Essays in Public & Urban Economics

By

Rowan Isaaks

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Economics

August 9, 2024

Nashville, Tennessee

Approved:

Lesley Turner, Ph.D.

Brian Beach, Ph.D.

Christopher Carpenter, Ph.D.

Bryan Weber, Ph.D.

Acknowledgements

I am grateful to my committee for their insights and for making the time to improve this dissertation. I especially want to thank Lesley Turner for her patience and wisdom throughout this process.

Thank you, Steph, for keeping me sane.

TABLE OF CONTENTS

Page

LIST OF TABLES . . . . **v**

LIST OF FIGURES . . . . **vii**

1 Revealed Preferences for Residential Traffic Calming: Evidence from Low Traffic Neighborhoods **1**

1.1 Introduction 1
1.2 Low Traffic Neighborhoods 6
1.3 Empirical Approach 8
1.4 Data 12
1.4.1 Assigning Neighborhoods to LTN Treatment 12
1.4.2 House Price Data 13
1.5 LTNs and House Prices in Waltham Forest 15
1.6 Heterogeneity & Local Amenities 22
1.6.1 Heterogeneity 22
1.6.2 Traffic Accidents 22
1.6.3 Crime 24
1.7 Discussion 25
1.8 References 27
1.9 Appendix 34
1.9.1 Technical Details of Low Traffic Neighborhoods 34
1.9.2 Polygon Construction for Neighborhood-Level Specifications 35
1.9.3 Valuation of Accident Reduction 37
1.9.4 Additional Figures 38
1.9.5 Additional Tables 41

2 Capitalization of 24-Hour Subway Service: Evidence from London's Night Tube **42**

2.1 Introduction 42
2.2 The Night Tube 44
2.2.1 Origins of the Night Tube 44
2.2.2 Justification & Rollout 46
2.2.3 Other Possible Impacts of Night Tube 47
2.3 Data & Empirical Approach 50
2.3.1 Data Sources 50
2.3.2 Empirical Strategy 50
2.3.3 Descriptive Statistics 51
2.4 Main Results 52
2.5 Heterogeneity & Mechanisms 59
2.5.1 Investigating Mechanisms 59
2.5.2 Heterogeneity Analysis 63
2.6 Conclusion 68
2.7 References 69
2.8 Appendix 73
2.8.1 Additional Figures 73
2.8.2 Additional Tables 78

3 Gender Diversity & Municipal Spending: Evidence from California City Councils 85

3.1 Introduction 85
3.2 City Councils in California 88
3.3 Data & Empirical Strategy 89
3.3.1 Data 89
3.3.2 Empirical Strategy 91
3.4 Results 94
3.4.1 Validity of Identifying Assumption 94
3.4.2 First-Stage Dynamic Effects 95
3.4.3 Regression Discontinuity Results 97
3.5 Mechanisms 104
3.5.1 Transportation Spending 104
3.5.2 Null Results 105
3.6 Conclusion 106
3.7 References 108
3.8 Appendix 113
3.8.1 Additional Figures 113
3.8.2 Additional Tables 116

LIST OF TABLES

Table		Page
1.1	Selected Descriptive Statistics (Means) by Treatment Status	16
1.2	Effects on House Prices	17
1.3	Robustness Tests and Callaway & Sant'Anna (2021) Specification	19
1.4	Investigating Spillover Effects	20
1.5	Effect on Supply: Number of Sales	21
1.6	Heterogeneity by Property Characteristics and Neighborhood Attributes	23
1.7	Effect of LTN on Traffic Accidents	23
1.8	Effect of LTNs on Crime	25
1.9	Borough Comparison by Mini-Holland Grant Status	41
1.10	Effects on House Prices	41
2.1	Selected Descriptive Statistics: Baseline Period (2010-2014)	53
2.2	Main Regression Results	57
2.3	Investigating Effects on Number of Sales	59
2.4	Investigating the Noise Mechanism	61
2.5	Testing the Crime Mechanism	63
2.6	Heterogeneity by Property Type	65
2.7	Heterogeneity by Location (Zone)	66
2.8	Interactions with Distance from Center of London	66
2.9	Station-Neighborhood Characteristics by Zone (2011 Census)	67
2.10	Event Study Estimates	78
2.11	Repeat Sales Specifications with Station Linear Time Trends	79
2.12	All Sales Regressions with Multiple Sales Sample	80
2.13	House-Price Premiums for Regular Tube Service	81
2.14	Specifications Absorbing Postcode Fixed Effects	82

2.15 Non-Windsorized Specifications . . . 83

2.16 Heterogeneity by Property Type and Zone . . . 84

2.17 Effects on Local Age Distribution . . . 84

3.1 Summary Stats and Differences in Means . . . 93

3.2 First-Stage Dynamic Effects by Subsequent Election Order . . . 96

3.3 First-Stage Incumbency Effect (Females) . . . 96

3.4 First-Stage Dynamic Effects on Non-Incumbent Female Victory . . . 97

3.5 Effect of a Marginal Female Victory on Expenditure Composition (Standard Deviation Measure) . . . 100

3.6 Effect of a Marginal Female Victory on Expenditure Composition (Per-Capita Measure) . . . 101

3.7 Differences in Spending Preferences by Gender . . . 107

3.8 Identification Assumption - RD with Baseline Variables as Outcomes . . . 116

3.9 Simple Incumbency Effects . . . 116

3.10 Effect of Marginal Female Victory on Expenditure & Composition (2-Year Lag) . . . 117

3.11 Effect of Marginal Female Victory on Expenditure & Composition (3-Year Lag) . . . 117

3.12 Effect of Marginal Female Victory on Expenditure & Composition (4-Year Lag) . . . 118

3.13 Effect of Marginal Female Victory on Expenditure & Composition (5-Year Lag) . . . 118

3.14 Effect of Marginal Female Victory on Expenditure & Composition (Aggregate) . . . 119

3.15 Heterogeneity of Marginal Female Victory (2-Year Lag) . . . 119

3.16 Heterogeneity of Marginal Female Victory (3-Year Lag) . . . 120

3.17 Heterogeneity of Marginal Female Victory (4-Year Lag) . . . 120

3.18 Heterogeneity of Marginal Female Victory (5-Year Lag) . . . 121

3.19 Heterogeneity of Marginal Female Victory (Aggregate) . . . 121

LIST OF FIGURES

Figure		Page
1.1	Illustration of a toy LTN	7
1.2	Examples of modal filters in Waltham Forest.	9
1.3	Location of LTNs in Waltham Forest implemented during Mini Holland pilot program (2015-2019), delineated at the postcode level	13
1.4	Year of LTN implementation in Waltham Forest, delineated at the postcode level.	14
1.5	Event Study of Stacked DID	18
1.6	Map of postcodes showing corresponding ring classification.	21
1.7	Modal filter that completely blocks car traffic (left) and a modal filter that is enforced by camera (right)	34
1.8	Example of implementing an LTN using modal filters in a grid-style road network.	35
1.9	Map of Treatment and Control Polygons used for panel regressions	36
1.10	Mini-Holland grant eligibility/outcome by borough	38
1.11	Map of (732) Output Areas in Waltham Forest	39
1.12	Event Study of Callaway & Sant'Anna (2021) Specification	40
2.1	Night Tube Service Map as of May 2020	47
2.2	Full Tube Service Map as of April 2020	48
2.3	Event Study - All Sales	54
2.4	Event Study - Repeat Sales Only	55
2.5	Distribution of Distance from Nearest Track Section	60
2.6	Event Study - Crime Count - All Sales	64
2.7	Event Study – All Sales – 0.5-Mile Radius	73
2.8	Event Study – Repeat Sales Only – 0.5-Mile Radius	74
2.9	Event Study (Quarter) - All Sales - 1 Mile Radius	75
2.10	Event Study (Quarter) - Repeat Sales - 1 Mile Radius	76
2.11	Crime Event Study - 2015 Treatment Definition	77

3.1 Test for Manipulation of Margin of Victory 94

3.2 RD Effects of a Marginal Female Victory 99

3.3 RD Heterogeneous Effects by Prior Council Composition 102

3.4 RD Effects on Transportation Spending by Type 104

3.5 Top Priority for State Spending - Males (0) vs. Females (1) 106

3.6 “Stacked” RD Estimates 113

3.7 Density Test - First Female Only 114

3.8 Density Test - Majority Female Only 114

3.9 RD Separate Heterogeneous Transportation Effects 115

3.10 RD Plot - Average Post-Election Transport Share by MV 115

CHAPTER 1

Revealed Preferences for Residential Traffic Calming: Evidence from Low Traffic Neighborhoods

1.1 Introduction

When transportation infrastructure experiences congestion, users of that infrastructure impose external costs onto residents. In the case of road networks, drivers that contribute to congestion impose time costs on other drivers (Anderson, 2014), which partially attenuates agglomeration economies of cities. Congestion also imposes costs on those who reside where congestion occurs: Previous research has found that congestion contributes to pollution and negative health outcomes (e.g. Knittel et al., 2016) and subjective measures of welfare (Conceição et al, 2023). Many countries experienced increased use of automobiles in recent decades, especially in residential neighborhoods. Between 2009 and 2019, total vehicle mileage in Great Britain increased by 15.4%, with mileage on minor roads increasing by over 25% (*Department for Transport*, 2020a). At sufficiently high levels, this traffic imposes a negative externality on residents (Hughes & Sirmans, 1992; Bateman et al., 2006; Ossokina & Verweij, 2015) that can operate through multiple channels such as increased air/noise pollution, accident risk, and crime. Theoretically, externalities can be be addressed by either price- or quantity-based interventions. Existing literature and policy tends to focus on price-based interventions such as congestion pricing and highway tolls (e.g. Gronau, 1999; Currie & Walker, 2011; Audretsch, Dohse, & Santos, 2020). However, these types of policies neither practical nor efficient solutions to address traffic in residential neighborhoods. While there is less evidence on interventions that restrict the quantity of traffic, a nascent option is the Low Traffic Neighborhood (LTN), which prohibits through-traffic while still permitting local access to the neighborhood, however, there is little evidence on the effectiveness of LTNs in ameliorating local disamenities from congestion.[1]

In this chapter, I explore the effects of an LTN pilot program that was rolled out between 2015 and 2019 in the London Borough of Waltham Forest on house prices. Specifically, I use a difference-in-differences (DID) research design, exploiting both spatial variation in the placement of LTNs and the staggered nature of rollout to analyze how residential property sale prices change in areas within LTNs compared to untreated neighborhoods in Waltham Forest, before and after implementation. In order to interpret the DID estimate as causal, LTN treatment must be exogenous with respect to housing price trends. While this assumption is not directly testable, I show event studies to test whether price trends are parallel prior to the introduction of

[1] Although LTNs do not explicitly limit the quantity of vehicle traffic, by eliminating a certain type of traffic the quantity significantly decreases.

LTNs. Additionally, in order to to interpret a change in house prices as the marginal WTP for an LTN, three other conditions must be met (Rosen, 1974; Greenstone, 2017):[2] First, all other characteristics of properties must be controlled for. In lieu of observing all relevant characteristics, I include property fixed effects in my model, so that identification comes from repeat sales of the same property before and after treatment. Second, housing supply should be relatively inelastic during the study period, otherwise changes in price could be the net effect of changes in demand and quantity demanded. I test this assumption empirically and find no evidence of changes in quantity of sales as a result of treatment. Finally, it cannot be the case that LTNs are funded by increases in property taxes. This assumption is satisfied in my setting.[3] I find that when a neighborhood receives an LTN, the price of houses sold increases by roughly 6.5% relative to untreated areas, providing the first evidence on net revealed preferences for residing within an LTN. These results are robust to alternative choices for the control group, time-varying neighborhood controls, and specifications not restricting the sample to properties with multiple sales. This effect is comparable to other local amenities such as being near a park in a safe area Albouy et al. (2020) or a power plant Davis (2011) and is larger than the house price premium attributed to the implementation of the London Congestion Charge ($\approx 3\%$) (Tang, 2021).[4]

I explore the mechanisms through which LTNs impact preferences over neighborhoods by investigating the extent to which measurable amenities also change as a result of LTN implementation. Previous descriptive analysis suggests that the presence of an LTN is associated with a decrease in crime (Goodman & Aldred, 2021) and traffic accidents (Laverty, Aldred, & Goodman, 2021). I find evidence that LTNs reduce the frequency of serious traffic accidents by more than half, but have negligible effects on most categories of crime.

The goal of an LTN is to internalize the externalities produced by commuters using neighborhood streets to reduce their own travel time, which is achieved by preventing use of these streets by commuters and disincentivizing driving altogether. However, it is not clear whether LTNs will ameliorate negative externalities in treated areas, or displace those externalities to other neighborhoods. One way that drivers may (partially) avoid internalizing the cost of traffic is by using untreated, adjacent neighborhoods as alternate routes. It could also be the case that the traffic is displaced from LTN neighborhoods to a more diffuse

[2]The relationship between house prices and a given characteristic, or the *hedonic pricing schedule* (Rosen, 1974), is a method of empirically estimating preferences for a local public good, as described by Tiebout (1956).

[3]Council (property) tax in the United Kingdom is fixed based on the value of the property as of April 1991, from which the property is assigned one of eight valuation bands. This means that council taxes are not related to current-day fluctuations in value. Additionally funding for the LTNs studied in this chapter were funded by a grant from the national government. This is discussed briefly in Section 1.2.

[4]The London Congestion Charge (LCC) is not strictly analogous to the Waltham Forest LTN program, since the LCC included both major roads and residential areas, and was based in Central London. However, it is probably the most comparable price-based policy to the LTN pilot.

area. While the latter case is hard to (dis)prove, I address the former case by running placebo tests where I assign a psuedo-treatment to neighborhoods that are close to ever-treated ones (and thus likely spillover candidates), and compare them to control neighborhoods that are further away from ever-treated areas. I do not find significant effects of this psuedo-treatment, suggesting that the effect on house prices is likely driven by improvements in treated areas.[5] Additionally, I directly try to estimate spillovers using a concentric ring method, and find evidence that properties within half a mile of LTNs experience a net-positive spillover effect, suggesting that there is a perceived benefit to living near an LTN even if I do not reside inside one myself.

There are broadly two types of interventions available to address congestion-based market failures: Price-based and quantity-based. Most research looks at price-based interventions such as congestion charging zones (e.g. Tang, 2021) and toll roads (e.g. Theisen, 2020) that charge a fee to the driver for driving on a specific road or in a specific area. However there are two broad motivations for analyzing quantity-based interventions: First, since policymakers operate with imperfect information, in cases when where the marginal cost of additional congestion can become very large, it may be preferable to set quantity directly rather than indirectly via a price mechanism.[6] Second, there are settings where price-based interventions may not be feasible and/or cost-effective. For example, when targeting specific residential neighborhoods that contain many entrances and exits via side streets, it is not practical to put cameras and tolls at every neighborhood entrance, since the quantity of traffic is unlikely to justify the of installing this infrastructure. In these cases, explicitly restricting quantity by installing physical barriers or changing road rules is a more viable approach. Less is known about quantity-based interventions, partly because they are less prevalent. One example of a quantity based intervention is an 'odd-even' rationing scheme, which restricts road access to half of all private vehicles on any given day based on the last digit of the license plate. This measure has been implemented in large cities including Delhi, Beijing, and Quito, and there is evidence that these policies reduce local pollution (Carillo, Malik, & Yoo, 2016) and increase the demand for access to mass transit (Jerch et al., 2023), but can also affect crime negatively in at least one instance (Carillo, Lopez-Luzuriaga, & Malik, 2018).[7] This type of policy is effective in reducing the maximum traffic volume in the city (or city center) as a whole, but is i) unlikely

[5]While there is no significant effect on house prices in spillover neighborhoods, the 95% confidence interval from this regression includes the DID coefficient of my actual treatment, so this result should be taken as suggestive evidence of no spillover effects rather than definitive proof. I also run specifications where I exclude the spillover candidates from the control group, and the magnitude of these estimates are not significantly different from my main results.

[6]Under perfect information, an urban planner could choose the optimal price-quantity pair and implement them with either price or quantity interventions. However, under imperfect information, this is not possible, and which approach will bring us closer to the social optimum is uncertain. In Weitzman's (1974) model, the curvature of the marginal cost of abatement determines whether price or quantity-based regulation is likely to be preferred. In a case where the marginal cost of congestion increases rapidly after a certain 'critical point', the penalty for 'choosing' the wrong quantity becomes large. In this case, it is preferable to restrict quantity to avoid this bad outcome. On the flip-side, when the slope of marginal cost is smaller, deviations from the optimal choice of price are less costly.

[7]This paper finds an increase in the total crime rate between 5 and 10% due to a vehicle restriction in Quito, Ecuador. The authors find suggestive evidence that a key mechanism is the diversion of policing resources towards enforcement of the policy, which correspondingly reduces resources available to prevent and disincentivize other crimes.

to be politically feasible in many metropolitan areas and ii) does not eliminate the phenomenon of through traffic in residential neighborhoods. This chapter is most closely related to Polloni (2019), who researches the effects of traffic irritants in Portland, Oregon (primarily in the form of speed humps and curb extensions) and their effect on house prices using a difference-in-differences design. Despite evidence of reductions in traffic speed and volume, there is no statistically-significant evidence of a causal treatment effect of these irritants on house prices.[8] I contribute to this literature by analyzing the effect of a policy that implements quantity-based traffic calming i) specifically to residential areas and ii) specifically *prohibits* through-traffic from neighborhoods rather than traffic irritants that affect both local and through-traffic, providing evidence of a significant increase in WTP.[9]

My findings also contribute to a nascent literature on LTNs. Previous analysis has looked at the effect of LTNs on several outcomes: There is descriptive evidence that LTNs in Waltham Forest are associated with a decrease in local crime (Goodman & Aldred, 2021) and an increase in road safety (Laverty, Aldred, & Goodman, 2021). Two studies have suggested that there is a modest decrease in car use and ownership in areas containing LTNs (Aldred & Goodman, 2020; Goodman et al., 2023). Finally, a recent paper by Yang et al. (2022) found evidence of LTNs in Islington, London decreasing pollution (NO_2) by around 6% and traffic flows by over 50% using a generalized difference-in-differences model.[10] LTNs are also a hugely controversial topic in local politics in the United Kingdom, generating strong opinions for and against their use: Disagreements about the implementation of LTNs have led to public protests (e.g. Seaward, 2023) and LTN infrastructure being vandalized (e.g. Taylor & Murray, 2023; Reynolds, 2023). In addition, LTNs have become a common debate topic between the two major political parties in the United Kingdom, with some representatives voicing concerns that LTNs are "anti-car" and that they make traffic worse in non-LTN areas. In 2023, prime minister Rishi Sunak ordered a review into LTNs, and is generally skeptical of them (Walker, 2023). Additionally, The Transport Secretary, Mark Harper, called for local authorities to review the suitability of their existing LTNs and has stopped funding from central government being used for the creating of new LTNs, meaning that local councils must use their own funds to implement future LTNs (Malnick, 2023). Due to the politicized nature of LTNs and the influence of local politics in their implementation, residents may be incentivized to misrepresent their preferences to local officials and representatives, making people's stated

[8]This chapter does show that there is a negative relationship between traffic volume and house prices, indicating that traffic calming projects should be expected to affect house prices if the resulting change in traffic volumes are large enough.

[9]Although LTNs do not set a specific quantity restriction on traffic flows, the ultimate effect of the policy is to significantly reduce traffic quantity. I do not directly observe traffic flows in this setting, but a suggestion of the magnitude in a similar setting is estimated to be a reduction of over 50% (Yang et al., 2022).

[10]It should be noted that the LTNs studied in this chapter were implemented in 2020 while COVID-19 lockdowns and travel restrictions were still active. Additionally, there is small amount of variation in treatment timing present in the LTNs studied in this chapter, but this is not addressed in the body of the paper.

preferences unreliable estimates of WTP. Therefore, information on revealed preferences for LTNs should be of interest to policymakers as well as economists. The value of housing is a useful metric that estimates the net-effect of LTNs and provides a well-established method to value public investments.[11] I contribute to the LTN literature by providing novel evidence on WTP for LTNs, showing a causal increase in house prices for neighborhoods that receive LTNs. I also evaluate whether relationships between LTNs and neighborhood amenities such as traffic accidents and crime are likely to be causal by using contemporary difference-in-differences methods. Finally, I address concerns that benefits accruing to treated neighborhoods come at the cost of making congestion worse in nearby untreated areas: I show that the my results are being driven by increases in WTP in treated areas, rather than being partially or wholly driven by decreases in house prices in nearby neighborhoods.

Finally, a characteristic of urban neighborhoods whose merits are debated is *permeability*; which can be defined as the extent to which the layout of the neighborhood permits the movement of people and vehicles through it. The go-to archetype of a low-permeability neighborhood is a cul-de-sac. There is some evidence to suggest that homes in cul-de-sacs fetch significant premiums compared to other homes (e.g. Asabere, 1990), but it is not certain whether this difference is caused by differences in permeability or by other differences correlated with permeability. For example, studies on neighborhood design in the United Kingdom have found that increased permeability is associated with increased crime (Johnson & Bowers, 2009; Armitage, Monchuk, & Rogerson, 2010). Aldred & Goodman (2021) mention two plausible channels for this effect: i) Increased permeability could lead to more "eyes on the street" a la Jane Jacobs (1961) which would increase the likelihood of being caught, or ii) LTNs could decrease ease of criminal access to a neighborhood via motor vehicle (Newman, 1973). LTNs are therefore of interest since they inherently reduce the permeability of a neighborhood.

The rest of the chapter is structured as follows: In Section 1.2, I give an overview of LTNs and the implementation of the Waltham Forest pilot program. In Section 1.3, I discuss my empirical approach and identification using variation in the timing of LTN rollout to different areas. In Section 1.4, I discuss the data sources that I use in my analysis. In Section 1.5, I present the main results on house prices and address the issue of spillovers. In Section 1.6, I provide evidence on the effects on local amenities that could contribute to house price increases and discuss heterogeneity in the main results. In Section 1.7, I wrap up with a discussion of findings and implications for the debate over LTNs and future urban planning policy.

[11] Examples of hedonic analysis being used to value investments include: Oates, 1969 (tax policy); Black, 1999 (school quality); Linden & Rockoff, 2008 (crime); Chay & Greenstone, 2005 (environmental quality); and Turner et al., 2014 (land use regulation).

1.2 Low Traffic Neighborhoods

Traffic in Great Britain has been trending upwards for many years, with traffic volume increasing disproportionally on minor urban roads in the last decade. Between 2009 and 2019, total vehicle mileage on urban, minor roads increased by 26.6%, compared to an increase of 15.7% among all roads (*Department for Transport*, 2020a). There are two main explanations for this increase. First, increased use of navigation technology and applications lead drivers to utilize residential roads as alternative routes, whereas previously they would have to be familiar with these routes and make informed guesses about the relative amount of traffic on each route (Kojima, Elfferding, & Kubota, 2014; Sustrans, 2020).[12] Second, given that vans and other cargo/delivery vehicles represent the largest increase in vehicle mileage by type (*Department for Transport*, 2020a), the rise of internet shopping and food/grocery delivery services may be a major factor in traffic increases. Long-standing concerns about rising road congestion in urban areas (Downs, 2004) provide motivation for Low Traffic Neighborhoods (LTNs) as a potential remedy.

Although the concept of an LTN is relatively recent, coming into public prominence in 2020 in the midst of COVID-19-related restrictions on travel and indoor activity, the elements that make up LTNs are not. LTNs are characterized by modal filters, which are placed into a road design to restrict passage to certain modes of transport.[13] Most commonly, modal filters are designed to prevent the passage of cars, vans, and trucks, while still allowing pedestrians and bicycles to use the road normally. Residential neighborhoods are typically bordered by main roads, so urban planners will define the boundaries of LTNs based on these borders. Figure 1.1 shows an illustration of a simple LTN structure: A residential neighborhood is bounded by main roads. Modal filters are placed at strategic locations such that the residential streets are still accessible to vehicles, but there are no through-routes available to drivers.

One of the major concerns voiced by LTN critics is that since drivers cannot use routes that contain roads within LTNs, they will simply switch to using the next best alternative route, which will exacerbate congestion on border roads and/or create through routes in adjacent residential neighborhoods. As part of my analysis, I assess whether there is any evidence of this effect driving my main results on price. Additionally, since LTNs are constructed such that they exclude main roads, I run specifications including/excluding main roads from the control group to test if my results are sensitive to or driven by this factor. In this chapter's

[12] *Google Maps*, the most popular mobile navigation application, was released in 2005. Evidence from the U.S. has found that rideshare entry into metropolitan areas leads to an increase in congestion intensity and duration Diao, Kong & Zhao (2021). The negative effects of increased congestion in residential areas via use of through-routes are compounded by the fact that a) residential streets are not designed for a high volume of traffic and b) drivers (particularly if routed by an app) are not familiar with street layouts (Cabannes et al., 2018).

[13] Modal filters have existed in the United Kingdom since the 1960s. However, until recently, they were typically implemented as stand-alone solutions to isolated issues on a road, rather than groups of modal filters being used systematically to treat a whole neighborhood.

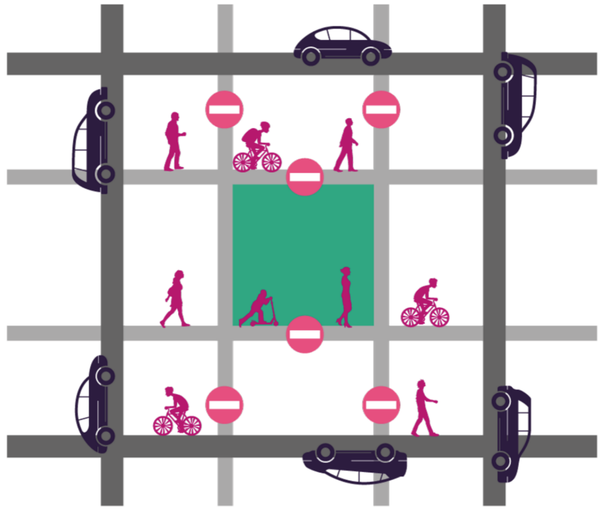


Figure 1.1: Illustration of a toy LTN

appendix, I discuss LTNs, types of modal filters and, implementation in more detail from the perspective of the urban planner.

The increasing prevalence of through-routes has led to residents reporting worse air quality and feeling less safe walking or cycling (Bosetti et al. 2022). In response to this trend, Boris Johnson, then the mayor of London, announced the *Mini Hollands Scheme*[14] in March 2014. This scheme's purpose was to reduce car dependency and encourage residents to cycle or use other forms of transportation. The end goal being that the scheme would "reduce pollution, traffic danger and congestion" (*Department for Transport*, 2020b). The scheme took the form of competitive grants that were available to be awarded to 20 of the 32 London boroughs classified as 'outer boroughs', since car dependency is typically higher compared to London's inner boroughs. Three boroughs (Enfield, Kingston-Upon-Thames, and Waltham Forest) were selected based on the perceived potential of their proposals to achieve reductions in car dependence and increases in walking and cycling as modes of transport.[15] The three boroughs were awarded £30 million each, with borough councils matching up to an additonal 50%.[16] Table 1.9 contains a comparison of boroughs with successful and unsuccessful grant applications, as well as those that did not apply. Overall, there is little difference in observed demographics, car ownership, and tenure status between boroughs who received grants and those who did not.

[14]The name is an ode to the modern Dutch style of urban planning, which prioritizes pedestrian safety and extensive cycle networks, and places vehicular traffic as a secondary priority.

[15]Out of the remaining outer boroughs, five applied but were not successful, those being Newham, Richmond, Bexley, Merton and Ealing.

[16]Figure 1.10 in the appendix shows a borough map of London and the locations of boroughs that were awarded these grants.

The locations of proposed LTNs in Waltham Forest were not chosen randomly: Neighborhoods were selected for traffic calming by the borough based on their propensity to achieve the goal of the Mini Holland scheme: Decreased car journeys and an increase in journeys made by foot and bicycle in the local area. Aside from having a recognized issue with through-traffic, preference was given to neighborhoods located in the vicinity of village high streets, local markets, and green areas.[17] This is concordant with generally-accepted practices for designing LTNs that they should be "in close proximity to key amenities and services, especially key transport interchanges", be bounded in some way by main roads, and be walkable within a reasonable amount of time, usually 15 minutes as a benchmark (London Living Streets, 2019). In addition, the boroughs proposals were required to be presented to residents and incorporate feedback, although I found no evidence that residents were able to vote for or against LTNs being implemented. Figure 1.2 shows examples of modal filters and infrastructure implemented as part of Waltham Forest LTNs: The top row modal filter where the bollard can be lowered for emergency vehicles (a) and a camera-enforced LTN entrance with planters to narrow the roadway (b). The middle row shows a before (c) and after (d) the installation of bollards and a raised pavement to prevent through traffic. The bottom row shows a before (e) and after (f) of of a pedestrianized segment of a village high street.

Previous literature suggests that LTNs can reduce traffic inside the zoned area (Goodman et al., 2023), and observational evidence on the impact of both the Waltham Forest and post-COVID LTNs indicates that there may be positive effects on crime, road accidents, and pollution. However, there is no research that estimates the net value that residents place on living in a neighborhood using a revealed preference approach, which is what I do in the remainder of this chapter.

1.3 Empirical Approach

I analyze how the introduction of LTNs into residential neighborhoods in Waltham Forest affected the WTP for properties in those neighborhoods. A simple difference in differences design would compare house sale prices of two groups before and after LTNs were implemented: A treated group, consisting of houses in neighborhoods that receive LTNs, and a control group of houses in neighborhoods that do not. This setup is described by Equation (1.1) below:

$$ln(y_{it}) = \beta_0 + \beta_1 LTN_i * POST_t + \beta_2 LTN_i + \beta_3 POST_t + \varepsilon_{it} \tag{1.1}$$

[17] The effect that these elements have on LTN location selecton, and how that affects our interpretation of my result, is discussed in Section 1.7.

Figure 1.2: Examples of modal filters in Waltham Forest.

Where $ln(y_{it})$ is the natural log of the sale price of a residential property with neighborhood treatment status $i \in \{LTN, Control\}$, and period $t \in \{Pre, Post\}$. β_1 is the regression coefficient for the DID treatment indicator $LTN_i * POST_t$. Under the assumption that sale prices of houses in LTN and control neighborhoods would have trended similarly in the absence of LTNs, β_1 represents the average increase in house prices in LTN neighborhoods caused by LTN treatment.

In reality, LTNs in Waltham Forest were introduced in a staggered fashion between 2015 and 2019, rather than all at once. To account for this variation in the timing of LTN treatment, I use a stacked DID approach in the vein of Deshpande & Li (2019). For each set of LTNs implemented in the same period t, a stack s is created such that those LTNs are the only treated areas. Areas that were treated later than t or never treated serve as the control group for stack s. Within each stack, the pre-period and post-period are defined by the timing of treatment for the LTN(s) being considered in the stack. For example if treatment in stack s occurs in 2016, then for all observations in that stack, the post-period is defined as 2016 onward. The overall average reduced-form effect of LTN treatment on house sale price can be estimated by appending all stacks into a single data set and estimating the following equation:

$$ln(y_{its}) = \beta_0 + \beta_1^{stacked} LTN\_Treat_{its} + \gamma_{is} + \delta_{ts} + \varepsilon_{its} \quad (1.2)$$

Where the outcome $ln(y_{its})$ is the natural log of the sale price of properties located in neighborhood i sold in time period t for the stack s, $\beta_1^{stacked}$ is the parameter of interest associated with $LTN\_Treat_{its} = LTN_{is} * POST_{ts}$, where LTN_{is} is the indicator for units that are treated in stack s and $POST_{ts}$ is the indicator for periods in or after the year of treatment in stack s. γ_{is} and δ_{ts} are neighborhood-stack and time-stack fixed effects respectively, which forces identification of $\beta_1^{stacked}$ to come from variation within a stack, and thus will not contain any problematic comparisons of just-treated units with earlier-treated units. The counterfactual parallel trends assumption is still present in this specification, i.e. in the absence of any LTN implementation, the trend in sale price would be similar for properties in neighborhoods that did receive LTNs vs. those that have not yet been treated and/or will never be treated. I obtain similar estimates using the Callaway & Sant'Anna (2021) group-time average treatment effects to estimate an average treatment effect for each group-time pair, and then aggregate these group-time ATEs to obtain an overall average treatment effect across all treatment timing groups.[18] Both the stacked DID and Callaway & Sant'Anna (2021) approaches

[18] Similar to the stacked approach, the Callaway & Sant'Anna method divides ever-treated neighborhoods into groups g based on the time of LTN treatment t. $\forall g$ and $\forall t \geq g$ I estimate a 2x2 DID regression: $ln(y_{it}) = \alpha_0 + \alpha_{1,gt} \mathbb{1}[G_i = g] * POST_t + \alpha_2 \mathbb{1}[G_i = g] + \alpha_3 POST_t + \varepsilon_{it}$, where $\mathbb{1}[G_i = g]$ is an indicator equal to 1 if property i is in group g and $\hat{\alpha}_{1,gt}$ is the estimate of the average treatment effect for group g in year t. I show the results of these estimations by showing an overall aggregate treatment effect and event study estimates.

are preferred to a generalized Two-Way Fixed Effects (TWFE) model. This is because the TWFE model will include some 2x2 comparisons where already-treated units are serving as controls for newly-treated units. This will only result in a consistent estimate if the effect of LTN treatment is constant over time (Goodman-Bacon, 2021). This is a strong assumption that is unlikely to be true in my setting because i) individuals may take time to update their preferences about LTNs since they are not widespread interventions, and ii) LTNs are not perfectly homogeneous treatments.

To claim that my estimates of WTP are plausibly causal, I address two additional challenges to my identification strategy: First, my data source for house prices does not allow me to observe all relevant characteristics of properties that could be correlated with LTN treatment status and could affect sale price (e.g. lot size, # of bedrooms, etc.). To overcome this issue, I use property-stack fixed effects (λ_{ps}) instead of γ_{is} in Equation (1.3), which controls for all time-invariant characteristics of properties.

$$ln(y_{pts}) = \beta_0 + \beta_1^{stacked} LTN\_Treat_{pts} + \lambda_{ps} + \delta_{ts} + \varepsilon_{pts} \tag{1.3}$$

This means that identification is coming from comparing repeat sales of properties before and after treatment. Since not all properties in my data are sold multiple times during the sample period, I utilize a subset of property sales in my preferred specifications. Second, interpreting a change in price as solely a change in WTP rests on the assumption that supply is relatively inelastic. To address this, I include regression specifications where the outcome variable is the number of sales observations in a neighborhood.[19] I find no statistical evidence that the number of sales is changing differentially between LTN and control neighborhoods as a result of LTN treatment.

LTNs are a local public good, and thus are related to the framework of the Tiebout (1956) model. In Tiebout's model, people make decisions about which locality to live in based partially on their preferences for a local public good. Conditional on people being geographically mobile and the LTN being excludable at the local level, people will "vote with their feet"; revealing their preference for living in a neighborhood within an LTN by their choice of where to live. The reduced-form parameter I estimate ($\beta_1^{stacked}$) is the effect on the sale price of a house due to the presence of an LTN in the neighborhood, which I obtain by applying a DID-version of the hedonic pricing framework first developed by Rosen (1974), which considers a housing unit as containing a host of characteristics, both physical characteristics such as size and number of bedrooms, and amenities specific to the local area. Consumers hold preferences for each of these attributes individually

[19]The results of this regression are shown in Table 1.5.

conditional on household budget. If we can control all other relevant attributes, then one can estimate the marginal price an individual is willing to pay for a given amenity.[20] In the case of LTNs, the underlying local amenity this chapter is attempting to value is (a lack of) traffic. However, in practice, an LTN does not completely isolate and mitigate traffic, leaving all else unchanged. For example, eliminating through-traffic could also locally reduce airborne pollutants (Yang et al., 2022). Since I do not observe neighborhood-level pollution in this research, I am unable to separate these two amenities.

1.4 Data

My main analysis uses publicly-available administrative data on residential housing transactions in Waltham Forest and previously-collected data on the locations of LTNs. In Section 1.6, I look at intermediate channels by which LTNs may be affecting house prices, using additional data on crime records and road accidents.

1.4.1 Assigning Neighborhoods to LTN Treatment

I use postcode-level data on LTN placement during the Waltham Forest Pilot program. A postcode is the lowest level of geographic enumeration in the United Kingdom Census. There are 1.9 million postcodes in the United Kingdom, with the average postcode containing 15 properties, although this can vary with average lot size.[21] Figure 1.3 shows a map of postcodes in Waltham Forest, with each dot representing the geographical centroid of a postcode.[22] The first LTNs were implemented in September 2015, with subsequent zones were introduced in 2016, 2018, and 2019. Figure 1.4 shows the distribution of LTNs in Waltham Forest across time. For my main results that use price as the outcome variable, the unit of observation is an individual sale of a specific property, since I need to be able to account for unobserved characteristics using property-stack fixed effects. However, for a subset of specifications that use non-price variables as outcomes (e.g. quantity of sales), I am interested in measuring outcomes that occur at the neighborhood level and are not mapped to a single property. Because of the way that LTN boundaries were drawn, they overlap political and census boundaries and as a result there is no prescribed way to define neighborhoods that are controls for LTN neighborhoods. To create valid controls for LTN neighborhoods, I follow a process of defining neighborhoods polygons that uses main roads and geographic features to bound neighborhoods. This process and a graphic displaying the resulting polygons are shown in the appendix. For specifications using price as the outcome variable I also use information from the 2011 UK Census to obtain characteristics of the local area. These variables are defined at the output area (OA), a census-specific geography. In Waltham Forest, there are 732

[20] This method has been used by many papers in the urban literature to estimate the value of various place-specific amenities (see Gibbons & Manchin, 2008).

[21] The closest equivalent in the U.S. is the 9-digit ZIP Code.

[22] Areas in Figure 1.3 that do not contain postcodes are mostly natural features, mainly rivers and reservoirs in the west, and wooded areas in the east.

