

Did the end-to-end opening of the Elizabeth line accelerate business churn near stations?

Tee Chin Min Benjamin

August 2026

This dissertation is submitted in part requirement for the MSc in the Centre for Advanced Spatial Analysis, Bartlett Faculty of the Build Environment, UCL.

Submitted: 20 August 2026

Supervisor: Dr Jens Kandt

Module ID: CASA0010 MSc Urban Spatial Science Dissertation

Student ID: 25049107

Word Count: 10,485

Abstract

Major transit investments such as the Elizabeth line in London are intended to deliver substantial economic and accessibility benefits. Official evaluations document positive effects on employment, travel times, and property values, but less is known about the line's impact on business churn in surrounding station catchments. Furthermore, while transit-induced commercial change is well-studied in North American light-rail contexts, the UK commuter-rail landscape remains relatively less studied. Using a Difference-in-Differences (DID) design with Coarsened Exact Matching (CEM), this dissertation investigates the extent to which the end-to-end opening of the Elizabeth line in May 2022 accelerated retail occupant churn within selected station catchments, relative to comparable station catchments. This study draws on a novel dataset, the OpenLocal ratepayers data assembled from the Valuation Office Agency (VOA) and local authority publications. In so doing, this dissertation also explores the extent to which this data source provides suitable panel data for evaluating major transport schemes. The analysis finds no statistically significant effect on retail occupant churn. This null result is robust across matching specifications, treatment definitions, and spatially clustered standard errors. While OpenLocal is useful for hereditament-level causal inference, it requires substantial preprocessing and may be better suited to targeted studies rather than city-wide analyses. The study's null findings suggest that the Elizabeth line generated offsetting commercial pressures in the short run, accelerating displacement for some incumbents while stabilising others. This short-term finding reassures policymakers that major transit investments may not inherently trigger immediate commercial displacement but highlights the need for longer-term monitoring and better tracking of retail composition outcomes.

Declaration of Authorship

I, Tee Chin Min Benjamin, hereby declare that this dissertation is all my own original work and that all sources have been acknowledged. Excluding annexes, this dissertation is approximately 10,485 words in length.

Signed: *Benjamin Tee*

Date: 20 August 2026

Table of Contents

Abstract.....	2
Declaration of Authorship	3
Table of Contents	4
Figures	6
Tables.....	7
Acronyms and Abbreviations.....	8
Acknowledgements.....	9
1. Introduction	10
2. Literature Review.....	13
2.1. How do transit investments impact surrounding businesses?	13
2.2. How can causal studies be used to estimate this impact?.....	16
3. Background on the Elizabeth line	19
4. Data	22
4.1. Data quality and coverage.....	22
4.2. Data cleaning and processing	22
4.3. Handling administrative artefacts.....	27
4.4. Additional data	27
4.5. Descriptive statistics.....	28
5. Empirical Strategy	31
5.1. Research design and study period.....	31
5.2. Defining treatment and control catchments.....	31
5.2.1. Identifying treatment stations.....	31
5.2.2. Identifying control stations.....	34
5.3. Matching controls on covariate characteristics	36
5.4. Difference-in-Differences (DID) analysis.....	41
6. Results	45
6.1. Coarsened exact matching (CEM) results.....	45
6.2. Simple Difference-in-Differences (DID) results.....	48
6.3. Difference-in-Differences (DID) regression results.....	49
6.4. Event study results.....	50

6.5.	Independent-business displacement	52
7.	Discussion.....	53
7.1.	Summary of regression results	53
7.2.	Interpreting regression findings	53
7.3.	Assessment of VOA OpenLocal ratepayer data	57
7.4.	Limitations	57
7.4.1.	Data quality and measurement.....	57
7.4.2.	Sample coverage	58
7.4.3.	Identification and confounding	58
7.4.4.	Economic inference and statistical power	59
7.5.	Future research	61
7.6.	Implications for policy and planning	62
8.	Conclusion	63
	References	64
	Annexes.....	72
	Annex A: Studies on impact of transit investments on commercial properties	73
	Annex B: Collinearity Diagnostics for CEM covariate candidates	75
	Annex C: Full covariate matching outcomes of Coarsened Exact Matching (CEM)	76
	Annex D: Mahalanobis nearest neighbour matching results	79
	Annex E: DID Regression with Conley Standard Errors	82
	Annex F: Event Study Coefficients.....	83
	Annex G: Leave-one-out analysis of event study specifications.....	85
	Annex H: Independent-business displacement analysis.....	86
	Research Log	89

Figures

Figure 1: Mechanisms affecting retail occupant churn in transit catchments.....	13
Figure 2: Geographical location of Elizabeth line stations in Greater London	21
Figure 3: Workflow for data cleaning and pre-processing of OpenLocal data	23
Figure 4: Frozen and suppressed turnover rate outcomes across six Local Authorities	24
Figure 5: Illustration of cleaning approach used to isolate genuine turnover events	25
Figure 6: Percentage of null in key fields, before vs after data quality exclusions	28
Figure 7: Retail turnover rate across Greater London (2Q2018 – 4Q2024)	29
Figure 8: Average quarterly retail turnover rate by LSOA (Pre/Post EZL opening)	30
Figure 9: 15-min walking isochrones and catchment areas of final treatment stations	33
Figure 10: Eligible treated and control catchments after exclusions	35
Figure 11: Distribution of pre-treatment covariates in treated and control catchments....	39
Figure 12: Covariate balance before and after CEM (Love plot)	47
Figure 13: Quarterly turnover rate for treated and matched controls stations.....	48
Figure 14: Binned Event-Study Plot (q-10 to q+10).....	51
Figure 15: Pre- vs post-treatment growth rates: turnover events and eligible UARNs.....	56
Figure 16: Pre-treatment catchment growth trajectories, extrapolated	56

Tables

Table 1: Top 5 Local Authorities (LA) with ratepayer name masking	24
Table 2: Evaluation and selection of primary dependent variable	26
Table 3: Composition of final treatment stations across two specifications	32
Table 4: Breakdown of exclusion criteria used to determine control stations.....	35
Table 5: Rationale for selected CEM covariates	38
Table 6: Finalized bin boundaries for covariate set and bin boundary rationale	40
Table 7: Summary of CEM outcomes and match rates	45
Table 8: Treatment and matched stations (MS1)	46
Table 9: Treatment and matched stations (MS2)	46
Table 10: Matched panel turnover rate, pre vs post treatment	48
Table 11: DID regression results of Elizabeth line opening on retail churn	50
Table 12: Joint pre-trend Wald tests for Event Study Window	52

Acronyms and Abbreviations

API	Application Programming Interface
ATT	Average Treatment effect on the Treated
CEM	Coarsened Exact Matching
CI	Confidence Interval
CBE	Crossrail Baseline Evaluation
DID	Difference-in-Differences
DLR	Docklands Light Railway
EZL	Elizabeth line
FE	Fixed Effects
GIS	Geographic Information System
HAC	Heteroskedasticity- and Autocorrelation-Consistent (standard errors)
IMD	Index of Multiple Deprivation
IPW	Inverse Probability Weighting
L1	L1 imbalance statistic (multivariate imbalance measure)
LA	Local Authority
LSOA	Lower-layer Super Output Area
MS1/MS2	Matching Specification 1 / Matching Specification 2
NaPTAN	National Public Transport Access Nodes
OSM	OpenStreetMap
PSM	Propensity Score Matching
SUTVA	Stable Unit Treatment Value Assumption
SE	Standard Error
SMD	Standardised Mean Difference
TWFE	Two-Way Fixed Effects
UARN	Unique Address Reference Number
UK	United Kingdom
VOA	Valuation Office Agency
VIF	Variance Inflation Factor
WEI	Wider Economic Impacts

Acknowledgements

First and foremost, I give thanks to God for granting me the strength and knowledge to complete this dissertation.

I am deeply grateful to my supervisor Dr Jens Kandt for his generous guidance, feedback, and encouragement throughout the course of this dissertation. I also appreciate the helpful suggestions from my peer group comprising Owen Hughes and Christian Mulder.

I extend this gratitude to staff at UCL's Centre for Advanced Spatial Analysis (CASA) for their assistance and dedication to teaching throughout the academic year.

I would like to acknowledge the Ministry of National Development (MND) Singapore, for granting me this opportunity to deepen my skills and expertise with CASA.

Finally, I would like to thank my wife and family for their continuous encouragement and support that has sustained me during this journey.

1. Introduction

In May 2022, after more than a decade of construction and around £19 billion in investment, the central section of the Elizabeth line opened through the heart of London, connecting Reading and Heathrow in the west to Abbey Wood and Shenfield in the east. Britain's largest transport project to date was projected to increase central London's rail capacity by 10 per cent and add over 1.5 million people to within 45 minutes of central London. By adding ten new stations and providing high-frequency services across 41 stations, the project aimed to connect London's main employment centres and support new journeys through central London out toward Essex, Buckinghamshire and Berkshire ([Transport for London, 2022b](#)). Yet, while the anticipated effects of the Elizabeth line on connectivity, employment, and property values have been extensively studied ([Arup and Department for Transport, 2025a, 2025b, 2025c](#)), comparatively less is known about how the line affects the commercial life of neighbourhoods served by its stations – specifically whether the shops, restaurants, and services around its stations are being displaced and replaced, and if so, how fast.

Transport infrastructure represents one of the largest investments a government can make, and the massive fiscal commitments motivate a desire to measure their returns. Over the past six decades, the framework for evaluating the returns on transport investment in the UK has evolved considerably, from early cost–benefit methodologies ([Foster and Beesley, 1963](#)), to the recognition that conventional appraisal did not fully capture transport's wider economic effects ([SACTRA, 1999; Banister and Berechman, 2001](#)), to modern frameworks that explore the transformational impacts ([Eddington, 2006](#)) and Wider Economic Impacts (WEIs) of transport investment ([Graham and Gibbons, 2019](#)). Official evaluations of the Elizabeth line commissioned by Transport for London and the Department for Transport thus far have documented its impact on jobs, journey times, property values and regeneration and development around stations ([Arup and Department for Transport, 2025a, 2025b, 2025c](#)). Empirical case studies of stations like Farringdon, Abbey Wood, and Slough highlight early transformational shifts in employment, productivity, housing delivery, and land use changes ([Arup and Department for Transport, 2025a, 2025b](#)). These studies mainly focus on documenting aggregate outcomes. However, less is known about business survival, displacement, and the speed of these shifts.

While new rail infrastructure and transit-oriented development promise enhanced connectivity and sustainable growth, they can bring about complex socioeconomic shifts (Padeiro, Louro and da Costa, 2019; Delmelle, 2021). Most studies on this topic have examined this from the perspective of residential gentrification and the displacement of lower-income households (Kahn, 2007; Zuk et al., 2018). However, the parallel impact on commercial businesses is relatively understudied (Bardaka, 2024). The introduction of new transit lines can add foot traffic and enhance the accessibility of a location, potentially benefiting existing, nearby businesses and attracting new ones (Schuetz, 2015; Credit, 2018; Sukaryavichute, Delmelle and Hammelman, 2021). On the cost side, however, small, independent businesses risk being disproportionately affected by rising commercial property values and rents or the shifting preferences of a new clientele (Meltzer, 2016). Effects can take time to materialise due to long commercial lease structures and multi-year construction disruption (Meltzer, 2016; Chapple and Loukaitou-Sideris, 2019; Wexler and Fan, 2022). Overall, the net impact of transit investment on retail businesses is uncertain. While some studies identify systematic displacement (Lin and Yang, 2019), others document mixed or null effects (Chapple et al., 2017; Liu and Bardaka, 2023) or protective effects where transit investments support the longevity and survival of certain service and retail sectors (Liu and Bardaka, 2025).

Studies examining the commercial dimensions of transit-induced change have employed a wide range of methodologies, from descriptive and spatial approaches (Ko and Cao, 2013; Seo et al., 2019) to qualitative interviews (Chapple and Loukaitou-Sideris, 2019; Tornabene and Nilsson, 2021) to causal inference studies (Pogonyi, Graham and Carbo, 2021; Liu and Bardaka, 2023, 2025). A common hurdle experienced by researchers attempting causal studies is the difficulty in accessing historical business microdata at sufficient granularity and frequency to accurately track individual firm births, deaths, and relocation (Bardaka, 2024). Consequently, causal studies on this topic largely focus on North American light-rail systems where such longitudinal datasets have become more available (Bardaka, Delgado and Florax, 2018; Credit, 2018; Wexler and Fan, 2022; Liu and Bardaka, 2023, 2025). Meanwhile, the UK context remains relatively understudied.

Two recent developments make it possible to revisit these challenges. First, the end-to-end opening of London’s Elizabeth line in May 2022 generated a discrete, spatially targeted accessibility improvement across station catchments, offering a well-defined quasi-experimental setting for causal evaluation. Second, computational advancements in automated data aggregation and algorithmic cleaning have enabled the formulation of the OpenLocal dataset (<https://openlocal.uk>) from raw Valuation Office Agency (VOA) non-domestic ratepayer and valuation lists (Whythawk, 2026). By tracking individual commercial retail units (hereditaments) in the UK over time, this longitudinal microdata dataset provides high-resolution panel visibility into establishment entry, exit, and structural vacancy.

Building on these developments, this study addresses two central questions:

1. Did the end-to-end opening of the Elizabeth line in May 2022 accelerate retail occupant churn within target station catchments relative to comparable, unserved control catchments?
2. What are the strengths and limitations of administrative ratepayer data as a panel source for evaluating this retail churn, and under what conditions can it support credible causal inference?

In answering these questions, this dissertation makes two key contributions. First, it provides a causal evaluation of the retail impact of a major commuter-rail investment in the UK, expanding upon the current literature that is largely dominated by North American light-rail and residential gentrification studies. Second, it evaluates the utility and constraints of a novel, establishment-level administrative dataset for urban policy research.

The remainder of this dissertation is structured as follows. [Section 2](#) synthesizes the theoretical mechanisms connecting transit investment to local business dynamics and reviews previous causal studies. [Section 3](#) provides context on the Elizabeth line and defines the station catchments analysed. [Section 4](#) introduces the OpenLocal dataset, detailing the data processing and cleaning pipeline constructed to convert administrative ratepayer records into a research-ready panel. [Section 5](#) presents the empirical strategy while [Section 6](#) reports the results. [Section 7](#) discusses key findings, limitations and implications for transport and land-use policy, and [Section 8](#) concludes.

2. Literature Review

2.1. How do transit investments impact surrounding businesses?

Major transit investments like the Elizabeth line can bring about complex socioeconomic shifts (Padeiro, Louro and da Costa, 2019; Delmelle, 2021). Drawing across the transit, gentrification, and small-business literature, this section outlines four mechanisms through which transit investment can affect retail occupant turnover or "churn". These mechanisms can exert opposing structural forces. While some accelerate churn by displacing incumbent businesses, others mitigate or delay churn by bolstering existing firm viability or buffering them against operational disruption (Liu and Bardaka, 2023). As a result of these competing pressures, empirical studies demonstrate complex and heterogeneous impacts, with net outcomes dependent on which mechanism dominates in each local context. [Figure 1](#) summarises the different mechanisms.

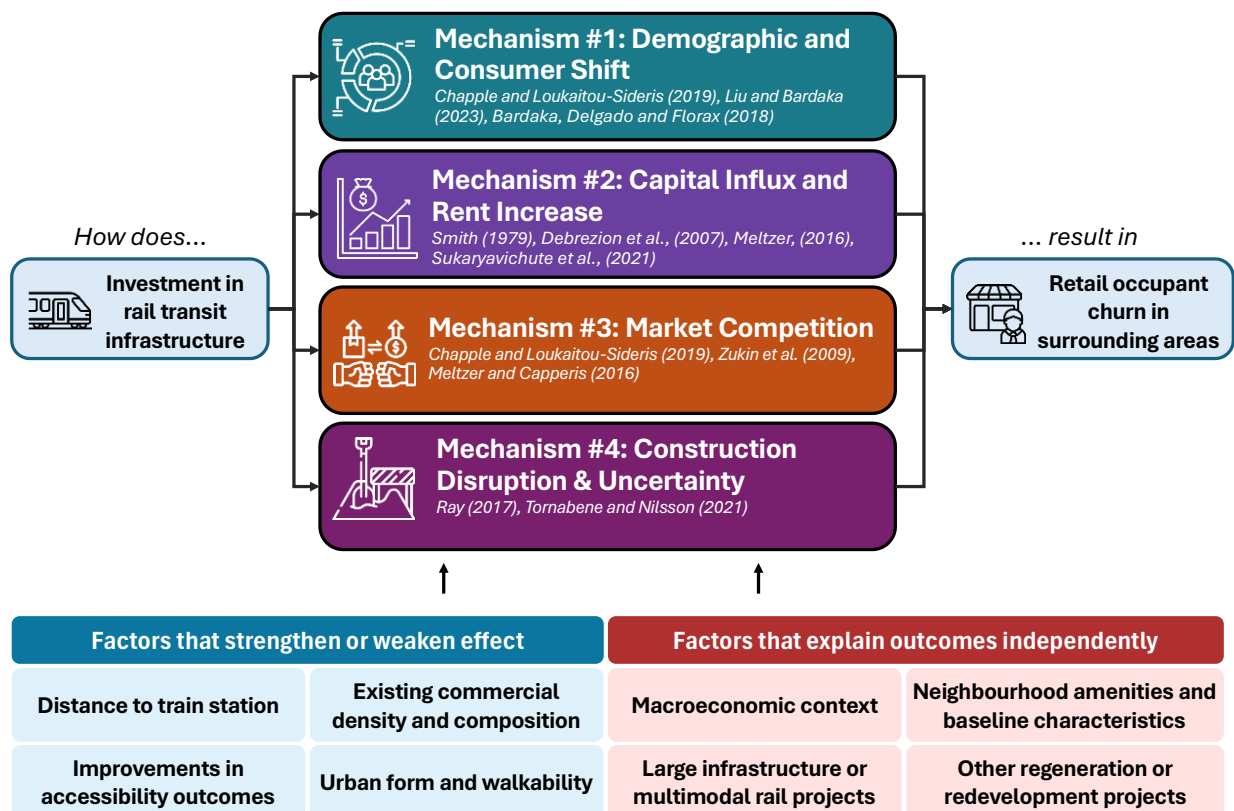


Figure 1: Mechanisms affecting retail occupant churn in transit catchments

Mechanism 1: Demographic and consumer shifts

The first mechanism through which transit investments can affect commercial churn is through demographic and consumer shifts, brought about by improvements in accessibility and reduction in transport costs (Schuetz, 2015). Expanded connectivity to key employment and leisure nodes broadens the potential labour market and retail catchment for nearby businesses (Knowles and Ferbrache, 2016; Lee and van Dijk, 2026). In tandem, transit investments often serve as a catalyst for physical improvements to the urban environment around a station, as developers capitalise on accessibility-driven value uplifts, and planners improve streetscapes to enable pedestrianization and improved access to stations (Cervero and Duncan, 2002; Knowles and Ferbrache, 2016). These spatial transformations often draw in a higher-income residential population, shifting local demand structures through a pathway commonly described as transit-induced commercial gentrification (Bardaka, Delgado and Florax, 2018; Liu and Bardaka, 2023).

This demographic transition exerts dual, opposing pressures on incumbent businesses. For some, the wider catchment and influx of higher-spending residents provides a revenue boost that strengthens business viability and reduces closures (Liu and Bardaka, 2025). For others, it may accelerate market departure, as firms unable to adapt to "retail upscaling" are displaced (Chapple and Loukaitou-Sideris, 2019).

Mechanism 2: Capital influx and rent increase

A second mechanism through which transit investments affect commercial change is private capital influx and rent appreciation, driven by speculative efforts to capture "land value uplift" (Debrezion, Pels and Rietveld, 2007; Medda, 2012). This dynamic is grounded in rent-gap theory, which posits that urban redevelopment is triggered when the disparity between a site's actual capitalized ground rent under current usage and its potential ground rent under a higher-and-better use becomes wide enough to incentivise reinvestment (Smith, 1979). While originally formulated around residential neighbourhood dynamics, scholars have extended this reasoning to commercial environments, where rent gaps arising from increased demand can encourage property owners to reposition retail space toward higher-yielding occupiers (Meltzer, 2016). As a location becomes more desirable, private capital flows in, causing property values and commercial rents to rise. For incumbent

businesses, this constitutes a negative effect especially if they cannot achieve a corresponding increase in revenue to offset higher rents (Meltzer, 2016; Ray, 2017). Sukaryavichute et al., (2021) found that increased rent was one of the top reasons for businesses to relocate following the construction of light rail stations in the US. Notwithstanding, Meltzer (2016) explains that displacement effects are often lagged given that commercial leases are typically longer than residential ones, lasting between three to five years, shielding existing tenants from rent hikes until their leases expire or are renewed.

Mechanism 3: Market competition

A third strand operates through heightened market competition. As elevated footfall and expanding consumer demand around improved transit nodes draw new commercial entrants, incoming businesses intensify competitive pressure on incumbents (Liu and Bardaka, 2023). High-accessibility locations are particularly attractive to multi-location chains and corporate franchises, which are typically better-capitalised than independent operators and can bid up commercial rents beyond levels that local, independent operators can sustain (Zukin et al., 2009; Meltzer and Capperis, 2016; Chapple et al., 2017).

Mechanism 4: Construction disruption and uncertainty

Prior to realizing the benefits of transit investments, local businesses must also navigate the challenges of multi-year construction phases. Reduced physical access, diminished parking capacity, and environmental disruptions directly threaten firm survival (Ray, 2017). Quantifying these effects, Ray (2017) found that station construction in Los Angeles elevated business closure risk by 46% for establishments within 400 meters. Prolonged construction timelines also introduce substantial uncertainty, leaving small businesses vulnerable to permanent customer attrition (Tornabene and Nilsson, 2021). Consequently, establishments that endure the initial construction phase may succumb to cumulative financial strain or fail to adapt to the post-completion urban landscape, resulting in delayed displacement (Tornabene and Nilsson, 2021).

Moderators of mechanism strength

The intensity with which these mechanisms operate is geographically non-uniform and is conditioned by the following catchment characteristics that either attenuate or amplify the transit effect.

First, accessibility gains, pedestrian footfall, and redevelopment pressures follow a spatial distance-decay pattern. These impacts typically concentrate within a 250- to 750-meter walking radius of a station, though measurable effects have been found to extend up to 1,500 meters ([Lin and Yang, 2019](#); [Wexler and Fan, 2022](#); [Deng and Li, 2024](#)).

Second, urban form and local walkability moderate the spatial depth and distribution of these shifts. Permeable, well-connected street networks facilitate the diffusion of footfall and commercial gentrification deeper into surrounding neighbourhoods ([Chapple and Loukaitou-Sideris, 2019](#)). Conversely, low-connectivity environments or physical barriers, such as grade separations and fenced rail corridors, confine commercial restructuring to the immediate station forecourt while insularly buffering adjacent blocks ([Tornabene and Nilsson, 2021](#); [Liu and Bardaka, 2023](#)).

Finally, baseline commercial density and tenant mix influence the overall volume of expected turnover. Catchments already saturated with high-value or corporate retail demonstrate minimal additional churn, as spatial and market constraints limit capacity for further repositioning or new entries ([Schuetz, 2015](#)). In contrast, areas characterised by mixed-use or lower-rent commercial stock have been found to experience significantly elevated churn rates, as low baseline rents present substantial rent-gap potential ([Meltzer and Capperis, 2016](#)).

