THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Essays in the economics of transportation, housing and discrimination

Cheng Keat Tang

October, 2018

A thesis submitted to the Department of Geography and Environment of the London School of Economics and Political Science for the degree of Doctor of Philosophy, London, 2018

Declaration

I certify that the thesis I have presented here for my PhD is solely my own work unless otherwise indicated. The copyright of this thesis rests with the author. While quotation from it is permissible, provided that full acknowledgement is made, content of this thesis shall not be reproduced without the prior written consent of the author. I warrant this authorisation does not infringe the rights of any third party. I declare that this thesis consist of 52,530 words.

Abstract

This thesis consists of three papers that are related to transportation economics, housing economics, and the economics of discrimination.

The first paper examines how much people are willing to pay (WTP), on average, to avoid road traffic near their residence using the housing market. The notion is that traffic confers substantial negative externalities such as congestion delays, air and noise pollution, and traffic accidents. Estimating these hedonic functions are, however, extremely challenging with omitted variable bias and sorting of households. Hence, to circumvent these challenges, I rely on the sharp variation in traffic conditions induced by the London Congestion Charge to recover the capitalization of road traffic on housing values.

The second paper examines whether installing speed cameras reduces traffic accidents and saves lives. Speeding is one of the major reasons why accidents occur, and the velocity of the vehicles affects the gravity of collisions. This paper sheds insights on how the government could intervene and nudge drivers from risky behaviours that could have serious consequences.

The third paper investigates whether facial attractiveness affects sentencing outcomes in courtrooms. I rely on a novel facial recognition system that locates various features from inmate mugshots to compute facial symmetry as a measure of attractiveness. This study is motivated by the burgeoning literature indicating that judges allow extraneous factors, such as race and gender of defendants, emotions and media attention, to influence their decisions, and the widespread discrimination of appearance in multiple contexts.

Acknowledgements

I would like to express my deepest gratitude to my supervisors, Professor Steve Gibbons and Professor Henry Overman for their guidance over the last four years. I have benefited tremendously from my in-depth discussions with them and would certainly not be able to complete my dissertation without their guidance.

I would also like to thank other faculty members in the department who have provided me helpful comments during the LSE work-in-progress seminars. This include Olmo Silva, Vernon Henderson, Sefi Roth, Felipe Carozzi, Gabriel Ahlfeldt and Christian Hilber.

Others who have been instrumental in my research progress include faculty members from other institutes who visited LSE. In particular, Hans Koster and Chris Cunningham have took their time to give me invaluable comments to improve my work.

Finally, I would also like to thank my peers, Ying, Hayoung, Pascal, Nathalie, Cong, Carl and Tanner for their constant support. It is through the constant discussions and arguments over concepts that have really deepen my understanding on research.

Last and most importantly, I would like to thank my fiancee, Shimin and my family for their support and understanding. This could never been possible without them.

Contents

CHAPTER 1

Preface

This thesis consists of three essays related to transportation economics, housing economics, and the economics of discrimination. Each essay addresses a topic that is surprisingly neglected by the existing literature despite its importance. This is largely due to the lack of quality data, the difficulty to measure the variable of interest, and the presence of endogeneity issues that impede causal inferences.

Bearing these challenges in mind, in my first paper, I am interested in:

"What is the average willingness to pay for less traffic near ones' residence ?"

Traffic congestion is an ubiquitous problem that many major cities face. Road traffic confers substantial negative externalities, including time delays being stuck in the traffic, air and noise pollution, traffic accidents etc. Thus, homeowners are likely to pay to avoid traffic near their residence. In this paper, I estimate the hedonic house price function relying on micro data on housing transactions and road traffic across Central London to measure the willing to pay (WTP) to avoid traffic. The idea is that traffic conditions (e.g. traffic flow or noise) vary across space and are capitalized into home prices. If one is able to control for the other differences between these property sales, the remaining variation in prices should reflect the WTP to avoid traffic. In reality, recovering these estimates is extremely challenging with omitted variable bias and the sorting of households across space.

To circumvent these challenges, this paper exploits the London Congestion Charge (CC) as an instrumental variable (IV) for changes in local traffic conditions to recover the WTP to avoid traffic. This charge forces drivers to internalize the negative traffic externalities imposed on others by taxing them when they drive into the zone during charge operating hours. Put differently, I am comparing the changes in traffic volume and house prices before and after the CC is implemented. I observe that the CC substantially improves traffic conditions and home prices after it is introduced. On average, traffic is around 8.5% lower in the charge zone after the CC is implemented and home buyers pay approximately 3.6% more. Putting these estimates together, the IV estimates suggest that the direct elasticity between traffic and house prices is around -0.43. These estimates are considerably larger compared to early OLS estimates, suggesting that previous studies have underestimated the cost of traffic. These results remain stable and robust even when I limit the analysis to sales around 500 metres from the charge perimeter, illustrating that the effects are not driven by unobserved neighbourhood differences across the boundary.

Additional analyses reveal that the charge attributes to a displacement of traffic across the charge boundary as drivers detour the zone to avoid the CC. Substantial capitalization gains imply that the CC has created a windfall for homeowners in the zone at the expense of poorer households living outside the zone, suggesting that the CC is regressive. Hence, to ensure that the policy is equitable, policy makers could consider removing the 90% discount for residents in the charge zone, taxing these capitalization gains for home-owners in the zone, and redistribute these revenues on enhancing and subsidizing the public transport system. Increasing the reliability and quality of the public transit could further improve the efficacy of the charge.

The next question I am interested in:

"Can speed cameras reduce accidents and save lives?"

Every year, based on estimates by the World Health Organization, approximately 50 million individuals are involved in traffic accidents, with 1.2 million eventually succumbing to these injuries around the world. To improve road safety, various laws and instruments have been introduced to ensure that drivers do not drive recklessly. These measures include texting bans, speed limits, drinking and seat belt laws with patrolling traffic police enforcing these regulations. Since the 1990s, the U.K government have been installing fixed speed cameras to penalize speeding. Chosen sites are often accident "black" spots - roads that have a high number of traffic collisions, injuries and deaths, and a sizeable percentage of drivers exceeding speed limits. These cameras measure commuting speed and penalize drivers when they exceed the stipulated limit. Over the next two decades, more than 3,000 cameras are introduced on the road network.

There are many other reasons why the efficacy of speed cameras is of policy interest. First, speed cameras are under the scrutiny of the public because of the huge amount of fines they managed to rake. Many oppose vehemently to these instruments, firmly believing that speed cameras are installed to generate revenues and that alternative instruments (such as speed limit signs) have the same desired impact. Second, there are concerns whether these devices cause more accidents as drivers unaware of the location of cameras could become a road hazard by abruptly dropping speed to avoid punishment. Finally, several areas in UK are forced to switch off their cameras due to budget cuts. It will be paramount to understand whether switching these devices off actually makes roads unsafe.

Conducting the analysis on a sample of speed cameras across England, Scotland and Wales, I document that speed cameras reduce both the number and severity of collisions. After a speed camera is installed, the number of accidents and minor injuries are 20% and 19% lower than pre-installation levels, corresponding to 1.13 fewer accidents and 1.33 fewer injuries per kilometre. As for seriousness of the crashes, the number of fatalities and serious injuries are 58% and 31% lower, which amount to 0.20 and 0.33 per kilometre. Further analyses suggest that the effects speed cameras are greater along roads with higher speed limits and are highly localised to within 500 metres from the camera. Putting these estimates into perspective, installing another 1,000 speed cameras reduce 1130 collisions, mitigate 330 serious injuries, and save 190 lives annually, generating net benefits of around £59 million. These welfare estimates are computed after considering a wide array of benefits and costs associated these devices.

In my final paper, I ask whether:

"Do judges discriminate against unattractive felons?"

The preferential treatment towards the attractive has been documented widely in multiple contexts that include labor markets, politics and financial markets. Research has shown that "beautiful" people have better labour market outcomes (Hamermesh & Biddle, 1994; Biddle & Hamermesh, 1998), attractive politicians are more likely to win more votes (Berggren *et al.*, 2010), and good-looking borrowers can secure loans more easily (Duarte *et al.*, 2012). At the same time, there is a burgeoning literature that has shown that judges allow extraneous factors to influence their decision making. Unimportant factors, such as outcome of football games (Eren & Mocan, 2016), duration from food breaks (Danziger *et al.*, 2011), media attention(Lim *et al.*, 2015), defendant's gender (Mustard, 2001) and race (Abrams *et al.*, 2012; Alesina & La Ferrara, 2014; Park, 2017), have been documented to sway judicial rulings. Hence, it would not be surprising to observe if judges indeed *"judge the book by its cover"*.

Yet, the literature has been remarkably limited on the impact of physiognomy on judicial outcomes due to empirical challenges. Firstly, given that "*Beauty is in the eyes of the beholder"*, different people might have different ideas on what is attractive. Thus, it is challenging to objectively define what is beautiful. Previous studies rely on multiple respondents to rate the same subjects to obtain an impartial measure of attractiveness. This resource-intensive research set up, however, meant that previous studies are limited to small sample sizes, questioning the external validity of these studies. Secondly, unobserved factors could correlate with facial attractiveness and influence sentencing outcomes. For instance, drug abuse could affect physiognomy and could push drug abusers to recidivate or commit more crimes that result in tougher punishments, biasing the effects between attractiveness and sentencing outcomes.

Bearing these challenges in mind, I apply facial recognition algorithms on more than 200,000 mugshots of convicted felons in Florida from 1998 to 2015 to locate various facial features (e.g eyes, nose, face-line, forehead etc.). Using these landmarks, I am able to compute facial symmetry as a measure of attractiveness. I further rely on the unusually rich set of information on inmates characteristics and case-related facts to mitigate the risk of omitted variables from biasing my results. The main finding is that judges hand out preferential sentences to felons with more symmetrical facial features. The disparity in punishment between criminals with more symmetric faces, at the 25th percentile, and criminals with less symmetric faces, at the 75th percentile, is around 1.0% to 1.9% of the mean sentence length, which amounts to between 17 and 32 days. Additional analyses reveal that this bias against felons with less proportional faces could vary across race, gender and type of crimes.

Concluding remarks

Although each of the essay is addressing a very different question, the unifying aim of the thesis is to perform rigorous empirical analysis on important questions using quality data to inform policy-making. This is challenging because researchers are required to not only obtain causal effects of the policy, but must also consider a wide spectrum of outcomes that the policy potentially affects to inform on social welfare.

To ensure that my estimates are causal, I emphasize on how to construct counterfactuals to ensure that they are similar to those that being treated. This requires a good understanding on how subjects are selected into treatment. I also pay close attention to potential spillovers across treatment and control groups that could bias the estimated effects of policy.

To inform whether policies improve welfare, I further consider a wide array of effects associated with the policy. An insight from these analyses is that policies often have unintended consequences, benefiting certain groups at the expense of others, or achieving policy objectives by incurring some costs. For instance, the Congestion Charge reduces traffic in the charge zone by displacing traffic to neighbourhoods outside while speed cameras improve road safety by reducing commuting speed and increasing travel time. Failure to account for these unintended effects could overstate welfare estimates. This is an area this thesis aims to improve on to better inform policy making.

Chapter 2

The Cost of Traffic: Evidence from the London Congestion Charge

2.1 Introduction

Traffic congestion is an urban disamenity from the agglomeration of economic activities. Attracted by productivity gains and amenities in cities, firms and individuals congregate in urban areas and compete for space, attributing to outward expansion of cities. With the proliferation of auto-mobiles, individuals are encouraged to drive and this surge in auto-mobiles on roads inevitably leads to traffic congestion, an ubiquitous problem many cities around the world faces. These traffic delays affected London as well. Average on-road commuting speed in the 1990s was slower than that at the beginning of twentieth century before car travel became prevalent (Newbery, 1990). By 2002, travel speed for motor vehicles during morning peak hours fell by almost 30% compared to that in 1974, from 14.2 to 10.0 miles per hour, and drivers spent, on average, 27.6% of their on-road time stationary (Department of Environment & the Regions, 1998).

Traffic is also a major source of air pollution. According to figures from Environmental Protection Agency, auto-mobiles contribute to more than 50% of the nitrogen oxide, 30% of the volatile organic compounds and 20% of the PM10 in US[1]. These emissions have detrimental effects on health outcomes, increasing infant mortality, reducing birth weight and inducing premature births (Currie & Walker, 2011; Knittel *et al.*, 2016). Heavier traffic can also increase traffic accidents and fatalities (Li *et al.*, 2012; Green *et al.*, 2016). Bottlenecks can also affect economic growth (Boarnet, 1997; Fernald, 1999; Graham, 2007), increase unemployment (Hymel, 2009) and reduce wages (De Borger, 2009). It is evident that traffic is undesirable and can affect the attractiveness of neighbourhoods, influencing household location decisions.

This paper measures the average marginal willingness to pay (MWTP) to avoid negative traffic externalities (e.g noise pollution, traffic exhaust, elevated traffic accident risk and congestion delays) in and around the location of residence using the

[1] For more information, refer to https://www.epa.gov/air-pollution-transportation/smog-soot-and-local-air-pollution

housing market[2]. Because an explicit market for traffic does not exist, the hedonic price method is broadly adopted in the literature to value non-market amenities[3]. The idea is that traffic varies across space and, holding all other factors constant, differences in home values should reflect the price paid to avoid traffic. While the concept is simple, attempts to estimate the casual effect of traffic on home prices have been fraught with difficulties. First, traffic is not randomly distributed across space and the heaviest traffic is usually around the city center where economic activities are congregated. Unobserved neighbourhood differences between these properties across space are likely to confound the estimates. Further, more affluent households who incur costlier time delays have incentives to sort themselves into the city center to reduce the need to commute. The concern is whether the WTP to avoid traffic could be confounded with the WTP for better neighbourhoods.

Bearing these challenges in mind, this paper exploits the substantial but localised changes in traffic conditions induced by the London Congestion Charge[4] (CC) to recover the cost of traffic. The charge boundary is drawn around the city centre to alleviate congestion from the most gridlocked roads in London. A flat fee of £5 is imposed for driving into the cordoned area during weekdays from 7:00am to 6:30pm, excluding public holidays. This Pigouvian tax equates the marginal private and social cost of transport to ensure that drivers incorporate congestion externalities into their private cost of travel (Pigou, 1924; Vickrey, 1963). The effects were immediate. Six months into implementation, the volume of cars into Central London fell by 27% and average travel speed was 20% higher than before (TfL, 2003a).

Estimation is based on a quasi-experimental instrumental variable (IV) approach. I exploit the introduction of the Congestion Charge as an instrumental variable for changes in local traffic conditions in hedonic house price regressions. Put differently, I am utilizing the sharp variation in traffic conditions in and around the charge zone and comparing the changes in traffic volume and house prices before and after the CC is implemented to recover the WTP to avoid traffic. To obtain consistent estimates for the MWTP to avoid traffic, several conditions must be satisfied. Other than the fact that the charge must significantly affect traffic flow, it is imperative that the mean differences in unobservables (e.g neighbourhood amenities, housing

[2]For convenience, this will be referred as the willingness to pay for less traffic or the cost of traffic across this paper.

[3]The hedonic approach has been used extensively in the literature to value non-market amenities since it is formalized by Rosen (1974). Some examples include school quality (Black, 1999; Bayer *et al.*, 2007; Gibbons *et al.*, 2013), air quality (Chay *et al.*, 2005), health hazards (Gayer *et al.*, 2000; Davis, 2004; Currie *et al.*, 2015), crime (Thaler, 1978; Gibbons, 2004) and transportation accessibility (Gibbons & Machin, 2005).

[4]Other cities that managed to introduce the CC include Singapore, Dubai, Milan, Stockholm, Gothenburg and Durham.

characteristics) between transactions across the CC boundary are not correlated with the implementation of the charge and the charge influences house prices only via traffic (also known as exclusionary restriction).

To attenuate unobservable differences between property sales, I partial out any time-invariant housing and neighbourhood characteristics by including postcode fixed effects. This is equivalent to comparing changes in sale prices and traffic conditions before and after the charge is implemented within a postcode. A postcode represents a building usually and there are approximately 17 sales in a postcode. In addition, I control for an extensive set of property and location characteristics surrounding each sale to reduce the risk of observable differences from confounding my estimates. Furthermore, I progressively limit the analysis to transactions that close to the charge zone (up to 500 metres in and out of the cordoned area) to mitigate unobserved neighbourhood differences across the CC boundary. This is only possible because Central London is densely built with many residential sales over the sample period and the CC generates sharp changes in traffic conditions across the CC boundary. By doing so, I am comparing properties sharing common amenities (e.g school quality, parks, crime rate) and neighbourhood demographics (e.g unemployment rate), but enjoying contrasting traffic conditions due to the CC.

I further identify several possibilities that the exclusionary restriction could be violated. First, home purchasers could be paying more to re-locate into the cordoned area because residents in the zone enjoy a 90% discount to the CC. Exploiting a sub-sample of sales outside the charge zone but are entitled to the CC discount, I show that this discount to the charge has a negligible effect on home prices. Second, affluent neighbours, who incur higher cost of being caught in the traffic, could sort themselves into the charge area. Relying on micro level census data in and around the charge zone, I demonstrate that there are no evidences of sorting of "better" neighbours into the CCZ/WEZ that could confound the WTP to avoid traffic.