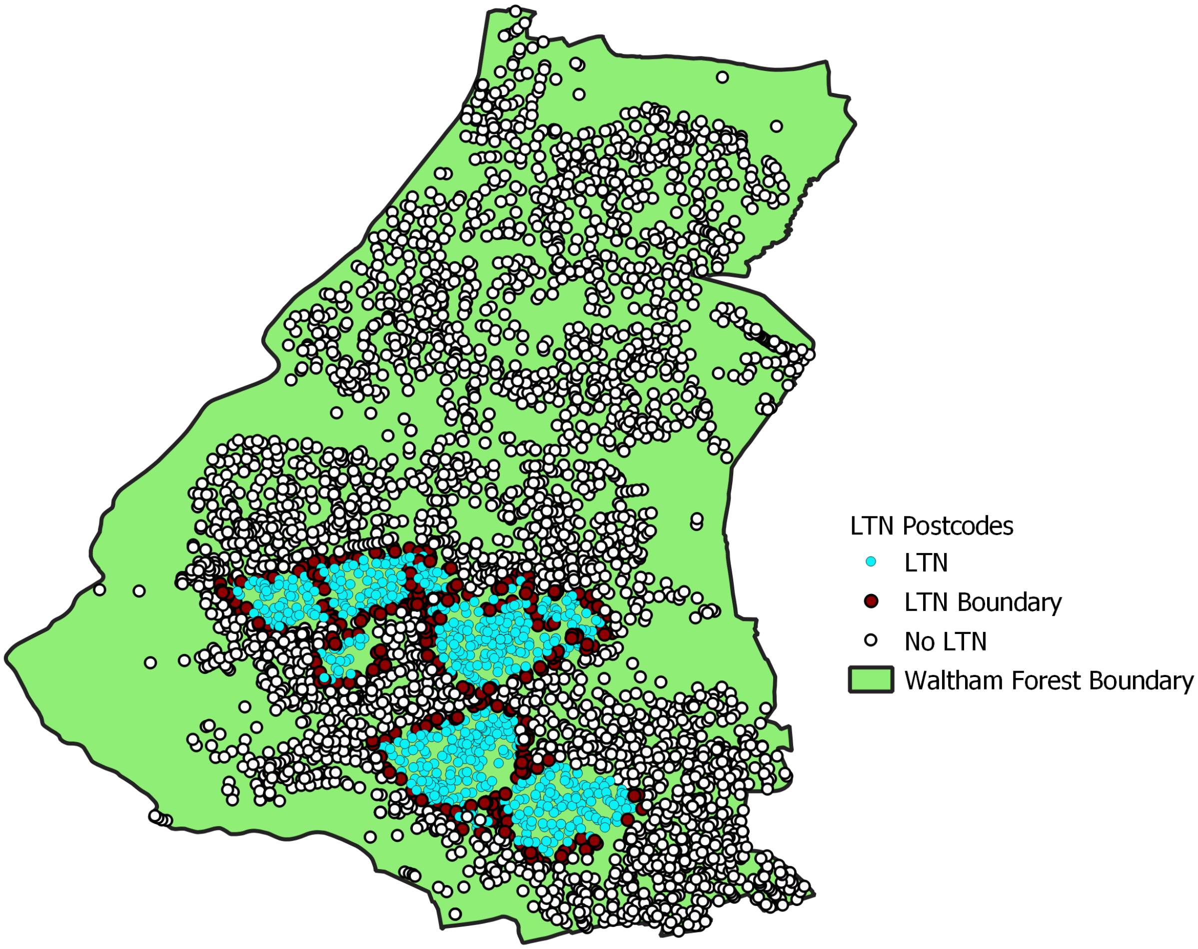


Figure 1.3: Location of LTNs in Waltham Forest implemented during Mini Holland pilot program (2015-2019), delineated at the postcode level

OAs containing 400 individuals on average. See Figure 1.11 in the appendix for a map displaying OAs in Waltham Forest.

1.4.2 House Price Data

I obtain data containing the near-universe of recorded property sales from *HM Land Registry* occurring between 2010 and 2019. Land Registry data is available up to the most recent month, but in order to avoid complicating the interpretation of results post-COVID, I end my analysis prior to 2020. I observe price, date of sale, full address of the property (processed via the Ordinance Survey's AddressBase), and some characteristics of the property, namely the type of property, the tenure (freehold vs. leasehold), and whether it involves

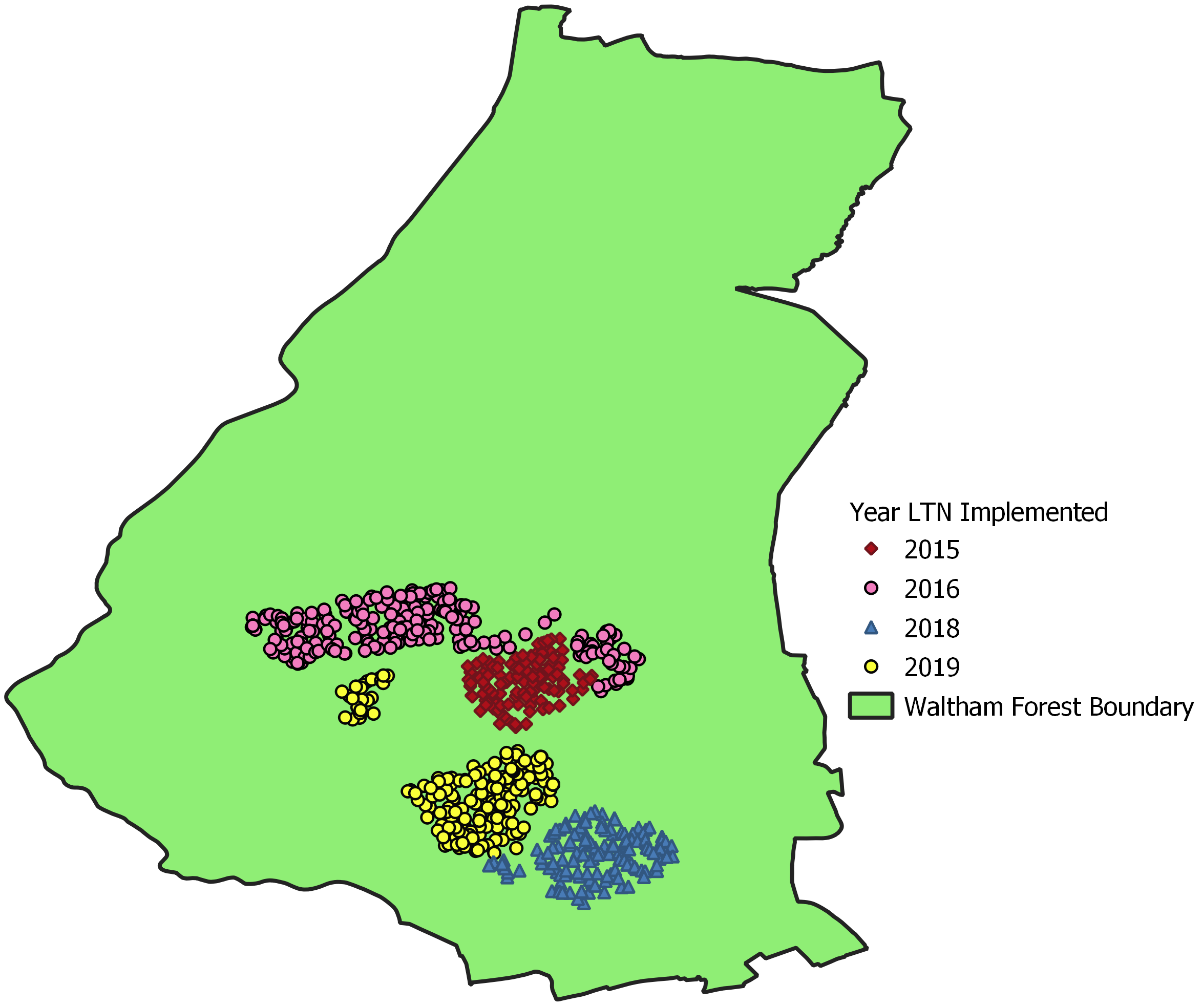


Figure 1.4: Year of LTN implementation in Waltham Forest, delineated at the postcode level.

new construction.[23] The data does not include commercial property sales or sales that occur through non-market means.[24] These records contain approximately 31,000 sales occurring in Waltham Forest between 2010 and 2019.[25]

Table 1.1 shows descriptive statistics for observations of house sales in Waltham Forest between 2010 and 2019, including sale price and property characteristics from *Land Registry*, and the corresponding neighborhood demographic characteristics from the 2011 UK Census, including a measure of economic deprivation, the presence of minor and elderly populations, and ethnic makeup. Prices throughout the chapter are adjusted to 2015GBP. Neighborhood characteristics are assigned to a sale by the Output Area that the property resides in. I divide neighborhoods into LTN, "Adjacent", and Controls. The LTN column is simply all treated neighborhoods. "Adjacent" refers to those that are never treated, but lie within 0.5 miles of an LTN, and thus are potentially impacted by spillover effects of treatment. Finally, controls refer to the remaining neighborhoods that are untreated and not sufficiently close to any LTNs. Average sales prices are similar between the three groups. However, there are some differences in other characteristics of sold properties: LTN sales are less likely to be detached properties, new builds, or on main roads. Control neighborhoods in Waltham Forest are also slightly less racially diverse than LTNs and adjacent areas.

1.5 LTNs and House Prices in Waltham Forest

Table 1.2 shows estimates from the stacked DID model shown in Equation (1.3). I limit the sample to the subset of properties that are transacted at least twice between 2010 and 2019. Column (1) does not include property-stack fixed effects or additional controls, and shows a large, positive effect, indicating that the introduction of an LTN increases sale prices by approximately 13%. Once I account for unobserved time-invariant differences in properties via the inclusion of property-stack fixed effects, the magnitude of this effect falls to 6.5% (Column (2)). This is my preferred specification, and based on the average price of ever-treated properties in Waltham Forest in the baseline period of 2010-2014 (£250,800), this corresponds to a premium of £16,302 that individuals are willing to pay to reside in an LTN. Finally, in Column (3) I perform the same regression but with the latest-treated stack (treated in 2019) dropped. This is done to be more directly comparable to the event study in Figure 1.5, where I drop the 2019 stack when estimating this event study in order to have more than 1 post-period year. The x-axis denotes the number of years since the LTN was implemented,

[23]Freehold means that the owner of the property also owns the land. Under a leasehold, ownership is held for a period usually between 40 and 120 years, and this lease must be renewed before expiration which is costly. Ground rents are usually also due annually to the landowner. Thus, leasehold properties have a lower market value than freehold properties. From 2019, the U.K. government has banned leaseholds for newly-built properties.

[24]Non-market sales could include transactions that are "not for value" such as gifted properties and administrative transfers, sales that occur through government schemes such as *Right to Buy*, and sales resulting from compulsory purchase orders.

[25]For context, Waltham Forest has a population of 275,000 (17th/32 boroughs in London) and is 15 square miles (16th/32) as of 2019.

Table 1.1: Selected Descriptive Statistics (Means) by Treatment Status

	(1) LTN	(2) Adjacent	(3) Control
Property Characteristics			
Price	£385,514	£364,582	£369,501
Flat	0.436	0.501	0.332
Detached	0.009	0.006	0.021
Semi-Detached	0.045	0.033	0.167
Terrace	0.494	0.410	0.462
New Build	0.014	0.080	0.073
Leasehold	0.541	0.442	0.658
On Main Road	0.016	0.127	0.090
Neighborhood Demographics			
Deprivation Percentile	48.9	43.8	55.4
HH w/ Dependent Child	0.334	0.340	0.328
Over 65 years old	0.081	0.079	0.125
White	0.482	0.479	0.644
Asian	0.257	0.240	0.135
Black	0.168	0.181	0.137
N	5,871	10,796	14,202

Table shows mean values of variables relating to properties and the neighborhood. Property characteristics come from the Land Registry data and neighborhood characteristics come from the 2011 UK Census and are enumerated at the Output Area level (see Figure 1.11 for map). Column (1) contains all properties that are ever within an LTN. Column (2) contains all properties that are never treated but fall within 0.5 miles of an LTN. Column (3) contains all propertiesthat are >0.5 miles from the nearest LTN. Prices are adjusted to 2015 pounds (£).

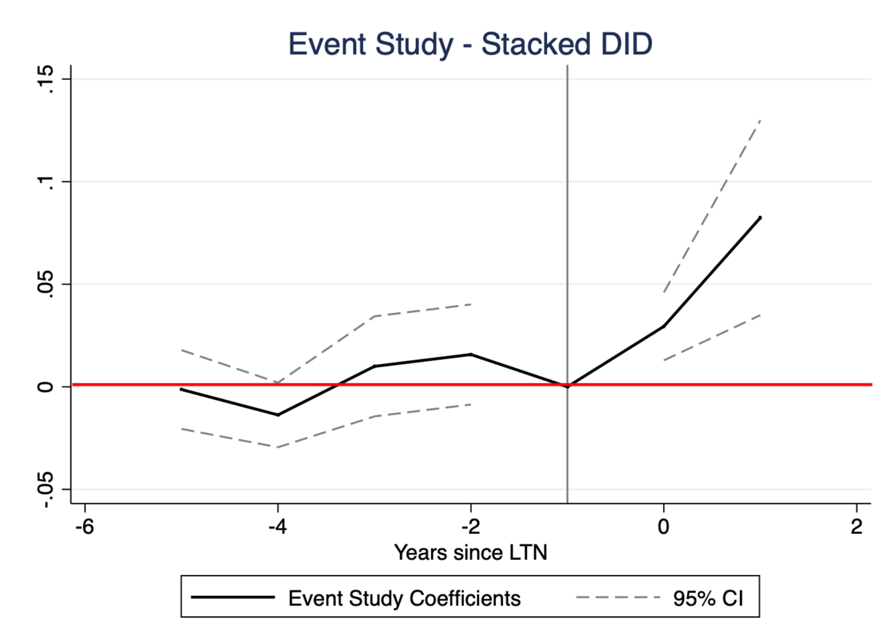
and the y-axis shows the point estimate $\beta_1^{stacked}$. The stacked DID estimate is not statistically different when dropping the 2019 stack.

Table 1.3 presents estimates from a series of alternative specifications to test the robustness of the main result. Column (1) shows estimates from the preferred specification (i.e. Table 1.2, Column (2)) with fixed effects and neighborhood controls. In the baseline results, I do not include properties that are located on major/main roads, since LTNs are interventions that target residential areas, so excluding them means I am comparing only residential areas in treated and control neighborhoods. Column (2) adds back in properties located on major roads, which mainly serves to change the composition of the control group. Column (3) restricts my control group to only neighborhoods that are never treated by an LTN. Column (4) uses the whole sample of house sales rather than only properties that sold multiple times, but omits property-stack fixed effects from the specification. Finally, Column (5) implements the Callaway & Sant'Anna (2021) estimator that uses a weighted average of group-time average treatment effects using inverse probability weighting. Point estimates in Columns (2)-(5) do vary, but none of them are statistically distinguishable from my preferred specification. Figure 1.12 in the appendix shows the corresponding event study for the Callaway & Sant'Anna (2021) specification from Table 1.3, Column (5).

Table 1.2: Effects on House Prices

	(1)	(2)	(3)
LTN Treat	0.132^{***}	0.065^{***}	0.059^{***}
	(0.0303)	(0.0159)	(0.0166)
Multiple Sales Sample	X	X	X
Property-Stack Fixed Effects		X	X
Drop 2019 Stack			X
N	43,401	43,401	35,701

Standard errors in parentheses and clustered at the property level.
Estimates are variations on the stacked DID design in Equation (1.3).
To allow me to account for unobserved differences at the property level, all specifications use the subsample of properties that sell multiple times in my sample period. Column (1) is the baseline estimate without fixed effects or controls. Column (2) adds the property-stack fixed effects. Column (3) excludes the last (2019) stack to match with the event study in Figure 1.5.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$



Notes: This figure depicts the event study from the stacked difference in differences model (Equation 1.3) comparing the sales prices of properties in areas that receive an LTN compared to properties in areas that do not. The latest-treated stack is dropped from the event study because there is only one year of observed post-treatment data for this stack. The figure shows coefficients and the 95% confidence intervals from the regression where the outcome variable is the natural log of house sale price adjusted to 2015 pounds (£). This specification includes property-stack and year-stack fixed effects. Standard errors are clustered at the property level.

Figure 1.5: Event Study of Stacked DID

Next, I investigate if there is any evidence that LTN treatment causes negative externalities to be displaced to nearby neighborhoods, a common argument of LTN critics, using two distinct strategies: First, I estimate specifications where I drop all ever-treated neighborhoods and assign a psuedo-treatment to adjacent neighborhoods with treatment timing based on when the LTN neighborhoods they are adjacent to were treated. I use 0.5 miles as a threshold for a neighborhood to qualify as adjacent, with neighborhoods more than 0.5 miles from the nearest LTN serving as the control group. If there was displacement of congestion-related externalities to immediate surrounding areas occurring, then we would expect to see a negative effect of this psuedo-treatment. Table 1.4 shows that there is no significant evidence of this being the case. In fact, the point estimates are positive, although too imprecise to make any statistical conclusions.

Second, I attempt to capture both direct treatment effects and spillover effect on nearby neighborhoods in a single regression by dividing properties into rings based on how far they are located from the nearest LTN.

Table 1.3: Robustness Tests and Callaway & Sant'Anna (2021) Specification

	(1)	(2)	(3)	(4)	(5)
LTN Treat	0.065***	0.073***	0.076***	0.030***	0.087**
	(0.0159)	(0.0171)	(0.0193)	(0.0084)	(0.0398)
Stacked DID	X	X	X	X	
Callaway & Sant'Anna (2021)					X
Multiple Sales Sample	X	X	X		X
Property-Stack Fixed Effects	X	X	X		
Include Major Roads		X			
Control: Never Treated Only			X		
Full Sample				X	
Property Fixed Effects				X	X
N	43,401	47,417	38,200	129,389	9,519

Standard errors in parentheses and clustered at the property level. Column (1) presents the preferred specification from Table 1.2 that includes property-stack FE (Column (2)).
Column (2) expands the set of controls by including properties that have postcodes corresponding to major roads (rather than residential ones). Column (3) limits the controls to neighborhoods that are never treated. Column (4) uses the full sample, including properties that are not sold multiple times. Finally, Column (5) implements the Callaway & Sant'Anna (2021) approach using a repeated cross-section and property fixed-effects.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 1.6 shows this division. The inner-ring contains ever-treated postcodes. The second ring is untreated postcodes within 0.5 miles of an LTN with properties in this ring being the most likely candidates for spillover effects. This ring is also the set of postcodes counted as treated in the analysis shown in Table 1.4. The third ring contains postcodes between 0.5 miles and 1.5 miles from an LTN, and the outer ring contains the controls more than 1.5 miles away, and the least likely (in theory) to be affected by spillovers. I follow the method described in Butts (2023) for analyzing spillover effects with multiple concentric rings by estimating the following equation:

$$ln(y_{pts}) = \beta_0 + \beta_1^{stacked} LTN\_Treat_{pts} + \gamma_1 Ring0.5_{pts} + \gamma_2 Ring1.5_{pts} + \lambda_{ps} + \delta_{ts} + \varepsilon_{pts} \tag{1.4}$$

Where $Ring0.5_{pts}$ is an indicator that turns on if property p is never-treated but within the $(0, 0.5]$ mile ring, and the corresponding LTN has been implemented as of t. The same applies for $Ring1.5_{pts}$, but for observations in the $(0.5, 1.5]$ mile ring. γ_1 and γ_2 measure the average spillover effect onto untreated properties in their respective rings. The outermost ring ($>$ 1.5 miles) serves as a reference group, and $\beta_1^{stacked}$ will be an unbiased estimate of the total treatment effect under the condition that the reference group does not experience spillover effects. The results of estimating Equation (1.4) are shown in Table 1.10 in the appendix, where estimates of $\beta_1^{stacked}$, γ_1, and γ_2 are shown with the same three specifications as the main results (Table

1.2). In both specifications that include property-stack fixed effects, Columns (2) and (3), $\beta_1^{stacked}$ is still positive and significant. The magnitudes iare larger than the estimates from the main specification in Table 1.2, but not statistically different due to the size of the standard errors. I find that the $(0,0.5]$ ring does exhibit significant a positive spillover in this specification, which matches the positive sign in Table 1.4. This result provides further evidence against the hypothesis that nearby neighborhoods suffer a welfare loss, and that there appears to be some perceived benefit to living near a low-traffic area, which results in a net-appreciation of prices.[26] Once we move further than 0.5 miles, this effect seems to dissipate, as the estimates of γ_2 are very close to 0.

Finally, in order to ascribe changes in price of properties to changes in willingness to pay, the hedonic approach requires that supply is relatively inelastic over the period being studied, or else changes in WTP will be at least partially capitalized into the quantity of housing rather than the price. One way to test this assumption is to see if the number of properties sold changes in treated areas relative to control areas (as defined by polygons, see appendix). Table 1.5 shows the results of this regression. I find no evidence that the volume of sales is changing differentially in LTN neighborhoods compared to untreated ones as a result of treatment.

Table 1.4: Investigating Spillover Effects

	(1)	(2)	(3)
LTN Treat	0.094	0.072	0.060
	(0.0611)	(0.0546)	(0.0451)
Multiple Sales Sample	X	X	X
Property-Stack Fixed Effects		X	X
Drop 2019 Stack			X
N	36,417	36,417	29,862

Standard errors in parentheses and clustered at the property level.
Estimates are variations on the stacked DID design in Equation (1.3).
Neighborhoods within 0.5 miles are psuedo-treated, and ever-treated units are dropped from the sample. To allow me to account for unobserved differences at the property level, all specifications use the subsample of properties that sell multiple times in my sample period. Column (1) is the baseline estimate without fixed effects or controls. Column (2) adds property-stack fixed effects. Column (3) excludes the 2019 stack.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

[26]This positive effect does not rule out the possibility that there is some negative element (e.g. rerouted traffic) to LTNs in these rings, but in this case, the positive effects would dominate.

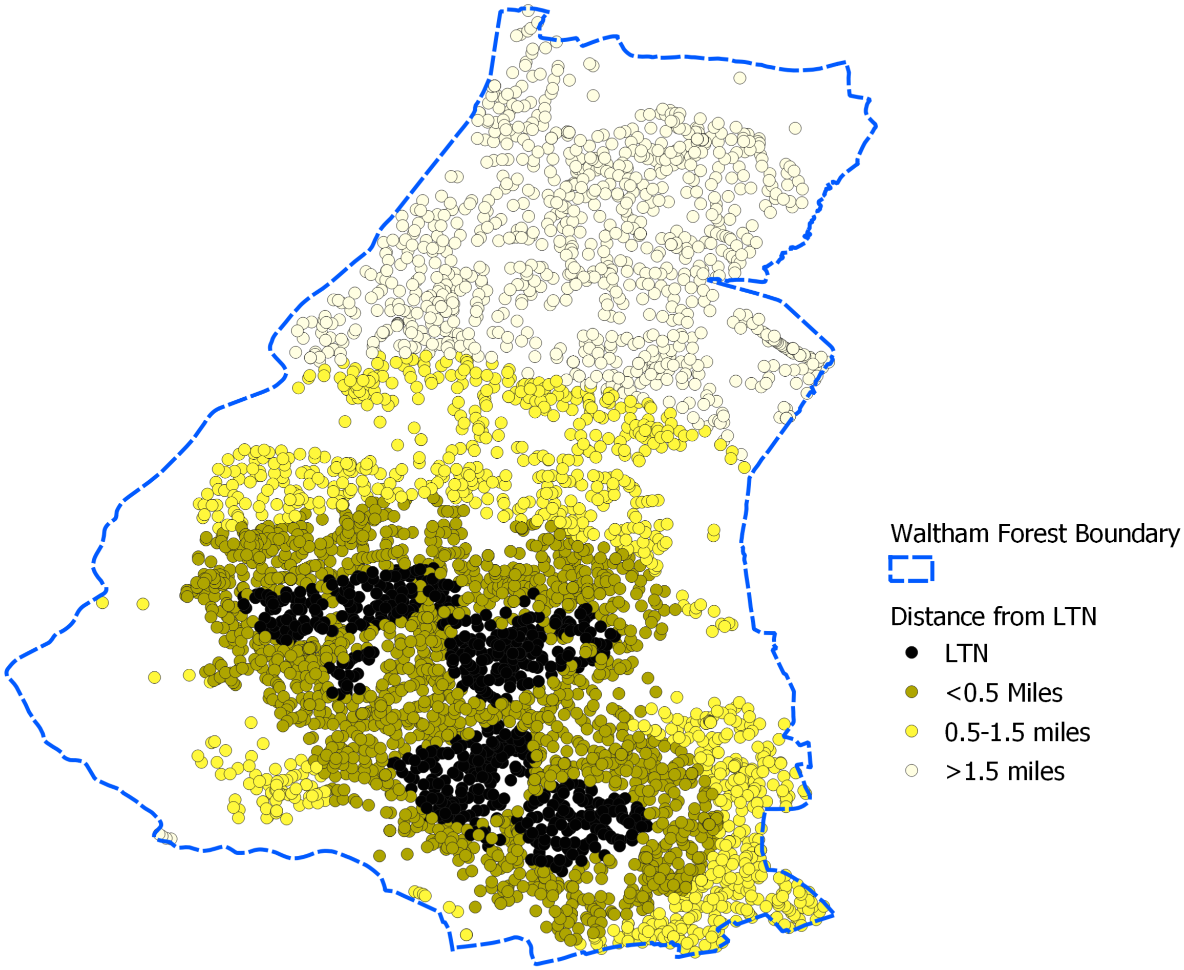


Figure 1.6: Map of postcodes showing corresponding ring classification.

Table 1.5: Effect on Supply: Number of Sales

	(1)	(2)
LTN Treat	-1.76	-1.55
	(1.781)	(1.968)
Exclude Adjacent Nhds		X
Mean	21.0	22.8
N	1,143	759

Standard errors are in parentheses and clustered at the polygon level. Regressions are panel versions of the Callaway & Sant'Anna (2021) method from Table 1.3, Column (5), with the number of sales on the LHS. Column (2) excludes any polygon that is adjacent to an ever-treated polygon.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1.6 Heterogeneity & Local Amenities

1.6.1 Heterogeneity

I investigate whether there is significant heterogeneity in my main estimates of the effect of LTNs on sales prices. If the positive effect is driven by certain property types or by neighborhoods with certain demographics, then it could indicate that the increase in implied willingness to pay is driven by certain subgroups. Table 1.6 shows the results from cutting the sample by some selected property and neighborhood characteristics. Columns (1) and (2) compare flats with houses. Houses, which are typically larger and more desirable, show a larger point estimate. Speculatively, this could be a result of the difference in preferences for LTNs and their associated amenities among households who are able to purchase houses in London. However, this difference is marginally insignificant at the 5% level. Columns (3) and (4) compare freehold and leasehold ownership status,[27] since changes in the price of a leasehold property may not fully reflect changes in the value of the land. However, this does not appear to be the case, as the point estimates are virtually identical. Finally, Columns (5) and (6) compare the treatment effect in neighborhoods that have a relatively high (above median) proportion of households with dependent children compared to those with a relatively low (below median) proportion, which may be relevant if adults with children value LTNs differently due to preferences for lower traffic, accident/crime risks, etc.. The estimated effect of LTNs on sale prices is almost double in areas that have an above-average prevalence of households with children, compared to below-average areas, although I am unable to say that their magnitudes are statistically different based on 95% confidence intervals.

1.6.2 Traffic Accidents

Data on traffic accidents comes from the *Department of Transport* (DfT). This data contains all accidents that (i) resulted in at least one personal injury and (ii) were reported to the police. Similar to housing sale data, these records contain the exact geographic coordinates of the accident. These data also contain severity of the accident, number of casualties, and variables relating to road conditions. I use data from 2010-2019; in this time-span, there are approximately $6,500$ such incidents recorded for Waltham Forest.

One of the benefits of LTNs are that they reduce (through) traffic, which may lead to lower vehicle accident rates. Table 1.7 shows estimates from regressions with the number of traffic accidents resulting in notable injury as the outcome. Regressions are specified similar to Table 1.5; a panel regression at the polygon level with standard errors clustered at the polygon-level. Estimates show a reduction of almost two serious accident per-polygon per-year on residential streets. This is roughly a 60% decrease in serious accidents.

[27] Freehold means that the homeowner also owns the land the property sits on, whereas leaseholders do not.

Given the total premium associated with living in a Waltham Forest LTN is £16,302 based on my preferred specification (Table 1.2, Column 2) and baseline period mean prices in ever-treated areas, I utilize statistics published by the *Department of Transport* on the costs associated with serious vehicle accidents to estimate the fraction of the increase in WTP that can be attributed to the reduction in the probability of being involved in such an accident. By making simplifying assumptions regarding the distribution of risk reduction and risk preferences,[28] I calculate that the net present value (NPV) associated with the reduced risk of accidents is £1,736.10 per household, which accounts for roughly 10.6% of the total premium. See the appendix for the details on how this number was reached.

Table 1.6: Heterogeneity by Property Characteristics and Neighborhood Attributes

	Flats	Houses	Leasehold	Freehold	High % Child	Low % Child
LTN Treat	0.035^{***}	0.091^{***}	0.068^{***}	0.069^{***}	0.096^{***}	0.053^{**}
	(0.0126)	(0.0239)	(0.0214)	(0.0202)	(0.0310)	(0.0238)
N	19,465	22,306	21,338	21,118	21,009	22,392
Difference	-0.056^{**}		-0.001		0.043	
P-Value	0.038		0.973		0.271	

Standard errors in parentheses and clustered at the property level. Leasehold and freehold refers to ownership of land the property sits on, with a freehold commanding a premium all other things equal. High/low dep. child nhd. refers to the percentage of households in the neighborhood (OA) that contain dependent children. Column (5) restricts the sample to properties in OAs that have above-median (> 33%) rates of households with dependent children, while Column (6) is the subset that has below-median rates of such households. The last two rows show the difference in point estimates between Columns (1/2), (3/4) and (5/6), and the corresponding p-values from a z-test of equality of coefficients.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.7: Effect of LTN on Traffic Accidents

	(1)
LTN Treat	-1.954^{**}
	(0.7814)
Outcome Mean	3.3
N	1,175

Standard errors in parentheses and clustered at the polygon level. Regressions are panel versions of the Callaway and Sant'Anna (2021) approach at the polygon-level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

[28] Specifically, I assume that households are risk neutral, have a 5% discount rate, and that the full benefit from reduction in accidents is spread out evenly among residents. See the appendix for a more detailed description.

1.6.3 Crime

Another local amenity that may be affected by the presence of an LTN is crime. Becker's (1968) cost of crime model posits that a rational criminal will consider the expected benefit gained from attempting a crime versus the cost (i.e. likelihood of getting caught and subsequent punishment) when deciding on whether to commit a crime. LTNs reduce the permeability of the neighborhood, which ex-ante has an ambiguous effect on crime: On one hand, lower permeability reduces access into and out of the neighborhood for criminals, and there is suggestive evidence in other settings that links this with lower crime (Cozens & Love, 2009). However, sacrificing permeability could come at the expense of "eyes on the street" (Jacobs, 1961) which could incentivize certain types of crimes. Additionally, the elimination of through-traffic could result in fewer vehicle-related crimes, although it is unclear to what extent this effect would be mitigated by substitution to other crimes.

To investigate the relationship between neighborhood LTN status and crime, I use open-source data on crime reports from 2011-2019 from the London Metropolitan Police. These data contain the precise location of the crime, month that the crime occurred, and the category of crime.[29] Table 1.8 shows the results of estimating Poisson regressions in the stacked DID framework in which crime count is the dependent variable. Column (1) uses the aggregate of all crimes reported, regardless of crime type. There is no significant effect of the LTN status of a neighborhood on the total number of crimes reported.[30] Columns (2) and (3) focus on less serious "antisocial" crimes and more serious crimes such as assault. There is no significant effect of LTNs on the incidence of crimes in either of these categories. Columns (4) and (5) compare the subset 'opportunistic' crimes such as street thefts and shoplifting to 'non-opportunistic' crimes that are pre-meditated such as carrying a weapon or drug charges. Consistent with Becker (1961), I find that only the opportunistic subcategory is significantly affected by LTNs.[31] However, this effect seems to be offset by a similar-sized decrease in vehicle-related crimes, which includes theft from or of vehicles, damage to a vehicle, or interference with a vehicle. This is expected given the mechanical effect of LTNs in preventing through-traffic.[32]

[29] To protect anonymity of victims, the coordinates reported in the data are very slightly altered. For example, recording the location as the middle of the street if the crime occurred in someone's house, which may reveal the identity of the victim.

[30] This is different to Goodman & Aldred (2021), who found close to a 10% decline in overall crime, but do not account for variation in timing of treatment. Although it is an obvious initial outcome to examine, total crime may not be as illustrative as looking at more disaggregated categories, since this measure implies that all crimes have equal weight, regardless of type, whereas we might instead care much more about a serious crime like assault than an incident of public drunkenness.

[31] However, due to the relatively large magnitude of the standard errors in these regressions, I am unable to definitively rule out equality of these coefficients at the 95% level.

[32] Although I cannot directly observe traffic flows in these neighborhoods, Yang et al. (2022) look at three LTNs in Islington, UK and find that traffic flows decrease by over 50% after treatment.

Table 1.8: Effect of LTNs on Crime

	All	Antisocial	Serious	'Non-Opp'	'Opp'	Vehicle
LTN Treat	-0.0119	0.0011	-0.0447	0.0075	0.1512^{**}	-0.1116^{***}
	(0.0426)	(0.0659)	(0.0293)	(0.0858)	(0.0759)	(0.0424)
N	279,036	279,036	279,036	279,036	279,036	279,036

Standard errors in parentheses and clustered at the output area. Parameters are from Poisson pseudo-maximum likelihood regressions with multiple fixed-effects. They are analogous to the baseline results but using count data as the outcome. Each column uses a different subset of crimes as the outcome. Antisocial refers to lower-level crimes such as public intoxication and disorder charges. Serious crimes include violent assaults, killings, and sexual assaults. Non-opportunistic ('non-opp') crimes include categories such as drugs and weapons charges, whereas opportunistic ('opp') crimes are things like shoplifting, theft of bikes, and muggings.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1.7 Discussion

In this chapter, I test for a causal relationship between the LTN status of a neighborhood and the willingness to pay to live in that neighborhood, as measured by recorded sales prices of residential properties. I use a difference in differences framework that exploits geographic and temporal variation in LTN status due to an LTN pilot program in the London Borough of Waltham Forest. My main result is that the implementation of an LTN results in a 6.5% sales price premium. This result is robust to different specifications. My analysis suggests that this effect is likely driven by improvements in amenities in treated areas, rather than by disamentities spilling over to nearby untreated areas. This indicates that concerns regarding negative externalities being passed-on to immediately-adjacent neighborhoods are not realized. I also look at a number of local amenities that may be driving some of the increase in WTP: I find a large reduction in serious road accidents, which I estimate to account for over 10% of the total effect, but no meaningful effects on crime, contrary to prior analysis.

These results have relevance for urban planning policy: I provide the first set of causal results that uncover revealed preferences for LTNs. Existing evidence on preferences for this type of traffic calming policy has been limited to stated preferences through surveys, polls, and consultations (e.g. Logan et al. 2021). However, residents may be incentivized to misrepresent their preferences since i) LTNs are a politically charged and divisive topic, and ii) LTNs are local policies, so individuals have more of a chance to influence decision-makers. Therefore, revealed preferences are important to good decision-making for policymakers in this space.

In the United Kingdom, LTNs have been contentious when first proposed and introduced, although anecdotal reports suggest that vocal opposition somewhat wanes over time (Bosetti et al., 2022). LTNs came to national attention in 2020, when residents were prohibited from traveling outside their immediate neighborhoods for non-essential reasons due to COVID-19. This led to an increased focus on the quality of amenities in proximity to one's home. To that end, 87 LTNs were implemented across all of London's boroughs between March 2020 and December 2020, and many more across the country Just like the Waltham Forest LTNs, there has been a mixed reaction to their presence, with some claiming that they are a poor use of public funds, or that they cause disruption to the road network and people that rely on car travel. There have also been cases of organized protests against LTNs and candidates for local offices running on an anti-LTN platform (e.g. Kouimtsidis, 2021). As a result, some post-COVID LTNs are still in operation, while others have been removed (e.g. Burford, 2021). The debate over LTNs has become prominent enough that the U.K. government led by Rishi Sunak has started to intervene: The 2022/23 edition of the Active Travel Fund, a major grant program run by the Department of Transport, explicitly excluded LTNs from the list of eligible projects over concerns about their downsides (Ames, 2023). Additionally, in July 2023, Sunak ordered the Department for Transport to carry out a review of LTNs, citing concerns that they were too "anti-motorist" (Walker, 2023).[33]

The size of the effects on house prices found in this chapter are fairly large when compared to related research looking at other policies that reduce traffic flows e.g. (Ossokina & Verweij, 2015; Polloni, 2019). It should be noted that the London borough of Waltham Forest selected neighborhoods to receive LTNs based on their promise to be implemented successfully (i.e. to discourage car use and to encourage walking/cycling). Therefore, one might not expect the same magnitude of effect if a hypothetical LTN was implemented in a random residential neighborhood. One factor mentioned by the borough was that the propensity of traffic calming to encourage walking and cycling may be higher if there are other relevant amenities nearby such as traditional village high-streets or urban parks. This idea is related to work by Albouy et al. (2020) on complementarity between public good. In the case of LTNs, the effect of an LTN on the desirability of a neighborhood may depend on existing infrastructure.

A current limitation of research into LTNs and other hyper-local urban planning policies is that getting data at both a fine geographic scale and a sufficiently high frequency can be difficult. An example of this in my setting is pollution, which is both an outcome of interest and often one of the main selling points of LTNs during the proposal stage. Traffic counts and transport modalities are also areas where current publicly-available

[33] It should be noted that there are likely political motivations behind the intervention from central government, which is currently under Conservative Party control, whereas a majority of London boroughs are controlled by the Labour Party.

data is sparse. Future research with these types of data could provide more insight into the effects of LTNs at a neighborhood level.

1.8 References

Albouy, D., Christensen, P., & Sarmiento-Barbieri, I. (2020). Unlocking amenities: Estimating public-good complementarity. Journal of Public Economics, 182(104110).

Anderson, M. L. (2014). Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion. American Economic Review, 104 (9), 2763–2796.

Aldred, Rachel, Ersilia Verlinghieri, Megan Sharkey, Irena Itova, & Anna Goodman (2021). Equity in new active travel infrastructure: A spatial analysis of London's new Low Traffic Neighbourhoods, Journal of Transport Geography, Volume 96, 2021, 103194, ISSN 0966-6923,

Aldred, Rachel, and Anna Goodman. 2020. "Low Traffic Neighbourhoods, Car Use, and Active Travel: Evidence from the People and Places Survey of Outer London Active Travel Interventions." Findings, September. https://doi.org/10.32866/001c.17128.

Ames, Chris (2023). £200m active travel cash allocated but LTNs blocked. *Highways Magazine*. Published May 22, 2023.

Asabere, Paul K. (1990): The Value of a Neighborhood Street with Reference to the Culde-Sac. Journal of Real Estate Finance and Economics, 3:185-193 (1990)

Audretsch, David B, Dirk Christian Dohse,& João Pereira dos Santos (2020). The effects of highway tolls on private business activity—results from a natural experiment, Journal of Economic Geography, Volume 20, Issue 6, November 2020, Pages 1331–1357, https://doi.org/10.1093/jeg/lbaa003

Bagby, D. Gordon (1980) The Effects of Traffic Flow on Residential Property Values, Journal of the American Planning Association, 46:1, 88-94, DOI: 10.1080/01944368008977020

Bateman, I.J., Day, B.H., Lake, I.R., & Lovett, A.A. (2006). The Effect of Road Traffic on Residential Property Values: A Literature Review and Hedonic Pricing Study.

Becker, G. S. (1968). Crime and Punishment: An Economic Approach. Journal of Political Economy, 76(2), 169–217. http://www.jstor.org/stable/1830482

Bosetti, Nicolas Kieran Connelly, Claire Harding and Denean Rowe (2022). Street Shift: The Future of Low-Traffic Neighbourhoods. Centre for London, June 2022

Braakmann, N. (2017). The link between crime risk and property prices in England and Wales: Evidence from street-level data. Urban Studies, 54(8), 1990–2007. https://doi.org/10.1177/0042098016634611

Burford, Rachael (2021). London borough of Ealing to rip out Low Traffic Neighbourhoods after resident 'referendum'. August 19, 2021.

Butts, Kyle (2023). Difference-in-Differences Estimation with Spatial Spillovers. arXiv:2105.03737v3 [econ.EM]. https://doi.org/10.48550/arXiv.2105.03737

Cabannes et al. (2018). The Impact of GPS-Enabled Shortest Path Routing on Mobility: A Game Theoretic Approach. Transportation Research Board 97th Annual Meeting Location: Washington DC, United States Date: 2018-1-7 to 2018-1-11

Callaway, Brantly & Pedro H.C. Sant'Anna (2021). Difference-in-Differences with multiple time periods, Journal of Econometrics, Volume 225, Issue 2, 2021, Pages 200-230, ISSN 0304-4076,

Carrillo, Paul E., Andrea Lopez-Luzuriaga, & Arun S. Malik (2018) Pollution or crime: The effect of driving restrictions on criminal activity, Journal of Public Economics, Volume 164, 2018, Pages 50-69.

