Mind the Gap: Windfall Gains in Housing Values Along London's Elizabeth Line

Candidate Number 603

University of Cambridge

Contents

1 Introduction

Cities remain a pillar of economies. Home to 56% of the world's population and generating over 80% of global GDP, urbanisation can be a key driver of individual and national development (World Bank, 2023). This is especially true in the UK, in which cities account for 9% of land, but outputs 63% of GVA (Centre for Cities, 2025). Transport infrastructure is vital and a necessary condition for the continuous operation of urban life, particularly in response to population growth. Beyond reducing travel times for its own and competing users, public transport helps boost worker productivity (Anas & Chang, 2023), relieve congestion (Anderson, 2013), reduce car-related accidents and injuries (Bauernschuster et al., 2017), improve health outcomes (Sener et al., 2016) and improve air quality at certain sites (Ma et al., 2021). The well-established socioeconomic and environmental benefits present clear grounds for public investment into improving inter- and intra-city connectivity; however, limited public funds naturally imposes a constraint on the number of transportation projects ultimately executed.

Another interesting potential impact is a capitalisation effect that manifests itself in uplifted property values. The existence of positive capitalisation effects from rail transit systems are very well documented in research covering many cities worldwide in the past five decades. Since 1970, there have been over 200 studies exploring transit-induced land value uplift (Rennert, 2022), with over 60 studies conducted in North America alone (Higgins & Kanaroglou, 2016). In addition, there exists strong evidence highlighting that accessibility effects are capitalised years before a line's opening, and can vary dramatically across stations. This has important implications on capturing these gains to finance ongoing or future projects via land value capture.

One of the most notable examples of transport infrastructure expansion in recent years is the construction of the Elizabeth line, a major public transport project aimed at improving connectivity and reducing travel times across London and its neighbours, as well as ameliorating regional disparities: an entrenched and severe issue rife across the UK.

Using a difference-in-differences model, we find that the Elizabeth line has resulted in positive capitalisation effects on housing values. However, the magnitude of land value uplift varies over space, time and regions. Across the entire study boundary, at the announcement phase, prices rise by 2.80% on average, which increases to 3.56% during the construction phase, before falling to 1.28% after opening. Yet, there is great spatial heterogeneity in results. The magnitude of land value uplift generally declines with distance from Elizabeth line stations.

Moreover, capitalisation effects in Central London and East regions, as well as in individual stations such as Canary Wharf, can exceed double that of the average.

The average treatment effect implies total land value uplift of £2.45 billion, or 13% of total construction costs, implying that a land value capture technique on residential properties, similar to the Business Rate Supplement applied to commercial properties, would aid in funding vital transport infrastructure in the future. Such a mechanism would move London closer to the beneficiary pays principle used by Hong Kong.

This topic lies at the crossroads between urban economics, real estate economics and spatial econometrics. The statistical techniques employed are relatively novel and continually refined in a growing body of research, but are nonetheless built upon well-established concepts and models. By combining economic and geospatial data using Geographic Information System (GIS), we are able to consider complex spatial interactions that would have otherwise been overlooked by traditional methods.

The paper is organised as follows. Section 2 reviews the related literature surrounding rail transport infrastructure expansion on residential property values. Section 3 provides a brief history and overview of the Elizabeth line. Section 4 describes the data and section 5 details the empirical methodology. Section 6 presents the results. Section 7 performs several robustness checks. Section 8 discusses the results and the policy implications. Section 9 concludes the paper.

2 Literature Review

2.1 Transport Network Expansion and Land Value Uplift

Bid-rent theory connects the expansion of transportation infrastructure to land value uplift (LVU). Building upon Von Thünen's (1826/1966) early work on farmland values and accessibility, Alonso (1964), Muth (1969) and Mills (1972) extended this idea to urban land markets in the monocentric city model. Under this framework, the spatial distribution of firms and households in a city is explained by the trade-off between land use and commuting costs, illustrated by the negative relationship between land prices and distance from the central business district known as bid-rent curves. This creates a utility maximisation problem whereby firms and households locate at a distance that balances travel costs and land prices based on their preferences. Since the physical characteristics of land are assumed given, this theory posits that location is the primary factor determining land values. As transport infrastructure reduces commuting costs, in terms of both time and money, improved accessibility should lead to increased bidding between users for properties close to stations, resulting in LVU and a localised, negatively sloped bid-rent curve away from stations.

There are two primary methods employed to estimate the magnitude of this location premium: hedonic pricing models (HPM) and difference-in-differences (DID) regressions. The majority of studies have found positive capitalisation effects (Mohammad et al., 2013; Rennert, 2022). Contemporaneously, negative capitalisation effects (Landis et al., 1995; Nelson, 1992; Wagner et al., 2017) and insignificant results (Clower & Weinstein, 2002; Du & Mulley, 2007; Gatzlaff & Smith, 1993) have also been identified in a small number of cases.

The broad scope of transportation allows researchers to study the effects of many transport modes — including light rail transit (LRT), heavy rail transit (HRT), commuter rail transit (CRT) and bus rapid transit (BRT) — on both residential and commercial property values. This is a source of some of the large variation in results across studies. The first major meta-analysis, by Debrezion et al. (2007), attempted to explain some of this heterogeneity, and concluded that LVU is stronger for commercial property than residential property for short distances from stations, but accessibility premiums persist across longer distances for residential property. In addition, CRT stations produce stronger uplift than stations pertaining to other transport modes, and highway accessibility diminishes transit-induced LVU due to its competing nature.

Mohammad et al. (2013) build upon this foundation by incorporating more contextual and methodological factors, and extend the geographical scope beyond North America by including studies in Europe and Asia. Their results confirm the key findings by Debrezion et al., and indicate stronger LVU effects in Europe and East Asia compared to North America, possibly due to greater reliance on public transit systems in these regions. Most recently, Rennert (2022) separates these two factors by controlling for public transport expenditure share, concluding that more expensive transit services, relative to income, reduce accessibility benefits, and that results from Europe, Asia and Oceania are statistically similar to North America, whereas East Asia produces lower results. Additionally, Rennert introduces further variables, including transit service quality factors, with larger transit networks generating marginally higher LVU.

Furthermore, Bowes and Ihlanfeldt (2001) theorise that LVU could be the net effect of positive benefits such as accessibility and retail development less negative externalities and crime, particularly for properties situated closest to stations. Capitalisation effects may also be dependent on the project's success; for example, the study by Wagner et al. (2017) was distinct in that the LRT system examined was among the worst performing LRT lines in the US at the time in terms of ridership, profitability and relative accessibility benefits, which contextualises the land value depreciation of 7.8%. Lastly, economic decline within the study area may contribute to insignificant or low capitalisation effects (Du & Mulley, 2007; Gatzlaff & Smith, 1993; Hess & Almeida, 2007).

2.2 Hedonic Analysis

Hedonic pricing theory, originating from the seminal work of Rosen (1974), has long been a popular method in the applied economics literature due to its applicability, enabling the valuation of non-market goods such as air quality or a statistical life with the appropriate dataset. The use of HPM to study the effects of transport infrastructure on housing prices is no exception. By incorporating housing and locational attributes along with an accessibility variable into an estimation, researchers can identify the implicit price of location. Differencing pre- and post-intervention regressions with a dummy variable identifying properties within the treatment zone reveals the magnitude of LVU.

However, using the ordinary least squares (OLS) estimator can potentially introduce endogeneity. This is because unobserved factors may simultaneously influence both property values and accessibility. An example of omitted variable bias in early studies involves the omission of other competing accessibility factors, such as distance to highways and freeways, which led to an overestimation of proximity effects (Debrezion et al., 2007; Mohammad et al., 2013). The omission of local crime statistics among other neighbourhood amenities could also lead to biased distance coefficients (Bowes & Ihlanfeldt, 2001).

Researchers use repeat-sales models to control for endogeneity in modelling housing prices. Although the repeat-sales approach removes biases caused by time-invariant omitted variables and requires lower-resolution data, this introduces further problems. Not only is the sample size is drastically reduced in the process, but there may also be sample selection biases if houses that are sold at least two times are materially different from those that sell just once (Gatzlaff & Haurin, 1997). Moreover, many housing and neighbourhood characteristics do change over time, even if slowly; therefore, differencing housing prices relies on a flawed assumption of time-invariant attributes that may not adequately address endogeneity.

2.3 Difference in Differences

DID studies have become increasingly adopted in the transit-induced LVU literature. The essence of DID lies in combining the strengths of before-and-after and with-and-without approaches by testing the pre- and post-intervention changes between treated and control groups following an exogenous shock, whilst controlling for observed and unobserved confounders in housing and neighbourhood characteristics. Using DID in lieu of OLS mitigates concerns of endogeneity, but in turn introduces the crucial new assumption of parallel trends. The ability to establish causality makes DID a powerful tool for impact evaluation. As with OLS, DID's versatility makes it a popular and suitable methodology for many fields beyond economics.

Gibbons and Machin (2005) follow the DID method by first-differencing housing prices between pre- and post-intervention periods for treated and control groups to estimate the effect of new DLR and Jubilee line stations on property values. Using data from Nationwide Building Society, the authors restrict their sample to properties within 30km of Holborn (a proxy for London's central business district), and find that a 1km reduction in distance to the nearest Jubilee line or DLR station increases property values by around 2.1% for properties within 2km of a station. This paper is one of the closest to our study given its focus on the London residential property market. However, its methodology can be extended by incorporating a treatment-year interaction term to test for anticipation effects, and by using property-level, rather than postcode-aggregated, data to improve spatial precision.

Mohammad et al. (2013) and Rennert (2022) demonstrate that LVU estimates derived from DID models are, on average, not statistically different compared to those from HPM. However, this position is complicated by individual studies employing both DID and HPM; for instance, Gibbons and Machin (2005) also estimate an OLS model which resulted in considerably higher LVU estimates than DID, whereas Mohammad et al. (2017) found the opposite in their analysis of the Dubai Metro on residential dwellings. Despite this discrepancy, the econometric limitations of OLS judged against the merits of DID justify preference towards the latter method.

2.4 Decomposing Treatment Effects

The average treatment effect across all stations is useful for inferring the aggregate impact of a transport line, but it does not tell the complete story. A number of studies focus on the timing of capitalisation effects, which is particularly relevant for informing policy decisions around project funding and land value capture (LVC) methods. Building upon McDonald and Osuji's (1995) study on the Midway Line in Chicago with a dataset spanning across more years, McMillen and McDonald (2004) show that capitalisation effects emerged up to 6 years before the line's opening. Yen et al. (2018) centre their focus on testing for anticipation effects across the entire lifespan of the transport project. By using data on housing transactions decades before the line's opening, the authors reveal that the majority of LVU occurred when there was evidence of firm commitment to the project, in addition to positive effects at the announcement and operation phases. Further studies also report LVU before the project's opening (Bae et al., 2003; Bao et al., 2021; Diao et al., 2017; Higgins, 2019; Huang et al., 2024; Hyun & Milcheva, 2019). Wagner et al. (2017) similarly found that negative capitalisation effects materialised both before and after operations. These results reinforce the notion that housing values indeed react relatively quickly to news, with accessibility benefits being capitalised long before a line's completion.

