

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

Dissertation Presentation

Lawrence Lee

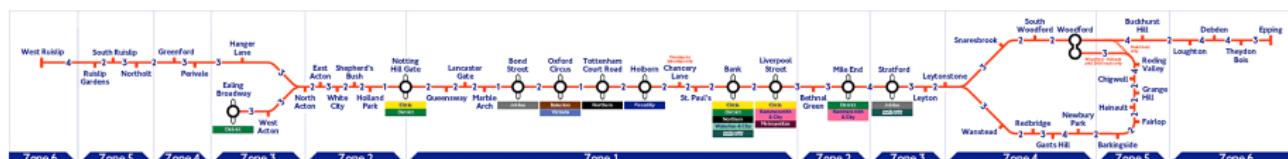
Friday, 21<sup>st</sup> March 2025

# Introduction

- **Aim:** To measure the capitalisation effects on residential property prices resulting from the Elizabeth line
- **Significance:** Important implications for land value capture
- Inspiration emerged whilst commuting on the Central line

## Central line

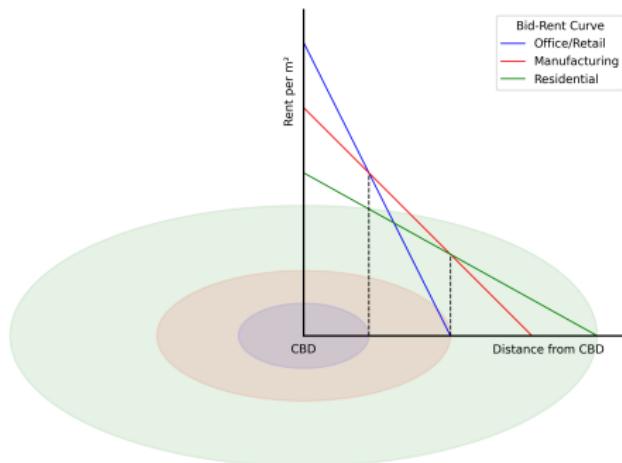
Off-peak running times between stations  
Please allow slightly longer for your journey during peak hours



# Literature Review

## Theoretical Foundations

- Trade-off between rent and distance under monocentric city model



- Implies accessibility is primary determinant of housing prices
- New transport infrastructure reduces commuting costs  
⇒ higher accessibility ⇒ land value uplift

# Literature Review

## Methods

- Earlier studies used hedonic pricing models
  - Accessibility captured by a distance from nearest station variable
  - Differencing regressions pre and post intervention
- Movement in favour of DID to isolate causality
  - Treatment zone often set at 800m or 1/2 mile
- Recent trend towards spatial DID models
- Many studies use the Euclidean distance for simplicity