Carrillo, P.E., Malik, A.S. and Yoo, Y. (2016), Driving restrictions that work? Quito's Pico y Placa Program. Canadian Journal of Economics/Revue canadienne d'économique, 49: 1536-1568.

Ceccato Vania, & Mats Wilhelmsson (2020) Do crime hot spots affect housing prices?, Nordic Journal of Criminology, 21:1, 84-102.
DOI: 10.1080/2578983X.2019.1662595

Chiaradia, Alain, Bill Hillier, Christian Schwander, and Yolande Barnes (2013). Compositional and urban form effects on residential property value patterns in Greater London Proceedings of the Institution of Civil Engineers - Urban Design and Planning 2013 166:3, 176-199

Collins, Courtney A., and Erin K. Kaplan. 2017. "Capitalization of School Quality in Housing Prices: Evidence from Boundary Changes in Shelby County, Tennessee." American Economic Review, 107 (5): 628-32.

Conceição, Marta Aranha, Mayara Moraes Monteiro, Dena Kasraian, Pauline van den Berg, Sonja Haustein, Inês Alves, Carlos Lima Azevedo & Bruno Miranda (2023) The effect of transport infrastructure, congestion and reliability on mental wellbeing: a systematic review of empirical studies, Transport Reviews, 43:2, 264-302, DOI: 10.1080/01441647.2022.2100943

Currie, Janet, and Reed Walker. 2011. "Traffic Congestion and Infant Health: Evidence from E-ZPass." American Economic Journal: Applied Economics, 3 (1): 65-90.

Davis, L. (2011). The effect of power plants on local housing values and rents. Review of Economics and Statistics, 93(4), 1391-1402.

Deshpande, Manasi, and Yue Li. "Who is screened out? Application costs and the targeting of disability programs." American Economic Journal: Economic Policy 11, no. 4 (2019): 213-48.

Department for Transport (DfT) (2020a). Road traffic estimates in Great Britain: 2019. https://www.gov.uk/government/statistics/road-traffic-estimates-in-great-britain-2019

Department for Transport (DfT) (2020b). Case study London Mini Hollands. https://www.gov.uk/government/case-studies/london-mini-hollands. Last accessed: January 29, 2023.

Diamond, Douglas B., & George S. Tolley (1982). The Economics of Urban Amenities. Academic Press, 111 5th Avenue, New York, NY 10003.

Diao, M., Kong, H. & Zhao, J. (2021) Impacts of transportation network companies on urban mobility. Nat Sustain 4, 494–500 (2021).
https://doi.org/10.1038/s41893-020-00678-z

Downs, Anthony (2004). Traffic: Why It's Getting Worse, What Government Can Do. *The Brookings Institution* Policy Brief #128. January 2004.

Gibbons & Manchin (2008). Valuing school quality, better transport, and lower crime: evidence from house prices. Oxford Review of Economic Policy, Volume 24, Number 1, 2008, pp.99–119

Goodman, Anna, and Rachel Aldred. 2021. "The Impact of Introducing a Low Traffic Neighbourhood on Street Crime, in Waltham Forest, London." Findings, February. https://doi.org/10.32866/001c.19414.

Goodman, Anna, Anthony A. Laverty, Jamie Furlong, and Rachel Aldred. 2023. "The Impact of 2020 Low Traffic Neighbourhoods on Levels of Car/Van Driving among Residents: Findings from Lambeth, London, UK." Findings, June. https://doi.org/10.32866/001c.75470.

Goodman-Bacon, Andrew (2021). Difference-in-differences with variation in treatment timing, Journal of Econometrics, Volume 225, Issue 2, 2021, Pages 254-277, ISSN 0304-4076, https://doi.org/10.1016/j.jeconom.2021.03.014.

Goodwin, Michael B., Matias Fontenla & Fidel Gonzalez (2021) Estimating the impact of pollution on wages and housing prices using satellite imagery, Applied Economics Letters, 28:20, 1750-1753, DOI: 10.1080/13504851.2020.1853665

Greenstone, M. (2017). The Continuing Impact of Sherwin Rosen's Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. Journal of Political Economy. https://doi.org/10.1086/694645

Gronau, R. (1999). The Economics of a Single Toll Road in a Toll-Free Environment. Journal of Transport Economics and Policy, 33(2), 163–172. http://www.jstor.org/stable/20053803

Hughes, W.T., Jr. and Sirmans, C.F. (1992), Traffic Externalitites and Single-Family House Prices. Journal of Regional Science, 32: 487-500. https://doi.org/10.1111/j.1467-9787.1992.tb00201.x

Ihlanfeldt, Keith, & Tom Mayock (2010a). Panel data estimates of the effects of different types of crime on housing prices, Regional Science and Urban Economics, Volume 40, Issues 2–3, 2010, Pages 161-172, ISSN 0166-0462, Ihlanfeldt, Keith, & Tom Mayock (2010b) "Property Values and Crime," in Handbook on the Economics of Crime, edited by B. Benson and P. Zimmerman, Edward Elgar, 2010.

Ioulia V. Ossokina & Gerard Verweij (2015) Urban traffic externalities: Quasi-experimental evidence from housing prices, Regional Science and Urban Economics, Volume 55, 2015, Pages 1-13, ISSN 0166-0462,

Jacobs, J. (1961) 2000. The Death and Life of Great American Cities. London: Pimlico.

Jerch Rhiannon , Panle Jia Barwick, Shanjun Li, and Jing Wu (2023) The Impact of Road Rationing on Housing Demand and Sorting. (Working Paper).

Knittel, C. R., Miller, D. L. and Sanders, N. J. (2016). Caution, Drivers! Children Present: Traffic, Pollution, and Infant Health. The Review of Economics and Statistics, 98 (2), 350–366

Kojima, A., Elfferding, S. & Kubota, H. Intelligent Rat-Runners: Impact of Car Navigation Systems on Safety of Residential Roads. Int. J. ITS Res. 13, 9–16 (2015). https://doi.org/10.1007/s13177-013-0075-7

Kouimtsidis, Dimitris (2021). GALLERY: Anti-LTN protest march in Ealing. https://ealing.nub.news/news/local-news/gallery-anti-ltn-protest-march-in-ealing. Accessed 01-29-23

Kroes, E. P., & Sheldon, R. J. (1988). Stated Preference Methods: An Introduction. Journal of Transport Economics and Policy, 22(1), 11–25.
http://www.jstor.org/stable/20052832

Laverty, A., Aldred, R. & Goodman, A. (2021). "The Impact of Introducing Low Traffic Neighbourhoods on Road Traffic Injuries." Findings, January. https://doi.org/10.32866/001c.18330.

Logan, Tracy, Rob McPhedran, Amy Young, & Emily King (2021). Low Traffic Neighbourhoods Residents' Survey. Kantar, January 2021.

London Living Streets (2019). A guide to low traffic neighborhoods. September 2019.

Malnick, Edward (2023). “Mark Harper calls for review of unpopular LTNs” *The Telegraph*, July 8, 2023.

McIlhatton, D., McGreal, W., Taltavul de la Paz, P. and Adair, A. (2016), ”Impact of crime on spatial analysis of house prices: evidence from a UK city”, International Journal of Housing Markets and Analysis, Vol. 9 No. 4, pp. 627-647. https://doi.org/10.1108/IJHMA-10-2015-0065

Newman, O. 1973. Defensible Space: People & Design in the Violent City. London: Architectural Press.

Ossokina, Ioulia V. & Gerard Verweij (2015). Urban traffic externalities: Quasi-experimental evidence from housing prices, Regional Science and Urban Economics, Volume 55, 2015, Pages 1-13, ISSN 0166-0462.

Polloni, S. (2019). Traffic calming and neighborhood livability: Evidence from housing prices in Portland. Regional Science and Urban Economics 74 (pp. 18-37).

Reynolds, Mark (2023). “Angry protesters set fire to hated LTN planters just hours after being installed” *Express*, March 28, 2023.

Rosen, S. 1974. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. J.P.E. 82 (1): 34–55.

Sant’Anna and Zhao (2020). Doubly robust difference-in-differences estimators, Journal of Econometrics, Volume 219, Issue 1, November 2020, Pages 101-122

Seaward, Tom (2023). “Anti-LTN and ’climate lockdown’ protest brings thousands to Oxford”. *Oxford Mail*, February 19, 2023.

Sheppard, Stephen (1999). Chapter 41: Hedonic analysis of housing markets, Handbook of Regional and Urban Economics, Elsevier, Volume 3, 1999, Pages 1595-1635,

Sustrans (2020). What is a low traffic neighbourhood? Published November 2nd, 2020. Available at: https://www.sustrans.org.uk/our-blog/get-active/2020/in-your-community/what-is-a-low-traffic-neighbourhood

Tang, Cheng Keat (2021). The Cost of Traffic: Evidence from the London Congestion Charge, Journal of Urban Economics, Volume 121, 2021.

Taylor, Matthew & Analisse Murray (2023). "Fixing vandalised LTN infrastructure costs London councils more than 850,000" *The Guardian*, August 16, 2023.

TfL (2020). Low Traffic Neighbourhoods: what, why and where? Accessed at: https://madeby.tfl.gov.uk/2traffic-neighbourhoods/

Theis Theisen (2020) The impact of an urban toll ring on housing prices, Research in Transportation Economics, Volume 82, 2020, 100882, ISSN 0739-8859, https://doi.org/10.1016/j.retrec.2020.100882.

Tiebout, Charles M. A Pure Theory of Local Expenditure. J.P.E. 64 (October 1956): 416-24.

Walker, Peter (2023). Rishi Sunak orders review of low-traffic neighbourhood schemes. *The Guardian*. Published July 30, 2023.

Weitzman, Martin L. (1974). Prices vs. Quantities. The Review of Economic Studies, Volume 41, Issue 4, October 1974, Pages 477–491
https://doi.org/10.2307/2296698

Wen, Haizhen, Zaiyuan Gui, Ling Zhang, Eddie C.M. Hui (2020). An empirical study of the impact of vehicular traffic and floor level on property price, Habitat International, Volume 97, 2020, 102132, ISSN 0197-3975

Wong, W.-C., Azhari, A., Abdullah, N.A.H. and Yip, C.Y. (2020), "Estimating the impact of crime risk on housing prices in Malaysia", International Journal of Housing Markets and Analysis, Vol. 13 No. 5, pp. 769-789. https://doi.org/10.1108/IJHMA-06-2019-0063

Yang, Xiuleng, Emma McCoy, Katherine Hough, & Audrey de Nazelle (2022). Evaluation of low traffic neighbourhood (LTN) impacts on NO2 and traffic. Transportation Research Part D: Transport and Environment, Volume 113, 2022, 103536, ISSN 1361-9209

1.9 Appendix

1.9.1 Technical Details of Low Traffic Neighborhoods

The modal filters that comprise LTNs can generally be categorized into three groups: The most drastic simply eliminates the roadway for cars and larger vehicles. This is often done either by raising the pavement itself, or placing obstacles such as bollards across the road. The latter still allows bikes to pass through easily, and some versions are temporarily removable to allow emergency services through. Second is signage restricting entry for vehicles that is enforced by camera. Fines can be issued to motorists that flout the restrictions.[34] This setup allows for restrictions to be enforced at busier times of the day, while not inconveniencing users in off-peak hours. An example of these first two types of filter are shown in Figure 1.7. Finally, there are road designs that do not prevent through-traffic explicitly, but instead disincentivize drivers from using a route as a rat-run.[35] These can include speed bumps, chicanes, and lower speed limits. Modal filters are not unique to the U.K., with municipalities in other countries such as the U.S. and the Netherlands, and Spain have also been known to have implemented modifications to roads to decrease the flow of traffic through residential areas.

Some LTNs, including those implemented in Waltham Forest are holistic in nature: As well as introducing modal filters and expanding cycling infrastructure, there are also instances where on-street parking spaces have been reduced in favor of expanded sidewalks, pedestrianization of streets, and the planting of trees to further improve the pedestrian/cyclist experience. Some policymakers express the view that only those that implement public space enhancements in addition to modal filters should be considered a 'true' LTN (Bosetti et al., 2022).

Figure 1.7: Modal filter that completely blocks car traffic (left) and a modal filter that is enforced by camera (right)

[34]The typical fine is £135, reduced to £65 if paid within 14 days.
[35]A 'Rat-run' is colloquial term referring to a residential or side road that drivers use as a through-route.

A graphical example of a hypothetical LTN before and after implementation is shown in Figure 1.8. The left-hand graphic shows a simple grid-style layout with the nine interior cells forming a contiguous neighborhood, denoted by red dots. Initially, there are no restrictions on the type of traffic that can traverse any of the roads. As a result, motorized vehicles such as cars and truck utilize roads within the neighborhood.

The right-hand graphic shows the effect of implementing an LTN on traffic flow. A series of modal filters are placed at strategic points within the neighborhood, represented by green squares. All road segments considered part of the new LTN are shaded in light green. An important element to note about all LTNs is that all road segments are still accessible to local traffic from at least one entry point, which is necessary not just for residents, but for service and emergency vehicles too. All non-local (or through traffic) must now divert around the neighborhood to traverse the map. Part of the urban planner's job when designing these LTNs is to achieve these outcomes without excessive use of modal filters. Pedestrians and cyclists can freely traverse all road segments, and the decrease in traffic volume is intended to encourage these modes of travel, especially when amenities exist within these neighborhoods such as parks and community centers.

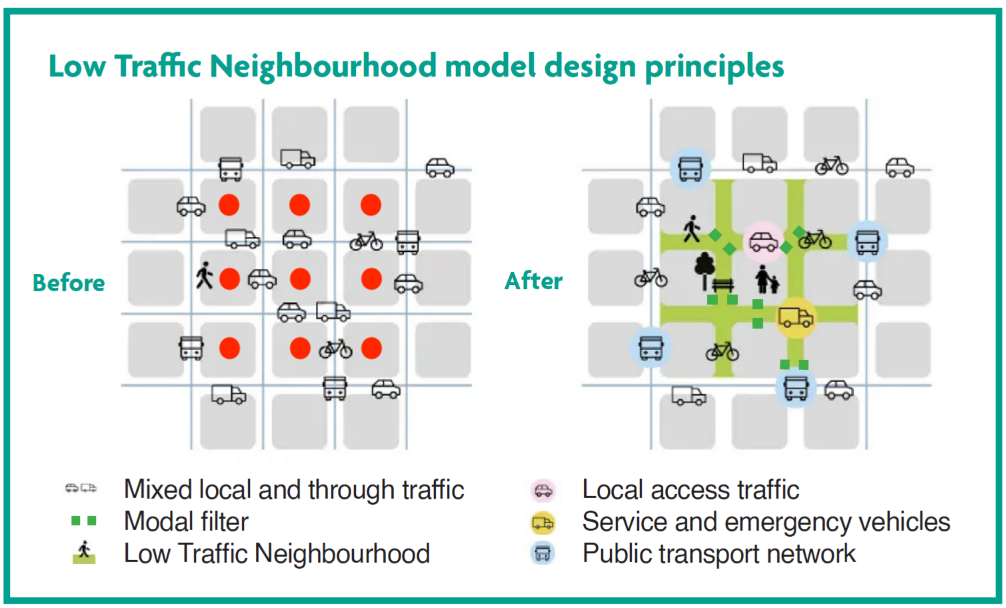


Figure 1.8: Example of implementing an LTN using modal filters in a grid-style road network.

1.9.2 Polygon Construction for Neighborhood-Level Specifications

Because LTNs are defined by the presence of modal filters, their boundaries are not necessarily drawn to match readily-available geographies (e.g. output areas). Using such geographies as cross-sectional units would therefore be subject to measurement errors such as identifying properties within an LTN as untreated and vice-versa. When I need to implement regressions at the neighborhood level that cleanly divide treatment and control areas, I aggregate individual observations to the 'polygon' level, a custom geography that does not overlap any LTN zones. Choosing large geographical areas leads to both treatment and control postcodes residing in the same cross-sectional unit, but choosing too fine an area results in many group-time pairings without observed sales, and thus an unbalanced panel. Therefore, I divide all of Waltham Forest into a set of geographically disjoint LTNs consisting of the actual LTN areas and counterfactual LTNs, considering residential areas that are defined by intersections between major roads, natural dividers (e.g. parks), and LTN boundaries themselves.[36] This method produces 16 LTN polygons and 52 control polygons, which are shown in Figure 1.9. For my main specifications, I use these polygons as cross-sectional units in certain regressions.

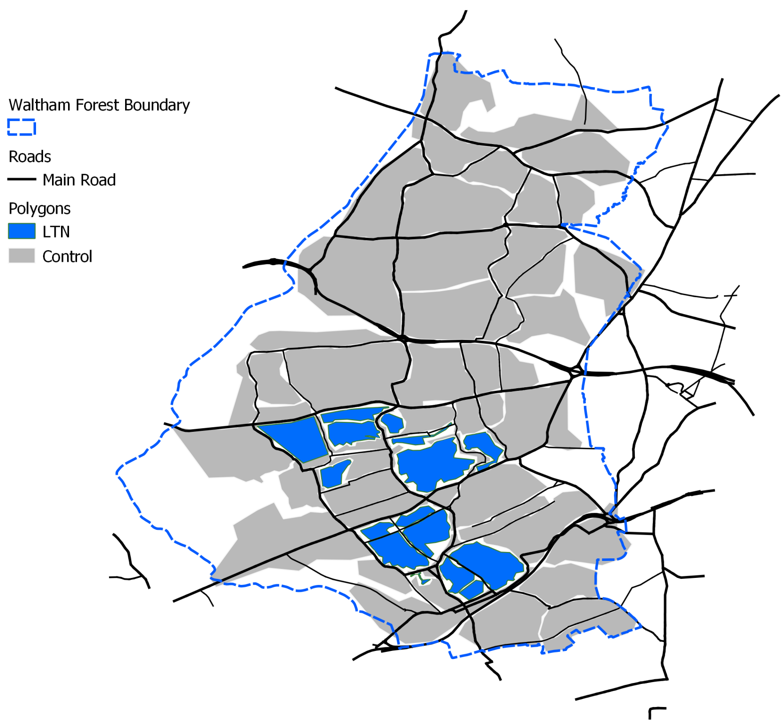


Figure 1.9: Map of Treatment and Control Polygons used for panel regressions

[36]If a larger section of residential streets appears to be dissect-able into two or more disjoint areas (minimal or no direct connection without passing through a main road), then they are drawn as separate polygons. The boundary of an LTN itself is an automatic boundary between polygons, although these are almost always defined by main roads themselves.

1.9.3 Valuation of Accident Reduction

I find that individuals are willing to pay a premium to live in an LTN-treated neighborhood in Waltham Forest. Based on my preferred specification (Table 1.2, Column 3), the increase in WTP is £16,302. Given the significant result I find in Table 1.7 regarding the effect of LTN treatment on serious accidents, some fraction of the £16,302 can be attributed to this reduction.

The UK *Department of Transport* publishes estimates of the various costs associated with road accidents and casualties in three categories: Human costs, lost output (wages), and medical costs. These costs are calculated separately for the three levels of severity of accident: Slight, serious, and fatal.[37] Since medical costs in the UK are socialized, I use only the first two to calculate benefit to households of reducing accident risk. Based on the frequency of each accident severity and the associated costs, the average recorded accident has a total cost of £67,498 (excluding medical costs).

From the empirical estimates of the effect of LTNs on accidents, treatment results in a reduction of 1.954 accidents/year/polygon. This corresponds to a total cost savings of £131,892/year/polygon. On average, there are approximately 4,100 people in each polygon in Waltham Forest. Making the simplifying assumptions that the benefit from accident reduction accrues wholly to residents and equally to each resident, this is a £32.17/year/person saving.

Assuming that this reduction in accidents is permanent and a discount rate of 5%, the NPV of this yearly saving is £643. There are an average of 2.7 individuals per household in Waltham Forest. Thus, £1,736.10/household (2015£) is the NPV of savings from accident reduction assuming risk-neutral preferences. This works out to roughly 10.6% of the total change in WTP for properties found from the introduction of LTNs.

[37] *Department for Transport*, Road Safety and Accident Statistics. Last updated 15 December 2022

1.9.4 Additional Figures

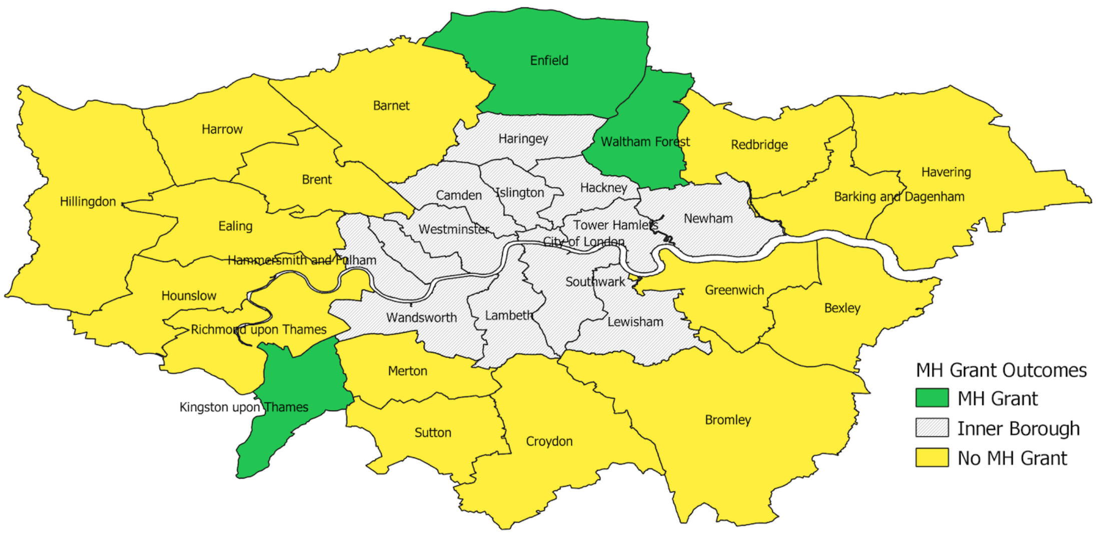


Figure 1.10: Mini-Holland grant eligibility/outcome by borough

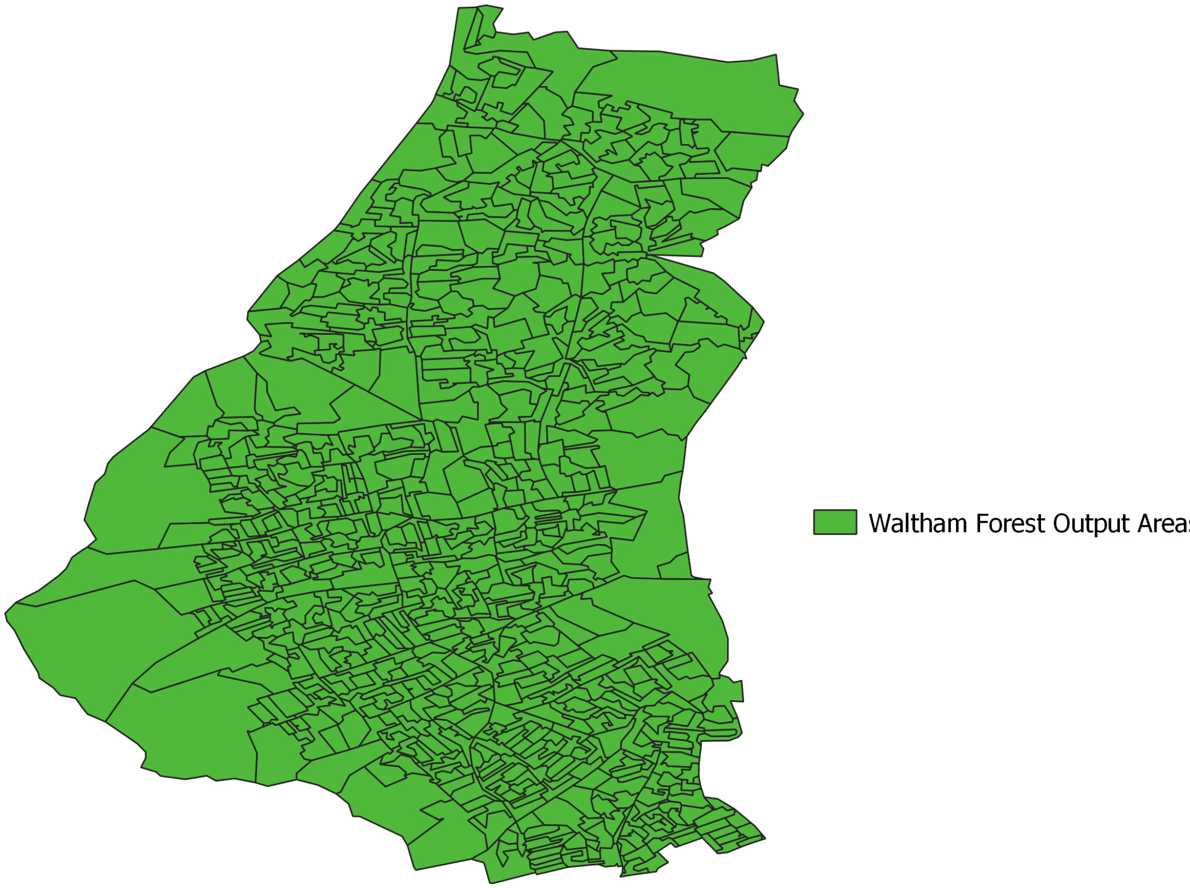
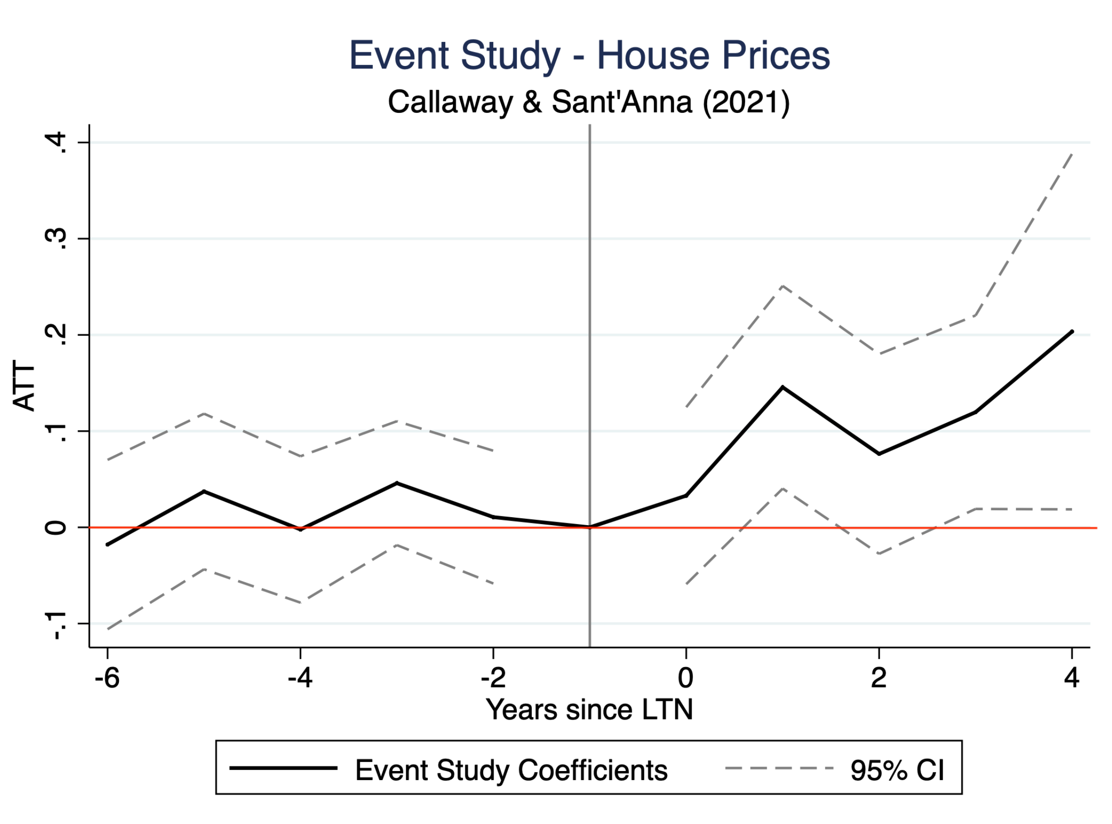


Figure 1.11: Map of (732) Output Areas in Waltham Forest



Notes: This figure shows the aggregated event study estimates from Callaway & Sant'Anna (2021). Each estimate is the average treatment effect on sales price for properties in neighborhoods that have been treated for $t - g$ periods, where t is the period that the house sale is observed and g is the period that the property is first treated by an LTN, compared to the period before treatment. Dotted lines above and below represent the 95% confidence intervals. Because the number of post-treatment years differs depending on the group g, some post-period estimates are aggregates of a subset of all treatment timing groups.

Figure 1.12: Event Study of Callaway & Sant'Anna (2021) Specification

1.9.5 Additional Tables

Table 1.9: Borough Comparison by Mini-Holland Grant Status

	(1) Waltham Forest	(2) Grant	(3) No Grant	(4) Others
Population	258,000	244,000	253,000	257,000
Avg. Deprivation Score (0-4)	1.06	0.93	0.90	0.96
Fraction White	0.52	0.63	0.62	0.61
Fraction Asian	0.21	0.16	0.21	0.17
Fraction Black	0.17	0.12	0.10	0.13
Cars Owned / Capita	1.27	1.06	1.13	1.48
Fraction Own Residence	0.50	0.57	0.56	0.46
N	1	3	5	24

Table shows mean values of variables relating to demographics and living circumstances of residents of London boroughs. Data come from the 2011 UK Census. Column (1) contains figures for Waltham Forest. Column (2) includes Waltham Forest and the two other Mini-Holland grant recipients, Enfield and Richmond. Column (3) is the set of eligible boroughs that applied for but did not receive the grant. Column (4) contains the set of boroughs that either did not apply or were not eligible (i.e. inner-London boroughs).

Table 1.10: Effects on House Prices

	(1)	(2)	(3)
LTN Treat ($\beta_1^{stacked}$)	0.149***	0.091***	0.085***
	(0.0260)	(0.0162)	(0.0169)
(0, 0.5] Ring "Treat" (γ_1)	0.025**	0.059***	0.059***
	(0.0129)	(0.0104)	(0.0106)
(0.5, 1.5] Ring "Treat" (γ_2)	0.026*	0.003	0.005
	(0.0157)	(0.0149)	(0.0150)
Multiple Sales Sample	X	X	X
Property-Stack Fixed Effects		X	X
Drop 2019 Stack			X
N	43,401	43,401	35,701

Standard errors in parentheses and clustered at the property level.
Estimates are variations on the ring-based regression in Equation (1.4).
LTN Treat denotes the standard stacked DID estimator.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

CHAPTER 2

Capitalization of 24-Hour Subway Service: Evidence from London's Night Tube

2.1 Introduction

Access to a metro or subway station is typically seen as an amenity by city-dwellers: Light rail represents a cost-effective method of commuting or getting around town. In some cases, urban road congestion can make taking a light rail train a faster and/or more convenient option than driving. Trains are usually preferable to buses in this context since in most cities buses are subject to the same road congestion issues as cars due to an unwillingness to prioritize mass transit on public roadways. Thus, it is plausible to think that a hedonic analysis would reveal a positive relationship between access to a subway station and the sale price of a residential property. Previous literature has indeed found that this prediction holds true for traditional daytime services (e.g. Baum-Snow & Kahn, 2000; Diao, Leonard, & Sing, 2017; McMillen & McDonald, 2004; Zhou et al. 2021). However, little attention has been paid to separately estimating the capitalization of nighttime service into property prices, likely because nighttime services are much less common, typically being restricted to larger cities that can justify running them. The modal users and uses of daytime and nighttime subway services are different, and so it is plausible to think that these amenities may be valued differently by current and prospective residents.

In this paper, I provide the first investigation of how nighttime service is valued by residents. I utilize the rollout of London's Night Tube from 2016 onwards to estimate the effect of receiving weekend nighttime service at a nearby station on the sale price of residential properties in Greater London. Using comprehensive data on residential property sales in Greater London, I exploit the fact that nighttime services were introduced on certain sections of the Tube network but not others, often due to technical and engineering considerations, and utilize a difference-in-differences approach to compare sale prices of homes located close to stations that did and did not receive nighttime service, before and after service begins. I find that Night Tube service is associated with a $1.3-2.6\%$ decrease in sale prices through the end of the decade. Based on the mean sales price of properties in the baseline period, this translates to average decreases of between £$6,500$ and £$13,000$, or roughly $15-30\%$ of the existing average premium of being located near any Tube station. I test my identification strategy by showing event studies and by running restrictive specifications that limit the sample to look at repeat sales of the same property. If anything, the repeat sales specification hints at a pre-trend that attenuates my result compared to an unbiased estimate of the treatment effect.[1]

I investigate two potential mechanisms that could plausibly explain my main result: First, rail services that operate at surface level produce noise pollution and thus could be (and have been) associated with reduced housing values (e.g. Diao, Qin, & Sing, 2016). To explore this channel, I use distance to the nearest relevant surface-level train track as a measure of the intensity of the potential disamenity and look for heterogeneity of my parameter estimate in this distance. However, I do not find any indication that this is driving my results. Second, the elevated foot traffic around stations at night could provide an increase in the supply of potential targets for criminals, which under a 'cost of crime' model (Becker, 1968) would increase the expected benefit of committing a crime, leading to an increase in crimes committed, which would be congruent with the main result. By replacing house prices with crime counts at the station level as my outcome variable, I find suggestive evidence that total crime counts increase near treatment stations by roughly 4%, which is driven by increases in thefts and anti-social behavior.[2] This result is consistent with previous research showing that crime is capitalized into house prices in Greater London (Gibbons, 2004; Gibbons & Manchin, 2008). To probe the crime mechanism further, I cut the sample by property type and by geography, and find that all the statistical significance comes from the subset of houses located outside the center of London. Household units that live in these kinds of areas/properties are more likely to contain dependent children, and are more likely to be homeowners, both groups that could plausibly be more sensitive to changes in crime than their counterparts.

My findings contribute to the existing literature that estimates the effect of access to light rail, or mass transit generally, on house prices, specifically when access is defined by being walkable (as opposed to 'park and ride' access that might be more common for intercity and/or regional rail services). While prior work looking at general service finds that access is capitalized positively into housing prices both generally (Baum-Snow & Kahn, 2000; Diao, Leonard, & Sing, 2017; Dube et al., 2014; McMillen & McDonald, 2004; Zhou et al. 2021) and specifically in London (Gibbons & Manchin, 2005), I show evidence that this does not necessarily extend to more specialized services, namely nighttime service. My investigation into crime as a suggestive mechanism is consistent with prior work that showing that crime near stations tends to decrease when service is suspended at those stations (Block & Block, 2000; Phillips & Sandler, 2015). From a policy perspective, I assert that these findings are of interest when considering equity of outcomes. Although a service such as

[1]When I add linear trends to the repeat sales specifications, the results look almost identical to the original specifications using all sales.

[2]I am only able to look at aggregate crime levels in the data, as opposed to crimes specifically occurring on weekend nights, so these results should not be taken as conclusive.

the Night Tube may pass a cost-benefit analysis (Volterra, 2014), there are some groups that view it as a net negative, at least to the extent that house prices in fact reveal true preferences. Under the assumption that crime is a factor contributing to my main result, practitioners may consider additional actions that can be taken to address the potential for increased crime near Night Tube stations during service hours.

The rest of the paper is structured as follows. In Section 2.2 I describe my setting of London's subway network and the implementation of the Night Tube. In Section 2.3, I describe the data used in the analysis and discuss my empirical strategy. Section 2.4 contains the main results and tests of the identifying assumption. Section 2.5 probes some potential mechanisms and demographic heterogeneity that could explain my main results. Section 2.6 concludes with a discussion of findings.

2.2 The Night Tube

2.2.1 Origins of the Night Tube

The "Tube" refers to the colloquial name of London's subway system that is operated by the quasi-governmental corporation Transport for London (TFL) (Croome & Jackson, 1993).[3] It is the oldest subsurface passenger railway in the world, with some sections of track having been originally built and operated as far back as 1863 (Day & Reed, 2008). Even these original sections of track were highly utilized and have been shown to have contributed to increases in land and property value (Heblich, Redding, & Sturm, 2020). Today, the network is one of the busiest in the world. Between the system's three main components: London Underground,[4] London Overground, and the Docklands Light Railway, there are about 1.6 billion journeys taken each year (Statistica, 2021a,b,c).

Despite having a large and lucrative nighttime economy, until recently London never had any form of all-night Tube service. Regular service ends around midnight, and services restart around 5-6am the next morning. There has always been a need for some form of late-night travel within the capital, especially on weekend nights to support the nightlife industry, and this need has historically been filled by the bus network and the use of private taxis. Multiple mayors and city administrations had expressed the desire to add Tube service to the suite of nighttime transport offerings, but hurdles existed that delayed the implementation of such a service. First, given the age of the Tube network, it requires a large amount of regular maintenance, much of which was (and still is) done overnight which is a barrier to night service. Second, despite the names Tube

[3] Although the origin of the nickname The Tube derives from the underground portion of the TFL network characterized by narrow tunnels, the term Tube as used in this paper will include many above-ground sections, including the London Overground and Docklands Light Railway (DLR) lines.

[4] The London Underground consists of 11 separate lines: Bakerloo, Central, Circle, District, Hammersmith & City, Jubilee, Metropolitan, Northern, Piccadilly, Victoria, and Waterloo & City.

and The Underground, almost half of the network is above ground, and so noise from trains and stations at surface level is a potential disamenity for residents. This is something that TFL is continuously aware of and working to remedy. Finally, operating the Tube is costly, and it was not always certain in the past that there was a strong enough economic case for a Night Tube service. In the aftermath of the 2012 London Olympic Games, where the volume of nighttime traffic made it apparent that nighttime transport infrastructure would eventually need to be expanded, TFL sought to develop a plan as to which Tube lines would be most viable for an initial rollout of a Night Tube service. This analysis focused on four main criteria (Roberts, 2016): Services where there is an existing demand for overnight transport, a minimum 6 trains per hour per direction through central London, each line to pass at least one train maintenance depot, and minimal impact to ongoing major upgrades.

Existing demand for overnight transport was gauged by looking at several factors, including night bus use and daytime Tube numbers. Only lines that passed through central London, essentially connecting the downtown area with outlying areas, were considered. This meant that the Waterloo & City line, which predominantly serves commuters, and some sections of the London Overground, which do not enter central London, were not considered. Accessible maintenance depots are required in the case that defective trains need to be swapped out on short notice. Although this did not rule out any lines per se, it did impact viability of at least one line on the intensive margin in the sense that the line had to be operated at least as far out as a suitable depot (Central line). Finally, and perhaps most consequentially, large parts of the Tube network were (and are still) in the midst of major upgrade programs. The most notable of these is named the Four Lines Modernization (4LM) program, which involves major long-term engineering upgrades to four Tube lines[5] (Stacy, 2016). These works prohibit the implementation of the Night Tube on these lines, even if there might be an economic case for their inclusion (especially in the case of the District line). Additionally, ongoing upgrades to Bank, a station that serves many lines, meant that the branch of the Northern line that serves Bank was not selected for Night Tube service, whereas the rest of the line was included. A quote around the time of the Night Tube launch from Mark Wild, the managing director for London Underground at the time, highlights the fact that the main factors behind the decision to leave out the Circle, District, Hammersmith & City, and Metropolitan lines from the original Night Tube proposal were related to technical constraints from engineering upgrades unrelated to Night Tube operation itself:

> *"The reason we cannot quickly put it (Night Tube) onto the other lines - District, Metropolitan, Hammersmith & City and Circle - is that we are presently, like open-heart surgery, replacing*

[5] These lines are the Circle, District, Hammersmith & City, and Metropolitan.