2.2. How can causal studies be used to estimate this impact?

While the existing literature conceptualizes the competing structural forces driving commercial churn, empirically isolating these mechanisms presents a distinct identification challenge. Empirical literature that have examined the impacts of transit rail investments fall into two broad categories:

- First, researchers have used hedonic price models ([Ko and Cao, 2013](#); [Seo et al., 2019](#)), spatial regression techniques ([Kim and Byun, 2021](#)), and qualitative case studies ([Sukaryavichute, Delmelle and Hammelman, 2021](#); [Tornabene and Nilsson,](#)

2021) using associational and descriptive approaches. While these studies offer valuable descriptive insights into spatial clustering and qualitative business experiences, they face attribution challenges and cannot disentangle the specific effect of transit improvements from concurrent urban trends.

- Second, a more recent body of work adopts quasi-experimental identification strategies grounded in [Rubin's \(1974\)](#) potential outcomes framework to estimate causal impacts. Difference-in-differences (DID) designs have emerged as the primary standard ([Schuetz, 2015](#); [Ray, 2017](#); [Wexler and Fan, 2022](#); [Liu and Bardaka, 2023](#)). Within the UK context, empirical evaluations of major transport infrastructure, such as the analysis of the Jubilee Line Extension ([Pogonyi, Graham and Carbo, 2021](#)) and evaluation of London's Northern Line and Barking Riverside extensions ([Lei, 2026](#)), highlight the use of causal frameworks to distinguish local infrastructure impacts from city-wide trends. Further advancing this literature, ([Kandt, Van Dijk and Todd, 2025](#)) deployed a quasi-experimental design using linked consumer register microdata across 19 UK transit schemes, demonstrating how causal frameworks applied to granular consumer data can effectively capture the timing, spatial extent, and local context of transit-induced demographic churn. A list of studies featuring both approaches can be found in [Annex A](#).

[Graham \(2025\)](#) argues that causal inference methods offer several advantages which are essential for moving beyond correlation to credible estimates of an intervention's effect.

- First, they mitigate selection bias and confounding. Infrastructure investments are non-randomly assigned as transit planners often deliberately target high-density corridors or areas undergoing planned urban renewal. Associational and descriptive models often mistake pre-existing momentum for a "transit effect," whereas causal identification strategies attempt to block these indirect back-door pathways to isolate the specific impact of the infrastructure itself.
- Second, they aim to recover a net causal signal relevant to policy. In settings where macroeconomic cycles, gentrification pressures, and local interventions occur concurrently, causal methods attempt to filter out surrounding temporal and spatial

shocks to estimate the intervention's specific impact. This enables researchers to develop a more realistic estimate of a project's impact, minimising the extent to which returns are overstated, and allowing for more informed decision-making.

However, [Graham \(2025\)](#) explains that these designs are not automatic solutions. They rest on “stringent identifying assumptions”, such as parallel trends for difference-in-differences, that need to be tested rather than assumed.

Operationalising a quasi-experiment approach to measure business churn

To effectively measure how transit infrastructure affects business churn, this study adopts a quantitative quasi-experimental design. Because randomized controlled trials are unfeasible for major urban infrastructure, a quasi-experiment offers the most credible available basis for constructing a counterfactual. By exploiting spatial variation in accessibility gains, the study pairs treated station catchments with control areas not directly served by the Elizabeth line, matched on observable pre-treatment characteristics. This framework is designed to identify the causal effect of transit investment on business churn, providing an empirical basis for evaluating the commercial turnover effects of large-scale transit investment.

3. Background on the Elizabeth line

The Elizabeth line represents one of the most significant transport infrastructure projects undertaken in the United Kingdom to date. Spanning over 100 kilometres and serving 41 stations, including ten newly constructed underground stations in central London, the scheme was delivered at a total project cost of £19 billion ([Benson and Robinson, 2024](#)).

Historical intent and rationale

The scheme traces back to the 1974 London Rail Study, which first proposed a high-frequency cross-London rail tunnel under the "Crossrail" name ([Hebbert, 2014](#)). The project established a major precedent in transport economics as its 2005 appraisal was among the first major UK infrastructure evaluations to formally integrate Wider Economic Benefits (WEBs). Grounded in agglomeration theory, the framework quantified productivity gains from enhanced spatial density and labour-market access in central London, elevating the benefit–cost ratio from 1.8 to 3.2 ([Department for Transport, 2005](#)).

Phased service openings

Eastern surface operations between Liverpool Street and Shenfield began under the "TfL Rail" brand in May 2015. Western surface services followed to Hayes & Harlington and Heathrow in May 2018, extending to Reading and Maidenhead in December 2019 and Heathrow Terminal 5 in July 2020. The critical milestone occurred on 24 May 2022 with the opening of the central tunnelled section, connecting Paddington and Abbey Wood. This enabled full end-to-end through-running and officially unified the network under the Elizabeth line brand in honour of Queen Elizabeth II ([Transport for London, 2022c](#)).

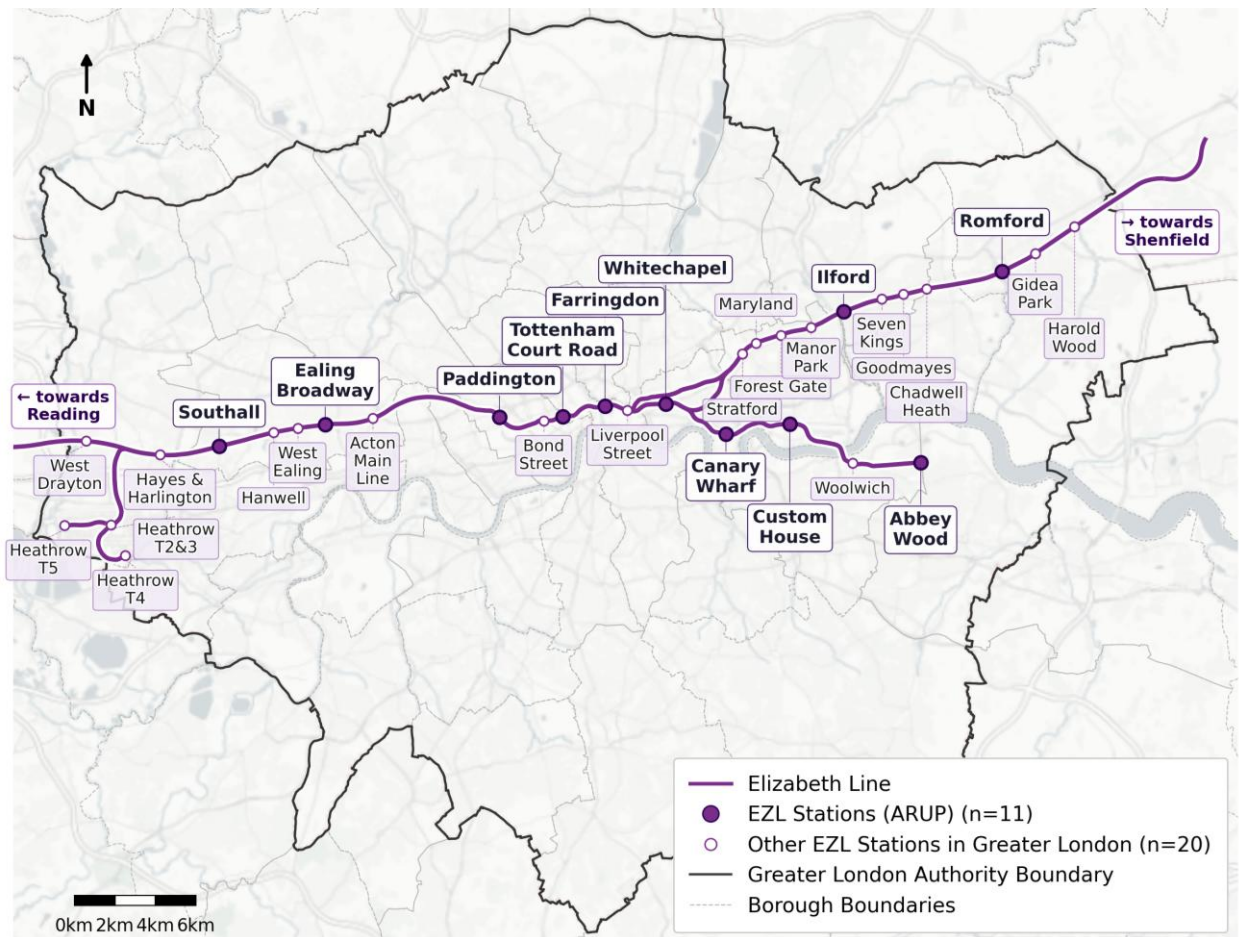
Post-opening evaluation outcomes

Early post-opening evaluation indicate substantial spatial and economic impacts. The line currently accommodates approximately 200 million annual passenger journeys, yielding corridor-specific journey-time reductions of more than 40% ([Lee and van Dijk, 2026](#)). Furthermore, development monitoring demonstrates that between 2015 and 2024, approximately 71,000 new homes were completed within 1 kilometre of Elizabeth line stations, alongside 125,000 new jobs registered within the same radius between 2022 and 2023 .

Focus of study

This study focuses on the May 2022 opening of the Elizabeth line's central tunnelled section, which delivered a step-change in regional accessibility and passenger throughput via through-running services to central London ([Arup and Department for Transport, 2025c](#)). To establish the candidate treated set, the study adopts the station-selection framework from Arup's post-opening evaluation, which updated the original Crossrail Baseline Evaluation (CBE). The CBE initially evaluated all 41 network stations against 13 criteria to derive a representative subset ([Arup and Department for Transport, 2025b](#)). The post-opening update retained this baseline and incorporated a twelfth station, Ilford, to capture the Eastern section. Restricting this set to Greater London (excluding Slough) yields 11 representative treated stations. [Figure 2](#) situates these within the full 41-station network, distinguishing the 11 representative stations, 20 other Elizabeth line stations within Greater London, and 10 stations outside it.

Aligning with this framework ensures comparability with official post-opening evaluation. However, because these stations were selected for descriptive representativeness rather than causal identification, [Section 5](#) subjects them to data-completeness and confounding filters to construct the final difference-in-differences samples. Treatment timing is defined not by initial surface-level operations, but by the commencement of central-tunnel through-running in May 2022, which represents the common accessibility shock granting direct, high-frequency access across all treated stations.



Map tiles © CARTO, Data © OpenStreetMap contributors.
 Linestrings and station coordinates for the Elizabeth Line obtained from Transport for London (TfL).
 Note: Elizabeth line stations not shown ($n = 10$, outside Greater London) include Iver, Langley, Slough, Burnham, Taplow, Maidenhead, Twyford, Reading, Brentwood and Shenfield.

Figure 2: Geographical location of Elizabeth line stations in Greater London

4. Data

Estimating the impact of the Elizabeth line opening on surrounding retail activity requires high-resolution spatial data. This study draws on a novel dataset, OpenLocal (<https://openlocal.uk/>), compiled from Valuation Office Agency (VOA) non-domestic rating lists and Local Authority (LA) ratepayer records. Specifically, OpenLocal integrates property-level (*hereditament*) attributes, such as rateable values, floor space, and detailed sub-valuation descriptions, with LA billing data, including account names, occupancy status, billing references, and rate relief records (Whythawk, 2026). Individual records are geocoded using Office for National Statistics (ONS) address databases and postcodes.

4.1. Data quality and coverage

OpenLocal offers good spatial and sectoral granularity, providing quarterly observations tracked by unique property identifiers (Unique Address Reference Numbers, or UARNs). Sub-valuation descriptions allow fine-grained categorization of business activities (e.g., retail, office, leisure). Account name and occupancy status are present for around 40%-60% of records. Where an account name is observed in consecutive quarters, a change in that name could signal a turnover event. Aggregating these events across a catchment's hereditaments yields the quarterly occupant turnover rate that forms the study's primary outcome.

However, overall data quality is mixed due to differences in reporting practices across LAs. For example, administrative lags were found, with timeliness dependent on LA update schedules. Certain LAs exhibited "frozen" records, where listed account names remain static over extended runs despite active occupancy, or complete name suppression possibly from privacy concerns. Most pertinently, VOA rating list updates such as the major revision in April 2023 can result in discontinuities with UARN altered or retired across revaluation boundaries.

4.2. Data cleaning and processing

[Figure 3](#) summarises the workflow used to construct a balanced panel, derive business turnover rates, and classify hereditaments into chain or independent businesses. All data pre-processing and cleaning was executed using Python.

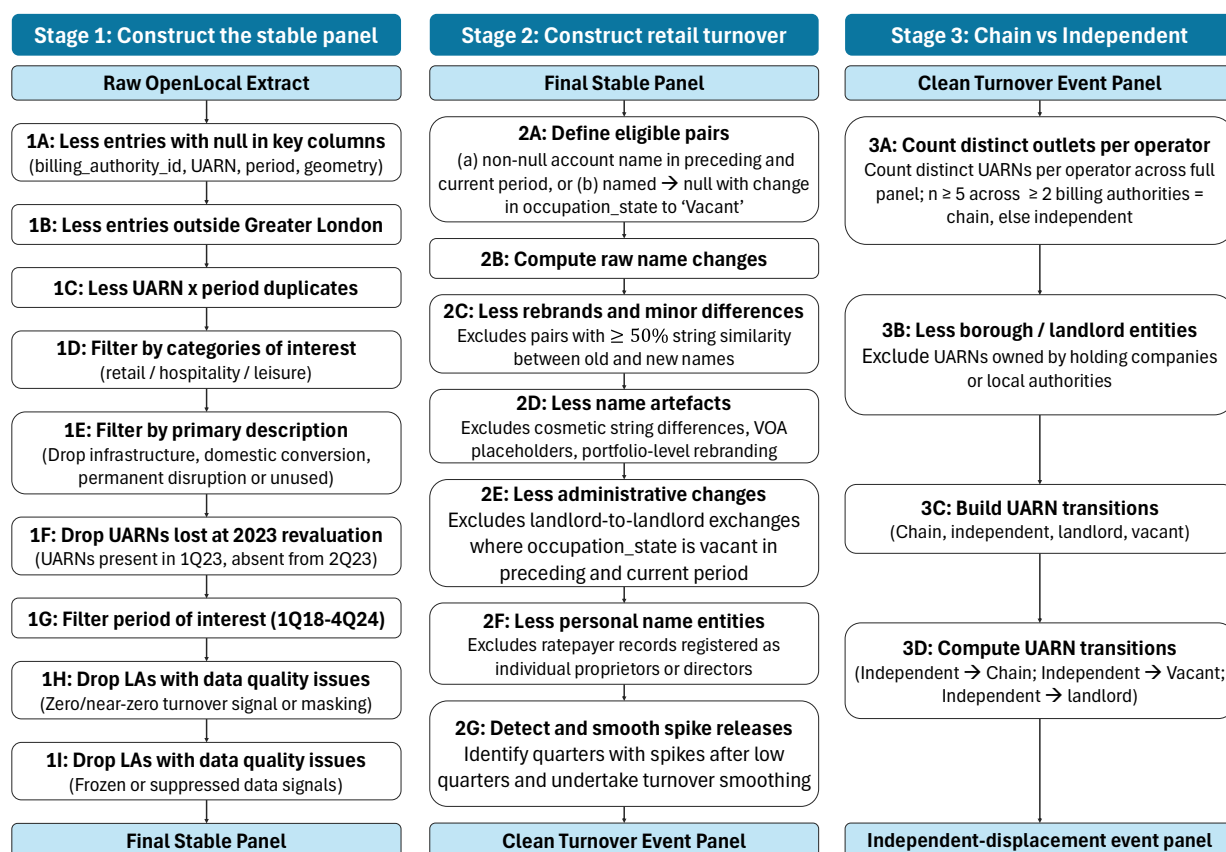


Figure 3: Workflow for data cleaning and pre-processing of OpenLocal data

Stage 1: Construction of stable panel

The raw dataset was restricted to retail-relevant commercial uses (e.g., high street retail, food and beverage, financial/professional services, food superstores, hotels, cinemas, and indoor leisure) (Stage 1D). Non-occupiable infrastructure (e.g., advertising hoardings, public kiosks) were excluded (Stage 1E). Hereditaments present prior to the April 2023 revaluation but absent from all subsequent quarters were dropped (Stage 1F). As revaluations can reassign UARNs, retaining these missing units would erroneously register identifier changes as business closures.

The overall panel was restricted to 1Q2018 – 4Q2024 (Stage 1G). Pre-2018 quarters were excluded due to low ratepayer name coverage (<50%), while post 4Q2024 records were omitted to avoid incomplete quarterly returns.

Of the 33 Greater London Local Authorities, 10 were excluded due to structural data limitations (Stage 1H and 1I). Four had severe ratepayer name masking ([Table 1](#)) and six had extended frozen or suppressed update patterns ([Figure 4](#)).

Table 1: Top 5 Local Authorities (LA) with ratepayer name masking

Local Authority	UARNs	Observations	Account name null (%)	Occupancy state null (%)	Outcome
Enfield	3,602	112,441	100	100	Dropped
Hackney	4,425	136,488	100	100	Dropped
Westminster	11,654	346,428	100	100	Dropped
Merton	2,344	71,534	83.6	90.3	Dropped
Haringey	3,817	119,705	55.3	22.8	Retained

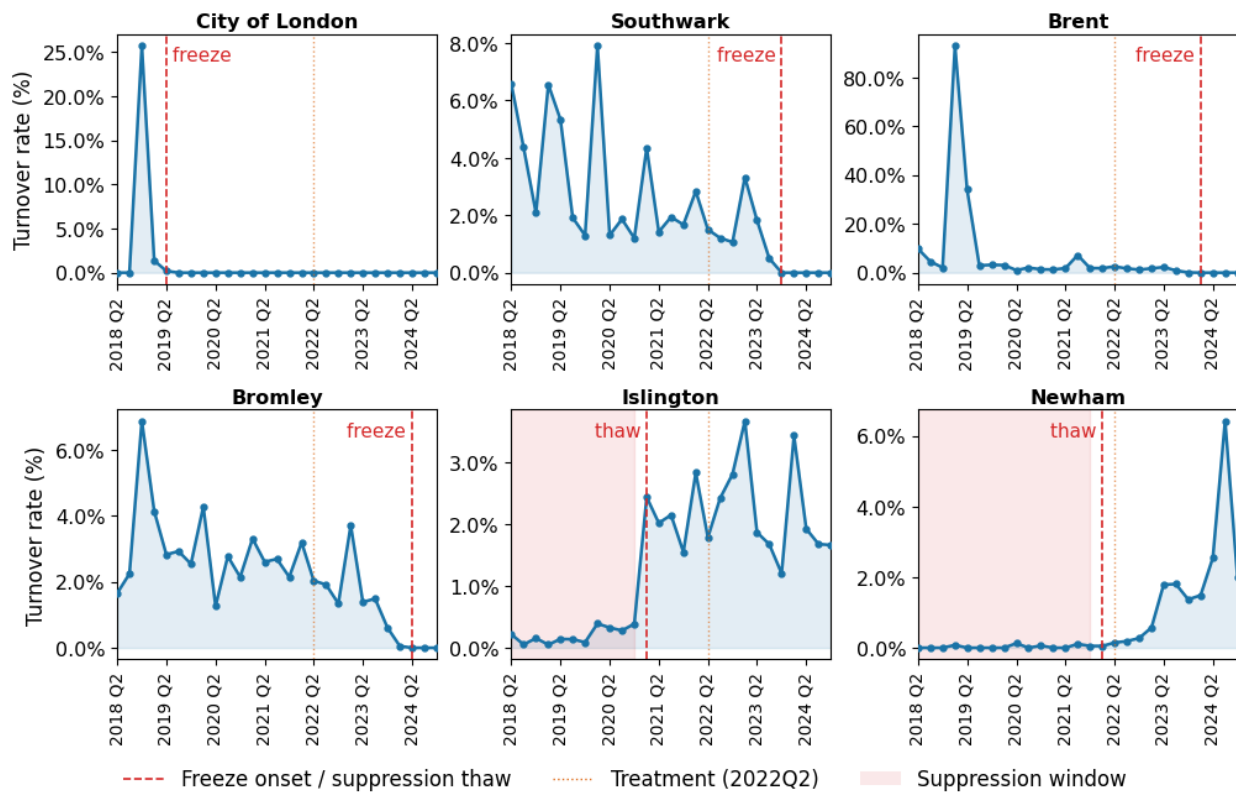


Figure 4: Frozen and suppressed turnover rate outcomes across six Local Authorities

Stage 2: Construction of retail turnover panel

A turnover event, serving as a proxy for occupant churn, was defined as an account name change between consecutive quarters for a given hereditament. To identify these ratepayer changes, consecutive quarterly observations were paired. Eligible pairs required either non-null account names in both periods or a transition to a null entry accompanied by an occupation state shift to 'Vacant' (Stage 2A). To isolate genuine tenant turnover from administrative noise, a multi-stage filtering pipeline removed account pairs with $\geq 50\%$ string similarity (Stage 2C) as well as cosmetic string differences, VOA placeholders, and portfolio-level rebranding (Stage 2D). [Figure 5](#) illustrates this process.

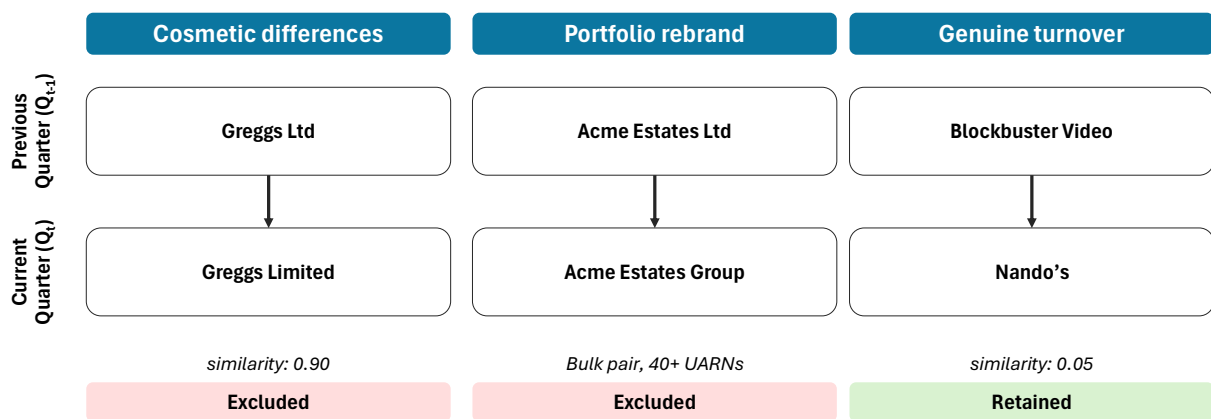


Figure 5: Illustration of cleaning approach used to isolate genuine turnover events

Further refinements removed non-tenancy changes and reporting anomalies such as landlord-to-landlord account exchanges (Stage 2E), and ratepayer records registered under individual proprietors (Stage 2F). Finally, to account for administrative reporting lags, quarters exhibiting artificial activity spikes following low-reporting periods were identified and smoothed (Stage 2G), yielding the final clean turnover event panel.

Turnover rate (Equation 1) was selected as the primary dependent variable over raw turnover counts to enable comparison and analysis between different spatial areas. [Table 2](#) sets out this rationale.

$$\text{Turnover Rate}_{i,t} = \frac{\text{Turnover Events}_{i,t}}{\text{Eligible Hereditaments}_{i,t}} \quad (1)$$

Table 2: *Evaluation and selection of primary dependent variable*

	Turnover counts	Turnover rate (proportion)
Definition	Number of occupant changes in LSOA i in quarter t	Number of occupant changes $\div$ total active hereditaments in LSOA i in quarter t
Output type	Non-negative integer	Continuous proportion, bounded [0,1]
Distributional properties	Frequently overdispersed; variance exceeds mean	Approximately continuous away from the bounds; heteroskedastic, as rate precision varies with stock size
Regression implications	Negative binomial DID or Poisson DID	OLS two-way fixed effects, weighted by eligible stock
Strengths	✓ Potentially more direct and intuitive interpretation in policy context ✗ May conflate stock size with turnover dynamics	✓ Normalises for stock size, measuring the propensity for turnover ✓ Directly comparable across LSOAs regardless of stock ✓ Simple to operationalise within a linear DID framework ✗ Rates from small-stock LSOAs are noisier, requiring stock weighting
Precedent in literature	Liu and Bardaka (2023)	No direct precedent

Stage 3: Classifying hereditaments into independent vs chain businesses

To enable the evaluation of displacement dynamics across business types, hereditaments were further categorised into independent and chain operators. Drawing from the methodology used by The Local Data Company, same-named entities operating five or more distinct outlets across the panel were classified as chain operators ([PwC, 2023](#)). An additional requirement that outlets must span two or more boroughs was used to prevent single-borough landlords, housing associations, or shopping centre freeholders from being misclassified as retail chains. Operators below the five-outlet threshold or operating within

a single borough were classified as independent operators. Single-borough entities with high property counts matching real estate or public-sector identifiers were isolated as ‘Landlord/Public Sector’.