# Literature Review

## Research Gap

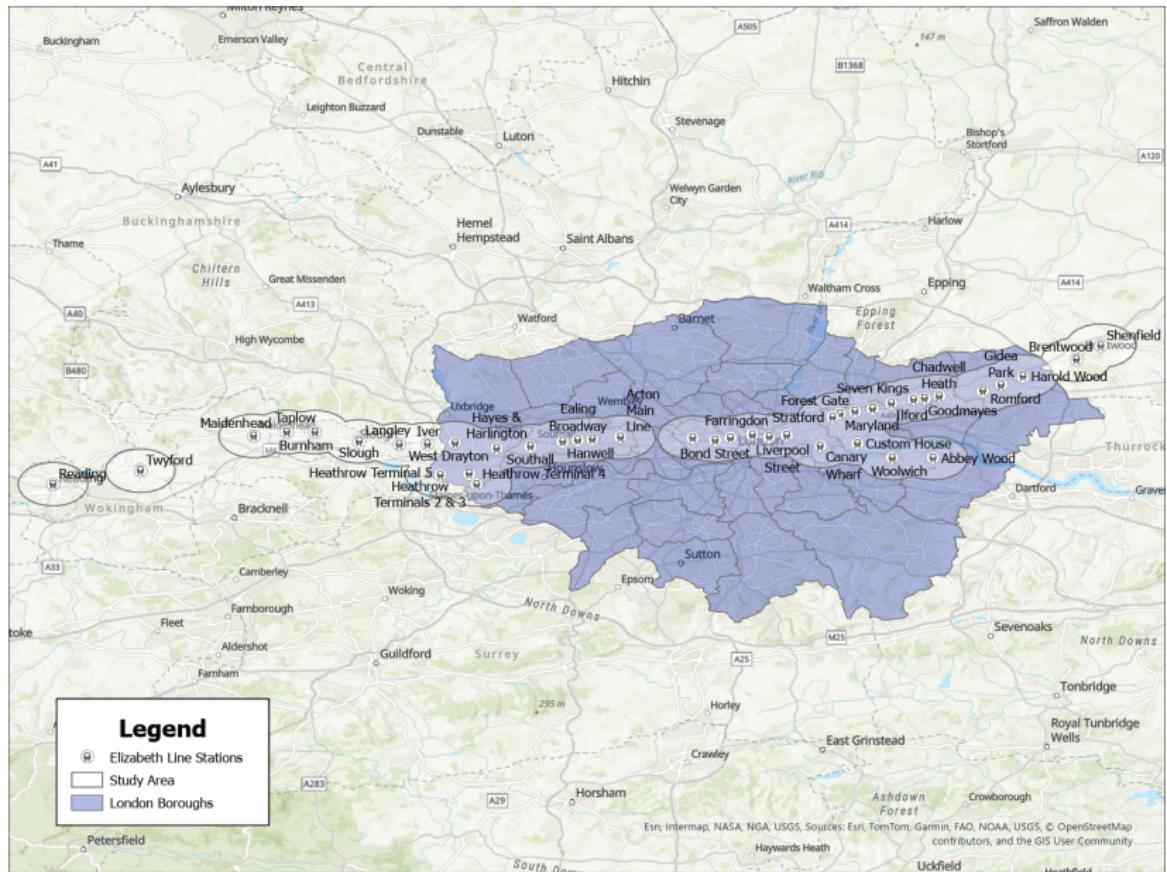
- ① Application of DID on the Elizabeth line
- ② Advocate the usage of Manhattan distance
- ③ Insight into cost-benefit analysis and land value capture

# Data

## Overview

- Study area: 3km buffer zone from all 41 Elizabeth line stations
- Time frame: 2006–2024
- Database: Zoopla
- Observations: 309,876