> *the signaling system there. We are presently using those weekends, as probably people know because they get their journeys interrupted, to install a brand new signaling system, which will be available only from 2019 onwards. It is very important that we do not try to put the Night Tube on top of a system that we know needs a lot of work and reliability. I predict, with the right business case, the Night Tube could be expanded but only to the other subsurface lines after the signaling upgrade."*

Although it is a possibility that at some point in the future these lines, or a subset of them, will be included in future Night Tube expansions, there are currently no indications that this will happen soon, and this was certainly also true prior to and during the initial implementation of the Night Tube.

2.2.2 Justification & Rollout

Large provisions of public goods, including mass transit, need to be justified to the public from a cost-benefit dimension. TFL commissioned a study in 2014 to analyze the costs and benefits of implementing a nighttime Tube service on the weekends. The original proposal included full service on two lines (Jubilee and Victoria) and partial service on three other lines (Central, Northern, and Piccadilly). This report estimated that the Night Tube would add around 2,000 permanent jobs, considering both direct employment effects of running the service and indirect effects on local businesses. It was also estimated that time savings on the average nighttime journey would be 20 minutes, but could be up to 1 hour on certain routes (Volterra, 2014). The study estimated the benefit-cost ratio to be 3.9/1.

In mid-2014, TFL announced the upcoming launch of the Night Tube on the five lines proposed in Volterra (2014). The plan was to begin service in September 2015. However, several labor unions representing employees working on the Tube were unhappy with the terms of labor laid out by TFL for workers on the Night Tube. This led to multiple strikes that occurred during the Summer of 2015 that put the launch of the Night Tube on hold. Eventually, in August 2015 it was announced that acceptable terms had been reached between TFL and the labor unions, and a new launch date was approximated for the second half of 2016. On August 19, 2016, service on two lines (Central and Victoria) began. Between then and December 2016, service on the three additional Night Tube lines (Jubilee, Northern, Piccadilly) began. Approximately one year later, between December 2017 and February 2018, an additional service was rolled out on a section of the London Overground. Figure 2.1 shows a map of current Night Tube service, and Figure 2.2 shows the full Tube network map (note that this map includes rail and tram lines that are not a part of my sample). Night Tube service operates all of Friday/Saturday nights, between the last 'regular service' train on Fridays and the first morning train on Sundays, to ensure continuous service through the whole weekend. In March 2020, Night

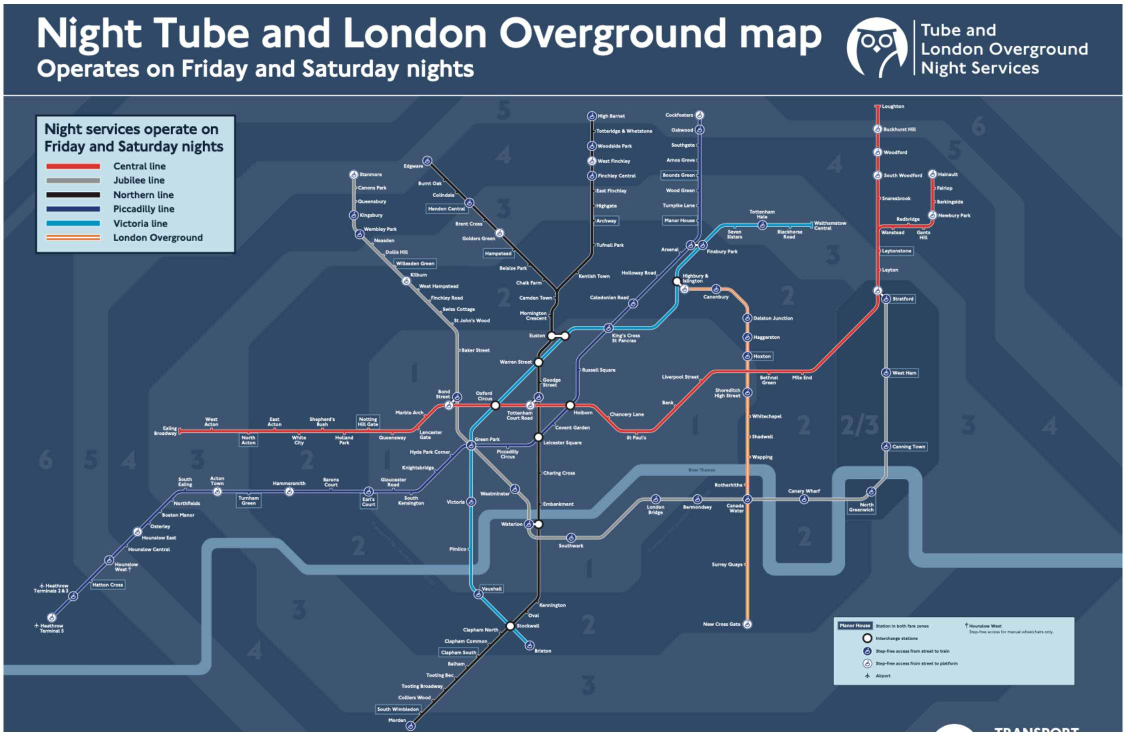


Figure 2.1: Night Tube Service Map as of May 2020

Tube service was suspended due to the shutdown of the nightlife and entertainment sectors in the light of COVID-19.

2.2.3 Other Possible Impacts of Night Tube

In addition to increased transit options, Night Tube service (and subway access more broadly) has the potential to impact house prices through effects on other relevant local amenities. The main specifications in this paper using house prices as an outcome thus represent the net effect of Night Tube treatment on local areas around stations. I can separately look at the effect of treatment of some, but not all, of these potential mechanisms. However, I argue that these unobserved mechanisms are likely less relevant when looking specifically at changes to transit access at night.

A common reason why households might refrain from locating in a neighborhood is the level of noise. In urban areas, transportation of various types can make up a significant fraction of noise. Previous literature has investigated the hedonic price schedule in relation to noise from road traffic (e.g. Andersson, Jonsson, & Ögren, 2010), aircraft (e.g. Nelson, 2004), and public transportation (e.g. Diao, Qin, & Sing, 2016), and

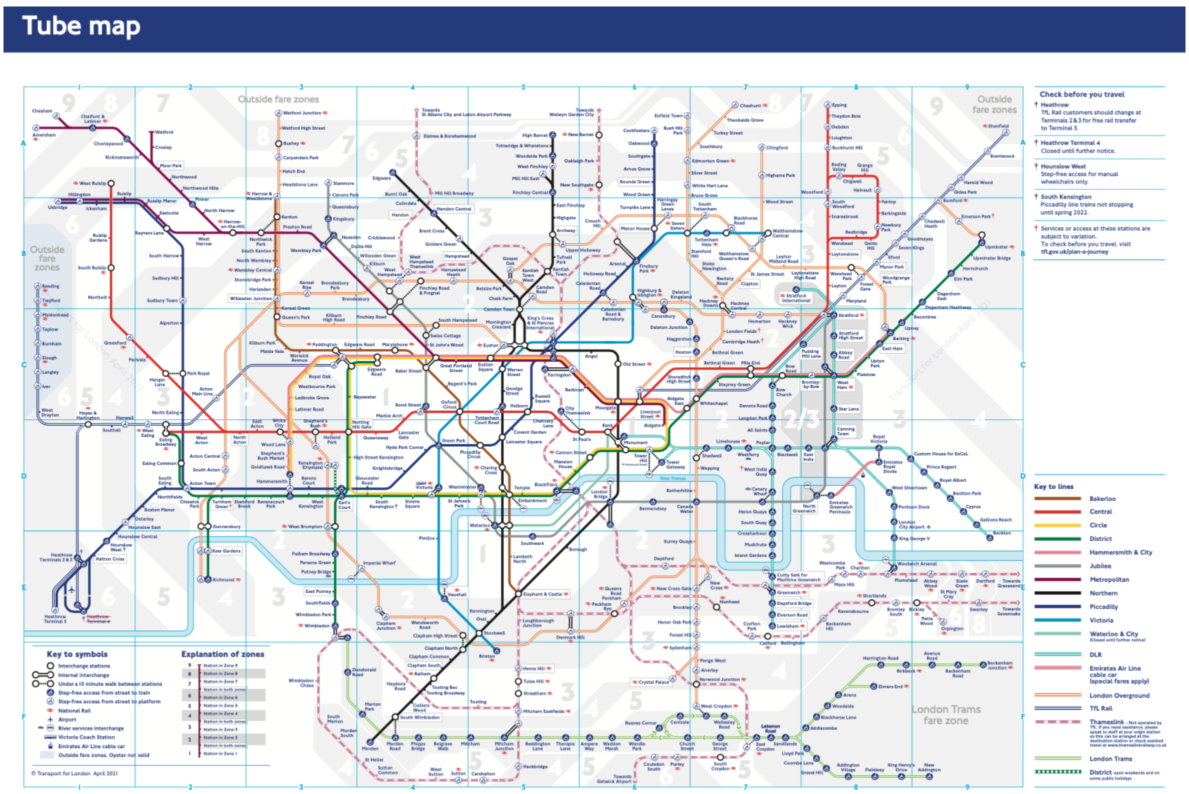


Figure 2.2: Full Tube Service Map as of April 2020

evidence predominantly points to noise being negatively capitalized into house prices. Although London's Tube network contains a lot of underground sections, around half of all stations are located at surface level, and thus being located close to either surface-level stations or the tracks in between surface-level stations could be associated with lower desirability of an area. In the case of the Night Tube, all stations, whether treated or untreated, already will have this noise present during the daytime, but noise during the night when people are trying to sleep could represent a further negative externality. TFL has been quite responsive to concerns regarding noise, making various upgrades that help reduce noise from tracks and trains, as well as reducing noise from announcements and noises at stations themselves (TFL, 2017).

Another well-studied local amenity related to house prices is the presence of crime. Linden & Rockoff (2008) find evidence that residents have a significant willingness-to-pay for a perceived decreased in the risk of crime in their neighborhood. Ceccato & Wilhelmsson (2020) find that house prices are sensitive to their distance from 'crime hotspots', specific areas with high concentrations for crime. Lynch & Rasmussen (2001) provide evidence that the frequency with which crimes occur in a neighborhood affects house prices. Thinking specifically about the Night Tube, it is possible that that the increased activity around Tube stations on weekend nights might attract crime in the local area, as previous literature has suggested that mass transit access and crime are positively related within a local area (Block & Block, 2000; Phillips & Sandler, 2015; Wu & Ridgeway, 2021). This evidence fits with the Becker model of rational crime (1968), but it is also possible that there is an 'eyes on the street' phenomenon where increased foot traffic actually discourages criminal activity (Jacobs, 1961). Both the noise and crime channels are discussed further in Section 2.5.1.

There exist other local amenities that could be affected by the presence of the Night Tube, but that are not assessed in this paper due to data limitations. First, since public transportation can be a substitute for car travel, increased access may decrease road traffic, which is also a local amenity that is plausibly related to house prices. However, empirical evidence suggests that this link is weak (e.g. Duranton & Turner, 2011) and since my treatment occurs at night and not during peak traffic hours, this channel is not a large concern. Second, the presence of transit-oriented development can influence land use patterns in the surrounding area, which can also affect the desirability of an area (Yang, Cao, & Zhu, 2021). Once again, my setting concerns treated and control stations that both already have full daytime service, and so a large portion of these effects on land use are likely already realized. However, night service may attract certain types of businesses or local investment (e.g. late night entertainment). Finally, it may be the case that the option value of transit access can be measured in terms of access to specific amenities. For example, Herskovic (2020) shows that improved public transit access increases the set of schools that can be feasibly attended by children in the household, which should weakly increase resulting school quality associated with a neighborhoods, which has been shown to be capitalized into house prices (Black, 1999).

2.3 Data & Empirical Approach

2.3.1 Data Sources

I get data containing virtually all recorded sales of residential properties from HM (His Majesty's) Land Registry that occur between 2010 and 2019. I observe price, date of sale, full address of the property (processed via the Ordinance Survey's *AddressBase*), and some characteristics of the property, namely the type of property, the tenure (freehold vs. leasehold), and whether it is a new construction. This data does not contain observations from sales that are "not for value" such as gifted properties and administrative transfers, sales that occur outside normal market forces such as Right to Buy sales and compulsory purchase orders. It does not include commercial property sales.

For data on stations, I manually collect coordinates and date of treatment by Night Tube line, if any. I include all stations in the (TFL) network that provide light rail services, meaning I exclude tram/bus-only stations and stations serving national rail routes only, to ensure the control stations are relevant. My final sample contains 387 stations, 154 of which receive NT service during the study period ($\approx 40\%$).[6] To quantify sales that occur "close" to a station, I match observations in the HM Land Registry data to their nearest-neighbor station. I then restrict the maximum allowed distance of the nearest neighbor for a sale to be attributed to that station. My default radii are 1 mile and 0.5 miles. The 0.5-mile radius is motivated by TFL's Public Transport Accessibility Levels (PTAL) standards, defining a station as being within a walkable 'catchment area' as being 960 meters in straight-line distance, often adjusted to 800m (0.49 miles) to account for road structures and obstacles (TFL, 2010). I prefer when possible to use a 1-mile radius in part because it increases power by reducing the number of observations dropped, but also because there is evidence of a "last mile" effect where Tube passengers will use taxis and rideshare services to get between the closest Tube station and their destination (Rao, 2016).

2.3.2 Empirical Strategy

My empirical strategy is a Difference-in-Differences specification that utilizes the fact that some Tube stations receive Night Tube service and some do not. As discussed in Section 2.2, this is partially due to economic considerations affecting the whole line, rather than stations individually, and partially due to preexisting upgrade programs and technical considerations of the network that prevent certain lines (and segments of

[6] I drop Chesham and Amersham stations (Metropolitan line) since I only consider Greater London postcodes.

lines) from being feasible at the current time. There is a small amount of variation in treatment timing, but variation is mainly coming from geography of treatment. In addition, my main specification does not utilize timing of treatment at all. Equation 2.1 shows the baseline specification:

$$ln(p_{ist}) = \beta_0 + \beta_1 TREAT_{st} + \overline{X_{ist}}\underline{\beta_2} + \gamma_s + \delta_t + \varepsilon_{ist} \tag{2.1}$$

Where $ln(p_{ist})$ is the natural log of sale price for house i that is a nearest-neighbor (and within the specified radii) to station s that occurs in period t. $TREAT_{st}$ is equal to 1 if the station receives Night Tube service and if treatment has started. Since house prices could respond to the expected value of future (dis)amenities, I define treatment either as the year of announcement (2015), or as the actual month that service starts contains available property characteristics, including the type of property,[7] whether the property is owned through a freehold or a leasehold,[8] and whether the property is a new or existing build. γ_s and δ_t denote station and time (year or month, depending on specification) fixed effects. The main identifying assumption of this model is that absent the introduction of the Night Tube, the growth of sale prices near treated and control stations should be similar. This assumption is tested in the next section. Because I do not observe all relevant property attributes, such as the number of bedrooms and lot size, I also include strict specifications where I include property fixed effects so that I am only identifying effects based on changes in the sale price of the same property before and after treatment.

2.3.3 Descriptive Statistics

Panel A of Table 2.1 contains some selected descriptive statistics broken down by treatment and control stations in the baseline period (2010-2014) for observations that meet the looser of the closeness criteria (1 mile). Property sales near treatment stations do have a higher average sale price in the pre-period, although there is a large amount of variation, mostly in the form of a positive skew since London does contain a small stock of highly expensive properties that are probably not very sensitive to public transport access. There does also appear to be some observations with implausibly low sale prices. These are likely either miscoded or should have been excluded from the data due to not being residential properties or not being market sales. This motivates me to take the natural log of price and windsorize the top and bottom 1 percent of the sample by price when running my regressions.[9] Treatment and control properties are reasonably similar in observable

[7]Property types include detached house, semi-detached house, terraced house, and flat.

[8]A freehold means that the owner holds the title to the land and all buildings on the land in perpetuity. A leasehold means that the owner has a long-term, but temporary ownership of the property only, typically at least 40 years but often 99 years or more. Leaseholders may have to pay ground rents to the landowner, may have some restrictions on making changes to the property, and must pay to extend the lease if close to expiration. These factors mean that properties under a leasehold command a lower asking price than an equivalent property that come with a freehold.

[9]For completeness, I include versions of my main results (Table 2.2) that do not windsorize in Table 2.15.

property characteristics, and the average distances to both the nearest station and the nearest train track are similar. In Panel B of Table 2.1 display statistics that look at demographic and household composition in the neighborhoods around treatment and control stations. I do this by matching each station to a Middle Super Output Area (MSOA) in the 2011 Census.[10] Treatment and control neighborhoods appear to be similar in all the observed categories in 2011. Most notably, measures of economic deprivation are not different and there is not a significant difference in the tenure status of households.

2.4 Main Results

To provide evidence supporting the identifying assumption behind the difference-in-differences model in Equation 2.1, I show event study versions of my main specification (treatment occurs in 2015). Figures 2.3 and 2.4 show these event studies for the 1-mile radius restriction. In Figure 2.3, there do not appear to be any significant pre-trends suggesting that the growth of prices would have been different in the treatment and control areas. When focusing in on repeat sales only in Figure 2.4, there is evidence of a possible pre-trend suggesting that treated units may have experienced a higher growth in prices in the absence of the Night Tube. If true, this would lead to my results to being attenuated compared to an unbiased estimate of the treatment effect. Visual inspection of the event studies indicate that the downward trend may have already started in 2015 when the announcement was made, but the effect is not statistically significant (possibly since the announcement was made mid-year). By 2016, Night Tube stations are experiencing significant decreases in price relative to other stations. In Figures 2.7 and 2.8, I show versions of the event studies using 0.5 mile radii, which are similar to my default 1-mile versions.[11]

Table 2.2 shows the results from estimating Equation 2.1. Columns (1)-(3) use 2015 as the start of treatment, corresponding to the announcement of NT, and Columns (4)-(6) use the actual month when service starts. In columns (2) and (5) I cluster standard errors at the station level, which is the level that treatment occurs at. Finally, some Tube stations, particularly near the center of London, are very close to other stations. If this is the case, then we could imagine a situation where a property is located close to both a treatment and a control station, but is closer to the control station, and so would be classified as a control observation in my analysis. To avoid this, in columns (3) and (6), I exclude “close stations”, which are defined by TFL as stations within a 10-minute walk from another station. Table 2.10 contains the event study coefficients for each pre-and post-period, and shows that in 2015 the estimated parameter is negative but not statistically-significant, but from 2016 onwards individual year estimates are largely significant. For specifications using

[10]There are around 1,000 MSOAs in my potential sample: 983 in the London region as defined by the Census, and a handful of MSOAs outside the London region that correspond to the most outlying stations on the Tube network

[11]I also show examples of event studies at the quarter-level in Figure 2.9, although due to the number of periods involved, the visual is somewhat noisier.

Table 2.1: Selected Descriptive Statistics: Baseline Period (2010-2014)

	Treatment		Control	
	Mean	S.D.	Mean	S.D.
Panel A: Sales				
Price (000's)	572.2	583.1	465.0	583.1
% Flat	0.63	-	0.58	-
% New	0.10	-	0.11	-
% Freehold	0.38	-	0.41	-
Distance (Station, mi.)	0.41	0.24	0.42	0.24
Distance (Track, mi.)	0.25	0.21	0.25	0.21
N	146,745		227,790	
Panel B: Neighborhoods				
Age				
% 0-17	17.3	4.9	19.6	5.0
% 18-59	68.1	6.9	66.0	7.5
% 60+	14.6	4.7	14.4	6.2
Deprivation				
Avg. Deprivation Index	0.95	0.20	0.95	0.23
Dwelling Type				
% Shared Dwelling	1.7	1.9	1.2	1.3
% House	32.3	24.0	39.0	26.9
Tenure				
% Own	40.4	17.2	43.3	20.3
% Social Housing	25.6	16.7	26.2	17.2
HH Composition				
% with Dep. Children	24.6	8.7	28.1	8.6
N	155		235	

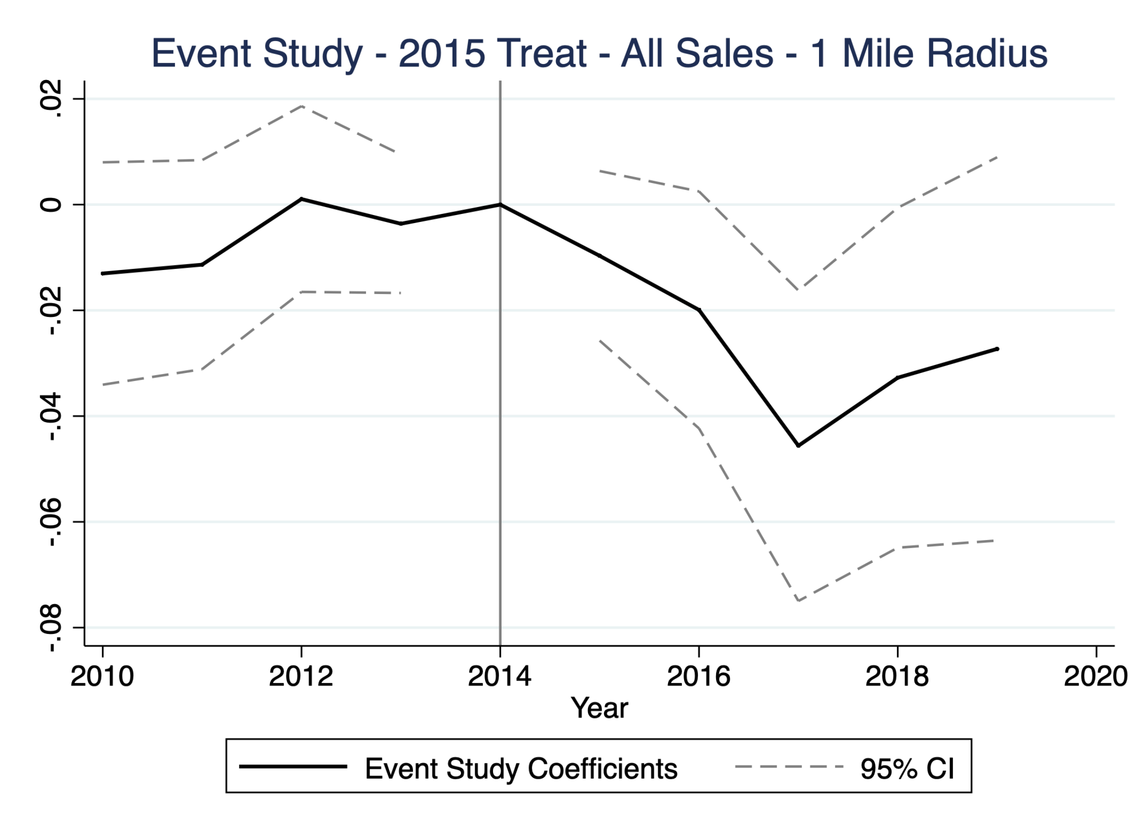


Figure 2.3: Event Study - All Sales

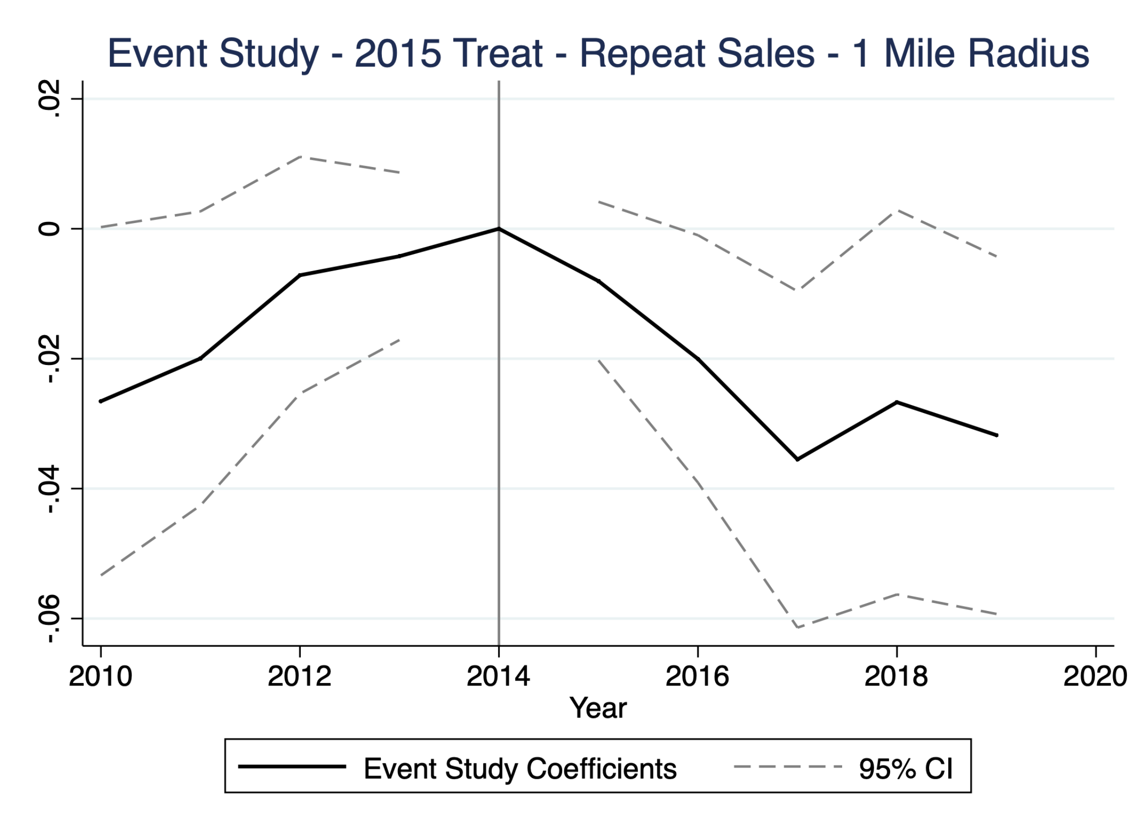


Figure 2.4: Event Study - Repeat Sales Only

all sales, treatment is associated with a 2-3 percent decrease in house prices, whereas when looking at only repeat sales, which has a more plausible argument for causal interpretation, estimates are slightly smaller, typically a 1.3-2.0 percent decrease, although due to potential attenuation as hinted at in Figure 2.4, the true magnitude may be greater since treatment stations appear to be differentially trending in a positive direction prior to treatment. To get a rough sense of this attenuation, I add linear time trends for each station to the right-hand side of repeat sales specifications of Equation 2.1. Results of these specifications are shown in Table 2.11. Adding the linear time trend brings the magnitude of the estimates in line with the specifications using all sales in Table 2.2, and greatly increases the statistical significance. An additional issue for interpretation is that including property fixed effects in the repeat sales specifications changes the sample of properties being examined. Thus, some of the differences in the estimated treatment effect may in fact come from this change in sample composition. To check if this is the case, I run a set of regressions using the "all sales" specification (i.e. no property fixed effects), but I limit the sample to properties that were transacted at least twice between 2010 and 2019. These specifications are shown in Table 2.12, and there is no meaningful difference in the magnitudes and/or statistical significance of the coefficients, indicating that it is not underlying differences in the types of properties transacted more than once that is driving the difference between the all sales and repeat sales regressions.

For interpretation of magnitudes, I focus on the results from the repeat sales samples, which are more plausibly causal, and possibly even biased towards zero. For the specifications using a 1-mile radius, estimates of the Night Tube treatment effect range from -1.3% to -2.3%. Based on the average sale price of all units in the pre-treatment period (£$\approx 507,000$), this translates to prices decreases of £$6,591$ to £$11,661$.[12] Under the assumption that the linear pre-trend observed in the repeat sales event studies would continue absent treatment, the coefficients from Table 2.11 give a range of treatment effects of -1.7% to -2.6% (£$8,619$ to £$13,182$ decrease). These decreases represent the net treatment effect associated with the neighborhood around a station receiving Night Tube service. The treatment effect is inclusive of any potential effects on other local amenities due to treatment, regardless of direction. Another way of framing the magnitude of this result is to compare it to the existing house price premium for living near a Tube/Overground/DLR station. In Table 2.13 I run an OLS regression of log price on binary and continuous variables relating to closeness from any station using the sample of properties that are within 3 miles of a station to avoid comparing areas that are too dissimilar.[13] Focusing on column (1), which uses a dummy for being within 1 mile of any station, I find that houses within 1 mile of a station are priced 8.5% higher compared to those located between 1 and 3

[12] The treatment effect based on the median price (£$340,000$) gives a range of £$4,420$-£$7,820$, which may be more instructive since house prices exhibit a positive skew.

[13] I rerun these regressions without the $<$ 3-miles restriction and the results do not significantly change.

Table 2.2: Main Regression Results

	Treat = 2015			Treat = Month Service Starts		
	(1)	(2)	(3)	(4)	(5)	(6)
All Sales						
Radius = 1 Mi.	-0.021***	-0.021**	-0.025***	-0.024***	-0.024**	-0.029**
N= 755,000	(0.002)	(0.011)	(0.011)	(0.002)	(0.012)	(0.012)
Radius = 0.5 Mi.	-0.020***	-0.020*	-0.024**	-0.021***	-0.021*	-0.026**
N= 503,000	(0.002)	(0.012)	(0.012)	(0.002)	(0.013)	(0.013)
Repeat Sales Only						
Radius = 1 Mi.	-0.013***	-0.013	-0.018*	-0.019***	-0.019*	-0.023**
N=211,000	(0.002)	(0.010)	(0.010)	(0.002)	(0.011)	(0.011)
Radius = 0.5 Mi.	-0.013***	-0.013	-0.017	-0.019***	-0.019	-0.021*
N=141,000	(0.002)	(0.011)	(0.011)	(0.002)	(0.012)	(0.012)
S.E. Clustering						
Robust	X			X		
Station		X	X		X	X
Robustness						
Exclude "Close Stations"			X			X

Notes: Standard errors are in parenthesis. Each estimate corresponds β1 in a variant of Equation (1). Each house sale is matched to the closest station in my sample. I windsorize the top and bottom 1% of sample to avoid extreme outliers impacting the estimate. The treatment indicator for an observation is equal to 1 if that station has received Night Tube service by the relevant month or year. I restrict the sample to observations whose nearest-neighbor station is within a certain distance, either 1 or 0.5 miles, to avoid properties that are not close to any station being identified as treated units. All specifications include station and time fixed effects. For repeat sales specifications, I include property fixed effects so that I am only comparing sales of the same property that happen before and after treatment. Columns (3) and (6) drop observations matched to stations that are within a 10-minute walk from another station, to avoid including sales that occur close to both treated and control stations. Around 6% of observations (N) fall under this definition.

$***p < 0.01$, $**p < 0.05$, $*p < 0.1$

miles away. Using the same mean price as before of £$507,000$, this is about a £$43,000$ difference. Therefore, one way of viewing the relative magnitude of the net Night Tube treatment effect is that it reduces the average existing premium by $\approx 15-30\%$ depending on the chosen specification.

As a midway point between my main specification that uses all sales, and my very restrictive one that only uses repeat sales for identification, I also run specifications that absorb postcode fixed effects.[14] This is still a highly dis-aggregated measure: There are only around a dozen addresses within a typical postcode, and the average number of observations per year per postcode in my data is 1.9. This specification will do a very good job at accounting for spatially-based amenities, but will not explicitly account for all possible property characteristics, except indirectly to the extent that unobserved property characteristics are spatially correlated. See Table 2.14 for these results. Encouragingly, the results are unchanged from the all-sales specifications, and in fact tend to have slightly larger magnitudes (although not significantly so). One other potential concern is that treatment might be affecting the rate at which properties are bought and sold. To look at this, I collapse observations to the station-year level and then put number of sales/observations on the left-hand side instead of price. The results of these specifications are shown in Table 2.3. The parameter estimates in this table translate the change in the number of sales that occur close to a station per period (year for columns 1 and 2, month for columns 3 and 4) because of treatment. Most specifications return insignificant decreases of between $2-5\%$, except for column (1) which is significant at the 10% level and represents roughly a 5% decrease from mean sales.[15]

A result showing a decrease in prices is somewhat surprising at first glance. The existing literature on the capitalization of rapid transit access into house prices overwhelmingly suggests that increased access should be associated with higher prices since access is an amenity (McMillen & McDonald, 2004; Gibbons & Manchin, 2005; Dube et al., 2014; Diao, Leonard, & Sing, 2017). One major difference between my setting and previous analysis is that I am looking specifically at an expansion of nighttime access rather than regular daytime use. In the next section, I investigate some possible stories that could explain a negative result.

[14]The closest equivalent in the U.S. is the 9-digit zip code.

[15]I can rule out decreases in sales larger than around 12% in column (1) using a 95% confidence interval.

Table 2.3: Investigating Effects on Number of Sales

	(1)	(2)	(3)	(4)
Treatment Effect	-11.89*	-7.98	-0.42	-0.31
	(6.80)	(5.57)	(0.50)	(0.44)
Mean Sales	207	137	18	12
Radius				
1-mile	X		X	
0.5 miles		X		X
Treatment				
Announcement	X	X		
Service Start			X	X

*Notes: Standard errors are in parenthesis. Each treatment effect estimate corresponds β1 in a variant of Equation (1) that collapses observations to the station-year level and replaces the dependent variable with number of sales. Columns (1) and (2) use the 2015 definition of treatment, and columns (3) and (4) use the actual month service begins. Mean sales refers to the average number of observations matched to station-time units. All specifications include station and time fixed effects. ***p < 0.01, **p < 0.05, *p < 0.1*

2.5 Heterogeneity & Mechanisms

2.5.1 Investigating Mechanisms

As mentioned in Section 2.2, a significant proportion of the Tube network is not actually underground (around half of stations). Thus, the level of noise emitted by Tube trains can be a potential negative externality, particularly if trains are running at night. Previous literature has found that stopping service on previously noisy train lines can increase property values located near tracks (Diao, Qin, & Sing, 2016). Even though TFL recognized the potential for this problem in advance of the launch of the Night Tube, and took (and continue to take) action to try and decrease the amount of noise that trains make, such as installing noise-absorbing track fastenings and grinding down the rails (TFL, 2017).

To assess whether noise from the Tube might be a contributing factor to house price decreases, I a) look specifically at the subsample of sales where the closest station is above-ground and b) calculate distance to the closest track on the Tube network.[16] Figure 2.5 shows the distribution of distance (in miles) from the nearest track section, specifically for the sample of observations within 1-mile of the nearest station, to stay consistent with the sample from the main analysis. Since a property should be at least as close to the nearest

[16] To calculate distance the closest section of track, I use the QGIS plugin NNJOIN, which allows me to perform nearest-neighbor matching between a vector layer of points (sales) and another vector layer of lines (tracks.)

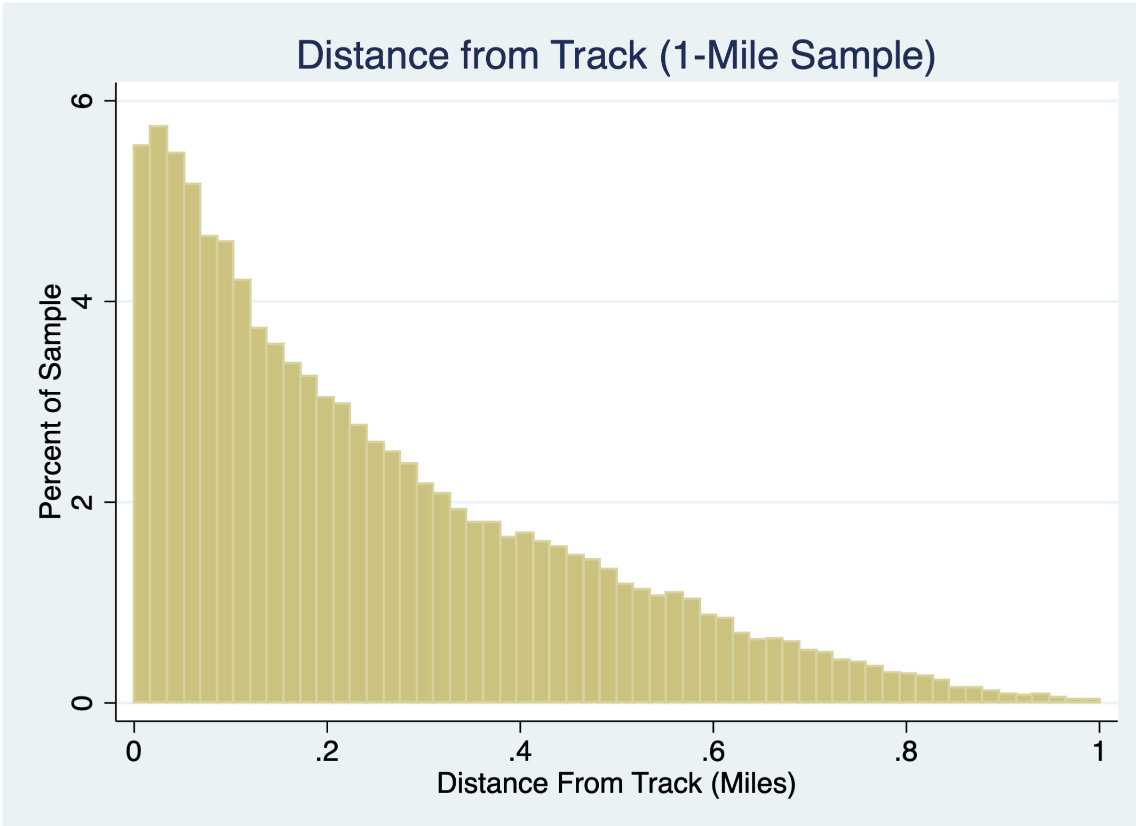


Figure 2.5: Distribution of Distance from Nearest Track Section

track as the nearest station, distance from track has a ceiling of 1 mile by design. However, most properties are significantly closer to a track section.

Table 2.4 shows the results of this analysis. Columns (1) and (4) replicate the all sales specifications in columns (3) and (6) in Table 2.2. Columns (2) and (5) indicate that there is no significant effect on prices for the subsample of surface-level stations, although this specification suffers from reduced power. However, column (3) suggests that in fact it is places further away from a train track (up to 1 mile) that are experiencing decreases in price, compared to properties closer to a track, although estimates are relatively imprecise. Note that distance to track is correlated with distance to station, but not the same thing. This relationship is not repeated in column (6) using a 0.5-mile radius. Regardless, there is no evidence that the negative estimates in Table 2.2 can be explained by noise pollution, since being closer to a surface-level train track does not appear to be driving the main effect.

Another potential mechanism that would square with a negative effect on prices is crime. Public tran-

Table 2.4: Investigating the Noise Mechanism

	Radius = 1 mile			Radius = 0.5 miles		
	(1)	(2)	(3)	(4)	(5)	(6)
Announcement						
Treat (2015)	-0.025**	-0.014	0.013	-0.029*	-0.004	-0.001
	(0.011)	(0.011)	(0.017)	(0.012)	(0.013)	(0.021)
Treat (2015) * Miles to Track	-	-	-0.093**	-	-	-0.024
			(0.038)			(0.087)
Service Start						
Treat (Month)	-0.024**	-0.012	0.019	-0.026*	0.0037	0.007
	(0.012)	(0.013)	(0.022)	(0.013)	(0.016)	(0.026)
Treat (Month) * Miles to Track	-	-	-0.104**	-	-	-0.021
			(0.044)			(0.096)
Surface-Level Sections Only		X	X		X	X

*Notes: Standard errors are in parenthesis. Each column corresponds to a variant of Equation (1). Columns (1) and (4) correspond to the first two rows of columns (3) and (6) in Table 2. Columns (2) and (5) restrict the sample to observations where the closest station is surface-level rather than underground. Columns (3) and (6) include an interaction term between treatment and the distance to the closest track section. ***p < 0.01, **p < 0.05, *p < 0.1*

sit does appear to positively affect (increase) crime in the literature, although it is not always the case. A common strategy is having been to look at temporary or permanent closures to service such as in Phillips & Sandler, (2015) and Wu & Ridgeway (2021), although these studies look at daytime subway service, which could produce a different result to nighttime service.