Other papers examine how the treatment effect varies across stations. Hess and Almeida (2007) find that the largest positive accessibility effects of LRT are felt in high-income neighbourhoods in Buffalo, New York, and the largest negative effects in low-income neighbourhoods, a result consistent with Bowes and Ihlanfeldt (2001) but contrary to Nelson (1992). Despite identifying negative and insignificant capitalisation effects for multi-family housing, Huang et al. (2024) found that one area outside the downtown core saw a 30.3% increase in multi-family housing prices at the announcement phase, whilst all other areas simultaneously experienced very low or negative LVU. Beyond finding area-specific variations in effects, a few studies have discovered network effects in the form of LVU for stations that did not receive new infrastructure, but nevertheless benefitted from improved network accessibility — whether through increased connectivity at a former terminus station gaining integration (Higgins, 2019), enhanced accessibility benefits at the regional scale extending beyond immediate catchment areas (He, 2020) or uplift along the existing line, especially at interchange stations (Zhu & Diao, 2024).

2.5 Spatial Econometrics

Researchers often assume away spatial interactions between observations by controlling for spatial fixed effects using postal or other administrative codes. However, there is a growing body of literature arguing for the need to control for possible spatial autocorrelation and spatial error in the interactions of property transactions in the study of transport infrastructure expansion and LVU. This is the premise behind the increasing adoption of spatial difference-in-differences (SDID) models, which is perhaps the most novel innovation in this field.

Much of the early work in spatial econometrics was pioneered by Anselin (1988). Spatial autoregressive models (SAR) incorporate a spatial lag term to capture the effect of neighbouring observations on the dependent variable; spatial error models (SEM) introduce a spatial error term to account for spatially-dependent omitted variables that could cause correlated errors between spatial units. Models including both terms are called spatial autocorrelation or spatial autoregressive combined models (SAC/SARAR). These are merely extensions of a general regression framework; if both of these parameters are statistically close to zero, the estimation in effect reverts back to the standard aspatial model. The existence of spatial lag is supported by the intuition that property valuation is not only based on its characteristics, but also on the transaction prices of nearby comparable properties.

Dubé et al. (2014) led the application of spatial econometrics to the study of transportation on housing values using SAR, but found insignificant results. Diao et al. (2017) use SARAR and report significant spatial lag and error parameters at the 1% level. Upon finding significant capitalisation effects of 10.6% and 8.6% from the Circle Line in Singapore using DID and SDID respectively, this implies that DID overestimates outcomes compared to SDID. Subsequent SAR and SARAR studies find mixed results, with some finding both significant spatial lag and error (Higgins, 2019; Higgins et al., 2024; Qiu & Tong, 2021; Zhu & Diao, 2024), others finding only significant spatial lag (Hyun & Milcheva, 2019) or only significant spatial error (Huang et al., 2024) and those with neither (Dubé et al., 2024).

The main issue with constructing the spatial weight matrix required for spatial regression is the lack of a mathematically-rigorous, first-best method that is universally accepted, and so largely remains an empirical question. Because authors use different methods to obtain the spatial weight matrix, this would explain part of the disparities in results. The memory-intensive nature of calculating the spatial weight matrix, owing to its quadratic time complexity, further stifles its diffusion into the literature. Though the lack of standardisation across spatial models may undermine their application in guiding policymaking, it is nonetheless worthwhile to include when modelling LVU to observe whether spatial autocorrelation truly exists in the housing market at hand, and to compare with conventional models.

2.6 Further Methodological Considerations

One empirical issue involves robustly defining the treatment zone of properties that are affected by the construction of new stations. Many studies arbitrarily choose a cutoff distance to sort between groups, with 1/2 mile (Billings, 2011; Hurst & West, 2014; McDonald & Osuji, 1995; Nelson & McCleskey, 1990) and 800m (Huang et al., 2024; Yen et al., 2018; Zhang, 2023) being a popular choice. The line of reasoning is that this figure represents a walkable distance of around 10 minutes (Guerra et al., 2012). Although this argument has merit, the true treatment zone may not be homogenous across study areas. A more robust method of selection would involve a statistical-based approach. Diao et al. (2017) use a local polynomial regression to identify a treated zone, which although allows for non-parametric relationships, it does not explicitly control for confounders. Xu and Zhang (2016), followed by Zhu and Diao (2024), construct discrete distance dummies and interpret their coefficients relative to a reference group using OLS to sort between groups.

When calculating the distance between a given observation and its nearest station, most studies elect between using either the Euclidean or network distance. Diao et al. (2017) argue that the Euclidean distance is problematic given that it ignores physical features of the urban environment, and thereby underestimates the true distance cost. Although the network distance reflects the true distance faced by commuters, the Euclidean distance is sometimes referred to as the perceived distance (Hess & Almeida, 2007). Both Hess and Almeida (2007) and Diao et al. (2017) compare the network and Euclidean distance measures and find the latter measure underestimating location premia. However, calculating the network distance requires shapefile data on road networks for the specific study area in GIS. A feasible compromise would be to use the Manhattan distance (also known as the city block or taxicab distance), which too has a one-size-fits-all formula that can be applied to entire datasets without the use of GIS and paints a more realistic depiction of urban reality.

2.7 Research Gap

On balance, it is clear that although LVU studies generally produce positive and significant outcomes, with meta-analyses revealing a consistent set of drivers influencing variations in results, the extent, timing and distribution of capitalisation effects are context dependent, and will ultimately depend on the specific nuances of the transportation project and study area.

There have been commercial reports on Crossrail before its inauguration. For example, GVA (2018) estimate realised LVU of 19% and 8% for residential and commercial properties within 1km of an Elizabeth line station respectively, representing an additional £10.6 billion in residential property value. Research by CBRE (2024) finds a growth premium of 6% in property values and 2% in rental values between 2008 and 2023 along London stations specifically (excluding those in Zone 1) compared to wider boroughs. Savills and KPMG found no evidence of anticipated LVU on residential property during Crossrail's construction, but uplifts on commercial values of around 1–2.5% per year (Transport for London & Greater London Authority, 2017). In a more recent report by TfL (2024), it is briefly mentioned that a 14% higher growth has been identified in East London and 8% in West London for properties within 1km of an Elizabeth Line station, compared to a 3km control area, between 2017 and 2022. While the results are positive, they show some variation, and the exact methodologies and datasets used in these commercial studies are less transparent than those used in academic research. Moreover, these reports focus on the shorter-term or marginal impacts from the line's post-announcement years, with none providing a full view of the Elizabeth line's entire life cycle, which requires data predating its official announcement in 2008.

To the best of my knowledge, there are no existing academic studies that analyse the capitalisation effects of the Elizabeth line on residential property values across its entire lifespan — from proposal to operation. By performing rigorous DID analysis on a comprehensive dataset spanning 18 years, we hope to contribute to the existing literature by 1) add another datapoint on the magnitude of capitalisation effects resulting from the Elizabeth line, 2) shed light on the existence of spatial dependencies using SDID and 3) compare the usage of the Manhattan distance over Euclidean distance. Analysing the timing of LVU and variation across regions, the results can help inform policymaking in terms of cost-benefit analysis and funding methods for future rail transport infrastructure and, lastly, show perspective on one potential driver of wealth and regional (in)equality.

3 The Elizabeth Line

Named after HM Queen Elizabeth II, the Elizabeth line is a 118km-long railway that runs in an east-west direction across London, with its eastern periphery extending to Shenfield and its western terminus in Reading. The line has been expected to expand central London's rail capacity by 10% and contribute £42 billion to the UK economy (Transport for London, 2022). Jointly sponsored by Transport for London and the Department for Transport, the project aligns with the government's broader transport decarbonisation goal and Levelling Up mission to ameliorate regional inequality. This is important given that, historically, East London has been less prosperous than its western counterpart, and most of economic output in the UK is heavily concentrated in London.

Under a different alias of Crossrail, plans first began in 2001 when the Crossrail Ltd joint venture formed between Transport for London and the Department for Transport. The business case, setting out its outcomes, benefits and multiple proposed routes, was published in 2005, with the Hybrid Bill's proposal to Parliament on 22^{nd} February 2005.[1] The system was formally approved by the Crossrail Act 2008, receiving royal assent on 22^{nd} July 2008.[2] Construction works began soon after on 15^{th} May 2009, starting in Canary Wharf. Though originally planned to open in 2018, the project faced repeated delays, including as a result of the COVID-19 pandemic. Service officially commenced on 24th May 2022. Bond Street station, however, opened five months later on 24th October 2022.

[1] For further details on Crossrail's initial plan, see Tucker (2017) and Bennett (2017).
[2] https://www.legislation.gov.uk/ukpga/2008/18/contents

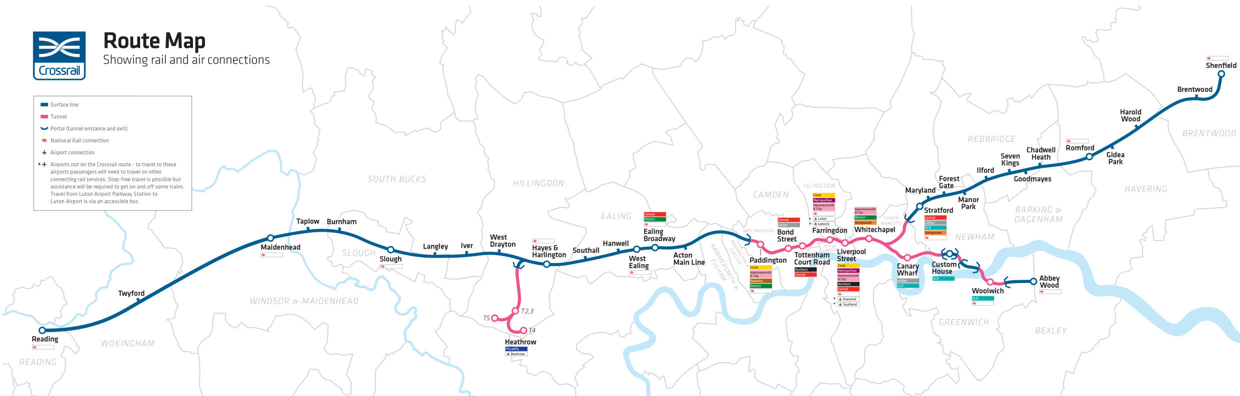


Figure 1: The Elizabeth Line Route

Since its opening, the line has made 350 million journeys. The cumulative benefits of improving train and highway journey times, congestion relief, accident savings, tax revenues, emissions, noise and air quality are estimated to outweigh capital and operating costs by a ratio of 1.9 (Transport for London, 2024).

