Title:

How does the extension of existing transport infrastructure affect land value? A case study of the Tyne and Wear Light Transit Metro system.

Abstract

This paper seeks to investigate the timing of land value uplift associated with an extension to an existing light rail transport system. It seeks to understand the amount of land value increase on residential properties at different stages of the infrastructure delivery process of the Tyne and Wear Metro (TWM) system in the North East of England. This is to shed light on the potential use of land value capture methods for subsequent extensions. Whilst there has been substantial international research into transport related land value capture methods, this tends to take place in atypical capital cities with buoyant land markets and/or entire transport systems. This is reflected in England, where research is less frequent in major conurbations outside of Central London, which are more typical of urban areas in this country. In response, the intention is to shed new light on this situation by updating historical research into the impact of the TWM. Its primary focus is the longitudinal performance of the 2002 Sunderland Metro Extension (SME). This historical appraisal of impact is important to understand when and how much uplift is generated. This in turn helps to evidence any justifications for subsequent extensions and funding via Land Value Capture. This study uses a time based analysis to quantify differences in property prices before, immediately after and fifteen years after the Sunderland Metro Extension (SME) was constructed. The results show that property values increased immediately after the extension becomes operational, though no significant results were found 15 years later.

Key words: Land value uplift, land value capture, transport accessibility, metro and light rail, Longitudinal study.

1. Introduction and rationale for study

Demand for mobility continues to increase in the international urban climate. Increasingly, locations are seeking to extend existing transport infrastructure to connect, and sustain, functional economic areas. However, most infrastructure projects are not profit making. Traditionally, investment in new transport infrastructure has required substantial funds from central government to be viable (Du and Mulley, 2007b; Pike et al, 2019). However, this situation exists against a broad backdrop of reduced government funding due to austerity and changing government ideology associated with reduced degrees of state intervention (Muldoon-Smith and Sandford, 2021). Although there are exceptions to this rule, for example, the Belt and Roads Initiative in China, 4G Roads and Highway Initiative in Columbia, and Nigeria's Lagos-Kano Railway.

In England, transport infrastructure has traditionally been funded through Central Government grant and traditional taxation methods. Over 85% of transport infrastructure in the pipeline for 2016-2021 is being funded via these methods. However, this will be the last era of this type of funding with state influence, particularly in provincial locations, reducing dramatically under narratives of devolution, fiscal decentralisation and more recently the impacts of Covid (Muldoon-Smith, 2019, Institute for Fiscal Studies, 2020). Already, in recent years, subdued economic growth has limited the availability of existing taxation for public transport infrastructure because it is difficult to justify the enactment of increased taxes (PwC, 2018). The increasing challenge for national, regional, and local governments (particularly in mature urban locations) is to explore and utilise more entrepreneurial and self-sustaining funding models (Fuller, 2018; Pike et al, 2019).

Consequently, land value capture (LVC) mechanisms have become an area of increasing interest. For example, as a means of funding public transport infrastructure through new tax revenue associated with additional value, as opposed to increased taxation associated with existing assets. In short, the premise of LVC is to utilise the increased value of nearby land because of increased accessibility through investment in transport infrastructure. In 2016, it was estimated that landowners pocketed £9.3 billion of the £12.3 billion of land value uplift that was generated via infrastructure improvements (CPRE, 2017). LVC methods, in the main, can be separated into two forms. The first is related to development rights around transport sites, which can be either sold or leased to generate and capture value. The second is typically through taxation of properties situated on land, where tenants or landlords who benefit from new infrastructure pay additional contributions. This additional taxation can be captured in a one-off payment or continually depending on the duration of the value uplift and the pragmatic choice of the local government. In England this would be via Council Tax (for residential properties) or Business Rates (for commercial and industrial properties).

An understanding of when land value increase occurs upon residential properties at different stages of the infrastructure delivery process is essential (Mohammad et al, 2013). This is because any design of an appropriate LVC mechanism should be predicated by an identification of when, and how, much uplift is generated. This is perhaps the most crucial aspect of LVC mechanisms. As the LVC must capture the uplift at the correct time to enable a successful funding strategy. For example, mechanisms implemented after construction start may miss a significant portion of the generated uplift, as value uplift effects can occur as early as a government's commitment to a scheme. Studies that have been undertaken seem to have a degree of disagreement (Yen et al, 2018) between whether value uplift occurs before construction (McDonald and Osuji, 1995), immediately afterward (Mohammed et al, 2013) or potentially sometime later once the infrastructure is established (Du and Mulley, 2007b).

Light rail (the category within which the TWM resides), has been shown to exhibit positive uplifts in both residential and commercial properties within the catchment area of stations (Cervero, 2004; Duncan, 2008; Pan and Zhang, 2008; Xu et al, 2016). However, carrying on the theme of lack of consensus, literature is full of studies failing to find any uplift (Hass-Klau et al, 2004), and even negative impact from light rail stations (Bollinger et al, 1998; Weinstein and Clower, 1999).

However, in England, most of the research has taken place in and around the atypical Central London land market where underlying land value is buoyant and land markets are responsive (Li and Love, 2020). Previous studies have found positive effects on property value from the London Victoria line, the Jubilee Line, the Channel Tunnel Rail Link and substantial positive effects from the Docklands Light Rail (see Chesterton, 2000 for a meta study of these lines), causing increases from £70,000 an acre in 1981, to £4.9 million an acre in 1988. There is less research into suboptimal, but more typical, locations outside of Central London that have land markets which are less buoyant and potentially less responsive to change (Crocker et al, 2000; Hass-Klau et al, 2004). Indeed, little evidence of land value uplift from new light rail infrastructure has been identified. This is an important gap in knowledge, as O' Brien and Pike (2015) note the increased importance of regional infrastructure outside of central locations which is needed to sustain economic development. While (Plimmer, 2017) notes the importance, but also difficulty, of funding infrastructure in regional and sub-regional locations.

Early studies on the TWM (Pickett, 1984), Manchester Metrolink (Forrest et al, 1996) and Sheffield Supertram (Crocker et al, 2000) failed to find evidence of land value uplift because of light rail infrastructure. However, more recent studies by Du and Mulley (2007; 2007b; 2012) did identify positive premiums on land value on the TWM, but, importantly, not around the extension. It can be theorised that such locations may take longer to register impact from new infrastructure or that the impact could be less pronounced in less buoyant locations. For example, in 1984, no significant land value uplift was found along the original TWM line. However, in the subsequent studies in 2006 and 2007, significant land value uplift was identified around some of the stations which had been served by the original line since 1980 (for 27 years).

This suggests that the beneficial impacts from metro systems may take rather a long time to occur outside of primary economic areas. As such, the empirical research presents the opportunity to test this theory, as the SME has now been in operation for approximately 17 years at the time of study. Building upon Du and Mulley (2007b) findings, this study aims to consider the longer-term impacts of the Sunderland Metro Extension (SME), to deduce whether extensions to long standing metro systems can have a positive impact on land value. In addressing this, a case study of the TWM system in North East England incorporating a time-based analysis of additional census points was constructed. This is to shed new scientific light on the funding of extensions to existing transport systems and the lack of consensus around the timing of value uplift in the academic community.

The remainder of this paper presents the underlying methodology for the paper, detailing the case study and hypothesis, time frame, method design and limitations. It then considers the results, testing the model, presenting findings and implications. The final part of the paper revisits the hypothesis and presents limitations to the study and some opportunities for further research.