The headline finding is that homeowners moving into the cordoned charge zone pay more to enjoy better traffic conditions. After the CCZ is implemented, I observe that traffic volume declined by about 8.5% (1,779 fewer vehicles every day) relative to neighbourhoods outside the cordoned area, illustrating the efficacy of the CC in reducing traffic. Corresponding to this improvement in traffic conditions, home prices are approximately 3.6% (£40,968) higher in the zone. Putting these results together, the instrumental variable (IV) estimates suggest that the elasticity of housing values with respect to traffic volume is around -0.43. These estimates of the average MWTP to avoid traffic are robust across a range of sensitivity analyses and when I constrain the analysis to sales just in and out of the CCZ/WEZ. I also observe that the MWTP to avoid traffic are much higher for residents moving into the WEZ. This could be because they incur higher cost of delay, are more likely to drive, and live further away from their workplace. Additional analyses reveal that this WTP for less traffic could stem from better air quality and reduction in traffic collisions in the charge zone after the CC is enforced.

These estimates imply that the implementation of the CC that has generated substantial windfall for homeowners relative to neighbours outside the zone. Multiplying the capitalization gains with the number of dwellings in the cordoned area, the aggregate increase in housing values in the zone CCZ and WEZ amount to £3.11billion and £11.91 billion respectively. These substantial gains measure the present value of the local benefits associated with the CC and is approximately 14% of the cost of implementing the charge. This is tenable considering the myriad of benefits with the CC that are not quantified in this study. These results, however, also suggest that the CC is regressive as it benefits richer homeowners inside the zone at the expense of poorer households living outside. Hence, to ensure that the CC is more equitable, policy makers could consider creaming off this windfall via taxes and remove the charge discounts given to residents in the zone. Revenues can be redistributed via public transport subsidies and investment on improving on public transport system. Finally, the elasticity of housing values with respect to traffic obtained from this study could be useful in estimating the potential welfare gains or losses associated with transportation infrastructures (e.g roads, congestion charges, public transit) before embarking on these projects given how cost intensive some of them are.

The remainder of this paper is structured as follows. Section 2 provides an overview on the Congestion Charge in London. Section 3 describes the existing literature on this subject. Section 4 outlines the data and Section 5 illustrates the identification strategy. Findings are then discussed in Section 6 and Section 7 concludes.

2.2 Road Pricing in London

The initial Congestion Charge Zone (CCZ[5]) covered a total of 21 square kilometres (slightly more than 1% of the Greater London Area) and encompassed the financial centre (Bank), parliament and government offices (Palace of Westminster), major shopping belts (Oxford Circus) and tourist attractions (Trafalgar Square, Westminster Abbey, Big Ben, St Paul Cathedral etc). Figure 2.1 shows the CCZ, the area shaded in green. The boundary was drawn to isolate the most congested areas in Central London and does not appear to be constrained by any physical features (rail lines, green spaces and rivers etc). It was bordered by major Inner Ring

[5]The initial Congestion Charge Zone will be abbreviated as the CCZ while the Western Extension Zone will be abbreviated as WEZ from this point onwards

Roads such as Edgeware, Vauxhall Bridge, Pentonville, Park Lane, Marylebone, Tower Bridge and Victoria to divert traffic displaced by the charge. Commuters travelling on these roads are not required to pay unless they turn into the zone. To protect residents and businesses outside the zone, off-street parking enforcement is improved to deter anyone from parking outside and walking into the charge zone to avoid paying the charge. The CCZ crosses the River Thames to the South and covers parts of the Lambeth and Southwark boroughs. Although this is an area not typically considered as Central London, it was incorporated for the ease of implementation and operation (Richards, 2006).

On the 17th of February 2003, a flat fee[6] of £5.00 was levied on commuters driving into the zone between 7:00am to 6:30pm from Monday to Friday, excluding public holidays. Residents living in the zone and some living outside but in discount zones are entitled to a 90% waiver[7] to the CC for their first registered vehicle. These discount zones are shaded in purple as shown in Figure 2.1. Residents residing in these areas are entitled to the discount because they are required to enter the CCZ or WEZ when driving home[8]. This policy was an outcome of extensive consultations with various stakeholders. Other than to reduce congestion, the CC is implemented to generate revenues to increase the frequencies and routes of buses and tube to enhance the public transit. Reduced travel time and enhanced reliability could encourage commuters to switch from private to public transport when commuting into the zone.

The tax levied was substantially increased to £8.00 on the 4th July 2005 to further reduce traffic and raise revenues. On the 19th of February 2007, charging was extended to Central West London (known as the Western Extension Zone - WEZ) because of congestion in that area. Operating hours of the CC were reduced by half an hour from 7:00am to 6:00pm. The westward extension is circumvented by Harrow Road, Scrubs Lane, West Cross Route, the Earls Court One-Way system, Chelsea Embankment and the River Thames [9] to the South. Refer to the area shaded in pink

[6]The rationale for levying a flat fee, other than the difficulty in imposing time varying fees to reduce congestion during peak hours, is that vehicular volume on roads seem fairly uniform across the day.

[7]Other groups excluded from the charge include public transport(taxis and buses), motorcycles, bicycles, environmentally friendly vehicles (battery powered or hybrid cars), vehicles driven by disabled individuals (blue badge holders), vehicles with 9 seaters or more and emergency service vehicles.

[8]This is a concern as home purchasers moving into the CCZ or WEZ could be paying more for homes for the CC discounts, violating exclusionary restriction. I will show in results later in Table 2.1 that home buyers are not paying more for these CC discounts by exploiting a unique part of the CC policy that permits home owners living outside by near to the CCZ/WEZ a 90% waiver of the charge.

[9]Unlike the Original CCZ, the WEZ is bounded by physical features. There is a concern whether the the neighbourhoods South of River Thames are different from those in the

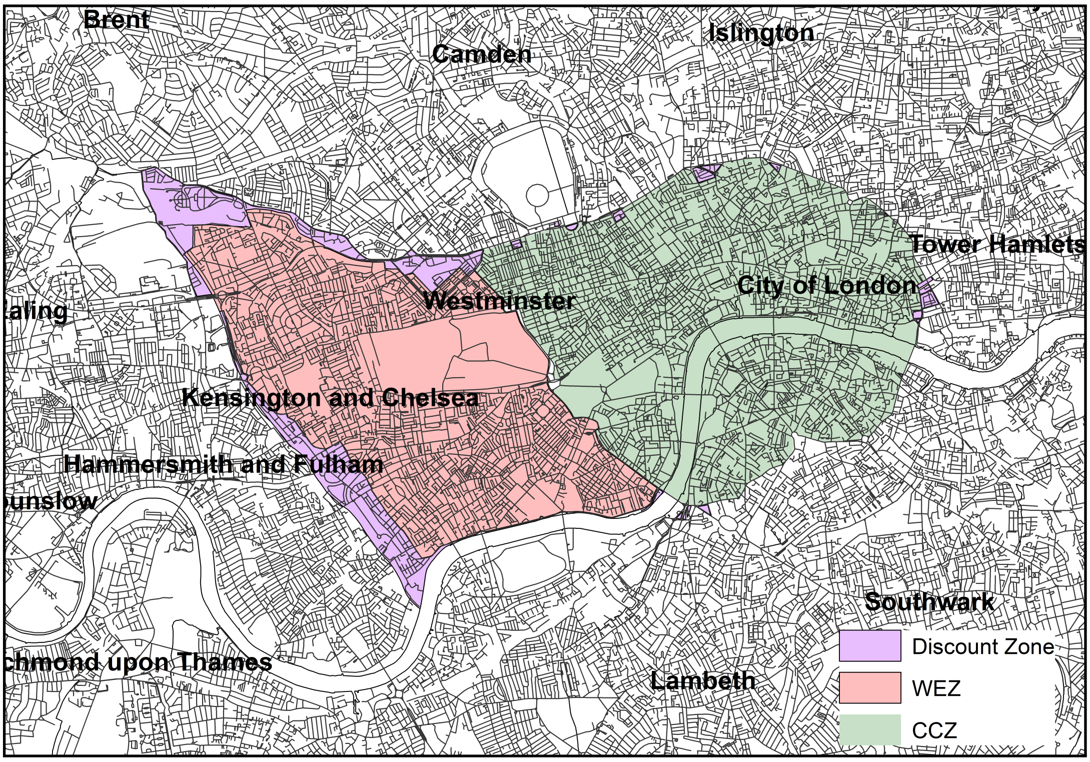


Figure 2.1: Map of the Original Congestion Charge Zone (CCZ) & Western Extension Zone (WEZ)

in Figure 2.1. However, under tremendous pressure from residents and businesses in West London, on the 24th of December 2010, the WEZ was scrapped. Between 2011 to 2015, the charge in the original CCZ underwent another two hikes. The CC was raised from £8 to £10 on the 4th January 2011 and from £10 to £11.50 on 16th June 2014. Overall, the CC experienced an average 10.83% growth per annum since introduction and this might have a compelling effect on commuters relying on private transport.

Initial impact assessment by Transport for London (TfL) showed significant improvement in traffic conditions after the charge is enforced in 2003. These results are very consistent with those reported in this study. All day travel speeds were almost 20% higher (from 14.3km to 16.7km per hour) and minutes of delay fell by 30% compared to uncongested traffic conditions (TfL, 2003a). This was largely due to a 27% overall drop in the number of private auto-mobiles into Central London. A change in composition of inbound traffic into the zone was observed: the volume of bicycles, buses and taxis went up by 28%, 21% and 22% respectively. Surveys conducted echoed similar findings with the majority of the drivers switching to pub-

North such that it might not be a suitable control group. Hence, I exclude transactions south of River Thames in my robustness test (refer to Table2.8). This has an immaterial effect on my estimates.

lic transport and others travelling during off-charging hours (TfL, 2005). Though the number of commuters using rail did not increase, the number of bus passengers during morning peak periods were 38% higher(TfL, 2004). There was no apparent displacement of traffic into neighbouring uncharged roads and weekends as traffic conditions were fairly similar compared to those during pre-charged periods. As for air quality, the CC led to a 12% reduction in both NO and PM_{10} in the cordoned area (TfL, 2004). Overall, evidences suggest that residents living in the charged zone are benefiting from the charge.

2.3 Literature Review

To estimate the marginal willingness to pay (MWTP) to avoid traffic, the hedonic property value approach is widely adopted in the existing literature. An association between traffic externalities, measured by traffic volume (Hughes & Sirmans, 1992) or noise (Palmquist, 1992; Andersson *et al.*, 2010), and housing prices are established using regression adjusted for differences in observable housing and neighbourhood characteristics. A review of the previous literature indicates that the doubling road traffic volume could reduce home values by 0.5%-3.0%, while every decibel increase in traffic noise corresponds to a 0.3%-0.6% reduction in transacted home prices. Estimates, however, appear to vary across studies that adopt different specifications and perverse relationships are sometimes reported. These results suggest that cross-sectional estimates could be biased due to unobserved differences in neighbourhood and housing quality between sales that are correlated with traffic conditions.

Several studies address the issue of omitted confounders by focusing on "natural experiments" that produces a shock to the amenity of interest[10]. Chay *et al.* (2005) rely on the implementation of the Clean Air Act in the 1970s to identify exogenous variation in air quality and examine its impact on housing prices. Davis (2004) take advantage of a sharp rise in paediatric leukaemia cases from a secluded county in Nevada to measure the health risk using home values. Gibbons & Machin (2005) appraises the price for better public transport accessibility by examining the impact of a new metro line on the housing market. Black (1999) and Gibbons *et al.* (2013) quantify the value of good schools by comparing sale prices of homes proximate to one another but on different school districts. In similar fashion, this study relies on the implementation of the CC that induces sharp variation in traffic conditions across the CC boundary to recover the MWTP to pay to avoid traffic.

Previous literature shows that the Congestion Charge reduce traffic jams and improve air quality. Beevers & Carslaw (2005) show that air quality inside the

[10]For the advantages associated with quasi-experimental approaches of hedonic methods for environmental valuation refer to Kuminoff *et al.* (2010).

charge zone improved after the CC is implemented. The levels of CO_2, NO and $PM10$ fell by 19.5%, 12% and 11.9% respectively. Similar results are echoed by Green *et al.* (2018) although they show that the concentration of NO_2 has increased after the CC is enforced. They propose that this could stem from the substitution of diesel-based vehicles, such as buses and taxis, in the zone as they are waived from the charge. This combustion of diesel contributes to more nitrogen oxides than the combustion of petrol.

Roads in the zone are also reported to be much safer after the CC is implemented. Li *et al.* (2012) reveal that car casualties fell by 5.2% although there are more fatalities associated with motorcycles (1.8%) and bicycles (13.5%). This could be driven by the switch to two wheelers that are not subjected to the charge. Larger effects are observed by Green *et al.* (2016). The CC coincides with a 32%-36% fall in accidents and 25%-35% decline in serious injuries and fatalities and no displacement of collisions to neighbouring areas outside the cordoned area are documented.

There have been several previous attempts to quantify the benefits associated with the charge using the housing market. Most of these studies have surprisingly documented insignificant or negative effects. The closest to this study is unpublished research conducted by Zhang & Shing (2006). They examine the effect of the CCZ in 2003 on a sample of residential sales in London from 2000 Q1 to 2006 Q1 and show that home prices are 8.5% lower in the zone after the charge is implemented. Percoco (2014) investigate the effect of the Milan EcoPass on housing prices. Examining average property values across 192 Micro-zones between 2006 and 2009, he reports that prices fell by 1.2% to 1.8% after the tax is introduced. Given that the CC improves local traffic conditions, it is surprising to observe that house prices are lower within the charge perimeter after the CC is implemented. The contradictory relationship documented in these studies could stem from omitted confounders due to the lack of controls, the incorporating of transactions fairly far from the charge boundary and the adoption of coarse spatial fixed effects[11]. Agarwal *et al.* (2015) improve the estimation by removing time-invariant neighbourhood unobservables with postcode fixed effects. They examine the effects of an increase in the Singapore Electronic Road Pricing (approximately £0.50) on retail, office and residential prices. While retail property values are adversely affected by the hike, residential property values remain unchanged. This is anticipated considering that an immaterial hike in the charge is unlikely to significantly improve traffic conditions to influence housing values[12].

[11] In unreported parsimonious specifications without micro-level fixed effects, I observe results that are fairly similar to these studies. Results are available upon request.

[12] This point is reinforced by my results in Table 2.10 summarized in Data Appendix. Most of the CC increments do not have perceptible effects on traffic and housing values.

In contrast, this research improves on the existing literature on several fronts. This is the first paper that links the effects of the CC on house prices via traffic using an instrumental variable framework. This is an important "first stage" that explains the mechanism for house price changes associated with the CC that is missing in the existing literature due to the absence of quality traffic flow data. Second, by relying on the CC as a natural experiment to tackle the issue of omitted confounders, this research is a significant improvement to the existing literature that rely on cross-sectional hedonic regressions. Third, this study draws inferences from a representative sample of more than 80,000 property sales from almost 10,000 postcodes in the vicinity of the CCZ/WEZ. This further allows the restriction of property sales physically close to the charge boundary to mitigate unobserved neighbourhood differences between properties in and out the charge zone.

2.4 Data

Average annual daily traffic flow (AADF)[13] collected at each count point (CP) from 2000 to 2014 is retrieved from Department of Transport (DfT). These count points are located along roads and traffic is manually counted at these locations to provide junction-to-junction traffic flow. There are a total of 2,774 CPs in London, most of them clustered around Central London as shown in Figure 2.2. To accurately measure the local traffic conditions for each transacted property, I first match the count points and roads based on location and road names . Subsequently, I draw 100 meter buffers[14] from this sample of matched-roads. The traffic conditions for each property will be determined by the traffic flow from the nearest road. Properties outside this 100 meters buffer will be omitted from the analysis as I could not reliably measure traffic conditions. For an illustration, refer to Figure 3.2.

Housing transactions from the 1st quarter of 2000 to the 4th quarter of 2015 are collected from Land Registry database. Property characteristics include sale price, property type (detached, semi-detached, terraced, flat or maisonette), tenure (leasehold or freehold) and whether the property is new or second-hand. Land Registry covers all the transactions made in United Kingdom. Given that terrace and flat housing constitute bulk of the transactions in Central London (close to 95%),

[13]Each site is counted by a trained enumerator on a *neutral day* in that year for a twelve hour period. A *neutral day* is a weekday between March and October, excluding all public holidays and school holidays. The idea is that traffic on these days are reflective of an "average" day across the year. There are a total of 10,000 manual count points across UK.

[14]Concerned that 100 meters buffer might be too big to accurately measure local traffic conditions, I reduce this buffer to 50 meters. I further re-weight my estimates, giving heavier weights to transactions that are closer to the roads with traffic data. None of these specifications appear to materially influence the results and are summarized in Table 2.8.