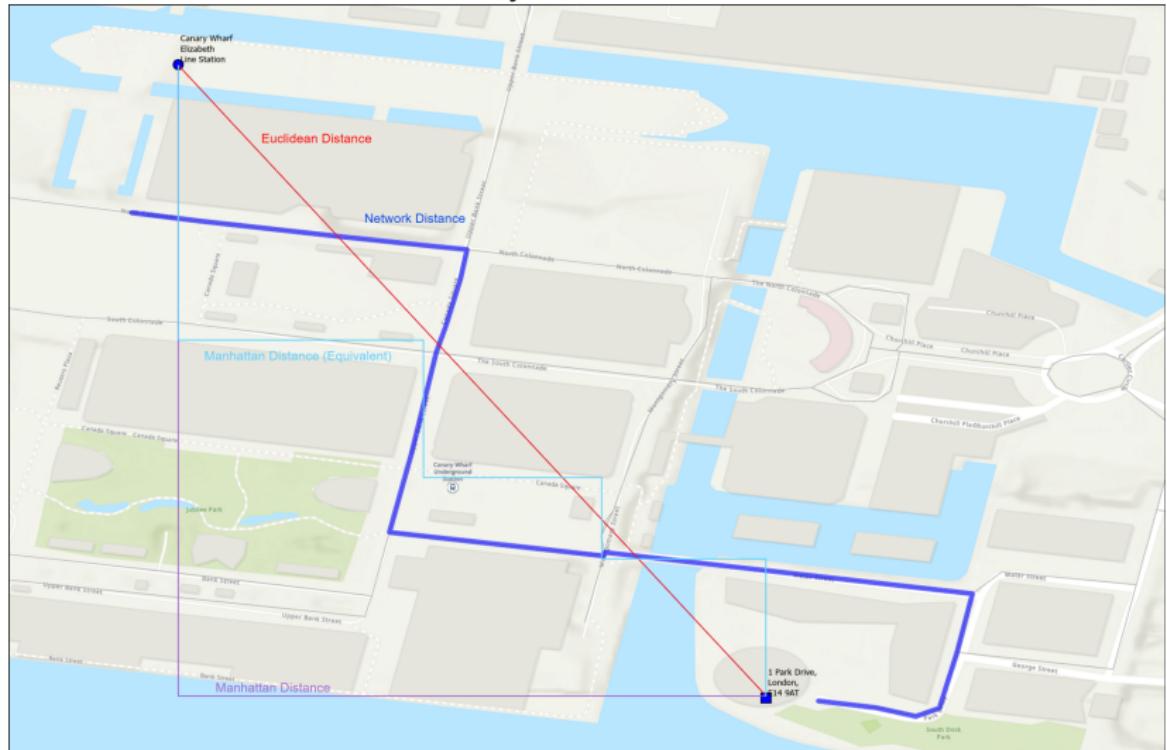
# Data



# Data

## Distance Measures

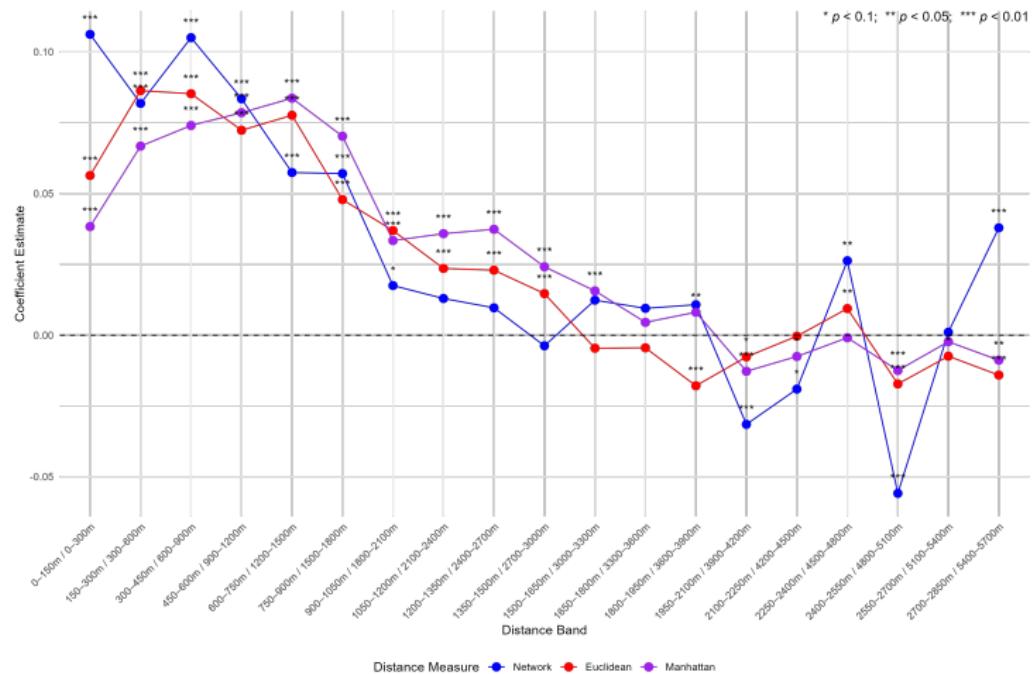
$$d_{ij}^{\text{Euclidean}} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} = 526, d_{ij}^{\text{Manhattan}} = |x_i - x_j| + |y_i - y_j| = 719,$$
$$d_{ij}^{\text{Network}} = 768$$



# Data

## Identifying the Treatment Zone

Used 150m/300m distance bands with >2850m/5700m as reference



Network: 2100m; Euclidean: 1500m; Manhattan: 1650m

# Empirical Estimation

## Baseline Model

### Base DID Model:

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

- Tested using network, Euclidean and Manhattan distance
- Hypothesis:  $\beta_1$ ,  $\beta_2$  and  $\beta_3$  are positive and significant

# Empirical Estimation

## Extension Models

### Treatment Effects over Space:

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

### Treatment Effects across Regions:

- Equation (1) further tested across 5 regions:
  - 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)
  - East (Brentwood and Shenfield)