4.3. Handling administrative artefacts

Because VOA ratepayer lists are built for tax administration rather than academic research, their updates reflect administrative processing practices that introduce artificial occupancy changes that may not reflect genuine market churn. Two such anomalies affect entire billing-authority quarters and cannot be directly corrected:

- a. **Sharp administrative spikes:** Borough-quarters where the recorded turnover rate exceeds 10 percentage points, far above any plausible commercial rate, reflecting bulk record updates rather than real-market churn. This affects seven borough-quarters across Harrow, Greenwich, Hammersmith & Fulham, and Tower Hamlets, In total, these spikes account for 19,614 hereditaments.
- b. **Unhandled suppression periods:** Intervals where data submission irregularities could not be reconciled into a clean series, such as a two-quarter data freeze in Bexley, affecting a further 4,468 hereditaments.

In total, 24,082 observations were affected by these administrative anomalies and were flagged and excluded from the restricted robustness sample, with results reported alongside baseline models that retain the unadjusted sample.

4.4. Additional data

This study draws additional data from several dimensions: transport networks, population density, Index of Multiple Deprivation (IMD) rankings and land-use composition, to enable identification, matching and comparison of treatment and control areas. The smallest common geography for which datasets are available is the Lower-layer Super Output Area (LSOA) level ([ONS, 2017](#)).

- **Transport network:** Station coordinates for underground, overground and national rail stations were compiled from Transport for London GIS Open Data Hub ([Transport](#)

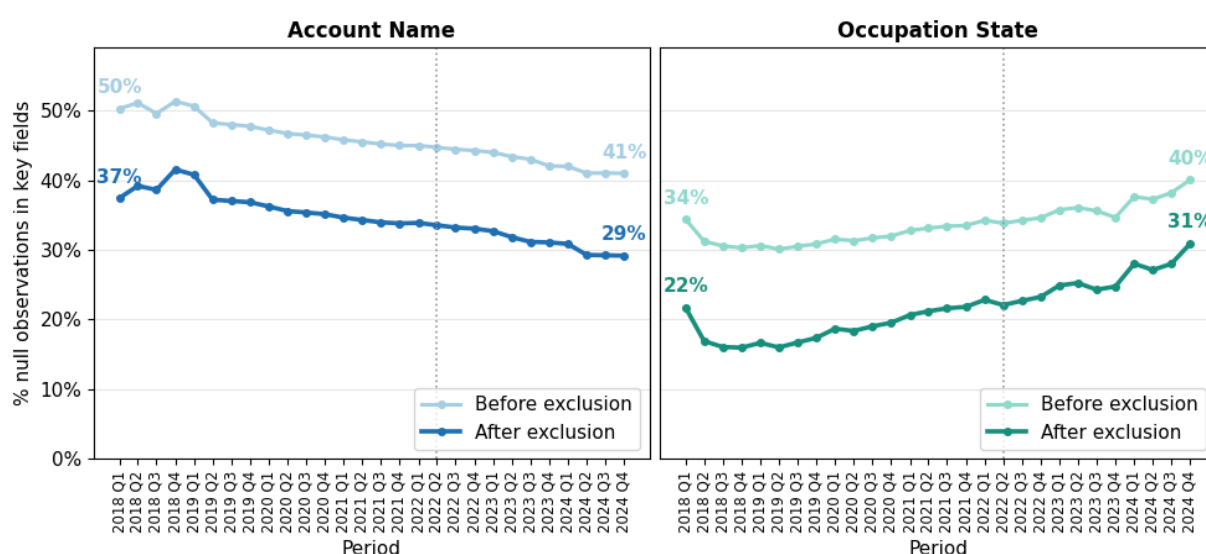
for London, 2025) and the National Public Transport Access Nodes (NaPTAN) dataset (Department for Transport, 2025).

- **Socioeconomic and built-environment controls:** Aggregated to the Lower Layer Super Output Area (LSOA) level:
 - **Deprivation:** Population-weighted mean Index of Multiple Deprivation scores from 2015 was compiled for each station catchment (MHCLG, 2015).
 - **Land Use:** Residential land area share for each station catchment in 2018 was calculated based on 8 use classifications (i.e. residential, commercial, industrial, civic, green, agricultural, transport, transitional) tagged by OpenStreetMap (OSM) with class definitions and tag criteria verified against OpenStreetMap Wiki (OSM, 2026).

4.5. Descriptive statistics

Data quality and missing values

Data cleaning and borough exclusions improved overall data completeness. Removing ten boroughs with structurally incomplete data reduced null rates for account names (50% to 37%) and occupation states (34% to 22%). While account-name nulls continued to decline to 29% by Q4 2024 due to better reporting, occupation-state nulls rose to 31%, indicating a gradual deterioration in recording occupancy status (Figure 6).



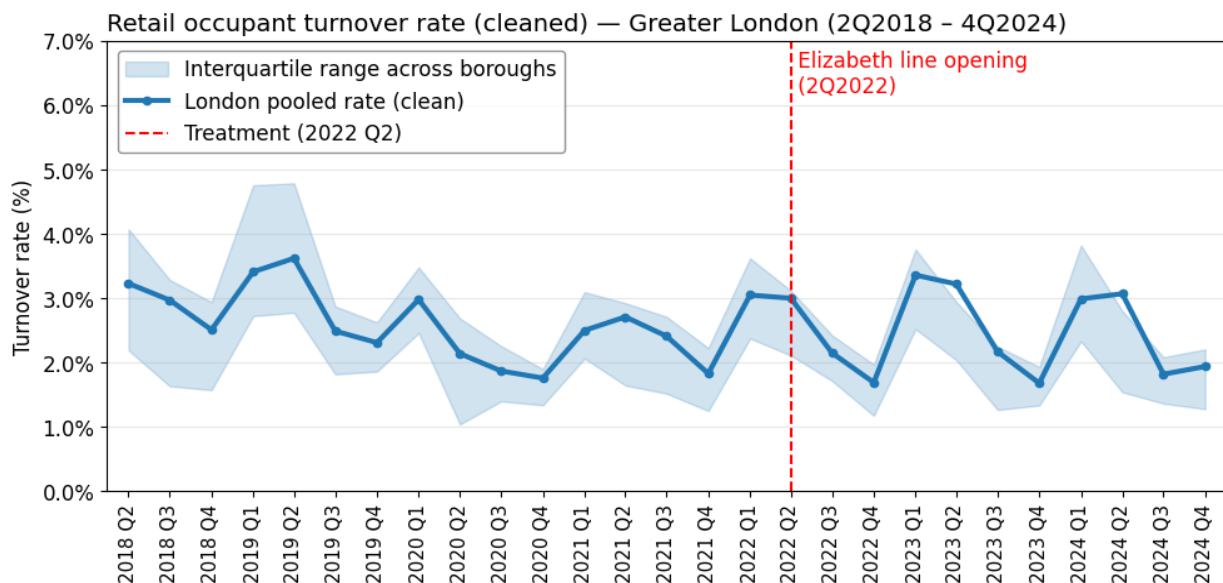
Note: 10 out of 33 boroughs were excluded due to structural data limitations. These include Westminster, Hackney, Enfield, Merton, City of London, Newham, Southwark, Brent, Bromley and Islington.

Figure 6: Percentage of null in key fields, before vs after data quality exclusions

Final panel summary and turnover

Following the data cleaning and processing pipeline, the final panel comprised 2,011,821 observations spanning 28 quarters (1Q2018–4Q2024). Across 1,254,548 eligible consecutive-quarter pairs, 31,757 turnover events were identified, corresponding to an average quarterly turnover rate of 2.53% across Greater London.

[Figure 7](#) illustrates the quarterly turnover trajectory bounded by the interquartile range across all boroughs. The pooled rate displays annual seasonality, with turnover peaking early in the calendar year. This likely reflects the administrative rhythm of ratepayer data with year-start revaluations and account updates concentrating recorded occupant changes. Because the pattern affects both treated and control catchments, it is absorbed by the quarter fixed effects in the Difference-in-Differences (DID) specification.



Note: Shaded band represents the interquartile range (25th–75th percentile) of borough-level turnover rates across the 23 boroughs retained in the panel. The pooled rate sits marginally above the IQR in 2Q2023 and 2Q2024, driven by elevated turnover in Tower Hamlets (10.5%) and Greenwich (17.0%) respectively.

Figure 7: Retail turnover rate across Greater London (2Q2018 – 4Q2024)

Average quarterly turnover across LSOAs

At the Lower Layer Super Output Area (LSOA) level, average quarterly occupant churn decreased modestly from 2.36% prior to the opening of the Elizabeth line (May 2022) to 2.11% in the post-opening period ([Figure 8](#)). This shift reflects broader commercial

consolidation trends as well as a national contraction in the retail outlet stock driven by the continued shift to online retail, and cost pressures (PwC, 2023).

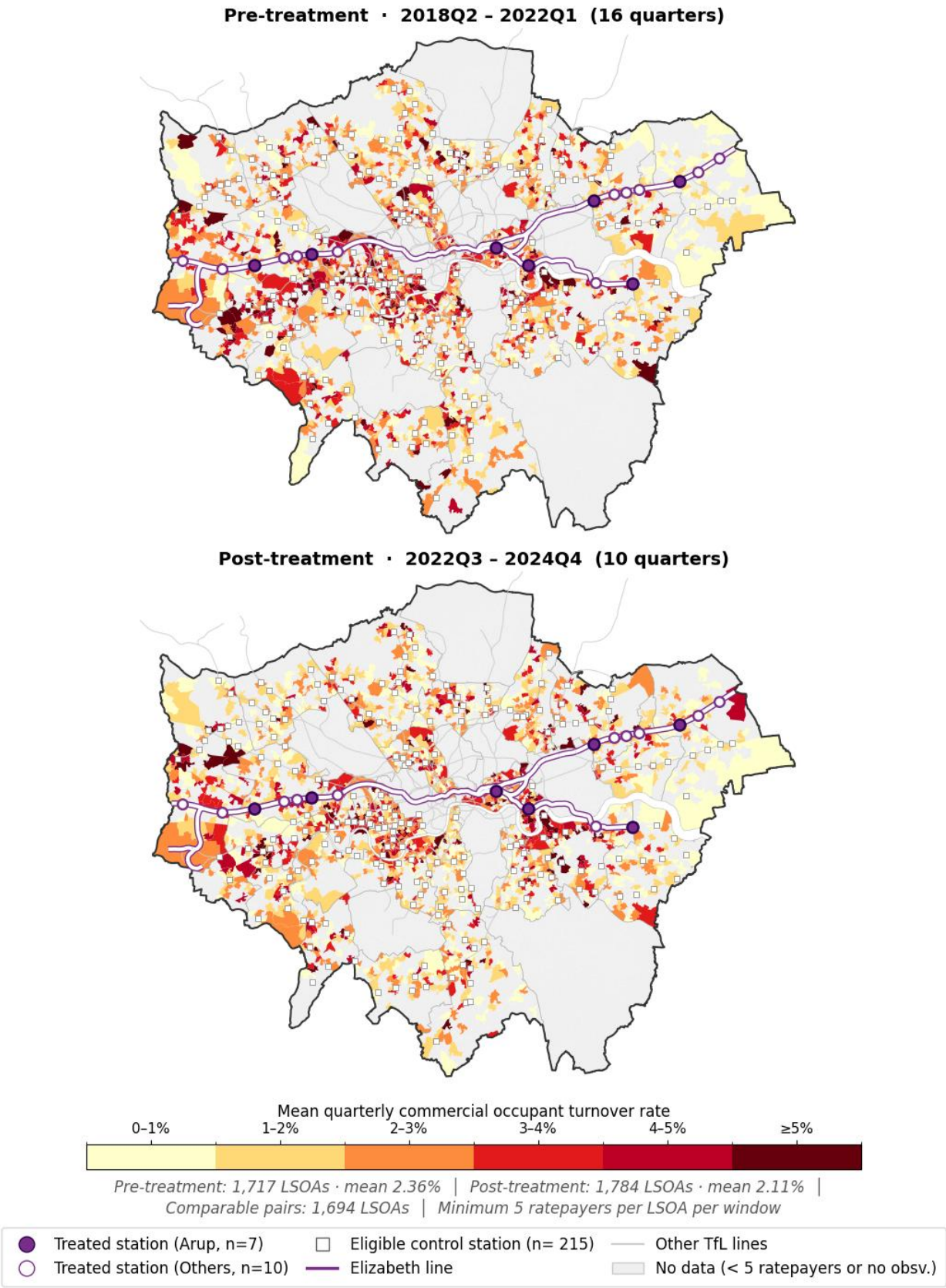


Figure 8: Average quarterly retail turnover rate by LSOA (Pre/Post EZL opening)

5. Empirical Strategy

5.1. Research design and study period

The objective of this paper is to estimate the impact of the opening of the Elizabeth line on business churn in surrounding catchments. To establish a causal link, a quasi-experimental approach using Difference-in-Differences (DID) with matching was adopted across three steps: (1) defining treatment and control catchments, (2) matching to reduce baseline characteristic imbalances, and (3) estimating the treatment effect using DID regression.

The study period spans 2018-2024, encompassing the latter part of the pre-treatment construction and operational phases for the Elizabeth line. May 2022 was taken as the treatment date across stations, on the basis that the Elizabeth line only delivered fast through-running services connecting the full corridor from May 2022. While there were service changes prior under the TfL rail brand in 2015 and 2018, these were taken as incremental improvements rather than treated as an intervention of interest.

5.2. Defining treatment and control catchments

5.2.1. Identifying treatment stations

Two specifications were formulated to identify treatment stations. First, a tighter subset of 11 stations was selected, mirroring the approach taken by Arup in its post-implementation study to establish a set of representative stations ([Arup and Department for Transport, 2025b](#)). Second, a broader set of 31 stations in Greater London, was used to evaluate the robustness of findings extended to the rest of the Elizabeth line.

Treatment catchments were established using 15-minute walking isochrones generated via the TravelTime API ([TravelTime, 2025](#)) from each station centroid. All 2011 LSOAs with centroids within this 15-minute walking distance were taken to comprise this catchment area, similar to the approach used in [Whearty, Murcio and Barros \(2026\)](#).

Next, treatment catchments were filtered to exclude stations with confounding events during the study period (e.g., Heathrow airport pandemic closures) and those with fewer than 50 hereditaments. Additionally, several central London stations, including Farringdon,

Paddington, Tottenham Court Road, and Custom House, were dropped due to incomplete OpenLocal borough data (City of London, Islington, Westminster, and Newham). [Table 3](#) lists final treatment stations, with catchment areas in [Figure 9](#).

Table 3: Composition of final treatment stations across two specifications

	Representative subset (Arup)	All EZL stations in Greater London
Eligible set	11 stations	31 stations
Dropped due to incomplete data	<u>-4 stations</u> Custom House, Paddington, Farringdon, Tottenham Court Road	<u>-10 stations</u> Bond Street, Custom House, Paddington, Farringdon, Forest Gate, Liverpool Street, Manor Park, Maryland, Stratford, Tottenham Court Road
Dropped due to confounding events	Nil	<u>-3 stations</u> Heathrow T2&3, Heathrow T4, Heathrow T5
Dropped due to small number of businesses	Nil	<u>-1 station</u> Harold Wood
Final treatment station set	=7 stations Abbey Wood, Canary Wharf, Ealing Broadway, Ilford, Romford, Southall, Whitechapel	=17 stations Acton Main Line, Abbey Wood, Canary Wharf, Chadwell Heath, Ealing Broadway, Gidea Park, Goodmayes, Hanwell, Hayes & Harlington, Ilford, Romford, Seven Kings, Southall, Whitechapel, Woolwich, West Ealing, West Drayton

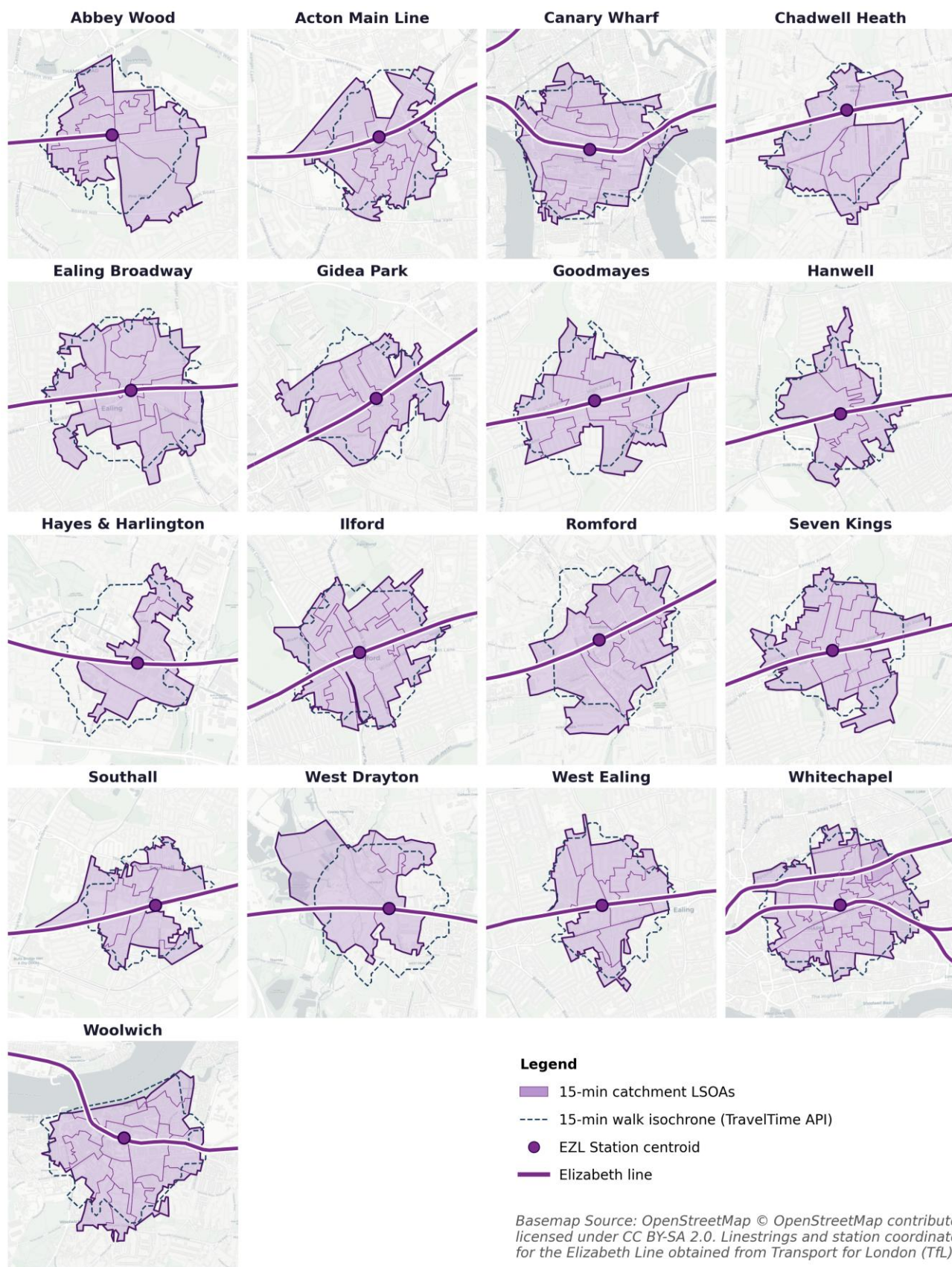


Figure 9: 15-min walking isochrones and catchment areas of final treatment stations

5.2.2. Identifying control stations

Candidate controls were drawn from all Tube and National Rail stations in Greater London. As with the treatment group, catchment areas were defined as 2011 LSOAs with centroids within a 15-minute walking distance from each station.

To ensure independence from treatment, prevent confounding, and maintain comparability, candidate stations were filtered based on the following criteria:

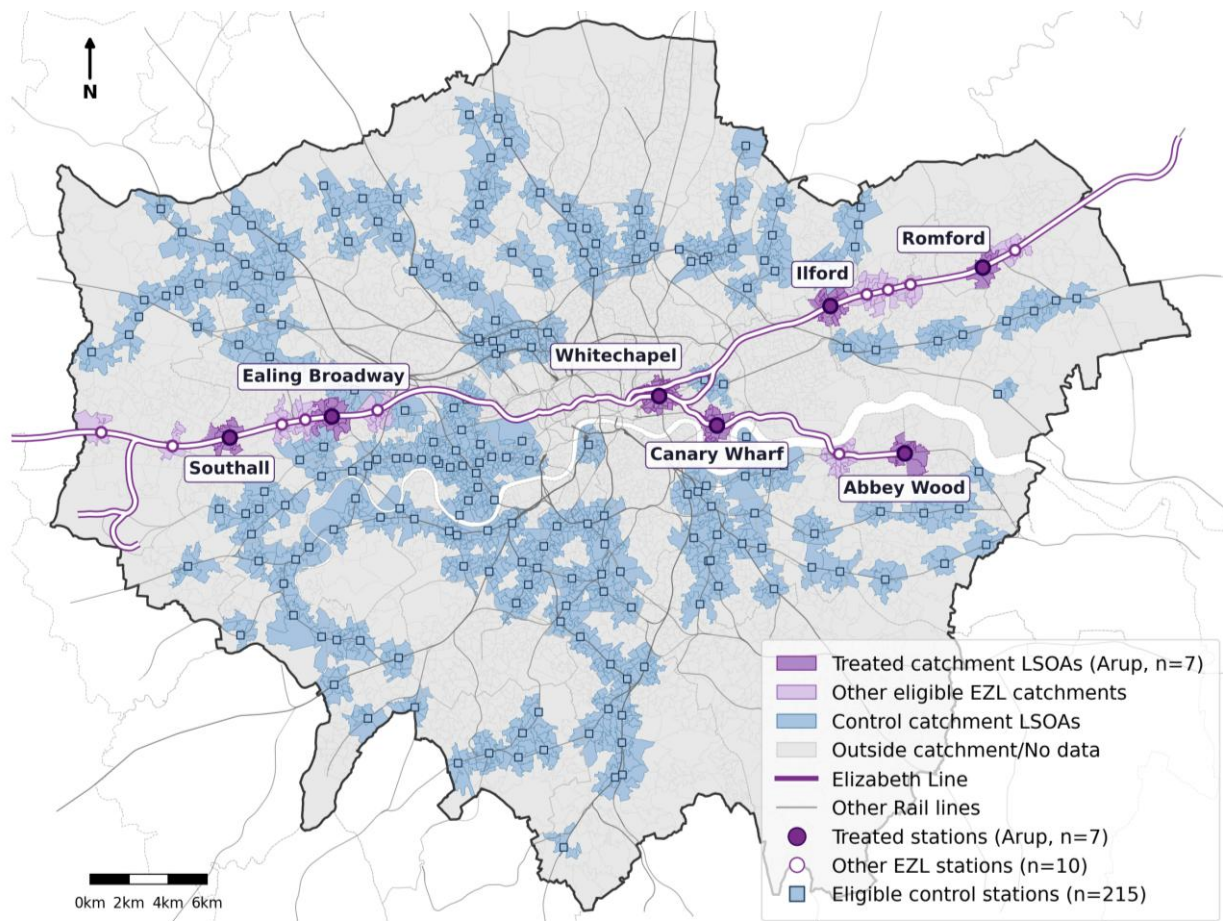
- a. **Elizabeth line overlap:** Station catchments intersecting Elizabeth line catchments.
- b. **Non-comparable transport modes:** DLR-only, London Trams, and Cable Car stations.
- c. **Orbital routes:** Stations exclusively on orbital lines (e.g., Suffragette, Liberty, Mildmay, and Windrush Overground lines), which differ fundamentally in character from radial routes.
- d. **Confounding major investments:** Catchments intersecting major network developments during the study window (e.g., the Northern Line Extension to Battersea Power Station/Nine Elms in 2021 ([Transport for London, 2021](#)), or the Overground extension to Barking Riverside in 2022 ([Transport for London, 2022a](#))), and catchments which saw major retail mall expansions (e.g. Westfield White City Expansion ([Retail Insight Network, 2018](#)))
- e. **Small stations:** Footfall below 500,000 annually or fewer than 50 hereditaments.
- f. **Data constraints:** Insufficient turnover signal due to OpenLocal data limitations.