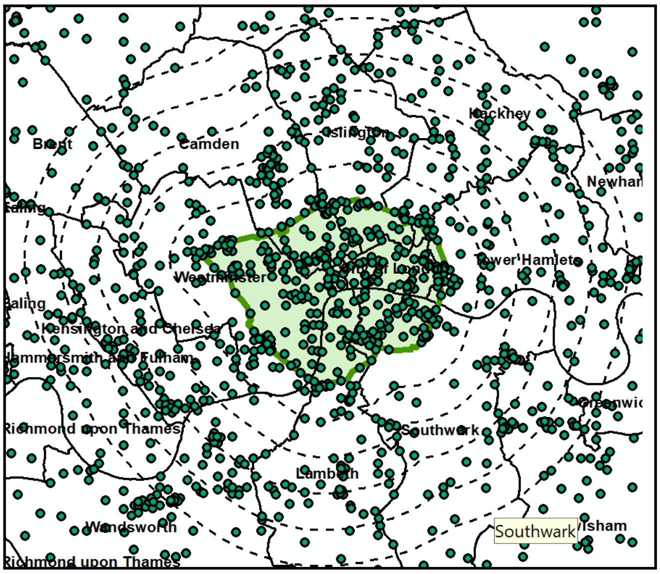


Figure 2.2: Illustration of count points 5km from the CCZ

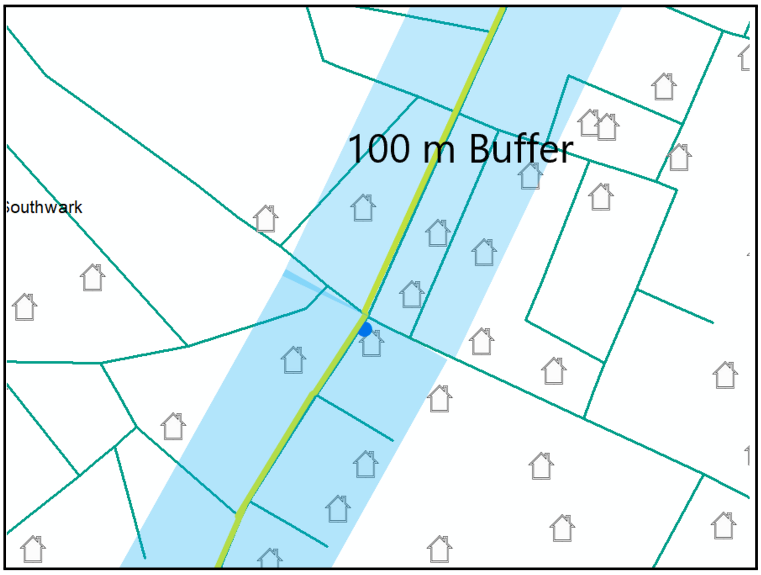


Figure 2.3: Illustration on how traffic conditions are measured for each property.

other property types are removed from the analysis to reduce heterogeneity in the sample that could raise endogeneity concerns. All the transactions are geo-coded using the address postcode. For a subset of transactions, more property information, such as floor area, number of bathrooms and bedrooms and age, are merged from Nationwide transaction database for balancing tests.

Information on the boundaries of the CCZ and WEZ and the areas entitled to 90% resident discount are from the shape-files provided by Transport of London (TfL). Using Geographic Information Systems (GIS) mapping, together with the official dates of implementation/announcement of the CC from TfL, I assign postcodes into treatment and control groups and compute nearest euclidean distance from the CC boundary. Further information on the locations of tube stations and bus stops are retrieved from TfL Open data source. I measure public transport accessibility based on the distance of each postcode from the nearest public transport node using GIS.

Census Data at Output Area[15] (OA) level are collected from 2001 and 2011 to measure the quality of neighbourhoods. This include the percentage of (1) minority residents and (2) uneducated residents, (3) unemployment rate and the percentage of (4) lone parent households. I assign the data from Census 2001 for any transactions before 2006 and data from Census 2011 for transactions made after 2006.

Shape files detailing the location of heritage buildings and parks are provided by MAGIC [16]. Using GIS, I measure the distance of each postcode from the nearest Grade 1 park - with international and historical significance. I further draw a 200 meter buffer around each postcode and compute the number of Grade 1 heritage buildings within these buffers. Designation is done by Historic England and is determined by the age, historical and architecture significance of the building. Only the top 2.5% of the buildings are classified as Grade 1. Maps for Thames River is obtained from Digimap. A buffer of 200 meters is drawn from Thames River and postcodes inside this area are assumed to have a river view.

2.5 Identification Strategy

2.5.1 Research Methodology

Traditionally, hedonic regressions estimating the effects of traffic on house prices adopt the following specification:

$$Y_{ijt} = \beta^{\mathbf{OLS}}\mathbf{T_{ijt}} + X'_{ij}\phi + V'_{jt}\omega + \tau_t + \varepsilon_{ijt}, \qquad \varepsilon_{ijt} = \alpha_i + \theta_{ijt} \tag{2.1}$$

where Y_{ijt} is the logarithm of price for property i in neighbourhood j sold at time t and $\mathbf{T_{it}}$ is the logarithm of local traffic conditions measured by local traffic volume near property i at time t. The key variable of interest, $\beta^{\mathbf{OLS}}$, denotes the percentage change in home prices from 1% change in local traffic flow. This exercise exploits the variation of traffic conditions and home prices across space and over time. To minimise salient differences between housing transactions, researchers usually control for observable property specific X'_i (e.g number of bedrooms, property size, garage) and

[15] The smallest geographical area if which Census data is collected. There are a total of 175,434 OAs across England and Wales (25,053 OAs in London) with around 110 to 140 households per OA.

[16] For more information, refer to http://magic.defra.gov.uk/.

neighbourhood characteristics V'_{jt} (e.g crime, unemployment rates). For consistent estimation, the least square estimator of $\beta^{\mathbf{OLS}}$ requires $E[\varepsilon_{ijt}, T_{ijt}] = 0$.

In reality, however, this assumption is likely to be violated if there are omitted time invariant (α_i) or time-variant unobservable (θ_{ijt}) that could covary with traffic conditions and influence home prices. The heaviest traffic are usually found in neighbourhoods around the Central Business District and they are quite different from areas further out. Properties near to the city center are usually better connected to transportation nodes and are closer to major shopping belts and business districts. If these differences are unaccounted for and enter into the specification, it is likely to underestimate the WTP to avoid traffic. The straightforward solution widely used in the literature is to include property fixed effects (α_i) to partial out these time-invariant unobservables. Put differently, I am now comparing changes in home prices with changes in traffic conditions over time.

There are major issues employing this strategy. First, it requires repeated transactions of the same property that is unlikely given the illiquid nature of real estate due to high transaction costs. Second, it is improbable to observe much variation of traffic in a particular location over time unless these areas experience major new developments that generate economic activities and attract more road traffic. The concern is whether these shocks also make neighbourhoods more attractive and influence local home prices. As a result, traffic conditions are likely to covary with unobserved time-variant shocks to house prices (θ_{ijt}) such that $E[\theta_{ijt}, T_{ijt}] \neq 0$.

Hence, to overcome these challenges, I instrument local traffic conditions ($\mathbf{T_{it}}$) using the London Congestion Charge (LCC). In other words, I am now exploiting the sharp variation in traffic conditions induced by the LCC to measure the cost of traffic. The system of equations to be estimated includes:

$$T_{ijkt} = \lambda_k + \gamma\mathbf{CC_{it}} + X'_{jt}\rho + V'_{jt}\kappa + \psi_t + \nu_{ijkt}, \tag{2.2}$$

$$Y_{ijkt} = \pi_k + \zeta\mathbf{CC_{it}} + X'_{jt}\delta + V'_{jt}\eta + \upsilon_t + \epsilon_{ijkt}, \tag{2.3}$$

$$Y_{ijkt} = \alpha_k^{IV} + \beta^{\mathbf{IV}}\widehat{\mathbf{T_{ijkt}}} + X'_{jt}\phi^{IV} + V'_{jt}\omega^{IV} + \tau_t^{IV} + \varepsilon_{ijkt}, \tag{2.4}$$

where CC_{it} is an indicator variable that takes the value of 1 when property i is located in the charge zone and sold after the LCC is implemented. I utilize both the implementation of the Congestion Charge Zone in 2003 (CCZ) and the Western Extension Zone (WEZ) in 2007 as instruments in separate regressions. I constrain the analysis to a sample of transactions two years before and after the charge is implemented to ensure that the various charge events do not overlap with one another. Refer to the time-line in Figure 2.4 for more information. I further examine the effects of the various charge increments in 2005, 2011 and 2014, and the removal of the WEZ in 2011 on traffic and house prices. Due to space constraints, I relegate these findings to the appendix. In short, these events do not affect traffic conditions and home prices. For more details, refer to Table 2.10 in Data Appendix. Given the lack of repeated sales[17] of the same unit over the sample period, postcode fixed effects ($\alpha_k; \lambda_k; \pi_k$) are included instead. There are, on average, 17 units sharing one postcode across United Kingdom and they are usually properties in the same building.

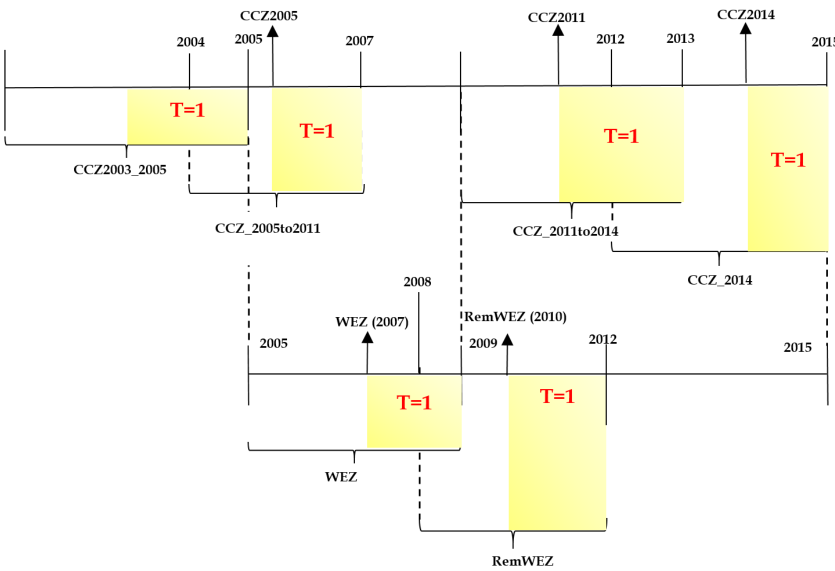


Figure 2.4: Sample window for the different CC events (T=1 denotes Treatment Period)

Equation 2.2 is the *first stage* regression that estimates the effectiveness of the CCZ/WEZ in reducing local traffic flow surrounding each property. The dependent variable, T_{ijkt}, is the natural logarithm of the average daily road traffic flow from vehicles with four or more wheels. The efficacy of the charge is captured by γ that measures the percentage change in the traffic flow in the charge zone after the CCZ/WEZ is implemented. Equation 2.3 is the *reduced form* regression that measures the impact of the CCZ/WEZ on home prices. ζ captures the percentage change in house prices in the charge zone after the CCZ/WEZ is introduced. If the implementation of the CC reduces traffic flow within the charge perimeter, and that new home buyers moving into the zone value this improvement in traffic conditions, I expect γ to be <0 and ζ to be >0.

These regressions combine to form the *instrumental variable regression* in equation 2.4 that identifies the causal effect of traffic on home prices. The main results of this paper come from the estimation of β_{IV}, which measures the direct elasticity

[17]Including address fixed effects significantly reduces the sample by more than 70% as there are limited repeated sales of the same property.

of traffic and house prices. $\widehat{T_{ijkt}}$ denotes the traffic conditions instrumented with CC_{it}. Since β_{IV}, is exactly identified, it is simply the ratio of the two reduced form parameters such that $\beta_{IV} = \frac{\zeta}{\gamma}$. For the instrumental variable estimator to provide a consistent estimator of the hedonic price schedule gradient, the conditions are that conditional on property and neighbourhood characteristics, and postcode fixed effects:

- CC_{it} affects local traffic conditions $[\gamma \neq 0]$ (**Relevance**)
- CC_{it} is as good as randomly assigned. (**Independence**)
- CC_{it} influences home prices only through changes in traffic conditions. (**exclusionary restriction**)

2.5.2 Testing the Identification Assumptions

In this section, I will highlight instances that could violate identification assumptions and address them to ensure that the instrumental variable regression framework is able to consistently estimate the MWTP to avoid traffic.

While it is straightforward to show the test the relevance from F-statistics from first-stage regressions (equation 2.2), it is challenging to ensure that the other conditions are not violated. To begin with, it is improbable that the charge zone is drawn exogenously as the policy is targeted towards curtailing traffic along the most congested roads in Central London. This is clearly the case as the charge zone overlaps with the Central Business Districts, major tourist attractions and shopping belts. Therefore, I progressively restrict the analysis to properties physically close to the charge zone, up to 500 meters left and right of the charge boundary. To visualize, refer to Figure 2.5. The assumption now is that the CCZ and WEZ are as good as randomly drawn between analogous neighbourhoods close to one another in and around the charge boundary. This strategy is possible because the charge induces a sharp discontinuous change in traffic around the boundary.

For exclusionary restriction to hold, the charge must only affect home prices only through traffic. There are at least three instances that this condition could be violated. First, the policy allows residents staying in the zone to a 90% waiver of the charge. The concern is whether new residents are paying more for homes in the zone to enjoy the CC discounts. Hence, the capitalization effects, if there are any, could be capturing the present value of these congestion charge savings. To address this concern, I exploit a feature of the CC policy that allows some homeowners close to but outside the zone a 90% discount to the charge. The reasons for extending the discount to these neighbourhoods are due to parking and severance issues (TfL, 2009). Some of the residents living outside the zone might have their designated

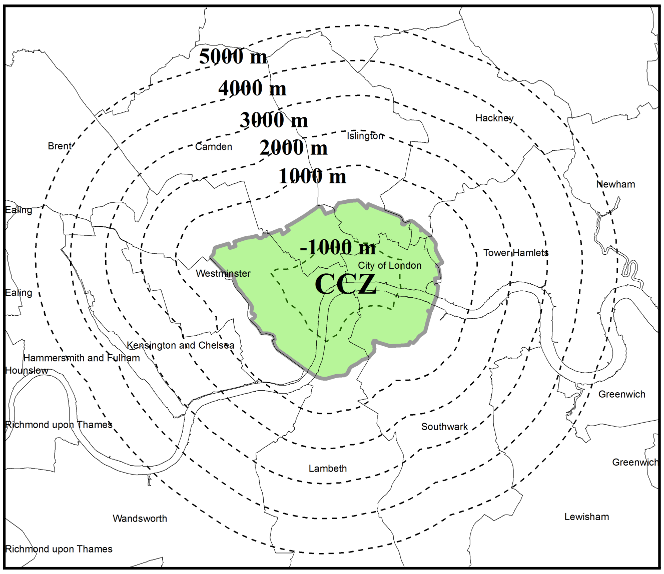


Figure 2.5: The CCZ (shaded) and 1 kilometre buffers from the CC boundary

parking lots inside the zone. Also, the nearest services and amenities (e.g hospitals, libraries etc) for these residents could also be in the charged area. These discount zones are shaded in purple-striped for the WEZ and in grey for the CCZ in Figure 2.1. To examine whether homeowners pay for the 90% waiver of the charge, I estimate the following regression:

$$Y_{ijkt} = \pi_k + \zeta \mathbf{CC_{it}} + \psi \mathbf{Dis_{it}} + X'_{jt}\delta + V'_{jt}\eta + \upsilon_t + \varepsilon_{ijkt}, \tag{2.5}$$

where Dis_{it} that denotes properties in the discount zone that were sold after the implementation of the CCZ/WEZ. The key parameter of interest, ψ, captures the willingness to pay for the 90% CC discount. If new homeowners are paying more to live in the CCZ/WEZ for better traffic conditions and not for the discounts, I expect ψ to be indistinguishable from zero. Looking closer, this regression resembles equation 2.3 other than the inclusion of properties sold in the discount zone.

As there are limited transactions[18] outside CCZ eligible for the CC savings, the focus will be on the discount zone of the WEZ that has a larger sample of 15,976 sales. Panel A of Table 2.1 presents the estimates. Home prices in the discount zone are not materially affected after the WEZ is enforced, suggesting that new home buyers are not paying more to enjoy the 90% discount of the CC. No house price changes are observed even when the discounts are taken away after the WEZ is removed. Taken together, these results suggest that homeonwers do not pay more

[18]In particular, there are only 936 sales in these areas. Considering the fact that home owners living near the WEZ, on average, earn higher income, have a higher tendency to drive and stay further from their work place compared to those bordering the CCZ, I would expect the WTP for the CC discounts to be more magnified for the homeowners in the WEZ discount zone.

for their homes to enjoy the CC discounts.

However, the traffic conditions in the discount zone might be affected by the charge, therefore confounding the WTP for these CC discounts. To verify, I conduct similar regressions but with traffic flow as the dependent variable. I relegate these results to Table 2.14 in Data Appendix due to space constraints. In short, the results show that the enforcement of the WEZ do not affect traffic conditions in these discount zones. Overall, evidence suggests that the CCZ/WEZ is not affecting home prices through the charge discounts.