2. Methodology

2.1 Case Study and Hypothesis

The hypotheses to be tested is:

'Land value uplift from metro extensions does occur over a fifteen-year period after the system becomes operational.'

The original research of Du and Mulley (2007b) predominantly concentrated on where land value capture occurred, this hypothesis focuses on when value capture occurred (Pagliara and Papa, 2011). This focus is adopted to test the contention of Du and Mulley (2007a; 2007b) that value change in the SME may take longer to materialise.

In order to carry out this research, quantitative data was collected relating to the TWM system. The system is operated and maintained by Nexus, the region's key public transport provider. It connects Newcastle upon Tyne, Sunderland, Gateshead and South Tyneside. The metro system was planned in the 1970's and the system's structure was mostly built at this time, whilst incorporating much of the existing Victorian rail infrastructure in the region. The system was opened in 1980 (the "Original Line"), followed by an extension to Newcastle International Airport in 1991. The system was further extended to South Tyneside and Sunderland, which was constructed and then opened in June 2000 and 2002 (the "Sunderland Extension"), respectively. At the time of study (2018), the original line has been in operation for 38 years and the Sunderland Extension has been operated for 16 years. A map of the Tyne and Wear Metro System is exhibited as Figure 1.

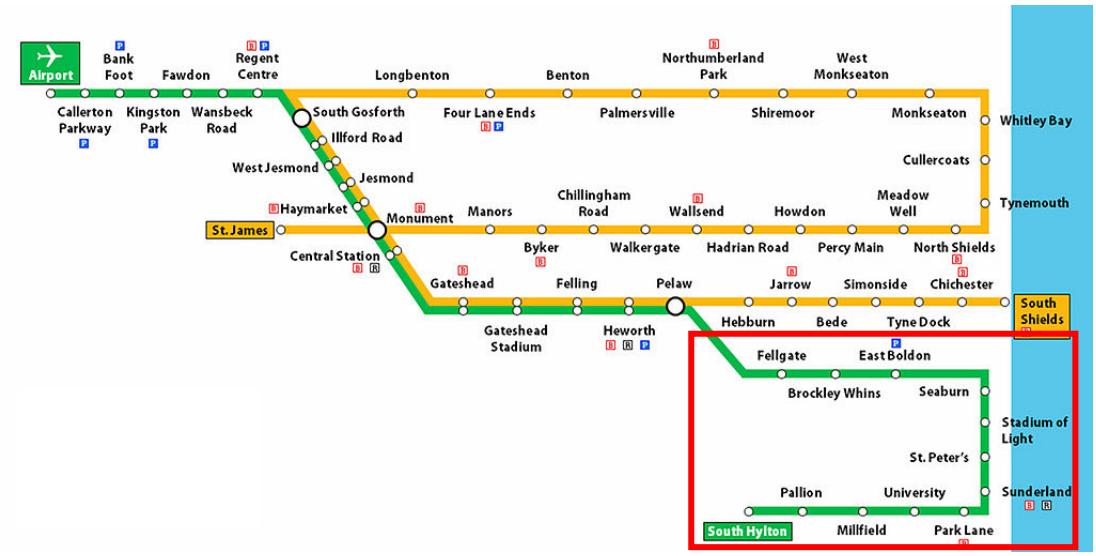


Figure 1. Map of the Tyne and Wear Metro (i.e., the Sunderland Metro Extension outlined in red)

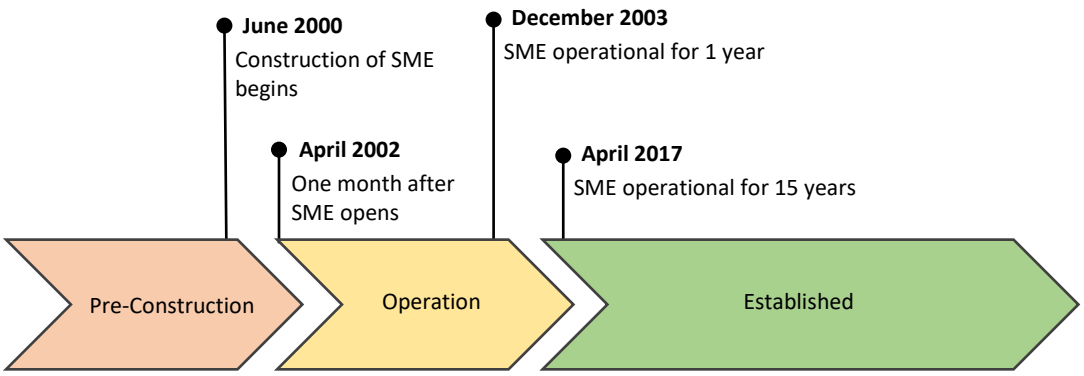
To do this, we use a hedonic 'global' model to assess the impact on the SME. Within this choice of method, we recognise the recent work engaging with the issue of local spatial variance – that is insensitivity to impact of specific variables at different points on the transport line (Dziauddin, 2015). As value uplift research has continued to advance, recent studies have sought to trial new methods of measuring the impact of transport accessibility on local land values (Yenn et al., 2018). These have included:

- Billings (2011) – incorporating a richer set of control variables in analyses.
- Cao and Porter-Nelson (2016) – location quotients to address spatial dependence.
- Du and Mulley (2007; 2007b) – geographic weighted regression to analyse spatial impacts.
- Yenn et al. (2018) – difference in difference model to analyse spatial impacts.

However, the focus of this paper is not on the 'local' impact upon where value uplift occurred. Rather, the intention is to reflect upon when value uplift occurred across the current lifetime of the SME. While we concede the volatility in hedonic models (Lieske et al, 2019). One of the ways of improving this situation, especially in time-based studies, is to improve the underlying transaction data and incorporate a systematic time element into it. This is to present an initial staging post for time-based enquiries into infrastructure extensions while remaining consistent with the original method used by Du and Mulley (2007a). Consequently, the authors are careful to be sceptical in their approach to the findings, conceding that there is scope for variance and bias within the global model. Indeed, there is scope to make further methodological innovation, and these options are considered in the concluding section as opportunities for further research. For example, in relation to spatial variance, impact of scale, urban design features and the incorporation of commercial real estate stock.

2.2 The Hedonic Model and Timeline

The timeline of study was selected to examine the lack of consensus around when value uplift occurred (Figure 2). Data was collected across all stations along the Sunderland Metro Extension. As this study utilises a global hedonic model (as opposed to a local model), the selection of the stations themselves are not explicitly important. The most important factor is the establishment of a property transaction data set that reflects time, supported by a suitable variety of data for the independent variables. As this will explore the interrelationship that exists between land value and the exogenous/ endogenous property factors.



NB: the underlying data for this study was captured at the following time periods. 1) Census in 2011 2) House price transaction data from Land Registry recorded in the following three census points:1999-2000, 2002-2003 and 2017-18.

Figure 2. Indicative timeline displaying the key events leading to the construction and operation of the Sunderland Metro Extension.