Stations benefiting from the Thameslink service frequency ramp-up (2019–2022) were flagged but retained ([AECOM, 2018](#)). Unlike the sharp geographical discontinuities created by new or extended lines, incremental enhancements like Thameslink frequency updates or Overground restructuring reflect broader background transit trends. While these background trends introduce minor precision constraints in isolating the Elizabeth line effect, they do not compromise the control group's validity.

[Table 4](#) summarises the systematic exclusions for controls and [Figure 10](#) depicts the spatial distribution of treatment and control catchments.

Table 4: Breakdown of exclusion criteria used to determine control stations

Step	Filter	Removed	Remaining
0	All TfL and National Rail Stations	-	589
1	Remove Elizabeth line stations	41	548
2a	Remove confounding infrastructure investments (Battersea, Nine Elms, Barking Riverside, King's Cross St. Pancras)	4	544
2b	Remove confounding infrastructure investments (Wood Lane, White City, Shepherd's Bush, Shepherd's Bush Market, Latimer Road)	5	539
3	Remove orbital-only stations (Suffragette, Mildmay, Windrush, Liberty)	39	500
4	Remove DLR-only, trams and cable car	38	462
5	Remove low footfall or low hereditament stations	52	410
6	Remove stations outside Greater London	15	395
7	Remove stations with catchment overlap with Elizabeth line	69	326
8	Remove stations without sufficient coverage	111	215
Final number of eligible control stations			215



Note: Stations excluded where the catchment falls largely in billing authorities dropped for poor data quality: Westminster, Hackney, Enfield, Merton (no ratepayer data submitted) and City of London, Newham, Southwark, Brent, Bromley (record freezes).

Figure 10: Eligible treated and control catchments after exclusions

5.3. Matching controls on covariate characteristics

An important precondition for effective DID analysis is that treated and control units have sufficiently comparable pre-treatment characteristics to support the parallel trends assumption ([Callaway and Sant'Anna, 2021](#)). Because Elizabeth line stations were not randomly placed, targeting areas with high baseline connectivity and employment density, a raw DID risks conflating structural differences with the treatment effect. Preprocessing data via matching reduces these pre-treatment confounder imbalances between groups ([Blackwell et al., 2009](#); [Huntington-Klein, 2021a](#)).

The literature offers several broad families of approaches for selecting and weighting comparison units, distinguished by how they handle covariate dimensionality and trade off bias reduction against efficiency.

- **Matching Estimators** (e.g., Nearest-Neighbour, Propensity Score Matching, Mahalanobis Distance) identify control units "close" on observed covariates to form a restricted sample ([Huntington-Klein, 2021a](#)). Propensity Score Matching (PSM) collapses multidimensional covariates into a single scalar representing treatment probability ([Rosenbaum and Rubin, 1983](#)). However, PSM can suffer from the "propensity score paradox" ([King and Nielsen, 2019](#)). In finite samples, it balances data only on average, leaving residual imbalances where small specification changes could drastically alter matches and potentially increase model dependence relative to no matching. Alternatively, Mahalanobis Distance Matching (MDM) computes distance directly across the multivariate covariate space, accounting for scale and covariance between variables without relying on a propensity score model ([Huntington-Klein, 2021a](#)).
- **Weighting Estimators** such as Inverse Probability Weighting (IPW) retain the full sample but reweight control units using the propensity score, creating an observationally equivalent pseudo-population on average ([Graham, 2025](#)). PSM and IPW are natural complements as both leverage the same propensity score model, but PSM discards unsuited units while IPW reweights the retained sample ([Huntington-Klein, 2021a](#)).

- **Coarsened Exact Matching (CEM)** takes a different approach. Instead of modelling treatment assignment, CEM coarsens continuous covariates into substantive bins (e.g., distance bands), exact-matches within strata, and discards unmatched strata ([Iacus, King and Porro, 2012](#)). This bounds maximum imbalance *ex ante*, a property termed Monotonic Imbalance Bounding (MIB), where tightening one covariate's coarsening never worsens imbalance on others ([Iacus, King and Porro, 2012](#)). Balance is thus explicitly dictated by user judgment rather than emerging as an incidental byproduct of a propensity score model.

This study adopts CEM as the preferred specification, as it is well-suited to the heterogeneity of London's station catchments, where treated and control areas vary substantially across multiple dimensions. An alternative matching specification using Mahalanobis distance matching with $k = 3$ nearest neighbours ([Huntington-Klein, 2021a](#); [Graham, 2025](#)) was used as a robustness check.

Coarsened Exact Matching (CEM) approach

Five covariates were selected for matching, each capturing a distinct structural dimension of catchment variability ([Table 5](#)). To mitigate the "curse of dimensionality" inherent to CEM ([Iacus, King and Porro, 2012](#)), covariates were coarsened into a small number of substantively meaningful bins rather than matched on continuous values. Because the number of strata scales exponentially with covariate count and bin granularity, overly fine coarsening generates empty or single-treatment strata, forcing otherwise viable units to be discarded. Restricting the specification to five deliberately coarsened dimensions keeps the total stratum count small relative to sample size, maximizing the retention of common-support strata containing both treated and control units ([Blackwell et al., 2009](#)).

Table 5: Rationale for selected CEM covariates

Covariate	Rationale
Distance to city (<i>distance_central_km</i>)	Proxies position along London's urban gradient. Rent levels, footfall, and retail mix vary systematically with centrality; controlling for distance prevents confusing centrality with treatment effects.
Number of rail lines (<i>num_lines</i>)	Captures baseline multimodal connectivity. Highly connected stations attract larger, more diverse commuter bases, which independently drives commercial demand and turnover dynamics.
Log hereditament density (<i>log_here_density</i>)	Measures baseline commercial saturation. Dense catchments experience structurally higher competition and churn; controlling for density isolates the transport intervention from underlying market scale.
Area deprivation – mean IMD across catchment LSOAs (<i>imd_mean</i>)	Proxies local consumer purchasing power across catchment LSOAs. Deprived areas face constrained demand, limiting store survival regardless of transit access. This is crucial given that several Elizabeth line stations span historically deprived areas.
Residential land-use percentage (<i>perc_residential</i>)	Differentiates destination commercial hubs from local neighbourhood retail centres, which exhibit fundamentally different retail mixes, customer bases, and structural turnover rates.

Alternative covariates were considered but excluded to avoid collinearity. Log population density and log employment density correlated strongly with log hereditament density ($r = 0.71$ and $r = 0.72$, respectively), as all three proxy baseline economic activity. Similarly, percentage commercial land use was omitted due to conceptual overlap with hereditament density. Conversely, percentage residential land use avoids this redundancy, distinguishing destination shopping hubs from neighbourhood retail, and exhibited low correlation across all candidate covariates ($|r| \leq 0.31$, [Annex B](#)). Variance inflation factor (VIF) diagnostics for the final five-covariate specification were below 2 for all variables, confirming no problematic multicollinearity.

Coarsened Exact Matching (CEM) configuration

Given the small number of treated stations relative to the control pool, strict one-to-one matching would unnecessarily discard control units. Instead, this study employs one-to-many matching, assigning each treated station all control units within its same CEM stratum, yielding 1–3 controls per treated unit. To prevent strata with larger control counts from disproportionately influencing estimates, observations were assigned CEM weights following (Iacus, King and Porro, 2012), where treated units received a weight of 1, while control units were weighted inversely to their stratum's relative size.

Bin edges were selected by combining domain knowledge regarding each covariate's natural units with empirical validation against the control distribution. Figure 11 illustrates the pre-treatment covariate distributions, and Table 6 outlines finalized bin boundaries.

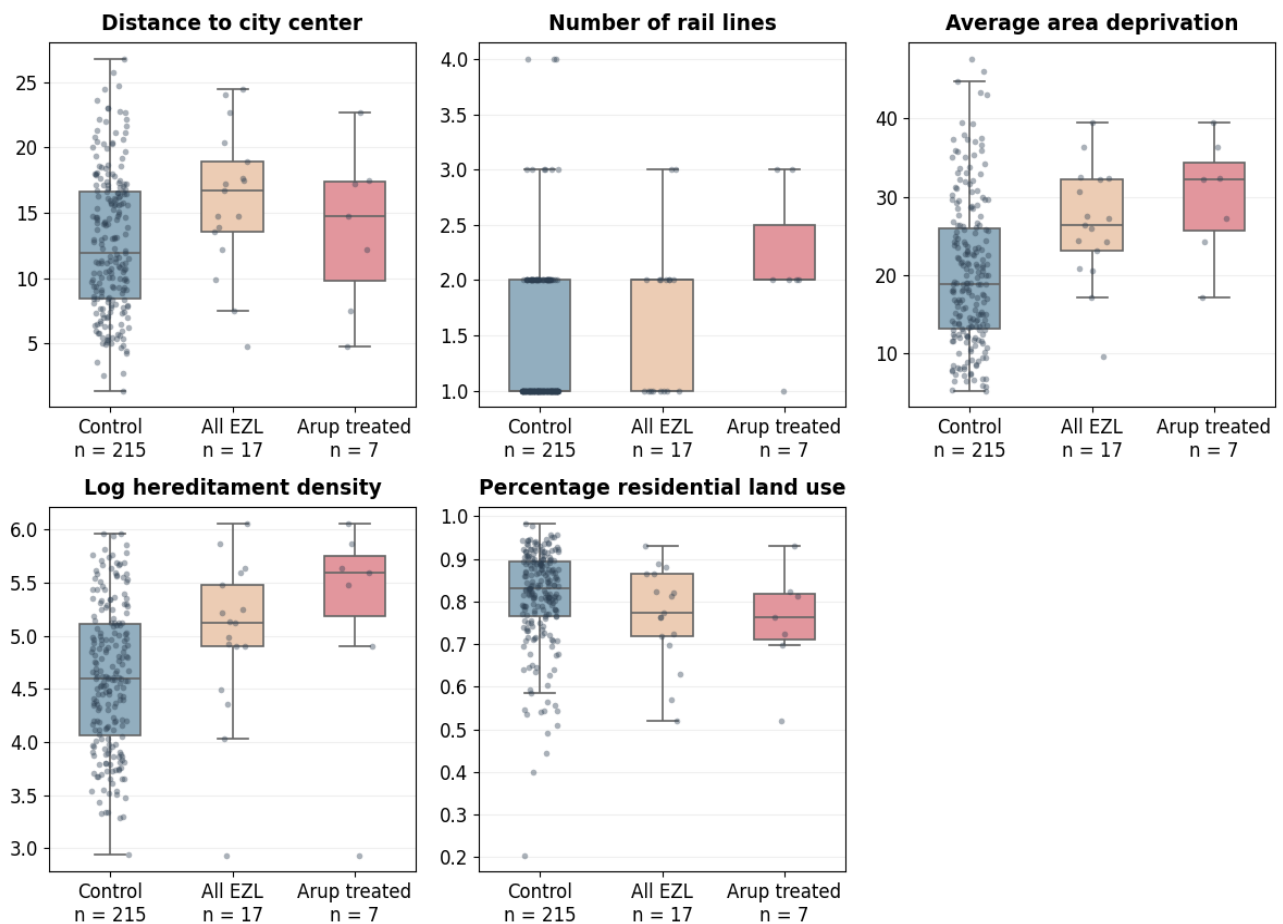


Figure 11: Distribution of pre-treatment covariates in treated and control catchments

Table 6: Finalized bin boundaries for covariate set and bin boundary rationale

	Covariate	Bins	Bin Rationale
1	Distance to city centre (<i>distance_central_km</i>)	[1.0, 8.0, 12.5, 20.0, 40.0]	[1-8km]: Inner London catchment [8-12.5km]: Inner-to-Outer transition [12.5km-30]: Outer Metropolitan Edge
2	Number of rail lines (<i>num_lines</i>)	[0, 1, 3, 5]	[0-1]: Single-line stations [2-3]: Moderate interchanges [4-5]: Large interchanges
3	Area deprivation (<i>imd_mean</i>)	[0, 10, 20, 40, 70]	[0-10]: Very low deprivation [10-20]: Low to moderate deprivation [20-40]: Moderate to high deprivation [40-70]: Very high deprivation
4	Log Hereditament Density (<i>log_here_density</i>)	[2.1, 4.1, 4.5, 5.5, 6.4]	No standard convention for commercial hereditament density. Edges were derived from the empirical distribution of the control pool. [2.1-4.1] Low density areas [4.1-4.5] Transitional bin [4.5-5.5] Modal commercial density bin [5.5-6.4] High commercial density areas
5	Residential land-use percentage (<i>perc_residential</i>)	[0.20, 0.8, 1.0]	Single split at the control-pool median distinguishes predominantly residential catchments from more mixed-use catchments [0.2-0.8] Mixed-use residential catchments [0.8-1.0] Predominantly residential catchments

5.4. Difference-in-Differences (DID) analysis

Difference-in-Differences (DID) is used to estimate causal effects by comparing changes in outcomes over time between a group exposed to an intervention (treatment group) and a group that is not (control group). By subtracting the pre-to-post change in the control group from the pre-to-post change in the treatment group, DID removes both baseline time-invariant differences between groups and shared temporal shocks, isolating the net impact of the policy intervention ([Callaway and Sant'Anna, 2021](#); [Huntington-Klein, 2021b](#); [Graham, 2025](#)).

The adoption of the DID framework is well-suited to the context of the Elizabeth line intervention for three main reasons. First, the opening of the central tunnel section in May 2022 serves as a well-defined treatment event separating the pre- and post-intervention periods. Second, while station locations were determined by historical planning decisions, the timing of the full end-to-end integration provides a clear spatial division between station catchment areas directly benefiting from the upgraded infrastructure and comparable control catchment areas. Third, the availability of quarterly observations allows for the tracking of commercial occupant churn across spatial units before and after the opening.

Key identification assumptions

The DID framework requires two core identification assumptions.

- First, the **Parallel Trends Assumption** states that in the absence of the Elizabeth line opening, the average change in retail occupant churn in the treatment catchment areas would have followed the same trajectory as in the control catchment areas ([Callaway and Sant'Anna, 2021](#); [Huntington-Klein, 2021b](#)). This assumption can plausibly be evaluated by examining pre-intervention quarterly trends via an event-study specification, in which pre-treatment coefficients jointly indistinguishable from zero support parallel trends, while a systematic pre-treatment divergence would indicate its violation.
- Second, the **Stable Unit Treatment Value Assumption (SUTVA)** requires that the treatment status of one spatial catchment area does not affect the outcome of another (i.e., no spatial spillovers), and there are no hidden variations in treatment intensity across units ([Roth et al., 2023](#); [Graham, 2025](#)).

Simple DID analysis

Following [Gertler et al., \(2010\)](#), a simple 2×2 difference-in-differences was first computed as the change in mean occupant churn rate in the treatment group relative to the change in the comparison group. These unweighted group means provide an indicative baseline for the treatment effect, prior to the CEM-weighted regression estimates.

DID regression analysis

To permit tests of statistical significance, DID regression analysis was conducted via a two-way fixed effects (TWFE) specification ([Angrist and Pischke, 2009](#); [Wooldridge, 2010](#); [Huntington-Klein, 2021b](#)). Turnover events were aggregated to the LSOA-quarter level as a turnover rate, giving a panel of LSOAs observed over multiple quarters. Because rate precision varies with the number of hereditaments in each LSOA, observations are weighted by eligible stock. The specification introduces two sources of unobserved heterogeneity that must be controlled for simultaneously:

- First, **LSOA fixed effects** absorb all time-invariant characteristics of each spatial unit such as its historical commercial composition, built environment characteristics, and structural position in the London economy.
- Second, **quarterly time fixed effects** absorb aggregate shocks that affect all LSOAs simultaneously in each period, including macroeconomic conditions, COVID-19 disruption, and London-wide commercial trends unrelated to the Elizabeth line.

The TWFE DID model can be expressed in Equation (2) as:

$$Y_{it} = \alpha + \beta(D_i \times T_t) + \mu_i + \gamma_t + \varepsilon_{it} \quad (2)$$

where Y_{it} denotes the commercial occupant churn rate in LSOA i in quarter t ; D_i is a binary indicator equal to 1 if LSOA i falls within the catchment area of an Elizabeth line station (the treatment group) and 0 otherwise; T_t is a time dummy equal to 1 for quarters following the end-to-end opening in May 2022 and 0 for pre-opening quarters; μ_i captures LSOA-level fixed effects; γ_t captures quarter-level time fixed effects; and ε_{it} is the error term.

Observations are weighted by eligible hereditament stock, so that LSOAs whose rates are measured more precisely contribute proportionally more, and by their CEM stratum weights, preserving the matched design. Standard errors are clustered at the station level to allow for correlated outcomes among LSOAs sharing a catchment.

The coefficient of interest, β , represents the average treatment effect on the treated (ATT), the differential change in retail occupant churn in Elizabeth line catchment LSOAs relative to matched control LSOAs following the opening, over and above any system-wide quarterly trends captured by γ_t . Because the outcome is a rate, β is interpreted directly as a percentage-point change in the quarterly churn rate.

TWFE DID robustness checks

To guard against the possibility that any estimated effect reflects covariate-correlated divergence rather than the opening itself, a doubly-robust specification (Ho et al., 2007) augments the model with interactions between the post-treatment indicator and the five matching covariates in Equation (3):

$$Y_{it} = \alpha + \beta(D_i \times T_t) + \sum_k \delta_k (T_t \times X_{ki}) + \mu_i + \gamma_t + \varepsilon_{it} \quad (3)$$

where X_{ki} is the k^{th} catchment covariate (distance to city centre, number of rail lines, log hereditament density, mean IMD, and residential land-use share), and the δ_k coefficients absorb differential post-opening trends associated with each covariate. This two-step matching-plus-adjustment procedure is doubly robust meaning that the estimate remains consistent if either the matching or the outcome model is correctly specified (Ho et al., 2007). β retains its ATT interpretation.

Inference is assessed for robustness to two features of the error structure. First, with a small number of treated stations, station-clustered standard errors may understate the true uncertainty, as cluster-robust inference is unreliable with few clusters. Second, turnover in neighbouring LSOAs may be spatially correlated across different catchments. Conley (1999) standard errors address both concerns. By permitting spatial dependence between any two units within a specified distance, they relax the cross-catchment assumption, and by not relying on clustering, they avoid the few-clusters problem. They are

computed between LSOA centroids at a cutoff at 2 km, and address correlation in the errors rather than treatment spillovers, which bear on the SUTVA assumption rather than on inference, and are considered separately.

Dynamic Event Study design

To examine whether retail adjustment in Elizabeth line catchments preceded the May 2022 opening, and to assess the parallel trends assumption underlying the main DID estimator, we estimate a dynamic (event-study) specification replacing the single post-treatment interaction with a separate coefficient for each quarter relative to treatment, following the standard practice of [Sun and Abraham \(2021\)](#) and [Schmidheiny and Siegloch \(2023\)](#) in Equation (4):

$$Y_{it} = \alpha + \sum_{\substack{b=-10 \\ b \neq -1}}^{+10} \beta_b (D_i \times \mathbf{1}[\text{bin}_b]) + \mu_i + \gamma_t + \varepsilon_{it} \quad (4)$$

Leads and lags beyond $q-10$ and $q+10$ are pooled into endpoint bins rather than dropped, so that panel-edge quarters continue to contribute to estimation without introducing a separate, noisy coefficient for each distant period ([Schmidheiny and Siegloch, 2023](#)).

Statistically significant and trending pre-treatment estimates indicate anticipatory commercial adjustment consistent with businesses and landlords pricing in connectivity gains ahead of opening, while pre-treatment coefficients jointly indistinguishable from zero provide supporting evidence for the parallel trends assumption. This approach follows precedent in the transit impact literature where event study specifications serve simultaneously as a primary results vehicle and pre-trends diagnostic ([Pogonyi, Graham and Carbo, 2021](#); [Wexler and Fan, 2022](#); [Kandt, Van Dijk and Todd, 2025](#)).

6. Results

6.1. Coarsened exact matching (CEM) results

CEM was undertaken for two definitions of the treated-station set, each against a common pool of 215 eligible non-treated stations in Greater London. Matching Specification 1 (MS1) comprised the 7 treated stations in Arup's representative sample, while Matching Specification 2 (MS2) comprised 17 eligible Elizabeth line stations.

6 of 7 MS1 treated stations (85.7%) and 14 of 17 MS2 treated stations (82.4%) were matched to at least one control station sharing a coarsened covariate cell ([Table 7](#)). Abbey Wood was unmatched under both specifications as its hereditament density (2.93) fell below the minimum observed across the entire 215-station control pool, placing it outside the region of common support. West Drayton and West Ealing were additionally unmatched under MS2, reflecting covariate combinations not represented among comparably-sized outer-London control stations.

Aggregating catchment LSOAs across matched stations, MS1 comprised 73 treated and 101 control LSOAs (174 in total), and MS2 130 treated and 161 control LSOAs (291 in total). Over the 27-quarter turnover panel (2Q2018–4Q2024), this yielded 4,670 and 7,699 LSOA-quarter observations respectively.

Station matches are in [Table 8](#) and [Table 9](#), with full covariate metrics in [Annex C](#).