# Results

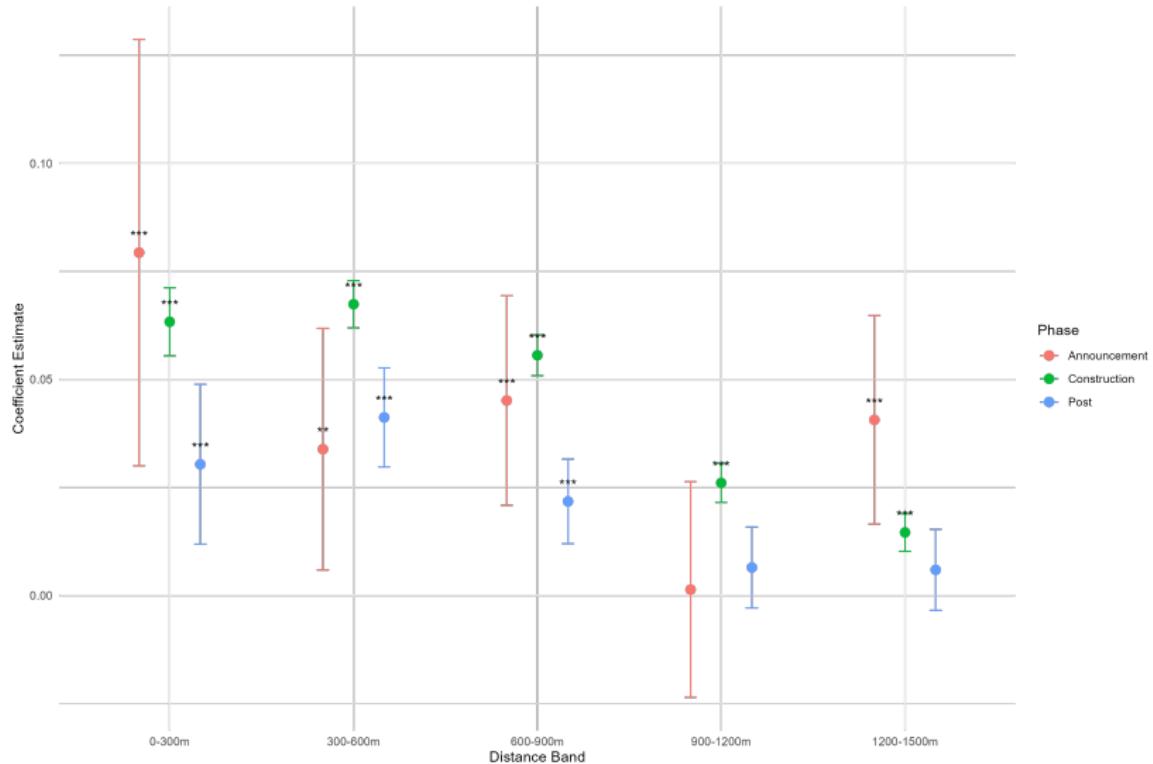
## Base DID Results

	(1)	(2)	(3)
Treat × Announcement	0.0313*** (0.0074)	0.0280*** (0.0075)	0.0180** (0.0078)
Treat × Construction	0.0372*** (0.0017)	0.0356*** (0.0017)	0.0358*** (0.0016)
Treat × Post	0.0177*** (0.0031)	0.0128*** (0.0031)	0.0142*** (0.0032)
Property Characteristics	Yes	Yes	Yes
Neighbourhood Attributes	Yes	Yes	Yes
Spatial Fixed Effects	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes
Distance Measure	Network	Euclidean	Manhattan
Observations	309,876	309,876	309,876
R <sup>2</sup>	0.7997	0.7997	0.7997
Adjusted R <sup>2</sup>	0.7995	0.7995	0.7995
Residual Std. Error (df = 309615)	0.2974	0.2975	0.2975
F Statistic (df = 260; 309615)	4,754.2500***	4,753.5150***	4,753.7940***

\* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$

# Results

## Treatment Effects over Space



# Results

## Treatment Effects across Regions

	(1)	(2)	(3)	(4)	(5)
Treat x Announcement	0.0412*** (0.0129)	-0.0261 (0.0174)	0.0249 (0.0197)	0.0327*** (0.0108)	0.1593*** (0.0396)
Treat x Construction	0.0307*** (0.0027)	0.0431*** (0.0039)	0.0749*** (0.0057)	0.0358*** (0.0023)	0.0872*** (0.0081)
Treat x Post	-0.0115** (0.0053)	0.0397*** (0.0067)	0.0593*** (0.0100)	0.0094** (0.0043)	0.0405*** (0.0155)
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 <sup>2</sup>	0.8281	0.8017	0.7672	0.7241	0.8210
Adjusted R <sup>2</sup>	0.8277	0.8012	0.7666	0.7238	0.8192
Residual Std. Error	0.2106 (df = 46851)	0.2687 (df = 54606)	0.3880 (df = 58653)	0.2668 (df = 139987)	0.2430 (df = 9136)
F Statistic	2,128.8630*** (df = 106; 46851)	1,824.2450*** (df = 121; 54606)	1,124.0380*** (df = 172; 58653)	2,551.0760*** (df = 144; 139987)	441.1201*** (df = 95; 9136)

\*p < 0.1; \*\*p < 0.05; \*\*\*p < 0.01

# Robustness Checks

## Parallel Trends

- Test for parallel trends between groups pre intervention:

$$\ln(P) = \alpha + \beta \text{Treat} + \sum_{i=1} \delta_i (\text{Treat} \times \tau_i) + H' \gamma + N' \theta + \varphi + \tau + \varepsilon \quad (3)$$

- All  $\delta_i$  must be statistically insignificant for assumption to be satisfied

# Robustness Checks

## Parallel Trends

Quarter	Network	Euclidean	Manhattan
2006 Q2	0.0053 (0.0099)	-0.0141 (0.0101)	0.0018 (0.0099)
2006 Q3	-0.0038 (0.0096)	0.0005 (0.0098)	0.0009 (0.0096)
2006 Q4	0.0185* (0.0096)	0.0106 (0.0098)	0.0120 (0.0096)
2007 Q1	0.0021 (0.0099)	-0.0054 (0.0101)	-0.0010 (0.0099)
2007 Q2	0.0026 (0.0097)	0.0054 (0.0099)	0.0052 (0.0097)
2007 Q3	-0.0007 (0.0096)	-0.0087 (0.0098)	-0.0037 (0.0096)
2007 Q4	-0.0050 (0.0101)	-0.0124 (0.0102)	-0.0124 (0.0101)
2008 Q1	-0.0050 (0.0110)	-0.0012 (0.0112)	0.0007 (0.0110)
2008 Q2	-0.0109 (0.0111)	-0.0143 (0.0113)	-0.0092 (0.0111)
2008 Q3	-0.0187 (0.0228)	-0.0013 (0.0231)	-0.0071 (0.0228)
Observations	50,995	50,995	50,995
R <sup>2</sup>	0.7887	0.7887	0.7885
Adjusted R <sup>2</sup>	0.7879	0.7879	0.7877
Residual Std. Error (df = 50800)	0.2438	0.2438	0.2439
F Statistic (df = 194; 50800)	977.3534***	977.2628***	976.3418***

\*p < 0.1; \*\*p < 0.05; \*\*\*p < 0.01

# Robustness Checks

## Alternative Treatment Boundaries

	(1)	(2)	(3)
Treat × Announcement	0.0380*** (0.0098)	0.0323*** (0.0086)	-0.0098 (0.0074)
Treat × Construction	0.0499*** (0.0019)	0.0467*** (0.0017)	0.0037** (0.0017)
Treat × Post	0.0279*** (0.0040)	0.0163*** (0.0035)	-0.0111*** (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 <sup>2</sup>	0.7998	0.7999	0.7994
Adjusted R <sup>2</sup>	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***

\* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$

# Discussion

## Policy Implications: Cost-Benefit Analysis

- Average windfall gain ( $\beta_2^{\text{Network}}$ ): 3.72%
- Average treated housing price: £481,357
- Average LVU = £481,357 × 3.72% = £17,906
- Number of treated properties: 148,082
- Total LVU = £17,906 × 148,082 = £2.65 billion
- Total cost of Elizabeth line: £18.9 billion  
⇒ 14% of total costs

# Discussion

## Policy Implications: Wealth Inequality

- Creates wealth gap between homeowners and non-homeowners
- Inter-regional wealth inequality
- However, must be contextualised with broader benefits

