from it (again this is in line with DfT's stance on strategy – DfT, 2017a);
- the potential role of land value capture (LVC); and
- the issues that exist in measuring the welfare changes to people in the study area.

Table 5.8: The roles of land value change in the Transport Business Case

| | *Type of analysis* | *Implications (from the study)* |
|---|---|---|
| Economic Case | Welfare economics, Social CBA | NPV=PVB-PVC<br>• Who benefits, who loses<br>• Dependent development<br>• Wider labour & productivity impacts<br>• LVC contributes to ↓C |
| Financial Case | Financial flows, Funding and financing arrangements, Financial sustainability check | • LVC contribution to Costs<br>• Borrowing against future LVC |
| Strategic Case | Policy fit, Objective achievement | • Policies relating to housing affordability<br>• Policies relating to land use and economic development |

CBA=Cost Benefit Analysis; NPV=Net Present Value; PVB=Present Value of Benefits; PVC= Present Value of Costs

When land values change, e.g. as a result of a transport investment, it is expected that there will be both benefits and disbenefits as shown in Table 5.9. The final incidence of benefits and disbenefits will depend on the interplay of the effects in Figure 5.1, involving households, businesses, the land and property markets and the labour market. A better understanding of these interactions between markets is needed.

Table 5.9: Initial benefits and disbenefits of land value change

| *Impact group* | *Impact* | *Potential magnitude (Northern context)* |
|---|---|---|
| Owner occupier residents | Near to the new facility: land value uplift<br>• **capital value gain (windfall) (+)**<br>• potential Council Tax and insurance cost increases (−) | Moderate: cross-sectional and quasi-experimental models point to approx. range 0 to +20% (for interventions considered) |
| | Distant from the new facility: land value change of undetermined sign | Small: effect spread over a wide area |
| Renting residents | Near to the new facility:<br>• **rent ↑ (−)**<br>• potential Council Tax and insurance cost increases | Related to property value uplift. |
| Investors (residential and commercial) | Near to the new facility: land value uplift<br>• **capital value gain (windfall) (+)**<br>• **rental income ↑ (+)** | Moderate: cross-sectional and quasi-experimental models point to approx. range 0 to +20% (for interventions considered) |
| Future movers (in) | • prices, rents, SDLT on purchase and insurance costs increased (increased cost of living) (−) | Related to property value uplift. |
| Business occupiers | • **rent ↑ (−)**<br>• capital gain (if owners) (+)<br>• NNDR & insurance costs (−) | Magnitude uncertain: some evidence of rail accessibility premium approx. 12%; uplift may exceed this locally. |
| Local Authority | Potential ↑ in Council Tax and NNDR revenue (net across areas) (+) | Total limited by infrequent revaluations and low % rate, and by loss of pooled part of NNDR. |
| Central Government | Potential ↑ SDLT and CGT revenue (net across areas) (+) | Total limited by applicability to movers/sellers but not to all existing property, and by low % rate and exemptions. |

SDLT = Stamp Duty Land Tax; NNDR – National Non-Domestic Rates (revenues shared 50% locally, 50% pooled nationally); CGT = Capital Gains Tax.

The wellbeing impact of these changes for residents will depend on their magnitude and the circumstances of the people impacted, including their household income (HM Treasury, 2018, outlines current official thinking on distributional weighting). Table ES10 highlights the potential complexity of these calculations. Relocation of households triggered by these changes might imply some further losses from the dislocation of an undesired move – potentially to a less well-connected location.

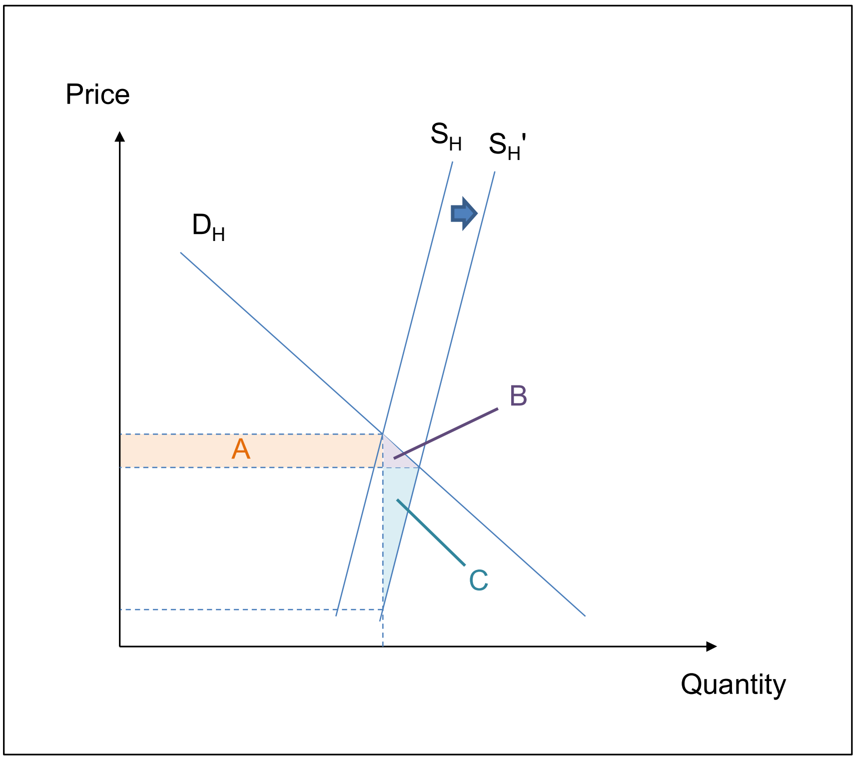
In the case of *dependent development*, where transport improvements unlock development on a particular site or sites, the net value of the additional development may be counted as a measure of benefit under certain conditions (TAG Unit A2.1, DfT, 2018; DfT 2017b). However, due to the uncertainty surrounding such estimates, these may only be included as 'indicative monetised impacts' which are not shown to the decision maker as part of the 'Initial Benefit:Cost Ratio (BCR)' or the 'Adjusted BCR'. Instead they should be presented as separate information[31]. Nevertheless, the dependent development (induced investment) case points towards the important role of development as a both a market response and policy response to rising accessibility premiums when transport investment occurs.

Seen at a regional level, dependent development represents the relaxation of the supply constraint on housing (in particular) and also commercial property through transport investment. Barker (2003) highlights that there may be a welfare loss from an inefficiently-low level of housing supply (Figure 5.11, areas B+C), and by implication a welfare gain from a strategy to help to ease these constraints. Targeted transport investments as part of an overall strategy should serve to ease these constraints if, for example:

- Land in existing cities is unprofitable to redevelop due to a lack of connectivity to workers, consumers or other market/non-market opportunities;
- The planning system is constraining the efficient supply of housing/commercial development;
- There are monopoly/oligopoly ownership concerns in the land market;
- There are co-ordination failures between developers in the provision of transport infrastructure[32].

A key concern for economists will be to establish the extent of any externalities or market failures in the land and property markets, such that the welfare implications can be quantified.