Table 2.1: Reduced form estimates of the Congestion Charge Discount on House Prices

	(1) 5km	(2) 4km	(3) 3km	(4) 2km	(5) 1km
Discount	0.0173	0.0117	0.0118	0.0104	0.0258
	(0.0207)	(0.0216)	(0.0228)	(0.0257)	(0.0310)
WEZ	0.0579^{a}	0.0534^{a}	0.0500^{a}	0.0477^{a}	0.0458^{b}
	(0.0141)	(0.0147)	(0.0155)	(0.0170)	(0.0195)
Obs	55849	47117	37603	28033	17997
R2	0.80	0.79	0.78	0.77	0.74
No.of Postcodes	9021	7503	5868	4329	2665
Discount	-0.0097	-0.0095	-0.0088	0.0046	0.0294
	(0.0213)	(0.0222)	(0.0236)	(0.0258)	(0.0298)
RemWEZ	0.0219	0.0211	0.0182	0.0242	0.0248
	(0.0148)	(0.0157)	(0.0170)	(0.0184)	(0.0209)
Obs	68415	57710	46289	33702	21236
R2	0.81	0.81	0.80	0.79	0.76
No.of Postcodes	9698	8034	6316	4653	2874

Each coefficient is from a different regression. Sample is constrained to sales within 5km (Column 1) to 1km (Column 5). *Discount* is a binary variable equals to one for sales made inside the discount zone after the WEZ is introduced. Dependent variable is the natural logarithm of the transacted property prices. Robust standard errors clustered at output area are reported in parenthesis.^{c} $p<0.10$, ^{b} $p<0.05$, ^{a} $p<0.01$

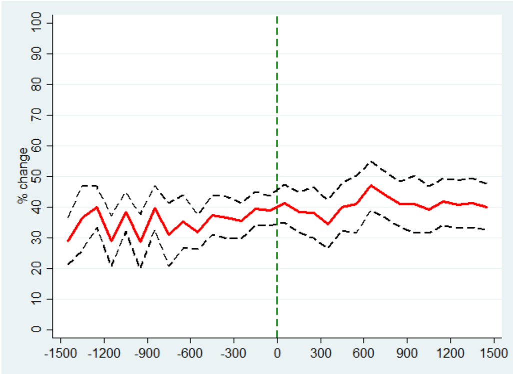
Secondly, there could be sorting of better households into the charge zone after the charge is implemented. If affluent homeowners, who incur higher congestion delays due to higher wages, are incentivised to move into the zone after the CC is introduced, the issue is whether the WTP to stay in the CCZ/WEZ could be confounded with the WTP to reside in better neighbourhoods, violating the exclusionary restriction. To investigate, I examine the changes in various neighbourhood characteristics across the boundary before and after the charge is implemented in Figure 2.6. This include percentage change (% Δ) of (1) residents who are ethnic minorities, (2) unemployment rate, (3) residents with no education, (4) lone-parent households, (5) households with cars and (6) residents driving to work. These figures are constructed by taking the long differences of neighbourhood characteristics from Census Data collected at Output Area[19] in 2001 (before) and 2011 (after), before regressing these changes on the interaction of the CCZ dummy with distance to boundary fixed effects. Each point in the figure, which is the coefficient of the respective distance dummies (in 100 meters bandwidth), denotes the conditional average change in neighbourhood characteristics at a given distance from the CC boundary. Negative distances, to the left of the dashed line, indicate neighbourhoods in the CCZ. As shown, there are no sharp changes across various demographics, driving habits and car ownership in and around the CC, suggesting that there is no sorting across the boundary after the charge is enforced. Due to space constraints, a similar set of figures for the WEZ is moved to Figure 2.10 in the Data Appendix. These results are fairly congruent to that reported for the CCZ.

Finally, by replacing address with postcode fixed effects, I am assuming that there are no changes in quality/characteristics for units sold in the same postcode (or building) after the charge is implemented. Exclusionary restriction could be violated if the WTP for the CCZ/WEZ are driven by the quality differences of the units sold in a postcode after the CC is enforced. This is possible if more affluent households move into the charged zone after the CC is implemented such that better units (e.g penthouses) in the same building are sold after the charge is enforced.

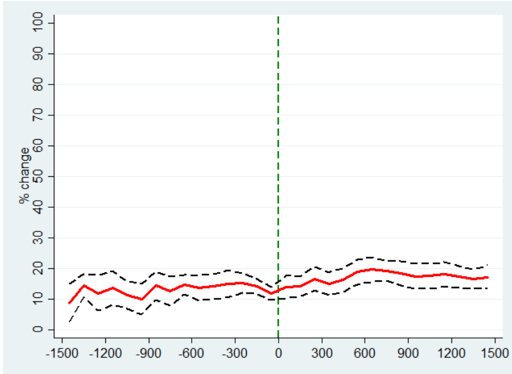
To address this concern, I conduct a battery of balancing tests on various observable housing characteristics. Results are summarized in Table 2.2. The specification is similar to that in Equation 2.3 but the dependent variable is replaced with housing characteristics, including flat dummy, leasehold dummy, floor area, availability of central heating and garage, number of bedrooms and bath, and the age of unit. Columns 1 to 2 summarize results from a larger sample from the Land Registry, while columns 3 to 8 entail findings from a sub-sample of residential sales from the Nationwide sales database with a richer set of housing characteristics. The analysis incorporates transactions within 3 kilometres of the CC boundary. As observed, there are no significant changes in the composition of transactions within a postcode before and after the introduction of the CC, mitigating the risk that estimates are driven by the change in quality of housing units[20].

[19]Output Area is the lowest geographical level at which census estimates are provided in UK. There are a total of 175,434 Output Areas in England and Wales.

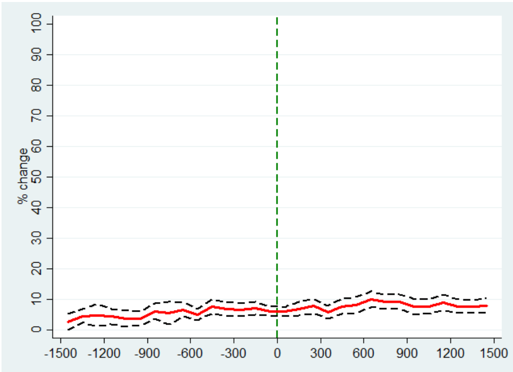
[20]I also estimated equation 2.3 with these hedonic characteristics as controls for the sample of transactions from Nationwide Database. The results are very similar to that reported in Table 2.5. However, due to the small sample size (less than 1,000 observations), I do not report the findings in this paper although it is available upon request.



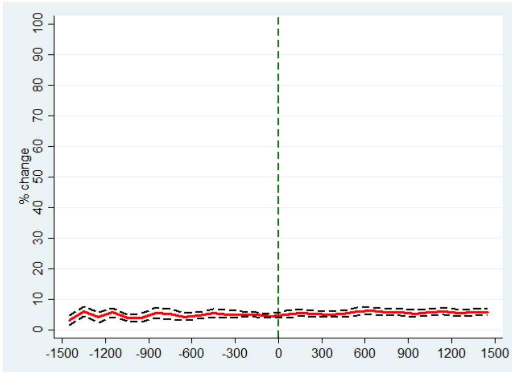
(a) Minority ethnicity



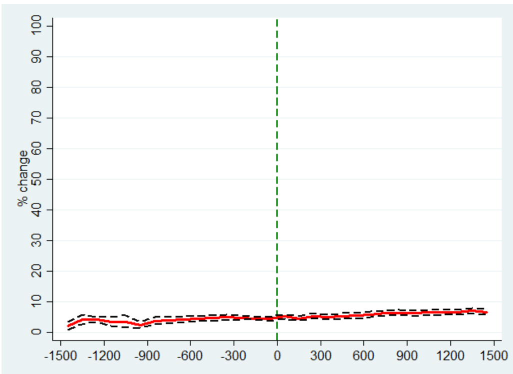
(d) Uneducated



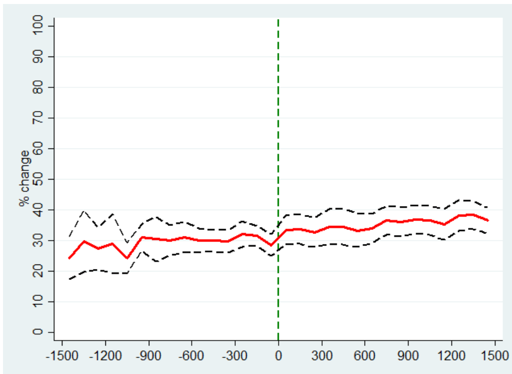
(b) Lone parent households



(e) Unemployment Rate



(c) Driving to work



(f) Car ownership

Figure 2.6: Census demographics around the CCZ. The solid line represents the conditional average change of various demographics at a given distance from the CC boundary and the dashed line represents the 95% confidence interval. It is constructed by regressing the % Δ in demographics at Census Output Area with boundary fixed effect and 100 meters distance bandwidths and coefficient of each distance dummy is plotted. Distance is negative when it is in the charged zone (Left of dashed Line). There are a total of 1,727 output areas within 1.5 kilometres in and out of the CCZ.

2.6 Empirical Results

In this section, I estimate the effects of the Congestion Charge on traffic and house prices. First, I describe the dataset with summary statistics. Next, I examine

Table 2.2: Balancing Test for Housing Characteristics for a subsample of transactions within 3km from the CC boundary

	(1) Flat	(2) Leasehold	(3) Floor Area	(4) Bathrooms	(5) Bedrooms	(6) Central Heat	(7) Garage	(8) Age
CCZ	-0.00498	-0.00361	-2.349	0.0338	0.00295	-0.0122	-0.113	-3.041
	(0.00352)	(0.00348)	(3.699)	(0.125)	(0.0709)	(0.0933)	(0.212)	(6.795)
N	110719	110719	5288	5288	5288	5288	5288	5288
WEZ	0.00188	0.00686	-5.041	-0.102	0.0288	-0.238	-0.381	-1.628
	(0.00533)	(0.00461)	(11.22)	(0.116)	(0.193)	(0.251)	(0.349)	(4.779)
N	62952	62952	3283	3283	3283	3283	3283	3283

Each coefficient is from a different regression. All regressions include post code and year quarter fixed effects. Dependent variable is the respective housing characteristics as labelled. Flat (1) is a binary variable indicating whether property sold is a flat. Leasehold (2) is a binary variable representing whether unit sold is leasehold. Floor area (3) is the size of unit in square meters. Bathrooms (4) and Bedrooms (5) is the count of Baths and Bedrooms in the unit. Central heating (6) and Garage (7) is a binary variable that denotes if unit has such facilities. Age (8) is the number of years since the unit is built. Columns 1 & 2 comprise of transactions from Land Registry while Columns 3 to 8 comprise of sales from Nationwide Database. Robust standard errors clustered at output area are reported in parenthesis.

the impact of the CCZ and WEZ on both traffic and home prices before combining the estimates to recover the MWTP to avoid traffic. Subsequently, I constrain the analyses to properties up to 500 metres from the CC boundary to minimize unobserved neighbourhood differences between sales across the CC boundary. Finally, I show that the results remain robust to a battery of tests that relaxes the identification assumptions, before discussing the policy implications associated with the findings.

There are two main results from the analysis. First, the implementation of the CCZ and the WEZ improve traffic conditions in the cordoned area relative to neighbourhoods outside. The effects are notably weaker associated with the WEZ, raising questions of its suitability as an instrument. New homeowners moving into the charge zone appear to pay more after the charge is implemented. Second, naive OLS specifications produce inconsistent estimates while instrumental variable estimates are much larger, more robust and stable, suggesting that previous studies have underestimated the willingness to pay to avoid traffic.

2.6.1 Descriptive Statistics

Table 4.2 reports summary statistics for the estimation sample of sales within 5 kilometres of the CCZ (Panel A) and WEZ (Panel B). I further breakdown the sample into inside and outside the charge zone. There are a total of 239,909 sales from 27,430 unique postcodes within 5 kilometres from the CCZ. The sample is slightly smaller for the WEZ. There are 136,375 transactions from 20,686 different postcodes within 5 kilometres from the WEZ. The sheer number of sales illustrates how densely built Central London is. Approximately 33% and 45% of the sales took place within the CCZ and the WEZ after the charge is implemented.

Table 2.3: Descriptive Statistics for Estimation Sample for the CCZ & WEZ

	Inside CCZ		Outside CCZ		Inside WEZ		Outside WEZ	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Sale Price	377424.44	354914.12	306406.18	253006.11	909496.17	1086279.62	394369.52	344095.99
Traffic Volume	18804.53	13546.04	19244.95	15326.21	17358.16	15279.52	16504.42	15327.01
CCZ/WEZ Treatment	0.33	0.47	0.00	0.00	0.45	0.50	0.00	0.00
New build	0.22	0.41	0.11	0.31	0.04	0.18	0.11	0.31
Flat/Mansionette	0.95	0.23	0.73	0.44	0.83	0.38	0.74	0.44
Terraced house	0.05	0.21	0.21	0.41	0.16	0.36	0.20	0.40
Leasehold	0.95	0.22	0.74	0.44	0.83	0.37	0.75	0.44
Dist to Park	1115.82	747.54	2419.50	1309.43	692.40	369.66	2451.62	1296.88
Heritage buildings (200m)	0.91	1.65	0.05	0.39	0.14	0.45	0.05	0.37
Thames River View	0.07	0.25	0.09	0.29	0.06	0.23	0.08	0.27
% with no education	12.99	9.84	16.42	10.58	8.16	6.64	13.07	9.26
Unemployment Rate	4.19	2.53	4.74	2.42	3.61	2.28	4.62	2.52
% of Lone Parent Households	3.99	4.38	6.01	4.88	3.24	3.39	6.27	5.15
% of Minority Race	25.41	12.90	27.69	15.66	25.36	13.10	31.99	16.81
Sample Size		239909				136375		
No. of Postcodes		27430				20686		

Sample of sales are 5km or less from the CCZ/WEZ boundary. Inside (Outside) CCZ/WEZ are property sales/roads within (outside) the charge zones.

The first row shows that the average transacted prices are much higher in the charged zone. The mean sale price in the CCZ (£377,424) is more than £70,000 higher than sales outside the zone (£306,406). This is expected given that properties in the CCZ are more desirable as they are better connected to transportation nodes, major shopping belts and the CBD. This disparity in transacted prices is even greater in the WEZ. The average house prices are more than twice inside the WEZ (£909,496) relative to outside (£394,370). The stark divergence in sale prices highlights the presence of "super-rich" neighbourhoods in the WEZ. The second row indicates that traffic conditions are about the same in and out of the CCZ and WEZ across the sample period. While the average daily traffic flow is just slightly higher in the WEZ (17,358) compared to outside (16,504), traffic is lighter in the CCZ (18,804) compared to neighbouring areas outside the zone (19,245). Furthermore, across the boundary, properties in the charge zone are more likely to be leasehold multi-family flats. They are also located closer to parks and are more likely to be surrounded by buildings with heritage value.

Surrounding neighbourhood characteristics, residents inside the CCZ and WEZ are more educated and are less likely to be unemployed. Also, households inside the zone are less likely to be single parent households with smaller minority race representation. Although there are differences in observable housing and neighbourhood characteristics, these disparities across the CC boundary will be minimized once I exploit (1) variation of property prices and traffic within postcodes and over time and (2) limit the analysis to sales and roads very close to the CC boundary.

To illustrate how neighbourhoods and properties are more similar when I limit the analysis to just in and out of the charge zone, I tabulate the difference-in-means of the average observable neighbourhood and housing characteristics between sales from 900 to 500 metres from the CC boundary. Results are summarized in

Panel A of Table 2.4 for the CCZ and Panel B for the WEZ. One can observe the convergence in house prices, neighbourhood and housing characteristics once as soon as I limit the analysis to sales closer to the charge zone. Conversely, the difference in traffic conditions exacerbates when I constrain the comparison to areas proximate to the charge perimeter. These results further emphasize the presence of traffic displacement across the boundary after the charge is implemented. Because Central London is so densely built, even when I constrain the analysis to 500 metres from the CCZ boundary, my sample is still fairly representative with 28,850 transactions from 3,344 postcodes. The sample size is slightly smaller for the WEZ with 18,571 sales from 1,839 postcodes. The sizeable number of transactions mitigates the concern that my study is drawing inferences from an unrepresentative sample of sales around the charge boundary.

To further show how the CC influences house prices and traffic conditions around the charge perimeter, I plot the conditional changes of house prices and traffic at every 100 metre from the CCZ boundary after the CC is implemented. These estimates are constructed by regressing traffic volume and house prices against postcode and year-month fixed effects, observable housing and neighbourhood covariates, 100 metre distance bandwidth dummies interacted with an indicator variable that takes the value of 1 for observations after the CC is implemented, and the interaction of these dummies with an indicator variable whether this observations are found in the CCZ, before plotting the coefficient of these distance dummies[21]. These results, analogous to those found in Figure 2.6, are summarized in Figure 2.7. Due to space constraints, similar results for the WEZ are reported in Figure 2.11 in Data Appendix. The general observation is that traffic volume is lower while home prices increase in the charge zone after the CCZ is implemented. These magnitude of the effects differs across space but the direction is largely consistent. Substantial traffic displacement is observed across the CC boundary as areas closest to the border experience a 12% increase in traffic flow.

[21] In other words, coefficients for the distance bandwidths inside the charge zone will be taken from the CC_{it}*Distance-bandwidth dummies (See equation 2.3), while those outside will be taken from $Post_{it}$*Distance Bandwidth dummies where $Post_{it}$ is equal to one for observations after the CCZ is enforced.