There are two broad channels by which we might expect the introduction of Night Tube service to be related to crime. First, from Becker's (1968) cost of crime model, a potential perpetrator will weigh up both the probability of being caught and the severity of punishment if caught against the reward for committing the crime and make a rational choice about whether to attempt the crime. One view in the urban literature is that crimes have both high transportation and search costs, increasing the risk of punishment, and that mass transit reduces both by providing inconspicuous mobility and a ready pool of potential targets respectively. By this logic, one would expect a positive relationship between mass transit access and crime rates. On the other hand, this clustering of people around Tube stations may decrease crime if there is an "eyes on the street" effect, which would increase the likelihood of being caught, and thus would decrease crime (Glaeser & Sacerdote, 1999).

The other way by which Night Tube service might be positively correlated with crime is behavioral in nature: Many nightlife activities and locations are strongly associated with alcohol consumption. Certain categories of crime might be more likely to be committed by those under the influence because of decreased inhibition and thus a decreased ability to rationally weigh up the pros and cons of one's actions. Night Tube service might work through this channel by either i) inducing more people to go out and drink in the first place or ii) substitution away from private transport such as taxis, where people have less opportunity to behave in a disorderly fashion or commit impulsive crimes.

To investigate the relationship between rollout of Night Tube service and crime, I utilize open-source police data on crime reports from 2011-2019 that contain crime location (longitude and latitude,[17] the month the crime occurred, and the type of crime committed. I collapse crime reports to the station-month level and run binomial regressions with the count of crimes occurring near a station on the left-hand side. Like with previous specifications, I include fixed effects for station and month, and cluster standard errors at the station level.

Table 2.5 shows the results of estimating a difference-in-differences regression of treatment on crime counts. For these regressions, I use the actual year of treatment rather than simply 2015, since it is more intuitive to think that crime would only increase once service has started, rather than in anticipation of service like is the case with house prices.[18] The first row shows that treatment is associated with roughly a 4% increase in total crime count. Looking at subsequent columns, it appears that this increase in crime is driven by incidents of theft and antisocial behavior, but not other types of crimes such as serious (violent) crimes and vehicle crimes. An increase in theft crimes would point to the Becker (1968) rational model of crime where the potential criminal is weighing up the costs and benefits of a crime, and having a Tube station open at night generates attractive targets. On the other hand, an increase in antisocial behavior is more indicative of a behavioral channel. I also run an event study for the crime outcome. Figure 2.6 shows the event study for a 1-mile specification.[19] There is no indication that there is a strong pre-trend in growth of crime counts before treatment. However, it should be noted that I am looking at aggregate crime counts, not just crimes committed at nighttime or during the weekends when the Night Tube would be operating. This is due to anonymization done by the London Metropolitan Police to protect the identity of those who are the victim of a crime. Therefore, these results are not conclusive, but do suggest a story that involves crime, since there is

[17] To protect anonymity of victims, the coordinates reported in the data are very slightly altered. For example, recording the location as the middle of the street if the crime occurred in someone's house, which may reveal the identity of the victim.

[18] For comparison and consistency with the house price results, I show a version of the crime event study that uses the 2015 definition of treatment in Figure 2.11.

[19] I also show that the event study looks similar if I use the 2015 definition of treatment, like I do for the house price event studies, in Figure 2.11.

Table 2.5: Testing the Crime Mechanism

	0.5 Mile	1 Mile
Total Crime	0.042***	0.038***
	(0.013)	(0.012)
Theft (All)	0.056**	0.059***
	(0.022)	(0.022)
Antisocial	0.072***	0.072***
	(0.017)	(0.016)
Serious Crime	-0.005	-0.007
	(0.013)	(0.011)
Vehicle	0.034	0.030
	(0.022)	(0.022)
Other	0.104	0.072
	(0.209)	(0.169)
Obs.	40,095	40,438

*Notes: Standard errors are in parenthesis. Each cell contains a Poisson parameter estimate of treatment at the station-month level on the count of crimes for a variety of categories. Antisocial behavior includes "personal, environmental, and nuisance antisocial behavior". Serious crime includes violent crimes, sexual assaults, possession and use of weapons, criminal damage, arson, and drug offenses. Vehicle crimes are defined as: "Including theft from or of a vehicle or interference with a vehicle". ***p < 0.01, **p < 0.05, *p < 0.1*

no specific reason why daytime or weekday crime counts should be changing differentially between treatment and control stations.

2.5.2 Heterogeneity Analysis

Finally, I cut the data by property type and location to see if there is any subgroup that is driving the main result. For property type, I look at whether a house or a flat (apartment) is being sold. For geography, I use the TFL zone that the nearest station resides in. The TFL network is divided into fare zones that are concentric in nature and usually used to determine the price charged for a journey: Zone 1 denotes the most central area of London, and zone 9 being the most outlying areas of the network.[20] For my purposes, I consider zones 1 and 2 to be central locations, zones 3 and above to be the 'suburb' areas, and zones 6 and above to be the outer suburb areas. Table 2.6 shows the cut by property type, where we can see that

[20]One station, Watford Junction, is technically outside of the fare zone system, and special fares are charged for journeys originating from or ending at this station. I simply code this station as being a part of a pseudo 'zone 10'.

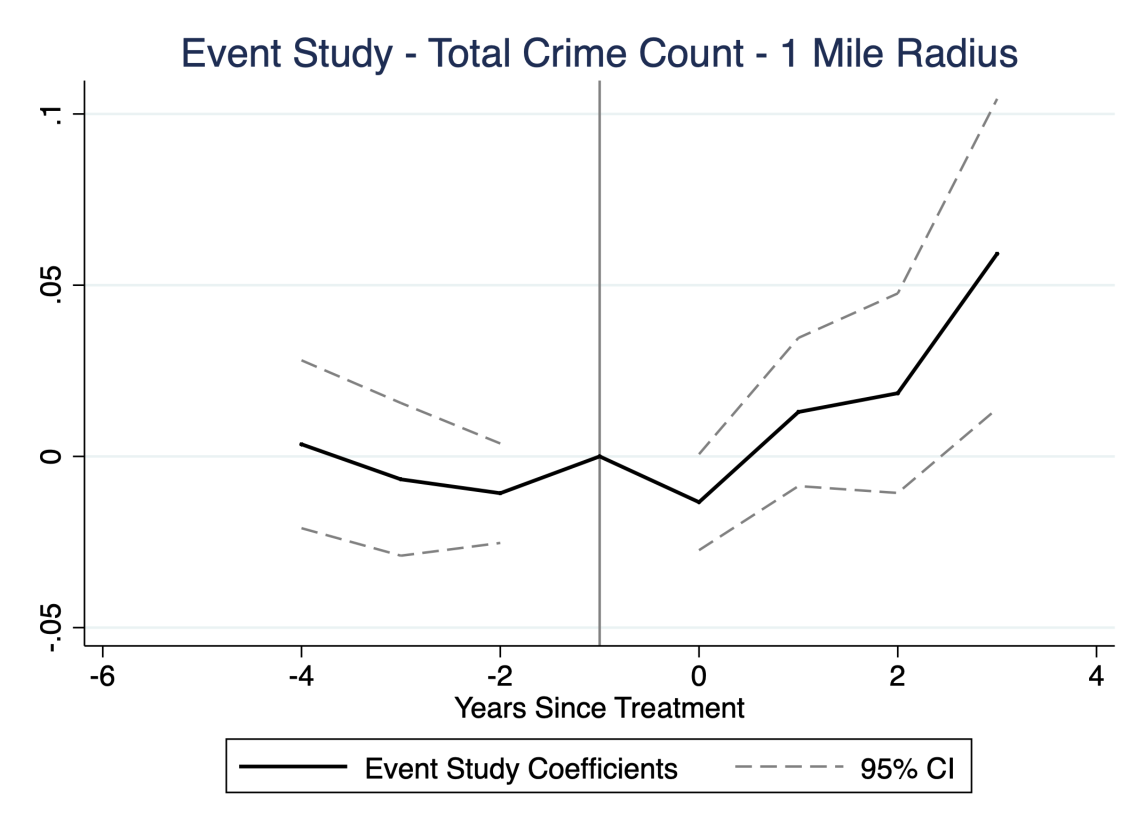


Figure 2.6: Event Study - Crime Count - All Sales

Table 2.6: Heterogeneity by Property Type

	(1)	(2)	(3)	(4)	(5)
2015 (All Sales)	-0.021**	-0.011	-0.034***	-0.029***	-0.028**
	(0.011)	(0.011)	(0.011)	(0.010)	(0.013)
Month (All Sales)	-0.021*	-0.014	-0.031***	-0.030***	-0.026*
	(0.011)	(0.013)	(0.011)	(0.011)	(0.014)
Heterogeneity					
Flats only		X			
Houses only			X		
Detached/Semi only				X	
Terraced only					X
N	755,417	471,642	283,775	98,009	185,766

*Notes: Standard errors are in parenthesis. Each estimate corresponds β_1 in a variant of Equation (1) using a particular subsample of the original data. Houses can include either detached, semidetached, or terraced properties. All specifications use a 1-mile radius. ***p < 0.01, **p < 0.05, *p < 0.1*

the samples of only houses, as well as the sub-types of houses, produce significant estimates. Similarly, in Table 2.7, all the significant estimates are coming from areas outside Central London. This suggests that the main results shown in Table 2.2 may be driven by houses located in neighborhoods outside of the center of London, and I further support this suggestion in Table 2.16 by cutting by property type and zone simultaneously, finding that specifications looking only at houses outside of zones 1 and 2 (columns 3 and 4) show consistently significant effects, although these regressions are lower-powered due to reduced sample sizes. Additionally, caution must be exercised here since there is still some overlap in the confidence intervals of estimates in these cuts. An alternative approach is to interact treatment with a function of distance from the city center. In Table 2.8 I show the results from running a regression interacting a quadratic function of distance from the latitude/longitude coordinates of the center of London.[21] This specification gives a slightly different conclusion: The treatment effect curve exhibits an upside-down U-shape, where properties in both very central areas and outlying areas exhibit negative house price effects, whereas areas outside the very center but not in the suburbs have a small or null effect. A story that would be consistent with this result is if very central areas experience negative effects due to increased nightlife, and outlying areas see an increase in crime around stations, but areas in-between do not, possibly because they were previously more likely to be served by existing transport like night buses or black cabs, so there is a less drastic change in foot traffic in these areas after treatment.

[21] This point has traditionally been defined as the Equestrian Statue of Charles I, located close to Charing Cross Tube station.

Table 2.7: Heterogeneity by Location (Zone)

	(1)	(2)	(3)	(4)	(5)
2015 (All Sales)	-0.021**	-0.025**	-0.034	-0.009	-0.027**
	(0.011)	(0.010)	(0.035)	(0.016)	(0.012)
Month (All Sales)	-0.021*	-0.025**	-0.032	-0.012	-0.025**
	(0.011)	(0.011)	(0.037)	(0.016)	(0.012)
Heterogeneity					
Postcode District FE		X			
Zone 1 (Very Central)			X		
Zone 1+2 (Central)				X	
Zones 3+ (Suburbs)					X
N	755,417	755,417	68,800	329,089	426,328

Notes: Standard errors are in parenthesis. Each estimate corresponds β_1 in a variant of Equation (1) using a particular subsample of the original data. Fr postcode district fixed effects, there are 328 such units. Zones 1 and 2 represent central areas of the city. Zones 3+ represent relatively more outlying areas.
****p < 0.01, **p < 0.05, *p < 0.1*

Table 2.8: Interactions with Distance from Center of London

	(1)	(2)
	All Sales	Repeat Sales
Treat (2015)	-0.100***	-0.115***
	(0.028)	(0.024)
Treat * Distance	0.022**	0.032***
	(0.009)	(0.008)
Treat * Distance^2	-0.0012*	-0.0021***
	(0.0006)	(0.0006)
N	710,000	198,000

*Notes: Standard errors are in parenthesis. Columns (1) and (2) are variations of Table 2, column (3) for the all sales and repeat sales specifications using a 1-mile radius. In addition to the main treatment effect, I also add a variable (Distance) representing the straight-line distance to the center point of London in miles. I model the treatment effect as a quadratic function of this distance. All specifications include station and time fixed effects. In column (2), I include property fixed effects so that I am only comparing sales of the same property that happen before and after treatment. I drop observations matched to stations that are within a 10-minute walk from another station, to avoid including sales that occur close to both treated and control stations. ***p < 0.01, **p < 0.05, *p < 0.1*

To see whether there is anything observably different about treated units in more outlying areas compared to those in more central areas prior to treatment, I use the same aggregate MSOA statistics from the 2011 UK Census as in Panel B of Table 2.1. Table 2.9 shows selected descriptive statistics from this analysis. I compare station-neighborhoods in central (zones 1-2) vs. more outlying areas (zones 3+). The first thing to note here is that the percentage of households living in house and in an owned property is significantly higher in suburb areas than in more central areas. How then do the demographics of households living in houses outside of central London compare to those who are more likely to rent apartments in central London? Although not a large difference, there are more children and more elderly people living in the suburb areas, whereas working-age adults are slightly more likely to live in the center of the city. We can also see that a larger percentage of households in living in zones 3+ contain dependent children. Despite the difference in tenure and property type, there is no perceptible difference in the average deprivation index score.

2.6 Conclusion

In this paper, I test for a causal relationship between subway access and property values in the novel context of nighttime service, using the rollout of the Night Tube and comparing changes in prices of sales near stations that did and did not receive night service using a differences-in-differences model. My most plausibly causal estimates show a $1.3-2.6\%$ decrease in prices associated with treatment. This effect essentially attenuates the preexisting premium associated with living close to any Tube station by $15-30\%$.

To look at possible mechanisms, I investigate noise pollution and crime. I find no evidence that noise is related to my main result, and show suggestive, but not conclusive evidence that an increase in certain types of crime increase, which has been shown in the literature to be negatively correlated with property prices. I also show that the result is significant specifically for houses (vs. flat) outside of the central areas of the city. The statistics in Table 2.9 present a possible unifying explanation for the crime and heterogeneity results. Households in outlying areas, that tend to also have a greater share of houses, also are more likely to be homeowners and are more likely to contain dependent children. Parents/guardians of dependent children may be more sensitive to crime levels in terms of willingness to pay for properties. It may also be the case that homeowners are more involved in their local communities and are also more sensitive to crime. However, I get a slightly different story when I interact distance from the center with treatment, and I am unable to definitively say which one is correct.

One possible knock-on effect of this result is a sorting of certain demographic groups into and out of ar-

Table 2.9: Station-Neighborhood Characteristics by Zone (2011 Census)

	Night Tube Stations			
	Zones 1-2		Zones 3+	
	Mean	S.D.	Mean	S.D.
	(1)	(2)	(3)	(4)
Age				
% 0-17	15.1	5.0	20.2	3.0
% 18-59	71.7	5.4	63.0	5.5
% 60+	13.2	3.9	16.8	4.9
Deprivation				
Avg. Deprivation Index	0.97	0.21	0.91	0.19
Dwelling Type				
% Shared Dwelling	2.3	2.1	0.8	0.9
% House	15.8	12.3	55.7	15.5
Tenure				
% Own	29.9	10.0	55.2	14.0
% Social Housing	32.9	16.7	15.2	9.8
HH Composition				
% with Dep. Children	19.8	7.4	31.4	5.4
N	91		64	

Notes: Each pair of columns represents a different subsample of stations and their associated neighborhoods, defined by the Middle Super Output Layer (MSOA) that they are contained within. Columns (1)-(2) and (3)-(4) show means and standard deviations for treatment stations in central areas (zones 1-2) and suburb (zones 3+) respectively. The deprivation index takes integer values between 0 and 4 for each household, and the score is based on how many of the four criteria for deprivation are met for that household. A household occupies a share dwelling if two or more household units occupy a single space.

eas with NT service. If, because of increased crime and access to the NT, families with children and older adults move away from NT areas, and younger adults without families move in the opposite direction, then we might expect to see a demographic shift by age group. I test this in Table 2.17 by running a simple difference-in-differences in NT service and time on the percentage of the population in an MSOA that is in each age range (0-20, 21-35, 36-64, 65+) using Office for National Statistics (ONS) yearly population estimates, but I find no evidence of any shift in age composition of treated vs. control neighborhoods, which is not entirely surprising given the relative inelasticity of housing stock in these areas.

This paper's results have relevance for both the urban literature on mass transit access and future policy-making decisions. This paper provides evidence that the positive effects of transit access on property prices cannot be automatically extended to nighttime service. On the policy side, there are many cities around the world that have subway service that does not run 24 hours a day. For these places, there is a relevant cost-benefit analysis about whether it is beneficial to extend the service to the nighttime, and in what capacity. To the extent that hedonic analysis can reveal consumer preferences for local amenities, practitioners can better understand who might be the winners and losers of 24-hour service and address the causes of these losses (e.g. crime).

2.7 References

Andersson, H., Jonsson, L. & Ögren, M. (2010) Property Prices and Exposure to Multiple Noise Sources: Hedonic Regression with Road and Railway Noise. *Environ. Resource Econ* 45, 73–89 (2010). https://doi.org/10.1007/s10640-009-9306-4

Baum-Snow, N. & Matthew E. Kahn. (2000). The effects of new public projects to expand urban rail transit. *Journal of Public Economics* 77 (2000) 241–263

Becker, G. S. (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy*, 76(2), 169–217. http://www.jstor.org/stable/1830482

Black, S. E. (1999). Do Better Schools Matter? Parental Valuation of Elementary Education. *The Quarterly Journal of Economics*, 114(2), 577–599. http://www.jstor.org/stable/2587017

Block, R., & Block, C.R. (2000). The Bronx and Chicago: Street Robbery in the Environs of Rapid Transit Stations. DOI:10.4135/9781452220369.N11

Ceccato, V. & M. Wilhelmsson (2020) Do crime hot spots affect housing prices?, *Nordic Journal of Criminology*, 21:1, 84-102, DOI: 10.1080/2578983X.2019.1662595

Croome, D. & Jackson, A. (1993). Rails Through the Clay — A History of London's Tube Railways. Capital Transport Publishing; 2nd edition (January 1, 1993). ISBN 978-1-85414-151-4.

Day, J.R. & Reed, J. (2008). The Story of London's Underground (10th ed.). Capital Transport Publishing; 10th edition. ISBN 978-1-85414-316-7.

Diao, M., Delon Leonard, & Tien Foo Sing (2017). Spatial-difference-in-differences models for impact of new mass rapid transit line on private housing values, *Regional Science and Urban Economics*, Volume 67, 2017, Pages 64-77, ISSN 0166-0462, https://doi.org/10.1016/j.regsciurbeco.2017.08.006.

Diao, M., Qin, Y. & Sing, T.F. (2016), Negative Externalities of Rail Noise and Housing Values: Evidence from the Cessation of Railway Operations in Singapore. *Real Estate Economics*, 44: 878-917. https://doi.org/10.1111/1540-6229.12123

Dubé, Jean, Diègo Legros, Marius Thériault, & François Des Rosiers (2014). A spatial Difference-in-Differences estimator to evaluate the effect of change in public mass transit systems on house prices, *Transportation Research Part B: Methodological*, Volume 64, 2014, Pages 24-40, ISSN 0191-2615, https://doi.org/10.1016/j.trb.2014.02.007.

Duranton, Gilles, & Matthew A. Turner. 2011. "The Fundamental Law of Road Congestion: Evidence from US Cities." *American Economic Review*, 101 (6): 2616-52.

Gibbons, S. (2004). The Costs of Urban Property Crime. *The Economic Journal*, 114(499), F441–F463. http://www.jstor.org/stable/3590167

Gibbons, S. & Machin, S. (2005). Valuing rail access using transport innovations, *Journal of Urban Economics*, Volume 57, Issue 1, 2005, Pages 148-169, ISSN 0094-1190, https://doi.org/10.1016/j.jue.2004.10.002.

Gibbons, S. & Machin, S. (2008). Valuing school quality, better transport, and lower crime: evidence from house prices, *Oxford Review of Economic Policy*, Volume 24, Issue 1, spring 2008, Pages 99–119, https://doi.org/10.1093/oxrep/grn008

Glaeser, E. L., & Sacerdote, B. (1999). Why is There More Crime in Cities? *Journal of Political Economy*, 107(S6), S225–S258. https://doi.org/10.1086/250109

Heblich, Stephan, Stephen J. Redding, & Daniel M. Sturm. (2020). The making of the modern metropolis: Evidence from London. *The Quarterly Journal of Economics* (2020), 2059–2133. doi 10.1093/qje/qjaa014

Herskovic, Luis. (2020). The Effect of Subway Access on School Choice, *Economics of Education Review*, Volume 78, 2020, 102021, ISSN 0272-7757,

Jacobs, Jane (1961). The Death and Life of Great American Cities. January 1, 1961, Random House. ISBN: 9780375508738

Linden, Leigh, & Jonah E. Rockoff. (2008). "Estimates of the Impact of Crime Risk on Property Values from Megan's Laws." *American Economic Review*, 98 (3): 1103-27

Lynch, Allen K. & David W. Rasmussen (2001) Measuring the impact of crime on house prices, *Applied Economics*, 33:15, 1981-1989, DOI: 10.1080/00036840110021735

McMillen, D.P. & McDonald, J. (2004), Reaction of House Prices to a New Rapid Transit Line: Chicago's Midway Line, 1983–1999. *Real Estate Economics*, 32: 463-486. https://doi.org/10.1111/j.1080-8620.2004.00099.x

Nelson, J. P. (2004). Meta-Analysis of Airport Noise and Hedonic Property Values: Problems and Prospects. *Journal of Transport Economics and Policy*, 38(1), 1–27. http://www.jstor.org/stable/20173043

Phillips, D.C. & Sandler, D. (2015). Does public transit spread crime? Evidence from temporary rail station closures, *Regional Science and Urban Economics*, Volume 52, 2015, Pages 13-26, ISSN 0166-0462, https://doi.org/10.1016/j.regsciurbeco.2015.02.001.

Rao, Santosh (2016). London's new late night alternative: The Night Tube + Uber. Accessed October 12, 2021 at: https://medium.com/uber-under-the-hood/londonsnew-late-night-alternative-the-night-tube-uber

-8f38e56de983#.v2f8853b7

Roberts, Johnathan (2016). Exploring The Night Tube Part 1: Making the Case. November 1 2016. Accessed October 12, 2021 at: https://www.londonreconnections.com/2016/night-tube-part-1-making-the-case/

Stacy, Mungo (2016). 4LM+1. August 24 2016. Accessed October 12, 2021 at: https://www.railengineer.co.uk/4lm1/

Statistica (2021a): Number of passenger journeys on the London Underground (UK) from 2000/01 to 2020/21. Published June 2021. Accessed October 12, 2021 at: https://www.statista.com/statistics/304852/passenger-journeys-on-the-london-underground/

Statistica (2021b): Number of journeys on the London Overground from January 2019 to April 2021 (in million journeys). Published June 2021. Accessed October 12, 2021 at: https://www.statista.com/statistics/412546/london-overground-journeys/

Statistica (2021c): Number of passenger journeys on the Docklands Light Railway between 2000/01 and 2020/21. Published June 2021. Accessed October 12, 2021 at: https://www.statista.com/statistics/305527/passenger-journeys-on-docklands-light-railwayuk/

TFL (2010). Measuring Public Transport Accessibility Levels (PTALs) - Summary. Accessed October 12, 2021 at: https://s3-eu-west-1.amazonaws.com/londondatastore-upload/PTAL-methodology.pdf

TFL (2017). Night Tube Implementation. Customer Service and Operational Performance Panel. 2 March 2017. Accessed October 12, 2021 at: https://content.tfl.gov.uk/csopp-20170302-part-1-item08-night-tube-implementation.pdf

Volterra (2014). TfL 90993 – Impact of the Night Tube on London's Night-Time Economy. Prepared by Volterra Partners for TfL and London First. September 2014 Accessed October 12, 2021 at: https://content.tfl.gov.uk/night-time-economy.pdf

Wu, Y. & Ridgeway, G. Effect of public transit on crime: evidence from SEPTA strikes in Philadelphia. *J. Exp. Criminology* 17, 267–286 (2021). https://doi.org/10.1007/s11292-020-09416-z

Yang, Jiawen, Jason Cao, & Yufei Zhou (2021). Elaborating non-linear associations and synergies of subway access and land uses with urban vitality in Shenzhen, *Transportation Research Part A: Policy and Practice*, Volume 144, 2021, Pages 74-88,

Zhou, Z., Chen, H., Han, L. & Zhang, A. (2021), The Effect of a Subway on House Prices: Evidence from Shanghai. *Real Estate Economics*, 49: 199-234. https://doi.org/10.1111/1540- 6229.12275

2.8 Appendix

2.8.1 Additional Figures

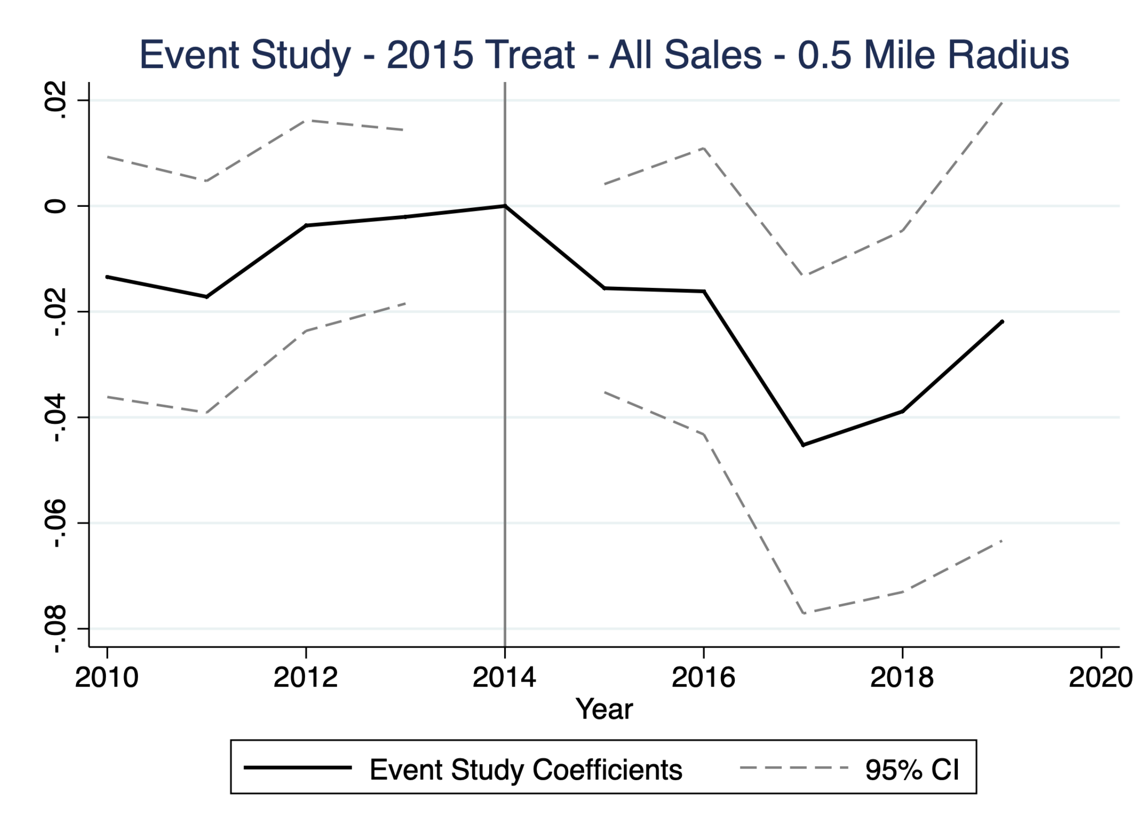


Figure 2.7: Event Study – All Sales – 0.5-Mile Radius

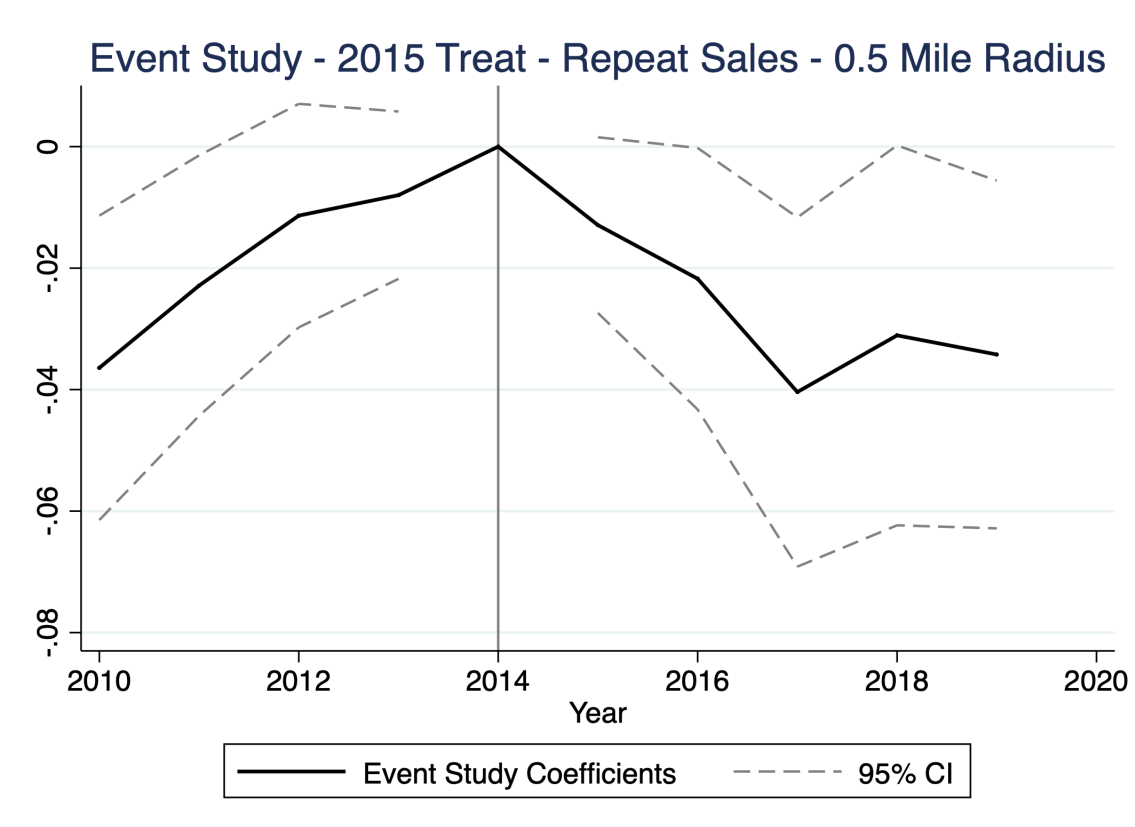


Figure 2.8: Event Study – Repeat Sales Only – 0.5-Mile Radius

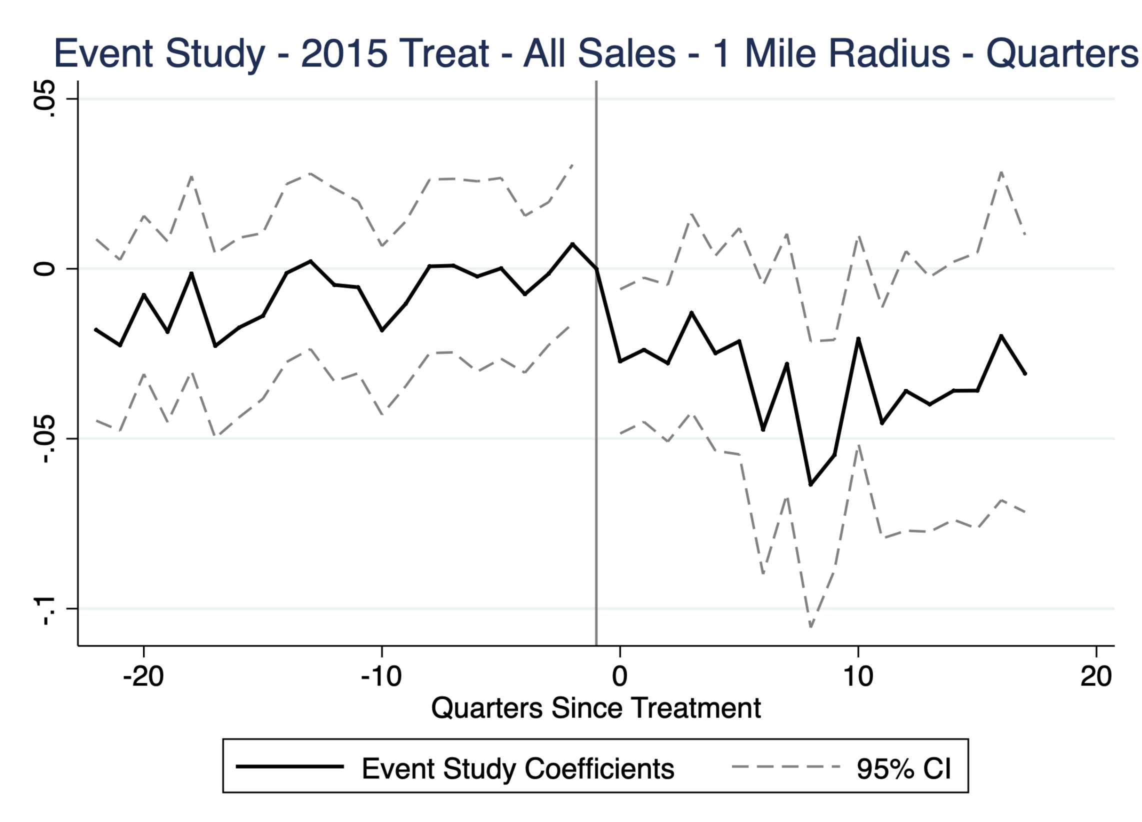


Figure 2.9: Event Study (Quarter) - All Sales - 1 Mile Radius

Figure 2.10: Event Study (Quarter) - Repeat Sales - 1 Mile Radius

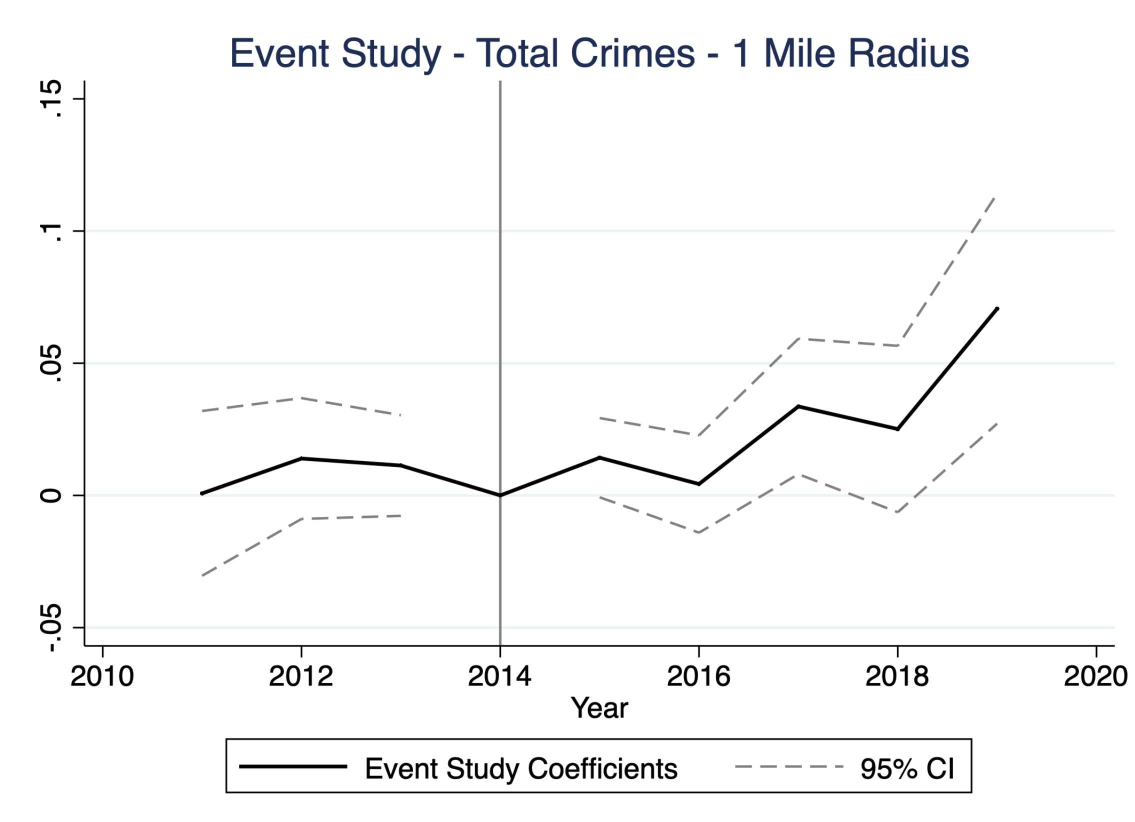


Figure 2.11: Crime Event Study - 2015 Treatment Definition

2.8.2 Additional Tables

Table 2.10: Event Study Estimates

	All Sales		Repeat Sales	
	(1)	(2)	(3)	(4)
Pre-treatment				
2010	-0.013	-0.011	**-0.027***	-0.022
	(0.011)	(0.011)	(0.014)	(0.014)
2011	-0.011	-0.010	**-0.020***	-0.015
	(0.010)	(0.010)	(0.012)	(0.012)
2012	0.001	0.004	-0.007	-0.004
	(0.009)	(0.009)	(0.009)	(0.010)
2013	-0.004	-0.002	-0.004	-0.002
	(0.007)	(0.007)	(0.007)	(0.007)
Post-Treatment				
2015	-0.010	-0.010	-0.008	-0.008
	(0.008)	(0.009)	(0.006)	(0.006)
2016	**-0.020***	**-0.020***	**-0.020****	**-0.023****
	(0.011)	(0.018)	(0.009)	(0.010)
2017	**-0.046*****	**-0.050*****	**-0.036*****	**-0.037*****
	(0.015)	(0.016)	(0.019)	(0.013)
2018	**-0.033****	**-0.035****	**-0.027***	**-0.029***
	(0.016)	(0.017)	(0.015)	(0.015)
2019	-0.027	-0.031	**-0.032****	**-0.033****
	(0.018)	(0.019)	(0.014)	(0.014)
N	755,000	710,000	211,000	198,000
S.E. Clustering				
Station	X	X	X	X
Robustness				
Exclude “Close Stations”		X		X

Table 2.11: Repeat Sales Specifications with Station Linear Time Trends

	Treat = 2015		Treat = Month	
	(1)	(2)	(3)	(4)
Repeat Sales Only				
Radius = 1 Mi.	-0.017**	-0.018**	-0.026**	-0.022*
	(0.009)	(0.009)	(0.011)	(0.011)
Radius = 0.5 Mi.	-0.022**	-0.024**	-0.028***	-0.023**
	(0.009)	(0.009)	(0.010)	(0.011)
S.E. Clustering				
Station	X	X	X	X
Robustness				
Exclude "Close Stations"		X		X

*Notes: Standard errors are in parenthesis and are clustered at the station level. Each estimate corresponds to β_1 in a variant of Equation (1) that absorbs property fixed effects so that I am only comparing sales of the same property that happen before and after treatment. Each house sale is matched to the closest station in my sample. I windsorize the top and bottom 1% of sample to avoid extreme outliers impacting the estimate. The treatment indicator for an observation is equal to 1 if that station has received Night Tube service by the relevant month or year. I restrict the sample to observations whose nearest-neighbor station is within a certain distance, either 1 or 0.5 miles, to avoid properties that are not close to any station being identified as treated units. Columns (2) and (4) drop observations matched to stations that are within a 10-minute walk from another station, to avoid including sales that occur close to both treated and control stations. ***p < 0.01, **p < 0.05, *p < 0.1*

Table 2.12: All Sales Regressions with Multiple Sales Sample

	Treat = 2015			Treat = Month Service Starts		
	(1)	(2)	(3)	(4)	(5)	(6)
All Sales						
Radius = 1 Mi.	-0.020***	-0.020*	-0.024**	-0.022***	-0.022**	-0.026**
N= 213,000	(0.003)	(0.011)	(0.011)	(0.003)	(0.011)	(0.011)
Radius = 0.5 Mi.	-0.023***	-0.023*	-0.028**	-0.024***	-0.024**	-0.028**
N= 142,000	(0.004)	(0.012)	(0.012)	(0.004)	(0.012)	(0.012)
S.E. Clustering						
Robust	X			X		
Station		X	X		X	X
Robustness						
Exclude "Close Stations"			X			X

Notes: Standard errors are in parenthesis. This table is a replication of the All Sales estimates from Table 2, but using the subsample of properties that are sold at least twice during the sample period (2010-2019) to be more comparable to the sample used for the Repeat Sales (i.e. w/ property fixed effects) specifications. Each estimate corresponds to β_1 in a variant of Equation (1). Each house sale is matched to the closest station in my sample. I windsorize the top and bottom 1% of sample to avoid extreme outliers impacting the estimate. The treatment indicator for an observation is equal to 1 if that station has received Night Tube service by the relevant month or year. I restrict the sample to observations whose nearest-neighbor station is within a certain distance, either 1 or 0.5 miles, to avoid properties that are not close to any station being identified as treated units. All specifications include station and time fixed effects. Columns (3) and (6) drop observations matched to stations that are within a 10-minute walk from another station, to avoid including sales that occur close to both treated and control stations. Around 6% of observations (N) fall under this definition.