Figure 5.11: Welfare loss from restricted housing supply




[31] From a financial perspective such dependent development may be a source of revenue through Land Value Capture mechanisms (e.g. in Crossrail 2 and the East Leeds Orbital Road – TfL, 2017; Leeds City Council, 2017) and hence reduce the Cost to the Broad Transport Budget (C) in the Economic Case.

[32] These issues are highlighted in TAG Unit A2.1 (DfT, 2018).

Area A in Figure 5.11 is a transfer from renters (and purchasers) to owners (and sellers) when prices rise due to a supply constraint. This may have welfare implications if the income of the owners and renters are different (or the seller and the purchasers are different). Distributive weighting then comes into play (HM Treasury, 2018).

In some cities, including Manchester (Manchester City Council, 2018) and Helsinki, for example, there are explicit policies on housing affordability, which emphasise the role that housing costs play in achieving the cities' strategic economic and social objectives (Table 5.10). For example, policy may target housing costs that are competitive with other cities (nationally/globally) to help attract mobile workers, and housing affordability may be part of achieving inclusive growth/social cohesion. These policies highlight that while accessibility improvement has economic benefits, the land value uplift it creates can be a variable to be managed rather than maximised. Increasing housing supply can be a part of that management – including policies for 'transit oriented development'.

Table 5.10: Manchester Housing Affordability Strategy

| *Manchester Strategy outcomes* | *Summary of the contribution to the strategy* |
|---|---|
| A thriving and sustainable city: supporting a diverse and distinctive economy that creates jobs and opportunities | Ensure Manchester has the right mix of housing that is affordable across a range of tenure and income levels to support a functioning Manchester and sub regional economy. |
| A highly skilled city: world class and home grown talent sustaining the city's economic success. | The new and existing homes will be well connected to employment opportunities and schools. |
| A progressive and equitable city: making a positive contribution by unlocking the potential of our communities | Increasing the supply of good quality affordable homes for sale and rent will provide the opportunity for Manchester residents to raise their individual and collective aspirations. |

Source: extract from Manchester City Council (2018).

As in Figure 5.1, the outcome from a transport investment will be the result of interactions between several linked markets, including labour, land and property, and business and industry. The outcome may be the reshaping of the economic geography, with relocation of businesses to more productive locations, households to locations where they can access more productive jobs, benefits through labour supply and agglomeration, etc. All of this is subject to economic analysis and modelling that goes beyond the scope of this study. The contribution of this study is to better understand the specific relationships between accessibility, place quality and value as measured in the property market – and as such to make a contribution to that wider modelling effort.

The relevance of land value uplift (LVU) in appraisal was outlined by the Phase 1 study (Nellthorp et al., 2016) as follows: *"it is not a substitute for transport user benefits in the appraisal of most rail or road projects. The main reasons for this are:*

*i. there are substantial doubts about whether the land market will fully capitalise all the transport benefits (or under/over-capitalise them);*

*ii. additionality: there is therefore no certainty over what part of the LVU is 'additional' and what part is 'double counting' the transport benefits for a transport project.*

*However, it does have two potentially important roles in appraisal:*

*i. as a source of values for place-related effects – a surrogate market – as it has been for noise and air quality valuation. This should be applied to … place-related* *effects, and will form part of the social benefits, and hence the Benefit:Cost Ratio (BCR). Note that this involves measuring the place-related benefits using a suitable model (such as HP/GWR), measuring the underlying welfare gain whilst controlling for variations in supply-demand conditions in local property markets. Using the 2nd stage of the hedonic model is one way which this could be addressed.*

*ii.* *as a source of funding through value capture, which reduces the cost to government of investment in infrastructure and hence reduces the Cost to the Broad Transport Budget in the BCR.*

*A third potential role for LVU is in measuring the value of dependent development enabled by a transport investment: this is covered in the consultation document DfT (2016), 'UVITI: updating wider economic impacts guidance' and the five linked TAG Units A2.1-2.4 and M5.3. The extent to which the development is dependent on the transport investment is a key input to that analysis.*

*In the Financial Case, the change in revenue may contribute to making the project viable and deliverable – as for Crossrail2, where 37% of the project funding comes through these channels.*

*Hence LVU has implications for the Economic Case and the Financial Case, as well as the Strategic Case, as part of the overall Transport Business Case".*

This Phase 2 study has added quantification to these matters. We are now in a better position to measure and predict the effects of accessibility improvements on property prices, and to measure the value of a range of place quality effects in the residential property market of the North of England (in Sections 4 and 5).

# 6. Conclusions

## 6.1 The models developed

New models of the property market have been developed, quantifying the links between accessibility and property values across a large urban and rural region[33] – the North of England. Three different types of model have been created:

- The Cross-Sectional Residential model is the most detailed, both spatially and in the way it represents the various influences on property values. It includes a set of variables measuring accessibility and place quality. It covers the whole North of England, an economically, environmentally and socially diverse area, at a fine level of spatial detail. As such it is a unique model (globally, not only in the UK).

  14 main variants of this model, and many more sub-variants, were tested. Across these models, the signs and magnitudes of the parameters are generally remarkably stable. Most parameters are highly significant, and the large dataset helps to support a model with a large number of significant variables. When applied, for example, to commuter rail stations in Greater Manchester, the model produces results that are comparable with previous well-regarded studies of that area (Nationwide, 2014/19). These results are encouraging for the robustness and wider applicability of the model.

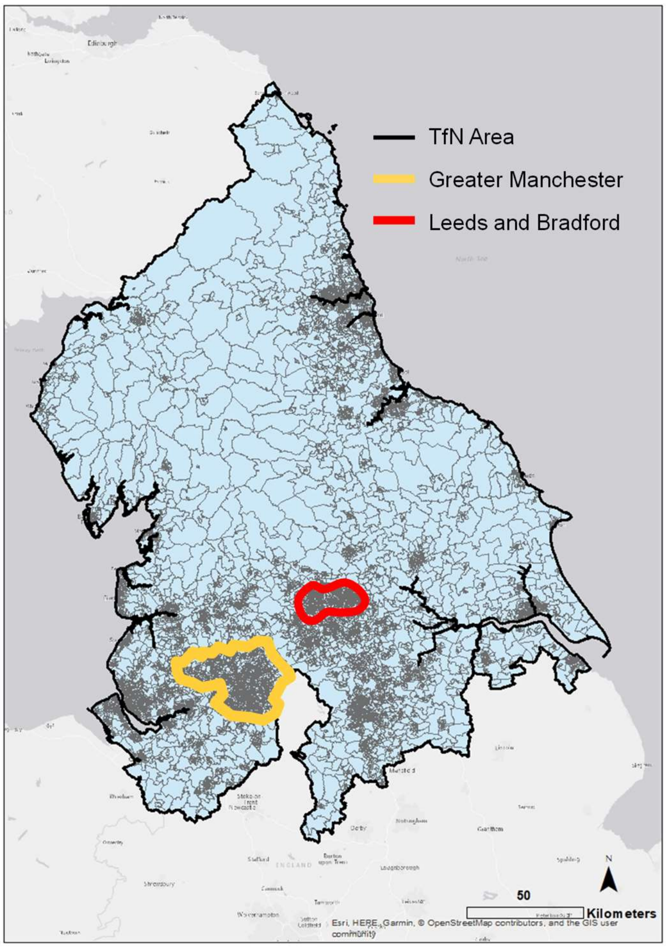



Figure 6.1:

Modelled areas: TfN area (whole); Greater Manchester; Leeds & Bradford

[33] 15 million people, 7.1 million homes and economic output of £350bn per annum (GVA).