Hedonic price model is applied for data analysis and the SPSS software package was adopted to identify the significance of the results through the analysis of variance. All stations along the Sunderland metro extension line were selected and property transaction data was collected within a catchment zone of 800 metres. The independent variables in the study were limited to property type, number of bedrooms (Sirmans et al, 2006), percentage of unemployment and higher managerial positions (Office for National Statistics, 2020), and distance to Sunderland and Newcastle City Centres. This conscious limitation occurred because the intention was to focus on the best fit of data to the catchment area. For example, the Census based 'unemployment' and 'higher managerial' figures were the only ones available at the catchment area geography. Rather than adding additional proxy data sets into the model (with various data attributes and granularities). While there were additional variables that may impact property value, they were only available at the municipal level, while information on 'age' and 'quality' of property does not exist. However, this is returned to in the

conclusion where the potential for the inclusion of additional independent variables is explored (for example proximity to green space and urban design features) alongside methodological innovation.

Thus, as put forward by Du and Mulley (2007b, p.12), property prices are a function of a group of variables:

$$P_i = f\,(H, N, T) \qquad \textit{Eq. (1)}$$

Where H, N and T is a vector of – (H) property features, (N) neighbourhood environment and (T) transport accessibility.

In the hedonic model, it is necessary to define the catchment areas for the residential property transaction data, as this represents the properties that have access to the metro extension. Figure 3 shows the radius within which the transacted property price data was collected within the identified timeline. Similar to many previous studies, the property transaction data located within 800m of the selected stations have been grouped into the catchment area. This selection method is based on two key themes that have emerged from the literature: first, that 800m is a walkable catchment area for rail based public transport (Yen et al, 2018) and second, that property prices appear to be greatest up to a distance of 800m from the transport station (Mohammad et al, 2013). As illustrated by the figure, the catchment areas overlap in Sunderland City Centre, providing an almost linear catchment area.

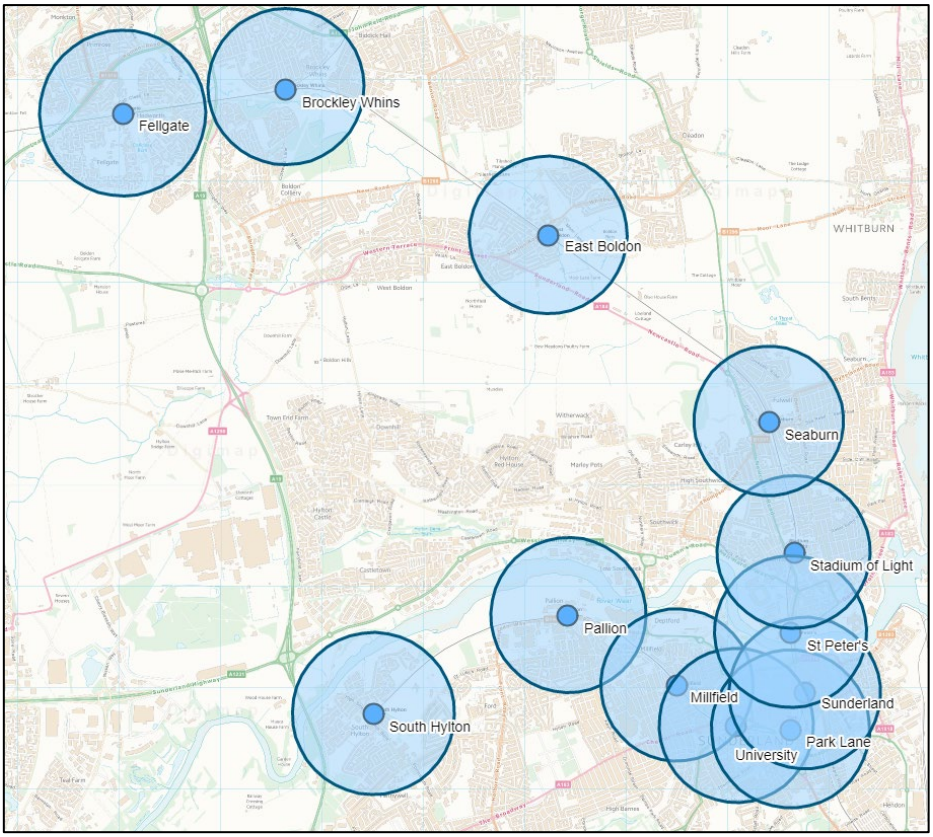


Figure 3. Map showing the stations of the Sunderland Metro Extension and the catchment area of each station

3. Data and Variables

3.1 Data Description

Table 2 summaries the details of each variable included in the model. The following subsections then present information of each variable in turn, outlining the data and the collection method. Additionally, Table 3 was also developed to show the price variability of different types of properties across the three delivery stages of the assets. Across the different unit types, the variability ranges from 30.85% to 46.09%, indicating that the property price data is dispersed approximately 32-48% around the mean values. This high degree of variability can be explained by the number of bedrooms, percentage of unemployment, percentage of higher managerial positions, distance from Sunderland City Centre and distance from Newcastle City Centre. Taking a 2-bedroom terraced property in areas of 0-2.9% unemployment as an example. The variability is calculated 18.23%, which is less than half of the lowest variability attributed to terraced housing. The remaining 18.23%, discussed in proceeding sections, can be attributed to the other independent variables in the model.

Table 2. Summary of the variables collected for the Hedonic Price Model

Vector	Variable	Type	Description	Data Source
Land Value (P_i)	PRICE	Continuous	Transacted price adjusted by House Price Index (HPI)	Land Registry
Property (P)	BEDROOM	Continuous	Number of bedrooms	Land Registry
	FLAT	Dichotomous	=1 if flat; =0 otherwise	
	TERR	Dichotomous	=1 if terraced; =0 otherwise	
	SEMI	Dichotomous	=1 if semi-detached; =0 otherwise	
	DETA	Dichotomous	=1 if detached; =0 otherwise	
	BUILD	Dichotomous	=1 if new build; =0 otherwise	
Neighbourhood (N)	%UNEMP	Categorical	% of long term unemployment. (1 = 0%-2.9%; 2 = 3%-5.9%; 4 = 6%-8.9%; 5 = 9-11.9%; 6 = 12%+)	2011 census
	%HMAN	Categorical	% of higher managerial positions. (1 = 0%-2.9%; 2 = 3%-5.9%; 4 = 6%-8.9%; 5 = 9-11.9%; 6 = 12%+)	
Amenities (A)	distSUN	Categorical	Distance from Sunderland City Centre (1 = 0 mi–0.9 mi; 2 = 1 mi–1.9 mi; 3 = 2 mi–2.9 mi; 4 = 3 mi+)	GSI
	distNEW	Categorical	Distance from Newcastle City Centre (1 = 9 mi–9.9 mi; 2 = 10 mi–10.9 mi; 3 = 11 mi–11.9 mi; 4 = 12 mi+)	
Transport accessibility (T)	extPRE	Dichotomous	Property sold between Jan 1999 and Mar 2000 (1=yes; 0=no)	Land Registry
	extOPE	Dichotomous	Property sold between Apr 2002 and Dec 2003 (1=yes; 0=no)	
	extEST	Dichotomous	Property sold between Apr 2017 and Aug 2018 (1=yes; 0=no)	

Table 3. Variability in price of unit types at different infrastructure delivery phases

Unit	extPRE		extOPE		extEST	
	StdDev	Variability	StdDev	Variability	StdDev	Variability
Flat	31903	32.41%	40624	42.68%	29806	39.46%
TERR	38806	39.86%	44786	44.41%	44645	46.09%
SEMI	42780	30.85%	57336	39.47%	49844	34.32%