Table 7: Summary of CEM outcomes and match rates

Metric	Match Spec 1 (Arup 7)	Match Spec 2 (All EZL 17)
Treated stations	7	17
Matched	6 (85.7%)	14 (82.4%)
Unmatched	Abbey Wood	Abbey Wood, West Drayton, West Ealing
L1 before / after	0.927 → 0.125	0.916 → 0.167
L1 reduction	87.0%	81.8%
Controls	~8 (3.7% of pool)	~18 (8.4% of pool)

Table 8: Treatment and matched stations (MS1)

Treated station	Matched control(s)	#
Canary Wharf	Goldhawk Road, West Brompton, West Hampstead	3
Ealing Broadway	Gunnersbury, Maze Hill	2
Ilford	East Croydon	1
Romford	Uxbridge	1
Southall	West Croydon	1
Whitechapel	Goldhawk Road, West Brompton, West Hampstead	3
Abbey Wood	(unmatched)	-

Table 9: Treatment and matched stations (MS2)

Treated station	Matched control(s)	#
Acton Main Line	Crofton Park, Kew Bridge	2
Canary Wharf	Goldhawk Road, West Brompton, West Hampstead	3
Chadwell Heath	Burnt Oak, South Croydon, South Harrow	3
Ealing Broadway	Gunnersbury, Maze Hill	2
Gidea Park	Hatch End, Northwood	2
Goodmayes	Burnt Oak, South Croydon, South Harrow	3
Hanwell	Hounslow Central, Hounslow East, Selhurst	3
Hayes & Harlington	Uxbridge	1
Ilford	East Croydon	1
Romford	Uxbridge	1
Seven Kings	Burnt Oak, South Croydon, South Harrow	3
Southall	West Croydon	1
Whitechapel	Goldhawk Road, West Brompton, West Hampstead	3
Woolwich	Hounslow Central, Hounslow East, Selhurst	3
Abbey Wood	(unmatched)	-
West Drayton	(unmatched)	-
West Ealing	(unmatched)	-

Matching substantially reduced multivariate covariate imbalance, measured via the L1 statistic ([Iacus, King and Porro, 2012](#)), from 0.927 to 0.125 (87.0% reduction) for MS1, and from 0.916 to 0.167 (81.8% reduction) for MS2. Standardised mean differences (SMDs) improved for the majority of matching covariates post-match ([Figure 12](#)). Consistent with the bias–variance trade-off in coarsened exact matching ([Iacus, King and Porro, 2012](#)), further tightening the bin specification to eliminate residual imbalance was found to substantially reduce the number of matched treated stations and available controls, inflating standard errors without a commensurate improvement in balance. Residual covariate imbalance was addressed using the regression specification with covariate-specific post-opening trends rather than through further coarsening.

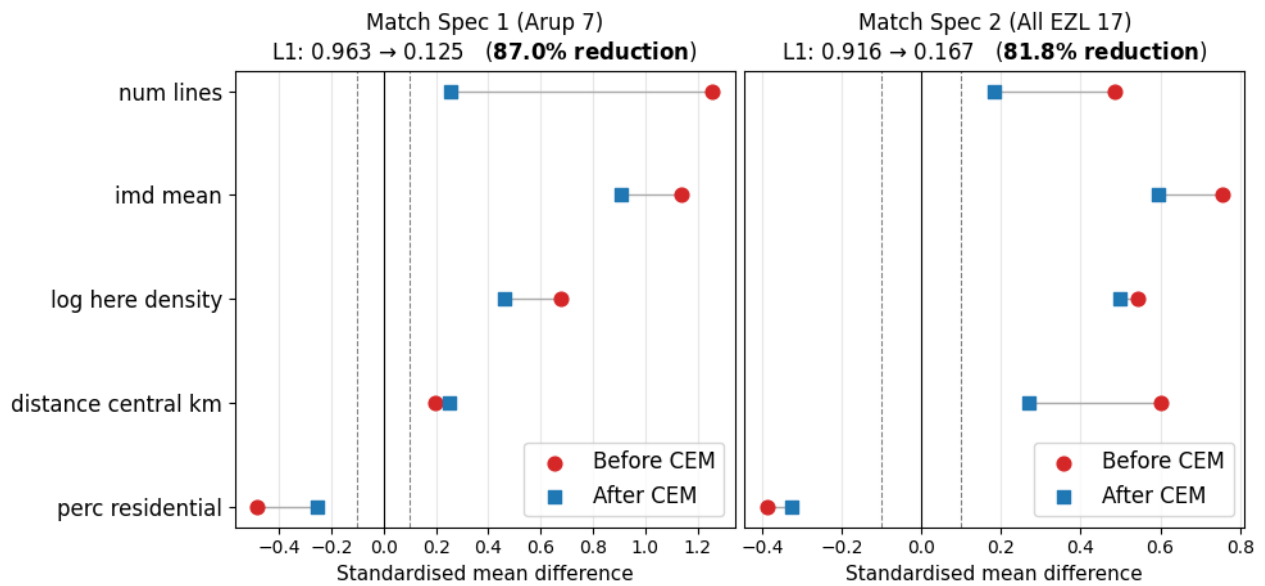


Figure 12: Covariate balance before and after CEM (Love plot)

To assess whether these match outcomes reflect the specific bin-width choices required by CEM rather than a genuine feature of the underlying covariate space, a Mahalanobis $k = 3$ nearest-neighbour matching procedure was independently run over the same covariates. [Annex D](#) contains the results. Both CEM and Mahalanobis matching converged on similar matches including confirming similar structural outliers (e.g. Abbey Wood) as the most challenging treated stations to match given high deprivation and retail density scores.

6.2. Simple Difference-in-Differences (DID) results

[Table 10](#) reports the matched panel's raw pre and post turnover rate by treatment status. The simple difference-in-differences is small and positive for both matched station sets (MS1: +0.35pp; MS2: +0.49pp).

Table 10: Matched panel turnover rate, pre vs post treatment

		Match Spec 1 (Arup 7)		Match Spec 2 (All EZL 17)	
		Pre-Treatment	Post-Treatment	Pre-Treatment	Post-Treatment
Treated	n_eligible	47,965	39,507	64,612	53,324
	Turnover rate	2.64%	2.94%	2.51%	2.79%
Control	n_eligible	49,933	39,194	67,023	53,016
	Turnover rate	3.31%	3.26%	3.28%	3.07%
Simple DID			(+0.35%)		(+0.49%)

[Figure 13](#) plots the quarterly turnover rate for treated and matched control stations across the full 2018Q2–2024Q4 panel. Pre-treatment trends are broadly similar between treated and control groups in both specifications, with comparable volatility and no visible divergence in level ahead of the treatment date.

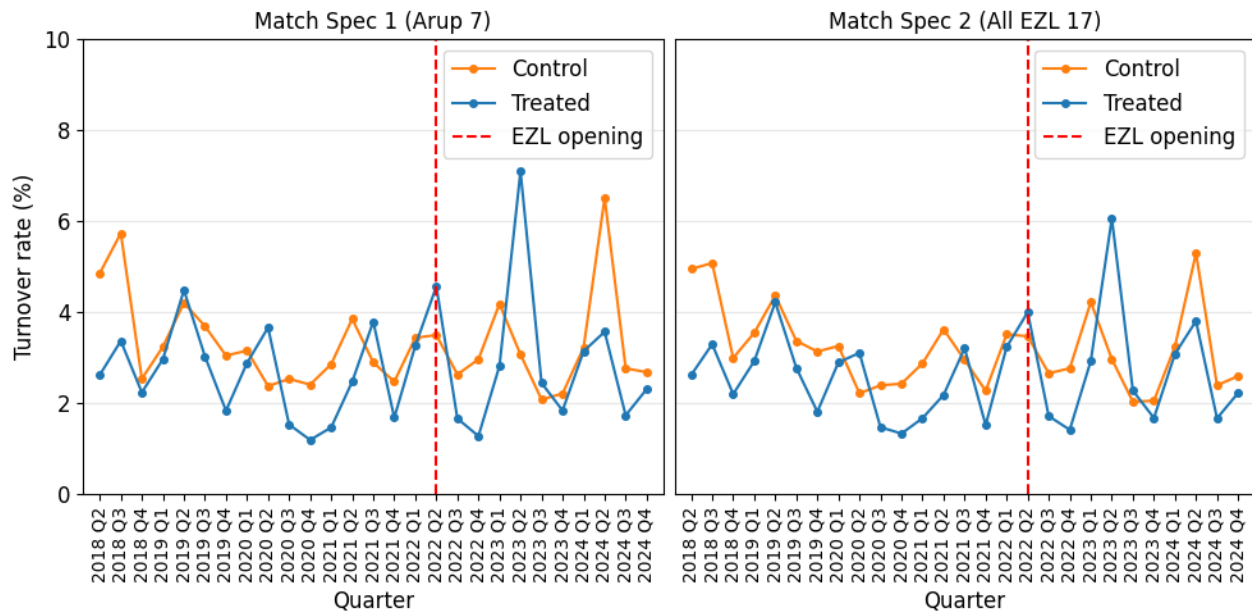


Figure 13: Quarterly turnover rate for treated and matched controls stations

6.3. Difference-in-Differences (DID) regression results

[Table 11](#) presents the difference-in-differences regression estimates of the effect of the Elizabeth line on quarterly retail occupant churn across four specifications and two matching sets. The preferred specification excludes artefact periods and uses a doubly-robust approach to control for remaining pre-treatment covariate imbalance ([Table 11, column 4](#)). It yields a small, positive, and statistically insignificant estimate across both matching sets (MS1: +0.203 percentage points, 95% CI [-0.324, +0.730]; MS2: +0.129 percentage points, 95% CI [-0.256, +0.514]). These intervals rule out turnover effects larger than roughly half a percentage point against a baseline quarterly turnover rate of approximately 2.5%, indicating a reasonably precise null, one that excludes moderate-to-large effects, rather than a merely underpowered estimate.

The only statistically significant coefficient ([Table 11, column 2](#), MS2: +0.449 percentage points, $p < 0.05$) appears in the specification that retains administrative artefacts documented in [Section 4](#). Once these artefact periods are excluded, the effect disappears, indicating that the result is likely due to data spikes rather than a genuine treatment response. Finally, inference is materially unchanged under [Conley \(1999\)](#) standard errors ([Annex E](#)), confirming the null is robust to spatial correlation across catchments.

Table 11: DID regression results of Elizabeth line opening on retail churn

	(1) Base incl. artefact	(2) Doubly Robust incl. artefact	(3) Base excl. artefact	(4) Doubly Robust excl. artefact
Panel A: Matching Specification 1 (Arup 7)				
treat × post (pp)	0.238	0.526	-0.231	0.203
clustered SE	0.322	0.339	0.273	0.244
p	0.473	0.145	0.412	0.420
95% CI (pp)	[-0.458, +0.934]	[-0.207, +1.260]	[-0.821, +0.359]	[-0.324, +0.730]
Observations	4,670	4,670	4,597	4,597
Clusters	14	14	14	14
Panel B: Matching Specification 2 (All EZL 17)				
treat × post (pp)	0.374	0.449	-0.034	0.129
clustered SE	0.271	0.216	0.249	0.189
p	0.177	0.047*	0.839	0.499
95% CI (pp)	[-0.178, +0.927]	[+0.007, +0.890]	[-0.542, +0.474]	[-0.256, +0.514]
Observations	7,699	7,699	7,575	7,575
Clusters	32	32	32	32
LSOA FE	Y	Y	Y	Y
Period FE	Y	Y	Y	Y
Covariates x post		Y		Y
Artefact excl.			Y	Y
CEM Weights	Y	Y	Y	Y

Note: Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors are clustered by station. Inference is materially unchanged under Conley (1999) spatial-HAC standard errors (2 km cutoff; see Annex E). Doubly-robust specifications interact post with distance to city centre, number of lines, log hereditament density, IMD, and per cent residential.

6.4. Event study results

Figure 14 shows the binned event-study plot (≤ -10 to $+10$) for both specifications. Point estimates for the pre-treatment quarters are individually small, and no single pre-treatment coefficient is statistically significant in either specification (Annex F). Joint Wald tests over the full pre-treatment window (Table 12) nonetheless reject the null that all leads are jointly zero (Arup 7: $\chi^2=30.48$, $p=0.0004$; All EZL 17: $\chi^2=27.73$, $p=0.0011$). Two features of this

rejection indicate it does not reflect a systematic differential trend. First, leave-one-out diagnostics ([Annex G](#)) localise the rejection to the three earliest leads, each of which accounts for 28–56% of the statistic. Second, the event-study trajectory shows no monotonic pre-treatment divergence, as would be expected if treated and control catchments were drifting apart on a sustained trend rather than diverging in a few isolated quarters. Because the treatment estimate is materially unchanged when these quarters are excluded, the full-window estimates are retained as the primary specification.

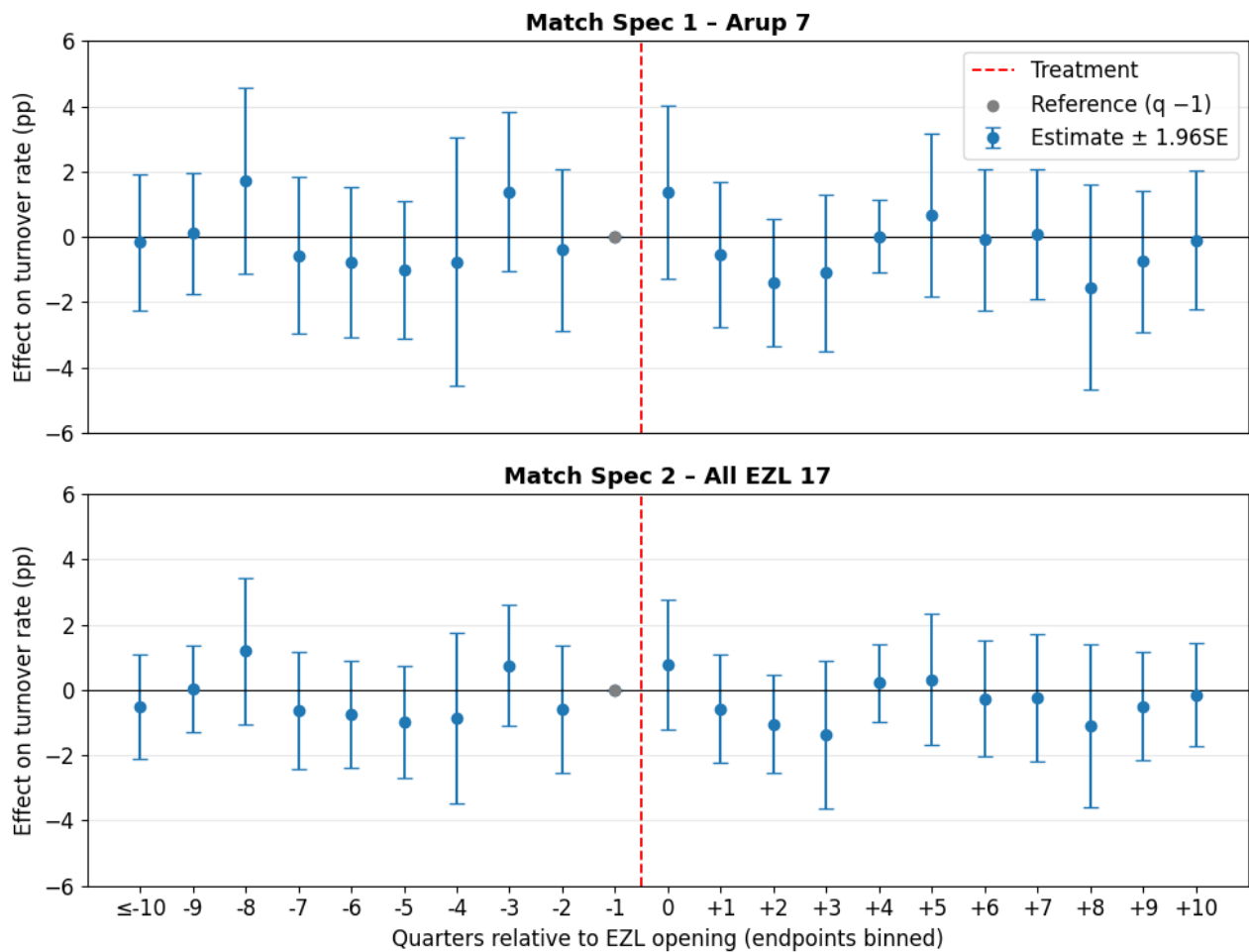


Figure 14: Binned Event-Study Plot (q-10 to q+10)

Table 12: Joint pre-trend Wald tests for Event Study Window

Specification	Leads tested	Window	χ^2	p
Arup 7	9	q-10 to q-2	30.48	0.0004
All EZL 17	9	q-10 to q-2	27.73	0.0011

6.5. Independent-business displacement

To test whether the Elizabeth line displaced independent businesses more rapidly in treated catchments, a difference-in-differences was also estimated on the independent-displacement rate, defined as the share of at-risk independent occupants displaced to a chain or vacancy. The analysis returned a positive and statistically significant coefficient across the selected specification (MS2: +0.24pp, $p=0.004$), suggesting faster independent displacement in treated areas. This apparent effect does not, however, survive scrutiny of the identifying assumption. The event-study specification rejects parallel trends for this outcome (joint pre-trend Wald test $p<0.001$), and the rejection persists after excluding the quarters that accounted for the violation in the turnover analysis. Treated and control catchments were therefore likely already diverging in independent displacement before the Elizabeth line opened. Thus, the coefficient is best understood as the continuation of a pre-existing trend rather than a causal effect. The approach and detailed results are in [Annex H](#).

7. Discussion

7.1. Summary of regression results

The opening of the Elizabeth line had no statistically significant effect on retail occupant churn in station catchments relative to matched controls. This null is consistent across the eight specifications reported in [Table 11](#), across both the representative (MS1) and broad (MS2) treatment sets, and under [Conley \(1999\)](#) standard errors. This affirms a reasonably robust pattern. The sole significant estimate arises only where administrative artefacts are retained and does not survive their exclusion.

Identification rests on parallel trends, which the full-window event study rejects. Leave-one-out diagnostics show the joint rejection is concentrated in the three earliest leads which spans the onset of the COVID-19 pandemic (2020Q1–Q2), when commercial activity contracted sharply and unevenly. Restricting the estimation window to exclude these quarters, the joint pre-trend test no longer rejects in either specification (MS1: $\chi^2=11.99$, $p=0.062$; MS2: $\chi^2=6.21$, $p=0.400$). Point estimates on this clean window remain small and insignificant and are materially unchanged from the full-window estimates. The central finding is therefore a precisely estimated null. Confidence intervals in the preferred specification exclude churn effects larger than roughly half a percentage point against a baseline quarterly rate of approximately 2.5%.

7.2. Interpreting regression findings

The observed net null effect is consistent with the theoretical account in [Section 2](#), where transit-induced commercial change can be understood as the outcome of opposing forces. While some mechanisms accelerate churn by displacing incumbents through rising rents and intensified competition, others could suppress or delay it by bolstering incumbent viability through higher footfall and revenues. Where accelerating and stabilising pressures roughly offset over the observation window, a net effect near zero can be expected.

Study findings are similar to [Liu and Bardaka \(2023\)](#) who find limited evidence of accelerated turnover in station catchments. However, findings diverge from other studies that detect significant effects, once distance is disaggregated. [Champagne, Dubé and](#)

[Legros \(2024\)](#) find increased survival probability for retail firms within 500–750m and 1,000–1,250m of new Montréal stations post-opening, though effects for other service sub-sectors were smaller or less robust. In contrast, [Lin and Yang \(2019\)](#) find that in Taipei, commercial gentrification is concentrated within a few hundred metres of transit stations, ranging from around 240m in the short run, and widening to 300m over time. Both findings suggest that transit's commercial effects, where they exist, tend to be spatially concentrated rather than uniform across a catchment. Given the pooled catchment-wide estimate used in this research design, a null or attenuated aggregate finding is a plausible result which could reflect either a genuine absence of effect or offsetting effects that vary by distance and cancel out in aggregate.

With the short post-treatment window of around two and a half years, the regression findings should also be interpreted as a short-run rather than long-run null. [Meltzer \(2016\)](#) notes that commercial leases typically run longer than residential ones which could shield existing tenants from rent-driven displacement in the short run. [Liu and Bardaka \(2025\)](#) further find that impacts on business survival, such as reduced closure risk for specialty retail and food services, typically only manifest during the medium-term operation phase, which they define as five to ten years post-opening. Consequently, the null finding should be read as the absence of a detectable short-term impact, leaving longer-term outcomes for future research.

Two patterns explain how this null could coexist with the intuition that Elizabeth line catchments experienced visible churn.

- First, quarterly turnover rates across London rose broadly over the study period, reflecting several pressures on physical retail. Online retail's share of London sales rose from 11% in 2014 to 26% by 2024 ([ONS, 2026](#)), and pandemic impacts on London's high streets were spatially uneven, with some centres benefiting from home-working while others lost anchor stores ([GLA, 2025](#)).
- Second, treated catchments recorded a higher absolute volume of turnover events post-opening, but also faster growth in their stock of units ([Figure 15](#)), so the rate was not correspondingly elevated. However, this latter pattern cannot be fully attributed to the opening, as the treated catchments' unit base was already growing faster than controls before May 2022 ([Figure 16](#)). As such, the higher post-opening volume can

be taken as a continuation of a pre-existing divergence rather than marking a treatment effect.

Taken together, the analysis quantifies a null net effect on aggregate churn with reasonable precision but is less equipped to detect compositional change beneath a stable aggregate rate, and delayed effects beyond the observation window.

The limitation with parallel trends suggests a broader identification challenge common to studies of purposively-routed transit infrastructure where new lines are directed toward areas already anticipated to grow ([Pogonyi, Graham and Carbo, 2021](#); [Kandt, Van Dijk and Todd, 2025](#)), making it difficult to fully separate a treatment effect from selection into treatment using covariate-based matching on static characteristics alone ([Baum-Snow and Ferreira, 2015](#)). CEM's covariate set balances treated and control catchments on cross-sectional characteristics but cannot, by construction, balance on pre-treatment growth trajectories unless such variables are explicitly included in the matching procedure. However, adding pre-treatment outcome trends to the matching is not a straightforward remedy, as matching on the pre-treatment path of the outcome itself can induce regression-to-the-mean bias, distorting an otherwise unbiased estimator ([Daw and Hatfield, 2018](#)).

Future research designs could explore the use of synthetic control methods to construct a weighted comparison unit to reproduce the treated area's pre-treatment trajectory directly, targeting pre-trend fit rather than static covariate balance ([Abadie, Diamond and Hainmueller, 2010](#)).