- The Quasi-Experimental Residential model addresses how property values change in response to a transport intervention, in order to understand how *'uplifts'* over time differ from cross-sectional value premiums. The model focuses on the Greater Manchester property market and the impact of Metrolink extensions 1995-2018.

  The findings show that there is a basic level of consistency between the two types of model. The average uplift due to Metrolink stations (6.3%) is on a par with the accessibility premium for rail stations in Greater Manchester (approximately 6%) calculated using the cross-sectional model. The cross-sectional model predicts that the premium will be greater in areas with a higher income population and vice versa – this pattern is also seen in the quasi-experimental model, e.g. larger uplifts in the higher-income South Manchester corridor and smaller or even insignificant ones in the lower-income East Manchester and Rochdale corridors. The range of uplifts is rather more exaggerated in the quasi-experimental model – this is potentially due to the further effects discussed in Section 5.1 and this is an area for further investigation.

- The prototype Cross-Sectional Commercial model shows how rents vary with: proximity to rail stations; walk accessibility to other jobs/effective density; property type (office/industrial/retail); quality; and size. The commercial dataset is not yet rich enough to obtain a deeper understanding of how rail connectivity to other cities affects values, but it is planned to address this in the next phase of modelling. Ways of explicitly including place quality in this model are also potentially worth considering in future.

### 6.2 The scale of uplift due to rail investment

Reflecting on the scale of the uplift found in this study compared with previous empirical research:

- The study by Gibbons and Machin (2005) is an important benchmark. It was a pioneering quasi-experimental study of two urban rail investments in London: the Jubilee Line Extension (JLE) and the Docklands Light Railway (DLR) Lewisham Extension. This provides for a London:Manchester comparison.

  Gibbons and Machin found that the uplift within a 2km new station catchment is +9.3%, with no effect beyond the 2km threshold. For comparison, results from the Quasi-Experimental Model in this study[34] show the uplifts in Table 6.1 for three Metrolink lines.

  - The South Manchester line is a good starting point for comparisons given the level of household income compared with London: Gross Household Disposable Income per capita at East Didsbury is only 4.5% below the London Borough of Newham, where the end of the JLE is located (ONS, 2018b). The uplift is +8.2% for Metrolink, versus +9.3% for the JLE/DLR extension in Gibbons and Machin's analysis.

  - The Airport line is a special case, in that it has a cluster of employment at both ends, which should have a positive impact through the rail accessibility premium. The uplift of +12.3% in the 0-2km catchment is larger than the Gibbons and Machin result for the JLE/DLR extension, which is an interesting finding.

[34] Section 4.3

- Not surprisingly, the 0-2km uplift for the East Manchester Metrolink line is lower than South Manchester (+6.2% vs +8.2%) and is not significant at 95% confidence, probably reflecting the lower income profile of the corridor and local property market conditions.

Table 6.1: London:Manchester uplift comparison: Gibbons & Machin (2005) versus the quasi-experimental model (Metrolink)

<table>
<tr><th>City</th><th>Route</th><th colspan="2">Uplift 0-2km</th><th>Source</th></tr>
<tr><td>London</td><td>JLE & DLR</td><td colspan="2">9.3%</td><td>Gibbons and Machin (2005)</td></tr>
<tr><td rowspan="9">Manchester</td><td rowspan="3">South Manchester</td><td colspan="2">8.2%</td><td rowspan="9">This Study – ITS (2019)</td></tr>
<tr><td>Uplift 0-1km</td><td>Uplift 1-2km</td></tr>
<tr><td>10.6%</td><td>7.4%</td></tr>
<tr><td rowspan="3">Airport</td><td colspan="2">12.3%</td></tr>
<tr><td>Uplift 0-1km</td><td>Uplift 1-2km</td></tr>
<tr><td>20.9%</td><td>9.4%</td></tr>
<tr><td rowspan="3">East Manchester †</td><td colspan="2">6.2%</td></tr>
<tr><td>Uplift 0-1km</td><td>Uplift 1-2km</td></tr>
<tr><td>7.5%</td><td>5.7%</td></tr>
</table>

Note: Manchester uplifts weighted by area; † East Manchester not sig. at 95% (t=1.6).

- Later analysis has produced larger uplifts in some cases in London, including for JLE. Notably, TfL (2017) reports work by Savills which shows the premium for the JLE 'zone of influence' compared with a 'control area', oscillating between +30% and +60% after opening in 1999 up to June 2004.

  Separately, however, Ahlfeldt's (2013) econometric analysis – which is quasi-experimental like our Metrolink analysis and like Gibbons and Machin's study – produces uplifts around stations on the JLE and DLR Lewisham Extension that are smaller than the Savills/TfL results. For example, North Greenwich station on the JLE shows some of the largest uplifts in Ahlfeldt's results: only a small area near the station (~300m or less) has uplifts between 16-28%; the rest of the 0-1km catchment shows uplifts between 5-16%; and the 1-2km buffer zone contains a uplifts from 1.5-16%. Bearing in mind that the JLE has a service frequency of 22-23 trains per hour in the AM peak into the some of the densest employment locations in Central London and Docklands, versus 9-10 per hour on the Metrolink South Manchester line into a less dense city centre, this difference in uplifts actually seems in tune with the theoretical model.

  Other stations have smaller uplifts in Ahlfeldt's analysis, e.g. Lewisham has (broadly) 10-16% in the 0-1km catchment and 5-10% in the 1-2km buffer; Deptford Bridge on the Lewisham Extension has uplifts mostly between 1.5-10% even in the 0-1km catchment. In both these cases the new DLR line serves Canary Wharf in Docklands, rather than providing direct service to Central London. The order of magnitude of these uplifts, relative to the Manchester findings, does seem to present a plausible overall picture.

- Finally, as a sense check, it is worth referring back to the Nationwide (2019) cross-sectional premiums for rail stations in London compared with Manchester: the comparison is +9.4% in London versus +7.8% in Manchester (at 500m from the station); or +4.1% in London versus +3.3% in Manchester (at 1000m from the station) (see Table 4.10). These differences between London and Manchester might be explained by differences in rail service (GJT), employment density and income (seen through the Cross-Sectional Residential model framework above – detailed calculations have not yet been done on this).
- The general narrative emerging is that:
    - Land value uplift is likely to be somewhat smaller (on average) in the North than in London, although there are exceptions;
    - As in any location, the size of the uplift will depend on the quality of the connectivity provided to economic and other opportunities;
    - Income is a factor which may limit uplifts in the North, unless/until the productivity and income gaps with London are closed; lower initial economic densities are another factor;
    - Low property demand in some corridors seems to be associated with low or insignificant uplifts in the short-to-medium term, however this does not preclude significant uplifts following later as housing demand grows with the regional economy over time – spatial strategy could support this.