3.2 The Dependant Variable

The transacted property price is the dependent variable in the hedonic price model as gathered from the Land Registry, a database holding all residential transactions in England. Land value is therefore represented in the transacted value of property. This is because, in England, land value (and tax) is a capitalised representation of the buildings constructed on top of it. As opposed to other countries that may utilise a purely land value tax system or split rate system combining differential tax rates for land and property. Therefore, it is important to note, throughout the rest of the paper, that residential

transaction value (of property) is taken as a reflection of land value in the English institutional context. This data set is one of the key opportunities for improvement on previous research. In property related analysis, the transacted sale price is considered as the ideal data to represent the property's value as it provides the true market value of the property (or in this case, the land) at the date of sale, giving the most accurate representation of the land's value (RICS, 2019). In this case the transaction related data set is also the most recent data available, b) has greater longitude and C) is the most accurate. The previous research from Du and Mulley (2007; 2007b) used an online website which listed properties for sale. The data was therefore partial because it only included those properties deposited on the website and limited by only detailing advertised price which includes expectation of value rather than real value – this is the least representative measurement of value (RICS, 2019). An alternative method would have been to use Council Tax bands (the basis for residential property tax in England) however these are based on out-of-date valuations (decades old).

All properties sold between January 1999 to June 2000, April 2002 to December 2003, and April 2017 to August 2017 within the station catchment areas were used for the time-based analysis. This provided 2957 records of transacted properties.

3.3 Property Data

The data for all property transactions in Sunderland since 1995 was downloaded from the database and imported to excel, which was then filtered by date to give property transactions within the three timeline phases. A House Price Index has been introduced to adjust for the price change over time due to fluctuations in the property market and due to inflation. This was calculated using January 1999 as a base and taking the average house price sale of each month within the time frame of the study. The postcode of each station was then entered into a postcode matching tool that identified all of the property postcodes within 800m. This was then used to filter out all properties outside of the catchment areas. Three key property attributes were gathered from the Land Registry database for the analysis including the number of bedrooms, the property type (flat, terraced, semi-detached or detached) and the build (whether the property is a new build). These attributes cover some of the key endogenous factors that are used in the marketing of residential property, which represents the most 'desirable' factors that may influence the value of the property. The property size is therefore limited and captured by the number of bedrooms variable. The property type and status of build are controlled by dichotomous/dummy variables, where 1 = present and 0 = not present. The variables FLAT, TERR and SEMI represent a subgroup of variables relating to the property type, wherein 1 represents a property of that type, e.g. a property with the TERR variable equal to 1 would represent a terraced property.

3.4 Neighbourhood Data

The neighbourhood data is used to capture the external characteristics of the area that may influence the transacted price of the properties. The purpose is to try to standardise these characteristics, so the effect of transport accessibility can be isolated and observed. The data selected for this study is the percentage of people 'Long term unemployed' and the percentage of people in 'Higher managerial and professional occupations'. These are grouped into five (5) categories: 1 (0% - 2.9%), 2 (3% - 5.9%), 3 (6% - 8.9%), 4 (9% - 11.9%) and 5 (12% +). As in the study by Du and Mulley (2007b), these two factors capture the household status of the area.

These attributes provide an indication of the area's demographic classification, as they are one of the most widely-used standard socio-economic classifications in the UK. These variables are derived from the Office for National Statistics (ONS) UK census at the electoral ward level for the year 2011. Similar

to the property attributes, a batch geo-coding tool was used to identify the postcodes within each ward, and then matched with the postcodes of each property transaction to attribute the values accordingly.

3.5 Amenities

Access to amenities has been captured by the distance from the property to the Central Business District (CBD), which has been considered as Newcastle City Centre and Sunderland City Centre. The data was grouped into four (4) categories, with each representing a 1-mile distance from the CBD. Following the bid-rent model, these are the destinations that most people living in the Tyne and Wear area are likely to require access to, as they are major employment areas and 'commercial hot spots'. These are basic variables to assess the impact of proximity to the CBD, measured by taking the distance from the postcode of each property to the postcode of the city centre, based on the longitude and latitude co-ordinates of each. It would be expected that there is a varying relationship between the amenities variable and house prices at different stages of the project. Prior to construction, property values would be expected to be higher as distance decreases as there is no metro link to provide increased accessibility. After the extension becomes operational, the magnitude of this relationship would be expected to decrease, as the increased accessibility would make land outside of the city more desirable by the bid-rent model.

3.6 Transport Accessibility

As previously identified, three separate periods are used in the investigation to represent the extent and nature of accessibility to a metro station, each representing a phase of the infrastructure delivery process within which land value uplift may occur. This has given rise to 3 (three) separate dummy variables, identified as extPRE (Pre-Construction), extOPE (Operation) and extEST (Established), which represent a subgroup denoting the phase of infrastructure delivery process. In other words, we present when there is no transport access (extPRE), whether there is transport access (extOPE) and whether there is established transport access (extEST). ExtOPE is the reference point for this analysis, where we can observe the results with no transport accessibility (extPRE) and with established transport accessibility (extEST) in comparison.

The interaction terms of the transport accessibility are used to measure the difference in property price between properties in the catchment area and given as a dichotomous value, where 1=present and 0=not present. There is only one catchment area, therefore there are 3 interaction terms (properties within 800m in the years 1999-2000, 2002-2003 and 2017-2018). As each property transaction inputted into the model has to have an extPRE, extOPE and extEST designation, we use '1' or '0' to indicate which stage the property is at. E.g. if the property was transacted in 1999 then extPRE = 1, extOPE = 0 and extEST = 0. If transacted in extOPE then extPRE = 0, extOPE = 1 and extEST = 0.

4. Findings

The results from the multiple regression model indicate that there is a clear positive relationship between the SME and property values within the catchment area. Immediately after the extension became operational, property values displayed an increase of £10,750 from the pre-extension values. However, once the extension had been operational for fifteen years, property values showed an increase of only £3,654 in comparison with the pre-extension values, which is a decrease of £7,056 in comparison with the values displayed immediately after the extension became operational. However, these latter results were not significant

Further analysis revealed that exogenous factors should be considered as a rational predictor for the increase in value caused by the metro stations. When constraining the variables of the regression model to the endogenous factors (I.e. property type and number of bedrooms), the percentage of value uplift appears to remain constant. As the mean price of the property increases, so does the increase from transport accessibility at the pre-extension phase. This suggest that the value uplift is not affected by the internal features of the property. Conversely, when considering the exogenous factors, there are clear associations between the data. In areas of low unemployment (below 6%), the value uplift is approximately 12%-14%. In areas of high unemployment (above 6%), value uplift is between 7.4%-8%, suggesting a clear relationship between the variables. This is also evident for distance from the CBD, areas located closer to the CBD (0 mile to 0.9 mile) experience a shallower uplift to those located further away (1 mile to 1.9 mile) through the bid-rent model.