Table 2.4: Differences in means for observable characteristics for properties in the CCZ (Panel A) & WEZ (Panel B)

	Panel A: CCZ					Panel B: WEZ				
	900m	800m	700m	600m	500m	900m	800m	700m	600m	500m
Sale Price	-54901.77^a	-48719.58^a	-46328.17^a	-43872.35^a	-43548.01^a	-375587.84^a	-356979.71^a	-328126.25^a	-294339.25^a	-271906.35^a
	(2661.96)	(2777.44)	(2899.17)	(3172.89)	(3455.90)	(8493.15)	(9056.37)	(9462.83)	(9536.82)	(10522.86)
Traffic Volume	4572.68^a	5520.30^a	5901.43^a	6342.89^a	6202.50^a	1010.08^a	646.00^a	579.81^a	742.36^a	1940.41^a
	(134.67)	(142.03)	(150.00)	(161.11)	(175.59)	(170.76)	(184.70)	(191.64)	(195.10)	(209.17)
New build	0.01^a	0.02^a	0.03^a	0.07^a	0.08^a	0.05^a	0.04^a	0.05^a	0.05^a	0.05^a
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Flat/Mansionette	-0.06^a	-0.06^a	-0.05^a	-0.03^a	-0.02^a	0.04^a	0.04^a	0.05^a	0.04^a	0.04^a
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Terraced house	0.06^a	0.06^a	0.05^a	0.03^a	0.02^a	-0.04^a	-0.04^a	-0.05^a	-0.04^a	-0.04^a
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Leasehold	-0.06^a	-0.06^a	-0.05^a	-0.03^a	-0.02^a	0.03^a	0.03^a	0.04^a	0.04^a	0.04^a
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Heritage buildings (200m)	-0.40^a	-0.34^a	-0.30^a	-0.25^a	-0.26^a	-0.07^a	-0.09^a	-0.08^a	-0.07^a	-0.07^a
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Thames River View	0.13^a	0.16^a	0.17^a	0.20^a	0.20^a	0.08^a	0.07^a	0.08^a	0.09^a	0.10^a
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
% of Minority Race	5.77^a	5.31^a	5.59^a	4.73^a	3.09^a	8.20^a	8.04^a	8.06^a	7.52^a	7.61^a
	(0.14)	(0.15)	(0.15)	(0.16)	(0.16)	(0.16)	(0.17)	(0.19)	(0.20)	(0.23)
% with no education	4.39^a	3.72^a	3.53^a	2.30^a	1.12^a	3.18^a	3.15^a	2.85^a	2.65^a	2.58^a
	(0.11)	(0.11)	(0.11)	(0.12)	(0.12)	(0.09)	(0.09)	(0.10)	(0.11)	(0.12)
% of Lone Parent Households	1.46^a	1.14^a	1.22^a	0.85^a	0.54^a	2.77^a	2.67^a	2.74^a	2.74^a	2.65^a
	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)
Unemployment Rate (%)	0.94^a	0.84^a	0.80^a	0.73^a	0.72^a	1.07^a	1.12^a	1.09^a	1.09^a	1.18^a
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
No.of Sales(Total)	45483	41591	38106	33304	28850	31828	28232	24966	21747	18571
No.of Sales(Treated)	22468	21067	19465	17097	14869	16983	15693	14298	12731	11064
Postcodes(Total)	5214	4786	4365	3900	3344	4943	4363	3818	3278	2773
Postcodes(Treated)	2777	2630	2458	2264	2006	2917	2663	2414	2137	1839

Coefficients reported are the differences in means for the various observable characteristics for properties inside and outside within distance bandwidth (900 to 500m). Standard errors are reported in parenthesis. ^{c} $p<0.10$, ^{b} $p<0.05$, ^{a} $p<0.01$.

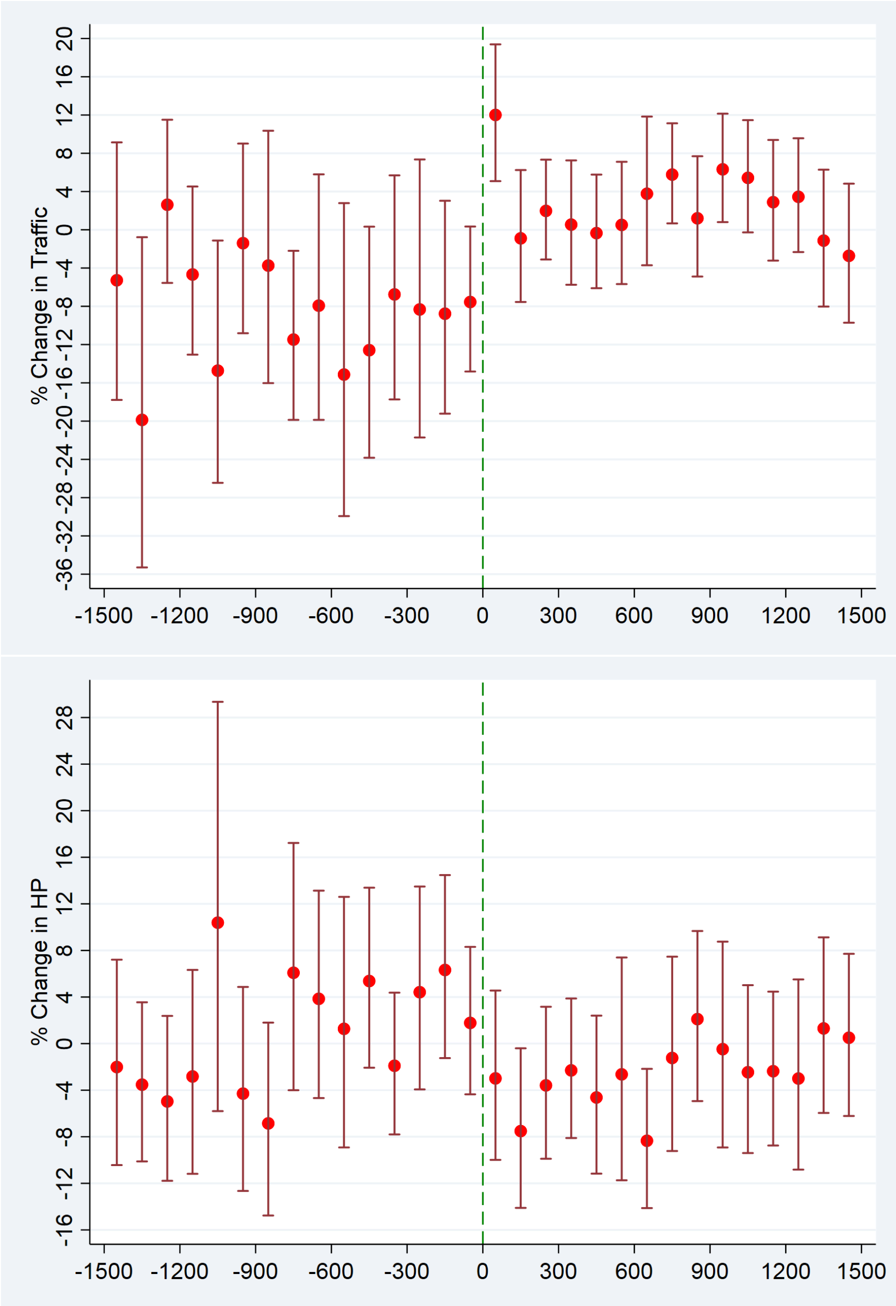


Figure 2.7: The effects of the CCZ on Traffic (Top) and House Prices (Bottom) across distance around the CC boundary. Distance is negative when it in the charged zone (Left of dashed line). The plotted coefficients denote the localised conditional changes of traffic/house prices at a given distance from the CC boundary. Tails marks the 95% confidence interval. They are constructed by regressing traffic volume or house prices against postcode and year-month fixed effects, observable housing and neighbourhood covariates and 100 metres distance-bandwidth dummies interacted with $Post_{it}$ and distance-bandwidth dummies interacted with CC_{it}. $Post_{it}$ is equal to one for observations after the CCZ is enforced, while CC_{it} takes the value of one for observations in the charge zone after the CCZ is enforced. This figure plots the coefficients associated with these distance dummies in and out the charged zone.

2.6.2 Effects of the London Congestion Charge on Traffic and Home Prices

Table 2.5: Estimates on the Impact of the CCZ/WEZ on Traffic & House Prices

	(1) 5km	(2) 4km	(3) 3km	(4) 2km	(5) 1km
Panel A - First Stage (Log Traffic)					
CCZ	-0.0523^{b}	-0.0581^{a}	-0.0711^{a}	-0.0751^{a}	-0.0816^{a}
	(0.0206)	(0.0209)	(0.0208)	(0.0206)	(0.0223)
R2	0.98	0.98	0.98	0.97	0.98
Δ Traffic	-1194	-1324	-1609	-1696	-1779
1st Stage F-statistics	6.42	7.76	11.66	13.34	13.38
WEZ	-0.0397^{a}	-0.0401^{a}	-0.0303^{a}	-0.0267^{a}	-0.0172^{c}
	(0.0075)	(0.0080)	(0.0081)	(0.0080)	(0.0100)
R2	0.99	0.99	0.99	0.99	0.99
Δ Traffic	-832	-839	-637	-562	-366
1st Stage F-statistics	28.06	24.94	14.06	11.00	2.99
Panel B - Reduced Form (Log House Price)					
CCZ	0.0305^{b}	0.0271^{b}	0.0245^{c}	0.0259^{c}	0.0349^{b}
	(0.0130)	(0.0132)	(0.0136)	(0.0137)	(0.0150)
R2	0.77	0.76	0.76	0.75	0.73
Δ HP	36320	32259	29089	30757	40968
Obs	85106	72001	54149	37433	23504
No.of Postcodes	9861	8329	6360	4258	2574
WEZ	0.0695^{a}	0.0679^{a}	0.0636^{a}	0.0658^{a}	0.0689^{a}
	(0.0159)	(0.0166)	(0.0174)	(0.0191)	(0.0215)
R2	0.80	0.80	0.79	0.78	0.75
Δ HP	109227	106635	99589	103141	106074
Obs	44056	36636	29126	21328	12490
No.of Postcodes	7222	5938	4639	3360	1896

Each coefficient is from a different regression. Sample is constrained to properties within 5 kilometres (Column 1) to 1 kilometre (Column 5) from the CCZ/WEZ boundary. Panel A reports first regression estimates (γ) from equation 2.2 and Panel B reports reduced form estimates (ζ) from equation 2.3 for both the **CCZ** and **WEZ**. Dependent variable is the logarithm of annual average daily traffic volume for vehicles with 4 wheels or more for Panel A and the logarithm of transacted house prices for Panel B. All regressions are estimated with postcode and year quarter fixed effects. Other control variables include housing characteristics (leasehold, newbuild and terrace dummies), neighbourhood characteristics by-year (% of residents with no education qualifications, % of residents with minority races, unemployment rate and % of lone parent households) and location characteristics by-year (Thames river view dummy, counts of heritage buildings within 200m, distance of the property from nearest park and from the CCZ/WEZ boundary). For more information on the variables, refer to Table 4.6 in Data Appendix. Δ Traffic is the absolute reduction in average daily traffic volume and Δ HP is the absolute effects on house prices converted to 2015 pound(£) value. Robust standard errors (in parenthesis) are clustered at output area. ^{c} $p<0.10$, ^{b} $p<0.05$, ^{a} $p<0.01$.

Panel A of Table 2.5 presents the effects of the CCZ and the WEZ on traffic volume. These estimates illustrate the efficacy of the CC in curbing congestion and the validity of the CC as an instrumental variable for traffic flow. Moving from column (1) to (5), I progressively restrict the sample to areas from 5 kilometres to 1 kilometre left and right of the CC boundary. After the introduction of the CC in

2003 (CCZ), I observe that traffic flow in the zone is 5.37%[22] lower when compared to neighbourhoods outside but within 5 kilometres from the boundary. The charge effects remain fairly stable when I streamline the sample to more comparable neighbourhoods in proximity to the zone. Within 4 kilometres, the effect increases to 5.98% and within 3 kilometres, the effect is 7.37%. These impact further increases to 7.80% when I constrain the sample to areas within 2 kilometres from the charged boundary and is even larger at 8.50% within 1 kilometre from the CC boundary. In absolute terms, I am looking at between 1,194 and 1,779 less vehicles[23] inside the zone everyday compared to areas outside the CCZ.

Magnified effects when constrained to areas near the boundary suggest the presence of traffic displacement across the CC boundary. Evidence implies the charge could have forced drivers to detour the charge area, inducing a surge in traffic for roads close to but outside the CCZ/WEZ. This displacement of traffic, although not an ideal outcome for the CC, induces substantial variation in traffic conditions between proximate neighbourhoods around the charge boundary. This makes the policy an ideal instrument for identifying the MWTP to avoid traffic because it permits the comparison of nearby properties around the charge perimeter to mitigate unobserved neighbourhood differences.

The implementation of the WEZ (WEZ) reduces traffic by 4.04% in the charged zone relative to areas outside but within 5 kilometres from the boundary. Within 4 kilometres, the effect is 4.09% and is 3.07% when I constrain the analysis to neighbourhoods 3 kilometres from the boundary. When I examine areas within 2 kilometres from the charged zone, the impact falls to 2.71%. As soon as I limit the analysis to districts 1 kilometres or less from the WEZ boundary, the impact further drops to 1.72%. In absolute terms, I am observing around 366 to 839 less vehicles every day after the WEZ is enforced. These effects are less than half compared to that reported in the CCZ. The WEZ is no longer a valid instrument when construed to areas 1 kilometre or less from the WEZ boundary, as evidenced by the low 1st stage F-statistics of less than 10.

Panel B of Table 2.5 summarizes the impacts of the CCZ and WEZ on transacted property values. Similar to before, I restrict the analysis to a sample of properties that are physically close to the CC boundary to mitigate unobserved heterogeneity in neighbourhood amenities between sales in and out the charged zone. I observe significant house price appreciation in the charge zone after the CCZ is introduced. When compared to residential sales within 5 kilometres from the boundary, house

[22] As it is a log-linear model, capitalization effects are computed by taking the exponential of the point estimates before subtracting by one. For instance, $Exp(0.0523) - 1 \approx 5.37\%$. The same conversion is applied for housing prices.

[23] This is obtained by multiplying the point estimates with the average pre-treatment traffic volume.

prices in the charge zone are 3.10% higher. Estimated effect drops to 2.75% relative to houses within 4 kilometres and 2.50% within 3 kilometres. Restricting the analysis to housing units just 2 kilometres in and out the CC boundary, I observe stable price responses at around 2.62%. Finally, looking at sales 1 kilometre or less from the CC boundary, which reduces the sample by almost 80%, I document that property values are 3.55% higher than before. In absolute monetary terms, the CCZ increases housing values in the charge zone by a magnitude of between £29,089 and £40,968[24]. All these estimates are significant at least at 5% level.

House prices in the charge zone also increase after the WEZ (WEZ) is introduced. Capitalization effects are around 7.20% when compared to transactions outside but within 5 kilometres of the boundary. Within 4 kilometres, the effect falls to around 7.02% and within 3 kilometres, impact further decreases to 6.57%. Restricting to housing units just 2 kilometres from the boundary increases price response marginally to 6.80%. Comparing units not more than 1 kilometre in and out the CC boundary, which cuts the sample size by about 75%, I observe that the house price appreciation is around 7.13%. All of the estimates are significant at least at 10% level. In monetary terms, homeowners are paying between £99,589 and £109,227 to enjoy better traffic in the WEZ. These absolute effects are much larger as home prices are, on average, much higher in Central West London.

Overall, these results indicate that the implementation of the CCZ and WEZ resulted in substantial improvement in traffic conditions and property values in the charge area relative to areas outside the zone. These findings confirm that the strength of the CC as an instrument for local traffic conditions.

2.6.3 Regression estimates of Marginal Willingness to Pay to avoid Traffic

Table 2.6 summarizes the estimates of the average MWTP to avoid traffic. Like before, I progressively restrict the sample of property sales from 5 kilometres to 1 kilometre from the CCZ/WEZ boundary moving from columns (1) to (5). In Panel A, I present naive Ordinary Least Square (OLS) estimates (β_{OLS}) from equation 2.1. These estimates are essential because they not only allow us to compare the instrumental variable (IV) estimates with the typical results reported in the literature, but they also illustrate how exploiting the exogenous variation in traffic conditions induced by the CC could improve identification of the WTP to avoid traffic.

[24]This is computed by multiplying the estimates on the pre-treatment average home prices adjusted to 2015 price levels in the cordoned area within the distance bandwidth from the CC boundary.

Table 2.6: OLS & IV estimates of the effect of Traffic on House Prices

	(1) 5km	(2) 4km	(3) 3km	(4) 2km	(5) 1km
Panel A - Naive OLS					
ln(Traffic) - CCZ	-0.0113	-0.0140	-0.0174	-0.0170	-0.0530[c]
	(0.0132)	(0.0144)	(0.0154)	(0.0193)	(0.0302)
R2	0.77	0.76	0.76	0.75	0.73
ln(Traffic) - WEZ	-0.0303	-0.0365	-0.0358	-0.0576	-0.0135
	(0.0312)	(0.0357)	(0.0441)	(0.0582)	(0.0625)
R2	0.80	0.80	0.79	0.78	0.75
Panel B - IV Regressions					
ln(Traffic) - CCZ	-0.5827[c]	-0.4664[c]	-0.3443	-0.3444[c]	-0.4276[b]
	(0.3229)	(0.2738)	(0.2110)	(0.2017)	(0.2035)
Obs	85106	72001	54149	37433	23504
R2	0.76	0.75	0.76	0.75	0.73
No.of Postcodes	9861	8329	6360	4258	2574
1st Stage F-statistics	6.42	7.76	11.66	13.34	13.38
ln(Traffic) - WEZ	-1.7498[a]	-1.6950[a]	-2.0990[a]	-2.4656[b]	-4.0071
	(0.5219)	(0.5361)	(0.8062)	(1.0347)	(2.6389)
Obs	44056	36636	29126	21328	12490
R2	0.77	0.77	0.75	0.73	0.61
No.of Postcodes	7222	5938	4639	3360	1896
1st Stage F-statistics	28.06	24.94	14.06	11.00	2.99

Each coefficient is from a different regression that measures the **direct elasticity between traffic volume and house prices**. Dependent variable is the logarithm of transacted house prices. Panel A reports naive OLS estimates (β_{OLS}) from equation 2.1 and Panel B reports IV estimates (β_{IV}) from equation 2.4. Sample is constrained to properties within 5 kilometres (Column 1) to 1 kilometre (Column 5) from the CCZ/WEZ boundary. For **CCZ**, the instrument is the binary variable that takes the value of 1 for properties in the CCZ that are sold after the charge is implemented on the 17th February 2003. For **WEZ**, the instrument is the binary variable that takes the value of 1 for properties in the WEZ that are sold after the charge is implemented on the 19th February 2007. See notes in previous tables for details on the control variables included. Robust standard errors (in parenthesis) are clustered at output area. [c] $p<0.10$, [b] $p<0.05$, [a] $p<0.01$.