### 6.3 The opportunity for value capture

In conclusion:

- As anywhere in UK, the set of LVC instruments is limited and this can be a barrier to implementation of land value capture (Section 5.3), however there is already experience with these LVC instruments in practice across the UK including the North of England;
- The size of the uplifts available may – as discussed above – be somewhat smaller than in London, however the evidence to date does not suggest that uplifts are orders of magnitude smaller for comparable projects. For example, evidence of uplifts for Manchester Metrolink extensions between 1995 and 2018 indicates that these average +6.3% and range from insignificant to +21% at this stage[35].
- The spatial pattern of uplifts means that for rail investment, authorities may need to consider LVC mechanisms to capture both a wide-area low-level uplift and a more concentrated uplift in sites close to the station. E.g. using CIL for new developments across a wide area (possibly differentiated by locality), and/or Business Rate Retention/TIF/Business Rate Supplements for commercial property, and/or Council Tax Precepts/SDLT[36] for residential property. Section 106 agreements are likely to be useful in the more concentrated area, and these are relatively flexible.

---

[35] in the 0-1km catchment.

[36] SDLT not being hypothecated or retained locally is an issue.

- Other types of intervention may have different spatial patterns of impact, and require a different mix of LVC mechanisms.
- Successful LVC involves a sharing of costs and rewards that is acceptable to all parties – hence consultation and negotiation are important, backed up by evidence. The issues of horizontal and vertical equity are seen as relevant in the international literature on LVC.

### 6.4 Integration with appraisal and policy

Integration of LVC with appraisal has been explored through the Northern Powerhouse Rail example, leading to preliminary estimates of the total uplift. The ability to capture this and feed it into the Financial and Economic Cases through a stream of contributions to Costs has been considered in Section 5.3.

The wider implications for the Benefits and for distributional analysis in the Economic Case were considered in Section 5.5 – this is work in progress and will require further input to reach theoretical and empirical conclusions, however it is clear there are potential welfare (wellbeing) impacts from land value uplift.

This work was focused mainly on measuring the impact of rail accessibility improvements, however the findings are relevant not only for new and existing stations but also for:

- access to stations – including pedestrian networks;
- service improvements;
- walkable neighbourhoods;
- new housing;
- commercial development;
- environmental improvements and green space.

An obvious question is 'are the parameters transferable'? Most of them were estimated on the whole TfN area using the Cross-Sectional Residential model, so the starting assumption is that they have wide applicability within this area. They do need to be viewed in the context of findings elsewhere – in London, for example (Section 6.2), where larger uplifts due to rail investment have been seen and predicted.

Another important question is whether these interventions including LVC are likely to attract widespread support, not only whether they generate useful financial flows. This depends on the distribution of the benefits and costs – i.e. who gains and who loses – including spatially (a form of horizontal equity) and socio-economically (vertical equity). A very important point emerging from the literature in this field, particularly Mulley et al. (2017) and Ahlfeldt (2013), is that network effects are a powerful mechanism for spreading the benefits of infrastructure investment across a wider area. Mulley et al. found that an initial uplift of 7% due to opening of a southern BRT in Brisbane was followed by a further uplift of 3% due to the opening of a connected northern busway. Ahlfeldt found uplifts of 1.5-5% in the 0-1km catchment around New Cross Gate station even though this was not even part of the JLE or DLR extensions – it was connected to the JLE via one interchange and the uplift is a clear network effect. This highlights the role of network planning across modes to ensure that as wide a set of residents as possible stand to gain from a relatively focused infrastructure investment.

### 6.5 Evidence gaps/needs

Finally, the following is a very brief summary of the evidence needs identified as remaining at the end of this Phase of the research:

- Commercial model – re-estimation and development using wide-area dataset;
- Built environment / urban realm / blue/green infrastructure – there is scope for further inclusion of these variables in the model.
- Disaggregation of employment to allow skills matching as a potential barrier to patronage and uplift in the short-medium term in some corridors.
- Dynamics – this is a substantial area that warrants further study, including the property market, labour market and business, and interactions between them.