4.1 Equation based on Hedonic Model

It is first essential to establish the best equation for the model. By employing a stepwise approach, each variable has been considered for addition or subtraction from the set of independent/explanatory variables. This can improve the R-square value of the model. The equation for the study has been identified and is defined as:

$$\mathbf{P_i} = B_0 + B_1\mathbf{FLAT_i} + B_2\mathbf{TERR_i} + B_3\mathbf{BUILD_i} + B_4\mathbf{BEDROOM_i} + B_5\mathbf{\%UNEMP_i} + B_6\mathbf{\%HMAN_i} + B_7\mathbf{distSUN_i} + B_8\mathbf{distNEWC_i} + B_9\mathbf{extOPE_i} + B_{10}\mathbf{extEST_i} \quad \text{Eq. (2)}$$

Where, P_i represents the transaction price of property i; B_0, where B_0 represents the constant; B_xV_i, where B_x represents the the coefficient of the variable for property i. As previously explained, FLAT, TERR, SEMI and DETA are a group of dummy variables that represent a subgroup relating to the property type. When computing the multiple regression model, one dummy variable should be omitted from the study to provide a 'reference class.' The coefficient values are then calculated based on the omitted variable, i.e. a coefficient of 10000 for DETA would represent a £10,000 premium on the price of detached properties from that of a semi-detached property, and vice versa. This is also true for the transport accessibility variables. In this case, the extPRE variable has been omitted as a base. This results in the value of extOPE (immediate operation) and extEST (established operation) being calculated in reference to no accessibility to transport (extPRE).

4.2 Regression Model

4.2.1 Model Testing

In multiple regression analysis, the variability of the dependent variable is expected to be high, as there are multiple independent variables influencing the result. This should be a particular case for this study, as a 1-bedroom flat will sell for much less than a 3-bedroom semi-detached house within similar areas. This is confirmed in Table 3 above.

Multicollinearity can negatively affect regression-based analysis. An effective method of assessing it is the Variance Inflation Factor (VIF). A VIF of 10 indicates a strong multicollinearity. Table 4 indicates that the VIF of all independent variables in the regression equation are less than 2, indicating that no independent variables are highly linear dependent, concluding that multicollinearity is not an issue in this study.

It is acknowledged that regression assumes that results should be normally distributed. This can be examined by using the histogram and Q-Q-Plot (see Figure 3). As indicated, the regression results of this study are normally distributed.

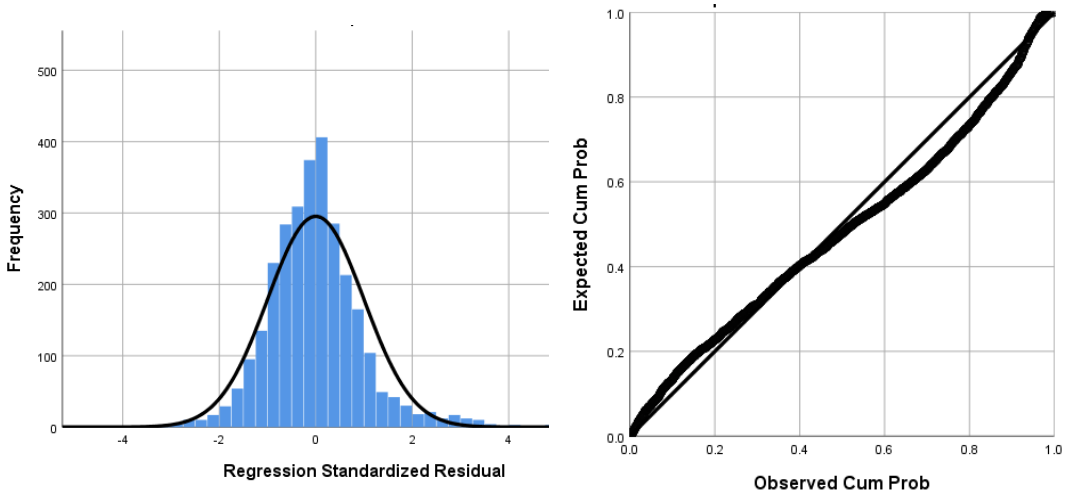


Figure 4. Histogram (left) and Q-Q Plot (right) showing normal data distribution

Testing of homoscedasticity is also essential for regression modelling. This can be achieved by plotting the standardised residuals against the predicted values of Y. As shown in Figure 4, there is no particular pattern to the distribution of variances, suggesting that homoscedasticity is present.

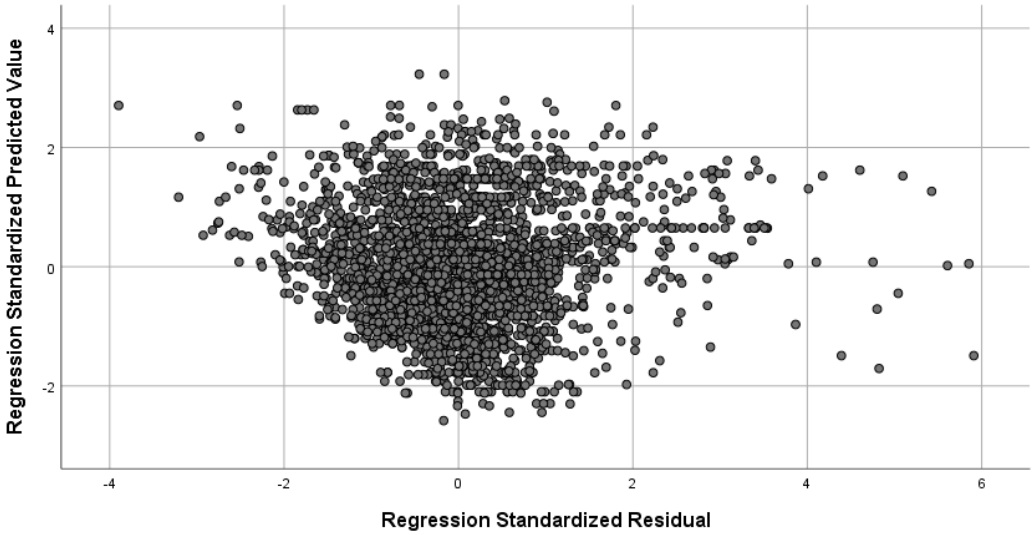


Figure 5. Homoscedasticity of data results

4.3 Results

The estimation results of the hedonic regression are presented in Table 4. A significant regression model was found of 'f' (10, 2946) = 241.424, $p < .05$), $R^2 = .449$ with the dataset. Overall, the model fit is represented by an adjusted R-square of 0.449, showing that 44.9% of the variation in the data is being explained. The unexplained variation is likely due to data limitations which led to the exclusion

of additional variables that may impact the value of property, including property attributes (number of bathrooms, car parking spaces, size in m^2) and amenities data (location to nearest schools, parks and other desirable areas). As this is a not a log model, the estimates are interpreted as the true measure of the dependent value, i.e. 29486.795 is equal to £29,486.80. The percentage of estimates provides the increase in house price on average for each variable, when all other variables are constant.

Table 4. Results of the multiple regression analysis

Parameter	**Estimate**	**Percentage of estimate / mean price**	**t value**	**VIF**
Constant	29486.795	N/A	5.056 *	
FLAT	-51826.306	N/A	-16.780 *	1.355
TERR	-35870.589	N/A	-22.340 *	1.260
BUILD	42153.467	38.27%	10.474 *	1.093
BEDROOM	17268.271	15.68%	18.323 *	1.080
%UNEMP	-4426.657	4.02%	-5.399 *	1.710
%HMAN	21043.463	19.10%	16.297 *	1.929
distSUN	1755.990	1.59%	2.542 *	1.073
distNEW	5839.930	5.30%	6.630 *	1.371
extOPE	10750.878	9.76%	6.774 *	1.394
extEST	3654.789	3.32%	1.850 non-sig	1.393
	Mean Price	110149		

** = significant at 5% level for one tailed tests*

In this hedonic price model, the endogenous factors are considered to be the variables that directly represent the property, including property type (FLAT, TERR, SEMI AND DETA) (whether the property is a new build) and the number of bedrooms. As the variable SEMI was dropped, the estimates for FLAT, TERR and DETA are in reference to the price of a semi-detached house. The price of a flat is £51,826 less expensive than a semi-detached house and a terraced property is £35,870 less expensive than a semi-detached house. Unsurprisingly, the data shows a new build premium for the properties, estimated at £42,153 or 38.27% of the mean property price. Finally, the estimate for number of bedrooms is £17,268 (15.68%) of the mean property price. This shows that each additional bedroom is worth £17,268, which takes up 15.68% of the mean property price.