Consistent with the previous literature, these OLS estimates are very small and highly unstable depending on the sample analysed. To interpret, a 1% increase in traffic is associated to a reduction in housing values that ranged between 0.01% to 0.05%. None of these estimates appear to be statistically distinguishable from zero except when I restrict the sample to just 1 kilometre in and out the CCZ boundary. The reported effects are now more than 5 times larger and are statistically significant at 10% level. Smaller effects for regressions incorporating transactions further away from the CC boundary is consistent with the idea that unobserved neighbourhood heterogeneity between properties could induce the underestimation of the cost of traffic. Next, I present OLS estimates for the properties in the WEZ. Likewise, the direct elasticity between traffic and house prices is very small between 0.014 and 0.058, and none of these estimates are statistically significant. Taken together, these results suggest that either home buyers do not care about local traffic conditions or conventional OLS estimates are severely biased by omitted variables.

Panel B of Table 2.6 summarizes IV estimates (β_{IV}) of the MWTP to avoid traffic from Equation 2.4 using either the CCZ or WEZ as instruments. These estimates are simply the ratio of ζ and γ from Table 2.5. Results reveals that a 1% increase in traffic volume corresponds to 0.34% - 0.58% lower transacted housing values. Compared to the traditional OLS estimates, the IV estimates are at least 10 times larger. This is a phenomenon that is congruent with the findings reported by Chay *et al.* (2005) for air quality. Most of the effects are significant at conventional levels except for sales within 3 kilometres in and out of the CCZ. Even so, the estimated effect remain similar in size although it is less precisely estimated. It is also reassuring to observe that the IV estimates are far less sensitive to sample chosen compared to the OLS estimates.

I further report IV estimates from the WEZ and these effects appear 4 to 5 times larger than those reported in the CCZ, suggesting that home owners in the WEZ are more willing to pay to avoid traffic. In particular, a 1% increase in traffic leads to a 1.70% to 2.47% decrease in housing values. This is not surprising given that earlier results from Table 2.5 show that home buyers pay much more for a negligible improvement in traffic from the WEZ. This much larger WTP to avoid traffic commands more attention.

Dwelling deeper into the demographics of home owners in both the CCZ and WEZ[25], I observe that residents in the WEZ are more likely to drive and incur much higher costs being stuck in the traffic. Homeowners living in the WEZ earn (£4,095), on average, much higher wages compared to those living in the CCZ (£3,517). It is also more likely for households in the WEZ (49%) to own a auto-mobile than those living in the CCZ (37%). There is also a higher tendency for those staying in the WEZ (25%) to drive to work when compared to residents in the CCZ (13%). This is probably because homeowners in the CCZ stay closer to their work place. 42% of the residents in the CCZ stay less than 2 kilometres from their workplace, compared to 25% of the residents in the WEZ. All these factors could explain why home owners in the WEZ are more willing to pay to avoid traffic. This disparity in the WTP is consistent with the idea that individuals have heterogeneous preferences on travel time (Small *et al.*, 2005).

Nevertheless, it is paramount to point out that as soon as I constrain the sample to 1 kilometre or less, the validity of the WEZ as an instrument is put into question as evidenced by the weak first-stage F-statistics. Moreover, earlier summary statistics reveal sizeable differences in sale prices between properties inside and outside the WEZ even when I restrict to sales around the charge perimeter (See Table 2.4).

[25] Data is collected from Census 2001 and 2011 and is weighted according to the geographical distribution of transactions analysed in this study.

This further exacerbate the risk that house price effects across the WEZ boundary could be driven by unobserved neighbourhood differences. Hence, from this point onwards, further analyses will rely only the variation in traffic conditions induced by the CCZ. Similar analyses for the WEZ can be found in the Tables 2.12 and 2.13 in Data Appendix.

2.6.4 Estimates Restricted To Proximate Transactions

Table 2.7: First Stage, Reduced form & IV estimates from sample 900m to 500m from the CCZ Boundary

	(1) 900m	(2) 800m	(3) 700m	(4) 600m	(5) 500m
Panel A: First Stage (Log Traffic)					
CCZ	-0.0833[a]	-0.0924[a]	-0.0883[a]	-0.0908[a]	-0.0847[a]
	(0.0229)	(0.0248)	(0.0263)	(0.0268)	(0.0263)
R2	0.98	0.98	0.98	0.98	0.98
Δ Traffic	-1818	-1966	-1879	-1852	-1780
Panel B: Reduced Form (Log House Price)					
CCZ	0.0349[b]	0.0434[a]	0.0373[b]	0.0365[b]	0.0390[b]
	(0.0159)	(0.0167)	(0.0169)	(0.0181)	(0.0188)
R2	0.72	0.72	0.72	0.72	0.72
Δ HP	41213	50203	43425	43003	46066
Panel C: IV Regressions					
ln(Traffic)	-0.4192[b]	-0.4697[b]	-0.4231[b]	-0.4023[c]	-0.4603[c]
	(0.2099)	(0.2049)	(0.2123)	(0.2199)	(0.2530)
Obs	21843	19719	17866	15775	14072
R2	0.72	0.72	0.72	0.72	0.72
No.of Postcodes	2380	2177	1962	1765	1555
1st Stage F-stats	13.22	13.90	11.27	11.45	10.37

Each coefficient is from a different regression. Sample is constrained to properties within 900m (Column 1) to 500m (Column 5) from the CCZ/WEZ boundary. See notes in previous tables for details on the control variables included. Robust standard errors (in parenthesis) are clustered at output area. [c] $p<0.10$, [b] $p<0.05$, [a] $p<0.01$.

Next, I restrict the analysis to transactions that are even closer to the CC boundary from 900 metres in Column (1) to 500 metres in Column (5). This strategy further abates the risk of unobserved neighbourhood differences from driving the results. Results are documented in Table 2.7. Panel A presents first stage estimates. It is comforting to observe effects that are not only comparable in size to earlier results, but are also stable across the various distance bandwidths. Overall, the enforcement of the CCZ attributed to significant reductions in traffic in the zone of between 8.70% and 9.68%, which amount to between 1780 and 1966 fewer vehicles every day. Strong first stage F-statistics across the regressions reinforce the validity of the instrument even for adjacent roads bordering the CCZ.

Panel B reports the reduced form estimates of the CCZ on housing prices. Conforming with earlier findings, I document significant house prices changes that are remarkably robust even when I confine the analysis to sales 500 metres from the CC boundary. Specifically, housing values are approximately 3.55% - 4.43% higher in the charge zone after the CCZ is implemented. In absolute terms, these increments range between £41,232 and £50,203, depending on distance from the CC boundary.

Putting these two estimates together in Panel C, I am not surprise to observe a robust association between traffic volume and house prices. In particular, the direct elasticity between traffic volume and housing values ranges between -0.40 to -0.47, analogous to earlier estimates in Table 2.6. These IV estimates are also appear strikingly stable in size across the board. This is because the magnitude of the house price changes vary with the relative changes in traffic flow across the distance bandwidths, lending support that what I am capturing from the house prices is the WTP to avoid traffic.

Overall, these results confirm that the estimates of the MWTP to avoid traffic are not susceptible to unobserved neighbourhood differences that are attenuated by limiting the analysis to properties just in out of the CCZ/WEZ.

2.6.5 Robustness and Placebo Tests

Table 2.8 summarizes the findings from a battery of robustness and placebo tests that further addresses the challenges that impede identification to provide more assuring evidences. It is shown earlier that estimates restricted to sales very close to the CC boundary (See Table 2.7) are fairly similar to the effects documented for sales within 1 kilometre from the CCZ. Therefore, the rest of the sensitivity analyses are conducted for sales within this distance bandwidth to balance between the representativeness of the findings and the potential bias driven by unobserved neighbourhood heterogeneity across the boundary.

Announcement Effects: In Column (1), I replicate earlier results but with announcement dates. This addresses the concern[26] whether there are any spurious house price or traffic responses to the release of the news for the charge before the actual implementation of the CC. The treatment period is defined as the day the CC event is officially announced by TfL and ends the day before the CC event is implemented[27]. Although I observe traffic is marginally heavier after the CCZ is

[26] Another concern is whether there are negative house price effects that predate the CC implementation such that any effects documented earlier is merely capturing mean reversion of home prices. As observed, this is not a concern as home prices are unaffected by the announcement of the CC.

[27] As hikes are announced only a few months before being enforced, there are insufficient pre-treatment property transactions. Hence, announcement effects are computed only for the initial implementation of the CCZ and WEZ (refer to figure 2.4 in Data Appendix)

announced, homebuyers do not respond to the news. This could be explained by the uncertainty of the residents over the effectiveness of the novel policy to curb congestion. This is consistent with the survey conducted by TfL that indicated that the respondents are unsure about whether the CC can reduce traffic and improve accessibility (TfL, 2003b).

CBD Effects: Another concern is whether the effects on traffic and house prices documented earlier could be associated with changes to the Central Business District (CBD), violating the exclusionary restriction. This could be an issue since the charge zone overlaps with the CBD. There is considerable decentralization of economic activities from the CBD with the emergence of Canary Wharf[28] around the implementation of the CC. This shift in economic activities could reduce the attractiveness of the CBD, leading to a fall in house prices and traffic in the zone unrelated to the CC and thereby confounding the average MWTP to avoid traffic. Although limiting the analysis to sales bordering the CCZ/WEZ could potentially mitigate this problem, to further allay this concern, I create artificial treatment areas by shrinking and expanding the CCZ by 1 kilometre. For the shrank zones, neighbourhoods at 0 to 1 kilometre from the boundary inside the CCZ are denoted as control areas (Shrank Control Area) and neighbourhoods beyond 1 kilometre from the boundary in the cordoned area are denoted as treated areas (Shrank Treatment Area). Conversely, for expanded CC zones, areas between 0 and 1 kilometre outside the actual CC zone are flagged as treated areas (Expanded Treatment Area) while areas between 1 and 2 kilometres outside the actual CC zone are denoted as control units (Expanded Control Area). For an illustration, refer to Figure 2.8. Column (2) and (3) report estimates associated with these shrank and expanded placebo areas. As observed, I do not document any spurious effects on traffic flow and house prices in these artificially created charge zones. This suggest that earlier findings are not confounded by the emergence of other commercial areas around London.

[28]From 1999 to 2005, the employment force in Canary Wharf surged by more than 100% from 40,000 to 87,000. This could be attributed to the development and opening of at least 10 commercial developments, including 8 Canada Street, One Churchill Place etc. For more information, refer to https://www.london.gov.uk/sites/default/files/gla_migrate_files_destination/londons-cbd-jan08.pdf

Table 2.8: Robustness Tests for the CCZ

	(1) Announce	(2) Shrank	(3) Expand	(4) Pcd>=5	(5) North	(6) Transport	(7) Rem Near	(8) 50m Houses	(9) IDW
Panel A: First Stage (Log Traffic)									
CCZ	0.0123^{c}	0.0041	-0.0047	-0.0783^{a}	-0.0736^{a}	-0.0761^{a}	-0.0971^{a}	-0.0945^{a}	-0.0739^{a}
	(0.0073)	(0.0496)	(0.0247)	(0.0239)	(0.0284)	(0.0222)	(0.0273)	(0.0275)	(0.0230)
R2	0.99	0.98	0.97	0.98	0.98	0.98	0.98	0.98	0.98
Δ Traffic	281	111	-117	-1723	-1493	-1663	-2005	-1990	-1618
F-stats	2.82	0.01	0.04	10.77	6.70	11.73	12.62	11.82	10.35
Panel B: Reduced Form (Log House Price)									
CCZ	-0.0037	0.0066	-0.0123	0.0443^{a}	0.0501^{a}	0.0382^{b}	0.0417^{b}	0.0475^{b}	0.0318^{c}
	(0.0194)	(0.0384)	(0.0211)	(0.0159)	(0.0192)	(0.0150)	(0.0172)	(0.0203)	(0.0174)
R2	0.73	0.71	0.76	0.70	0.73	0.73	0.72	0.72	0.74
Δ HP	-4260	8452	-15429	50542	64013	44921	49179	54700	37281
Panel C: IV Regressions									
ln(Traffic)	-0.3011	1.6164	2.6455	-0.5629^{b}	-0.6812^{b}	-0.5020^{b}	-0.4293^{b}	-0.5024^{b}	-0.4300^{c}
	(1.5930)	(22.8552)	(14.5219)	(0.2472)	(0.3441)	(0.2303)	(0.1952)	(0.2398)	(0.2520)
Obs	14283	12323	25391	20268	17315	23504	18612	12826	23504
R2	0.73	0.60	0.57	0.69	0.72	0.72	0.72	0.71	0.73
No.of Postcodes	1905	1335	2957	1403	2022	2574	2067	1402	2574

Each coefficient is from a different regression. Sample is constrained to sales 1 kilometres from the CC boundary unless otherwise stated. In (1), the treatment period (CC_{it}) is defined by the announcement window and begins the day the CCZ is announced officially by the Transport for London (TfL) and ends the day before the CCZ is implemented. In (2) and (3), I create artificial treatment zones by expanding and shrinking the CCZ by 1 kilometre. To visualize, refer to Figure 2.8. In (4), I remove any sales in postcodes with less than 5 repeated transactions over sample period. In (5), I remove any sales south of the Thames River. In (6), I include distance to tube-by-year and number of buslines-by-year fixed effects. In (7), I exclude sales that are 100 meters or less from the charge boundary (both inside and outside the zone). In (8), I remove any transactions that are beyond 50 meters from the nearest roads that I can reliably measure traffic flow. In (9), estimates are weighted inversely according the distance from transacted property from matched road. Robust standard errors (in parenthesis) are clustered at output area. ^{c} $p<0.10$, ^{b} $p<0.05$, ^{a} $p<0.01$.

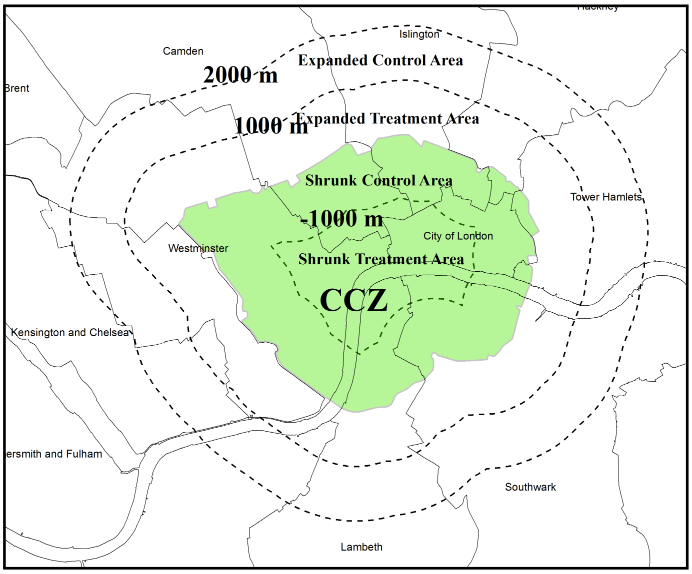


Figure 2.8: The Shrank and Expanded Placebo CCZ

Insufficient Transactions: Another issue is that there could be inadequate repeated observations within some postcode and outliers could be driving the estimates. Thus, I drop any postcodes with less than 5 repeated transactions over the sample period in Column (4). This reduces the number of observations marginally by about 14%. Again, this did not matter as results are similar to those reported earlier.

Physical Barriers: An additional concern is whether the CCZ boundary overlaps with physical constraints (hills, rivers, forest etc.) or major infrastructures (railways, flyovers etc.). If the CC boundary coincides with these features, even restricting to proximate areas on different side of these features might not eliminate unobserved neighbourhood differences. While the CCZ crosses the Thames River due to the ease of charge implementation, the south of the WEZ is bounded by Thames River. The concern that properties to the south of the river are different from those in the north is not unfounded as these areas are typically not considered as part of Central London. Thus, I exclude housing transactions located south of the Thames River from the estimation in Column (5). Doing so has no discernible impact on the estimates although the CCZ is no longer a relevant instrument as reflected by the slightly lower first stage F-statistics.

Removal of Sales closest to the CC boundary: I further remove property sales than are within 100 metres from the CC boundary. The notion is that although restricting to properties closest to the charge boundary can minimize unobserved neighbourhood differences, the spillover effects could be greater as well. For instance, pollutants from traffic emissions could travel across the boundary to neighbouring areas inside the zone. Moreover, properties very close to the CC boundary but inside the CCZ could be near to these congested ring roads circumventing the zone. Home-owners living in these areas, despite being inside the CCZ, might receive considerable negative traffic externalities that cross the boundary, resulting in an underestimation of the WTP to avoid traffic. Results are summarized in Column (6). Removing these roads and property sales close to the CC boundary do not matter much. It appears that although home owners pay more for properties further inside the CCZ, the impact of the charge on traffic is larger as well, resulting in an elasticity that is within the range reported previously.

Public Transport Capitalization Effects: One of the correlated effects associated with the implementation of the CC is the channelling of charge revenues on improving public transport facilities. This could increase the values of homes that are better connected to public transportation nodes and is especially the case for houses outside the cordoned area as driving into the zone is more expensive. To partial out these effects, I add a vector of controls that include: (1) a binary variable denoting whether postcode j is within 200 metres of a tube station and (2) the count of bus lines from bus stops within 200 metres of the postcode. Both are interacted with year dummies as they are time-invariant. As seen in Column (7), upon controlling for these covariates, the effects on housing prices are marginally larger now compared to earlier results from Table 2.5. This is consistent with the idea that house prices outside of the CCZ but close to transportation nodes have appreciated more and accounting for this attribute to larger MWTP to avoid traffic.