# Appendix I: Correlation matrix

Table I.1: Full correlation matrix

| Variable Number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 LN_Price | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2 Rail Jobs (P) | 0.01 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3 Walk Jobs (P) | -0.09 | 0.39 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 4 Car Jobs (P) | 0.00 | 0.56 | 0.60 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 5 PriSchool rate (1) | -0.13 | -0.02 | -0.01 | -0.04 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 6 PriSchool rate (Avg) | -0.21 | -0.05 | -0.04 | -0.07 | 0.48 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 7 PriSchool rate (5) | -0.07 | -0.01 | -0.02 | -0.03 | 0.04 | 0.48 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | | |
| 8 SecSchool rate (1) | -0.16 | -0.01 | 0.01 | 0.02 | 0.01 | 0.09 | 0.05 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | | |
| 9 SecSchool rate (Avg) | -0.24 | 0.02 | 0.01 | 0.02 | 0.02 | 0.09 | 0.06 | 0.49 | 1.00 | | | | | | | | | | | | | | | | | | | | | | | |
| 10 Parks (P) | -0.14 | 0.22 | 0.31 | 0.17 | 0.01 | -0.02 | -0.02 | 0.03 | 0.08 | 1.00 | | | | | | | | | | | | | | | | | | | | | | |
| 11 Playground (P) | -0.12 | 0.06 | 0.03 | 0.04 | 0.05 | 0.04 | 0.00 | 0.01 | -0.01 | 0.05 | 1.00 | | | | | | | | | | | | | | | | | | | | | |
| 12 Bank (P) | -0.10 | 0.18 | 0.18 | 0.05 | 0.02 | -0.02 | -0.01 | 0.02 | 0.01 | 0.28 | 0.11 | 1.00 | | | | | | | | | | | | | | | | | | | | |
| 13 TownCentre (P) | -0.05 | 0.20 | 0.18 | 0.09 | 0.02 | -0.03 | -0.03 | -0.01 | 0.01 | 0.30 | 0.07 | 0.50 | 1.00 | | | | | | | | | | | | | | | | | | | |
| 14 Landfill (dist) | 0.05 | -0.02 | 0.01 | -0.14 | 0.03 | 0.03 | 0.01 | -0.02 | -0.02 | 0.03 | -0.05 | 0.10 | 0.04 | 1.00 | | | | | | | | | | | | | | | | | | |
| 15 Students | 0.03 | 0.07 | 0.23 | 0.06 | 0.02 | 0.01 | -0.01 | -0.01 | -0.05 | 0.09 | -0.01 | 0.08 | 0.10 | 0.05 | 1.00 | | | | | | | | | | | | | | | | | |
| 16 Tourists | 0.05 | -0.02 | 0.01 | -0.07 | -0.01 | -0.02 | -0.02 | -0.02 | -0.03 | 0.01 | -0.01 | 0.02 | 0.04 | 0.06 | 0.03 | 1.00 | | | | | | | | | | | | | | | | |
| 17 PM 2.5 | -0.20 | 0.20 | 0.24 | 0.19 | 0.05 | 0.05 | 0.02 | 0.03 | 0.05 | 0.18 | 0.09 | 0.19 | 0.17 | 0.00 | 0.02 | -0.04 | 1.00 | | | | | | | | | | | | | | | |
| 18 Crime | -0.26 | 0.13 | 0.34 | 0.13 | 0.06 | 0.11 | 0.06 | 0.06 | 0.08 | 0.18 | 0.04 | 0.22 | 0.26 | -0.02 | 0.09 | 0.19 | 0.12 | 1.00 | | | | | | | | | | | | | | |
| 19 Noise Road_55_59 | 0.02 | 0.08 | 0.06 | 0.19 | -0.02 | -0.03 | 0.00 | -0.01 | 0.00 | -0.02 | -0.02 | -0.05 | -0.02 | -0.06 | -0.02 | -0.01 | 0.00 | 0.03 | 1.00 | | | | | | | | | | | | | |
| 20 Noise Road_60_64 | 0.03 | 0.03 | 0.02 | 0.11 | -0.01 | 0.00 | 0.00 | -0.01 | -0.01 | -0.01 | 0.00 | -0.03 | -0.01 | -0.02 | -0.01 | 0.00 | 0.00 | 0.01 | -0.08 | 1.00 | | | | | | | | | | | | |
| 21 Noise Road_65_69 | 0.00 | 0.03 | 0.02 | 0.07 | 0.00 | 0.01 | 0.01 | 0.00 | 0.00 | -0.01 | 0.00 | -0.01 | 0.00 | -0.02 | -0.01 | 0.00 | 0.01 | 0.02 | -0.04 | -0.03 | 1.00 | | | | | | | | | | | |
| 22 Noise Road_70_75 | 0.00 | 0.00 | 0.01 | 0.02 | 0.01 | 0.02 | 0.00 | 0.00 | 0.00 | -0.01 | -0.02 | -0.01 | 0.00 | -0.01 | 0.00 | 0.00 | 0.00 | 0.00 | -0.02 | -0.01 | -0.01 | 1.00 | | | | | | | | | | |
| 23 Floor area | 0.61 | -0.02 | -0.07 | -0.06 | -0.06 | -0.08 | -0.02 | -0.06 | -0.07 | -0.02 | -0.06 | -0.01 | -0.01 | 0.05 | 0.03 | 0.03 | -0.08 | -0.09 | 0.00 | 0.04 | 0.01 | 0.00 | 1.00 | | | | | | | | | |
| 24 New_Build | 0.19 | -0.05 | -0.04 | -0.06 | 0.00 | 0.01 | 0.01 | 0.00 | 0.00 | -0.09 | -0.08 | -0.11 | -0.05 | -0.01 | -0.02 | 0.01 | -0.07 | 0.05 | 0.08 | 0.08 | 0.04 | 0.05 | 0.08 | 1.00 | | | | | | | | |
| 25 Freehold | 0.06 | -0.14 | -0.15 | -0.25 | 0.04 | 0.08 | 0.02 | -0.01 | -0.03 | -0.07 | -0.01 | -0.03 | -0.07 | 0.08 | -0.03 | -0.01 | -0.04 | -0.06 | -0.05 | -0.05 | -0.03 | -0.01 | 0.11 | -0.11 | 1.00 | | | | | | | |
| 26 Terrace | -0.40 | 0.09 | 0.17 | 0.07 | 0.04 | 0.05 | 0.01 | 0.03 | 0.05 | 0.17 | 0.13 | 0.20 | 0.17 | 0.01 | 0.04 | 0.00 | 0.16 | 0.18 | -0.03 | 0.00 | 0.02 | 0.01 | -0.15 | -0.10 | 0.03 | 1.00 | | | | | | |
| 27 Detatched | 0.53 | -0.10 | -0.20 | -0.10 | -0.04 | -0.04 | -0.01 | -0.05 | -0.06 | -0.19 | -0.11 | -0.17 | -0.14 | 0.00 | -0.04 | 0.01 | -0.14 | -0.15 | 0.03 | 0.02 | 0.00 | 0.00 | 0.40 | 0.19 | 0.05 | -0.39 | 1.00 | | | | | |
| 28 Flat | -0.17 | 0.01 | 0.12 | 0.02 | -0.02 | -0.04 | -0.02 | -0.01 | -0.01 | 0.09 | -0.01 | 0.08 | 0.08 | 0.01 | 0.08 | 0.07 | -0.01 | 0.10 | 0.00 | 0.01 | -0.01 | 0.00 | -0.19 | -0.03 | -0.39 | -0.17 | -0.12 | 1.00 | | | | |
| 29 % Owner Occupied (LA) | 0.11 | -0.05 | -0.35 | -0.24 | -0.01 | -0.05 | -0.01 | -0.03 | -0.06 | -0.08 | 0.04 | -0.02 | -0.02 | -0.08 | -0.09 | 0.02 | -0.03 | -0.11 | -0.03 | -0.02 | -0.02 | -0.02 | 0.06 | -0.01 | -0.01 | -0.07 | 0.08 | -0.04 | 1.00 | | | |
| 30 % Privately rented (LA) | -0.01 | 0.07 | 0.36 | 0.11 | 0.05 | 0.06 | 0.02 | 0.01 | -0.03 | 0.05 | -0.01 | 0.09 | 0.02 | 0.16 | 0.09 | 0.01 | 0.08 | 0.11 | -0.02 | -0.01 | 0.00 | -0.02 | -0.02 | -0.04 | 0.07 | 0.08 | -0.09 | 0.02 | -0.28 | 1.00 | | |
| 31 Dwelling Stock (5-year %Δ) | 0.07 | -0.01 | -0.07 | 0.00 | 0.03 | 0.02 | -0.01 | -0.04 | -0.03 | -0.12 | -0.03 | -0.04 | -0.05 | 0.07 | 0.01 | 0.04 | -0.08 | -0.01 | 0.05 | 0.03 | 0.01 | -0.01 | 0.02 | 0.11 | 0.07 | -0.05 | 0.09 | 0.00 | 0.01 | -0.03 | 1.00 | |
| 32 Income LSOA | 0.62 | -0.07 | -0.24 | -0.09 | -0.16 | -0.24 | -0.09 | -0.17 | -0.24 | -0.20 | -0.12 | -0.15 | -0.14 | 0.00 | -0.02 | 0.06 | -0.24 | -0.38 | 0.02 | 0.03 | 0.00 | 0.01 | 0.26 | 0.05 | 0.00 | -0.27 | 0.31 | 0.02 | 0.17 | -0.10 | 0.10 | 1.00 |

Notes: (P) = Potentiality indicator, using an function of journey time; (1) = closest facility; (5) = 5th closest facility; (Avg) = Average among 5 closest facilities; (dist) = distance; (LA) = Local Authority level