4 Data

4.1 Study Area

Our dataset consists of residential property transactions within a 3km buffer zone from all 41 Elizabeth line stations between January 2006 and December 2024. This dataset comes from Zoopla, one of the largest real estate property portals in the UK, and consists of 309,876 housing transactions in total.

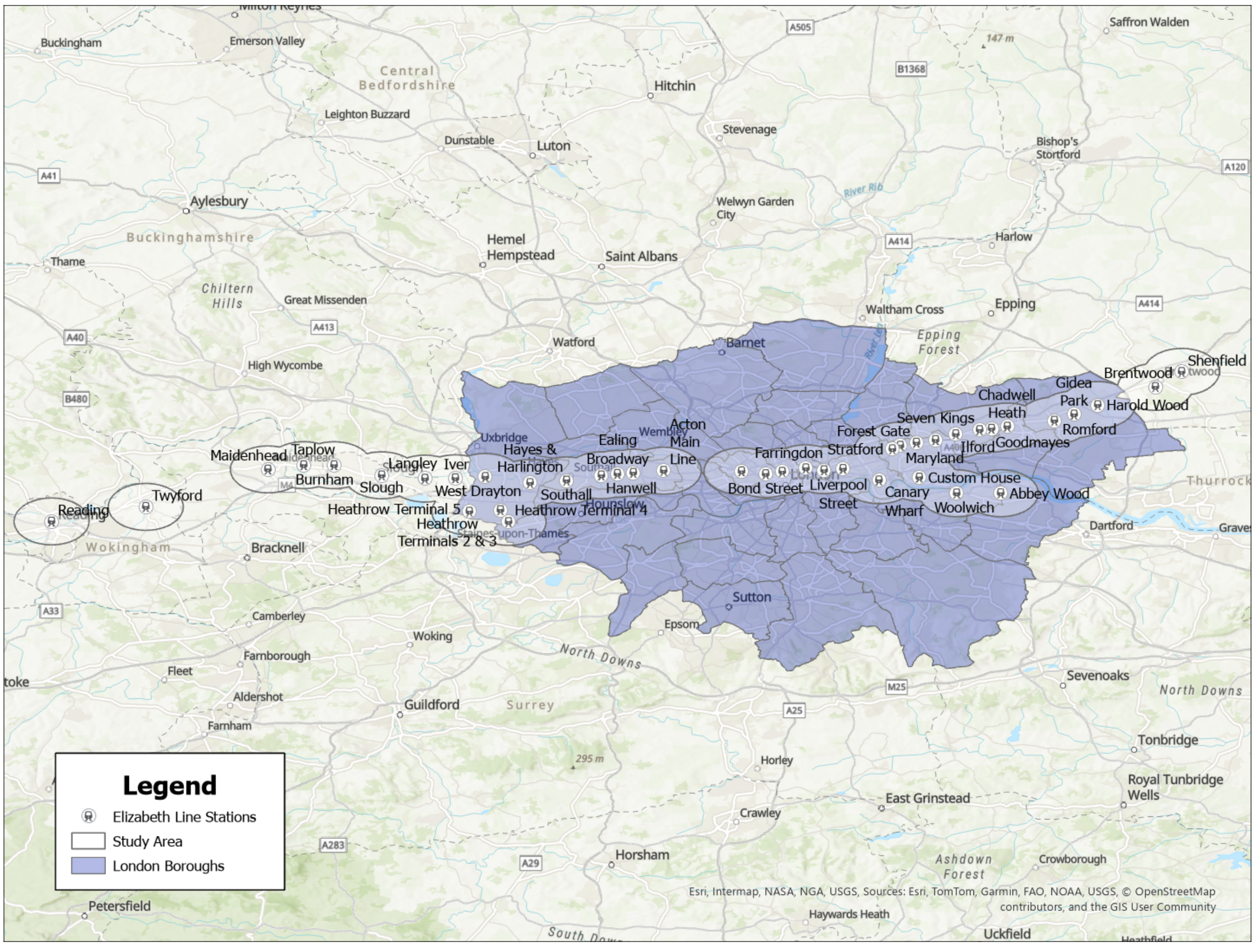


Figure 2: Study Boundary (Scale: 1:525,000)

4.2 Housing Characteristics

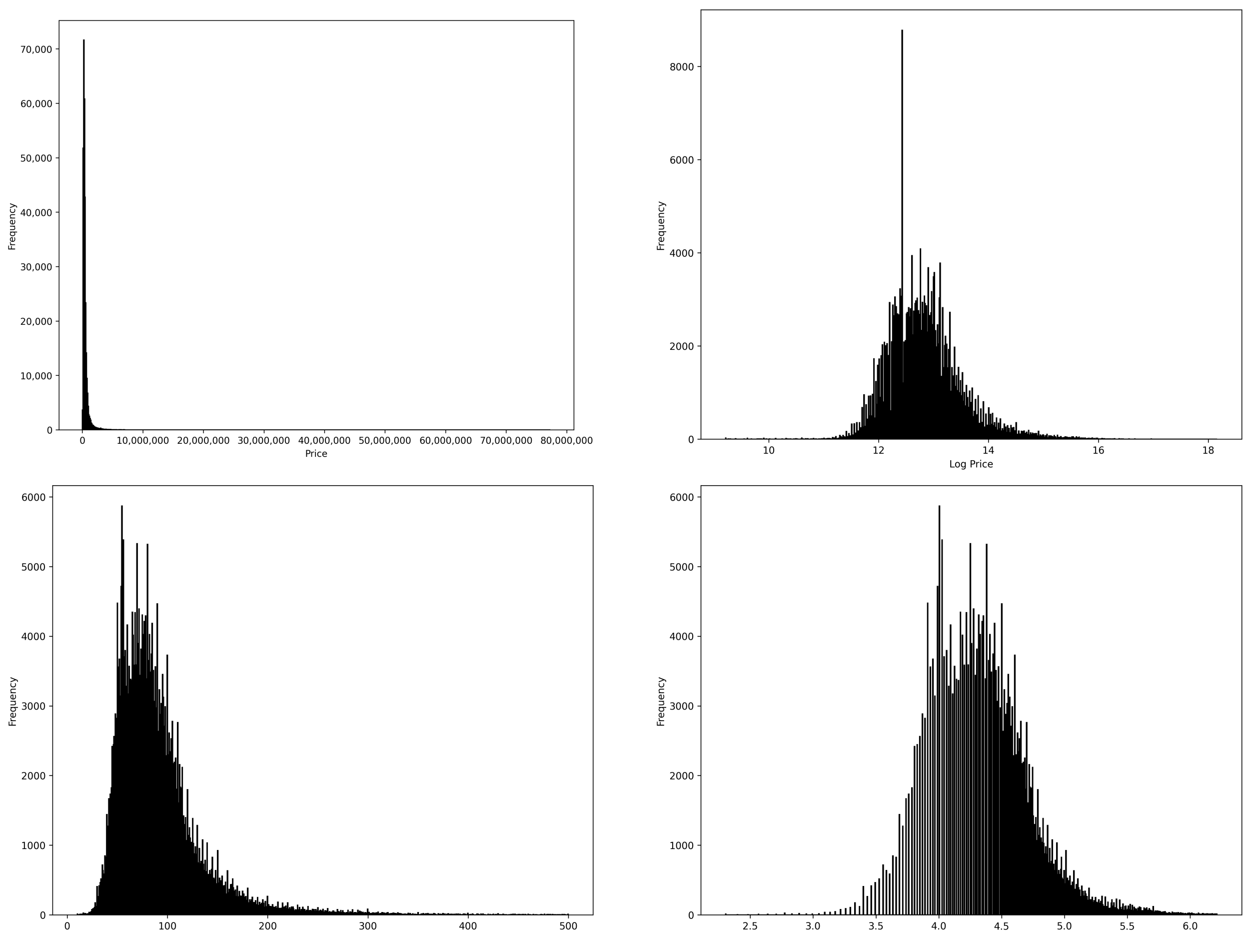
Each observation includes the floor area (in square metres), number of bedrooms, bathrooms and living rooms, property type, tenure, Energy Performance Certificate (EPC) rating and longitude and latitude coordinates.

The Property Type category includes Bungalow, Flats, Terraced, Semi-Detached and Detached properties; Tenure is either Leasehold or Freehold. EPC ranges from A to G, where A represents the best possible energy efficiency rating.

In addition, the distance to the nearest Elizabeth line station is calculated for each property using the network, Euclidean and Manhattan distance measures. Elizabeth line station coordinates are obtained from the Transport for London website. Road network shapefile data is available from Ordnance Survey.

4.2.1 Semi-Log Model

The natural logarithm of the Price and Floor Area variables is taken in order to normalise the distributions, and to capture diminishing marginal utility in attributes.



4.3 Neighbourhood Characteristics

We match population density and population percentage over 65 data from the Office for National Statistics using each property's Lower Layer Super Output Area (LSOA), a small geographical unit created for statistical purposes, primarily for the Census.

The Index of Multiple Deprivation (IMD) is matched to each property using their LSOA. This index captures various dimensions of socioeconomic disadvantage, including income, employment, education and crime, providing a comprehensive measure of deprivation. Given the IMD's multifaceted nature, median income and crime statistics are not required; including them results in strong multicollinearity.

4.4 Fixed Effects

To control for unobserved locational attributes uncontrolled for by existing neighbourhood controls, 168 spatial fixed effects are introduced in total using the outward postcode. Time fixed effects are included on a quarterly basis, with 76 quarterly time dummies spanning from Q1 2006 to Q4 2024. The first outward code, CM13, and first quarter, Q1 2006, will serve as the reference categories.

4.4.1 Multicollinearity

Due to high correlation between Floor Area and Bedrooms, we drop Bedrooms from the models given its discrete nature. We also drop the Tenure category given the high correlation between Flats and Freehold and Leasehold.

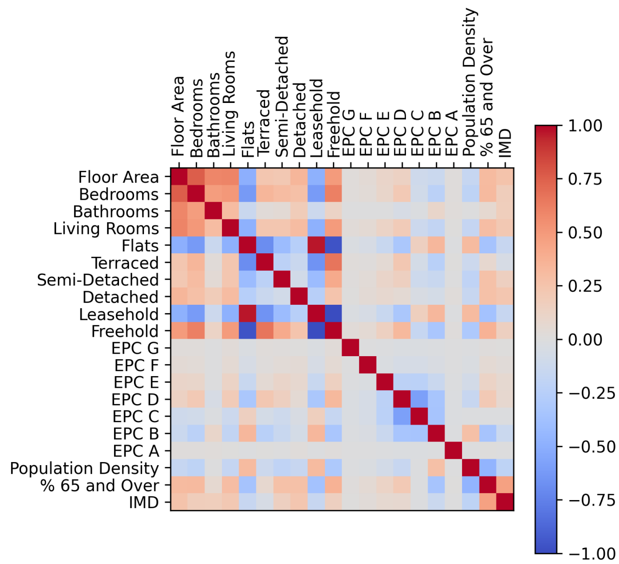


Figure 3: Covariate Correlation Matrix

For each categorical variable, one dummy is dropped to serve as a reference category, preventing perfect multicollinearity. For Property Type, the Flats category is dropped, and for EPC, the G rating.