Measurement Error: One may also argue that the local traffic for each property could be inaccurately measured by assigning road traffic conditions that are up to 100 metres away from the property (See Figure 2.3). This measurement error could lead to attenuation bias. To mitigate this concern, I adopt several strategies to more reliably quantify local traffic flow. In Column (8), I only incorporate sales that are within 50 metres from the roads that I could accurately measure traffic conditions. Here, I observe more pronounced effects of the CCZ on both traffic and housing prices. Putting them together, a 1% increase in traffic corresponds to a 0.50% fall in home prices, which is congruent to earlier findings. In Column (9), I re-weight the estimates inversely based on the euclidean distance of the property from the nearest road. Put differently, like before, I am placing more emphasis on sales that I can more precisely determine traffic conditions. Again, the impacts on the estimates are modest.

Spurious time effects: Next, I address the concern whether effect of the CCZ on home prices[29] could be documented spuriously during pre-treatment periods. To do so I generate rolling 1-year pre-treatment placebo windows for the CCZ. Placebo

[29] As I only have yearly traffic flow from 2000 onwards, I am unfortunately not able to perform a similar analysis for traffic flow.

treatment period is between t_{false} and $t_{false} + 1year$ and the placebo window is from $t_{false} - 1year$ to $t_{false} + 1year$ where t_{false} represents every quarter from 1996Q1 onwards till 2002Q1. For instance, for 1996Q1, the pre-treatment period is from 1995Q1 to 1995Q4 and the treatment period is from 1996Q1 to 1996Q4. The new key regressor - $CCZ * t_{false}$ - is the interaction of a binary variable of whether the property i in the CCZ is sold during the false treatment period. This falsification test incorporates transactions within 1 kilometre from the CCZ boundary.
Placebo estimates are summarized in Figure 3.8. Each dot represents estimate from a different placebo regression and the tails denote the 95% confidence interval. The dashed line denotes the implementation effects from Column (5) of table 2.5. As observed, none of the placebo estimates, except for 1998Q1, is statistically different from zero and most of the estimates are smaller than the implementation effects. These findings increase the confidence that effects documented earlier are not spuriously reported in non-treatment periods.

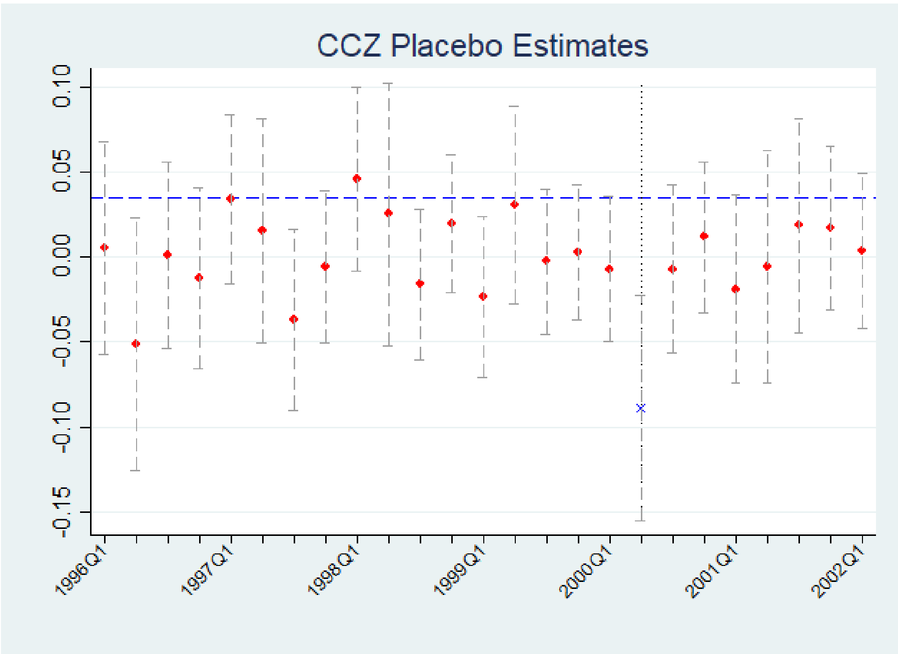


Figure 2.9: The CCZ Placebo Estimates during pre-treatment period. Each point represents a different regression where the treatment period is 1-year rolling window from the corresponding quarter and the pre-treatment period is 1 year before the quarter. The tails represent 95% confidence interval. Cross indicates that the estimate is significant at least at 10%, while dot shows otherwise.

2.6.6 Discussion

In this section, I employ earlier estimates to compute the localised economic benefits associated with the charge. To simplify things, I make the following assumptions: (1) the preferences for traffic are identical across individuals living in the same cordoned area but could differ between the CCZ and WEZ and (2) the relationship between traffic and house price is linear. The implementation of the charge, on average, induces home prices to increase by £33,879 and £104,933[30] in the CCZ and WEZ respectively. Based on the Census estimates on the number of dwellings, which indicates that there are around 91,848 and 113,535 houses in the CCZ and WEZ in 2011 respectively, this implies that the CCZ and WEZ have generated an aggregate windfall of around £3.11 billion and £11.91 billion relative to those outside the zone. This figure is meaningful as it presents monetary measure of the local benefits associated with the charge.

Although these effects seem large at first sight, it is not as it measures the WTP for improvement in traffic conditions in perpetuity given the long-lived nature of real estate. But are they tenable? To answer this, I compare the benefits to the cost of implementing the charge. I did some adjustments to the operating costs of running the London Congestion Charge provided by Leape (2006). Estimating the first year cost to be around £163 million and the subsequent annual operating cost equal to £140 million (£23 million is the set up cost), the present value net cost of implementing the charge for the next 30 years at 2015 £value is around £4.15 billion. This is computed by assuming an inflation rate of 2.7% and a discount rate of 3.0%. The net house price gains, which measures the benefits for home owners in the zone, covers almost 75% of the net cost. This is just about right considering the array[31] of benefits enjoyed by others that are not quantified in this study.

There are policy implications associated with this study. The main reason why many individuals are against the Pigouvian tax is that it is regressive. The huge windfall enjoyed by residents in the zone at the expense of poorer households living outside the cordoned area suggests that this is true. Hence, policy makers should ensure that the charge is more equitable for individuals residing outside the zone. For one, the 90% waiver of the charge that is given to homeowners living in the CCZ/WEZ should be either removed or reduced since they have benefited tremendously from better traffic conditions and higher home values. Furthermore, policy makers could consider implementing a tax to cream off these capitalization gains. Channelling these additional revenues or taxes to enhance the reliability and quality of public transit could further improve the efficacy of the charge and provide a more equitable redistribution of benefits to home-owners living outside.

Finally, I investigate the impact on the LCC on traffic accidents and air quality. As this is out of the purview of this study, I move these results to Table 2.15

---

[30] This is computed by simply taking the average of the capitalization effects across the different distance bandwidths from Table 2.5.

[31] Other benefits that are not localised include the time savings for those living outside the zone and the overall improvement in air quality with less traffic.

and Table 2.16 in Data Appendix. In short, these results show that the CC not only reduces the number of traffic collisions, but also improves air quality in the cordoned area. Specifically, compared to areas outside but within 3 kilometres from the charged zone, the number of accidents and injuries decline by around 5%, while the concentration of PM10 decrease by 4% after the CC is implemented. These effects are weaker associated with air quality as, depending on wind speed and direction, air pollutants could travel across the boundary into the charge zone. This is observed for lighter air pollutants such as NOX and NO2. Another explanation purported by Green *et al.* (2018) for the increase in these nitrogen oxides is the substitution of diesel based vehicles into the cordoned area as taxis and buses are not charged. Combustion of diesel could lead to higher content of nitrogen oxides. Nevertheless, these findings largely support the notion that the cordoned area has become a more pleasant area after the charge is implemented, which could explain why home prices are higher.

2.7 Conclusion

This paper exploits the sharp but localised changes in traffic conditions induced by the London Congestion Charge (LCC) in the Congestion Charge Zone (CCZ) and the Western Extension Zone (WEZ) to estimate the cost of traffic by estimating the hedonic house price function. Using the LCC as an instrumental variable for traffic conditions, this study is an improvement from the typical cross-sectional approaches that are blighted by omitted variable bias and sorting.

The evidence suggests that the introduction of the CC in the CCZ and WEZ are associated with declines in traffic volume and increments in housing values. Comparing properties just inside and outside the Congestion Charge (CC) boundary to reduce unobserved neighbourhood differences, I observe that new homeowners pay, on average, 3.6% (£40,968) more for their homes to enjoy 8.5% (1,779 vehicles) reduction in traffic in the CCZ. Putting these results together, the instrumental variable estimates imply that the elasticity of housing values with respect to traffic is -0.43. These results are robust across a battery of sensitivity analyses and placebo tests. Compared to the previous literature, these estimates on the average marginal willingness to pay to avoid traffic are much larger and are far less sensitive to specifications. Additional results indicate that home buyers could have paid more for better air quality and safer roads in the cordoned area.

My estimates indicate that the tolls generated substantial local wealth gains of around £3.11 billion and £11.91 billion for home-owners in the CCZ and the WEZ respectively relative to neighbours residing outside the zone. These gains measures the local benefits associated with the charge and suggested that the policy created a windfall for residents in the zone by creating a less congested and more conducive living environment.

Given that congestion is fast becoming a salient issue for many cities around the world, this problem has drawn considerable interests from policy makers and economists. Yet, solutions such as constructing more roads (Duranton & Turner, 2011) and implementing fuel taxes (Anas & Lindsey, 2011) are notoriously ineffectual in reducing traffic jams. My findings suggest that although congestion tolls successfully reduce in traffic in the cordoned area, this may be at the expense of neighbouring areas outside as substantial displacement of traffic across the boundary is detected. Hence, to ensure that the policy is effective in abating bottlenecks, there must be proper management of traffic around and beyond the charge zone. Also, it is imperative to provide a reliable and comprehensive public transport system to encourage commuters to switch from driving.

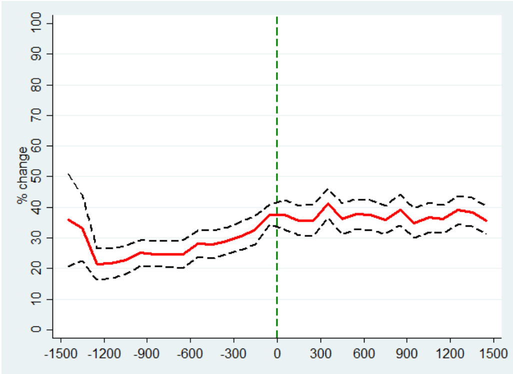
2.8 Data Appendix

A1. Description of Data

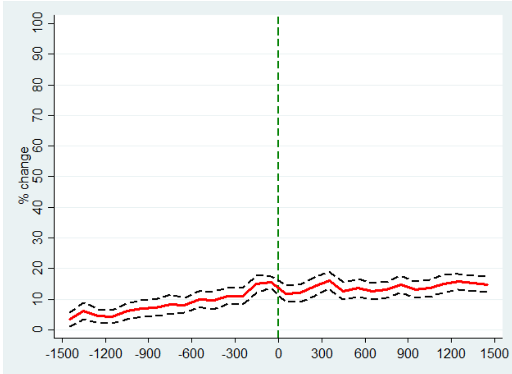
Table 2.9: Description of Variables used in the analysis

Variable	Source	Description
Panel A: Main Specification Variables		
Dependent Variable		
Housing Price (Y_{ijkt})	Land Registry	Natural logarithm of property price of transaction i at postcode k, neighbourhood j at quarter q of year t
Traffic Flow (T_{ijkt})	Department Of Transport	Natural logarithm of traffic flow from vehicles with 4 or more wheels for transaction i at postcode k at year t
Collision Outcomes (A_{rt})	STATS19	Counts of collisions outcome (Accidents, Slight injuries, Serious injuries and Deaths) at road section r at year-month t
Air Pollutant (P_{mt})	London Air Quality Network	Natural logarithm of air pollutant (NO_2, NOX & $PM10$) at monitoring station m at year-month t
Housing Characteristics(X'_{it})		
New Sales	Land Registry	Dummy denoting whether transaction i is new build
Terrace	Land Registry	Dummy denoting whether the property type for transaction i is terrace
Leasehold	Land Registry	Dummy denoting whether the tenure for transaction i is leasehold
Location/Neighbourhood Characteristics (V'_{jt})		
Distance to the CCZ/WEZ boundary	-	Elucidian distance of postcode j from the perimeter of the CCZ/WEZ
Distance to nearest Grade 1 Park	Magic	Elucidian distance of nearest Grade 1 Park from postcode j in km
Counts of Heritage Buildings	Magic	Number of Heritage buildings within 200m from postcode j
Thames River View	Digimap	Binary variable = 1 if postcode j within 200m from Thames River, 0 otherwise
Minority race residents	Census 2001 & 2011	% of Asian/African/Middle Eastern and other minority race residents in OA
Unemployment rate	Census 2001 & 2011	% of unemployed working adults in OA
Uneducated residents	Census 2001 & 2011	% of residents in OA with no education qualifications
Lone parent households	Census 2001 & 2011	% of single-parent households in OA

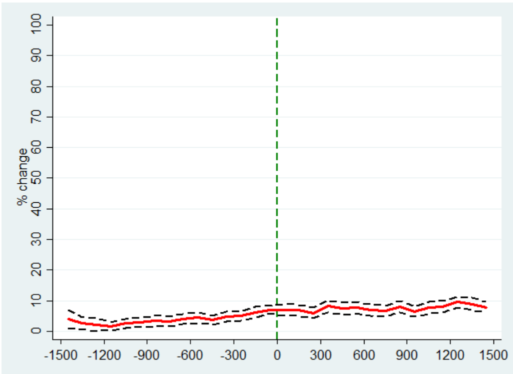
A2. Additional Descriptive Statistics



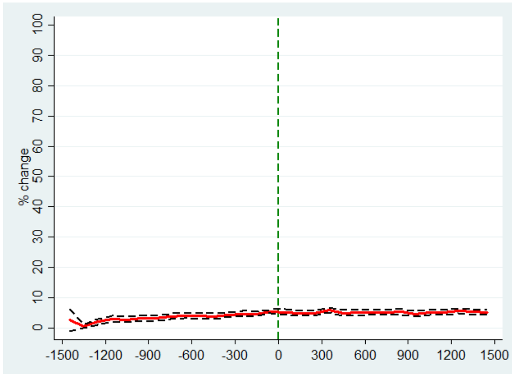
(a) Minority ethnicity



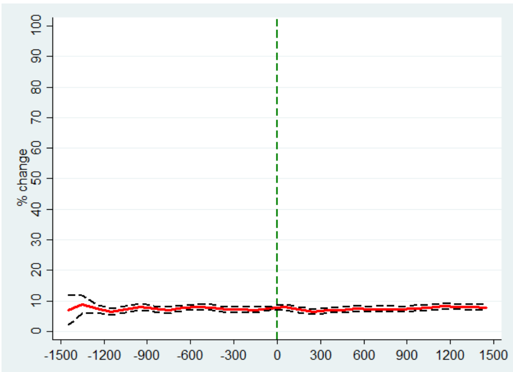
(d) Uneducated



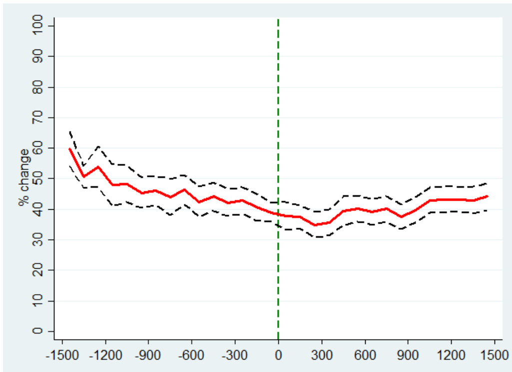
(b) Lone parent households



(e) Unemployment Rate



(c) Driving to work



(f) Car ownership

Figure 2.10: Census demographics around the WEZ. The solid line represents the conditional average change of various demographics at a given distance from the CC boundary and the dashed line represents the 95% confidence interval. It is constructed by regressing the % Δ in demographics at Census Output Area with boundary fixed effect and 100 meters distance bandwidths and coefficient of each distance dummy is plotted. Distance is negative when it is in the charged zone (Left of dashed Line). There are a total of 1,727 output areas within 1.5 kilometres in and out of the WEZ.