# Appendix II: Local Authority dummies

Table II.1: Local Authority dummies (estimates from Model 12b; see Table 4.2)

| LA code | LA name | Est. | t-ratio | Δ%P | LA code | LA name | Est. | t-ratio | Δ%P |
|---|---|---|---|---|---|---|---|---|---|
| E08000012 | Liverpool (Base category) | N/A | N/A | N/A | E06000011 | East Riding of Yorkshire | -0.025 | -3.47 | -3% |
| E08000009 | Trafford | 0.359 | 49.58 | 36% | E08000015 | Wirral | -0.029 | -4.31 | -3% |
| E07000165 | Harrogate | 0.330 | 38.56 | 33% | E07000123 | Preston | -0.031 | -3.92 | -3% |
| E06000014 | York | 0.315 | 39.58 | 32% | E08000005 | Rochdale | -0.036 | -4.85 | -4% |
| E08000007 | Stockport | 0.263 | 41.7 | 26% | E08000013 | St. Helens | -0.040 | -5.11 | -4% |
| E07000031 | South Lakeland | 0.244 | 21.88 | 24% | E08000036 | Wakefield | -0.040 | -6.06 | -4% |
| E08000003 | Manchester | 0.237 | 33.2 | 24% | E07000118 | Chorley | -0.041 | -5.08 | -4% |
| E07000035 | Derbyshire Dales | 0.184 | 13.9 | 18% | E07000030 | Eden | -0.042 | -2.92 | -4% |
| E07000163 | Craven | 0.175 | 16.09 | 18% | E08000037 | Gateshead | -0.051 | -5.88 | -5% |
| E08000035 | Leeds | 0.161 | 28.37 | 16% | E08000018 | Rotherham | -0.053 | -7.09 | -5% |
| E07000167 | Ryedale | 0.153 | 11.09 | 15% | E08000023 | South Tyneside | -0.053 | -5.52 | -5% |
| E08000006 | Salford | 0.147 | 19.09 | 15% | E07000027 | Barrow-in-Furness | -0.065 | -5.79 | -7% |
| E07000168 | Scarborough | 0.144 | 15.45 | 14% | E08000010 | Wigan | -0.067 | -9.9 | -7% |
| E07000164 | Hambleton | 0.144 | 12.34 | 14% | E06000057 | Northumberland | -0.078 | -9.72 | -8% |
| E07000128 | Wyre | 0.126 | 4.52 | 13% | E07000166 | Richmondshire | -0.094 | -4.86 | -9% |
| E08000019 | Sheffield | 0.113 | 18.09 | 11% | E06000009 | Blackpool | -0.096 | -11.69 | -10% |
| E08000008 | Tameside | 0.093 | 13.06 | 9% | E08000016 | Barnsley | -0.107 | -14.45 | -11% |
| E08000002 | Bury | 0.091 | 12.19 | 9% | E06000005 | Darlington | -0.118 | -12.73 | -12% |
| E07000038 | North East Derbyshire | 0.087 | 8.88 | 9% | E08000017 | Doncaster | -0.119 | -16.5 | -12% |
| E06000007 | Warrington | 0.070 | 10.12 | 7% | E06000006 | Halton | -0.131 | -16.48 | -13% |
| E07000034 | Chesterfield | 0.069 | 8.15 | 7% | E06000008 | Blackburn with Darwen | -0.134 | -14.64 | -13% |
| E07000119 | Fylde | 0.063 | 6.68 | 6% | E07000028 | Carlisle | -0.137 | -14.43 | -14% |
| E06000049 | Cheshire East | 0.059 | 8.78 | 6% | E07000033 | Bolsover | -0.139 | -13.89 | -14% |
| E08000034 | Kirklees | 0.049 | 7.57 | 5% | E07000120 | Hyndburn | -0.142 | -12.58 | -14% |
| E08000021 | Newcastle upon Tyne | 0.043 | 5.42 | 4% | E07000117 | Burnley | -0.143 | -13.3 | -14% |
| E07000169 | Selby | 0.033 | 3.76 | 3% | E08000024 | Sunderland | -0.156 | -19.03 | -16% |
| E07000124 | Ribble Valley | 0.031 | 3.02 | 3% | E07000125 | Rossendale | -0.168 | -15.4 | -17% |
| E06000050 | Cheshire West and Chester | 0.028 | 4.32 | 3% | E06000002 | Middlesbrough | -0.170 | -20.02 | -17% |
| E08000022 | North Tyneside | 0.020 | 2.49 | 2% | E07000171 | Bassetlaw | -0.177 | -20.46 | -18% |
| E08000032 | Bradford | 0.020 | 3.05 | 2% | E06000003 | Redcar and Cleveland | -0.184 | -19.07 | -18% |
| E07000127 | West Lancashire | 0.019 | 1.95 | 2% | E06000010 | Kingston upon Hull, City of | -0.191 | -25.27 | -19% |
| E08000033 | Calderdale | 0.007 | 0.91 | 1% | E07000122 | Pendle | -0.199 | -17.28 | -20% |
| E08000014 | Sefton | 0.002 | 0.22 | 0% | E07000026 | Allerdale | -0.214 | -15.72 | -21% |
| E08000004 | Oldham | -0.007 | -0.85 | -1% | E06000004 | Stockton-on-Tees | -0.231 | -28.67 | -23% |
| E07000121 | Lancaster | -0.011 | -1.21 | -1% | E06000012 | North East Lincolnshire | -0.239 | -28.31 | -24% |
| E08000011 | Knowsley | -0.013 | -1.54 | -1% | E06000047 | County Durham | -0.268 | -35.34 | -27% |
| E08000001 | Bolton | -0.014 | -2.11 | -1% | E06000013 | North Lincolnshire | -0.276 | -34.07 | -28% |
| E07000126 | South Ribble | -0.014 | -1.89 | -1% | E07000029 | Copeland | -0.298 | -20.82 | -30% |

# References

Adair, A., McGreal, S., Smyth, A., Cooper, J. and Ryley, T. (2000). House prices and accessibility: The testing of relationships within the Belfast urban area. *Housing Studies,* 15(5), 699-716.

Ahlfeldt, G.M. (2013). "If we build it, will they pay? Predicting property price effects of transport innovations". *Environment and Planning A,* 45(8), 1977-1994.

Ahlfeldt, G.M. and Holman, N. (2018). Distinctively different: a new approach to valuing architectural amenities. *The Economic Journal*, 128(608), 1-33.

Barker, K. (2003). *Review of Housing Supply, Interim Report – Analysis, Securing our Future Housing Needs.* London: HM Treasury.

Batley, R., Bates, J., Bliemer, M. et al. (2017). New appraisal values of travel time saving and reliability in Great Britain. *Transportation*. https://doi.org/10.1007/s11116-017-9798-7

Bayer, P., Ferreira, F. and McMillan, R. (2007). A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy*, 115(4), 588-638.

Ben-Akiva, M. and Lerman, S.R. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand.* Cambridge, Massachusetts: MIT Press.

Buck, M. (2017). Crossrail project: finance, funding and value capture for London's Elizabeth line. *Proceedings of the Institution of Civil Engineers*, 170(6), 15-22.

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

Coppola, P., Ibeas, Á., dell'Olio, L. and Cordera, R. (2013). LUTI model for the metropolitan area of Santander. *Journal of Urban Planning and Development*, 139(3), 153-165.