The dummy variables extPRE, extOPE and extEST capture the property price difference at the various stages of the infrastructure delivery process. As in the property type subgroup, a variable has been omitted to provide a reference variable, which was extPRE. The estimates of extOPE and extEST therefore represent the estimated effect of SME accessibility on property values for the respective phase of the infrastructure delivery process, using extPRE as a base. ExtPRE represents the pre-construction phase of the extension line (Jan 1999 - May 2000), extOPE is immediately after the opening of the extension line (Apr 2002 – December 2003), and extEST represents the established (or longer-term) operation of the metro (Apr 2017 – Aug 2018).

As indicated by the regression model, the estimates for the variables BUILD, BEDROOM, %HMAN, extOPE and extEST were expected to be positive (β_3, β_4, β_6, β_9, β_{10} >0) while FLAT, TERR, %UNEMP, distSUN and distNEWC were expected to be negative (β_1, β_2, β_5, β_7, β_8 <0). To test whether the actual results can match expectations, a comparison of the outcomes and expectations has been undertaken,

for a purpose of identifying an alternative hypothesis (i.e. β_x (or H_1) > or < 0) and whether the outcome is significantly different from 0. If the result is not significant or $\beta_x = 0$, the alternative hypothesis is rejected, and the null hypothesis is accepted. The variables that do not match the expected outcome are marked in bold.

Considering the internal factors, FLAT and TERR were expected to have a negative value because the regression equation uses SEMI as a reference group, and flats and terraced properties should have a lower value than semi-detached properties. The BUILD and BEDROOM variables were expected to have a positive value on the constant, due to a new build premium that would be expected on all properties and the added value that is accrued from extra bedrooms in a property. As shown by Table 7, the actual estimates for these variables match the outcomes.

Regarding the neighbourhood or the socio-economic factors, it is expected that %UNEMP would be negative and %HMAN would be positive. This is anticipated as, generally, it would be expected that areas which possess lower percentage of unemployment and higher percentage of higher managerial positions would be regarded as more desirable neighbourhoods to live in and therefore increasing the price of properties in that area – with the converse relationship also expected to be true. Again, Table 7 indicates that these expected relationships match the actual outcomes from the regression model. The %UNEMP variable confirms that property price estimates decrease as the percentage of unemployment rate increases within the neighbourhood. For every one-point increase in the %UNEMP variable (i.e. every 2.9% increase in unemployment), property price decreases by £4,426. Conversely, the %HMAN variable is significant in confirming that property price estimates increase as the percentage of higher managerial positions increase. As a result of one-point increase in the %HMAN variable (i.e. every 2.9% increase in higher managerial position), property price increases by £21,043.

In terms of amenities, it was predicted that as the distance from Sunderland City Centre and Newcastle City Centre increases, the property price would decrease. This follows the premise of the bid-rent model, in that as distance from the desirable area increases the price paid for the land will decrease due to less demand. However, Table 7 indicates that the actual outcome from the regression model does not fit these expectations. For every 1 mile further from Sunderland City Centre and Newcastle City Centre, property values increase by £1755 and £5839, respectively.

When considering the results for each phase of the infrastructure process, it was expected that both extOPE and extEST would present positive values. This would suggest that properties with access to either a newly operational metro station or an established metro station would have greater property values than those without access, all else held constant. Both expected results were confirmed in the outcomes presented in Table 7. However, what is somewhat surprising is that the impact on property prices due to newly operational metro stations is almost three times the amount than that of established stations (£10,750 compared to £3,654).

4.3.1 Implications

Findings show that the *B* values of the variable extOPE is greater than 0, which was proven to be significant ($p<0.05$). This means that land value uplift does in fact occur immediately after the metro extension line becomes operational. This is different from the historical research which a) found no immediate change in value and b) hypothesised that value increase may come later. The estimate of the regression model also shows that in terms of transport accessibility, property values sharply increase immediately after the SME. Before decreasing once the line becomes established, at the

extPRE phase, property values were recorded at an average of £113,523. As identified from the regression model, the estimated increase at the extOPE phase is £10,750.

This suggests that an introduction of the newly operational metro stations caused an increase in property values of 9.53%. At the extEST phase, the increase in property values was estimated at £3,655. Put simply, after the metro station becomes operational, property values suffer a -5.71% decrease from the uplifted property value, all else held constant. However, from the extPRE to extEST phase, there is still an overall increase of 3.22% to property values. However, as confirmed by the regression estimate, these findings were not significant, and therefore cannot be accepted.

5. Discussion and Conclusion

Previous research considered a timeframe appropriate to measure short-term impacts of changes of accessibility on land values, as measured by changes in property prices. This paper progressed this original research to test a hypothesis that any impact may take longer to feed into house prices. A time-based study was conducted as opposed to a control group to compare properties without access (before the extension construction) and with access (immediately after and fifteen years after). The prices were adjusted using a house price index to bring all values in line with the present value, which meant that the 'time' aspect was the variable measuring transport accessibility.

This study hypothesised that land value uplift for extension lines occurs over a fifteen-year period as opposed to immediately after the line becomes operational. The results obtained from the regression model showed that in fact the greatest uplift in land value appeared to occur immediately after the line becomes operational. An average increase of £10,750, or 9.76%, accrues for properties at the extOPE phase, whereas no significant uplift was found at the extEST phase. This concludes that land value uplift does not in fact occur over a long period of time in the Tyne and Wear region, but rather immediately after the extension line opens. This potentially disproves the contention from Du and Mulley (2007a; 2007b), and the underlying hypothesis in this study, that uplift in value may occur later once the infrastructure is established. It also sheds light on the ambiguity found in literature around when value uplift occurs. Findings support the contention of (Mohammad, 2013) that value uplift occurs immediately after construction, indicating some insight into when land value capture mechanisms should be implemented. However, the authors caveat this contention by arguing that all locations, and underlying land and property markets are unique and demand their own analysis (Levinson and Istrate, 2011).

When considering the endogenous factors affecting property values, the regression model identified that there has been little effect on the premium generated by transport accessibility. The different property types (flat, terraced and semi-detached) and number of bedrooms (1, 2, 3 and 4) made little impact on land value uplift at the extOPE phase. All property types exhibited an increase of 7.80%-8.30%, and similarly the number of bedrooms exhibited an increase of between 9.68%-10.59%. This is an additional finding to the original research that did not consider property differentiation but suggests, in any event, that the endogenous factors are poor predictions of land value uplift in this case and would therefore not be required when identifying areas for land value capture implementation.