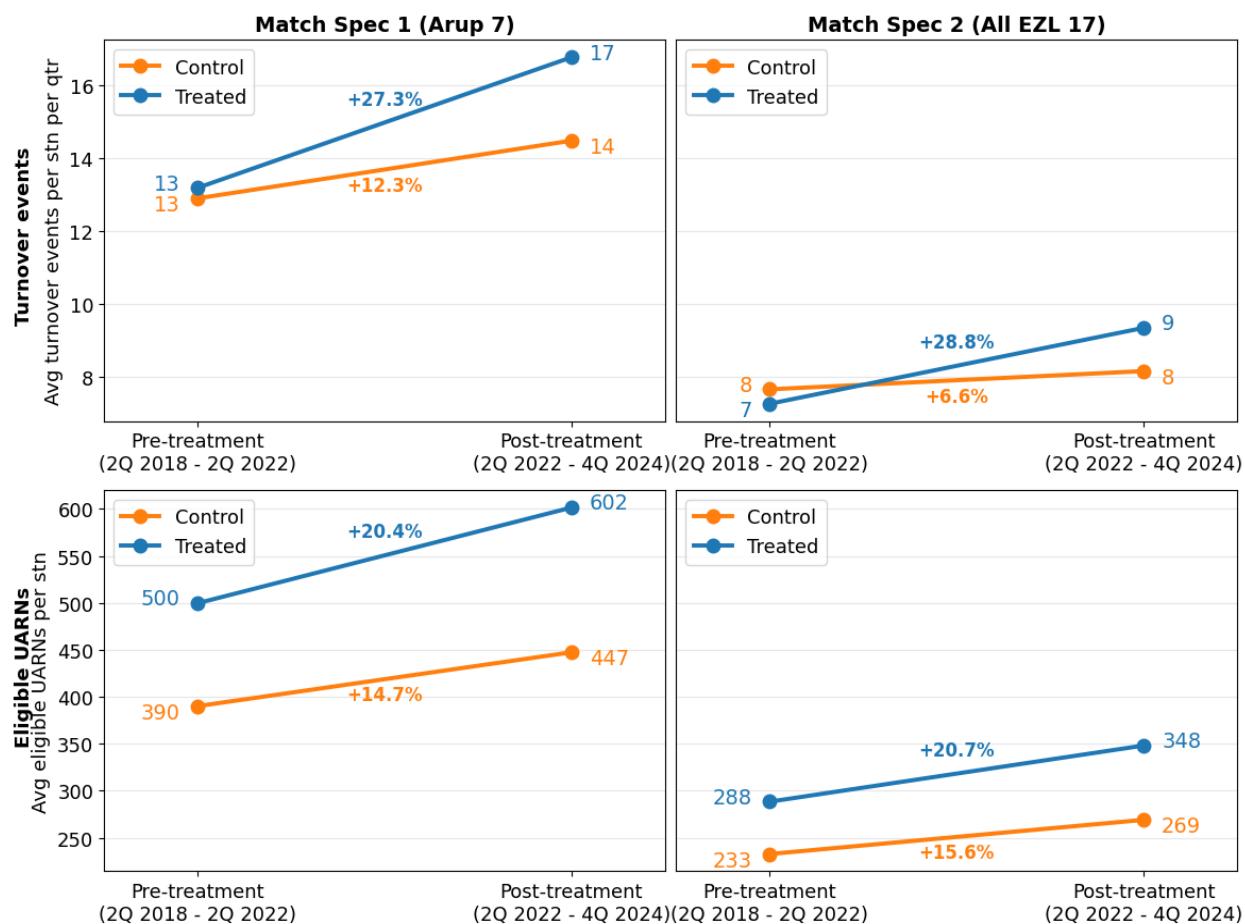


Figure 15: Pre- vs post-treatment growth rates: turnover events and eligible UARNs

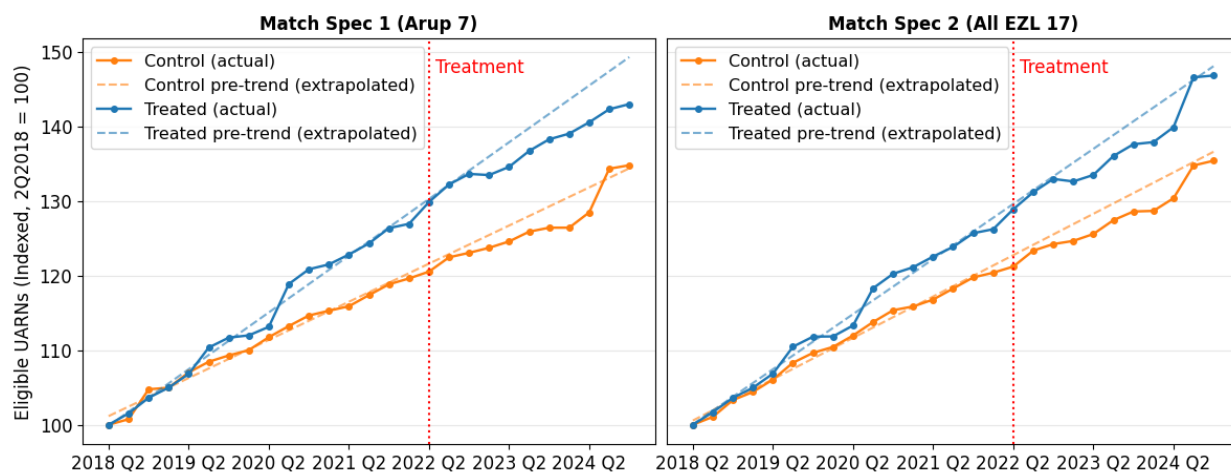


Figure 16: Pre-treatment catchment growth trajectories, extrapolated

7.3. Assessment of VOA OpenLocal ratepayer data

VOA OpenLocal ratepayer data provides a novel, hereditament-level basis for estimating commercial occupant churn, well-suited to localised studies of specific catchments or corridors. However, realising this requires substantial preprocessing and handling of revaluation-cycle boundaries. Data artefacts such as bulk name-contamination events, billing-authority submission freezes, and swap-revert record pairs discovered during the study and must be corrected before turnover can be reliably measured. Data masking and uneven billing-authority record quality limit the dataset's usefulness for London-wide analysis. Overall, OpenLocal may be better suited to targeted, station- or corridor-level studies than to city-wide commercial change analysis.

7.4. Limitations

7.4.1. Data quality and measurement

While OpenLocal data provides a granular view of retail activity at the hereditament level, the effectiveness of the dataset depends on proper reporting. Due to account-name masking and unresolved record-instability issues, 10 of 33 London boroughs were excluded from the panel. Overall variability in turnover in the remaining sample was also shaped by billing-authority specific administrative practices, such as periodic batch updates, that were only partially correctable without over-smoothing genuine turnover signal. To the extent this residual noise is uncorrelated with treatment status, it should not bias point estimates but adds unexplained variance to the turnover series.

Catchment boundaries were defined using LSOAs rather than raw 15-minute isochrone polygons directly, since the deprivation and land-use covariates used in matching are only available at LSOA resolution. While necessary to make catchment-level covariates computable, LSOA boundaries do not follow the isochrone's shape exactly, so a catchment may include areas just outside the true 15-minute walk or exclude areas just inside it.

Finally, the turnover measure captures when a change in ratepayer name is administratively recorded by the VOA, which may lag the point at which a business vacates, or a new occupant begins trading. This includes instances where a lease runs to its contractual end before a name change is filed, or where administrative processing itself is

delayed. Visual inspection of Hammersmith & Fulham's turnover series suggested a roughly three-year periodicity consistent with typical commercial lease-renewal cycles, implying that the timing of recorded turnover reflects lease structure as much as underlying occupant change.

7.4.2. Sample coverage

Two independent processes reduced the number of treated stations entering the DID estimates.

- First, the borough exclusions disproportionately removed high-profile treated catchments such as Paddington, Farringdon, and Tottenham Court Road, which were also areas plausibly expected to show more pronounced turnover effects. 6 of the 10 excluded boroughs were in Inner London, which additionally reduced the pool of suitable inner-London controls for the stations that remained in scope.
- Second, because Elizabeth line stations generally sit in areas of higher pre-existing retail density and deprivation than the broader London control pool, Coarsened Exact Matching (CEM) could not find a comparable control for every treated station that survived the first exclusion. Some treated stations with unusually high density and deprivation had no equivalent stratum among London's control stations. This produced a match rate of 6 of 7 stations in the narrower specification and 14 of 17 in the wider, with common support absent for Abbey Wood, West Drayton, and West Ealing.

These reductions mean that the treatment effect estimated describes a subset of EZL catchments skewed away from the most data-compromised boroughs and the most atypical (highest-density, highest-deprivation) stations.

7.4.3. Identification and confounding

The Elizabeth line's phased rollout, including early TfL Rail surface services, predates the central tunnel opening in May 2022, complicating the assignment of a single treatment date. However, rather than using a staggered framework, this study defines the treatment event in May 2022, as prior incremental phases did not deliver the step-change in cross-

London connectivity offered by the full route. Additionally, OpenLocal data constraints prior to 2016 precluded analysing pre-treatment trends around the initial announcement.

Three further sources of confounding proved difficult to fully isolate. First, rail operates as an integrated network, and the Elizabeth line was partly designed to relieve central London congestion; its opening inevitably redistributed passenger flows across adjacent lines in ways that cannot be fully separated from its direct commercial impact. Second, rail investment often coincides with wider regeneration. While catchments overlapping London Plan Opportunity Areas were excluded where possible, given the spatial concentration of regeneration in corridors also served by the line, residual co-investment confounding cannot be ruled out entirely. Third, smaller concurrent network changes, such as incremental Thameslink frequency increases, may have affected some control catchments. This study excludes major line openings likely to have material commercial effects, treating smaller incremental changes as unavoidable background noise.

A fourth confounder is COVID-19. The study window spans the pandemic disruption, which had an uneven commercial impact across London. If COVID-19 affected Elizabeth line catchments differently from control catchments, which could be plausible given their higher commercial density and commuter orientation, this could bias the estimated treatment effect even after the inclusion of time fixed effects.

7.4.4. Economic inference and statistical power

The observation window, concluding in the last quarter of 2024, presents a limited post-treatment horizon. The literature emphasises that commercial responses to transit shocks are not instantaneous. Retail leases, which often span three to five years, act as a buffer, temporarily shielding incumbent firms from escalating market rents ([Meltzer, 2016](#); [Shepherd Commercial, 2023](#)). With the noise in quarterly turnover reporting, which widens the standard errors, this short window means that while the analysis can rule out effects larger than roughly half a percentage point, it may lack the power to detect a genuinely small one.

A further limitation of the research design is its inability to estimate distance-decay effects. The DID specification treats the 15-minute walking catchment as a single binary unit, coding all hereditaments identically regardless of proximity to the station entrance. If commercial gains are concentrated near station entrances, averaging across the entire

catchment risks diluting localized effects toward statistical insignificance. Partitioning catchments into concentric distance bands could theoretically capture spatial decay ([Pogonyi, Graham and Carbo, 2021](#); [Wexler and Fan, 2022](#)) but introduces other econometric trade-offs. Finer bands reduce sample sizes per unit, eroding statistical power already constrained by the short post-treatment window. For example, ([Pogonyi, Graham and Carbo, 2021](#)) acknowledged that using many distance-based instrumental variable (IV) bands can lead to "imprecise estimation" due to the limited number of observations in each band. Expanding outer boundaries to compensate increases potential overlap with neighbouring rail catchments, contaminating potential controls. Finally, spatial differencing using outer rings as internal controls is potentially problematic, as these areas may follow different underlying growth trajectories, violating parallel trends ([Graham, 2025](#)). This also risks absorbing displacement or spillover effects, violating the clean control assumption ([Baum-Snow and Ferreira, 2015](#)).

7.5. Future research

Three extensions follow from the limitations discussed. First, future work could consider extending the current two-way fixed effects (TWFE) specification by formalizing the hierarchical, spatial structure of the panel within a mixed-effects (multi-level) framework. While the TWFE approach successfully controls for time-invariant omitted variable bias, it does so by absorbing group-level variance into static entity dummy variables ([McNeish and Kelley, 2019](#)). A mixed-effects specification which would entail nesting LSOAs within Local Authorities (boroughs), could allow the researcher to partition and explicitly estimate this variance ([Dong and Harris, 2015](#); [McNeish and Kelley, 2019](#)). This is particularly valuable given that London boroughs act as independent rating billing authorities. A multilevel model can isolate how much turnover variance is driven by localized LSOA-level economic activity versus administrative or reporting variations at the borough level. Furthermore, because transit catchments may not strictly align with administrative lines, borough and station groupings can be crossed rather than nested. A mixed-effects framework can accommodate these crossed random effects, offering a computationally efficient way to model complex, overlapping spatial hierarchies without the severe degree-of-freedom loss or collinearity issues inherent in multi-way fixed effects ([McNeish and Kelley, 2019](#); [Graham, 2025](#)).

Second, to address the catchment-definition limitations noted in [Section 7.4.1](#) and [Section 7.4.4](#), future research could replace LSOA proxies with raw isochrones to isolate hereditaments within the 15-minute walking zone more accurately. Because VOA data are recorded at the hereditament level, researchers can compute exact network walking distances to station entrances. This would enable replacing the binary treatment indicator with a continuous distance metric or partitioning catchments into concentric bands, directly testing distance-decay effects. However, such an approach must balance gains in spatial resolution against reduced statistical power per band while carefully mitigating spatial contamination risks ([Baum-Snow and Ferreira, 2015](#); [Graham, 2025](#)).

Third, future studies could incorporate a longer post-treatment time horizon. A panel extending further past 2024Q4 could be better positioned to capture treatment effects that take time to materialise, rather than the short-run window available here.

7.6. Implications for policy and planning

The finding of a null short-run effect offers reassurance to local authorities that major rail investments do not inevitably trigger immediate commercial displacement or abrupt market destabilisation. Nonetheless, this study does not preclude possible effects in the longer term as commercial pressures from rising land values and rents intensify. Consequently, local authorities should establish robust, long-term monitoring mechanisms to track catchment-level business churn and tenant composition over the longer term. Station masterplans, regeneration frameworks, and business support policies can then be more effectively directed to safeguard commercial diversity before structural displacement occurs.

Valuation Office Agency (VOA) OpenLocal ratepayer data provides transport authorities with a high-resolution asset for scheme evaluation, provided its operational constraints are recognised. Its primary strength lies in its hereditament-level granularity, which allows evaluators to construct exact walk-sheds and parcel-level catchments, reducing the geographic aggregation bias inherent in fixed administrative boundaries. Furthermore, tracking individual hereditaments offers critical visibility into retail churn dynamics, such as distinguishing whether displacement affects independent versus chain businesses.

However, potential users must treat OpenLocal as a complex analytical dataset requiring substantial cleaning and preprocessing prior to use. Because the dataset relies on reporting from individual local authorities, it inherently suffers from administrative processing lags and cross-jurisdictional variations in data quality. Consequently, potential users may find it more effective to scope analyses toward targeted, localized corridors, while advocating for standardised reporting practices to establish consistent submission timeframes across local authorities.

8. Conclusion

This study evaluated whether the May 2022 opening of the Elizabeth line's central section accelerated retail occupant churn in station catchments, while assessing the capacity of administrative ratepayer data to support such evaluations. Utilizing a difference-in-differences design with Coarsened Exact Matching on a novel establishment-level panel derived from VOA OpenLocal records, the analysis yields no statistically significant short-run effect on aggregate retail churn. The null result is precisely estimated, ruling out moderate-to-large impacts. It remains robust across alternative matching estimators, treatment definitions, and spatially clustered standard errors.

This absence of a net effect is not evidence that the Elizabeth line left its commercial surroundings unchanged. Rather, it is consistent with the literature, which suggests that transit investments may exert countervailing forces on surrounding retail establishments, with displacement pressures on some, set against stabilizing gains for others. Over the short horizon, these countervailing forces may broadly offset producing the observed null.

The study demonstrates the potential benefits and constraints of hereditament-level ratepayer data for commercial-change evaluation. Its granularity supports precise, catchment-level measurement that administrative-boundary data cannot. However, inconsistencies in data reporting and revaluation discontinuities demand extensive preprocessing and make it better suited to targeted studies.

These findings reassure policymakers that major rail infrastructure investments do not trigger large, immediate swings in short-run commercial restructuring. However, post-implementation studies will need to consider an appropriate horizon to account for possible longer-term structural adjustments. Future studies constrained to the near-to-medium term should look instead to capture compositional change within catchments, such as shifts in the mix of independent and chain occupants, or in the types of retail present. Such shifts may register the transport line's influence well before any net change in the overall rate of churn becomes detectable and could enable targeted near-term support measures.

References

- Abadie, A., Diamond, A. and Hainmueller, J. (2010) “Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California’s Tobacco Control Program,” *Journal of the American Statistical Association*, 105(490), pp. 493–505. Available at: <https://doi.org/10.1198/jasa.2009.ap08746>.
- AECOM (2018) *Thameslink Programme Evaluation: Baseline Report*. Available at: <https://assets.publishing.service.gov.uk/media/5aa8038240f0b66b5fb4b8a0/thameslink-programme-baseline-evaluation-report.pdf>.
- Angrist, J.D. and Pischke, J.-S. (2009) “Parallel Worlds: Fixed Effects, Differences-in-Differences, and Panel Data,” *Mostly Harmless Econometrics*. Princeton University Press (An Empiricist’s Companion), pp. 221–248. Available at: <https://doi.org/10.2307/j.ctvcm4j72.12>.
- Arup and Department for Transport (2025a) *Elizabeth line post-opening evaluation - Appendix: Transformational Impacts Assessment*. Available at: <https://content.tfl.gov.uk/elizabeth-line-post-opening-evaluation-transformational-impacts-appendix.pdf>.
- Arup and Department for Transport (2025b) *Elizabeth line post-opening evaluation - Case study area profiles appendix*. Available at: <https://content.tfl.gov.uk/elizabeth-line-post-opening-evaluation-case-study-area-profiles-appendix.pdf>.
- Arup and Department for Transport (2025c) *Elizabeth line post-opening evaluation - Full Report*. Available at: <https://content.tfl.gov.uk/elizabeth-line-post-opening-evaluation-full-report.pdf>.
- Banister, D. and Berechman, Y. (2001) “Transport investment and the promotion of economic growth,” *Journal of Transport Geography*, 9(3), pp. 209–218. Available at: [https://doi.org/10.1016/S0966-6923\(01\)00013-8](https://doi.org/10.1016/S0966-6923(01)00013-8).
- Bardaka, E. (2024) “Transit-induced gentrification and displacement: future directions in research and practice,” *Transport Reviews*, 44(3), pp. 567–571. Available at: <https://doi.org/10.1080/01441647.2023.2282285>.
- Bardaka, E., Delgado, M.S. and Florax, R.J.G.M. (2018) “Causal identification of transit-induced gentrification and spatial spillover effects: The case of the Denver light rail,” *Journal of Transport Geography*, 71, pp. 15–31. Available at: <https://doi.org/10.1016/j.jtrangeo.2018.06.025>.
- Baum-Snow, N. and Ferreira, F. (2015) “Causal Inference in Urban and Regional Economics,” *Handbook of Regional and Urban Economics*. Elsevier, pp. 3–68. Available at: <https://doi.org/10.1016/B978-0-444-59517-1.00001-5>.
- Benson, M. and Robinson, T. (2024) *Future transport infrastructure projects and the Elizabeth Line*. Available at: <https://commonslibrary.parliament.uk/research-briefings/cdp-2024-0146/> (Accessed: July 22, 2026).

- Blackwell, M. et al. (2009) "Cem: Coarsened Exact Matching in Stata," *The Stata Journal*, 9(4), pp. 524–546. Available at: <https://doi.org/10.1177/1536867X0900900402>.
- Callaway, B. and Sant'Anna, P.H.C. (2021) "Difference-in-Differences with multiple time periods," *Journal of Econometrics*, 225(2), pp. 200–230. Available at: <https://doi.org/10.1016/j.jeconom.2020.12.001>.
- Cervero, R. and Duncan, M. (2002) *Transit's Value-Added Effects: Light and Commuter Rail Services and Commercial Land Values*. Available at: <https://doi.org/10.3141/1805-02>.
- Champagne, M.-P., Dubé, J. and Legros, D. (2024) "Standing strong? The causal impact of metro stations on service firms' survival," *Transportation Research Part A: Policy and Practice*, 181, p. 103994. Available at: <https://doi.org/10.1016/j.tra.2024.103994>.
- Chapple, K. et al. (2017) *Transit-oriented development & commercial gentrification: exploring the linkages*. Available at: <https://dot.ca.gov/-/media/dot-media/programs/research-innovation-system-information/documents/final-reports/ca17-2833-finalreport-a11y.pdf> (Accessed: July 25, 2026).
- Chapple, K. and Loukaitou-Sideris, A. (2019) "Commercial Gentrification and Displacement," *Transit-Oriented Displacement or Community Dividends?: Understanding the Effects of Smarter Growth on Communities*. The MIT Press. Available at: <https://doi.org/10.7551/mitpress/11300.001.0001>.
- Conley, T.G. (1999) "GMM estimation with cross sectional dependence," *Journal of Econometrics*, 92(1), pp. 1–45. Available at: [https://doi.org/10.1016/S0304-4076\(98\)00084-0](https://doi.org/10.1016/S0304-4076(98)00084-0).
- Credit (2018) "Transit-oriented economic development: The impact of light rail on new business starts in the Phoenix, AZ Region, USA," *Urban Studies*, Vol. 55(No. 13), pp. 2838–2862 (25 pages) <https://www-jstor-org.libproxy.ucl.ac.uk/stable/26510440>.
- Daw, J.R. and Hatfield, L.A. (2018) "Matching and Regression to the Mean in Difference-in-Differences Analysis," *Health Services Research*, 53(6), pp. 4138–4156. Available at: <https://doi.org/10.1111/1475-6773.12993>.
- Debrezion, G., Pels, E. and 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), pp. 161–180. Available at: <https://doi.org/10.1007/s11146-007-9032-z>.
- Delmelle, E.C. (2021) "Transit-induced gentrification and displacement: The state of the debate," *Advances in Transport Policy and Planning*. Elsevier, pp. 173–190. Available at: <https://doi.org/10.1016/bs.atpp.2021.06.005>.
- Deng, K.K. and Li, L. (2024) "Rail transit and the growth of consumer amenities: evidence from Chinese cities," *Transportation*, 51(5), pp. 1611–1634. Available at: <https://doi.org/10.1007/s11116-023-10379-z>.

- Department for Transport (2005) *Economic Appraisal of Crossrail*. Available at: <https://learninglegacy.crossrail.co.uk/wp-content/uploads/2023/04/Economic-Appraisal-of-Crossrail-2005.pdf>.
- Department for Transport (2025) *National Public Transport Access Nodes (NaPTAN) dataset*. Available at: <https://beta-naptan.dft.gov.uk/> (Accessed: July 22, 2026).
- Dong, G. and Harris, R. (2015) “Spatial Autoregressive Models for Geographically Hierarchical Data Structures,” *Geographical Analysis*, 47(2), pp. 173–191. Available at: <https://doi.org/10.1111/gean.12049>.
- Eddington (2006) “The Eddington Transport Study The case for action: Sir Rod Eddington’s advice to Government.” Available at: http://news.bbc.co.uk/1/shared/bsp/hi/pdfs/01_12_06_eddingtonreport.pdf.
- Foster, C.D. and Beesley, M.E. (1963) “Estimating the Social Benefit of Constructing an Underground Railway in London,” *Royal Statistical Society. Journal. Series A: General*, 126(1), pp. 46–78. Available at: <https://doi.org/10.2307/2982446>.
- Gertler, P.J. et al. (2010) *Impact Evaluation in Practice*. The World Bank. Available at: <https://doi.org/10.1596/978-0-8213-8541-8>.
- GLA (2025) *At a crossroads: London’s high streets*. London Assembly Economy, Culture and Skills Committee. Available at: <https://www.london.gov.uk/who-we-are/what-london-assembly-does/london-assembly-work/london-assembly-publications/put-local-communities-heart-high-street-regeneration>.
- Graham, D.J. (2025) “Causal inference for transport research,” *Transportation Research Part A: Policy and Practice*, 192, p. 104324. Available at: <https://doi.org/10.1016/j.tra.2024.104324>.
- Graham, D.J. and Gibbons, S. (2019) “Quantifying Wider Economic Impacts of agglomeration for transport appraisal: Existing evidence and future directions,” *Economics of Transportation*, 19, p. 100121. Available at: <https://doi.org/10.1016/j.ecotra.2019.100121>.
- Hebbert, M. (2014) “Crossrail: the slow route to London’s regional express railway,” *Town Planning Review*, 85(2), pp. 171–190. Available at: <https://doi.org/10.3828/tpr.2014.11>.
- Ho, D.E. et al. (2007) “Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference,” *Political Analysis*, 15(3), pp. 199–236. Available at: <https://doi.org/10.1093/pan/mpi013>.
- Huntington-Klein, N. (2021a) *Chapter 14 - Matching | The Effect*. New York. Available at: <https://theeffectbook.net/ch-Matching.html> (Accessed: July 23, 2026).
- Huntington-Klein, N. (2021b) *Chapter 18 - Difference-in-Differences | The Effect*. Available at: <https://theeffectbook.net/ch-DifferenceinDifference.html> (Accessed: July 23, 2026).