The Variance Inflation Factor (VIF) is calculated for all covariates in Table 1. As no values exceed 5, this suggests that multicollinearity is not a significant issue.

Variable	VIF
Treat × Announcement	1.6728
Treat × Construction	2.1124
Treat × Post	1.8894
ln(Floor Area)	2.7213
Bathrooms	1.6948
Living Rooms	1.6469
Property Type	1.2228
EPC	1.0454
Population Density	1.7992
% over 65	2.8352
IMD	2.1401
Outcode	1.0090
Quarter	1.0104

Table 1: VIF

4.5 Distance Measures

Suppose we have two observations i and j with respective longitude and latitude coordinates (x_i, y_i) and (x_j, y_j). The Euclidean and Manhattan distance measures are calculated in two dimensions using the formulas:

$$d_{ij}^{\text{Euclidean}} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

$$d_{ij}^{\text{Manhattan}} = |x_i - x_j| + |y_i - y_j|$$

Except for when i and j are identical or perfectly aligned, the Manhattan distance must necessarily be greater than the Euclidean distance due to the triangle inequality. Given that the Euclidean distance underestimates the true distance faced by commuters, using this measure may lead to biased results.

In the below example, the network distance between the 1 Park Drive residential skyscraper and Canary Wharf Elizabeth line station is calculated as 670m in ArcGIS Pro. The Euclidean distance is calculated as 526m, and the Manhattan distance as 719m. The latter measure is a closer approximation of the network distance and, unlike the former, does not necessarily require the commuter to traverse through water and buildings.

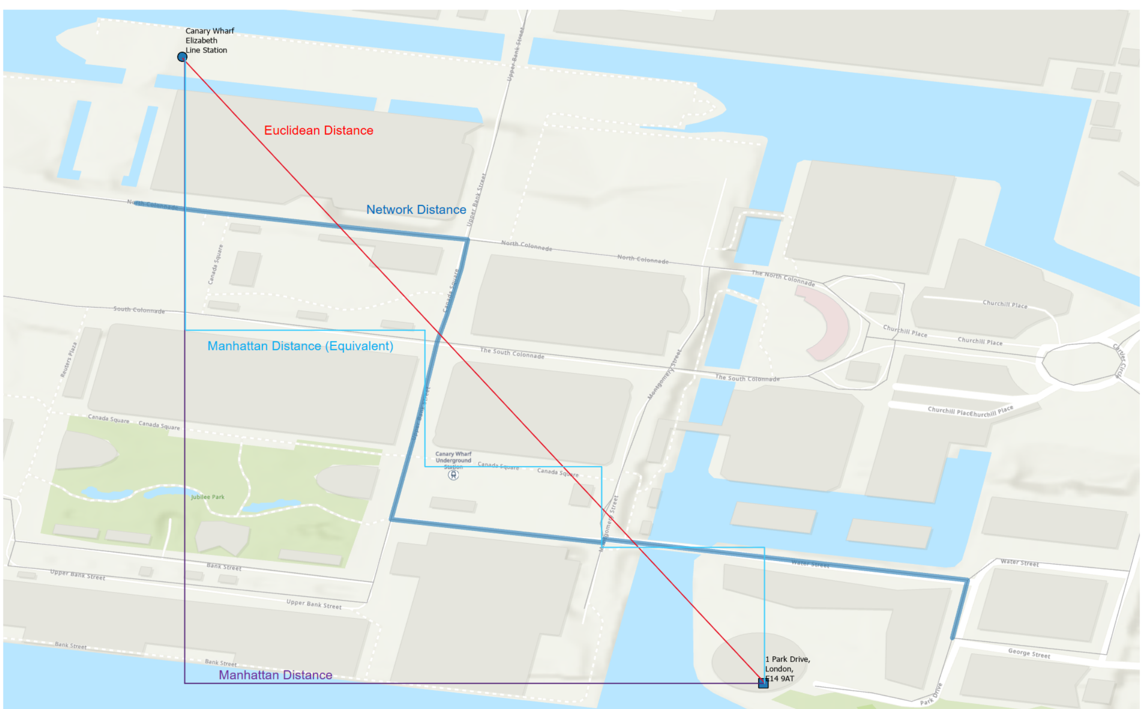


Figure 4: Distance Measure Comparison

Generally, the Manhattan distance becomes a better proxy for the network distance the more grid-like the study area's layout is. However, both measures will fall short if there is a large insurmountable obstacle, such as crossing the River Thames in areas without a nearby bridge or tunnel. Thus, all models will be based on the network distance unless otherwise stated.

4.6 Treatment and Control Zones

Rather than arbitrarily assigning a cutoff distance to sort between treatment and control groups, we can dynamically set a boundary based on the data. We adapt Xu and Zhang's (2016) methodology by sorting observations into 150-metre distance bands and set properties beyond 2850m as the reference category; using HPM with housing and neighbourhood controls alongside spatial and time fixed effects enables us to identify the price premium associated with each band. It is expected that the coefficients should be systematically positive and significant up until a certain group, whereupon it remains insignificant or negative thereafter. It is also very likely that the treatment zone will differ depending on the distance measure used.

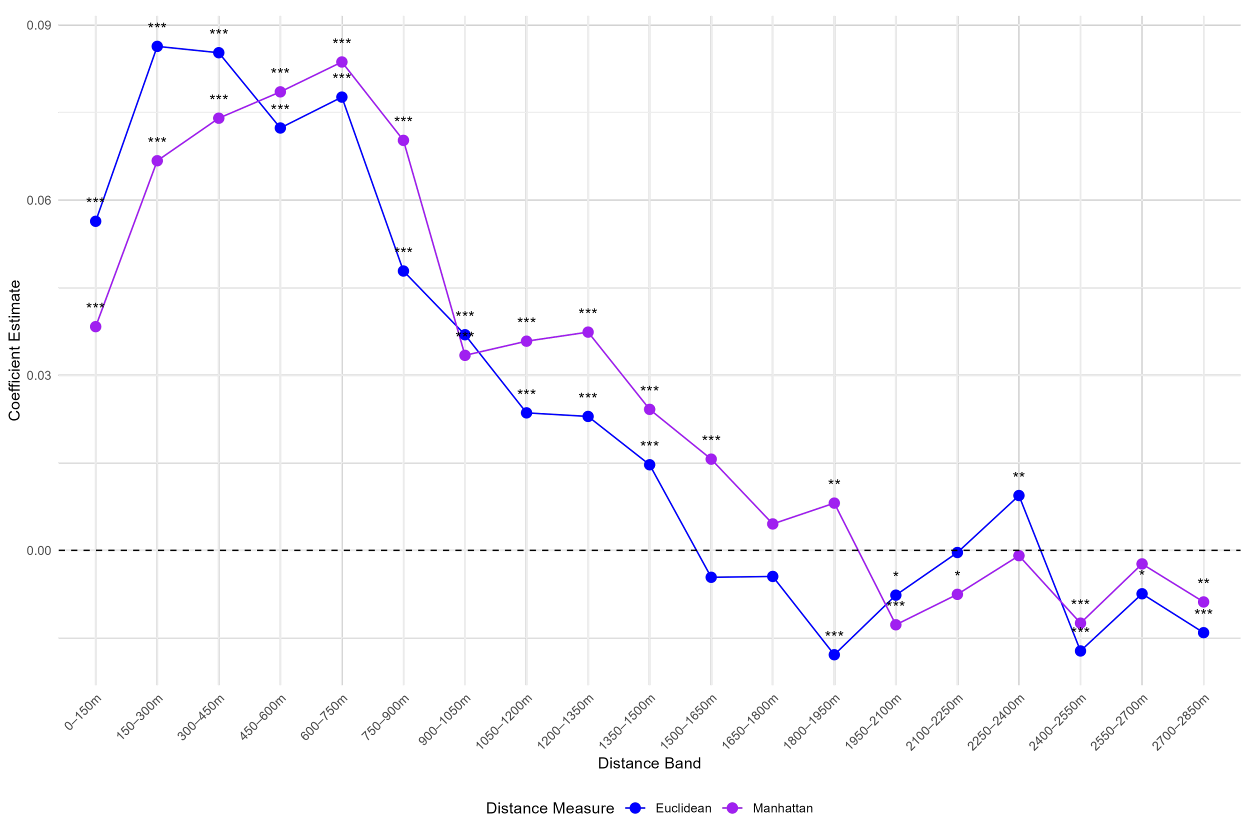


Figure 5: Location Premium by Distance Band

Figure 5 shows that the hedonic price for proximity to stations initially increases with distance, before gradually descending. The coefficients are positive and significant up until 1500m, 1500m and 1650m for the network, Euclidean and Manhattan distance measures respectively. This result is similar to a report by Nationwide (2021), finding that the location premium associated with London Underground stations extends up to 1500 metres.

4.7 Descriptive Statistics

Table 2 presents the descriptive statistics for the entire dataset and the two groups.

	Full Dataset			Treatment Group			Control Group		
Statistic	N	Mean	St. Dev.	N	Mean	St. Dev.	N	Mean	St. Dev.
Price	309,876	508,688.00	739,005.70	147,310	467,571.10	617,438.40	162,566	545,946.20	832,248.00
ln(Price)	309,876	12.85	0.66	147,310	12.81	0.63	162,566	12.89	0.69
Floor Area	309,876	90.92	45.51	147,310	88.11	43.08	162,566	93.47	47.45
ln(Floor Area)	309,876	4.42	0.42	147,310	4.39	0.41	162,566	4.44	0.42
Bedrooms	309,876	2.36	1.05	147,310	2.28	1.04	162,566	2.43	1.06
Bathrooms	309,876	1.41	0.66	147,310	1.40	0.63	162,566	1.42	0.68
Living Rooms	309,876	1.26	0.54	147,310	1.24	0.52	162,566	1.28	0.57
Population Density	309,876	12,103.23	10,486.51	147,310	12,447.94	11,334.50	162,566	11,790.86	9,643.38
% over 65	309,876	0.10	0.06	147,310	0.10	0.06	162,566	0.11	0.06
IMD	309,876	15,006.35	7,489.59	147,310	15,111.65	7,230.71	162,566	14,910.94	7,715.45
Euclidean Distance	309,876	1,588.09	772.03	147,310	897.03	365.14	162,566	2,214.29	436.39
Manhattan Distance	309,876	1,977.45	963.87	147,310	1,133.02	471.62	162,566	2,742.64	581.11

Table 2: Summary Statistics

5 Estimation Strategy

5.1 Difference-in-Differences Model Specification

For the baseline results, we start with a standard DID model of housing prices:

$$\begin{aligned}\ln(P) = \alpha &+ \beta_1(\text{Announcement} \times \text{Treat}) + \beta_2(\text{Construction} \times \text{Treat}) \\ &+ \beta_3(\text{Post} \times \text{Treat}) + H'\gamma + N'\theta + \varphi + \tau + \varepsilon\end{aligned} \tag{1}$$

where $\ln(P)$ is log property price, $H'\gamma$ represents housing characteristics, $N'\theta$ represents neighbourhood characteristics, φ are spatial fixed effects, τ are time fixed effects and ε is the i.i.d. error term.

