Mind the Gap: Windfall Gains in House 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 Plan

Rationale: Topic relates to real estate, urban and housing economics and is of personal interest to me. Given the novelty of this intervention, there is a large enough gap in the existing literature for this to fill, as well as to advise on its policy implications.

Aim: With the opening of the Elizabeth Line in 2022, the study aims to test whether its inauguration has increased house prices in the areas surrounding the line's stations and, if so, by exactly how much on average.

Methodology: Regression difference-in-difference model.

Structure: Abstract, introduction, literature review, methodology, results, discussion, conclusion and bibliography.

2 Estimation

The regression difference-in-difference (DiD) model takes the form:

$$Y_{it} = \alpha + \beta_1 \text{Post}_t + \beta_2 \text{Treatment}_i + \beta_3(\text{Post}_t \times \text{Treatment}_i) + \gamma X_{it} + \epsilon_{it}$$

where:

- Y_{it} is the property price for unit i at time t,
- Post_t is the indicator variable for the post-construction period,
- Treatment_i is the indicator variable for the treatment group,
- $\text{Post}_t \times \text{Treatment}_i$ is the interaction term to estimate the treatment effect and
- X_{it} is other control variables.

The key coefficient of interest is $\beta_3(\text{Post}_t \times \text{Treatment}_i)$. The estimated value will answer the question of whether there have been a statistically significant increase in house values as a result of the construction of the Elizabeth Line, and also by how much on average.

The main control variables would be those of the hedonic regression below:

$$\ln(\text{Price}_i) = \alpha + \beta_1(\text{Size}_i) + \beta_2(\text{Bedrooms}_i) + \beta_3(\text{Bathrooms}_i) + \beta_4(\text{Age}_i) + \beta_5(\text{PropertyType}_i) + \beta_6(\text{EPC}_i) + \beta_7(\text{DistancetoStation}_i) + \beta_8(\text{Zone}_i)$$