$***p < 0.01$, $**p < 0.05$, $*p < 0.1$

Table 2.13: House-Price Premiums for Regular Tube Service

	(1)	(2)	(3)	(4)
Less than 1-Mile	0.085***			
	(0.022)			
Less than 0.5 Miles		0.066***		
		(0.013)		
Miles to station			-0.064***	-0.129***
			(0.017)	(0.030)
Miles to station^2				0.024**
				(0.012)
N	1,011,423	1,011,423	1,011,423	1,011,423

*Notes: Standard errors are in parenthesis. Estimates correspond to treatment effect parameters from regressions of the natural log of sale price on a distance variable, station and year fixed effects, and controls for property characteristics. I restrict the sample to a maximum of 3 miles to the nearest station to avoid comparing properties is vastly different areas, although results do not change drastically if I use the whole sample. In the first two columns, the parameter of interest is an indicator for if the property is within 1 or 0.5 miles from a tube station respectively. Column (3) uses the straight-line distance, and column (4) adds a squared term. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$*

Table 2.14: Specifications Absorbing Postcode Fixed Effects

	Treat = 2015			Treat = Month Service Starts		
	(1)	(2)	(3)	(4)	(5)	(6)
Within Postcode						
Radius = 1 Mi.	-0.026***	-0.026**	-0.030***	-0.028***	-0.028***	-0.032***
N= 763,328	(0.002)	(0.010)	(0.010)	(0.002)	(0.011)	(0.011)
Radius = 0.5 Mi.	-0.027***	-0.027**	-0.030**	-0.029***	-0.029**	-0.032***
N= 509,851	(0.002)	(0.011)	(0.019)	(0.002)	(0.012)	(0.012)
S.E. Clustering						
Robust	X			X		
Station		X	X		X	X
Robustness						
Exclude "Close Stations"			X			X

*Notes: Standard errors are in parenthesis. Each estimate corresponds β_1 in a variant of Equation (1). Each house sale is matched to the closest station in my sample. I windsorize the top and bottom 1% of sample to avoid extreme outliers impacting the estimate. The treatment indicator for an observation is equal to 1 if that station has received Night Tube service by the relevant month or year. I restrict the sample to observations whose nearest-neighbor station is within a certain distance, either 1 or 0.5 miles, to avoid properties that are not close to any station being identified as treated units. All specifications include station, time, and postcode fixed effects. Columns (3) and (6) drop observations matched to stations that are within a 10-minute walk from another station, to avoid including sales that occur close to both treated and control stations. Around 6% of observations (N) fall under this definition. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$*

Table 2.15: Non-Windsorized Specifications

	Treat = 2015			Treat = Month Service Starts		
	(1)	(2)	(3)	(4)	(5)	(6)
All Sales						
Radius = 1 Mi.	-0.014***	-0.014	-0.018	-0.016***	-0.016	-0.021*
N= 773,035	(0.002)	(0.011)	(0.011)	(0.002)	(0.012)	(0.012)
Radius = 0.5 Mi.	-0.010***	-0.010	-0.013	-0.010***	-0.010	-0.014
N= 516,469	(0.003)	(0.012)	(0.012)	(0.003)	(0.013)	(0.014)
Repeat Sales Only						
Radius = 1 Mi.	-0.014***	-0.014	-0.019*	-0.019***	-0.019*	-0.022**
N= 217,248	(0.002)	(0.010)	(0.010	(0.002)	(0.011)	(0.011)
Radius = 0.5 Mi.	-0.014***	-0.014	-0.018	-0.017***	-0.017	-0.020*
N= 144,874	(0.003)	(0.011)	(0.011)	(0.003)	(0.012)	(0.012)
S.E. Clustering						
Robust	X			X		
Station		X	X		X	X
Robustness						
Exclude "Close Stations"			X			X

*Notes: Standard errors are in parenthesis. Each estimate corresponds β_1 in a variant of Equation (1). Each house sale is matched to the closest station in my sample. The treatment indicator for an observation is equal to 1 if that station has received Night Tube service by the relevant month or year. I restrict the sample to observations whose nearest-neighbor station is within a certain distance, either 1 or 0.5 miles, to avoid properties that are not close to any station being identified as treated units. All specifications include station and time fixed effects. For repeat sales specifications, I include property fixed effects so that I am only comparing sales of the same property that happen before and after treatment. Columns (3) and (6) drop observations matched to stations that are within a 10-minute walk from another station, to avoid including sales that occur close to both treated and control stations. Around 6% of observations (N) fall under this definition. ***p < 0.01, **p < 0.05, *p < 0.1*

Table 2.16: Heterogeneity by Property Type and Zone

	Houses				**Flats**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treat=2015 (All Sales)	-0.034***	-0.028	-0.034***	-0.046***	-0.011	-0.010	-0.011	0.019
	(0.011)	(0.023)	(0.012)	(0.011)	(0.011)	(0.015)	(0.012)	(0.021)
Treat=Month (All Sales)	-0.031***	-0.023	-0.032***	-0.024*	-0.014	-0.013	-0.014	0.028***
	(0.011)	(0.022)	(0.012)	(0.013)	(0.013)	(0.017)	(0.015)	(0.009)
N	[283,784]	[56,860]	[226,924]	[43,073]	[471,642]	[272,229]	[199,413]	[19,086]
Property Type								
Houses only	X	X	X	X				
Flats only					X	X	X	X
Geography								
Zones 1/2 (Central)		X				X		
Zones 3+ (Suburbs)			X				X	
Zones 6+ (Outer-Suburbs				X				X

Notes: Standard errors are in parenthesis. Each estimate corresponds β_1 in a variant of Equation (1) using a subsample of the original data. Houses can include either detached, semidetached, or terraced properties. Zones 1 and 2 represent central areas of the city. Zones 3/4/5 I define as the "inner suburbs", and zones 6 and above as the "outer suburbs". ***p < 0.01, **p < 0.05, *p < 0.1

Table 2.17: Effects on Local Age Distribution

	(1)	(2)	(3)	(4)
Age	(0-20)	(21-34)	(35-64)	(65+)
Treat=2015 (All Sales)	0.0011	-0.0009	0.0004	-0.0005
	(0.0013)	(0.0026)	(0.0019)	(0.0009)
Post	0.0002	-0.0011	0.0014	-0.0005
	(0.0006)	(0.0012)	(0.0009)	(0.0004)
NT	-0.0287***	0.0476***	-0.0104**	-0.0085
	(0.0054)	(0.0095)	(0.0046)	(0.0054)
N	2,601	2,601	2,601	2,601

*Notes: Standard errors are in parenthesis and are clustered at the MSOA-level.. Estimates correspond to treatment effect parameters from pre-post/NT diff-in-diff regressions where percentage (%) of the population in an MSOA that is in the given age range is the outcome variable. Treat=2015 indicates the DID coefficient, using the announcement date for treatment timing and the 1-mile radius criterion for inclusion in the sample. ***p < 0.01, **p < 0.05, *p < 0.1*

CHAPTER 3

Gender Diversity & Municipal Spending: Evidence from California City Councils

3.1 Introduction

Although women's economic and social standings have improved over the last several decades, a telling sign of persistent and structural inequality is that women are underrepresented in positions of decision-making and power. As of 2018, fewer than 10% of CEOs and 24% of national-level politicians worldwide are women (Hora, n.d.; IPU, 2019). The political arena is especially imbalanced: In the latest World Economic Forum (WEF) Global Gender Gap Index, which ranks 153 countries on various gender parity indices, although health and educational attainment gaps now stand at 3 and 4% respectively, the political empowerment gap is a substantial 75% (WEF, 2020). Aside from the moral and social justifications for having diverse leadership in politics, an interesting question to ask is whether having more equal representation will actually change the decisions that are made by those with policy-making power. Political economy models differ in their predictions about the importance of the characteristics of representatives on policy outcomes, so it is uncertain whether there should be an effect. A modest literature has examined the impact of politician gender on public expenditure outcomes, but the results are mixed, uncertain, and limited to specific settings.

In this paper, I estimate the effect of gender diversity in municipal government councils on the size and composition of local public spending, using comprehensive election data and yearly expenditure figures from California municipalities. My empirical strategy is a regression discontinuity design (RDD) that exploits close elections between a male and a female candidate for open seats on city councils. Since city councils in this setting are relatively small,[1] a single close election where the female candidate wins (vs. a male candidate) causes a meaningful swing in the gender composition of the council. In addition to looking at total spending, I analyze the impact on specific sectors, including police/public safety, transportation, health, culture, and parks & recreation. For most outcomes I do not find significant and/or robust effects. However, I do find that a narrow female victory vs. a male candidate increases the share of spending allocated to transportation by around 2 percentage points across the four-year term of the new female councilor.

Digging deeper into the mechanisms of diversity, I look specifically at the effect of a city council gaining its first female councilor, as well as the effect of a council going from a male majority to a female majority. This analysis indicates that the transportation results are driven by councils shifting from a male majority to

[1] Relative to city councils both in other states and other countries (Mathews, 2020)

a female majority. Further disaggregation of the transportation category indicates that the observed effect is concentrated in capital outlays rather than changes in operational expenditures. Finally, I explore differences in preferences for public spending between men and women as a potential explanation for the transportation results, as well as the lack of an effect in other categories (including overall spending) using California survey data.

These results speak to the theoretical ambiguity about the effects of candidate characteristics like gender. Increasing women's representation may lead to an allocation of resources that benefits or matches the preferences of women. This is predicated on the idea that politicians will pursue agendas that benefit certain segments of the population (in this case, their own gender), so it is then natural to ask whether this is the case: For example, in one of the classic models of voting, the Median Voter Theorem, the preferences of politicians should be of no consequence (Downs, 1957). However, other models such as the Citizen-Candidate model conclude that the underlying preferences of candidates will ultimately drive their policy decisions (Besley & Coate, 1997). Under the Citizen-Candidate model, a shift in the gender composition of politicians would be expected to shift the composition of public spending to better reflect the preferences of women. A good example of evidence for the citizen-candidate model in the empirical literature is Lee, Moretti, & Butler (2004).

However, in order for there to be any effect in this model, we still need two more conditions to hold: First, there has to be a difference in the relative preferences of men and women, otherwise there is no shift that can occur. Using data on preferences for proposed public spending changes & increases, I do find divergence in spending priorities between men and women, but I am unable to detect statistically significant differences for transportation spending specifically to complement and/or explain the results of my main analysis.

Second, representation has to translate into influence that affects female councilors' ability to guide spending decisions; if there are additional power structures in politics that prevent women from making changes, then we may not see any effect. Much of the older political science literature on representation implicitly assumed that representation would automatically translate into policy change that represents that group, but leaves the exact mechanism by which this occurs as a "black box" (Kanthak & Krause, 2012). One of the more prominent theories about how/if diversity causes representative change is that of "critical mass", whereby a certain proportion of women on a legislative body is required before the policy priorities of body as a whole change. As noted by one study: "It is not clear whether sheer numbers of women should have a proportional impact, a curvilinear impact, or an absolute numbers impact on policy-making around women's interest" (Beckwith & Cowell-Meyers, 2007). To this end, I utilize variation in the prior composition of councils and show that the significant effect on transportation spending occurs when the council shifts to a female majority, which is consistent with the idea that the marginal impact of an additional woman on the council is not constant.

These results are also related to the existing economics literature studying the effect of politician gender on government policy and spending. Early literature used lab games dictator games and found that hypothetical policy choices made by women favored a higher degree of equality of outcome, whereas male subjects placed a higher emphasis on maximizing efficiency. More recently, results from natural and field experiments are more mixed, and come from several disparate settings. Quasi-experimental research in developed countries has repeatedly found no effect of politician gender on composition of expenditures (Rigon & Tanzi, 2012; Chen, 2013; Ferreira & Gyourko, 2014; Bagues & Campa, 2017; Geys & Sorenson, 2019), even though non-experimental studies by Svaleryd (2009) and Funk & Gathmann (2015) had found evidence that politician gender matters in Sweden and Switzerland respectively. In developing countries, there is more evidence suggesting that gender composition matters for direct spending outcomes. The seminal paper is Chattopadhyay & Duflo (2004), who found effects in India on the likelihood that infrastructure projects that women had specifically brought up in village meetings were completed, suggesting that female politicians implemented policies that were of particular interest to their female constituents. Clots-Figueras (2012) also provides evidence from India, showing that electing female politicians increases educational metrics in urban areas (but not rural ones).

In many cases, gender quotas and affirmative action policies have been used in an attempt to address the issue of low gender diversity. However, unlike papers such as Chattopadhyay & Duflo (2004) and Rigon & Tanzi (2012), I am studying changes in female representation in the absence of a government policy or quota, so changes in female representation over this period may be more likely to persist in the long-term, aligning this paper more closely with Clots-Figueras (2012) and Funk & Phillips (2019). In the U.S., the main piece of evidence comes from Ferreira & Gyourko (2014), who examine mayoral elections and find that female mayors do not affect various measures of spending and employment. In my setting, instead of looking at binary elections of a single office, looking at local councils adds the explicit and observable dimension of the interaction between men and women when making public spending decisions, where the composition of the council may have heterogeneous effects on spending depending on whether representation has reached a "critical mass". To my knowledge, this is the first paper to analyze this dimension of heterogeneity empirically. The idea of looking at local government composition has received recent attention from papers looking at both gender diversity in Brazil (Funk & Philips, 2019) and ethnic diversity in California (Beach & Jones, 2017; Beach et al., 2019; Kogan et al., 2020). I also explicitly observe preference data (in the form of surveys) from the specific setting that I am estimating effects in, whereas most other studies tend to rely on traditional gender roles and national-level data to make assumptions about differences in preferences for their setting.

More broadly, these results are related to other literature that studies whether women make different decisions to men: Do women make more efficient and/or more equitable choices? In the development literature, women tend to spend cash transfers differently from men (Bobonis, 2009; Armand et al., 2016). For example, Duflo & Udry (2004) found that women will direct more of a cash transfer into expenditure on children than men. Additionally, Doepke & Tertilt (2017) show that whether targeting transfers to women rather than men leads to higher growth depends on which factors of production are most important: Directing household transfers to women will be preferable if human capital is more important, compared to physical capital or labor. There is also direct evidence on differences in political belief between men and women: In the U.S., Lott & Kenny found that an increase in female voting through suffrage led to increases in public goods expenditures (1999), and Edlund & Pande (2001) document that women have become more "left-wing" in their political beliefs, which is broadly associated with a suite of preferences on government spending. In the literature on diversity in companies at the board level, where performance is perhaps easier to measure than in other settings, there is little causal evidence that increasing gender diversity improves firm performance directly. However, there is suggestive evidence that moving more women into positions of power can reduce gender bias and allow women's contributions to be recognized and more optimally utilized (e.g. Alexander, 2012; Boutchkova et al., 2017; Bertrand et al., 2017; Smith, 2018).

The rest of the paper is structured as follows. In Section 3.2 I describe the structure of California city councils and elections. In Section 3.3, I describe the data used in the analysis and discuss my empirical strategy. Section 3.4 contains the main results and robustness checks. Section 3.5 probes some potential underlying factors behind the results in Section 3.4. Section 3.6 concludes with a discussion of findings.

3.2 City Councils in California

The structure of city councils and elections in California is relatively uniform. This is because the state has guidelines concerning how local governments are set up, although a small minority of municipalities have passed measures that deviate from the state's template. The majority of councils consist of five councilors (88% of municipalities) that serve four year terms. Staggered elections are the norm, with 2-3 seats on the council coming up for election every 2 years (Beach & Jones, 2017). These elections are non-partisan, meaning that voters do not observe candidate party affiliations. A possible consequence of this is that voters will pay more attention to the characteristics of the candidate as a replacement for using party affiliation as a heuristic. In the large majority of cities (92%), councilors are elected 'at-large', meaning they are elected by the entire electorate, rather than representing a particular district/ward within the city. This means that we can assume that seats on the council are identical and not worry about the district a candidate represents being correlated with your influence on the council. An interesting feature of councils is that although there is technically a city mayor, they are not elected in a conventional mayoral election. Rather, this "mayor" is usually simply chosen from one of the elected councilors (93% of cities), and the title provides no additional powers or status (98% of cities). The council as a whole decide policy, which is then implemented by an appointed city manager. This is notable when relating my analysis to that of Ferreira & Gyourko (2014), who look nationwide at the impact of the election of a female mayor: In Californian cities, the mayor is simply one of the councilors, and not elected in a separate election, so looking at the effect of the mayor's gender alone as in Ferreira & Gyourko (2014) would not capture variation in spending attributable to the gender composition of the rest of the council. It is technically possible for municipalities to change the aforementioned structure of government, but as Beach & Jones note, this does not happen very often, and even when it does, in a large number of cases the proposed measures have not been passed (2017).

The importance and involvement of the city council in expenditure decisions varies by sector: First of all, there are separate elections for school board members, who make decisions on education spending, so city councils play little role in this arena. However, councils play a role in many other important areas of public good provision. Some of these are typically provided by the city itself (e.g. parks and recreation), whereas others are often contracted out to private operators or provided jointly by a set of neighboring municipalities (e.g. transportation). Even if the city is not directly providing these services, the council can still affect the expenditure on these items, since they are the ones that contract with these outside parties. A final point is that in 1978, California passed Proposition 13, which placed a limit on the growth of property taxes. This means that it is difficult for a council to drastically increase spending in one area without cuts to another area, which means that a priori, we should expect any effect of female representation to have a smaller impact on *total* municipal spending than in places that do not have such a restriction.

3.3 Data & Empirical Strategy

3.3.1 Data

In order to estimate the effect of gender composition of city councils on various measures of spending, I use data from two main sources: 1) Municipality-level expenditures and 2) information on candidates that ran for election in California municipalities and the outcomes of those elections

Expenditure data is sourced from the *California State Controller's Office*. Currently, data from 2003-2016 is available. This data has both aggregate spending measures and spending broken down into various categories and by types such as operational spending and total spending by category. I use the total spending measures, as well as those for general, police, public safety, transportation, health, parks & recreation, and culture & leisure spending. The types of public goods that municipalities provide varies, and so not all city councils will be responsible for all of the aforementioned categories. Some categories, such as transportation, are usually provided at a higher geographic level than the typical municipality. However, this does not mean that the city council does not have any influence over spending in these categories: City councils still have the ability to choose who they contract out services to and the scope/terms of these arrangements. The expenditure files also contain population estimates by year which I use to construct per-capita measures of spending figures. Expenditures are adjusted to January 2020 dollars.

Data on election candidates and results comes from the *California Election Data Archive* (CEDA). These files contain the universe of candidates that ran in local elections from 1995-present. The files contain the first and last names of the candidates, as well as their stated occupation and the voting results for each candidate. Unfortunately, the crucial piece of information that is not present is gender. To get around this, I predict gender using first names by using the package **gender()** in R (Blevins & Mullins, 2015). This package takes a first name as an input, and then predicts the likelihood that the person is male or female by looking at the frequency at which is occurs in a chosen government data set.[2] I use Social Security Administration data on names that can be called by using a companion package **genderdata()**. One of the potential problems with a method like this highlights the strength of this package: Some names may have different gender associations at different points in time. Blevins & Mullins (2015) highlight an example of this: In 1900, the name 'Leslie' was almost exclusively a male name (90%), but starting in the 1920s, it gradually gained popularity as a female name, to the point where in 2000, someone born with the name has a 96% chance of being female. Since city council candidates can be any age above 18, whether a Leslie observed on the ballot is female or male depends on the year of birth. The **gender()** package allows me to do robustness checks by specifying the years of the SSA data to be referenced when predicting gender. As an additional test of prediction quality, I use this method to predict genders of candidates from Beach & Jones (2017), who used Amazon's *Mechanical Turk* service to identify demographic characteristics of candidates, including gender. Because some observations have unusual first names or nicknames, I manually code around 2% of candidate genders by researching them via a web search. Some candidates have information missing from the name variables

[2] I define a candidate as female if the reported probability of being female if greater than 50%, since almost no names have a probability close to 50%. I use the SSA data from 1950-1990 to construct the predictions.

and so cannot be identified, but among those that I can identify, my gender prediction matches the gender indicator used in Beach & Jones (2017). Another issue that comes up when analyzing heterogeneous effects of a female win by prior gender is that I do not have a direct measure for the council size, which is needed to determine which gender has a majority in a council, so I get a proxy for it by summing the number of 'winning' candidates in the year of the election and the three years prior since terms are four years long.

3.3.2 Empirical Strategy

Suppose we are in a world where there is only one seat up for election for each council in each time period. A naive comparison of the the effect of gender diversity on the composition of local public spending would be to regress expenditure in different sectors on whether a female or a male won the election for city council c. Then we could write the regression as

$$E_{ct} = \beta_0 + \beta_1 d\_female\_win_{ct} + \varepsilon_{ct} \tag{3.1}$$

Where E_{ct} is the expenditure outcome for city c in year t, $d\_female\_win_{ct}$ is a dummy for whether a female candidate won the election, and ε_{ct} is the error term that contains all other variation in expenditure. Due to the possibility that the error term contains variables that also influence whether a female candidate wins the election, the Ordinary Least Squares (OLS) estimate may be biased. My strategy to account for this unobserved heterogeneity is to use an RDD, pioneered by Thistlethwaithe and Campbell (1960), and popularized in the election setting in a series of papers by Lee (2001, 2008) and Lee et al. (2004). Although the winner of an election is endogenous, it is a deterministic function of the vote margin between the winner(s) and loser(s). In this setting, the RDD compares "sufficiently close" elections between a man and a woman, and argues that for such close elections, the winner is as good as randomly selected. As long as there are no agents (candidates, councils, parties, etc.) that can precisely control the outcome of elections, then the RDD will produce unbiased estimators of β_1. A basic specification of this type can be written as:

$$E_{ct} = \beta_0 + \beta_1 + d\_female\_win_{ct} + \beta_2 MV_{ct} + \beta_3 (d\_female\_win_{ct} * MV_{ct}) + \varepsilon_{ct} \tag{3.2}$$

$$d\_female\_win_{ct} = \mathbb{1}[MV_{ct} \geq 0] \tag{3.3}$$

Where MV_{ct} is the margin of victory for the female candidate and the treatment occurs if this running variable is positive.

However, as discussed in the previous section, California city councils almost always conduct 'at-large' elections, which means that there are multiple seats up for grabs in a given election among a single pool of candidates. Therefore, I focus on the candidate that was elected with the lowest fraction of the votes ('last winner'), and compare them to the candidate that was not elected with the highest fraction of the votes ('first loser'), and more specifically, those cases where these two candidates are of the opposite sex. Another wrinkle is that the marginal effect of the female candidate winning vs. the male candidate could theoretically contain not only a direct element, which is just the effect on spending that I am interested in, but also a dynamic element where a female win changes the likelihood that other female candidates in future elections are successful. My main specifications disregard this dynamic element by using the "differences in discontinuity" approach taken previously by Grembi, Nannicini, & Troiano (2016) and Beach & Jones (2017): Within my sample of elections, I consider only the first year where I can observe a male vs. female election for each city, and truncate all observations that occur after a second male vs. female close election in that municipality. Note that the first election that I observe is not the first election *ever* between a male and female candidate, rather just the first one that I can observe in my data. This leads to a base specification of the following form:

$$E_{ct} = \beta_0 + \beta_1 d\_female\_win_{ct} + \beta_2 MV_c + \beta_3 (d\_female\_win_{ct} * MV_c) + \gamma_t + \varepsilon_{ct} \tag{3.4}$$

Where the treatment $d\_female\_win_{ct}$ is now defined by:

$$d\_female\_win_{ct} = \mathbb{1}[MV_c > 0] * POST_t \tag{3.5}$$

Where $POST_t$ is an dummy variable that indicates if an observation occurs after the year that the first female candidate is elected (as observed in the sample period). Note that the margin of victory term is now time-invariant since I am only focusing on the *first* instance of potential treatment, even though the data is still a panel. γ_t denotes time fixed-effects. This model is only preferred if there are in fact no dynamic effects to worry about, otherwise a dynamic RD strategy akin to that of Cellini, Ferreira, & Rothstein (2010) is preferred. I check this assumption in the next section by running first-stage variants of Equation (3.4).[3]

Although I provide multiple pieces of evidence in support of the internal validity of the empirical strategy in the next section, the external validity of my analysis is harder to pin down. This is because the election data themselves do not contain demographic information regarding the candidates. Therefore, if gender of a candidate is correlated with other characteristics, then the effect being identified by β_1 in Equation (3.4)

[3] Although I do not find any evidence of dynamic effects in the next section, which could justify using all male vs. female elections in a municipality as separate observations in a 'stacked' approach, I stick with the difference in discontinuity approach as my preferred specification to be certain there are no dynamics, and because the precision on the transportation result is slightly better. I show the stacked results as as robustness test later on.

Table 3.1: Summary Stats and Differences in Means

	A: Full Sample		B: "Relevant" Elections		C: Potential Female Majority		Difference	
Variable	Mean (1)	S.D. (2)	Mean (3)	S.D. (4)	Mean (5)	S.D. (6)	A vs. B (7)	B vs. C (8)
Year	2007	3.7	2007	3.0	2007	3.1	0.56	-0.39
							(0.28)	(0.42)
Population	44,044	118,000	43,915	55,340	44,989	54,902	129	-1,074
							(8,763)	(7,583)
Per-Capita Spend	1,864	1,766	1,811	1,685	1,780	1,414	52.96	31.13
							(133.48)	(221.62)
Share: General	.121	.079	.129	.093	.127	.087	-0.008	0.002
							(0.006)	(0.013)
Share: Police	.195	.062	.199	.064	.204	.075	-0.004	-0.006
							(0.005)	(0.009)
Share: Public Safety	.269	.089	.270	.094	.274	.100	-0.0002	-0.004
							(0.007)	(0.013)
Share: Transport	.153	.098	.139	.086	.147	.105	0.014	-0.007
							(0.007)	(0.013)
Share: Health	.108	.125	.107	.127	.093	.125	0.0007	0.014
							(0.010)	(0.017)
Share: Parks & Rec	.064	.045	.067	.049	.067	.049	-0.003	-0.0001
							(0.003)	(0.007)
Share: Culture & Leisure	.090	.068	.089	.064	.088	.061	0.0006	0.001
							(0.005)	(0.009)
N		2,506		187		74		

Sample A contains all city-year elections where the marginal candidates are a male and a female. Sample B consists of only the first observed male vs. female election in each municipality where spending data is available. Sample C contains observations where either i) the council went from majority male to majority female because of that election, or ii) such a swing would have occurred if the female candidate had won the election, but did not. Columns (7) and (8) show the results of comparison of means t-tests between columns (1) & (3) and (3) & (5) respectively. Standard errors are shown in brackets.

is the joint effect of that bundle of characteristics that are correlated with female candidates as well as the effect of female representation *ceteris paribus*. Therefore, one should be cautious extrapolating results in this paper to settings where the relationship between gender and other demographic characteristics is significantly different from California.[4]

Table 3.1 provides descriptive statistics for year, estimated population, and the outcomes of interest, which include per-capita spending and the share of total spending allocated to one of seven categories. I use shares rather than dollar or per-capita amounts for the spending categories. This is because a change in the dollar amount of spending on a particular category could come from reallocating resources from other categories, or from an increase in total spending that affects all categories proportionally. If the variable is measured in dollars, then to isolate the former effect, I would have to control for total spending. However, this would be a 'bad control' since total spending is a main outcome of interest (Cinelli, Forney, & Pearl, 2022).

Statistics are shown for three distinct samples. The full sample is the set of city-year pairs where a male and female candidate are the last winner and first loser (or vice versa). The set of relevant elections refer to the first observed male vs. female elections in a municipality with non-missing expenditure data. Finally, the 'potential female majority' sub-sample contains elections where either the council flipped from male majority

[4]Beach & Jones (2017) collect data on both ethnicity and gender, so I use the data on marginal candidates from their replication files to run simple regressions of ethnicity on gender. There appears to be no significant correlation between the two, suggesting that any results I do find are not simply an artifact of changes in ethnicity expressing themselves through gender.

to female majority, or were a single additional female councilor away from doing so but did not. Across all samples, municipalities in California spend the largest share of their budget on law enforcement (police and public safety), and the least on local amenities (parks & recreation and culture & leisure). Columns (7) and (8) show the results of tests for equality of means between different samples. There turns out to be little difference between mean values in these samples. The main difference is that the relevant election sample tends to contain earlier years, since by definition it contains the earliest valid male vs. female election.

3.4 Results

3.4.1 Validity of Identifying Assumption

In order to assert that the occurrence of a female victory in a male vs. female election is an exogenous event, and thus claim that the parameter β_1 from Equation (3.4) is causal, it needs to be the case that precise manipulation of the running variable, in this case MV_c, is not occurring. I provide evidence for this claim by showing that the density of MV_c does not 'jump' at the cutoff of 0. I follow the method of Cattaneo, Jansson, & Ma (2018) by plotting two separate density functions of MV_c on each side of the cutoff. This plot is shown in Figure 3.1. Graphically the density looks similar on each side of the cutoff: The 95% confidence intervals are shown by the shaded grey areas, and there is significant overlap of these confidence intervals as we approach $MV_c = 0$ from both sides. This is confirmed numerically, as the binomial null hypothesis tests that the probability of a female win are not statistically different from 0.5 are not rejected (p-values range: $0.3374 - 0.8644$).

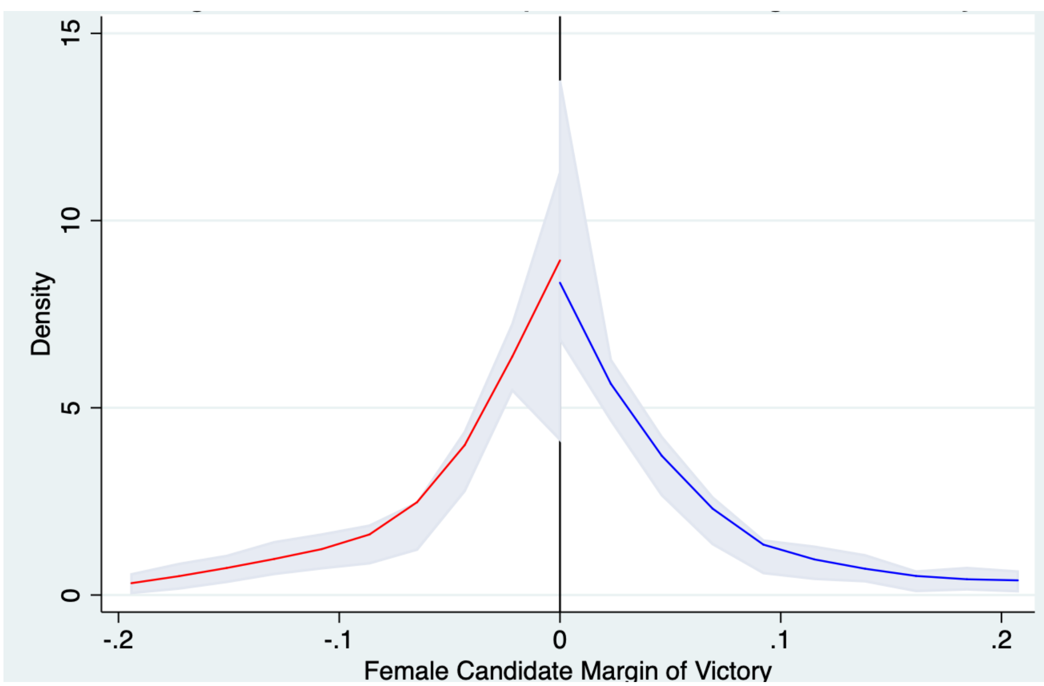


Figure 3.1: Test for Manipulation of Margin of Victory

As an additional test of the identifying assumption, it should be the case that other variables do not discontinuously jump at the cutoff. Since my data is limited in terms of the number of true control variables at the city-year level, I use baseline levels of my outcome variables as additional controls, since they should not discontinuously jump prior to treatment. For an election in year t, I use expenditure variables for fiscal year t as controls, since fiscal year t typically begins on July 1 of year $t-1$ and ends on June 30 of year t, so certainly should not be affected by councilors newly elected in year t. Table 3.8 estimates Equation 3.4 using baseline levels of outcomes, and I do not find any significant effects as expected.

3.4.2 First-Stage Dynamic Effects

As discussed in Section 3.2, the difference in discontinuity specification is only preferred if there are no dynamic effects of a close female victory on the probability of other female candidates winning in the future. I test whether this is the case by replacing the left-hand side of my main equation (3.4) with the share of elected candidates that are female in *future* elections. I run specifications looking at the 1st through the 5th subsequent election after the election of interest that represents the potential treatment in the difference-in-discontinuities approach. Table 3.2 shows the results of this test. The only election where the outcome of the election of interest has a significant effect is in the second subsequent election; if the female candidate wins against a male candidate, then the fraction of future winners that are female increases by 15%-22%. The explanation for finding a significant effect in the second subsequent election and not in others is that the majority of municipalities hold elections every two years, and so the second subsequent election will fall when the female winner from our election of interest is up for re-election. This hypothesis is confirmed by Tables 3.3 and 3.4. Table 3.3 replaces the dependent variable with an indicator for whether the 'last winner' in the nth subsequent election is an *incumbent* female. We see that the parameter estimate is still significant in most specifications for the 2nd subsequent election; a female victory increases the probability of a female incumbent victory by 10-15%. However, in Table 3.4, the dependent variable is instead an indicator for a *non-incumbent* female victory. We see that the estimates becomes insignificant. Thus, we can conclude that the first-stage effects observed in Table 3.2 are driven by incumbency effects, not by a dynamic effect on election probabilities of other female candidates. Therefore, implementing a dynamic RD is not necessary.

A tangential question of interest is whether men and women experience different levels of incumbency advantage. It may be the case that there is an implicit bias against women in elections, which would result in there being a higher 'competency threshold' for female candidates to meet before being considered viable candidates by the electorate. If true, we might expect women to have a higher incumbency advantage since the average quality of *elected* women is higher. On the flip side, it might instead be the case that female council members are more harshly judged for their performance, in which case we would expect the opposite

Table 3.2: First-Stage Dynamic Effects by Subsequent Election Order

Table 2: First Stage Dynamic Effects by Subsequent Election Order

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Female Share of Winners in n^{th} election after MvF election of interest						
1st Election	-0.041	-0.041	-0.027	-0.145	-0.105	-0.124**
	(0.036)	(0.036)	(0.037)	(0.089)	(0.076)	(0.058)
2nd Election	0.147***	0.147***	0.144***	0.130	0.225***	0.165***
	(0.037)	(0.037)	(0.038)	(0.081)	(0.066)	(0.053)
3rd Election	0.015	0.015	0.027	-0.069	0.035	0.022
	(0.039)	(0.038)	(0.039)	(0.090)	(0.076)	(0.060)
4th Election	-0.058	-0.058	-0.064	-0.131	-0.069	-0.074
	(0.041)	(0.040)	(0.041)	(0.089)	(0.075)	(0.060)
5th Election	-0.001	-0.001	-0.011	0.035	-0.026	0.033
	(0.054)	(0.056)	(0.057)	(0.129)	(0.111)	(0.088)
Cluster SE		X	X	X	X	X
Year FE			X	X	X	X
Bandwidth				3%	5%	10%

Standard errors are in parenthesis. The majority of councils elect ½ of the council every two years, but there are some councils with alternative schedules so models for the n^{th} election after the election of interest are used instead of actual years.

**** p<0.01, ** p<0.05, * p<0.1*

Table 3.3: First-Stage Incumbency Effect (Females)

Table 3: First Stage Incumbency Effect (Females)

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Indicator for an incumbent female victory						
1st Election	0.066*	0.066	0.073*	-0.040	0.011	0.076
	(0.039)	(0.041)	(0.041)	(0.112)	(0.091)	(0.068)
2nd Election	0.100**	0.100**	0.108**	0.109	0.159**	0.128**
	(0.044)	(0.043)	(0.043)	(0.091)	(0.074)	(0.063)
3rd Election	0.069	0.069	0.079	0.133	0.133	0.077
	(0.048)	(0.048)	(0.048)	(0.121)	(0.095)	(0.078)
4th Election	-0.015	-0.015	-0.002	-0.087	-0.021	-0.041
	(0.050)	(0.049)	(0.050)	(0.102)	(0.087)	(0.069)
5th Election	-0.017	-0.017	-0.044	-0.194	-0.105	-0.109
	(0.058)	(0.062)	(0.062)	(0.169)	(0.134)	(0.101)
Cluster SE		X	X	X	X	X
Year FE			X	X	X	X
Bandwidth				3%	5%	10%

Standard errors are in parenthesis. The majority of councils elect ½ of the council every two years, but there are some councils with alternative schedules so models for the n^{th} election after the election of interest are used instead of actual years. Standard errors are clustered at the city level in Columns (2)-(6).