Treat is a binary variable which takes on the value of 1 if the property is within the treatment zone threshold distance from an Elizabeth line station, and 0 otherwise. Announcement is another binary variable that equals 1 if the transaction date occurs between 22$^{\text{nd}}$ July 2008 and before 15$^{\text{th}}$ May 2009, and 0 otherwise. Construction equals 1 if the transaction date is between 15$^{\text{th}}$ and 24$^{\text{th}}$ May 2024. Post equals 1 if the transaction date occurs on 24$^{\text{th}}$ May 2024 or afterwards, and 0 otherwise. The pre-announcement phase will include all transactions falling before 22$^{\text{nd}}$ July 2008. The coefficients of interest are the average treatment effect estimators: β_1, β_2 and β_3. Our hypothesis is that they are all positive and significant.

We will test this base model using the network, Euclidean and Manhattan distance measures to compare their results.

5.1.1 Treatment Effects over Space

To test for spatial variations in treatment effects, we divide the treatment group into 5 zones:

$$\begin{aligned}\ln(P) = \alpha &+ \sum_{i=1}^{5} \beta_i(\text{Zone}_i \times \text{Announcement}) + \sum_{j=1}^{5} \delta_j(\text{Zone}_j \times \text{Construction}) \\ &+ \sum_{k=1}^{5} \sigma_k(\text{Zone}_k \times \text{Post}) + H'\gamma + N'\theta + \varphi + \tau + \varepsilon\end{aligned} \tag{2}$$

where each Zone covers 300m distance, except for Zone_5, which ranges 1200–1500m/1650m based on treatment threshold identified for each distance measure earlier.

5.1.2 Treatment Effects across Regions

To identify whether capitalisation effects are stronger in some areas compared to others, we group the 41 stations into 5 regions and rerun equation (1) on the five subdatasets.

The regions are divided into West (Reading to Iver), West London (West Drayton and Heathrow to Acton Main Line), Central London (Paddington to Liverpool Street), East London (Whitechapel to Harold Wood and Abbey Wood) and East (Brentwood and Shenfield).

5.2 Spatial Difference-in-Differences Model Specification

Combining the baseline DID regression with the STAR model produces the SDID specification:

$$\begin{aligned} \ln(P) &= \rho W \ln(P) + X'\eta + u \\ u &= \lambda W u + \varepsilon \end{aligned} \tag{3}$$

where ρ is the spatial autoregressive parameter, $W \ln(P)$ is the spatially lagged dependent variable, X represents all the explanatory variables in equation (1) with coefficient vector η, u is the spatially correlated error term, λ is the spatial error coefficient and Wu is the spatially lagged error term.

5.2.1 The Spatial Weight Matrix and Spatial Weights

The spatial weight matrix, S, is an $n \times n$ positive symmetric and non-stochastic matrix with element s_{ij} at location (i, j) acting as the spatial weight. The magnitude of the elements measure the degree of proximity for each pair of locations. By convention, the diagonal cells equal 0, as self-neighbourhood is excluded. This is the key term that formalises spatial dependence into a model:

$$S = \begin{bmatrix} 0 & s_{12} & \cdots & s_{1n} \\ s_{21} & 0 & \cdots & s_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n1} & s_{n2} & \cdots & 0 \end{bmatrix}$$

For the distance-based spatial weights, we use the negative exponential model for calculating weights:

$$s_{ij} = \begin{cases} \exp(-d_{ij}) & \text{if } i \neq j \text{ and } d_{ij} \leq \overline{d}, \\ 0 & \text{otherwise,} \end{cases}$$

where d_{ij} is the Euclidean distance in kilometres and $\overline{d}$ is a threshold distance determined by fitting an empirical semivariogram. As we are dealing with a temporal dimension, the data needs to be sorted chronologically before any calculations are performed.

5.2.2 The Temporal Weight Matrix

Given the pooled cross-sectional nature of the data, temporal constraints must be imposed onto the spatial matrix. The temporal weight matrix is structurally analogous to the spatial weight matrix, with each element representing the temporal distance between any two observations:

$$T = \begin{bmatrix} 0 & t_{12} & \cdots & t_{1n} \\ t_{21} & 0 & \cdots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ t_{n1} & t_{n2} & \cdots & 0 \end{bmatrix}$$

The temporal distance between observations i and j can be calculated as follows:

$$t_{ij} = \begin{cases} (v_i - v_j)^{-1} & \text{if } 0 < v_i - v_j \leq \overline{v}, \\ 1 & \text{if } v_i = v_j \text{ and } i \neq j, \\ 0 & \text{otherwise,} \end{cases}$$

where v_i is the number of months since the earliest month of the dataset (January 2006) and $\overline{v}$ is the cutoff distance set at 6 months. As $v_i - v_j$ gives us the number of months between i and j, a negative value will correspond with i occurring before j, hence the 0 weight. This results in an lower triangular matrix.

The spatio-temporal weight matrix is computed as the Hadamard product of the spatial and temporal weight matrices:

$$W = S \odot T$$

which is then subjected to row standardisation.

6 Results

6.1 DID Model Results

Table 3 presents the results from equation (1). All coefficients are positive and significant as hypothesised, and vary over time. Under Euclidean distance, capitalisation effects start at 2.80% upon announcement, then rise to 3.56% during the construction years, and then fall to 1.28% upon the line's inauguration. The results follow a similar trend when using Manhattan distance, although the magnitude is higher for the construction and post phases, and lower for announcement.

	(1)	(2)
Treat × Announcement	0.0280***	0.0180**
	(0.0075)	(0.0078)
Treat × Construction	0.0356***	0.0358***
	(0.0017)	(0.0016)
Treat × Post	0.0128***	0.0142***
	(0.0031)	(0.0032)
Property Characteristics	Yes	Yes
Neighbourhood Attributes	Yes	Yes
Spatial Fixed Effects	Yes	Yes
Time Fixed Effects	Yes	Yes
Distance Measure	Euclidean	Manhattan
Observations	309,876	309,876
R^2	0.7997	0.7997
Adjusted R^2	0.7995	0.7995
Residual Std. Error (df = 309615)	0.2975	0.2975
F Statistic (df = 260; 309615)	4,753.5150***	4,753.7940***

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 3: Capitalisation Effects over Time

6.1.1 Hedonic Prices

Although the coefficients for housing and neighbourhood characteristics are not the primary focus of this study, they are not without interest. The coefficients for Floor Area, Bathrooms, Living Rooms and Property Type are positive and significant, as expected. The EPC coefficients are more peculiar. Whilst an EPC rating of B suggests a 13.47% premium compared to properties with the lowest, G, rating, all the other grades are statistically insignificant, implying that homebuyers do not value energy efficiency in general. This result could be because EPC ratings can be expired, or perhaps prospective buyers place value on the potential, rather than current, EPC rating. Regarding neighbourhood characteristics, population density and % over 65 negatively impact price, whereas a higher IMD rating marginally increases values.

	Dependent Variable: $\ln(P)$
Floor Area	0.6111***
	(0.0021)
Bathrooms	0.0994***
	(0.0011)
Living Rooms	0.0359***
	(0.0013)
Terraced	0.1413***
	(0.0017)
Semi-Detached	0.1787***
	(0.0023)
Detached	0.2484***
	(0.0032)
EPC F	0.0176
	(0.0154)
EPC E	0.0103
	(0.0140)
EPC D	−0.0058
	(0.0139)
EPC C	−0.0080
	(0.0139)
EPC B	0.1347***
	(0.0140)
EPC A	−0.0005
	(0.0241)
Population Density	−0.000001***
	(0.000000)
% over 65	−0.2072***
	(0.0153)
IMD	0.00001***
	(0.000000)
Constant	8.9081***
	(0.0178)
Observations	309,876
R^2	0.7997
Adjusted R^2	0.7995
Residual Std. Error	0.2975 (df = 309615)
F Statistic	4,753.5150*** (df = 260; 309615)
	$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 4: Hedonic Prices

6.1.2 Average Treatment Effect over Space

Across all distance measures, Table 5 exhibits a clear pattern where capitalisation effects are stronger for properties closest to stations. Properties within 0–300m and 300–600m experience the greatest LVU, and those between 0 and 900m enjoy above-average LVU. Prima facie, this suggests the lack of negative externalities associated with stations. However, the 0–300m group is the only distance band to observe a decrease in capitalisation effects when moving from the announcement to construction phase, implying that negative externalities, most likely those arising from construction efforts, negatively impacted prices for the most proximate stations. The Manhattan distance measure predicts higher LVU than Euclidean distance, except for the 0–300m and 300–600m categories.

	(1)	(2)
0-300m × Announcement	0.0793***	0.0657**
	(0.0252)	(0.0304)
0-300m × Construction	0.0633***	0.0469***
	(0.0040)	(0.0048)
0-300m × Post	0.0303***	0.0146
	(0.0094)	(0.0118)
300-600m × Announcement	0.0338**	0.0262
	(0.0143)	(0.0184)
300-600m × Construction	0.0674***	0.0651***
	(0.0028)	(0.0032)
300-600m × Post	0.0412***	0.0450***
	(0.0058)	(0.0071)
600-900m × Announcement	0.0451***	0.0543***
	(0.0124)	(0.0137)
600-900m × Construction	0.0556***	0.0669***
	(0.0024)	(0.0026)
600-900m × Post	0.0217***	0.0370***
	(0.0050)	(0.0057)
900-1200m × Announcement	0.0013	0.0034
	(0.0127)	(0.0136)
900-1200m × Construction	0.0260***	0.0262***
	(0.0023)	(0.0024)
900-1200m × Post	0.0065	−0.0003
	(0.0048)	(0.0053)
1200-End × Announcement	0.0406***	0.0038
	(0.0123)	(0.0115)
1200-End × Construction	0.0146***	0.0199***
	(0.0022)	(0.0021)
1200-End × Post	0.0059	0.0083*
	(0.0048)	(0.0043)
Property Characteristics	Yes	Yes
Neighbourhood Attributes	Yes	Yes
Spatial Fixed Effects	Yes	Yes
Time Fixed Effects	Yes	Yes
Distance Measure	Euclidean	Manhattan
Observations	309,876	309,876
R^2	0.8000	0.8000
Adjusted R^2	0.7998	0.7998
Residual Std. Error (df = 309603)	0.2972	0.2972
F Statistic (df = 272; 309603)	4,552.4240***	4,551.8610***

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 5: Spatial Variation in Effects

6.1.3 Average Treatment Effect across Regions

Table 6 reveals large disparities in LVU across regions. Most notably, Central London and East benefit the most, with 7.43% and 10.56% in uplift at the construction stage — over double the average treatment effect for the entire study area. Furthermore, capitalisation effects remain strongest post opening for Central London compared to other regions. East London and West experience below-average LVU, and experience negative effects after opening.