- Iacus, S.M., King, G. and Porro, G. (2012) "Causal Inference without Balance Checking: Coarsened Exact Matching," *Political Analysis*, 20(1), pp. 1–24. Available at: <https://doi.org/10.1093/pan/mpr013>.
- Kahn, M.E. (2007) "Gentrification Trends in New Transit-Oriented Communities: Evidence from 14 Cities That Expanded and Built Rail Transit Systems," *Real Estate Economics*, 35(2), pp. 155–182. Available at: <https://doi.org/10.1111/j.1540-6229.2007.00186.x>.
- Kandt, J., Van Dijk, J. and Todd, J. (2025) "How do major public transport investments induce gentrification? Evidence from consumer data," *Transportation Research Part D: Transport and Environment*, 142, p. 104684. Available at: <https://doi.org/10.1016/j.trd.2025.104684>.
- Kim, S. and Byun, J. (2021) "Identifying spatiotemporally-varying effects of a newly built subway line on land price: Difference and correlation between commercial and residential uses," *International Journal of Sustainable Transportation*, 15(5), pp. 364–374. Available at: <https://doi.org/10.1080/15568318.2020.1789248>.
- King, G. and Nielsen, R. (2019) "Why Propensity Scores Should Not Be Used for Matching," *Political Analysis*, 27(4), pp. 435–454. Available at: <https://doi.org/10.1017/pan.2019.11>.
- Knowles, R. and Ferbrache, F. (2016) "Evaluation of wider economic impacts of light rail investment on cities," *Journal of Transport Geography*, 54, pp. 430–439. Available at: <https://doi.org/10.1016/j.jtrangeo.2015.09.002>.
- Ko, K. and Cao, X. (2013) "The Impact of Hiawatha Light Rail on Commercial and Industrial Property Values in Minneapolis," *Journal of Public Transportation*, 16(1), pp. 47–66. Available at: <https://doi.org/10.5038/2375-0901.16.1.3>.
- Lee, V. and van Dijk, J. (2026) "Connecting or dividing London? Evaluating the impact of the Elizabeth Line on job accessibility and spatial equity," *Case Studies on Transport Policy*, 24, p. 101768. Available at: <https://doi.org/10.1016/j.cstp.2026.101768>.
- Lei, C. (2026) "Differentiated neighborhood change pathways: A comparative study of London's Northern Line Extension and Barking Riverside Extension," *Journal of Transport Geography*, 133, p. 104621. Available at: <https://doi.org/10.1016/j.jtrangeo.2026.104621>.
- Lin, J.-J. and Yang, S.-H. (2019) "Proximity to metro stations and commercial gentrification," *Transport Policy*, 77, pp. 79–89. Available at: <https://doi.org/10.1016/j.tranpol.2019.03.003>.
- Liu, C. and Bardaka, E. (2023) "Transit-induced commercial gentrification: Causal inference through a difference-in-differences analysis of business microdata," *Transportation Research Part A: Policy and Practice*, 175, p. 103758. Available at: <https://doi.org/10.1016/j.tra.2023.103758>.
- Liu, C. and Bardaka, E. (2025) "Exploring the effects of urban rail on the survival of nearby businesses," *Transport Policy*, 169, pp. 209–226. Available at: <https://doi.org/10.1016/j.tranpol.2025.04.011>.

- McNeish, D. and Kelley, K. (2019) “Fixed effects models versus mixed effects models for clustered data: Reviewing the approaches, disentangling the differences, and making recommendations,” *Psychological Methods*, 24(1), pp. 20–35. Available at: <https://doi.org/10.1037/met0000182>.
- Medda, F. (2012) “Land value capture finance for transport accessibility: a review,” *Journal of Transport Geography*, 25, pp. 154–161. Available at: <https://doi.org/10.1016/j.jtrangeo.2012.07.013>.
- Meltzer, R. (2016) “Gentrification and Small Business: Threat or Opportunity?,” *Cityscape*, 18(3), pp. 57–86. Available at: <https://www-jstor-org.libproxy.ucl.ac.uk/stable/26328273>
- Meltzer, R. and Capperis, S. (2016) “Neighbourhood differences in retail turnover: Evidence from New York City,” *Urban Studies*, 54(13). Available at: <https://doi.org/10.1177/0042098016661268>.
- MHCLG (2015) *English indices of deprivation 2015*, GOV.UK. Ministry of Housing, Communities and Local Government. Available at: <https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015> (Accessed: July 22, 2026).
- ONS (2017) *Lower layer Super Output Areas (December 2011) Boundaries EW*. Office for National Statistics licensed under the Open Government Licence v.3.0. Available at: <https://geoportal.statistics.gov.uk/datasets/ons::lower-layer-super-output-areas-december-2011-boundaries-ew-bfc-v3/about> (Accessed: July 22, 2026).
- ONS (2026) *Retail sales, Great Britain - Office for National Statistics*. Available at: <https://www.ons.gov.uk/businessindustryandtrade/retailindustry/bulletins/retailsales/may2026> (Accessed: July 30, 2026).
- OSM (2026) *OpenStreetMap Wiki - landuse*. Available at: <https://wiki.openstreetmap.org/wiki/Key:landuse> (Accessed: July 22, 2026).
- Padeiro, M., Louro, A. and da Costa, N.M. (2019) “Transit-oriented development and gentrification: a systematic review,” *Transport Reviews*, 39(6), pp. 733–754. Available at: <https://doi.org/10.1080/01441647.2019.1649316>.
- Pogonyi, C.G., Graham, D.J. and Carbo, J.M. (2021) “Metros, agglomeration and displacement. Evidence from London,” *Regional Science and Urban Economics*, 90, p. 103681. Available at: <https://doi.org/10.1016/j.regsciurbeco.2021.103681>.
- PwC (2023) *Retail Store Openings and Closures 2023*. Available at: <https://www.pwc.co.uk/industries/retail-consumer/insights/store-openings-and-closures-2023.html>.
- Ray, R.S. (2017) “Open for business? Effects of Los Angeles Metro Rail construction on adjacent businesses,” *Journal of Transport and Land Use*, 10(1). Available at: <https://doi.org/10.5198/jtlu.2017.932>.

- Retail Insight Network (2018) *Westfield London expansion: Location, design and facilities*. Available at: <https://www.retail-insight-network.com/projects/westfield-london-expansion/> (Accessed: July 23, 2026).
- Rosenbaum, P.R. and Rubin, D.B. (1983) "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70(1), pp. 41–55. Available at: <https://doi.org/10.2307/2335942>.
- Roth, J. et al. (2023) "What's trending in difference-in-differences? A synthesis of the recent econometrics literature," *Journal of Econometrics*, 235(2), pp. 2218–2244. Available at: <https://doi.org/10.1016/j.jeconom.2023.03.008>.
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688–701. <https://doi.org/10.1037/h0037350>
- SACTRA (1999) *Standing Advisory Committee on Trunk Road Assessment (SACTRA)*. Department of the Environment, Transport and the Regions (DETR). Available at: https://www.ffue.org/wp-content/uploads/2016/08/SACTRA_Full-report.pdf.
- Schmidheiny, K. and Siegloch, S. (2023) "On event studies and distributed-lags in two-way fixed effects models: Identification, equivalence, and generalization," *Journal of Applied Econometrics*, 38(5), pp. 695–713. Available at: <https://doi.org/10.1002/jae.2971>.
- Schuetz, J. (2015) "Do rail transit stations encourage neighbourhood retail activity?," *Urban Studies*, 52(14), pp. 2699–2723. Available at: <https://www-jstor-org.libproxy.ucl.ac.uk/stable/26146188>
- Seo, K. et al. (2019) "Hedonic modeling of commercial property values: distance decay from the links and nodes of rail and highway infrastructure," *Transportation*, 46(3), pp. 859–882. Available at: <https://doi.org/10.1007/s11116-018-9861-z>.
- Shepherd Commercial (2023) "What is the average lease term for a commercial property? - Shepherd Commercial," 28 February. Available at: <https://shepcom.com/averageleaselengthforcommercialproperty/> (Accessed: July 25, 2026).
- Smith, N. (1979) "Toward a Theory of Gentrification A Back to the City Movement by Capital, not People," *Journal of the American Planning Association*, 45(4), pp. 538–548. Available at: <https://doi.org/10.1080/01944367908977002>.
- Sukaryavichute, E., Delmelle, E. and Hammelman, C. (2021) "Opportunities and challenges for small businesses in new transit neighborhoods: Understanding impacts through in-depth interviews," *Regional Science Policy & Practice*, 13(3), pp. 1025–1042. Available at: <https://doi.org/10.1111/rsp3.12412>.
- Sun, L. and Abraham, S. (2021) "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of Econometrics*, 225(2), pp. 175–199. Available at: <https://doi.org/10.1016/j.jeconom.2020.09.006>.

- Tornabene, S. and Nilsson, I. (2021) "Rail transit investments and economic development: Challenges for small businesses," *Journal of Transport Geography*, 94, p. 103087. Available at: <https://doi.org/10.1016/j.jtrangeo.2021.103087>.
- Transport for London (2021) *New Northern line stations open today as Tube extends to Battersea Power Station*, Transport for London. Available at: <https://tfl.gov.uk/info-for/media/press-releases/2021/september/new-northern-line-stations-open-today-as-tube-extends-to-battersea-power-station> (Accessed: July 23, 2026).
- Transport for London (2022a) *Barking Riverside station to open 18 July*, Transport for London. Available at: <https://tfl.gov.uk/info-for/media/press-releases/2022/july/barking-riverside-station-to-open-18-july> (Accessed: July 23, 2026).
- Transport for London (2022b) *Elizabeth line benefits framework*, Transport for London. Transport for London (tfl.gov.uk). Available at: <https://www.tfl.gov.uk/corporate/publications-and-reports/elizabeth-line-benefits-framework> (Accessed: July 27, 2026).
- Transport for London (2022c) *Elizabeth line to open on 24 May 2022*, Transport for London. Available at: <https://tfl.gov.uk/info-for/media/press-releases/2022/may/elizabeth-line-to-open-on-24-may-2022> (Accessed: July 22, 2026).
- Transport for London (2025) *Transport for London GIS Open Data Hub*. Available at: <https://gis-tfl.opendata.arcgis.com/maps/5dcb79b6817b4bf89732de68a0337312> (Accessed: July 22, 2026).
- TravelTime (2025) "Isochrone API Docs | High Performance, Great Value." TravelTime. Available at: <https://docs.traveltime.com/api/overview/isochrones> (Accessed: July 23, 2026).
- Wexler, N. and Fan, Y. (2022) "Transitway investment and nearby commercial gentrification," *Multimodal Transportation*, 1(2), p. 100015. Available at: <https://doi.org/10.1016/j.multra.2022.100015>.
- Whearty, M. Murcio, R. & Joana Barros, J. (2026) The influence of passenger rail interventions in London on relative deprivation in affected station catchment areas, *Urban, Planning and Transport Research*, Available at: <https://doi.org/10.1080/21650020.2026.2612861>
- Whythawk (2026) *OpenLocal Explorer - Business rates for England and Wales*. Available at: <https://openlocal.uk/about> (Accessed: July 22, 2026).
- Wooldridge, J.M. (2010) "Basic Linear Unobserved Effects Panel Data Models," *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, pp. 281–344. Available at: <https://www.jstor.org/stable/j.ctt5hhcfr.14> (Accessed: July 24, 2026).
- Zuk, M. et al. (2018) "Gentrification, Displacement, and the Role of Public Investment," *Journal of Planning Literature*, 33(1), pp. 31–44. Available at: <https://doi.org/10.1177/0885412217716439>.

Zukin, S. *et al.* (2009) "New Retail Capital and Neighborhood Change: Boutiques and Gentrification in New York City," *City & Community*, 8(1), pp. 47–64. Available at: <https://doi.org/10.1111/j.1540-6040.2009.01269.x>.

Annexes

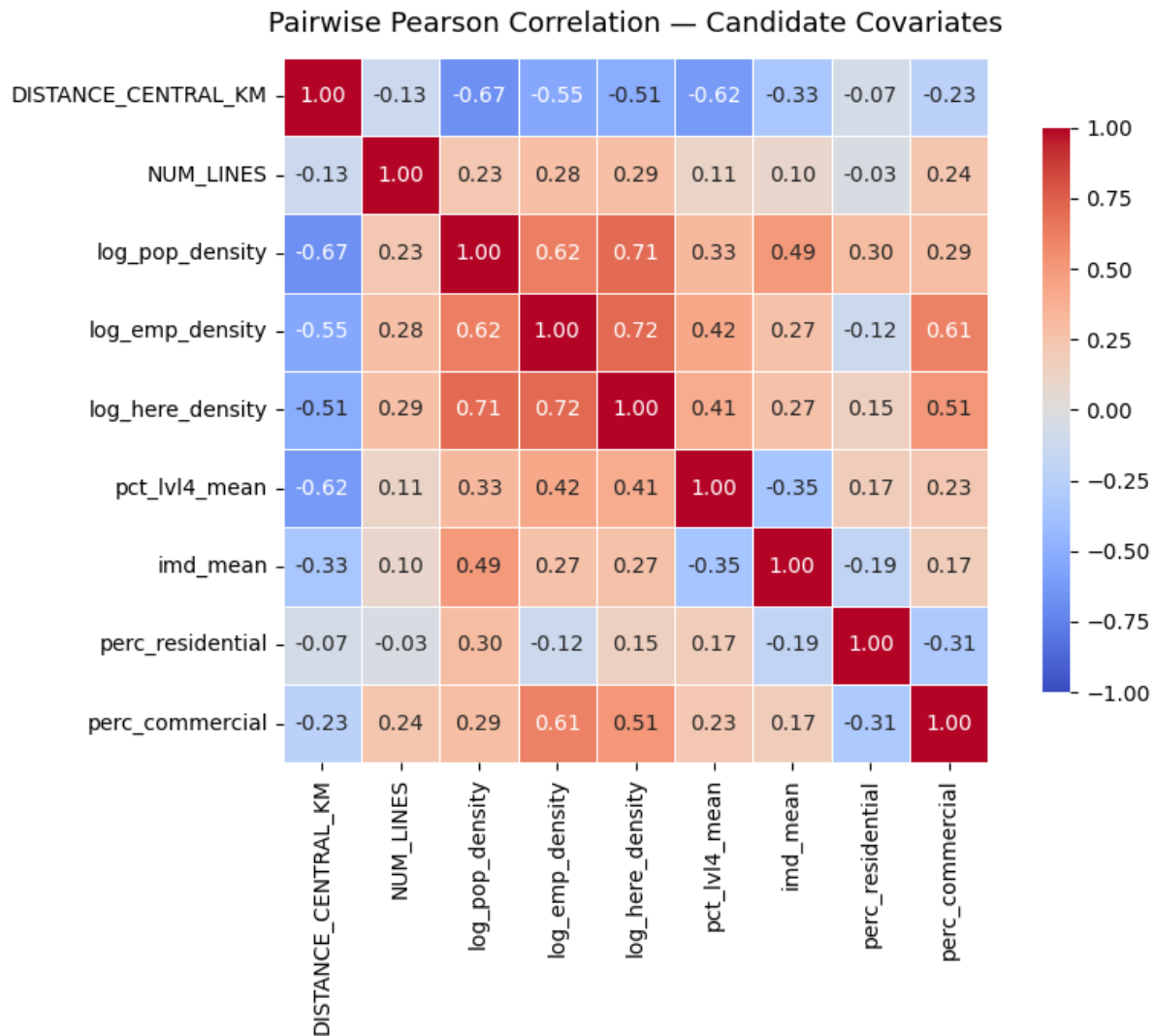
Annex A: Recent studies on impact of transit investments on commercial properties	73
Annex B: Collinearity Diagnostics for CEM covariate candidates	75
Annex C: Full covariate matching outcomes of Coarsened Exact Matching (CEM)	76
Annex D: Mahalanobis nearest neighbour matching results	79
Annex E: DID Regression with Conley Standard Errors.....	82
Annex F: Event Study Coefficients.....	83
Annex G: Leave-one-out analysis of event study specifications.....	85
Annex H: Independent-business displacement – investigation and results	86

Annex A: Studies on impact of transit investments on commercial properties

Study	Location	Method	Rail Type	Data	Notable Finding
<i>Commercial gentrification</i>					
Chapple et al. (2017)	San Francisco Bay and Los Angeles, CA	Probit Regression	Light rail	Business microdata (1990–2013)	Transit proximity was generally not a statistically significant predictor of commercial gentrification in the regional econometric models.
Lin and Yang (2019)	Neihu District, Taipei City, Taiwan	Binary logit model	Metro	Google Streetview (2009–2015)	Probability of business displacement increases as distance to metro stations decreases (with significant influence ranges of ~240m for short-term and ~300m for long-term changes).
Liu and Bardaka (2023)	Charlotte, NC	DID negative binomial regression	Light rail (LYNX Blue Line)	Business microdata (1998–2019)	No strong evidence of transit-induced commercial gentrification was found along the line; instead, businesses near stations experienced lower turnover and fewer exits compared to controls.
Wexler and Fan (2022)	Minneapolis and St. Paul, MN	Local polynomial regressions and event study Difference-in-Differences (DiD)	Light Rail	INFOUSA business survey (2000–2019)	Establishments within 750m were negatively affected post-opening. Single-location firms saw a 13% reduction in sales volume starting two years after operations began, while multi-location firms and employment levels remained unaffected.
<i>Business formation and survival</i>					
Credit (2018)	Phoenix Metropolitan Area, Arizona	Adjusted-Interrupted Time Series (AITS) Poisson regression models	Light rail	Business microdata (1990–2014)	Station opening led to immediate gains in new business formation near stations: +88% in knowledge sector, +40% in service sector, and +28% in retail.
Deng and Li (2024)	6 Chinese Cities	DID regression	Rail Transit	Point of Interest (POI) data (2015–2018)	Opening a new station induced an ~4% increase in the total number of consumer amenities nearby within the first year, an impact that strengthened substantially after two years.

Study	Location	Method	Rail Type	Data	Notable Finding
Ray (2017)	Los Angeles, CA	Quasi-experimental Difference-in-Differences (DID) & Survival Analysis	Metro Rail / Heavy Rail	NETS Business Microdata (1992–2011)	Construction significantly reduced business survival within 400m of stations where intrusive cut-and-cover construction occurred
Tornabene and Nilsson (2021)	Charlotte, NC	Mixed-methods approach	Light rail	Survey data, in-depth interviews	Construction disruption and lack of mitigation strategies caused severe accessibility and visibility losses. This led to customer drop-off and commercial displacement <i>before</i> property values or rent increases could even take effect.
Bardaka et al. (2025)	Charlotte, NC	Quasi-experimental spatial survival/ cox hazard analysis	Light rail	Longitudinal business microdata (1998–2019)	Proximity to light rail and TOD overlay zones substantially reduces the risk of business closure for key commercial sectors, indicating that transit proximity enhances business longevity rather than triggering widespread displacement.
<i>Land value uplift / commercial property values</i>					
Arup / Volterra (2022)	London, UK	DID regression	Commuter Rail	CoStar and JLL commercial rental data	3%–8% pre-opening impact on office rents within 500m–1km of central section stations
Seo et al. (2019)	Phoenix, AZ	Hedonic price model	Light 4ail	Commercial property sales data (2009-2014)	For office properties, retail and service properties, proximity to the light rail stations has a positive association with value, proximity to light rail links has a little or negative relationship with value.
Ko and Cao (2013)	Minneapolis and St. Paul, MN	Hedonic price model	Light rail	Commercial property sales data (2000–2008)	"LRT effect" on property values far-reaching, extending approximately 0.875 miles (1,400 metres) from station sites. Proximity to a station increases property values in a non-linear fashion

Annex B: Collinearity Diagnostics for CEM covariate candidates



Variance Inflation Factor (VIF) Analysis

Final Covariate Set	VIF
log_here_density	1.51
Distance_central_km	1.44
imd_mean	1.21
num_lines	1.10
perc_residential	1.10

Annex C: Full covariate matching outcomes of Coarsened Exact Matching (CEM)

Match Specification 1 (Arup 7)

Role	Station	Distance from City Centre	Number of Lines	Hereditament Density (Log)	Average Deprivation (IMD)	Percentage Residential Land Use	Stratum Cell	CEM Weight
T	CANARY WHARF	7.467	2	5.594	32.30	0.725	0_1_3_2_0	1.000
C	GOLDHAWK ROAD	7.009	2	5.964	25.52	0.712	0_1_3_2_0	0.889
C	WEST BROMPTON	5.338	3	5.672	24.98	0.818	0_1_3_2_0	0.889
C	WEST HAMPSTEAD	6.271	2	5.552	22.69	0.810	0_1_3_2_0	0.889
T	EALING BROADWAY	12.196	3	4.899	17.19	0.813	1_1_2_1_0	1.000
C	GUNNERSBURY	10.505	2	4.738	14.35	0.831	1_1_2_1_0	0.667
C	MAZE HILL	9.381	2	4.870	17.93	0.780	1_1_2_1_0	0.667
T	ILFORD	14.711	1	5.635	32.17	0.823	2_0_3_2_0	1.000
C	EAST CROYDON	14.892	1	5.713	26.08	0.750	2_0_3_2_0	1.333
T	ROMFORD	22.686	2	5.476	24.25	0.763	3_1_2_2_0	1.000
C	UXBRIDGE	24.769	2	4.676	21.24	0.645	3_1_2_2_0	1.333
T	SOUTHALL	17.479	2	5.873	36.28	0.698	2_1_3_2_0	1.000
C	WEST CROYDON	14.473	2	5.811	35.93	0.635	2_1_3_2_0	1.333
T	WHITECHAPEL	4.780	3	6.057	39.46	0.520	0_1_3_2_0	1.000
C	GOLDHAWK ROAD	7.009	2	5.964	25.52	0.712	0_1_3_2_0	0.889
C	WEST BROMPTON	5.338	3	5.672	24.98	0.818	0_1_3_2_0	0.889
C	WEST HAMPSTEAD	6.271	2	5.552	22.69	0.810	0_1_3_2_0	0.889

Match Specification 2 (All EZL 17)

Role	Station	Distance from City Centre	Number of Lines	Hereditament Density (Log)	Average Deprivation (IMD)	Percentage Residential Land Use	Stratum Cell	CEM Weight
T	ACTON MAIN LINE	9.849	1	4.493	26.41	0.63	1_0_1_2_0	1.000
C	CROFTON PARK	8.534	1	4.431	22.66	0.827	1_0_1_2_0	0.643
C	KEW BRIDGE	11.362	1	4.205	24.04	0.675	1_0_1_2_0	0.643
T	CANARY WHARF	7.467	2	5.594	32.3	0.725	0_1_3_2_0	1.000
C	GOLDHAWK ROAD	7.009	2	5.964	25.52	0.712	0_1_3_2_0	0.857
C	WEST BROMPTON	5.338	3	5.672	24.98	0.818	0_1_3_2_0	0.857
C	WEST HAMPSTEAD	6.271	2	5.552	22.69	0.81	0_1_3_2_0	0.857
T	CHADWELL HEATH	18.902	1	5.249	32.46	0.864	2_0_2_2_1	1.000
C	BURNT OAK	14.262	1	4.51	28.48	0.837	2_0_2_2_1	1.286
C	SOUTH CROYDON	16.260	1	4.897	22.91	0.886	2_0_2_2_1	1.286
C	SOUTH HARROW	16.905	1	4.561	20.62	0.84	2_0_2_2_1	1.286
T	EALING BROADWAY	12.196	3	4.899	17.19	0.813	1_1_2_1_0	1.000
C	GUNNERSBURY	10.505	2	4.738	14.35	0.831	1_1_2_1_0	0.643
C	MAZE HILL	9.381	2	4.87	17.93	0.78	1_1_2_1_0	0.643
T	GIDEA PARK	24.479	1	4.036	9.61	0.89	3_0_0_0_1	1.000
C	HATCH END	20.268	1	3.733	7.95	0.863	3_0_0_0_1	0.643
C	NORTHWOOD	23.633	1	3.742	8.58	0.936	3_0_0_0_1	0.643
T	GOODMAYES	17.689	1	5.123	26	0.866	2_0_2_2_1	1.000
C	BURNT OAK	14.262	1	4.51	28.48	0.837	2_0_2_2_1	1.286
C	SOUTH CROYDON	16.26	1	4.897	22.91	0.886	2_0_2_2_1	1.286
C	SOUTH HARROW	16.905	1	4.561	20.62	0.84	2_0_2_2_1	1.286

Role	Station	Distance from City Centre	Number of Lines	Hereditament Density (Log)	Average Deprivation (IMD)	Percentage Residential Land Use	Stratum Cell	CEM Weight
T	HANWELL	14.747	1	4.907	20.85	0.774	2_0_2_2_0	1.000
C	HOUNSLOW CENTRAL	17.164	1	4.965	23.51	0.801	2_0_2_2_0	0.857
C	HOUNSLOW EAST	16.440	1	5.073	21.61	0.807	2_0_2_2_0	0.857
C	SELHURST	13.144	1	4.854	37.08	0.738	2_0_2_2_0	0.857
T	HAYES & HARLINGTON	20.411	2	5.215	27.58	0.569	3_1_2_2_0	1.000
C	UXBRIDGE	24.769	2	4.676	21.24	0.645	3_1_2_2_0	2.571
T	ILFORD	14.711	1	5.635	32.17	0.823	2_0_3_2_0	1.000
C	EAST CROYDON	14.892	1	5.713	26.08	0.750	2_0_3_2_0	1.286
T	ROMFORD	22.686	2	5.476	24.25	0.763	3_1_2_2_0	1.000
C	UXBRIDGE	24.769	2	4.676	21.24	0.645	3_1_2_2_0	2.571
T	SEVEN KINGS	16.669	1	4.928	24.35	0.881	2_0_2_2_1	1.000
C	BURNT OAK	14.262	1	4.510	28.48	0.837	2_0_2_2_1	1.286
C	SOUTH CROYDON	16.26	1	4.897	22.91	0.886	2_0_2_2_1	1.286
C	SOUTH HARROW	16.905	1	4.561	20.62	0.84	2_0_2_2_1	1.286
T	SOUTHALL	17.479	2	5.873	36.28	0.698	2_1_3_2_0	1.000
C	WEST CROYDON	14.473	2	5.811	35.93	0.635	2_1_3_2_0	1.286
T	WHITECHAPEL	4.780	3	6.057	39.46	0.520	0_1_3_2_0	1.000
C	GOLDHAWK ROAD	7.009	2	5.964	25.52	0.712	0_1_3_2_0	0.857
C	WEST BROMPTON	5.338	3	5.672	24.98	0.818	0_1_3_2_0	0.857
C	WEST HAMPSTEAD	6.271	2	5.552	22.69	0.810	0_1_3_2_0	0.857
T	WOOLWICH	13.878	1	5.132	30.69	0.764	2_0_2_2_0	1.000
C	HOUNSLOW CENTRAL	17.164	1	4.965	23.51	0.801	2_0_2_2_0	0.857
C	HOUNSLOW EAST	16.44	1	5.073	21.61	0.807	2_0_2_2_0	0.857
C	SELHURST	13.144	1	4.854	37.08	0.738	2_0_2_2_0	0.857

Annex D: Mahalanobis nearest neighbour matching results

As a robustness check on the matching framework, treatment stations were additionally matched to controls using Mahalanobis-distance nearest-neighbour matching on the five covariates. Each treated station was matched to its $k = 3$ nearest control neighbours without replacement, subject to a distance threshold of 2.0 beyond which matches were discarded to avoid pairing dissimilar units.