# Discussion

## Policy Implications: Land Value Capture

- Windfall gains suggests that redistribution would be ideal
- Business Rates Supplement already targets commercial properties
- Current tax system too unresponsive in capturing residential gains
- TfL have considered several potential capture methods using levies

# Discussion

## Limitations

- Data:
  - Additional housing controls
  - Unusable transactions
  - Network distance calculation (e.g. Google Distance Matrix API)

# Conclusion

- ① The Elizabeth line did cause windfall gains in housing values
- ② However, the magnitude varies greatly across space, time and regions
- ③ Manhattan distance is a reasonable approximation of network distance
- ④ Land value capture mechanisms will be stifled by political opposition

# Conclusion

- Highlights: Using R and ArcGIS Pro
- Lowlights: Web scraping Zoopla
- Extensions: Improve neighbourhood controls and implement SDID

# Conclusion

Thank you for listening. Any questions?

# Appendix

## A Brief Timeline

22<sup>nd</sup> July 2008: Formally approved by Crossrail Act 2008

15<sup>th</sup> May 2009: Start of construction

24<sup>th</sup> May 2022: Operations began

- Key dates:

- Pre Announcement (Baseline): 01-01-2006 to 21-07-2008
- Announcement: 22-07-2008 to 14-05-2009
- Construction: 15-05-2009 to 23-05-2022
- Post: 24-05-2022 onwards

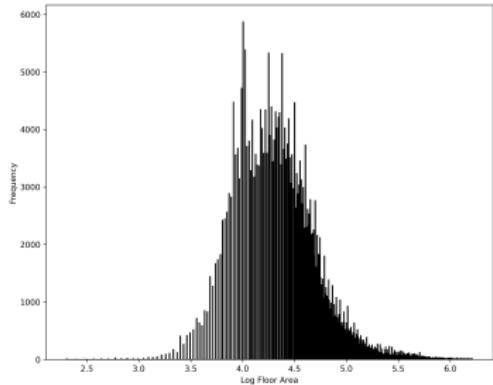
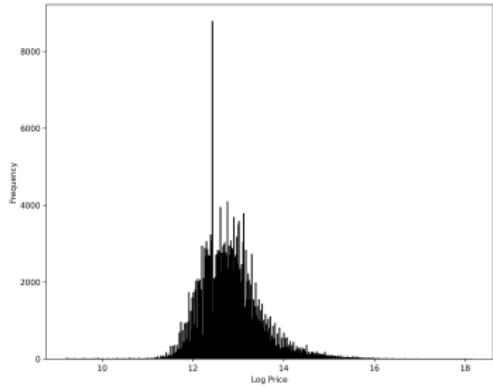
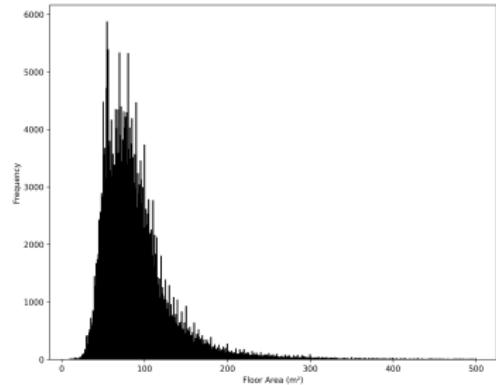
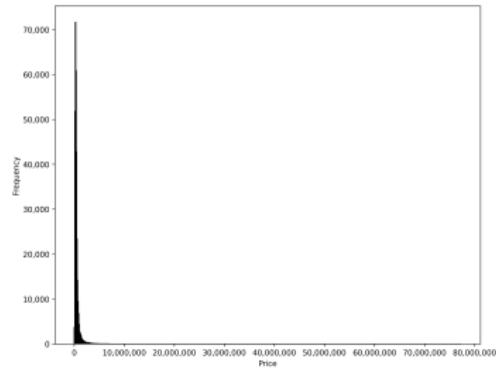
# Appendix