	(1)	(2)	(3)	(4)	(5)
Treat × Announcement	0.0365^{***}	-0.0274	0.0186	0.0285^{***}	0.1641^{***}
	(0.0129)	(0.0174)	(0.0207)	(0.0108)	(0.0388)
Treat × Construction	0.0283^{***}	0.0353^{***}	0.0743^{***}	0.0319^{***}	0.1056^{***}
	(0.0027)	(0.0036)	(0.0061)	(0.0022)	(0.0077)
Treat × Post	-0.0102^{*}	0.0301^{***}	0.0523^{***}	-0.0075^{*}	0.0389^{***}
	(0.0053)	(0.0066)	(0.0105)	(0.0042)	(0.0146)
Property Characteristics	Yes	Yes	Yes	Yes	Yes
Neighbourhood Attributes	Yes	Yes	Yes	Yes	Yes
Spatial Fixed Effects	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes
Region	West	West London	Central London	East London	East
Observations	46,958	54,728	58,826	140,132	9,232
R^2	0.8280	0.8016	0.7671	0.7240	0.8223
Adjusted R^2	0.8276	0.8011	0.7664	0.7237	0.8205
Residual Std. Error	0.2106 (df = 46851)	0.2688 (df = 54606)	0.3881 (df = 58653)	0.2669 (df = 139987)	0.2421 (df = 9136)
F Statistic	$2{,}127.3740^{***}$ (df = 106; 46851)	$1{,}822.8140^{***}$ (df = 121; 54606)	$1{,}123.3600^{***}$ (df = 172; 58653)	$2{,}550.4980^{***}$ (df = 144; 139987)	445.0768^{***} (df = 95; 9136)

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 6: Regional Variation in Effects

6.2 SDID Model Results

	Model 1
Treat × Announcement	0.0825
	(0.1442)
Treat × Construction	0.0237
	(0.0308)
Treat × Post	-0.0438
	(0.0643)
Num. obs.	499
Parameters	97
Log Likelihood	134.2697
AIC (Linear model)	-73.9639
AIC (Spatial model)	-74.5394
LR test: statistic	4.5755
LR test: p-value	0.1015

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 7: SDID Results

The opening date of 24^{th} October is accounted for Bond Street.

		DID					SDID						
Station Name	N	β_1	β_2	β_3	R^2	AIC	β_1	β_2	β_3	ρ	λ	R^2	AIC
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													
33													
34													
35													
36													
37													
38													
39													
40													
41													

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 8: DID and SDID by Station

7 Robustness Checks

7.1 Parallel Trends

Because the choice of station location is almost always influenced by factors such as population density, deprivation levels, potential for development and zoning changes, the non-random nature of station placement complicates proving parallel trends and establishing causal relationships compared to randomised control trials. Some studies use propensity score matching to ensure comparability between treatment and control groups; however, this comes at the cost of reducing the sample size. Other studies compare descriptive statistics between groups or plot their respective time trends. A number of authors have included a formal test for parallel trends prior to the intervention. The idea is to add interaction terms between the treatment variable and time fixed effects to estimate the temporal trend in the treated group relative to the control group (Huang et al., 2024; Zhu & Diao, 2024):

$$\ln(P) = \alpha + \beta \text{Treat} + \sum_{i=1} \delta_i (\text{Treat} \times \tau_i) + H'\gamma + N'\theta + \varphi + \tau + \varepsilon \tag{4}$$

If the coefficients of the interaction terms are statistically insignificant, this suggests no diverging patterns between the treatment and control groups in the pre-treatment period, thereby validating the parallel trends assumption critical to DID models.

Table 9 shows no significant interaction coefficients. Therefore, the fundamental assumption of parallel trends is satisfied.

Quarter	Coefficient
2006 Q2	0.0018
	(0.0099)
2006 Q3	0.0009
	(0.0096)
2006 Q4	0.0120
	(0.0096)
2007 Q1	−0.0010
	(0.0099)
2007 Q2	0.0052
	(0.0097)
2007 Q3	−0.0037
	(0.0096)
2007 Q4	−0.0124
	(0.0101)
2008 Q1	0.0007
	(0.0110)
2008 Q2	−0.0092
	(0.0111)
2008 Q3	−0.0071
	(0.0228)
Observations	50,995
R^2	0.7885
Adjusted R^2	0.7877
Residual Std. Error	0.2439 (df = 50800)
F Statistic	976.3418*** (df = 194; 50800)

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 9: Parallel Trends Test

We can visually represent this by constructing a housing price index for the treatment and control groups based on the quarterly fixed effects.

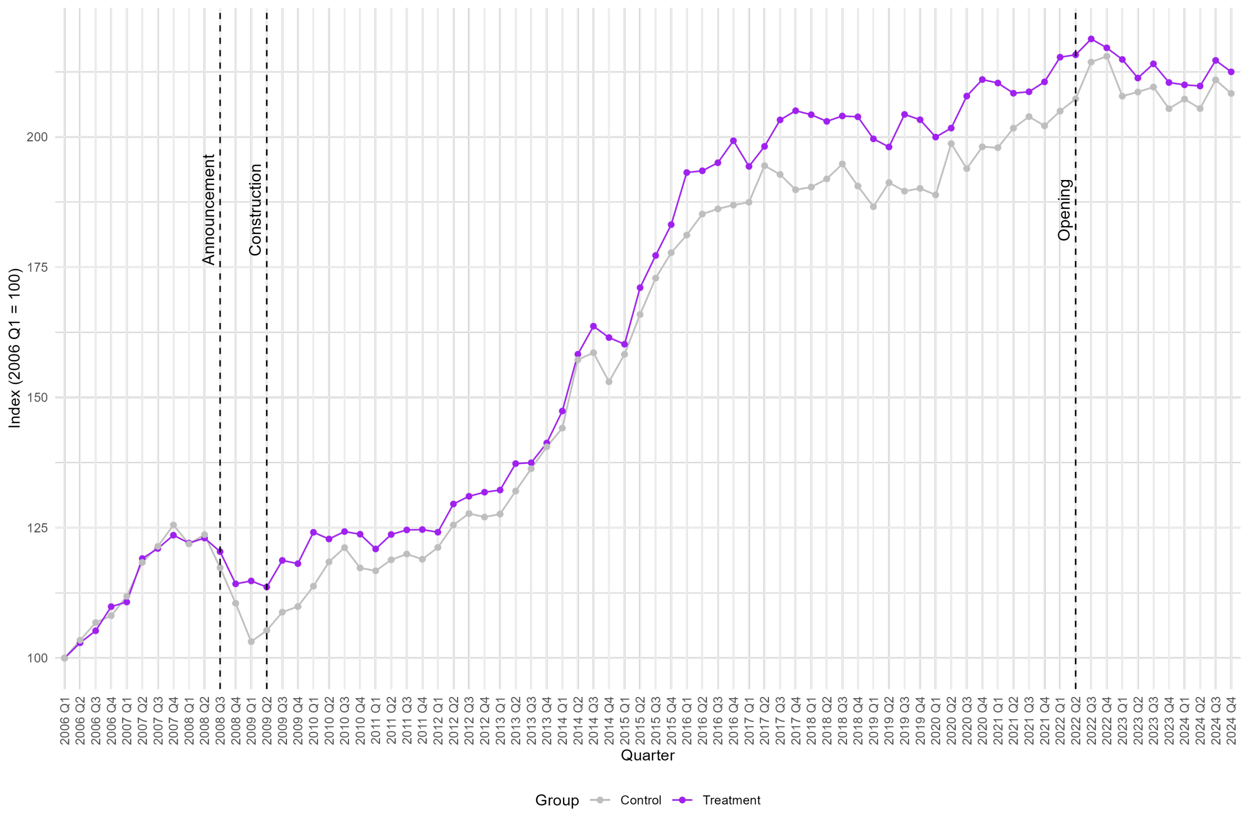


Figure 6: Quarterly Housing Price Index

7.2 Treatment Zone

To test the sensitivity of the treatment effects, we also test equation (1) using several other treatment thresholds. We use 800m given its wide usage in the literature, 1000m based on a few academic and commercial studies that use this as the boundary and 2000m as used by Gibbons and Machin (2005).

Table 10 displays a clear trend of treatment effects increasing with potency as the treatment boundary declines. This is to be expected given the results in Figure 5 and Table 5, whereby the location premium and capitalisation effects generally fall with distance from the nearest station. As the 2000m boundary includes properties unaffected by new stations within the treatment group, the results become diluted to the point where it is near zero.

Treatment Boundary	800m	1000m	2000m
Treat × Announcement	0.0380***	0.0323***	−0.0098
	(0.0098)	(0.0086)	(0.0074)
Treat × Construction	0.0499***	0.0467***	0.0037**
	(0.0019)	(0.0017)	(0.0017)
Treat × Post	0.0279***	0.0163***	−0.0111***
	(0.0040)	(0.0035)	(0.0032)
Property Characteristics	Yes	Yes	Yes
Neighbourhood Attributes	Yes	Yes	Yes
Spatial Fixed Effects	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes
Treatment Boundary	800m	1000m	2000m
Observations	309,876	309,876	309,876
R^2	0.7998	0.7999	0.7994
Adjusted R^2	0.7997	0.7997	0.7992
Residual Std. Error (df = 309615)	0.2973	0.2973	0.2977
F Statistic (df = 260; 309615)	4,758.5590***	4,758.8800***	4,744.9860***

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 10: Sensitivity Analysis

8 Discussion

8.1 Cost-Benefit Analysis

Using capitalisation effects during construction from Table 3 based on network distance, this tells us that residential properties near an Elizabeth line station experienced, on average, capitalisation effects of 3.56%. Given that the average housing value within the treatment group is £467,571, this implies an average LVU of £16,645.53. Since there are 147,310 residential properties in the treatment group, the total LVU resulting from the Elizabeth line amounts to £2.45 billion. With an estimated final cost of £18.9bn, this results in approximately 13% of total costs.