Cortright, J. (2009), *Walking the Walk: How Walkability Raises Housing Values in U.S. Cities*, Washington, DC: CEOs for Cities.

Debrezion, G., Pels, E. and Rietveld, P. (2007). The impact of railway stations on residential and commercial property value: A meta-analysis. *The Journal of Real Estate Finance and Economics*, 35(2), 161-180.

Defra (2015). Round 2 Strategic Noise Mapping – Data Production Process. Technical Report P054. [Online]. [Accessed 14.8.2019].
randd.defra.gov.uk/Document.aspx?Document...ENDR2Dataproductionprocessfinal.pdf

DfT (2013). *The Transport Business Cases.* London: DfT. [Online]. [Accessed 14.8.2019]. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/85930/dft-transport-business-case.pdf

Department for Transport (DfT) (2016), *Understanding and Valuing Impacts of Transport Investment: updating wider economic impacts guidance.* London: DfT.

DfT (2017a). Strategic Case Supplementary Guidance – Transport Investment Strategy.

DfT (2017b). *Value for Money Supplementary Guidance on Categories*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/627490/value-for-money-supplementary-guidance-on-categories.pdf

DfT (2018). Transport Analysis Guidance (TAG). Units A1.3, A2.1-2.4, M5.3. https://www.gov.uk/guidance/transport-analysis-guidance-webtag

Dziauddin, M.F., Powe, N., and Alvanides, S., (2015), Estimating the Effects of Light Rail Transit (LRT) System on Residential Property Values Using Geographically Weighted Regression (GWR), *Applied Spatial Analysis*, 8, 1.25.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). *Geographically weighted regression: the analysis of spatially varying relationships*. Chichester: Wiley.

Geurs, K.T. and Van Wee, B. (2004). Accessibility evaluation of land-use and transport strategies: review and research directions. *Journal of Transport Geography*, 12, 127-140.

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

Gibbons, S. & Machin, S. (2008). Valuing school quality, better transport, and lower crime: evidence from house prices. *Oxford Review of Economic Policy*, 24(1), 99-119.

Gibbons, S., Lyytikainen, T., Overman, H. G., & Sanchis-Guarner, R. (2012). New road infrastructure: the effects on firms. http://eprints.lse.ac.uk/58527/1/__lse.ac.uk_storage_LIBRARY_Secondary_libfile_shared_repository_Content_SERC%20discussion%20papers_2012_sercdp0117.pdf

Graham, D.J. (2007). Agglomeration, Productivity and Transport Investment, *Journal of Transport Economics and Policy*, 41(3), 317-343.

Graham, D.J., Gibbons, S, and Martin, R (2010). *The spatial decay of agglomeration economies: estimates for use in transport appraisal.* Department for Transport.

Greater London Authority (2010). Working Paper 42: Valuing housing and green spaces: Understanding local amenities, the built environment and house prices in London.

HM Treasury (2018). *The Green Book: Central Government Guidance on Appraisal and Evaluation.* London: TSO.

Handy, S., and Niemeier, D. (1997). Measuring accessibility: An exploration of issues and alternatives. *Environment and Planning A,* 29, 1175-94.

Hanushek, E. A. (2003). The failure of input-based schooling policies. *The Economic Journal*, 113(485), F64-F98.

He, S., Tao, S., Hou, Y. and Jiang, W. (2018). Mass Transit Railway, transit-oriented development and spatial justice: The competition for prime residential locations in Hong Kong since the 1980s. *Town Planning Review*, 89(5), 467–493.

Higgins, C.D. (2019). A 4D spatio-temporal approach to modelling land value uplift from rapid transit in high density and topographically-rich cities. *Landscape and Urban Planning*, 185, 68-82.

House of Commons (2018). *Land Value Capture*. Tenth Report of Session 2017–19. Housing, Communities and Local Government Committee. London: TSO.

Ibanez, M.R. and Pennington-Cross, A. (2013). Commercial Property Rent Dynamics in U.S. Metropolitan Areas: An Examination of Office, Industrial, Flex and Retail Space, *The Journal of Real Estate Finance and Economics,* 46(2), 232–259.

Kenny, G. (1999). Modelling the demand and supply sides of the housing market: evidence from Ireland. *Economic Modelling*, 16(3), 389-409.

Kong, F., Yin, H. and Nakagoshi, N. (2007). Using GIS and landscape metrics in the hedonic price modeling of the amenity value of urban green space: A case study in Jinan City, China. *Landscape and Urban Planning*, 79(3-4), 240-252.

Koster, H.R.A., van Ommeren, J. N., & Rietveld, P. (2016). Historic amenities, income and sorting of households. *Journal of Economic Geography*, 16(1), 203-236.

Koster, H.R.A. & van Ommeren, J. (2019). Place-based policies and the housing market. *Review of Economics and Statistics,* 101(3), 400-414.

Kuminoff, N. V., Smith, V. K., & Timmins, C. (2013). The new economics of equilibrium sorting and policy evaluation using housing markets. *Journal of Economic Literature*, 51(4), 1007-62.

Lee, S.S. and Senior, M.L. (2013), Do light rail services discourage car ownership and use? Evidence from Census data for four English cities. *Journal of Transport Geography*, 29 (May), 11-23.

Leeds City Council (2017). East Leeds Orbital Road – Land Assembly & Financing. Report to Executive Board, 21 June 2017. [Online]. [Accessed 26.6.2019]. https://democracy.leeds.gov.uk/documents/s162232/ELOR%20Cover%20Report%20120617.pdf

LeSage, J. and Pace, R.K. (2009). *Introduction to Spatial Econometrics.* New York: CRC Press

Manchester City Council (2018). Delivering Manchester's Affordable Housing Strategy – update on the delivery of the Council's affordable housing strategy. Report for Information, 5 Dec 2018. [Online]. [Accessed 21.6.2019]. Available from: https://democracy.manchester.gov.uk/documents/s2929/Delivering%20Manchesters%20Affordable%20Housing%20Strategy%20Update%20on%20the%20delivery%20of%20the%20Councils%20afford.pdf

Ministry of Housing, Communities & Local Government (MHCLG) (2018). *Land Value Estimates for Policy Appraisal.* London: MHCLG.

Millard, T., Nellthorp, J. and Ojeda Cabral, M. (2018). 'What is the value of urban realm? - a cross-sectional analysis in London'. Paper presented at the International Transportation Economics Association Conference in Hong Kong, 27-29 June 2018.

Mohammad, S.I., Graham, D.J., Melo, P.C. and Anderson, R.J. (2013). A meta-analysis of the impact of rail projects on land and property values. *Transportation Research Part A: Policy and Practice*, 50, 158-170.

Mulley, C. (2014). Accessibility and Residential Land Value Uplift: identifying Spatial Variations in the Accessibility Impacts of a Bus Transitway. *Urban Studies*, 51(8), 1707–1724.

Mulley, C. (2017). 'What do we know about value uplift and what sort of uplift do we see in Australia?'. Presentation slides.

Mulley, C., Tsai, C.-H. and Ma, L. (2018). Does residential property price benefit from light rail in Sydney? *Research in Transportation Economics,* 67, 3-10.