Further analysis beyond the hypothesis and a consideration of limitations and opportunities for further research is required to fully understand this situation but it could be contended that this different reading of events immediately after construction could be down to the quality of the underlying transaction data, which better reflects the market conditions at this time. However, there is a large amount of unexplained variation in the data. Hence the conclusions should be viewed

circumspectly. In order to reflect on this variation, there is scope for further analysis to measure access to transport as aerial distance to metro stops in combination with the time variable, or by including control groups outside the catchment area. In addition, different catchment area sizes (rather than the singular catchment area used in this study) could be deployed.

Furthermore, the property transaction data set was relatively rich in its currency, the Census data was restricted by its decennial nature. Although exhaustive in its coverage and geographic granularity, a weakness of the research is that this data is being used as a proxy over a long period of time. The next Census takes place in 2021 and should be included in any subsequent research. In addition, this paper relates to residential properties only. Although a more heterogeneous property type, there is also considerable scope to model potential uplift to existing commercial properties. The Valuation Office Agency have recently made freely available national rateable valuation information that can be geo referenced to local areas. Historically, this information has not been available in the public domain. This may help to determine whether value capture from commercial property would be worthwhile, as it provides the opportunity to capture value from potentially higher value properties around key areas of Tyne and Wear for further infrastructure projects.

Furthermore, for the purposes of this study, a prescribed set of variables were included to compute a hedonic price model. Instead, the emphasis was placed on developing a time-based data set. There is a long list of variables that could impact the value of property which may have implications for value uplift generated by the extension to metro lines, particularly at the local level. As this study identified some impact from exogenous factors, it is recommended that a more comprehensive set of exogenous factors are tested to fully understand the relationship with a view to developing rational predictors for identifying areas benefitting from transport accessibility and therefore potential land value capture methods. For example, in this study, it appears that unemployment can be considered a predictor variable to identify areas for land value capture. Although, due care should be taken in this area due to the potential gentrification impacts of targeted value uplift strategies in areas of high unemployment.

In addition, the study focused on three points in time, immediately before, after, and once established. This was an appropriate means of testing the contention of Du and Mulley (2007a; 2007b), that value uplift occurred later and to test the findings from literature that value uplift occurs immediately before or after construction. In this study, house price change between the census points is accounted for via a House Price Index that calculates the present value at each census point in the study – adjusting for inflation. However, there is opportunity going forward to also investigate each individual year within the date range, to investigate any variability between respective date ranges and to open the possibility of additional local methodological enquiry.

For example, although many studies have identified a 'positive' uplift in land value through improved accessibility, some have identified both an uplift and downlift around different stations within the same public transport system (Tsai et al, 2017). This was caused by specific local contextual factors, which prevented certain areas of land from producing the beneficial impacts from the infrastructure investment. Hence, it is important for further research to look at not only the global impact of an extension over time per se, but also the dynamics between stations – this is to analyse and counter any potential bias in the global model. Taking this geographical theme forward. It is recommended that further methodologies be used to understand the SME extension further. Geographic weighted regression (GWR) offers scope for additional research (Efthymiou & Antoniou, 2013; Mulley, 2014; Dziauddin et al, 2015; Wu et al, 2018; Dziauddin, 2019; Locurcio, 2020; Munshi, 2020) – while the difference in difference model can also be used to counter spatial autocorrelation (Forouhar & Hasankhani, 2018). GWR is based on a global regression model, such as a hedonic price model, which

is then modified by GWR to calibrate local regression parameters. The distance between data points is weighted through the geographical coordinates of the data (Du & Mulley, 2007b). Using the GWR method, there is scope for the internal heterogeneity of land value across the TWM system to become clearer. The information can be overlapped on a GIS to clearly present the spatial dispersion of land value uplift. In previous studies, both positive and negative impacts on land values were recorded around different stops in the TWM (Du & Mulley, 2007b) and further afield on the Buffalo light rail system (Hess & Almeida, 2007) and the Brisbane ferry system (Tsai, et al, 2017). This identified that while, on the global level, uplift across all stations was identified, specific local contextual factors prevented some stations from producing beneficial impacts from the investment.

Notwithstanding these limitations, the authors hope that these findings will also help inform the debate into, and scrutiny of, 'the everywhere else' – those locations around the world with sub-optimal conditions. In this sense, the findings are considered to have broader relevance in relation to debates involving the economic, social, political, and environmental dimensions of urban and regional (subnational) change, uneven development, and territorial evolution. The research also provides a basis for policy makers, governments and practitioners when evaluating ideas for financing and funding transport infrastructure systems in sub-optimal locations. For those transport and property professionals involved in the day-to-day management of transports systems, infrastructure development and real estate assets in the developed world, findings identified in this research will help contribute to more knowledgeable and effective practice regarding transport infrastructure development. Also, the developed approach will provide city leaders and property professionals with an insight into dealing with the demands of accelerating urbanisation in the less developed world. Such locations have the potential to learn from and avoid the pitfalls and complications seen in more mature urban areas. With direct pertinence for the empirical research area, the TWM system is currently being considered for further extension. The North East Joint Transport Committee responsible for strategic transport stated that: 'Some communities across the region were disadvantaged because they do not have access to the rail network. Let's put any financial surplus towards understanding the future of the network and where we can expand' (Councillor Carl Johnson, 2019).

The approach developed in this study has the potential to lead to a better understanding of the factors determining land value changes in the TWM system – particularly in relation to when it occurs. This can inform the type of land value capture mechanisms that would be appropriate to fund any subsequent extension. It can help tailor the financial planning of infrastructure within the landscape of fiscal uncertainty that demand intricate alignment of complex capital and revenue funding streams within locally specific conditions. As Levinson and Istrate (2011) state, there is no pre-conceived silver bullet for the selection and implementation of best practice for LVC mechanisms through accessibility value creation. Echoing Knowles & Ferbrache (2016), for an LVC mechanism to be effective, accessibility value must be measured and married with local planning and fiscal policy in order to successfully capture land value uplift at the correct place and time.

References

Bae, C. C., Jun, M. & Park, H., 2003. The impact of Seoul's subway Line 5 on residential property values. Transport Policy, 10, pp. 85-94

Billings, S. B., 2011. Estimating the value of a new transit option. Regional Science and Urban Economics, 41(6), pp. 525-536.

Bollinger, C., Ihlanfeldt, K. & Bowes, D., 1998. Spatial variation in office rents within the Atlanta Region. Urban Studies, 35, pp. 1097-1118.

Bowes, D. R. & Ihlanfeldt, K. R., 2001. Identifying the Impacts of Rail Transit Stations. Journal of Urban Economics, 50, pp. 1-25.

Cao, X. & Porter-Nelson, D., 2016. Real estate development in anticipation of the Green Line light rail transit in St. Paul. Transport Policy, 51, pp. 24-32.

Cervero, R., 2004. Effects of Light and Commuter Rail Transit on Land Prices: Experiences in San Diego County. Journal of the Transportation Research Forum, 43, pp. 121-138.

Clower, T. L. & Weinstein, B. L., 2002. The impact of Dallas (Texas) area rapid transit light rail stations on taxable property valuations. Australasian Journal of Regional Studies, 8, pp. 389-400.

CPRE., 2017. A response by the Campaign to Protect Rural England to the Raynsford Review of Planning Call for Evidence. UK Government.

Crocker, S., Dabinett, G., Gore, T., Haywood, R., Hennebury, J., Herrington, A., Kirkpatrick, A., Lawless, P. & Townroe, P., 2000. Monitoring the Economic and Development Impact of South Yorkshire Supertram, 2000: Centre for Regional Economic and Social Research, Sheffield Hallam University.