The total LVU will of course vary depending on which model's results one chooses to apply, and a more refined analysis would involve repeating this exercise for each phase, for each distance band within each region. Nevertheless, £2.45 billion in windfall gains arising from accessibility improvements is a substantial number, and represents a non-trivial proportion of the line's construction costs.

8.2 Relative Beneficiaries

Despite the aggregate benefits derived from the Elizabeth line, these gains do not necessarily benefit everyone, nor are they universally distributed. At a basic level, the groups who have benefitted from the intervention include: users of the line, firms involved in its construction and operation, ancillary beneficiaries (of improved air quality, reduced congestion elsewhere, housing development, macroeconomic growth) and, of course, property owners and firms located near the line. Although the taxpayer bears a cost, the net effect will ultimately depend on which group they belong to. Non-users who do not experience any indirect benefits will bear the greatest burden — most likely those living particularly far from the line.

As we have seen, even for homeowners within the treatment zone, windfall gains are not homogenous. Properties relatively closer to stations experience greater LVU, and properties in Central London, Canary Wharf and in the East region benefit more than twice over the average LVU observed across the entire study area. Given that the median annual income in 2023–24 was £37,430, this windfall gain represents almost 45% of the median person's yearly income. Within the favourable areas, where the average treated housing price will almost certainly be higher than the sample's average, LVU will easily exceed 100%.

If it is assumed that wealthier individuals are more likely to own housing, as is very likely the case in London and its neighbouring areas, such a result would indicate that, all else equal, the line exacerbates wealth inequality via LVU. If rents were to rise in treated areas, this would represent an additional redistribution of funds from tenants to landlords, which has additional implications for income inequality. Again, the wealth disparities will be unevenly concentrated across regions, with Central London and East generating more wealth than other Elizabeth line areas.

It is important to note, however, that these results must be contextualised with the other socioeconomic and environmental benefits produced by the Elizabeth line. The line ultimately helps connect previously underserved communities in East and South East London to the city centre, and in the process, stimulates local economies and offers easier access to opportunities and resources elsewhere. While property-driven wealth inequalities may arise, they should not overshadow the overarching benefits and goal of regional development.

8.3 Land Value Capture

The prior discussion presents a clear case for capturing some of these windfall gains. Designing a perfect LVC mechanism goes beyond the scope of this paper; however, it is worth discussing the current state of affairs.

There are various arguments for implementing LVC. While the marginal social costs of taking public transport typically exceed its marginal private cost, fares are oftentimes both capped and subsidised in order to keep public transport as accessible as possible; this means that fare revenues alone may be insufficient to recoup construction and maintenance costs in the short term. From a fairness perspective, homeowners near Crossrail stations experience a windfall gain in the value of their property — gains funded by taxpayers. As these homeowners benefit disproportionately, without any personal effort, it is fair to expect them to contribute towards the financing of such projects, thus helping to distribute the benefits more equitably. Monetising LVU may also facilitate speed and efficiency by mitigating any financing-related delays, and create incentives for completion, thereby possibly even preventing cancellations.

The UK currently has four property taxes: Council Tax, Business Rates, Stamp Duty Land Tax and Capital Gains Tax. However, all four are ineffective at capturing LVU because they are relatively unresponsive to changes in property values. Furthermore, London already employs a Business Rates Supplement on commercial properties to help fund transport improvements, whereby a 2p per £1 levy on high-value business properties was used to raise over £4 billion towards Crossrail. It should be noted that LVC, if implemented, will unlikely be a silver bullet for public financing. With our estimate of LVU covering 13% of project costs, this represents the upper bound of what LVC can hope to capture in the ideal scenario where it taxes 100% of windfall gains. However, this 13% figure is by no means inconsequential, and would be amplified with BRS among other LVC techniques on the rental market.

Hong Kong's "Rail + Property" joint development model offers a celebrated example of LVC through the integration of transit and real estate development. The Mass Transit Railway Corporation obtains development rights for land around new stations at pre-rail values, and uses the post-rail values and profits to recoup its costs. More than half of the Corporation's income is from property development rather than fares, enabling the transit system to remain profitable and even pay dividends to the government, rather than requiring subsidies (Cervero & Murakami, 2009). As a result, this model may be regarded as the gold standard for LVC, and can also be applied to finance large infrastructure projects in mainland China; however, its viability relies on the areas surrounding the stations possessing ample potential for development. In the case of Hong Kong's West Island Line extension, the mature character of neighbourhoods hindered the programme's implementation. Furthermore, Hong Kong has a public leasehold system, in which the government owns virtually all land, empowering them to easily grant long-term leases. Due to private ownership and differing planning laws, transposing the Rail + Property model to the UK would be challenging, but nevertheless offers an alternative development-based approach to LVC.

To capture windfall gains on the residential side, in their LVC report, TfL (2017) proposes several mechanisms. This includes Community Infrastructure Levy and Section 106 requirements for new developments near major transport projects to help capture uplift without directly taxing homeowners. Additionally, allowing a portion of Stamp Duty Land Tax growth in transport-enhanced areas to be reinvested locally would provide a steady revenue stream without imposing new levies. Finally, TfL could expand its joint venture model to retain long-term stakes in development on public land, mirroring a scaled-down version of Hong Kong's approach. Political opposition and disputes over taxes based on property valuations will inevitably rise against these proposals, paralleling some of the demerits of a land value tax, which has been proposed since the late 19th century, but never implemented in the UK. Nonetheless, a hybrid strategy that leverages these mechanisms could help London systematically capture the value it creates, ensuring future transport projects are more self-sustaining.

8.4 Limitations

Despite robustness checks and careful model design, there are nonetheless limitations inherent to this paper.

First, there are many arbitrary choices that have to be made: from selecting the study boundary to deciding how to model spatio-temporal weights. Whilst we have tried to limit the number of arbitrary decisions, such as using HPM to determine a treatment boundary and using an empirical semivariogram to set the cutoff distance for spatial weights, many arbitrary choices had to be made elsewhere. Despite justifications, the decision to choose one method or value over another will inevitably impact the model's results. Within the transit-induced LVU literature, many questions remain an empirical question.

Second, constraints in housing data often plague many real estate studies. Although all housing transactions are publicly available through HM Land Registry, the data lacks sufficient structural characteristics. Using repeat-sales or DID is suboptimal given they cannot difference out time-varying traits. Although Zoopla provides housing characteristics, two issues arise. One is that a subset of the total database of transactions on Zoopla contains all housing characteristics. Thus, even with a large sample size, if properties that have data on all characteristics on Zoopla are fundamentally different from those which are missing at least one variable, the results will not be truly representative of the population. Furthermore, Zoopla does not provide all time-varying characteristics, such as age, new build status and potential EPC rating, which could lead to higher model explanatory power and more detailed insights. In a similar vein, interpolating and extrapolating neighbourhood data is less reliable compared to a first-best world where it is regularly updated.

Regarding model design, would have been possible to include even more detail. For instance, we could have separately accounted for Bond Street's late opening in the earlier models, examined treatment effects over space across regions or stations, subdivided the dataset by property type and added more phases, specifically within the construction phase to separate out earlier years from the later years that experienced delays and the pandemic. Whilst these enhanced methods would indeed result in additional insights, they also risk overfitting the models, reducing their simplicity and generalisability.

Finally, although running station-level SDID models was necessary given the large sample sizes and memory constraints, conducting 41 separate tests increases the risk of Type I and Type II errors. With optimisations to spatial regression libraries and advancements in computer hardware, one can hope that spatial regression techniques will become more feasible for larger datasets in the future.

9 Conclusion

Using DID to identify the effects of the Elizabeth line on residential property prices across its lifespan, our models show that the line has increased housing values for those within the treatment zone. Although the average capitalisation effects are valued at 2.80%, 3.56% and 1.28% for the announcement, construction and operational phases respectively, there is significant variation across space and time. Properties closest to stations experience greater LVU compared to those further away but still within the treatment zone, and properties within Central London, Canary Wharf and East experience more than double the average LVU found across the entire dataset. The largest uplift is typically experienced during the construction phase, followed by post operation and announcement.

These findings corroborate with the broader transit-induced LVU literature, accentuating the channel in which accessibility benefits are capitalised into property values. The results are robust upon changing the treatment zone boundary and satisfy the parallel trends assumption. We also advocate for the Manhattan distance measure as a middle ground between Euclidean and network distance.

With an estimated aggregate LVU of £2.54 billion, or 13% of the line's total construction costs, there is reason to implement LVC mechanisms for residential properties experiencing uplift in order to help finance future transport infrastructure that go on to improve and shape the cities we live in. In terms of future direction, one can go beyond and also estimate the network effects arising from the Elizabeth line to obtain a fuller picture of its impacts beyond our study boundary. Furthermore, exploring pre-opening capitalisation effects of HS2 can shed more light on LVU dynamics within the UK, which in turn builds more evidence towards greenlighting proposed transport lines such as Crossrail 2.

10 Bibliography

Alonso, W. (1964). *Location and land use: Toward a general theory of land rent*. Harvard University Press.

Anas, A., & Chang, H. (2023). Productivity benefits of urban transportation megaprojects: A general equilibrium analysis of «Grand Paris Express». *Transportation Research Part B: Methodological*, *174*, 102746. https://doi.org/10.1016/j.trb.2023.03.006

Anderson, M. L. (2013). *Subways, strikes, and slowdowns: The impacts of public transit on traffic congestion* (Working Paper No. 18757). National Bureau of Economic Research. https://doi.org/10.3386/w18757

Anselin, L. (1988). *Spatial econometrics: Methods and models*. Kluwer Academic Publishers.