**** p<0.01, ** p<0.05, * p<0.1*

Table 3.4: First-Stage Dynamic Effects on Non-Incumbent Female Victory

Table 4: First Stage Dynamic Effects on Non-incumbent Female Victory

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Indicator for a non-incumbent female victory						
1st Election	-0.058	-0.058	-0.047	-0.103	-0.051	-0.125*
	(0.046)	(0.046)	(0.047)	(0.098)	(0.090)	(0.069)
2nd Election	0.025	0.025	0.010	-0.119	-0.032	-0.068
	(0.045)	(0.046)	(0.047)	(0.103)	(0.085)	(0.069)
3rd Election	0.013	0.013	0.038	0.071	0.021	0.052
	(0.048)	(0.051)	(0.052)	(0.120)	(0.099)	(0.072)
4th Election	0.046	0.046	0.035	0.076	-0.078	-0.047
	(0.056)	(0.056)	(0.059)	(0.130)	(0.108)	(0.086)
5th Election	0.015	0.015	0.047	0.199	0.111	0.165
	(0.071)	(0.074)	(0.072)	(0.170)	(0.142)	(0.110)
Cluster SE		X	X	X	X	X
Year FE			X	X	X	X
Bandwidth				3%	5%	10%

Standard errors are in parenthesis. The majority of councils elect ½ of the council every two years, but there are some councils with alternative schedules so models for the nth election after the election of interest are used instead of actual years. Standard errors are clustered at the city level in Columns (2)-(6).
****** $p<0.01$, ** $p<0.05$, * $p<0.1$*

to be true. Table 3.9 shows simple regressions of the effect of a having a male or a female incumbent competing in the marginal election of interest on the probability of a male or female victory in that election. The coefficients in columns (3) and (4), corresponding to the effect of the male candidate being the incumbent of the likelihood of a male victory are slightly larger than their female counterparts in columns (1) and (2), but they are not significantly different.

3.4.3 Regression Discontinuity Results

Figure 3.2 shows the results of estimating Equation (3.4), specifically parameter estimates of β_1. Bandwidth selection for these main results are chosen using the method of Calonico, Cattaneo, & Titiunik (2014).[5] The top-left panel plots these estimates for the log of per-captia spending. The other panels plot the estimates using the share of spending that was allocated to each category. In each figure, I run specifications with the dependent variable measured with a 2, 3, 4, & 5 year lead, as well as the aggregate of those four years. For the log spending panel, the aggregate is the sum of years 2-5, whereas in the other panels, it represents the average share across the four year period. I use the 2 year lead as the earliest period because the outcomes are fixed to a fiscal year, whereas the election results are fixed to calendar years. The fiscal year in California typically runs July through June (occasionally giving or taking a month), whereas elections would usually be held in November. That means that by the time a new councilor is elected in year t, the fiscal year $t+1$ has already begun. Terms are four years long, so I use a 5 year lead as the last period. I also include two baseline periods, 0 and 1 as falsification tests. Few of the sub-figures show significant results. The one that stands

[5] Bandwidths chosen by this method are included in Tables 3.10-3.14.

out the most is transportation, where the share of total spending 2 and 5 years later, as well as aggregated over the four year period, is around 2 percentage points higher if the female candidate wins in the election of interest. The other two categories that are worth noting are Parks & Recreation and Culture & Leisure. Parks & Recreation spending is negative and significant four and five years out, although not statistically-significant over the councilor's full term. Culture & Leisure is the other way around: None of the individual years themselves are statistically significant (although the last three years are barely insignificant at the 95% level), but the aggregate is negative and significant; a female victory decreases the share of total spending dedicated to Culture & Leisure by 1.6%.

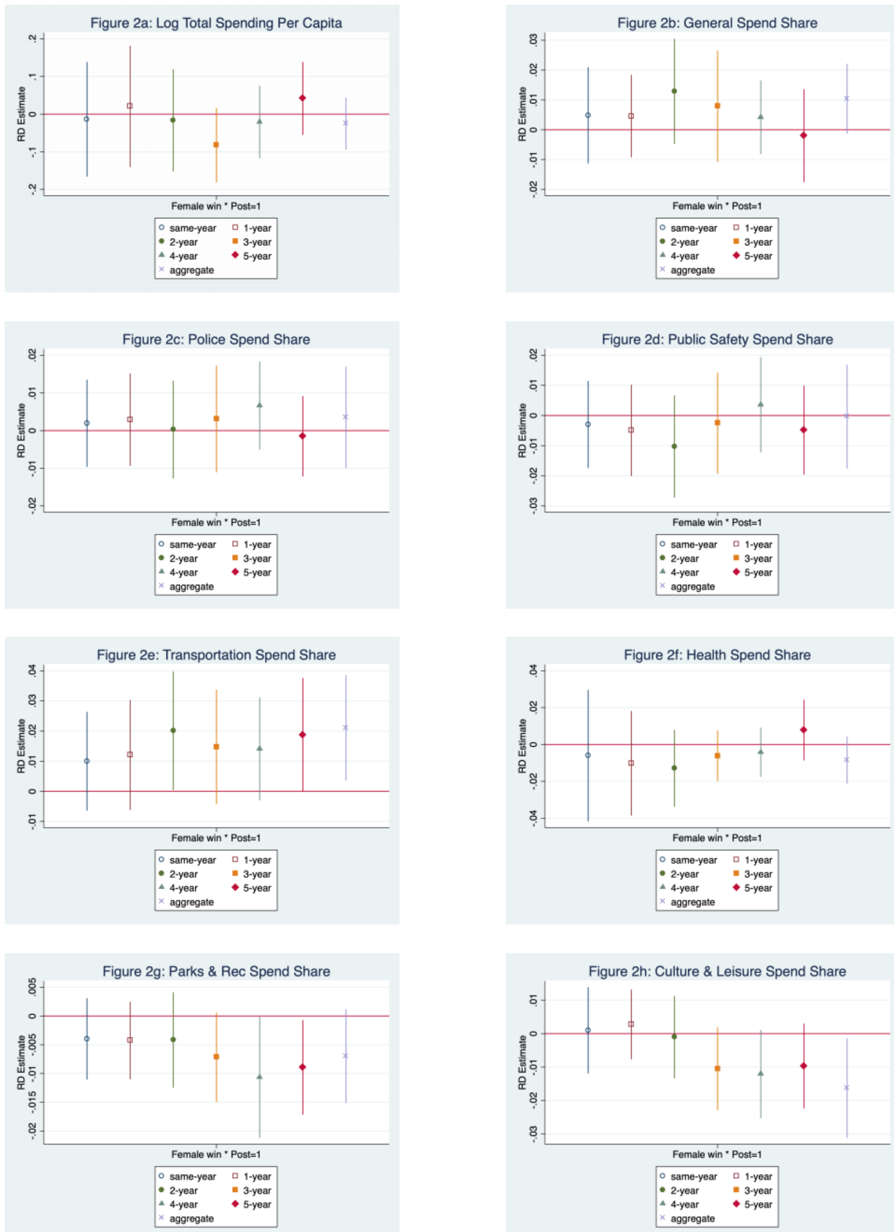


Figure 3.2: RD Effects of a Marginal Female Victory

The parameter estimate from average transportation shares indicate that the share of spending goes up 2pp. during the four-year term of a marginally-elected female councilor. To frame estimates for category shares more intuitively, I also convert the outcome variable to units of standard deviations and in $/capita terms. Table 3.5 shows the parameter estimates for share outcomes in Equation (3.4) but with the shares scaled to be in terms of standard deviations. For the average 2-5 column transportation outcome, the marginal female victory is associated with a 0.23 standard deviation increase in transportation share. Table 3.6 replicates Table 3.5 but uses the natural log of spending/capita, similar to overall spending in Figure 3.2. For the average 2-5

Table 3.5: Effect of a Marginal Female Victory on Expenditure Composition (Standard Deviation Measure)

Table 5: Effect of a marginal female victory (vs. a male) on expenditure composition (standard deviation share measure)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Pre-period		Post-period				
Years Since Election:	0	1	2	3	4	5	Average 2-5
General	0.062	0.059	0.166	0.102	0.054	-0.026	0.135*
	(0.107)	(0.091)	(0.116)	(0.123)	(0.082)	(0.103)	(0.077)
Police	0.033	0.049	0.005	0.053	0.113	-0.025	0.060
	(0.100)	(0.106)	(0.112)	(0.123)	(0.101)	(0.092)	(0.116)
Public Safety	-0.035	-0.058	-0.120	-0.029	0.041	-0.056	-0.004
	(0.086)	(0.091)	(0.101)	(0.100)	(0.094)	(0.088)	(0.103)
Transportation	0.110	0.133	0.221**	0.162	0.155	0.206*	0.232**
	(0.092)	(0.102)	(0.110)	(0.107)	(0.096)	(0.106)	(0.098)
Health	-0.050	-0.085	-0.107	-0.051	-0.035	0.065	-0.070
	(0.151)	(0.120)	(0.088)	(0.059)	(0.056)	(0.070)	(0.054)
Parks & Rec	-0.091	-0.098	-0.096	-0.165*	-0.245**	-0.205**	-0.160*
	(0.083)	(0.079)	(0.097)	(0.091)	(0.123)	(0.096)	(0.096)
Culture & Leisure	0.015	0.044	-0.017	-0.167*	-0.192*	-0.154	-0.259**
	(0.105)	(0.085)	(0.100)	(0.101)	(0.107)	(0.103)	(0.121)

Standard errors are in parenthesis. Each row/column combination is an estimate of β_1 from Equation (4). Each column contains estimates from a given number of years since the election of interest, and column (7) is an average of all post-periods (2-5 years after election). All specifications use the default (CCT1) bandwidth (see Tables A3-A7) with standard errors clustered at the council level. The outcome is the share of total spending allocated to a given category in the given year, standardized such that a parameter estimates of X is interpreted as a X standard deviation increase in the share of spending allocated to a given category when the female candidate marginally wins the election.
**** p<0.01, ** p<0.05, * p<0.1*

transportation specification, the marginal female victory corresponds to a 17pp. increase in spending/capita, off a median value of around \$790/capita.

As mentioned in Section 3.1, one of the assumptions of the Citizen Candidate model specific to this setting is that the marginal influence of a female councilor is independent of the prior composition of the council. In reality, this may not be the case. One possible reality is that female preferences in the council carry less weight because of implicit or explicit gender biases. It may also be the case that majority rule is a factor, and in most cases it is women who are a minority on the council. For these reasons, there might be heterogeneity in the marginal effect of a female councilor (and corresponding marginal decrease in male councilors) based on the genders of the other councilors (Beckwith & Cowell-Meyers, 2007). I focus on two points along the male-female power balance where there might be different effects: The first is when a council gains its first female councilor. In this case, the incoming councilor may bring a new voice that was absent prior to the election. The second is when the council flips from being majority male to majority female. In this case, even if all male councilors are hostile to all potential spending decisions that are relatively preferred by the female councilors, it is possible that the female majority could ensure their preferences are reflected in actual expenditure, although this would vary issue to issue depending on the exact preference orderings for different spending categories.[6] To do this, I construct indicator variables for a) if there were no female councilors prior to the focal election but at least one is elected in the election, and b) if prior to the election of interest the

[6] In an ideal data setting, I would be able to look at a finer degree of heterogeneity by looking at the heterogeneous effect of the 1st, 2nd, 3rd, etc. female councilor. Unfortunately, I do not have the sample sizes to do this.

Table 3.6: Effect of a Marginal Female Victory on Expenditure Composition (Per-Capita Measure)

Table 6: Effect of a marginal female victory (vs. a male) on expenditure composition (log $/capita measure)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Pre-period		Post-period				
Years Since Election:	0	1	2	3	4	5	Average 2-5
General	-0.006	0.042	0.126	0.087	0.108*	0.117*	0.132**
	(0.089)	(0.083)	(0.080)	(0.075)	(0.062)	(0.063)	(0.067)
Police	0.010	0.046	0.053	0.051	0.044	0.029	0.038
	(0.046)	(0.054)	(0.052)	(0.052)	(0.047)	(0.040)	(0.059)
Public Safety	-0.032	-0.003	0.020	0.045	0.041	0.033	0.033
	(0.040)	(0.047)	(0.050)	(0.052)	(0.046)	(0.040)	(0.059)
Transportation	0.084	0.092	0.153**	0.139*	0.081	0.155*	0.169**
	(0.074)	(0.074)	(0.074)	(0.078)	(0.066)	(0.080)	(0.068)
Health	0.093	0.032	0.026	-0.046	0.025	0.119	0.093
	(0.205)	(0.176)	(0.148)	(0.107)	(0.079)	(0.092)	(0.136)
Parks & Rec	-0.125*	-0.081	-0.033	-0.092	-0.207**	-0.186*	-0.012
	(0.068)	(0.071)	(0.080)	(0.073)	(0.103)	(0.105)	(0.103)
Culture & Leisure	-0.068	-0.017	0.036	-0.045	-0.084	-0.128	-0.114
	(0.083)	(0.092)	(0.101)	(0.081)	(0.099)	(0.103)	(0.089)

Standard errors are in parenthesis. Each row/column combination is an estimate of β_1 from Equation (4). Each column contains estimates from a given number of years since the election of interest, and column (7) is an average of all post-periods (2-5 years after election). All specifications use the default (CCT1) bandwidth (see Tables A3-A7) with standard errors clustered at the council level. The outcome is the natural log of spending per-capita, thus parameter estimates are interpreted as the percentage point change in per-capita spending as a result of a marginal female victory.
****p<0.01, **p<0.05, *p<0.1*

council was a male majority but after the focal election it is a female majority. I interact these indicators with the original treatment $d\_female\_win_{ct}$ to check for heterogeneous treatment effects. The new specification is shown in Equation (3.6). β_6 and β_7 are the parameters of interest for these effects.

$$\begin{aligned} E_{ct} = \beta_0 + \beta_1 d\_female\_win_{ct} + \beta_2 MV_c + \beta_3(d\_female\_win_{ct} * MV_c) \\ + \beta_4 first\_female_{ct} + \beta_5 majority\_female_{ct} \\ + \beta_6(d\_female\_win_{ct} * first\_female_{ct}) + \beta_7(d\_female\_win_{ct} * majority\_female_{ct}) + \gamma_t + \varepsilon_{ct} \end{aligned} \tag{3.6}$$

The results of these regressions are shown in Figure 3.3. One important point is that a council flipping from a male majority to a female majority is not a common event. This means that the estimates for β_7 are more imprecise that for β_1 and even β_6. Additionally, a larger fraction of missing data for the per-capita spending outcome means I have to omit $majority\_female_{ct}$ from that specification. Looking specifically at those outcomes for which there were significant effects in the main specification, we can see that our main effect is now a lot smaller are hovering just around or under the significance at the 5% level. However, the increase in transportation spending seems to be driven by cases where the election of a new female councilor occurs and creates a female majority. For Parks & Recreation and Culture & Leisure spending, the main effects are still just on the cusp of 5% significance. The effects of a female majority (β_7) show positive, significant effects in a couple of specifications, but are quite imprecise, so it is harder to be confident about inferring anything from these estimates.

Figure 3.3: RD Heterogeneous Effects by Prior Council Composition

I also conduct robustness tests on the specifications in Figures 3.2 and 3.3. The results of these alternative specifications are shown in Tables 3.10-3.14 (corresponding to Figure 3.2) and 3.15-3.19 (corresponding to Figure 3.3). Specifically, I use a slightly different bandwidth selection: The default method of Calonico, Cattaneo, & Titiunik (2014) is to use a regularization term to avoid choosing a too-large bandwidth. This method can produce very conservative bandwidths in some cases. I rerun all my original specifications without using this regularization term. The selected bandwidths with no regularization term is shown in column

(2) of Tables 3.10-3.14, next to the original bandwidths in column (3) of the same tables. Column (4) shows the parameter estimates for β_1 using no regularization term and the results from the default bandwidth with the regularization term are in column (5). I also test using a polynomial of degree 2 in column (6). Overall, the transportation results are fairly robust and the Parks & Recreation and Culture & Leisure are somewhat less so. Finally, column (1) in Tables 3.10-3.14 show the mean share of total spending in that category across all city-year pairs. For example: In Table 3.10, column 1 for transport shows on average a 15% share of spending. The point estimate in column (5) (the default specification from Figure 3.2 is 2%, implying a 13% increase in total transportation spending, which is a large magnitude.

I try re-estimating treatment effects with different models. First, for the main effect estimates from Figure 3.2, instead of a difference in discontinuity approach, I implement a cross-sectional stacked RD where I look at every male vs. female election, regardless of whether it is the first in a municipality or not, and so do not truncate observations. The appeal of this specification is that the number of observations increases significantly. The downside is that it is difficult to interpret the parameter, since the effect of past or future female victories could muddy the treatment. These estimates are shown in Figure 3.6. Focusing on the previously-robust transportation share, although the aggregate magnitude is similar, the estimate is a little less precise and thus marginally loses significance. I also show the RD Plot for the average 2-5 year transportation result in Figure 3.10.

Finally, I re-estimate the heterogeneous effects for transportation share separately. Specifically, I estimate the effect of electing the first female councilor and electing a majority female council from Equation (3.6) separately in the equations below:

$$E_{ct} = \beta_0 + \beta_2 MV_c + \beta_4 first\_female_{ct} + \beta_8 (first\_female_{ct} * MV_c) + \varepsilon_{ct} \tag{3.7}$$

$$E_{ct} = \beta_0 + \beta_2 MV_c + \beta_5 majority\_female_{ct} + \beta_9 (majority\_female_{ct} * MV_c) + \varepsilon_{ct} \tag{3.8}$$

Figure 3.7 and 3.8 show the density plots corresponding to the samples used in Equations (3.7) and (3.8) respectively, analogous to the one shown in Figure 1 for the full analytic sample. Figure 3.9 shows the results of this analysis. The top-left panel is a replication of the transportation result in Figure 3.2, the main specification effect for transportation. The top-right panel shows estimates of β_4 in Equation (3.7), and the bottom-left panel shows estimates of β_5 in Equation (3.8). Comparing these three figures together to the joint estimation from Figure 3.2, the estimates do not seem to change a lot when estimating separately.

3.5 Mechanisms

In the previous section, I found that transportation spending was affected by a change in the gender composition of city councils, and this result was robust. A couple other categories of spending were suggestive, but not robust. On the other hand, the majority of spending categories, as well as total spending per-capita was unaffected. A null result is consistent with the quasi-experimental literature in developed countries such as Ferreira & Gyourko (2014) in the United States. In this section, I attempt to a) explore the significant effect in the transportation sector further, and b) Discuss a potential explanation for null effects.

3.5.1 Transportation Spending

Finding that transportation spending as a category changes when council gender composition changes is interesting, but not fully informative: Transportation compromises many things, such as road/highway maintenance and landscaping, but also public transit, and even ports and airports for larger municipalities. Some of these sub-categories are ongoing, regular expenditures and some of them might be large, one-off capital investments. To see if one type of spending might be driving this effect. I split transportation spending into operating expenditures and capital outlay, and re-estimate Equation (3.5). The results are shown in Figure 3.4. We can see that the original effect found in Figure 3.2 is being driven entirely by changes in capital outlay, rather than operational expenditure. The parameter estimate for the aggregate 4 year period (2.1pp) is very similar to the original estimate in Figure 3.2 (2.0pp) This suggests that increased female presence on councils is leading to higher investment in long-term projects. Unfortunately, I do not have more dis-aggregated data on exactly what this investment is comprised of.

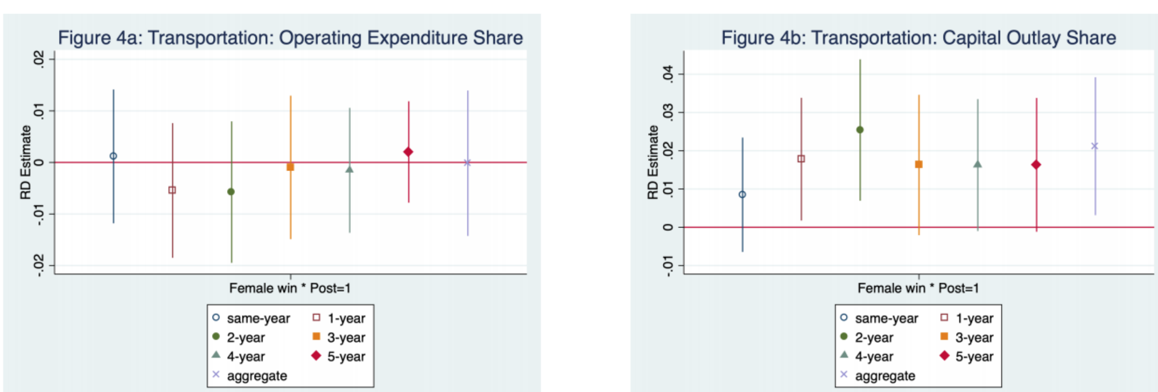


Figure 3.4: RD Effects on Transportation Spending by Type

3.5.2 Null Results

In Section 3.4.3, I discussed one of the implicit assumptions of the Citizen-Candidate model, namely that of power dynamics within a council. The other assumption is that men and women actually have different preferences, otherwise there is no other set of preferences that would change spending decisions. Many other papers investigating this research question rely on either broad national data, research on preferences from other settings, or simply traditional stereotypes of gender (e.g. women care more about social welfare) when determining ex ante hypotheses about which sectors should be affected by increased diversity. I am able to utilize California survey data specifically asking about public expenditures at the local and state level in a similar period as my main data. The Public Policy Institute of California (PPIC) conducts statewide surveys throughout the year, asking all types of questions, including ones about governance and public spending (PPIC, 2020).[7] Panel A of Table 3.7 shows responses to questions about whether respondents agree with increasing expenditure in four different categories: K-12 education, higher education, health and human service, and prisons and corrections. Columns (1) and (2) show the fraction of male and female respondents that said yes respectively. Column (3) shows the difference, and are accompanied by stars that indicate the results of a simple t-test for equality of means. The negative and significant values indicate that women are more likely to want to increase spending in the first three categories than men by several percentage points in each case. Panel B contains some specific questions regarding proposals to increase funding for various transportation projects. Surprisingly, given the results in Section 3.4, there is no statistically significant difference between men's and women's responses for any of the four questions. However, we should note that these questions are binary and in some cases quite specific measures, which may not tease out true differences in preferences. Another question asks respondents to choose one of same four areas from Table 3.7, Panel A, that should be the top priority for increased state funding. The results of this question are shown in Figure 3.5. Again, we see that there are differences by gender, especially for higher education and prisons, which seems intuitive, as K-12 and health spending could be viewed as more 'essential' particularly given the political leanings of the state overall. What this graph and table show, at least in Panel A of Table 3.7, is that there do seem to be different preferences between men and women for allocating public spending in this specific setting, albeit in the limited scope of the survey questions available. A final possibility is that city councilors simply do not have the ability to, or do not increase, decrease, or reallocate municipal budgets. This argument however is quickly rebutted by the fact that other papers such as Beach & Jones (2017) did find statistically significant changes in public spending as a result of changes in the composition of different ethnicities serving on California city councils, as well as the fact that this paper finds impacts of gender com-

[7] Public Policy Institute of California provides this data free of charge. PPIC bears no responsibility for the interpretations presented or conclusions reached based to the analysis of their data.

positions on transportation spending.

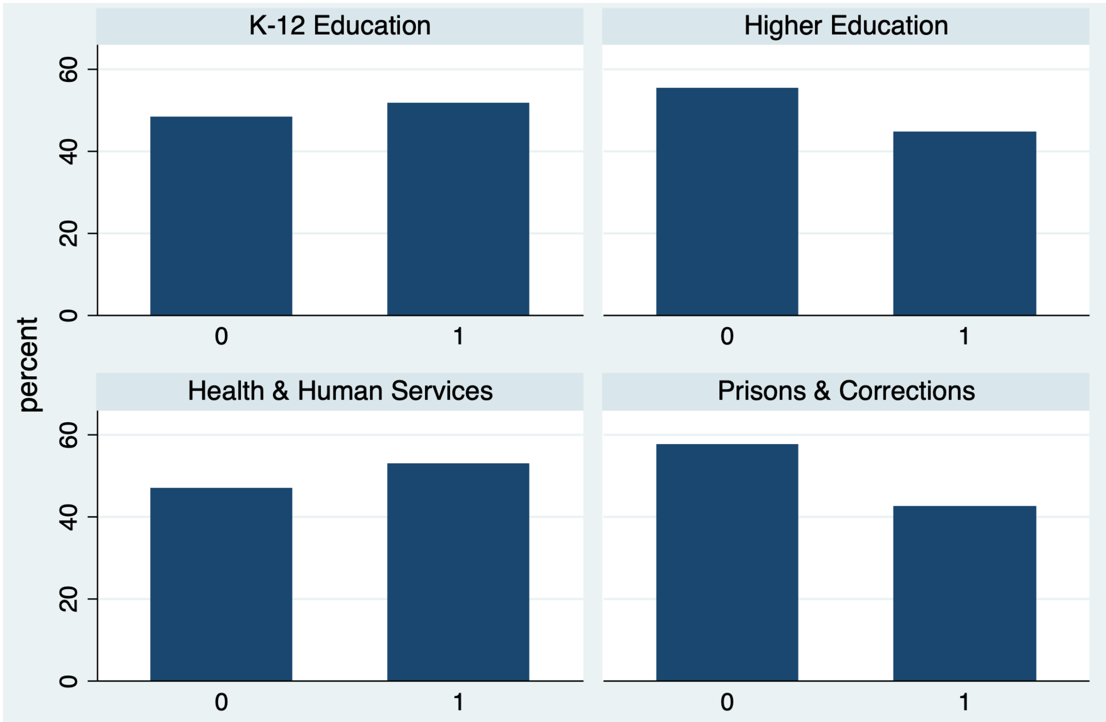


Figure 3.5: Top Priority for State Spending - Males (0) vs. Females (1)

3.6 Conclusion

In this paper, I use plausibly-exogenous variation in the gender composition of local municipal councils in California stemming from close elections in order to analyze the impact on aggregate spending and the composition of spending between different categories. Broadly, I find few significant effects of electing female councilors, including on total spending per-capita. However, when a female candidate wins a marginal, close election against a male candidate, the share of spending directed at transportation increased by around two percentage points. Analyzing the heterogeneity of this effect shows that instances of councils shifting from male majority to female majority seemed to be driving this result. Digging deeper, these increases come from capital outlays on long-term projects rather than increases in regular operating costs. I show some evidence that men and women in California indeed have different perspectives on allocation of public funds to different sectors, but am unable to find statistically significant differences in responses to survey questions regarding the transportation sector specifically.

Caution should be exercised when interpreting the transportation effect since many potential outcomes are analyzed. Additionally, although I find that this results comes from capital outlays, the underlying reason for

Table 3.7: Differences in Spending Preferences by Gender

Table 7: Differences in spending preferences by gender

Question	Male Fraction Yes (1)	Female Fraction Yes (2)	Difference (3)
Panel A: Desire for increased spending			
1. Favor increasing spending on K-12?	0.739	0.809	-0.070***
2. Favor increasing spending on higher ed?	0.577	0.673	-0.096***
3. Favor increasing spending on health?	0.679	0.718	-0.039*
4. Favor increasing spending on prisons?	0.220	0.205	0.015
Panel B: Transportation Funding/Spending			
5. Favor $43B in additional funding for state/local transportation projects?	0.436	0.401	0.035
6. Would you vote yes on a state bond measure to improve roads/transportation?	0.646	0.622	0.023
7. Would you vote yes on a local bond measure to improve roads/transportation?	0.519	0.502	0.018
8. Should we increase state funding for roads and surface transportation?	0.488	0.463	0.025

Notes: Columns (1) and (2) show the percentage of respondents (including those that refused or answered 'don't know') that answered `yes' or `in favor' or agreed with/to the statement presented. Column (3) shows the raw difference between the first two columns. Stars indicate the significance level for a simple t-test for equality of means.
****** p<0.01, ** p<0.05, * p<0.1*

this is still unknown. One possible explanation is that some public transportation systems run across multiple municipalities, and female councilors are more adept at securing and/or funding these joint ventures. Future analysis could look more specifically at whether each each public good is provided internally or jointly with other municipalities, and whether this distinction matters for observing an effect of gender composition.

This paper extends the existing literature in two main ways. First, I provide the first analysis of the effect of gender in municipal councils in the U.S., as opposed to binary positions such as mayors or state representatives. Second, I am the first to examine the heterogeneity of gender diversity by prior council composition, speaking to the question posed by Beckwith & Cowell-Meyers on constant vs. non-constant effects of female representation (2007). Although I did find significant effects in one sector in particular, most of my specifications draw similar conclusions to previous studies done in developed countries that gender composition of councils does not have a major impact on allocation of financial resources, at least in the aggregate or in broad categories (Ferreira & Gyourko, 2014; Bagues & Campa, 2017; Geys & Sorenson, 2019). This is in contrast to the literature in developing countries such as India (Chattopadhyay & Duflo, 2004; Clots-Figueras, 2012) and Brazil (Funk & Phillips, 2019) where significant effects are found more frequently. There are various possible explanations for this trend, but these are beyond the immediate scope of this paper.

Even though the link between gender diversity and expenditure allocation in developed countries may be broadly weak, there are pertinent, non-monetary outcomes that could be affected by electing more women to political office. Hessami & Lopes da Fonseca (2020) summarize a literature that has largely taken place in developing countries, but has tackled topics like the impact of female representation on corruption, which is relevant for efficiency of dollars spent by local government (Baskaran et al., 2018; Brollo & Troiano, 2016). In addition, since women are underrepresented at the top level in many powerful positions, the presence of additional leaders and role-models could have many indirect effects on the future behavior of both men and women in ways that are not observable directly by looking at financial outcomes. It may be that spending is not the most important outcome that we should be looking at when searching for impacts of representation. Future economic research, particularly in developed settings, should try to focus more on these 'softer' outcomes despite the difficulty of obtaining concrete measures of such outcomes.

3.7 References

Alexander, Amy C. (2012). Change in Women's Descriptive Representation and the Belief in Women's Ability to Govern: A Virtuous Cycle. Politics and Gender 8(4), pp.437-464.

Armand, Alex., Orazio Attanasio, Pedro Carniero, & Valérie Lechene (2016): The Effect of Gender-Targeted Conditional Cash Transfers on Household Expenditures: Evidence from a Randomized Experiment. IZA Discussion Paper No. 10133, August 2016.

Bagues, M., & P. Campa (2017): "Can gender quotas in candidate lists empower women? Evidence from a regression discontinuity design," CEPR Discussion Paper No. 12149

Baskaran, T., S. Bhalotra, B. Min, & Y. Uppal (2018): "Women Legislators and Economic Performance," IZA Discussion Paper No. 11596.

Beach, Brian & Daniel B. Jones (2017). Gridlock: Ethnic Diversity in Government and the Provision of Public Goods. American Economic Journal: Economic Policy 2017, 9(1): 112–136.

Beach, Brian, Daniel B. Jones, Tate Twinam, & Randall Walsh (2019). Minority Representation in Local Government. NBER Working Paper 25192

Beckwith, Karen & Kimberly Cowell-Meyers (2007). Sheer Numbers: Critical Representation Thresholds and Women's Political Representation. Perspectives on Politics 5(3), pp.553-565.

Bertrand, Marianne., Sandre E. Black, Sissel Jensen, & Adriana Lleras-Muney (2017): Breaking the Glass Ceiling? The Effect of Board Quotas on Female Labor Market Outcomes in Norway. NBER Working Paper 20256.

Besley, Timothy & Stephen Coate (1997). An economic model of representative democracy. Quarterly Journal of Economics 112, 85–114.

Blevins, C. & Cameron Mullins (2015). Jane, John ... Leslie? A Historical Method for Algorithmic Gender Prediction. Digital Humanities Quarterly 9(3), 2015.

Bobonis, Gustavo J. (2009): Is the Allocation of Resources within the Household Efficient? New Evidence from a Randomized Experiment. Journal of Political Economy 117(3).

Boutchkova, Maria, Angelica Gonzalez, Brian G. M. Main, & Vathunyoo Sila (2017). Gender Diversity and the Spillover Effects of Women on Boards. Available at SSRN: https://ssrn.com/abstract=2536887 or http://dx.doi.org/10.2139/ssrn.2536887

Brollo. F & U. Troiano (2016) "What happens when a woman wins a close election? Evidence from Brazil," Journal of Development Economics, 122, 28–45.

Cattaneo, Jansson and Ma (2018): Manipulation Testing based on Density Discontinuity. Stata Journal 18(1): 234-261.

Cellini, Stephanie R., Fernando Ferreira, & Jesse Rothstein (2010): The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design. Quarterly Journal of Economics 125(1), 215-261.

Chattopadhyay, Raghabendra & Esther Duflo (2004). Women as policy makers: evidence from a randomized policy experiment in India. Econometrica 72 (5), 1409–1443.

Chen, Li-Ju (2013). Do Female Politicians Influence Spending? Evidence From Taiwan. International

Journal of Applied Economics 10(2), 32-51.

Cinelli, C., Forney, A., & Pearl, J. (2022). A Crash Course in Good and Bad Controls. Sociological Methods & Research, May 20, 2022

Clots-Figueras, Irma (2012). Are female leaders good for education? Evidence from India. American Economic Journal: Applied Economics. 4 (1), 212–244.

Doepke, Matthias & Tertilt, Michèle (2017): Does Female Empowerment Promote Economic Development? (November 1, 2017). Global Poverty Research Lab Working Paper No. 17-112. Available at SSRN: https://ssrn.com/abstract=3127032 or http://dx.doi.org/10.2139/ssrn.3127032

Downs, Anthony (1957): An Economic Theory of Political Action in A Democracy. Journal of Political Economy 65(2), April 1957, pp. 135-150

Edlund, L. & R. Pande (2001): "Why have women become left-wing? The political gender gap and the decline in marriage." *Quarterly Journal of Economics*, 117, 917-961.

Ferreira, Fernando & Joseph Gyourko (2014). Does gender matter for political leadership? The case of U.S. mayors. Journal of Public Economics, Volume 112, 2014, Pages 24-39.

Funk, P., & C. Gathmann (2015): "Gender gaps in policy making: Evidence from direct democracy in Switzerland," Economic Policy, 30(81), 141–181.

Funk, K. D., & Philips, A. Q. (2019). Representative Budgeting: Women Mayors and the Composition of Spending in Local Governments. Political Research Quarterly 72(1), 19–33.
https://doi.org/10.1177/1065912918775237

Geys, B., & R. J. Sorensen (2019): "The impact of women above the political glass ceiling: Evidence from a Norwegian executive gender quota reform," Electoral Studies, 60.

Grembi, Veronica, Tommaso Nannicini, & Ugo Troiano. 2016. "Do Fiscal Rules Matter?" American Economic Journal: Applied Economics 8 (3): 1–30.

Hessami & Lopes da Fonseca (2020) Female Political Representation and Substantive Effects on Policies: A Literature Review. IZA DP No. 13125, April 2020.

Hora, Gurdeep S. (n.d.). Where are the female CEOs. A global study: Myths & Reality. Retrieved from https://www.ircsearchpartners.com/thought-leadership/where-are-women-ceos-myths-and-reality/

Inter-Parliamentary Union: Women in parliament in 2018: The year in review. Inter-Parliamentary Union, Geneva, Switzerland (March 2019). Retrieved from https://www.ipu.org/resources/publications/reports/2019-03/women-in-parliament-in-2018-year-in-review

Kanthak, Kristin & George A. Krause (2012). The Diversity Paradox: Political Parties, Legislatures, and the Organizational Foundations of Representation in America. Oxford University Press, USA, Apr 11, 2012

Kogan, Vladimir, Stephane Lavertu, & Zachary Peskowitz (2020): How does Minority Political Representation Affect School District Administration and Student Outcomes? American Journal of Political Science. Forthcoming. https://doi.org/10.1111/ajps.12587

Lee, David (2001): The Electoral Advantage to Incumbency and Voter's Valuation of Politicians' Experience: A Regression Discontinuity Analysis of Elections to the U.S. House. NBER Working Paper N8441.

Lee, David (2008): Randomized Experiments from Non-Random Selection in U.S. House Elections. *Journal of Econometrics* 142, 675-697.

Lee, David., Enrico Moretti, & Mathew Butler (2004): Do Voters Affect or Elect Policies? Evidence from the U.S. House. *Quarterly Journal of Economics* 119 (3): 807–59.

Lott, J.R. & L.W. Kenny (1999): "Did women's suffrage change the size and scope of government?" *Journal of Political Economy*, 107, 1163-1198.

Mathews J. (2020): "Madrid had 57 council members. Seoul has 100. Why does L.A. have 15?" September 1, 2020. Accessed at: www.zocalopublicsquare.org/2020/09/01/strong-mayors-california-city-government-city-council/ideas/connecting-california/

PPIC (2020). Public Policy Institute of California Statewide Survey. Survey Director: Mark Baldassare. Accessed at www.ppic.org/survey.

Rigon, Massimiliano & Giulia Tanzi (2012). Does Gender Matter for Public Spending? Empirical Evidence from Italian Municipalities. Bank of Italy Temi di Discussione (Working Paper) No. 862, 15 May 2012.

Smith, Nina (2018). Gender Quotas on Boards of Directors. IZA World of Labor 2018 7v2, doi: 10.15185/izawol.7.v2

Svaleryd, Helena (2009). Women's representation and public spending. European Journal of Political Economy vol. 25, issue 2, 186-198.

Thistlethwaithe, Donald. L., & Donald T. Campbell (1960): Regression-Discontinuity Analysis: An Alternative to the Ex Post Facto Experiment. *Journal of Educational Psychology* 51(6), 309-317.

WEF (2020): The Global Gender Gap Report 2020. Discussion Paper, World Economic Forum.