## Descriptive Statistics

Statistic	Full Dataset			Treatment Group			Control Group		
	N	Mean	St. Dev.	N	Mean	St. Dev.	N	Mean	St. Dev.
Price	309,876	508,688.00	739,005.70	148,082	481,356.80	654,097.20	127,291	542,816.70	802,040.40
Floor Area	309,876	90.92	45.51	148,082	89.46	44.15	127,291	92.75	47.01
Bedrooms	309,876	2.36	1.05	148,082	2.32	1.05	127,291	2.41	1.05
Bathrooms	309,876	1.41	0.66	148,082	1.40	0.64	127,291	1.42	0.67
Living Rooms	309,876	1.26	0.54	148,082	1.26	0.53	127,291	1.27	0.56
Population Density	309,876	12,103.23	10,486.51	148,082	12,250.68	10,929.49	127,291	11,753.98	9,154.60
% over 65	309,876	0.10	0.06	148,082	0.10	0.06	127,291	0.11	0.06
IMD	309,876	15,006.35	7,489.59	148,082	15,307.52	7,310.38	127,291	14,752.70	7,726.93
Network Distance	309,876	2,232.12	1,143.95	148,082	1,279.53	503.54	127,291	3,327.60	794.74
Euclidean Distance	309,876	1,588.09	772.03	148,082	942.47	415.56	127,291	2,313.41	457.39
Manhattan Distance	309,876	1,977.45	963.87	148,082	1,190.12	531.27	127,291	2,860.35	613.58

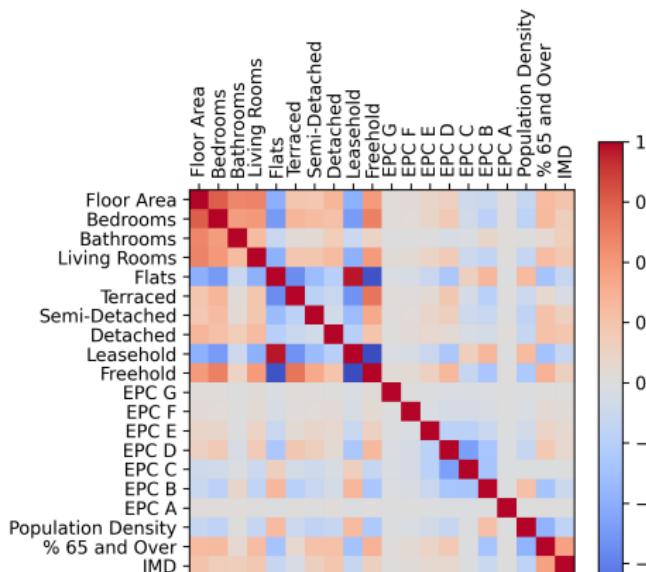
# Appendix

## Semi-Log Model



# Appendix

## Multicollinearity



Variable	VIF
Treat x Announcement	1.7129
Treat x Construction	2.1374
Treat x Post	1.9383
In(Floor Area)	2.7219
Bathrooms	1.6948
Living Rooms	1.6471
Property Type	1.2230
EPC	1.0453
Population Density	1.7992
% over 65	2.8297
IMD	2.1397
Outcode	1.0091
Quarter	1.0107

Omitted Bedrooms and Tenure given high correlation

# Appendix

## Hedonic Prices

Dependent Variable: $\ln(P)$	
In(Floor Area)	0.6106*** (0.0021)
Bathrooms	0.0996*** (0.0011)
Living Rooms	0.0358*** (0.0013)
Terraced	0.1414*** (0.0017)
Semi-Detached	0.1790*** (0.0023)
Detached	0.2490*** (0.0032)
EPC F	0.0187 (0.0154)
EPC E	0.0116 (0.0140)
EPC D	-0.0042 (0.0139)
EPC C	-0.0063 (0.0139)
EPC B	0.1372*** (0.0140)
EPC A	0.00003 (0.0241)
Population Density	-0.000001*** (0.000000)
% over 65	-0.2080*** (0.0152)
IMD	0.00001*** (0.000000)
Constant	8.9066*** (0.0178)
Observations	309,876
R <sup>2</sup>	0.7997
Adjusted R <sup>2</sup>	0.7995
Residual Std. Error	0.2974 (df = 309615)
F Statistic	4,754.2500*** (df = 260; 309615)

\* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$

# Appendix

## Quarterly Housing Price Index