Bae, C.-H. C., Jun, M.-J., & Park, H. (2003). The impact of Seoul's subway Line 5 on residential property values. *Transport Policy*, *10*(2), 85–94. https://doi.org/10.1016/S0967-070X(02)00048-3

Bao, H. X., Larsson, J. P., & Wong, V. (2021). Light at the end of the tunnel: The impacts of expected major transport improvements on residential property prices. *Urban Studies*, *58*(14), 2971–2990. https://doi.org/10.1177/0042098020967308

Bauernschuster, S., Hener, T., & Rainer, H. (2017). When labor disputes bring cities to a standstill: The impact of public transit strikes on traffic, accidents, air pollution, and health. *American Economic Journal: Economic Policy*, *9*(1), 1–37. https://doi.org/10.1257/pol.20150414

Bennett, S. (2017). Crossrail project to deliver London's Elizabeth line: the parliamentary bill process. *Proceedings of the Institution of Civil Engineers - Civil Engineering*, *170*(6), 10–14. https://doi.org/10.1680/jcien.17.00014

Billings, S. B. (2011). Estimating the value of a new transit option. *Regional Science and Urban Economics*, *41*(6), 525–536. https://doi.org/10.1016/j.regsciurbeco.2011.03.013

Bowes, D. R., & Ihlanfeldt, K. R. (2001). Identifying the impacts of rail transit stations on residential property values. *Journal of Urban Economics*, *50*(1), 1–25. https://doi.org/10.1006/juec.2001.2214

CBRE. (2024). *The Elizabeth line: The impact on London's housing market*. https://www.cbre.co.uk/insights/reports/the-elizabeth-line-the-impact-on-london-s-housing-market

Centre for Cities. (2025). *Cities outlook 2025*. Centre for Cities. https://www.centreforcities.org/wp-content/uploads/2025/01/Cities-Outlook-2025.pdf

Cervero, R., & Murakami, J. (2009). Rail and property development in Hong Kong: Experiences and extensions. *Urban Studies*, *46*(10), 2019–2043. https://doi.org/10.1177/0042098009339431

Clower, T. L., & Weinstein, B. L. (2002). The impact of Dallas (Texas) area rapid transit light rail stations on taxable property valuations. *The Australasian Journal of Regional Studies*, *8*(3), 389–400. https://search.informit.org/doi/10.3316/ielapa.028492369068664

Debrezion, G., Pels, E., & 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. https://doi.org/10.1007/s11146-007-9032-z

Diao, M., Leonard, D., & Sing, T. F. (2017). Spatial-difference-in-differences models for impact of new mass rapid transit line on private housing values. *Regional Science and Urban Economics*, *67*, 64–77. https://doi.org/10.1016/j.regsciurbeco.2017.08.006

Du, H., & Mulley, C. (2007). The short-term land value impacts of urban rail transit: Quantitative evidence from Sunderland, UK. *Land Use Policy*, *24*(1), 223–233. https://doi.org/10.1016/j.landusepol.2005.12.003

Dubé, J., Le Gallo, J., Des Rosiers, F., Legros, D., & Champagne, M.-P. (2024). An integrated causal framework to evaluate uplift value with an example on change in public transport supply. *Transportation Research Part E: Logistics and Transportation Review*, *185*, 103500. https://doi.org/10.1016/j.tre.2024.103500

Dubé, J., Legros, D., Thériault, M., & Des Rosiers, F. (2014). A spatial difference-in-differences estimator to evaluate the effect of change in public mass transit systems on house prices. *Transportation Research Part B: Methodological*, *64*, 24–40. https://doi.org/10.1016/j.trb.2014.02.007

Gatzlaff, D. H., & Haurin, D. R. (1997). Sample selection bias and repeat-sales index estimates. *The Journal of Real Estate Finance and Economics*, *14*(1), 33–50. https://doi.org/10.1023/A:1007763816289

Gatzlaff, D. H., & Smith, M. T. (1993). The impact of the Miami Metrorail on the value of residences near station locations. *Land Economics*, *69*(1), 54–66. https://doi.org/10.2307/3146278

Gibbons, S., & Machin, S. (2005). Valuing rail access using transport innovations. *Journal of Urban Economics*, *57*(1), 148–169. https://doi.org/10.1016/j.jue.2004.10.002

Guerra, E., Cervero, R., & Tischler, D. (2012). Half-mile circle: Does it best represent transit station catchments? *Transportation Research Record*, *2276*(1), 101–109. https://doi.org/10.3141/2276-12

GVA & Crossrail Ltd. (2018). *Crossrail property impact & regeneration study 2012-2026*. https://learninglegacy.crossrail.co.uk/wp-content/uploads/2018/07/4D-003-crossrail_property_impact_regeneration_study.pdf

He, S. Y. (2020). Regional impact of rail network accessibility on residential property price: Modelling spatial heterogeneous capitalisation effects in Hong Kong. *Transportation*

Research Part A: Policy and Practice, *135*, 244–263. https://doi.org/10.1016/j.tra.2020.01.025

Hess, D. B., & Almeida, T. M. (2007). Impact of proximity to light rail rapid transit on station-area property values in buffalo, new york. *Urban Studies*, *44*(5-6), 1041–1068. https://doi.org/10.1080/00420980701256005

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. https://doi.org/10.1016/j.landurbplan.2018.12.011

Higgins, C. D., Arku, R. N., Farber, S., & Miller, E. J. (2024). Modelling changes in accessibility and property values associated with the King Street Transit Priority Corridor project in Toronto. *Transportation Research Part A: Policy and Practice*, *190*, 104256. https://doi.org/10.1016/j.tra.2024.104256

Higgins, C. D., & Kanaroglou, P. S. (2016). Forty years of modelling rapid transit's land value uplift in North America: Moving beyond the tip of the iceberg. *Transport Reviews*, *36*(5), 610–634. https://doi.org/10.1080/01441647.2016.1174748

Huang, Y., Parker, D. C., Babin, R., & Kong, F. (2024). Causal identification of transit-induced property value uplift in Canada's Waterloo Region: A spatio-temporal difference-in-differences method application. *Cities*, *145*, 104676. https://doi.org/10.1016/j.cities.2023.104676

Hurst, N. B., & West, S. E. (2014). Public transit and urban redevelopment: The effect of light rail transit on land use in Minneapolis, Minnesota. *Regional Science and Urban Economics*, *46*, 57–72. https://doi.org/10.1016/j.regsciurbeco.2014.02.002

Hyun, D., & Milcheva, S. (2019). Spatio-temporal effects of an urban development announcement and its cancellation on house prices: A quasi-natural experiment. *Journal of Housing Economics*, *43*, 23–36. https://doi.org/10.1016/j.jhe.2018.09.008

Landis, J., Guhathakurta, S., Huang, W., & Zhang, M. (1995). *Rail transit investments, real estate values, and land use change: A comparative analysis of five California rail transit systems* (University of California Transportation Center, Working Papers). University of California Transportation Center. https://escholarship.org/uc/item/2hf9s9sr

Ma, L., Graham, D. J., & Stettler, M. E. (2021). Air quality impacts of new public transport provision: A causal analysis of the Jubilee Line Extension in London. *Atmospheric Environment*, *245*, 118025. https://doi.org/10.1016/j.atmosenv.2020.118025

McDonald, J. F., & Osuji, C. I. (1995). The effect of anticipated transportation improvement on residential land values. *Regional Science and Urban Economics*, *25*(3), 261–278. https://doi.org/10.1016/0166-0462(94)02085-U

McMillen, D. P., & McDonald, J. (2004). Reaction of house prices to a new rapid transit line: Chicago's midway line, 1983–1999. *Real Estate Economics*, *32*(3), 463–486. https://doi.org/10.1111/j.1080-8620.2004.00099.x

Mills, E. S. (1972). *Studies in the structure of the urban economy*. Johns Hopkins University Press.

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*, 158–170. https://doi.org/10.1016/j.tra.2013.01.013

Mohammad, S. I., Graham, D. J., & Melo, P. C. (2017). The effect of the Dubai Metro on the value of residential and commercial properties. *Journal of Transport and Land Use*, *10*(1), 263–290. https://doi.org/10.5198/jtlu.2015.750

Muth, R. F. (1969). *Cities and housing: The spatial pattern of urban residential land use*. University of Chicago Press.

Nationwide. (2021). *Transport special feature: June 2021*. https://www.nationwidehousepriceindex.co.uk/reports/london-sees-biggest-house-price-premium-as-households-still-attach-significant-value-to-transport-links-despite-the-pandemic

Nelson, A. C. (1992). Effects of elevated heavy-rail transit stations on house prices with respect to neighborhood income. *Transportation Research Record*, (1359). http://onlinepubs.trb.org/Onlinepubs/trr/1992/1359/1359-016.pdf

Nelson, A. C., & McCleskey, S. J. (1990). Improving the effects of elevated transit stations on neighborhoods. *Transportation Research Record*, (1226), 173–180. http://onlinepubs.trb.org/Onlinepubs/trr/1990/1266/1266-017.pdf

Qiu, F., & Tong, Q. (2021). A spatial difference-in-differences approach to evaluate the impact of light rail transit on property values. *Economic Modelling*, *99*. https://doi.org/10.1016/j.econmod.2021.03.015

Rennert, L. (2022). A meta-analysis of the impact of rail stations on property values: Applying a transit planning lens. *Transportation Research Part A: Policy and Practice*, *163*, 165–180. https://doi.org/10.1016/j.tra.2022.06.013

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, *82*(1), 34–55. https://doi.org/10.1086/260169

Sener, I. N., Lee, R. J., & Elgart, Z. (2016). Potential health implications and health cost reductions of transit-induced physical activity. *Journal of Transport & Health*, *3*(2), 133–140. https://doi.org/10.1016/j.jth.2016.02.002

Transport for London. (2022). *Elizabeth line: Evidencing the value*. https://tfl.gov.uk/corporate/publications-and-reports/elizabeth-line-benefits-framework

Transport for London. (2024). *Evidencing the value of the Elizabeth line: An update on the outcomes of London's transformational railway*. https://tfl.gov.uk/corporate/publications-and-reports/elizabeth-line-benefits-framework

Transport for London & Greater London Authority. (2017). *Land value capture: Final report*. https://www.london.gov.uk/sites/default/files/land_value_capture_report_transport_for_london.pdf

Tucker, W. (2017). Crossrail project: the execution strategy for delivering London's Elizabeth line. *Proceedings of the Institution of Civil Engineers - Civil Engineering*, *170*(5), 3–14. https://doi.org/10.1680/jcien.16.00021

Von Thünen, J. H. (1966). *The isolated state: An English translation of Der isolierte Staat* (P. Hall, Ed.; C. M. Wartenberg, Trans.). Pergamon Press. (Original work published 1826)

Wagner, G. A., Komarek, T., & Martin, J. (2017). Is the light rail "Tide" lifting property values? Evidence from Hampton Roads, VA. *Regional Science and Urban Economics*, *65*, 25–37. https://doi.org/10.1016/j.regsciurbeco.2017.03.008

World Bank. (2023). *Urban development* [Accessed: December 19, 2024]. https://www.worldbank.org/en/topic/urbandevelopment/overview

Xu, T., & Zhang, M. (2016). Tailoring empirical research on transit access premiums for planning applications. *Transport Policy*, *51*, 49–60. https://doi.org/10.1016/j.tranpol.2016.03.003

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*, 412–422. https://doi.org/10.1016/j.landusepol.2018.02.007

Zhang, M. (2023). Value uplift from transit investment-Property value or land value? A case study of the Gold Coast light rail system in Australia. *Transport Policy*, *132*, 88–98. https://doi.org/10.1016/j.tranpol.2022.12.005

Zhu, Y., & Diao, M. (2024). The local and network effects of rail transit network expansion on retail property values. *Journal of Planning Education and Research*, *44*(3), 1820–1834. https://doi.org/10.1177/0739456X221121247