3.8 Appendix

3.8.1 Additional Figures

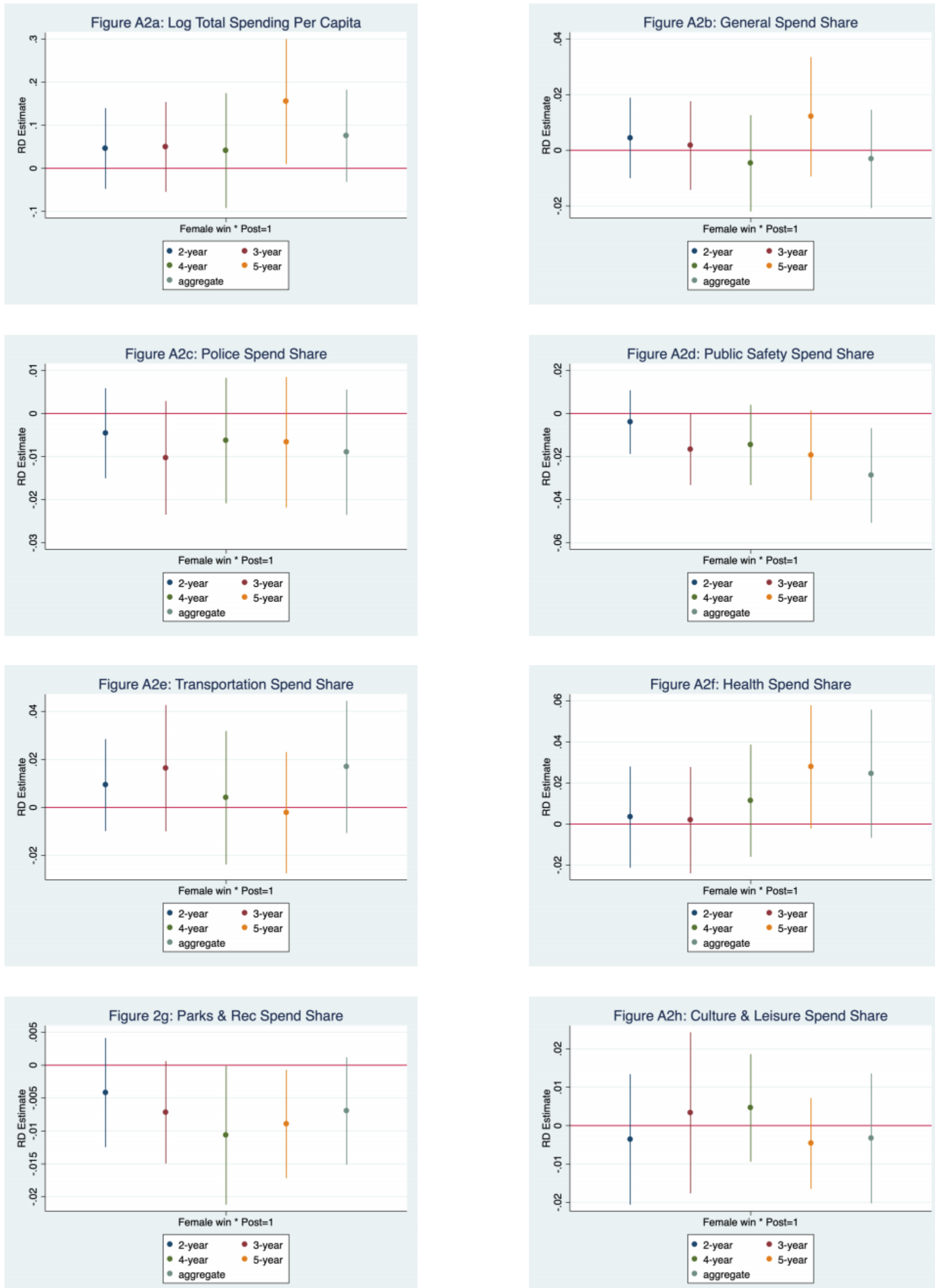


Figure 3.6: “Stacked” RD Estimates

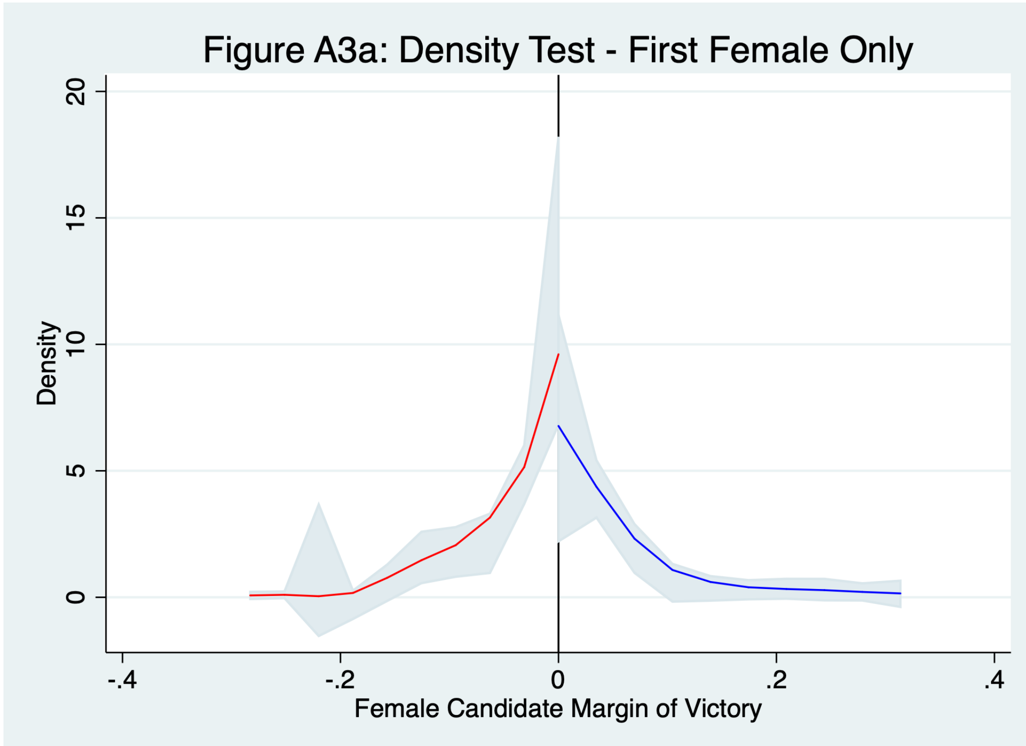


Figure 3.7: Density Test - First Female Only

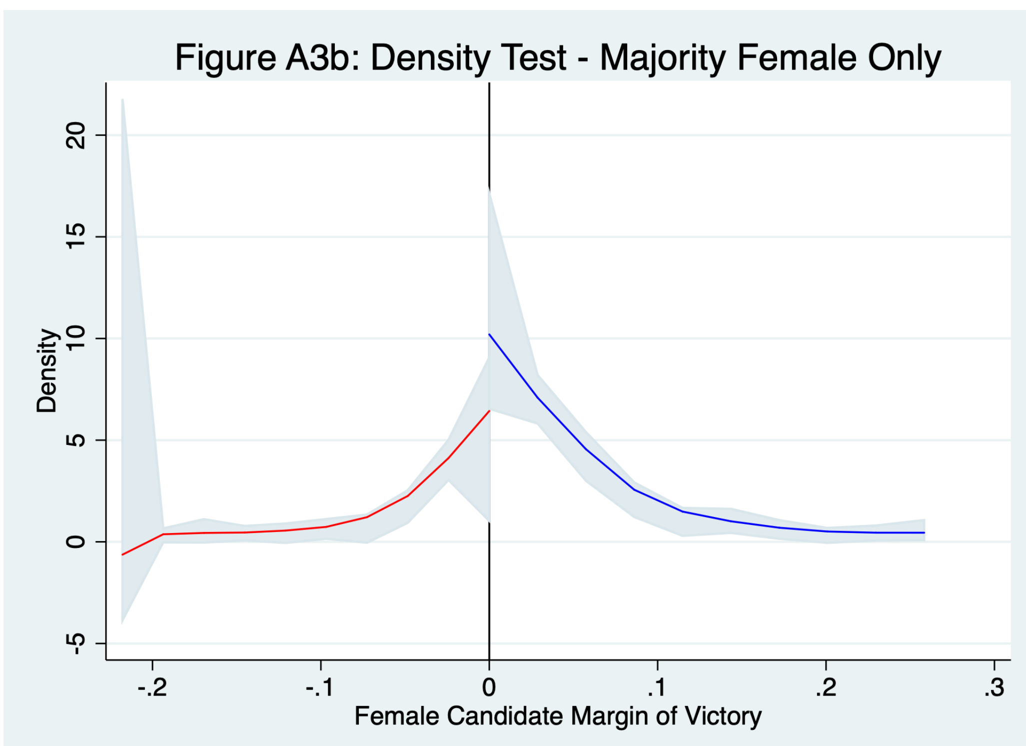


Figure 3.8: Density Test - Majority Female Only

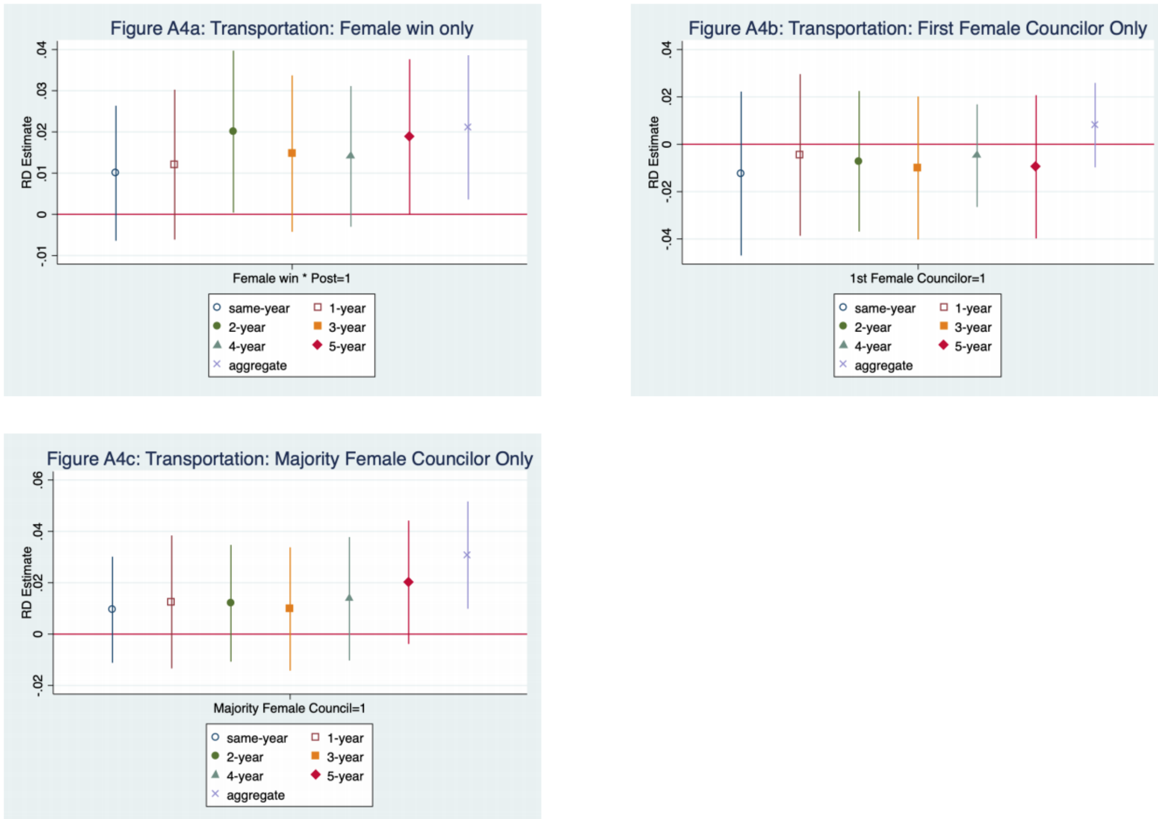


Figure 3.9: RD Separate Heterogeneous Transportation Effects

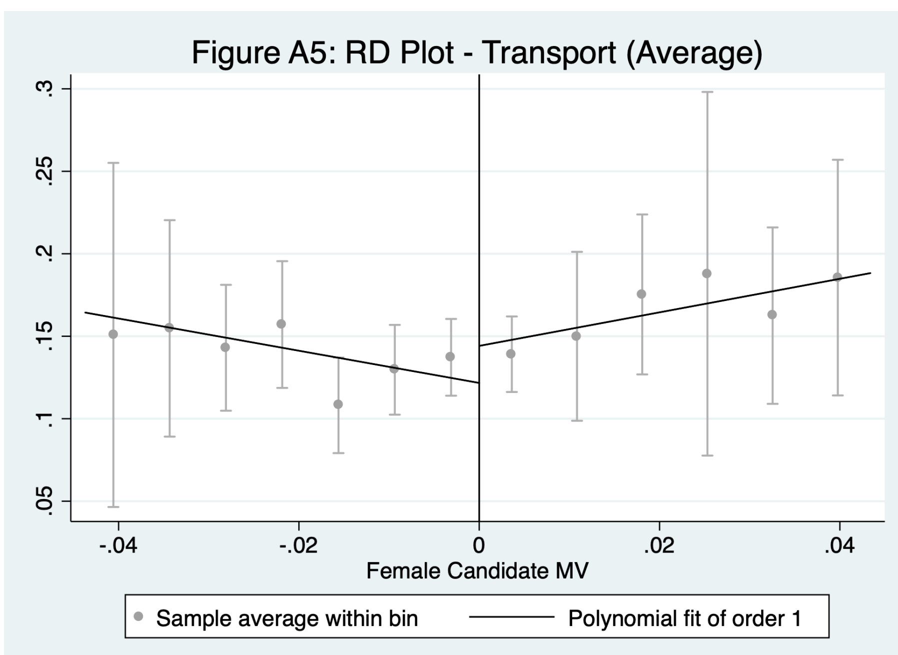


Figure 3.10: RD Plot - Average Post-Election Transport Share by MV

3.8.2 Additional Tables

Table 3.8: Identification Assumption - RD with Baseline Variables as Outcomes

Table A1: Identification Assumption - RD with Baseline Variables as Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Population	Ln Exp./Cap.	Share General	Share Police	Share P Safe	Share Transport	Share Health	Share Parks/ Rec	Share Culture
F Win	491.188	-0.025	0.012	0.001	-0.004	0.000	-0.008	-0.003	0.003
	(898.143)	(0.039)	(0.008)	(0.005)	(0.007)	(0.009)	(0.015)	(0.004)	(0.005)
Outcome Mean	47,515	7.28	0.128	0.197	0.274	0.152	0.098	0.063	0.087
Obs.	1,657	1,631	1,658	1,658	1,658	1,658	1,658	1,658	1,658

Standard errors are in brackets. For all columns except (1), the top 1% of observations by total spending are trimmed from the sample in the first row, and 0 values are trimmed for all outcome variables, so I am capturing the treatment effect for cities that actually provide that public good. All outcomes are measured in the fiscal year ending in the year of the election (i.e. if election occurs in November of year t, then the baseline outcome is measure for the fiscal year ending in t, usually on June 30). The first row 'F Win' shows the RD treatment effect. To allow for comparison and similar sample sizes for each column, I use a 7% bandwidth on the running variable. My main results in Figures 2 and 3 are robust to using this bandwidth across the board. I include year FE in all specifications. Standard errors are clustered at the city level in all specifications.

**** p<0.01, ** p<0.05, * p<0.1*

Table 3.9: Simple Incumbency Effects

Table A2: Simple Incumbency Effects

	(1)	(2)	(3)	(4)
	Marginal Female Win		Marginal Male Win	
Female Incumbent	0.249***	0.204***		
	(0.028)	(0.036)		
Male Incumbent			0.282***	0.266***
			(0.028)	(0.036)
Constant	0.446***	0.364***	0.397***	0.547***
	(0.016)	(0.088)	(0.015)	(0.092)
Obs.	1429	1429	1429	1429
Year FE		X		X
City FE		X		X
Clustered SE (City)		X		X

Standard errors are in parenthesis

**** p<0.01, ** p<0.05, * p<0.1*

Table 3.10: Effect of Marginal Female Victory on Expenditure & Composition (2-Year Lag)

Table A3: Impact of the first observed female (vs. male) election win on expenditure and composition of public spending (2 year lag)

	(1)	(2)	(3)	(4)	(5)	(6)
		Optimal Bandwidths				
Dependent Variable	Mean Share	CCT(0)	CCT(1)	CCT(0)	CCT(1)	CCT(1)
Ln(PC Total Spending)	-	1.6%	1.6%	-0.017	-0.017	-0.022
				(0.069)	(0.069)	(0.099)
Share: General	12%	48.3%	5.7%	0.007	0.013	0.009
				(0.009)	(0.009)	(0.011)
Share: Police	20%	6.0%	5.2%	0.003	0.0003	-0.004
				(0.006)	(0.007)	(0.007)
Share: Public Safety	27%	4.8%	4.8%	-0.010	-0.010	-0.013
				(0.009)	(0.009)	(0.011)
Share: Transport	15%	5.0%	4.4%	0.017*	0.020**	0.010
				(0.010)	(0.010)	(0.012)
Share: Health	11%	3.4%	3.2%	-0.011	-0.013	-0.016
				(0.011)	(0.011)	(0.011)
Share: Parks & Rec	6%	10.3%	7.9%	-0.004	-0.004	-0.006
				(0.004)	(0.004)	(0.006)
Share: Culture & Leisure	9%	4.6%	4.1%	-0.002	-0.001	0.007
				(0.006)	(0.006)	(0.009)
Polynomial Degree				1	1	2

Standard errors are in brackets. The top 1% of observations by total spending are trimmed from the sample in the first row, and 0 values are trimmed for all outcome variables, so I am capturing the treatment effect for cities that actually provide that public good. Ln(PC Total Spending) is the log of per-capita expenditure in period t+2. Other dependent variables are the share of total expenditure on public goods allocated to the given category. Column (1) shows the average share of total public spending each category constitutes. Columns (2) and (3) shows the mean squared error (MSE) optimal bandwidth generated by the Stata package 'rdbwselect'. In Column (2) there is no regularization term. In Column (3) the default setting is used (scale parameter = 1) The parameter of interest in Columns (4)-(6) is a binary variable that 'turns on' in cities where the female candidate wins the first close male vs. female election, in years once the election has taken place. I include year FE in all specifications. Standard errors are clustered at the city level in all specifications.
**** p<0.01, ** p<0.05, * p<0.1*

Table 3.11: Effect of Marginal Female Victory on Expenditure & Composition (3-Year Lag)

Table A4: Impact of the first observed female (vs. male) election win on expenditure and composition of public spending (3 year lag)

	(1)	(2)	(3)	(4)	(5)	(6)
		Optimal Bandwidths				
Dependent Variable	Mean Share	CCT(0)	CCT(1)	CCT(0)	CCT(1)	CCT(1)
Ln(PC Total Spending)	-	1.8%	1.8%	-0.082	-0.082	-0.055
				(0.051)	(0.051)	(0.072)
Share: General	12%	11.5%	5.6%	-0.0002	0.008	0.008
				(0.007)	(0.010)	(0.011)
Share: Police	20%	4.9%	4.6%	0.003	0.003	0.009
				(0.007)	(0.007)	(0.008)
Share: Public Safety	27%	4.6%	4.5%	-0.002	-0.002	0.002
				(0.009)	(0.009)	(0.009)
Share: Transport	15%	4.1%	3.9%	0.016*	0.015	0.010
				(0.010)	(0.010)	(0.013)
Share: Health	11%	4.3%	3.6%	-0.005	-0.006	-0.020**
				(0.007)	(0.007)	(0.009)
Share: Parks & Rec	6%	14.0%	6.3%	-0.004	-0.007*	-0.011**
				(0.004)	(0.004)	(0.005)
Share: Culture & Leisure	9%	4.1%	3.8%	-0.008	-0.010*	-0.004
				(0.007)	(0.006)	(0.007)
Polynomial Degree				1	1	2

Standard errors are in brackets. The top 1% of observations by total spending are trimmed from the sample in the first row, and 0 values are trimmed for all outcome variables, so I am capturing the treatment effect for cities that actually provide that public good. Ln(PC Total Spending) is the log of per-capita expenditure in period t+3. Other dependent variables are the share of total expenditure on public goods allocated to the given category. Column (1) shows the average share of total public spending each category constitutes. Columns (2) and (3) shows the mean squared error (MSE) optimal bandwidth generated by the Stata package 'rdbwselect'. In Column (2) there is no regularization term. In Column (3) the default setting is used (scale parameter = 1) The parameter of interest in Columns (4)-(6) is a binary variable that 'turns on' in cities where the female candidate wins the first close male vs. female election, in years once the election has taken place. I include year FE in all specifications. Standard errors are clustered at the city level in all specifications.
**** p<0.01, ** p<0.05, * p<0.1*

Table 3.12: Effect of Marginal Female Victory on Expenditure & Composition (4-Year Lag)

Table A5: Impact of the first observed female (vs. male) election win on expenditure and composition of public spending (4 year lag)

	(1)	(2)	(3)	(4)	(5)	(6)
		Optimal Bandwidths				
Dependent Variable	Mean Share	CCT(0)	CCT(1)	CCT(0)	CCT(1)	CCT(1)
Ln(PC Total Spending)	-	1.9%	1.9%	-0.021	-0.021	-0.136**
				(0.049)	(0.049)	(0.069)
Share: General	12%	10.7%	6.4%	-0.004	0.004	0.011
				(0.006)	(0.006)	(0.009)
Share: Police	20%	4.5%	4.3%	0.007	0.007	0.014*
				(0.006)	(0.006)	(0.008)
Share: Public Safety	27%	4.7%	4.6%	0.004	0.004	0.004
				(0.008)	(0.008)	(0.011)
Share: Transport	15%	4.3%	4.1%	0.014	0.014	0.021*
				(0.009)	(0.009)	(0.011)
Share: Health	11%	5.7%	5.2%	-0.002	-0.004	-0.006
				(0.007)	(0.007)	(0.008)
Share: Parks & Rec	6%	5.1%	5.0%	-0.011**	-0.011**	-0.015**
				(0.005)	(0.005)	(0.007)
Share: Culture & Leisure	9%	5.8%	4.7%	-0.008	-0.012*	-0.016*
				(0.006)	(0.007)	(0.009)
Polynomial Degree				1	1	2

Standard errors are in brackets. The top 1% of observations by total spending are trimmed from the sample in the first row, and 0 values are trimmed for all outcome variables, so I am capturing the treatment effect for cities that actually provide that public good. Ln(PC Total Spending) is the log of per-capita expenditure in period t+4. Other dependent variables are the share of total expenditure on public goods allocated to the given category. Column (1) shows the average share of total public spending each category constitutes. Columns (2) and (3) shows the mean squared error (MSE) optimal bandwidth generated by the Stata package 'rdbwselect'. In Column (2) there is no regularization term. In Column (3) the default setting is used (scale parameter = 1) The parameter of interest in Columns (4)-(6) is a binary variable that 'turns on' in cities where the female candidate wins the first close male vs. female election, in years once the election has taken place. I include year FE in all specifications. Standard errors are clustered at the city level in all specifications.
**** p<0.01, ** p<0.05, * p<0.1*

Table 3.13: Effect of Marginal Female Victory on Expenditure & Composition (5-Year Lag)

Table A6: Impact of the first observed female (vs. male) election win on expenditure and composition of public spending (5 year lag)

	(1)	(2)	(3)	(4)	(5)	(6)
		Optimal Bandwidths				
Dependent Variable	Mean Share	CCT(0)	CCT(1)	CCT(0)	CCT(1)	CCT(1)
Ln(PC Total Spending)	-	2.0%	2.0%	0.042	0.042	-0.094
				(0.049)	(0.049)	(0.071)
Share: General	12%	4.6%	4.2%	-0.003	-0.002	-0.006
				(0.008)	(0.008)	(0.012)
Share: Police	20%	4.3%	4.2%	-0.002	-0.002	0.007
				(0.005)	(0.005)	(0.007)
Share: Public Safety	27%	4.7%	4.5%	-0.004	-0.005	-0.001
				(0.008)	(0.008)	(0.010)
Share: Transport	15%	5.0%	4.6%	0.015	0.019*	0.043***
				(0.010)	(0.010)	(0.011)
Share: Health	11%	4.3%	4.4%	0.007	0.008	-0.010
				(0.008)	(0.008)	(0.009)
Share: Parks & Rec	6%	5.6%	5.2%	-0.008**	-0.009**	-0.013**
				(0.004)	(0.004)	(0.006)
Share: Culture & Leisure	9%	5.9%	4.6%	-0.008	-0.010	-0.020**
				(0.005)	(0.006)	(0.009)
Polynomial Degree				1	1	2

Standard errors are in brackets. The top 1% of observations by total spending are trimmed from the sample in the first row, and 0 values are trimmed for all outcome variables, so I am capturing the treatment effect for cities that actually provide that public good. Ln(PC Total Spending) is the log of per-capita expenditure in period t+5. Other dependent variables are the share of total expenditure on public goods allocated to the given category. Column (1) shows the average share of total public spending each category constitutes. Columns (2) and (3) shows the mean squared error (MSE) optimal bandwidth generated by the Stata package 'rdbwselect'. In Column (2) there is no regularization term. In Column (3) the default setting is used (scale parameter = 1) The parameter of interest in Columns (4)-(6) is a binary variable that 'turns on' in cities where the female candidate wins the first close male vs. female election, in years once the election has taken place. I include year FE in all specifications. Standard errors are clustered at the city level in all specifications.
**** p<0.01, ** p<0.05, * p<0.1*

Table 3.14: Effect of Marginal Female Victory on Expenditure & Composition (Aggregate)

Table A7: Impact of the first observed female (vs. male) election win on expenditure and composition of public spending (Aggregate)

	(1)	(2)	(3)	(4)	(5)	(6)
		Optimal Bandwidths				
Dependent Variable	Mean Share	CCT(0)	CCT(1)	CCT(0)	CCT(1)	CCT(1)
Ln(PC Total Spending)	-	1.7%	1.6%	-0.069	-0.025	-0.003
				(0.042)	(0.035)	(0.050)
Share: General	12%	4.2%	3.9%	0.006	0.010*	0.009
				(0.006)	(0.006)	(0.007)
Share: Police	20%	4.3%	3.9%	0.001	0.004	0.003
				(0.007)	(0.007)	(0.006)
Share: Public Safety	27%	3.3%	3.2%	0.001	-0.0003	-0.006
				(0.009)	(0.009)	(0.011)
Share: Transport	15%	3.9%	3.8%	0.020**	0.021**	0.024*
				(0.009)	(0.009)	(0.013)
Share: Health	11%	3.2%	3.1%	-0.008	-0.008	-0.015*
				(0.006)	(0.007)	(0.008)
Share: Parks & Rec	6%	11.4%	6.9%	-0.004	-0.007*	-0.014***
				(0.004)	(0.004)	(0.005)
Share: Culture & Leisure	9%	3.2%	3.1%	-0.016**	-0.016**	0.004
				(0.007)	(0.008)	(0.006)
Polynomial Degree				1	1	2

Standard errors are in brackets. The top 1% of observations by total spending are trimmed from the sample in the first row, and 0 values are trimmed for all outcome variables, so I am capturing the treatment effect for cities that actually provide that public good. Ln(PC Total Spending) is the sum of log of per-capita expenditure 2, 3, 4, & 5 years after the election year of interest . Other dependent variables are the *average* shares of total expenditure on public goods allocated to the given category across the four years. Column (1) shows the average of these average shares of total public spending. Columns (2) and (3) shows the mean squared error (MSE) optimal bandwidth generated by the Stata package 'rdbwselect'. In Column (2) there is no regularization term. In Column (3) the default setting is used (scale parameter = 1) The parameter of interest in Columns (4)-(6) is a binary variable that 'turns on' in cities where the female candidate wins the first close male vs. female election, in years once the election has taken place. I include year FE in all specifications. Standard errors are clustered at the city level in all specifications.
**** p<0.01, ** p<0.05, * p<0.1*

Table 3.15: Heterogeneity of Marginal Female Victory (2-Year Lag)

Table A8: Examining heterogeneity (by council composition) in the effect of the first observed female (vs. male) election win (2 year lag)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ln Exp./Cap.	Share General	Share Police	Share P Safe	Share Transport	Share Health	Share Parks/ Rec	Share Culture
F Win	-0.052	0.018**	0.005	-0.008	0.027*	-0.019	-0.007	-0.010
	(0.086)	(0.009)	(0.008)	(0.012)	(0.015)	(0.015)	(0.005)	(0.008)
MV	-26.087*	0.470	-0.107	-0.510	0.677	-0.095	0.057	-0.312
	(13.618)	(0.363)	(0.270)	(0.455)	(0.490)	(1.186)	(0.137)	(0.366)
F Win * MV	9.511	-0.636	-0.249	0.277	-0.500	0.504	0.180	0.232
	(8.745)	(0.398)	(0.280)	(0.415)	(0.441)	(0.555)	(0.126)	(0.342)
1st F Councilor	-0.118***	0.028	0.010*	0.019*	-0.042***	-0.011	0.009	0.005
	(0.045)	(0.020)	(0.006)	(0.010)	(0.009)	(0.009)	(0.007)	(0.018)
Maj F Council	0.086	-0.045***	-0.033***	-0.018**	-0.068***	0.222***	-0.022***	-0.037***
	(0.067)	(0.007)	(0.005)	(0.008)	(0.010)	(0.022)	(0.003)	(0.006)
F Win * 1st F Councilor	0.076	-0.049**	-0.005	-0.007	0.028*	0.021	-0.010	0.005
	(0.065)	(0.022)	(0.009)	(0.013)	(0.016)	(0.014)	(0.010)	(0.021)
F Win * Maj F Council		0.043***	0.021**	0.009	0.058***	-0.216***	0.031***	0.051***
		(0.013)	(0.009)	(0.013)	(0.015)	(0.025)	(0.007)	(0.010)
Obs.	615	1599	1523	1450	1382	1133	1827	1340

Standard errors are in parenthesis. The top 1% of observations are trimmed from column (1), and 0 values are dropped for all columns. Ln Exp./Cap. is the log of per-capita expenditure in period t+2. Other dependent variables are the share of total expenditure on public goods allocated to the given category. 1st F Councilor is an indicator that =1 (in the post period) if the council previously had no female councilors, and elected at least 1 female councilor in the election of interest. Maj F Council is an indicator that =1 (in the post period) if the council was a male majority before the election, and switches to female majority after the election. The interaction terms F Win * 1st F Councilor and F Win * Maj F Council are the parameters that measure the effect of a **marginal** female victory that results in the 1st female councilor and a majority female councilor respectively on spending outcomes. I use the CCT bandwidths from column (3) in the previous table, which are slightly more conservative than those in column (2). I include year FE in all specifications. Standard errors are clustered at the city level in all specifications.
**** p<0.01, ** p<0.05, * p<0.1*

Table 3.16: Heterogeneity of Marginal Female Victory (3-Year Lag)

Table A9: Examining heterogeneity (by council composition) in the effect of the first observed female (vs. male) election win (3 year lag)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ln Exp./Cap.	Share General	Share Police	Share P Safe	Share Transport	Share Health	Share Parks/ Rec	Share Culture
F Win	-0.085	0.011	0.006	-0.005	0.023	-0.005	-0.009*	-0.021**
	(0.065)	(0.009)	(0.009)	(0.012)	(0.014)	(0.009)	(0.005)	(0.009)
MV	-9.968	0.392	-0.174	-0.303	-0.004	-0.017	0.084	-0.401
	(10.558)	(0.367)	(0.291)	(0.475)	(0.450)	(0.810)	(0.173)	(0.378)
F Win * MV	17.192***	-0.526	-0.301	0.049	-0.093	0.202	0.213	0.749**
	(6.487)	(0.439)	(0.316)	(0.398)	(0.400)	(0.356)	(0.171)	(0.368)
1st F Councilor	-0.042	-0.021*	0.017**	0.026	-0.005	-0.010	0.016	0.001
	(0.067)	(0.011)	(0.009)	(0.018)	(0.012)	(0.010)	(0.011)	(0.013)
Maj F Council	0.057	-0.037***	-0.034***	-0.037***	-0.028***	0.147***	-0.007***	-0.019***
	(0.066)	(0.005)	(0.003)	(0.004)	(0.006)	(0.007)	(0.002)	(0.003)
F Win * 1st F Councilor	-0.075	-0.004	-0.011	-0.007	-0.010	0.013	-0.009	0.012
	(0.085)	(0.015)	(0.013)	(0.022)	(0.018)	(0.013)	(0.013)	(0.018)
F Win * Maj F Council		0.038***	0.026***	0.034***	0.017	-0.150***	0.011*	0.035***
		(0.014)	(0.010)	(0.012)	(0.015)	(0.011)	(0.006)	(0.009)
Obs.	756	1695	1515	1498	1410	1344	1810	1384

Standard errors are in parenthesis. The top 1% of observations are trimmed from column (1), and 0 values are dropped for all columns. Ln Exp./Cap. is the log of per-capita expenditure in period t+3. Other dependent variables are the share of total expenditure on public goods allocated to the given category. 1st F Councilor is an indicator that =1 (in the post period) if the council previously had no female councilors, and elected at least 1 female councilor in the election of interest. Maj F Council is an indicator that =1 (in the post period) if the council was a male majority before the election, and switches to female majority after the election. The interaction terms F Win * 1st F Councilor and F Win * Maj F Council are the parameters that measure the effect of a **marginal** female victory that results in the 1st female councilor and a majority female councilor respectively on spending outcomes. I use the CCT bandwidths from column (3) in the previous table, which are slightly more conservative than those in column (2). I include year FE in all specifications. Standard errors are clustered at the city level in all specifications.
****** p<0.01, ** p<0.05, * p<0.1*

Table 3.17: Heterogeneity of Marginal Female Victory (4-Year Lag)

Table A10: Examining heterogeneity (by council composition) in the effect of the first observed female (vs. male) election win (4 year lag)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ln Exp./Cap.	Share General	Share Police	Share P Safe	Share Transport	Share Health	Share Parks/ Rec	Share Culture
F Win	-0.027	0.003	0.014*	0.004	0.022*	-0.003	-0.015**	-0.018**
	(0.063)	(0.007)	(0.008)	(0.011)	(0.012)	(0.009)	(0.007)	(0.009)
MV	-0.629	0.273	-0.039	-0.115	0.298	-0.317	-0.037	-0.126
	(8.897)	(0.247)	(0.316)	(0.445)	(0.470)	(0.458)	(0.197)	(0.304)
F Win * MV	4.027	-0.280	-0.493	-0.084	-0.588	0.253	0.599**	0.665*
	(5.351)	(0.249)	(0.337)	(0.399)	(0.375)	(0.352)	(0.278)	(0.373)
1st F Councilor	0.022	-0.001	0.003	0.008	0.018	-0.009	0.016	-0.004
	(0.065)	(0.014)	(0.006)	(0.010)	(0.021)	(0.007)	(0.011)	(0.017)
Maj F Council	0.027	-0.014	0.009	0.033***	-0.060***	-0.019	0.012	0.010
	(0.073)	(0.017)	(0.017)	(0.004)	(0.007)	(0.012)	(0.008)	(0.026)
F Win * 1st F Councilor	-0.059	-0.001	-0.014	0.005	-0.041*	0.014	-0.011	0.015
	(0.077)	(0.017)	(0.016)	(0.018)	(0.024)	(0.011)	(0.014)	(0.021)
F Win * Maj F Council		0.018	-0.020	-0.040***	0.055***	0.014	-0.003	-0.001
		(0.018)	(0.018)	(0.010)	(0.014)	(0.014)	(0.011)	(0.027)
Obs.	805	1860	1462	1513	1438	1634	1585	1528

Standard errors are in parenthesis. The top 1% of observations are trimmed from column (1), and 0 values are dropped for all columns. Ln Exp./Cap. is the log of per-capita expenditure in period t+4. Other dependent variables are the share of total expenditure on public goods allocated to the given category. 1st F Councilor is an indicator that =1 (in the post period) if the council previously had no female councilors, and elected at least 1 female councilor in the election of interest. Maj F Council is an indicator that =1 (in the post period) if the council was a male majority before the election, and switches to female majority after the election. The interaction terms F Win * 1st F Councilor and F Win * Maj F Council are the parameters that measure the effect of a **marginal** female victory that results in the 1st female councilor and a majority female councilor respectively on spending outcomes. I use the CCT bandwidths from column (3) in the previous table, which are slightly more conservative than those in column (2). I include year FE in all specifications. Standard errors are clustered at the city level in all specifications.
****** p<0.01, ** p<0.05, * p<0.1*

Table 3.18: Heterogeneity of Marginal Female Victory (5-Year Lag)

Table A11: Examining heterogeneity (by council composition) in the effect of the first observed female (vs. male) election win (5 year lag)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ln Exp./Cap.	Share General	Share Police	Share P Safe	Share Transport	Share Health	Share Parks/ Rec	Share Culture
F Win	0.044	-0.002	0.001	-0.008	0.029**	0.013	-0.012**	-0.018**
	(0.072)	(0.010)	(0.007)	(0.009)	(0.012)	(0.012)	(0.006)	(0.008)
MV	-0.208	0.494	-0.236	-0.272	0.281	-0.121	0.004	-0.147
	(8.594)	(0.455)	(0.303)	(0.458)	(0.388)	(0.613)	(0.183)	(0.316)
F Win * MV	-0.971	0.208	-0.220	0.165	-0.659	-0.418	0.435**	0.594
	(5.746)	(0.395)	(0.244)	(0.313)	(0.409)	(0.424)	(0.198)	(0.367)
1st F Councilor	-0.077	-0.001	-0.004	-0.006	0.008	-0.011	0.015*	-0.002
	(0.064)	(0.017)	(0.012)	(0.017)	(0.021)	(0.010)	(0.009)	(0.010)
Maj F Council	0.154	-0.018	0.052***	0.081***	-0.089***	-0.100***	0.015***	0.026
	(0.105)	(0.023)	(0.017)	(0.018)	(0.009)	(0.036)	(0.002)	(0.021)
F Win * 1st F Councilor	0.075	-0.005	0.000	0.026	-0.034	0.006	-0.012	0.015
	(0.078)	(0.019)	(0.015)	(0.021)	(0.024)	(0.014)	(0.011)	(0.015)
F Win * Maj F Council	-0.158	0.021	-0.057***	-0.084***	0.079***	0.091**	-0.008	-0.013
	(0.117)	(0.024)	(0.018)	(0.019)	(0.014)	(0.037)	(0.006)	(0.022)
Obs.	796	1396	1396	1441	1457	1432	1574	1457

Standard errors are in parenthesis. The top 1% of observations are trimmed from column (1), and 0 values are dropped for all columns. Ln Exp./Cap. is the log of per-capita expenditure in period t+5. Other dependent variables are the share of total expenditure on public goods allocated to the given category. 1st F Councilor is an indicator that =1 (in the post period) if the council previously had no female councilors, and elected at least 1 female councilor in the election of interest. Maj F Council is an indicator that =1 (in the post period) if the council was a male majority before the election, and switches to female majority after the election. The interaction terms F Win * 1st F Councilor and F Win * Maj F Council are the parameters that measure the effect of a **marginal** female victory that results in the 1st female councilor and a majority female councilor respectively on spending outcomes. I use the CCT bandwidths from column (3) in the previous table, which are slightly more conservative than those in column (2). I include year FE in all specifications. Standard errors are clustered at the city level in all specifications.
****** *p<0.01, ** p<0.05, * p<0.1*

Table 3.19: Heterogeneity of Marginal Female Victory (Aggregate)

Table A12: Examining heterogeneity (by council composition) in the effect of the first observed female (vs. male) election win (Aggregate)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ln Exp./Cap.	Share General	Share Police	Share P Safe	Share Transport	Share Health	Share Parks/ Rec	Share Culture
F Win	-0.033	0.012*	0.008	0.005	0.031**	-0.005	-0.011**	-0.030***
	(0.046)	(0.007)	(0.009)	(0.012)	(0.015)	(0.009)	(0.005)	(0.011)
MV	-23.796*	1.041**	-0.192	-0.260	0.494	-0.752	0.063	-0.688
	(13.116)	(0.527)	(0.352)	(0.713)	(0.514)	(0.851)	(0.152)	(0.526)
F Win * MV	9.184**	-0.718*	-0.341	-0.430	-0.943**	0.809**	0.220*	1.236**
	(4.115)	(0.369)	(0.288)	(0.401)	(0.415)	(0.375)	(0.130)	(0.580)
1st F Councilor	0.005	-0.016*	0.011***	0.015**	-0.008	0.000	0.006*	-0.003
	(0.019)	(0.008)	(0.003)	(0.007)	(0.013)	(0.006)	(0.003)	(0.009)
Maj F Council	0.025	-0.046***	-0.005	0.013	-0.088***	0.178***	-0.015***	-0.029***
	(0.036)	(0.011)	(0.006)	(0.014)	(0.010)	(0.016)	(0.003)	(0.009)
F Win * 1st F Councilor	-0.043	-0.001	-0.006	-0.006	-0.024	-0.002	0.002	0.029*
	(0.034)	(0.012)	(0.007)	(0.011)	(0.018)	(0.011)	(0.009)	(0.017)
F Win * Maj F Council		0.050***	-0.007	-0.026	0.082***	-0.185***	0.021***	0.048***
		(0.013)	(0.009)	(0.017)	(0.017)	(0.017)	(0.006)	(0.012)
Obs.	545	1154	1154	1002	1134	985	1556	985

Standard errors are in parenthesis. The top 1% of observations are trimmed from column (1), and 0 values are dropped for all columns. Ln Exp./Cap. is the sum of log of per-capita expenditure 2, 3, 4, & 5 years after the election year of interest. Other dependent variables are the *average* shares of total expenditure on public goods allocated to the given category across the four years. 1st F Councilor is an indicator that =1 (in the post period) if the council previously had no female councilors, and elected at least 1 female councilor in the election of interest. Maj F Council is an indicator that =1 (in the post period) if the council was a male majority before the election, and switches to female majority after the election. The interaction terms F Win * 1st F Councilor and F Win * Maj F Council are the parameters that measure the effect of a **marginal** female victory that results in the 1st female councilor and a majority female councilor respectively on spending outcomes. I use the CCT bandwidths from column (3) in the previous table, which are slightly more conservative than those in column (2). I include year FE in all specifications. Standard errors are clustered at the city level in all specifications.
**** p<0.01, ** p<0.05, * p<0.1*