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

Lawrence Lee

University of Cambridge

Note

This is still a very early work in progress. I will be continuously working on this throughout the next few months (starting from late October until March/April), and so will hopefully become much more populated very soon.

Contents

1 Literature Review **4**

2 The Elizabeth Line **7**

3 Estimation Strategy **8**

4 Bibliography **9**

1 Literature Review

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 geocoded data using Geographic Information System (GIS), we are able to consider complex spatial effects that would have otherwise been overlooked by traditional methods.

1.1 Theoretical Foundations of the City

Alonso (1964), Muth (1969) and Mills (1972) famously argued that households trade-off between commuting cost and housing consumption in their residential location decisions. This formed the basis of the AMM model and is the fundamental premise for the monocentric city model that predicts a negative housing price gradient with respect to distance to the employment centre. In urban economics, this negative bid rent curve influences the spatial distribution of households and firms across a city and shapes the urban landscape.

Many empirical studies have shown that new urban travel infrastructure investments could significantly flatten the bid rent gradient, and reduce the housing price gap between urban and rural areas. The increased accessibility brought by new travel investment is translated into housing wealth accrued to local residents via capitalisation effects.

1.2 Hedonic Pricing Models

Modern hedonic pricing theory, originating from the work of Rosen (1974), has become a widely used methodology in the applied economics and real estate literature. Groundbreaking at the time, this technique enables researchers to quantify the implicit price of ordinarily unobservable attributes simply by running an Ordinary Least Squares (OLS) regression on observable variables. As a result, difficult questions such as determining the value of a statistical life, environmental quality and crime could now be answered with the appropriate dataset.

Unsurprisingly, this approach was adopted into studies on the effects of transport infrastructure on housing prices. By including a rich set of housing and spatial attributes, researchers are able to determine the market price of location.

However, OLS unfortunately suffers from serious endogeneity issues when applied to the study of RTS effects. This is because the models are unable to separate unobserved factors that could influence the covariance between housing price and RTS accessibility. This stems from the fact that the consumer simultaneously chooses both the quantity and marginal price of the characteristic. For OLS, the explanatory variables must truly be exogenous and uncorrelated with the error term.

Researchers use repeated sales data to control for endogeneity in modelling housing price changes. While the repeated-sales approach removes biases caused by time-invariant omitted variables by taking the first differencing in housing prices, the sample size is drastically reduced in the process. Furthermore, there may be selection biases if house that are sold at least two times possess different attributes from those that sell just once. Changes in structural characteristics and amenities in local areas can occur over a long period of time, and sometimes randomly, thus differencing the repeated housing prices alone may not adequately remove intertemporal effects of both observed and unobserved factors.

1.3 Difference in Differences

The quasi-experimental approach has become increasingly popular in the regional and urban economics literature. In a randomised experiment, sample houses are sorted into a treatment group and control group. A random event — such as the opening of a new line — is used to simulate an exogenous shock to housing prices. The difference-in-difference (DID) model then tests the before-and-after price changes between the two groups following the shock whilst controlling for observed and unobserved variations in housing and spatial factors. If the pre-existing within-group price variations change after the treatment, causality of the new stations on housing prices can be established.

1.4 Spatial Difference in Differences

In the context of areas undergoing changes following the opening of a new station, there is a need to control for possible autoregressive lag and error in the spatial interactions of houses. This is why spatial difference-in-difference (SDID) models have been increasingly used. Diao et al. (2017) calibrate the SDID models to estimate the capitalisation effect of a new transit line in Singapore that accounts for the spatial dependencies in the dependent variable and the error term.

To move from DID to SDID, one must incorporate 1) a spatial lag term (SAC) to account for the spatially dependent responses and 2) a spatial error term (SARAR).

1.5 Defining the Treatment Area and Distance Measures

Most of the early studies use a Euclidean distance measure from the property to the closest station as a variable in hedonic pricing models to capture cross-sectional variations in housing prices.

Furthermore, earlier studies define the treatment area using proximity measures such as linear distance or buffer zone, which suffers from omitted variable problems, most particularly when the selection of the station location is not exogenous, but correlated with some other factor including population density. The effect in a densely populated city centre may co-vary with other amenities in the area.

Diao et al. (2017) were among the first to propose using network distance as a more appropriate distance measure, as it accounts for topographic features including natural obstacles and spatial constraints. They also use a local polynomial regression to identify the treatment zone.