Table D1: Summary of Mahalanobis Matching Outcomes

Metric	Match Spec 1M (Arup 7)	Match Spec 2M (All EZL 17)
Treated stations	7	17
Matched	5 (71.4%)	14 (82.4%)
Unmatched	Abbey Wood, Whitechapel	Abbey Wood, West Drayton, West Ealing
Mean SMD Reduction	0.749 → 0.309 (58.7%)	0.554 → 0.279 (49.6%)
Controls	21 (9.8%)	51 (23.7%)

Table D2: Summary of Mahalanobis Matching Outcomes (MS1M)

Treated	1 st Match	2 nd Match	3 rd Match
ILFORD	WALTHAMSTOW QUEENS ROAD (0.864)	EAST CROYDON (1.134)	HOUNSLOW CENTRAL (1.259)
SOUTHALL	WEST CROYDON (0.935)		
EALING BROADWAY	RICHMOND (1.165)	KENSINGTON (OLYMPIA) (1.281)	WEST BROMPTON (1.454)
CANARY WHARF	GOLDHAWK ROAD (1.166)	KENTISH TOWN (1.202)	DEPTFORD (1.211)
ROMFORD	UXBRIDGE (1.526)	RUISLIP (1.955)	
ABBEY WOOD			
WHITECHAPEL			

Note: Figures in parenthesis indicate matched Mahalanobis distance. Cells in grey indicate no suitable matches within distance of 2.000.

Table D3: Summary of Mahalanobis Matching Outcomes (MS2M)

Treated	1st Match	2nd Match	3rd Match
SEVEN KINGS	SOUTH CROYDON 0.218	HOUNSLOW 0.610	HOUNSLOW CENTRAL 0.772
ACTON MAIN LINE	KEW BRIDGE 0.626	COLINDALE 1.032	ALEXANDRA PALACE 1.084
HANWELL	HOUNSLOW EAST 0.636	HIGH BARNET 0.730	EDGWARE 0.766
GIDEA PARK	NORTHWOOD 0.771	HORNCHURCH 0.925	ELM PARK 0.933
WEST EALING	SUTTON 0.789	HARROW-ON-THE-HILL 1.018	CARSHALTON 1.162
WOOLWICH	WOOD GREEN 0.804	WALTHAMSTOW QUEENS ROAD 0.891	TURNPIKE LANE 0.908
SOUTHALL	WEST CROYDON 0.920		
CHADWELL HEATH	RAINHAM 0.979	COULSDON SOUTH 1.634	HOUNSLOW WEST 1.722
WEST DRAYTON	UXBRIDGE 1.073		
GOODMAYES	WALLINGTON 1.137	SOUTH HARROW 1.202	WELLING 1.388
CANARY WHARF	KENTISH TOWN 1.183	GOLDHAWK ROAD 1.189	BLACKHORSE RD 1.213
EALING BROADWAY	RICHMOND 1.185	KENSINGTON (OLYMPIA) 1.270	WEST BROMPTON 1.465
ILFORD	ST JAMES STREET 1.551	SELHURST 1.738	HORNSEY 1.933
ROMFORD	RUISLIP 1.947		
ABBEY WOOD			
HAYES & HARLINGTON			
WHITECHAPEL			

Note: Figures in parenthesis indicate matched Mahalanobis distance. Cells in grey indicate no suitable matches within distance of 2.000.

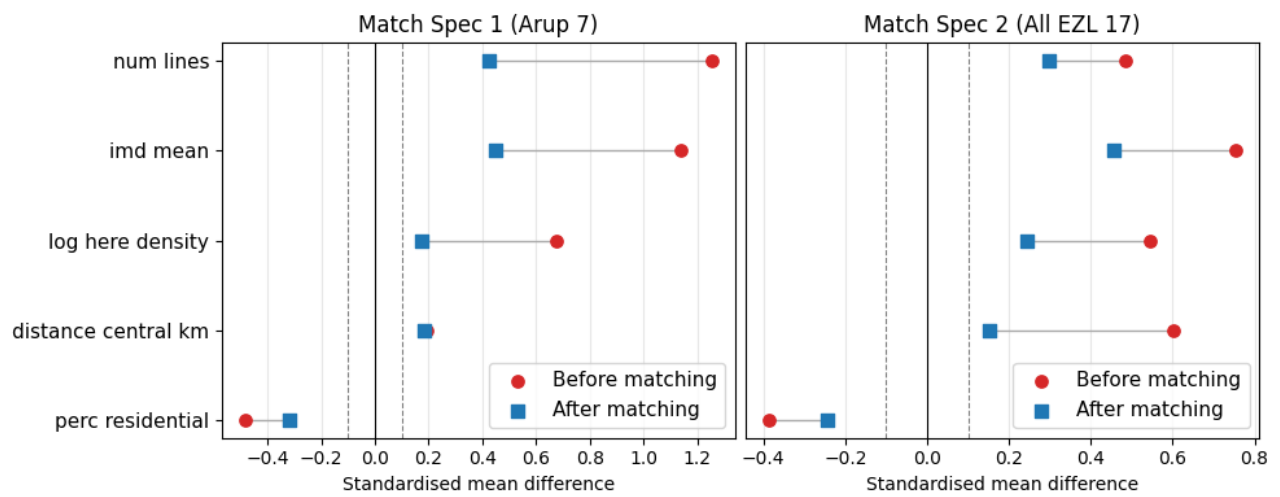


Figure D1: Covariate balance before and after Mahalanobis Matching

Annex E: DID Regression with Conley Standard Errors

	(1) Base incl. artefact	(2) Doubly Robust incl. artefact	(3) Base excl. artefact	(4) Doubly Robust excl. artefact
Panel A: Matching Specification 1 (Arup 7)				
treat × post (pp)	0.238	0.526	-0.231	0.203
Conley SE (2km)	0.308	0.329	0.254	0.238
p	0.439	0.110	0.361	0.394
95% CI (pp)	[-0.365, +0.841]	[-0.118, +1.171]	[-0.728, +0.266]	[-0.264, +0.671]
Observations	4,670	4,670	4,597	4,597
Panel B: Matching Specification 2 (All EZL 17)				
treat × post (pp)	0.374	0.449	-0.034	0.129
Conley SE (2km)	0.284	0.222	0.260	0.193
p	0.187	0.043*	0.896	0.505
95% CI (pp)	[-0.182, +0.930]	[+0.014, +0.884]	[-0.544, +0.476]	[-0.249, +0.507]
Observations	7,699	7,699	7,575	7,575
LSOA FE	Y	Y	Y	Y
Period FE	Y	Y	Y	Y
Covariates x post		Y		
Artefact excluded			Y	Y
CEM Weights	Y	Y	Y	Y

*Note: Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors are clustered using [Conley \(1999\)](#) spatial-HAC standard errors (2 km cutoff). Doubly-robust specifications interact post with distance to city centre, number of lines, log hereditament density, IMD, and per cent residential.*

Annex F: Event Study Coefficients

Table F1: Event Study Coefficients for Matched Specification 1 (Arup 7)

Quarter	Estimate (pp)	Std. Error	t	p	2.5%	97.5%
≤-10	-0.162	1.070	-0.151	0.882	-0.025	0.021
-9	0.103	0.947	0.108	0.915	-0.019	0.021
-8	1.729	1.450	1.192	0.254	-0.014	0.049
-7	-0.569	1.226	-0.464	0.650	-0.032	0.021
-6	-0.784	1.172	-0.669	0.515	-0.033	0.017
-5	-1.005	1.074	-0.936	0.366	-0.033	0.013
-4	-0.760	1.942	-0.391	0.702	-0.050	0.034
-3	1.387	1.244	1.115	0.285	-0.013	0.041
-2	-0.406	1.259	-0.322	0.752	-0.031	0.023
-1	0	-	-	-	-	-
0	1.366	1.358	1.006	0.333	-0.016	0.043
1	-0.538	1.140	-0.472	0.645	-0.030	0.019
2	-1.400	1.002	-1.397	0.186	-0.036	0.008
3	-1.101	1.227	-0.898	0.386	-0.038	0.015
4	0.011	0.570	0.020	0.985	-0.012	0.012
5	0.683	1.275	0.536	0.601	-0.021	0.034
6	-0.088	1.104	-0.079	0.938	-0.025	0.023
7	0.074	1.013	0.073	0.943	-0.021	0.023
8	-1.553	1.601	-0.970	0.350	-0.050	0.019
9	-0.746	1.101	-0.678	0.510	-0.031	0.016
10	-0.107	1.086	-0.098	0.923	-0.025	0.022

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Coefficients are quarter-specific treatment interactions from the dynamic event-study specification, in percentage points. Quarter -1 is the omitted reference period. Endpoint bins (≤ -10 , 10) pool all more distant leads/lags. No individual coefficient is statistically significant at the 5% level.

Table F2: Event study coefficients for Matched Specification 2 (All EZL 17)

Quarter	Estimate (pp)	Std. Error	t	p	2.5%	97.5%
≤-10	-0.505	0.824	-0.613	0.545	-0.022	0.012
-9	0.029	0.676	0.043	0.966	-0.013	0.014
-8	1.201	1.146	1.048	0.303	-0.011	0.035
-7	-0.629	0.909	-0.692	0.494	-0.025	0.012
-6	-0.756	0.837	-0.903	0.374	-0.025	0.010
-5	-0.975	0.871	-1.119	0.272	-0.028	0.008
-4	-0.863	1.332	-0.648	0.522	-0.036	0.019
-3	0.758	0.949	0.798	0.431	-0.012	0.027
-2	-0.590	1.000	-0.590	0.559	-0.026	0.014
-1	0	-	-	-	-	-
0	0.798	1.015	0.786	0.438	-0.013	0.029
1	-0.578	0.846	-0.683	0.50	-0.023	0.011
2	-1.039	0.758	-1.370	0.180	-0.026	0.005
3	-1.378	1.153	-1.195	0.241	-0.037	0.010
4	0.215	0.609	0.352	0.727	-0.01	0.015
5	0.326	1.025	0.318	0.753	-0.018	0.024
6	-0.263	0.900	-0.292	0.772	-0.021	0.016
7	-0.243	0.997	-0.244	0.809	-0.023	0.018
8	-1.095	1.272	-0.86	0.396	-0.037	0.015
9	-0.495	0.842	-0.587	0.561	-0.022	0.012
10	-0.140	0.804	-0.175	0.862	-0.018	0.015

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Coefficients are quarter-specific treatment interactions from the dynamic event-study specification, in percentage points. Quarter -1 is the omitted reference period. Endpoint bins (≤ -10 , 10) pool all more distant leads/lags. No individual coefficient is statistically significant at the 5% level.

Annex G: Leave-one-out analysis of event study specifications

The three earliest leads (q-10 to q-8, corresponding to 4Q2019-2Q2020 and comprising the onset of COVID-19 pandemic) jointly account for most of the baseline statistic in both specifications.

Table G1: Matched Specification 1 (Arup 7), baseline $\chi^2 = 30.48$

Dropped Lead	χ^2 (remaining)	Δ from baseline	% of baseline
q-10	15.17	15.31	50
q-9	22.06	8.42	28
q-8	21.07	9.41	31
q-7	29.04	1.44	5
q-6	29.01	1.47	5
q-5	28.54	1.94	6
q-4	29.23	1.25	4
q-3	27.45	3.03	10
q-2	24.99	5.49	18

Note: Each row reports the joint Wald statistic for the remaining pre-treatment leads when the indicated lead is dropped, the change from the full-window baseline, and that change as a percentage of the baseline.

Table G2: Matched Specification 2 (All EZL 17), baseline $\chi^2 = 27.73$

Dropped Lead	χ^2 (remaining)	Δ from baseline	% of baseline
q-10	17.34	10.39	38
q-9	18.73	9.00	33
q-8	12.16	15.57	56
q-7	27.64	0.09	0
q-6	27.73	0.00	0
q-5	27.38	0.35	1
q-4	27.04	0.68	3
q-3	24.38	3.35	12
q-2	27.06	0.67	2

Note: Each row reports the joint Wald statistic for the remaining pre-treatment leads when the indicated lead is dropped, the change from the full-window baseline, and that change as a percentage of the baseline.

Annex H: Independent-business displacement analysis

Motivation and definitions

While the aggregate analysis found no effect of the Elizabeth line on overall retail churn, this approach only measures how often occupants change and not the change in composition. This study further investigates whether independent occupants were displaced more rapidly in treated catchments.

Among at-risk independent occupants, every hereditament-quarter whose outgoing occupant was a resolved independent, an event is recorded where the incoming occupant is a chain (i.e. independent → chain) or the unit becomes vacant (i.e. independent → vacant). Independent → independent transitions were not counted as displacement. The rate is:

$$\text{displacement rate} = \frac{\text{independent displaced to chain or vacancy}}{\text{at risk independents}}$$

Across the full panel, 3,117 of 948,673 at-risk independents were displaced under this definition (0.33%).

Estimation

The rate was aggregated to LSOA × quarter and estimated with the same design as the turnover analysis with two-way (LSOA and period) fixed effects, CEM matching weights, and station-clustered standard errors, across both matching specifications (MS1, MS2) and both artefact-inclusion samples. Because displacement is rare and the at-risk denominator varies widely across cells (median 13, ~40% below 10), cells were weighted by the at-risk count so that precisely-measured rates dominate. This weight was combined with the CEM stratum weight.

Results

Difference-in-differences returned a positive and significant effect in every specification: +0.21 to +0.32pp across the four MS1/MS2 × artefact combinations, all $p < 0.05$. Table H1 reports all specifications.

Robustness checks

Three checks confirmed the pooled estimate was not an artefact of a single influential unit or of covariate imbalance. Leave-on-out diagnostics for Canary Wharf which accounted for 28-35% of treated post-opening displacement events left the estimate essentially unchanged. The effect also survived the doubly-robust specification remaining positive and significant. Finally, the at-risk pool was also checked for endogeneity. Rather than shrinking under displacement, the treated independent base grew slightly faster than control post-opening (+20–22% vs +16%), indicating the rate is not an artefact of a collapsing denominator.

Table H1: Difference-in-differences regression estimates of the effect of the Elizabeth line on quarterly independent retail churn (percentage points).

	(1) Base incl. artefact	(2) Doubly Robust incl. artefact	(3) Base excl. artefact	(4) Doubly Robust excl. artefact
Matching Specification 1 (Arup 7)				
treat × post (pp)	0.310	0.238	0.212	0.161
clustered SE	0.092	0.078	0.088	0.069
p	0.005	0.010	0.031	0.037
95% CI (pp)	[+0.112, +0.508]	[+0.068, +0.407]	[+0.022, +0.402]	[+0.012, +0.311]
Observations	4,602	4,602	4,529	4,529
Clusters	14	14	14	14
Matching Specification 2 (All EZL 17)				
treat × post (pp)	0.319	0.306	0.242	0.234
clustered SE	0.078	0.064	0.079	0.068
p	0.000	0.000	0.004	0.002
95% CI (pp)	[+0.161, +0.478]	[+0.175, +0.436]	[+0.082, +0.403]	[+0.094, +0.373]
Observations	7,540	7,540	7,419	7,419
Clusters	32	32	32	32
LSOA FE	Y	Y	Y	Y
Period FE	Y	Y	Y	Y
Covariates x post		Y		Y
Artefact excluded			Y	Y
CEM Weights	Y	Y	Y	Y

*Note: Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors are clustered by station. Doubly-robust specifications interact post with distance to city centre, number of lines, log hereditament density, IMD, and per cent residential.*

Parallel-trends failure

Despite this robustness, the identifying assumption fails. The event-study specification (binned $q-10$ to $q+10$) rejects the joint null that pre-treatment leads are zero, decisively, in both specifications (MS1 $\chi^2=56.2$; MS2 $\chi^2=36.3$; both $p<0.001$). Leave-one-out diagnosis of the pre-trend shows the violation is not localised to the specific early quarters that drove the analogous turnover violation. Contributions are spread across the pre-period and include near-treatment quarters

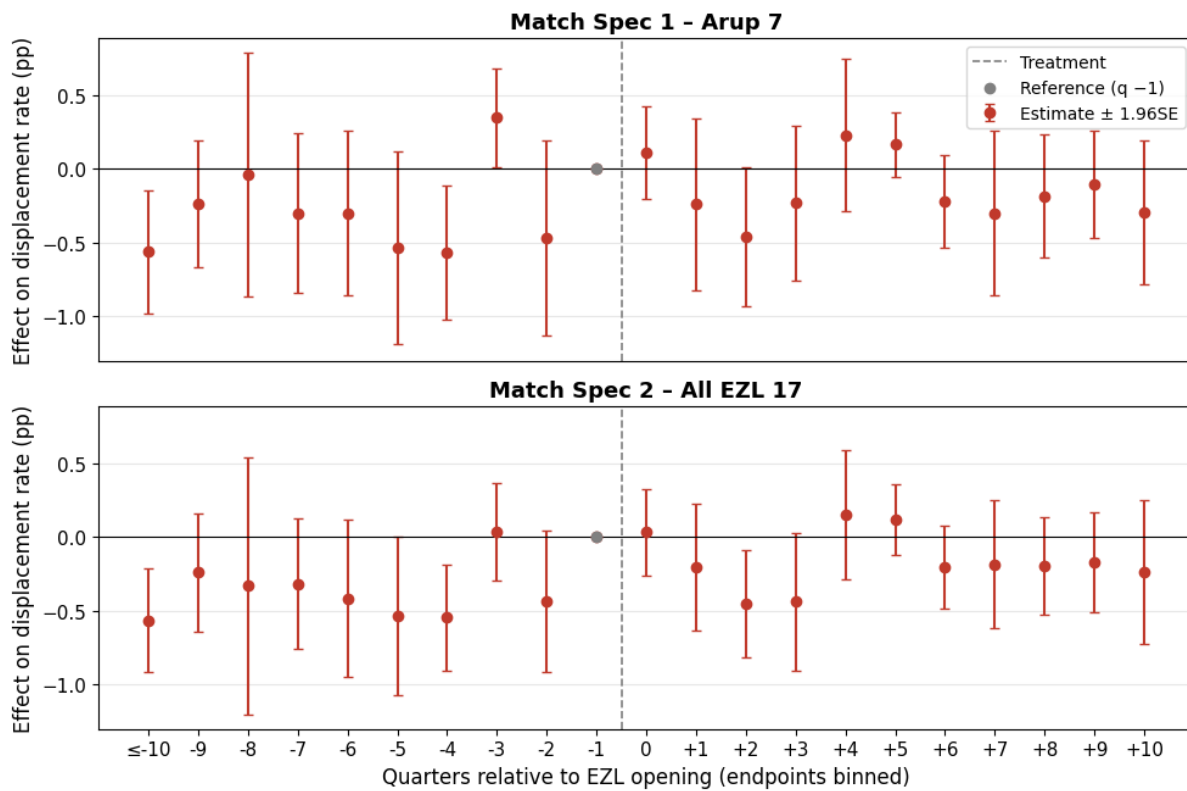


Figure H1: Binned Event-Study Plot (q-10 to q+10) for independent-displacement event study

Conclusion

Treated and control catchments were already diverging in independent-displacement rates before the Elizabeth line opened. The coefficient therefore reflects the continuation of a pre-existing trend and cannot be given a causal interpretation. The result suggests that there is possible displacement pressure in Elizabeth line catchments, but the design cannot separate that pressure from a pre-existing trajectory.

Research Log

Date	Venue	Subject	Attendees
10 Apr 26	Online, MS Teams	Dissertation Kick-off Meeting	Dr Jens Kandt, Benjamin Tee, Owen Hughes, Christian Mulder
1 May 26	Maple House, Room 503	Dissertation symposium 1: Research Question and focus	Dr Jens Kandt, Benjamin Tee, Owen Hughes, Christian Mulder
1 May 26	Maple House, Room 511	OpenLocal Data Meeting	Dr Jens Kandt, Benjamin Tee, Gavin Chait
22 May 26	Online, MS Teams	Refinements to research question and methods	Dr Jens Kandt, Benjamin Tee
29 May 26	Maple House, Room 503	Dissertation symposium 2: Literature review and research design	Dr Jens Kandt, Benjamin Tee, Owen Hughes, Christian Mulder
8 Jun 26	Maple House, Room 501	Challenges with OpenLocal dataset	Dr Jens Kandt, Benjamin Tee
26 Jun 26	Maple House, Room 503	Dissertation symposium 3: Research design and prelim results	Dr Jens Kandt, Benjamin Tee, Owen Hughes, Christian Mulder
6 Jul 26	Maple House, Room 501	Update on research results and progress	Dr Jens Kandt, Benjamin Tee
31 Jul 26	Maple House, Room 503	Dissertation symposium 4: Key findings, limitations and future work	Dr Jens Kandt, Benjamin Tee, Owen Hughes, Christian Mulder