Murphy, L. (2018). *The Invisible Land: The hidden force driving the UK's unequal economy and broken housing market.* London: IPPR.

NCHRP (2018). *Guidebook to Funding Transportation Through Land Value Return and Recycling,* Research Report 873. Washington: TRB.

Nationwide (2014). *House Price Index Special Report – June 2019: Tram & rail links in Greater Manchester attract premium amongst homebuyers.* Swindon: Nationwide Building Society.

Nationwide (2018). House Price Index. [Online]. [Accessed 22.12.2018]. Available from: https://www.nationwide.co.uk/about/house-price-index/download-data#tab:Downloaddata

Nationwide (2019). *House Price Index Special Report – June 2019: Londoners pay highest premium to live near a station, but tram/rail links also attractive to homebuyers in Manchester.* Swindon: Nationwide Building Society.

Nellthorp, J., Bristow, A.L. and Day, B. (2007), Introducing willingness-to-pay for noise changes into transport appraisal: an application of benefit transfer. *Transport Reviews,* 27(3), 327-353.

Nellthorp, J. (2016). *Business Case Development Manuel Review: Stage 2 Urban realm.* Report to Transport for London. Leeds: ITS.

Nellthorp, J., Johnson, D., Ojeda Cabral, M. and Laird, J. (2016). *Transport and Land Value Uplift: Evidence and Implications for Appraisal, Modelling and Strategy*. Briefing Note. Phase 1 Study sponsored by the West Yorkshire Transport Research Innovation Fund. Leeds: ITS.

Nellthorp, J., Ojeda Cabral, M., Johnson, D. and Leahy, C. (2018). *Transport and Land Value: Modelling and Appraisal – Advisory Panel Discussion Paper 1*. Leeds: ITS.

Nikolopoulou, M., J. Kleissl, P.F. Linden, S. and Lykoudis, S. (2011). Pedestrians' perception of environmental stimuli through field surveys: focus on particulate pollution. *Science of the Total Environment,* 409, 2493–2502.

Northall, S., (2014). Improving the political case for transport investment: an ex-post evaluation of the external economic benefits of the Nottingham Express Transit LRT Scheme, Mott MacDonald TPS Bursary Paper 2014.

Office for National Statistics (ONS) (2015). *Family spending in the UK: calendar year 2014.* Cardiff: ONS.

Office for National Statistics (ONS) (2015). *The UK national balance sheet estimates: 2018.* Cardiff: ONS.

Office for National Statistics (ONS) (2018a). *The UK national balance sheet.* Cardiff: ONS.

Office for National Statistics (ONS) (2018b). *Regional Gross Disposable Household Income (GDHI) by Local Authority.* Cardiff: ONS.

Office of the Deputy Prime Minister (2004). Producing Boundaries and Statistics for Town Centres, England and Wales 2000, Interim Report.

Paas, B., Schmidt, T., Markova, S., Maras, I., Ziefle, M., Schneider, C. (2016). Small-scale variability of particulate matter and perception of air quality in an inner-city recreational area in Aachen, Germany. *Meteorologische Zeitschrift*, 25(3), 305-7.

Panduro, T.E. and Veie, K.L. (2013). Classification and valuation of urban green spaces—A hedonic house price valuation. *Landscape and Urban Planning*, 120, 119-128.

Rail Delivery Group (RDG) (2018). *Passenger Demand Forecasting Handbook (PDFH),* version 6.0.

Powe, N. A., Garrod, G. D. and Willis, K. G. (1995). Valuation of urban amenities using an hedonic price model, *Journal of Property Research.* 12(2), 137-147

Ricardo Energy & Environment (2016). Technical report on UK supplementary assessment under the Air Quality Directive (2008/50/EC), the Air Quality Framework Directive (96/62/EC) and Fourth Daughter Directive (2004/107/EC) for 2014.

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy*, 82(1), 34-55.

Rouwendal, J. (1992). The hedonic price function as an envelope of bid functions; an exercise in applied economic theory. *Netherlands Journal of Housing and the Built Environment*, 7(1), 59-80.

Rouwendal, J. (1998). On housing services. *Journal of Housing Economics*, 7(3), 218-242.

Rouwendal, J. and van Der Straaten, W. (2008). The costs and benefits of providing open space in cities. Tinbergen Institute Discussion Paper TI 2008-001/3. Amsterdam: Tinbergen Institute.

Salon, D., Sclar, E. and Barone, R. (2019). Can Location Value Capture Pay for Transit? Organizational Challenges of Transforming Theory into Practice. *Urban Affairs Review*, 55(3) 743–771.

SDG (2016), *Funding and Financing Public Transport Infrastructure, Report to CBT*. http://cbtthoughtleadership.org.uk/SDGFundingAndFinancingReport.pdf

Simmonds, D.C. (2004). Assessment of UK land-use and transport strategies using land-use/transport interaction models. *European Journal of Transport and Infrastructure Research*, 4, 273-293.

Smith, T., Noble, M., Noble, S., Wright, G., McLennan D. and Plunkett, E. (2015). *The English Indices of Deprivation 2015 - Research report.* London: DCLG.

SQW & CE (2016), *Northern Powerhouse Independent Economic Review*. Report to Transport for the North.

Talen, E. (1997). The social equity of urban service distribution: An exploration of park access in Pueblo, Colorado, and Macon, Georgia. *Urban Geography*, 18(6), 521-541.

Taylor, L. O. (2008). Theoretical foundations and empirical developments in hedonic modeling. In *Hedonic methods in housing markets* (pp. 15-37). Springer, New York, NY.

TfL (2017). *Land Value Capture:, Final Report.* London: TfL. https://www.london.gov.uk/sites/default/files/land_value_capture_report_transport_for_london.pdf

TfL (2019a). Mayoral Community Infrastructure Levy. https://www.london.gov.uk/what-we-do/planning/implementing-london-plan/mayoral-community-infrastructure-levy Accessed 23.7.2019.

TfL (2019b). Barking Riverside extension. https://tfl.gov.uk/travel-information/improvements-and-projects/barking-riverside-extension Accessed 23.7.2019.

Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of Political Economy*, 64(5), 416-424

Tsai, I. C. (2018). The cause and outcomes of the ripple effect: housing prices and transaction volume. *The Annals of Regional Science*, 1-23.

Tse, R. Y., Ho, C. W., & Ganesan, S. (1999). Matching housing supply and demand: an empirical study of Hong Kong's market. *Construction Management & Economics*, 17(5), 625-633.

Tyndall, J. (2017). Waiting for the R train: Public transportation and employment. *Urban Studies* 54, 520–537.

Van Wee, B. (2016). Accessible accessibility research challenges. *Journal of Transport Geography*, 51, 9-16.

Venables A.J., Laird J.J. and Overman H.G. (2014). *Transport investment and economic performance: Implications for project appraisal.* London: DfT.

Xue, Y. (2016), The economics of residential property prices: a case study in Leeds Region. MSc Dissertation, Institute for Transport Studies, University of Leeds.

UNIVERSITY OF LEEDS

ITS

www.its.leeds.ac.uk