Duncan, M., 2008. Comparing rail transit capitalization benefits for single-family and condominium units in San Diego, California. Transportation Research Record: Journal of the Transportation Research Board, 1, pp 120-130.

Du, H. & Mulley, C., 2007a. The short-term land value impacts of urban rail transit: Quantitative evidence from Sunderland, UK. Land Use Policy, 24, pp 223-233.

Du, H. & Mulley, C., 2007b. Transport accessibility and land value: a case study of Tyne and Wear. RICS Research paper series. RICS. UK.

Du, H. & Mulley, C., 2012. Understanding spatial variations in the impact of accessibility on land value using geographically weighted regression. Journal of Transport and Land Use, 5, pp 46-59.

Dziauddin, M.F., Powse, N. & Alvanides, S., 2015. Estimating the Effects of Light Rail Transit (LRT) System on Residential Property Values Using Geographically Weighted Regression (GWR). Applied Spatial Analysis and Policy, 8, pp1–25.

Dziauddin, M.F. (2019) Estimating land value uplift around light rail transit stations in Greater Kuala Lumpur: An empirical study based on geographically weighted regression (GWR). Research in Transportation Economics, 74,10-20. https://doi.org/10.1016/j.retrec.2019.01.003

Efthymiou, D., Antoniou, C. (2013) How do transport infrastructure and policies affect house prices and rents? Evidence from Athens, Greece. Transportation Research Part A: Policy and Practice. 52, 1–22. https://doi.org/10.1016/j.tra.2013.04.002

Forouhar, A. & Hasankhani, M. (2018) The effect of Tehran metro rail system on residential property values: A comparative analysis between high-income and low-income neighbourhoods. Urban Studies, 55(16), 3503-3524. https://doi.org/10.1177/0042098017753089.

Forrest, D., J. Glen & R. Ward., 1996. The impact of a light rail system on the structure of house prices - A hedonic longitudinal study. Journal of Transport Economics and Policy, 30, pp 15-29.

Fuller, C., 2018. Entrepreneurial urbanism, austerity and economic governance. Cambridge Journal of Regions, Economy and Society, 11, pp 565–585.

Hass-Klau, C., Crampton, G. R. & Benjari, R., 2004. Economic impact of light rail: the results of 15 urban areas in France, Germany, UK and North America, Brighton: Environmental and Transport Planning.

Institute for Fiscal Studies (2020) The outlook for the public finances under the long shadow of COVID-19. IFS Briefing Note BN295. Published June 2020. IFS, London.

Johnson, R. 2019. New study into extension of Metro trains to Metrocentre, Washington and beyond could launch in 2020. Newcastle Evening Chronicle. 24th November 2019.

Knowles. R. D. & Ferbrache, F., 2016. Evaluation of wider economic impacts of light rail investment on cities. Journal of Transport Geography, 54, pp 430-439.

Lieske. S. N., Nouwelant, R. & Han. J.H. 2019. A novel hedonic price modelling approach for estimating the impact of transportation infrastructure on property prices. Urban Studies, 58, pp 182-202.

Levinson, D., & Istrate, E., 2011. Access for Value: Financing Transportation Through Land Value Capture, BROOKINGS, April 2011, [Online] Available from: http://www.brookings.edu/~/media/Files/rc/papers/2011/0428_transportation_funding_levinson_istrate/0428_transportation_funding_levinson_

Locurcio, M.; Morano, P.; Tajani, F.; Liddo, F.D. (2020) An innovative GIS-based territorial information tool for the evaluation of corporate properties: An application to the Italian context. Sustainability, 124), 5836. https://doi.org/10.3390/su12145836

McDonald. J.F, & Osuji.C.I. 1995. The effect of anticipated transportation improvement on residential land values. Regional Science and Urban Economics, 25, pp 261 – 278.

Mohammad, S. I., Graham, D. J., Melo, P. C. & 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, pp. 158-170.

Muldoon-Smith, K. & Sandford, M, 2021. Grasping the Nettle: the central-local constraints on local government funding in England. Territories, Politics, Governance, forthcoming.

Muldoon-Smith, K., 2019. Is it possible to incentivise and capture local wealth? The business rate challenge. Local Economy, 34, pp 213-220.

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. https://doi.org/10.1177/0042098013499082

Munshi, T. (2020) Accessibility, Infrastructure Provision and Residential Land Value: Modelling the Relation Using Geographic Weighted Regression in the City of Rajkot, India. Sustainability, 12(20), 8615. https://doi.org/10.3390/su12208615

O'Brien, P. & Pike, A., 2015. The governance of local infrastructure funding and financing. Infrastructure Complexity, 2 DOI 10.1186/s40551-015-0007-6.

Office for National Statistics (2020) A guide to labour market statistics. Explanation of the major concepts that exist within the labour market and their relationship to each other.

Pagliara, F. & Papa, E., 2011. Urban rail systems investments: an analysis of the impacts on property values and residents' location. Journal of Transport Geography, 19, pp 200-211.

Pan, H. & Zhang, M., 2008. Rail transit impacts on land use: evidence from Shanghai, China. Transportation Research Record: Journal of the Transportation Research Board, 2048, pp. 16-25.

Pickett. M. W. 1984. The effect of the Tyne and Wear Metro in residential property value. Transport and Road Research Laboratory (TRRL) Wokingham, Berkshire United Kingdom.

Pike, A., O'Brien., Strickland, T., Thrower, P. & Tomaney, J. (2019) *Financialising City Statecraft and Infrastructure*. Elgar – UK.

Plimmer, G. (2017) UK infrastructure projects struggle to attract pension funds. Financial Times, April 5, 2017.

PwC., 2018. Insights on Infrastructure: Perspectives across the project life cycle series. [Online]

Available at: https://www.pwc.co.uk/industries/capital-projects-infrastructure/insights/funding-uk-infrastructure.html [Accessed 24 January 2020].

RICS (2019) Comparable evidence in real estate valuation. 1st Edition.

Sirmans. G.S., MacDonald.L., Macpherson.D.A., & Zietz.E.N. 1996. The Value of Housing Characteristics: A Meta Analysis. Journal of Real Estate Investment and Economics. 33, 215-240.

Tsai, C.H., Mulley, C., Burke, M. & Yen, B., 2017. Exploring property value effects of ferry terminals: evidence from Brisbane, Australia. Journal of Transport and Land Use, 10, pp. 1-19.

Weinstein, B. L. & Clower, T. L., 1999. The Initial Economic Impacts of the DART LRT System, Texas: University of North Texas Centre for Economic Development and Research.

Wu, H.; Jiao, H.; Yu, Y.; Li, Z.; Peng, Z.; Liu, L.; Zeng, Z. (2018) Influence Factors and Regression Model of Urban Housing Prices Based on Internet Open Access Data. Sustainability, 10 (5), 1676. https://doi.org/10.3390/su10051676

Xu, T., Zhang, M. & Aditjandra, P., 2016. The impact of urban rail transit on commercial property value: New evidence from Wuhan, China. Transportation Research Part A: Policy and Practice, 91, pp. 223-235.

Yen, B. T., Mulley, C., Shearer, H. & Burke, M., 2018. Announcement, construction or delivery: When does value uplift occur for residential properties? Evidence from the Gold Coast Light Rail system in Australia. Land Use Policy, 73, pp. 412-422.