1.6 Related Studies

A large number of empirical studies find evidence of positive capitalisation effects of proximity to stations in housing values. Evidence of positive capitalisation effects of urban rail transit systems is shown in many studies across many countries over time. Debrezion et al. (2007) provide a meta-analysis of 57 cities and suggest that property values increase by 2.3% for every 250m closer to a railway station.

There exists a few studies that find insignificant or even negative effects of proximity to selected stations in cities. The negative externalities are usually caused by noise and high crime rates found in these areas near stations.

2 The Elizabeth Line

The Elizabeth line is a heavy rail transit service that runs across London, with its peripheries extending to Reading and Shenfield. Named after HM Queen Elizabeth II, the line opened on 17th May 2022 during her Platinum Jubilee year. Operations commenced soon after on 24th May 2022.

Plans for the Elizabeth line first began in 2001. Under a different alias of Crossrail, the system was approved in 2007. The Crossrail Act 2008 authorised the construction project, receiving royal assent on 22nd July 2008.[1] Construction works began on 15th May 2009, starting in Canary Wharf. Though originally planned to open in 2018, the project faced repeated delays, including for several months as a result of the COVID-19 pandemic.

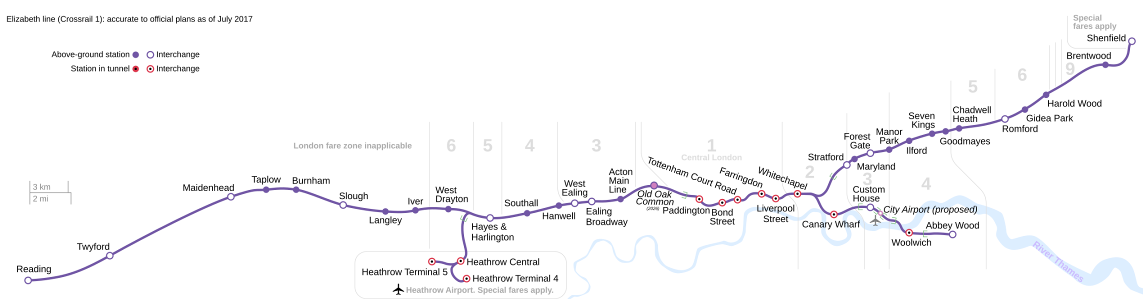

Figure 1: The Elizabeth Line Map

The line reached over 200 million trips annually in its second year of operation. Today, it carries one seventh of all trips by rail in the United Kingdom.

[1]https://www.legislation.gov.uk/ukpga/2008/18/contents

3 Estimation Strategy

For our baseline results, we start with the difference-in-differences model:

$$\ln(P) = \alpha + \beta_1 \text{Post} + \beta_2 \text{Treat} + \beta_3 (\text{Treat} \times \text{Post}) + H'\gamma + N'\theta + \varphi + \tau + \epsilon \quad (1)$$

where $\ln(P)$ is the log of property price, α is the intercept, Post is a post-treatment indicator, Treat is a treatment indicator, Treat $\times$ Post captures the average treatment effect, $H'\gamma$ represents property characteristics, $N'\theta$ represents neighbourhood characteristics, φ are spatial fixed effects, τ are time fixed effects and ϵ is the error term.

The spatial lag model is formally defined as:

$$\begin{aligned} \ln(P) &= \rho W \ln(P) + X'\eta + u \\ u &= \lambda W u + \epsilon \end{aligned} \quad (2)$$

where ρ is the spatial lag coefficient, $W \ln(P)$ is the spatially lagged dependent variable, $X'\eta$ represents a vector of explanatory variables with coefficients η, u is the spatially correlated error term, λ is the spatial error coefficient, Wu is the spatially lagged error term and ϵ is the random error.

The spatial weight matrix is represented as:

$$W_{ij} = \begin{cases} \frac{1}{d_{ij}} & \text{if } i \neq j \text{ and units } i \text{ and } j \text{ are neighbors,} \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

where W_{ij} is the spatial weight between units i and j, d_{ij} is the distance between them, with weights assigned based on inverse distance for neighbouring units and set to 0 otherwise.

4 Bibliography

Alonso, W. (1964). *Location and Land Use: Toward a General Theory of Land Rent*. Harvard University Press. https://doi.org/doi:10.4159/harvard.9780674730854

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*, 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

Mills, E. S. (1972). Studies in the Structure of the Urban Economy.

Muth, R. (1969). *Cities and Housing: The Spatial Pattern of Urban Residential Land Use*. University of Chicago Press.

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