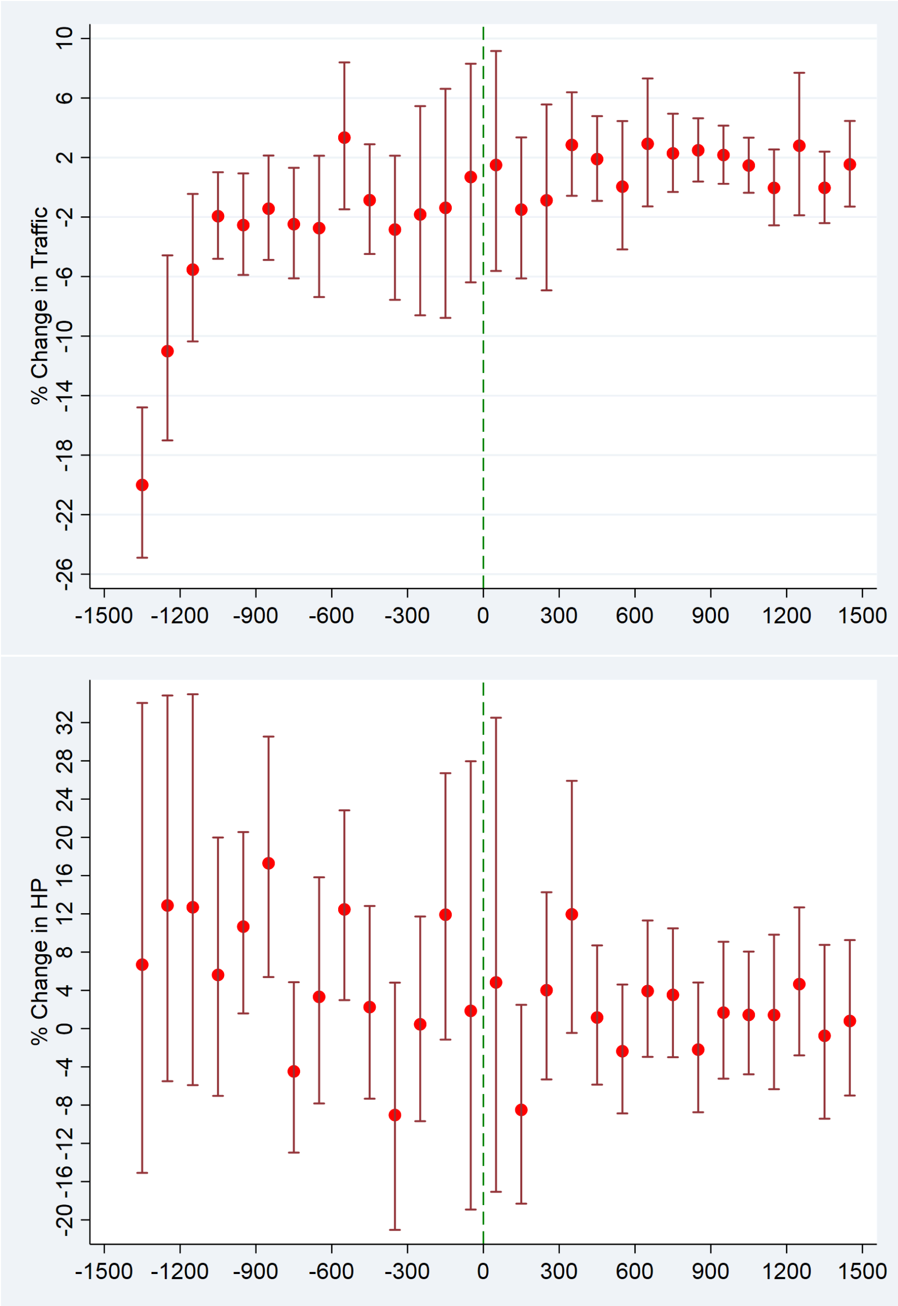


Figure 2.11: The effects of the WEZ on Traffic (Top) and House Prices (Bottom) across distance around the CC boundary. Distance is negative when it in the charged zone (Left of dashed line). The plotted coefficients denote the localised conditional changes of traffic/house prices at a given distance from the WEZ boundary. Tails denote 95% confidence interval. They are constructed by regressing traffic volume or house prices against postcode and year-month fixed effects, observable housing and neighbourhood covariates and 100 metres distance-bandwidth dummies interacted with $Post_{it}$ and distance-bandwidth dummies interacted with CC_{it}. $Post_{it}$ is equal to one for observations after the WEZ is enforced, while CC_{it} takes the value of one for observations in the cordoned area after the WEZ is enforced. This figure plots the coefficients associated with these distance dummies.

A3. Effects of Other Congestion Charge Events on Traffic & House Prices

In this section, I report the effects of the other Congestion Charge events on traffic and housing values. These events include (1) the increase in the charge from £5 to £8 from the 4th of July 2005 ($CCZ2005$) and (2) from £8 to £10 from the 4th of January 2011 ($CCZ2011$), (3) the removal of the WEZ from the 24th December 2010 ($RemWEZ$) and (4) the increase in the charge from £10 to £11.50 from the 16th of June 2014 ($CCZ2014$). The sample windows for these events are defined by 2 years before and after the respective event dates. Refer to Figure 2.4 for more details.

Overall, as reflected in Panel A of Table 2.10, most of the charge increments do not materially improve traffic conditions. This could explain why these hikes have no effects on house prices, as documented in Panel B. The only exception is during the charge increment in 2005 ($CCZ2005$). Restricting the analysis to areas 1 kilometre in and out the CCZ, I observe significant reductions of traffic flow at around 2.98%. This works out to around 523 less vehicles every day. Corresponding to these reductions, home prices are 1.28% higher in the CCZ although this estimate is not statistically significant at any conventional levels. Interpreting these estimates collectively, the elasticity of housing prices with respect to traffic volume, as shown in Column (5) of Table 2.11, is around -0.43, which is fairly comparable to earlier findings. Negligible impact of these hikes explain why none of the MWTP to avoid traffic estimates are statistically significant in Table 2.11. The immaterial effects of the charge increment are consistent with the findings reported by Agarwal *et al.* (2015) who also show that the increase in the CC in Singapore do not affect residential transacted prices.

In other results, I observe a slight rebound in traffic flow that ranges between 2.47% and 4.08% (482 - 798 vehicles) in response to the removal of the WEZ. This surge in the traffic after the WEZ (366 - 839 vehicles) is taken away is remarkably comparable to the effects documented after the implementation of the WEZ (See Table 2.5). These results reinforce the effectiveness of the implementation of the CC in reducing congestion. Contrary to expectations, I observe effects of around 3.24% to 3.77% that are too imprecise to be statistically significant after the removal of the WEZ. The muted effects on home prices explain why the average MWTP on traffic are imprecisely estimated[32] across the board for the removal of the WEZ as observed in Table 2.11.

[32]Even though the removal of the WEZ has a significant effect on traffic conditions, the weak first stage F-statistics (See Table 2.10) suggest that the instrument might have the strength to obtain consistent estimate of the MWTP to avoid traffic.

Table 2.10: Estimates of the Impact of the other Congestion Charge events on Traffic & House Prices

	(1) 5km	(2) 4km	(3) 3km	(4) 2km	(5) 1km	(6) 5km	(7) 4km	(8) 3km	(9) 2km	(10) 1km
	Panel A: First Stage (Log Traffic)					*Panel B: Reduced Form (Log House Price)*				
CCZ2005	-0.0132	-0.0143	-0.0180	-0.0190	-0.0294^{b}	0.0027	0.0045	0.0022	0.0083	0.0127
	(0.0133)	(0.0132)	(0.0128)	(0.0129)	(0.0149)	(0.0130)	(0.0135)	(0.0143)	(0.0151)	(0.0186)
R2	1.00	1.00	1.00	0.99	0.99	0.79	0.78	0.78	0.77	0.75
Δ Traffic	-234	-253	-318	-337	-527	.	.	.	.	.
Δ HP	.	.	.	.	.	2055	3427	1685	6357	9690
F-stats	0.97	1.17	1.97	2.17	3.89	.	.	.	.	.
CCZ2011	-0.0156	-0.0095	-0.0107	-0.0016	0.0102	-0.0121	-0.0104	-0.0132	-0.0094	0.0280
	(0.0147)	(0.0148)	(0.0146)	(0.0144)	(0.0185)	(0.0276)	(0.0289)	(0.0308)	(0.0310)	(0.0303)
R2	0.99	0.99	0.99	0.99	0.99	0.77	0.77	0.77	0.77	0.73
Δ Traffic	-239	-147	-165	-25	156	.	.	.	.	.
Δ HP	.	.	.	.	.	-6596	-5696	-7220	-5153	15462
F-stats	1.12	0.42	0.54	0.01	0.31	.	.	.	.	.
RemWEZ	0.0321^{b}	0.0394^{a}	0.0366^{b}	0.0400^{a}	0.0244	0.0366	0.0347	0.0319	0.0328	0.0370
	(0.0144)	(0.0146)	(0.0145)	(0.0146)	(0.0163)	(0.0245)	(0.0255)	(0.0263)	(0.0271)	(0.0295)
R2	0.99	0.99	0.99	0.99	0.99	0.80	0.79	0.78	0.77	0.74
Δ Traffic	639	785	729	798	482	.	.	.	.	.
Δ HP	.	.	.	.	.	36790	34775	31952	32874	37019
F-stats	4.99	7.31	6.35	7.55	2.25	.	.	.	.	.
CCZ2014	-0.0032	-0.0061	0.0022	0.0015	0.0028	-0.0610	-0.0622	-0.0572	-0.0698	-0.1329^{b}
	(0.0057)	(0.0057)	(0.0053)	(0.0058)	(0.0041)	(0.0427)	(0.0430)	(0.0451)	(0.0464)	(0.0574)
R2	1.00	1.00	1.00	1.00	1.00	0.79	0.79	0.81	0.80	0.78
Δ Traffic	-51	-96	34	23	44	.	.	.	.	
Δ HP	.	.	.	.	.	-49002	-49958	-46041	-55819	-94761
F-stats	0.32	1.15	0.17	0.07	0.47	.	.	.	.	.

Each coefficient (γ) is from a different first stage estimation of equation 2.2. Sample is constrained to properties within 5 kilometres from the CCZ/WEZ boundary. Panel A reports the first stage regressions with the natural logarithm of annual average traffic volume as the dependent variable. Panel B summarizes reduced form regressions with the natural logarithm of house prices. Four other events are examined: (1) the increase in the charge from £5 to £8 in 2005 ($CCZ2005$), (2) the increase in the charge from £8 to £10 in 2011 ($CCZ2011$) (3) the removal of the WEZ in 2010 ($RemWEZ$) and (4) the increase in the charge from £10 to £11.50 in 2014 ($CCZ2014$). Robust standard errors (in parenthesis) are clustered at output area. ^{c} p<0.10, ^{b} p<0.05, ^{a} p<0.01.

Table 2.11: Instrumental Variable Estimates of the impact of the other charge events on House Prices

	(1) 5km	(2) 4km	(3) 3km	(4) 2km	(5) 1km
ln(Traffic) - CCZ2005	-0.2041	-0.3135	-0.1225	-0.4352	-0.4304
	(1.0233)	(1.0112)	(0.8029)	(0.8587)	(0.6845)
Obs	43169	34873	26588	18743	11060
R2	0.79	0.78	0.78	0.76	0.75
No.of Postcodes	7361	6042	4667	3259	1888
ln(Traffic) - CCZ2011	0.7741	1.0902	1.2334	5.7398	2.7304
	(1.8886)	(3.3888)	(3.2590)	(53.9422)	(5.9623)
Obs	21098	17952	13729	9499	5508
R2	0.76	0.74	0.75	0.19	0.62
No.of Postcodes	3967	3440	2756	1964	1108
ln(Traffic) - RemWEZ	1.1398	0.8800	0.8714	0.8194	1.5180
	(0.9530)	(0.7470)	(0.8183)	(0.7652)	(1.6835)
Obs	27122	22920	18372	13486	8545
R2	0.77	0.77	0.77	0.76	0.70
No.of Postcodes	4190	3474	2748	2010	1235
ln(Traffic) - CCZ2014	18.9003	10.1860	-26.5141	-47.3229	-47.5710
	(35.7493)	(11.6536)	(67.0420)	(186.2897)	(71.3023)
Obs	16064	12882	9349	6705	4068
R2	-2.18	-0.07	-1.62	-5.67	-3.49
No.of Postcodes	3092	2550	1886	1301	748

Each coefficient is the IV estimate (β_{IV}) from a different regression that measures the **direct elasticity between traffic volume and house prices** using the different CC events as instruments that include 1) the increase in the charge from £5 to £8 in 2005 ($CCZ2005$), (2) the increase in the charge from £8 to £10 in 2011 ($CCZ2011$) (3) the removal of the WEZ in 2010 ($RemWEZ$) and (4) the increase in the charge from £10 to £11.50 in 2014 ($CCZ2014$). Dependent variable is the natural logarithm of transacted house prices. Sample is constrained to properties within 5 kilometres (Column 1) to 1 kilometre (Column 5) from the CCZ/WEZ boundary. See notes in previous tables for details on the control variables included. Robust standard errors (in parenthesis) are clustered at output area. ^{c} p<0.10, ^{b} p<0.05, ^{a} p<0.01.

A4. Additional Results for the WEZ

In this section, I provide results for the WEZ analogous to earlier specifications that I have conducted for the CCZ. In particular, I constraint the sample of transactions from 900 to 500 meters from the WEZ boundary in Table 2.12. Although there are some evidences indicating that the introduction of the WEZ reduces traffic flow and increases home prices in the cordoned area, these results fluctuate across the different samples. Moreover, low F-statistics (<10) suggests that WEZ is a weak instrument with limited impact on improving traffic conditions. Using the WEZ as an instrument could produce unreliable and overly inflated estimates of the hedonic price schedule gradient.

Table 2.12: Reduced form & IV estimates of the WEZ on Traffic & House Prices

	900m	800m	700m	600m	500m
Panel A: First Stage (Log Traffic)					
WEZ	-0.0180[c]	-0.0212[c]	-0.0207	-0.0098	-0.0316[c]
	(0.0107)	(0.0120)	(0.0139)	(0.0147)	(0.0170)
R2	0.99	0.99	0.99	0.99	0.99
Δ Traffic	-419	-505	-469	-218	-717
F-stats	2.84	3.12	2.22	0.44	3.47
Panel B: Reduced Form (Log House Price)					
WEZ	0.0573[b]	0.0334	0.0412	0.0352	0.0031
	(0.0254)	(0.0272)	(0.0305)	(0.0345)	(0.0419)
R2	0.75	0.75	0.74	0.74	0.73
Δ HP	87387	49606	56745	47444	4107
Panel C: IV Regressions					
ln(Traffic) - WEZ	-3.1821	-1.5741	-1.9944	-3.6093	-0.0989
	(2.2443)	(1.4907)	(1.8881)	(6.3134)	(1.3213)
Obs	11110	9938	8770	7496	6388
R2	0.66	0.72	0.70	0.59	0.73
No.of Postcodes	1675	1469	1296	1110	921

Each coefficient is from a different regression. Sample is constrained to properties within 900m (Column 1) to 500m (Column 5) from the WEZ boundary. See notes in previous tables for details on the control variables included. Robust standard errors (in parenthesis) are clustered at output area. [c] p<0.10, [b] p<0.05, [a] p<0.01.

Similar observations can be made in Table 2.13. Noisy and small effects from the introduction of the WEZ on traffic flow could explain why the willingness to pay to avoid using the WEZ as an instrument is inflated and imprecisely estimated. One notable result is the announcement effects associated with the WEZ in Column (1). It appears that homeowners are fairly optimistic about the impact of the WEZ, as evidenced by the 2.7% (£34,532) increase in home prices. These positive beliefs could be driven by the effectiveness of the CCZ in curbing traffic congestion. It is also intriguing to observe that traffic volume went down by 4.77% (1037 vehicles). One explanation is the possible spillover effects from the implementation of the CCZ. There could be less traffic passing through the WEZ towards the CCZ because of the charge imposed.

Table 2.13: Robustness Tests for the WEZ

	(1) Announce	(2) Pcd>=5	(3) North	(4) Transport	(5) Rem Near	(6) 50m Houses	(7) IDW
Panel A: First Stage (Log Traffic)							
WEZ	-0.0466[a]	-0.0160	-0.0301[a]	-0.0240[b]	-0.0187[c]	-0.0085	-0.0212[b]
	(0.0103)	(0.0106)	(0.0104)	(0.0099)	(0.0113)	(0.0135)	(0.0100)
R2	0.99	0.99	0.99	0.99	0.99	0.99	0.99
Δ Traffic	-1037	-362	-669	-534	-392	-193	-472
F-stats	20.30	2.29	8.41	5.84	2.76	0.40	4.44
Panel B: Reduced Form (Log House Price)							
WEZ	0.0266	0.0639[a]	0.0654[a]	0.0681[a]	0.0647[a]	0.0582[c]	0.0731[a]
	(0.0174)	(0.0242)	(0.0227)	(0.0215)	(0.0221)	(0.0319)	(0.0234)
R2	0.75	0.70	0.76	0.75	0.75	0.74	0.76
Δ HP	34532	91592	100413	104768	103284	88204	112700
Panel C: IV Regressions							
ln(Traffic)	-0.5710	-3.9998	-2.1686[b]	-2.8407[c]	-3.4598	-6.8249	-3.4534[c]
	(0.3951)	(3.0329)	(1.0496)	(1.4920)	(2.3799)	(11.7901)	(1.9555)
Obs	19303	9987	10914	12490	10598	6453	12490
R2	0.74	0.54	0.72	0.68	0.66	0.35	0.66
No.of Postcodes	2951	1014	1776	1896	1635	975	1896

Each coefficient is from a different regression. Sample is constrained to sales 1 km from the CC boundary unless otherwise stated. In (1), the treatment period (CC_{it}) is defined by the announcement window and begins the day the WEZ is announced officially by the Transport for London (TfL) and ends the day before the CCZ is implemented. In (2), I remove any sales in postcodes with less than 5 repeated transactions over sample period. In (3), I remove any sales south of the Thames River. In (4), I include distance to tube-by-year and number of buslines-by-year fixed effects. In (5), I exclude sales that are 100 meters or less from the charge boundary (both inside and outside the zone). In (6), I remove any transactions that are beyond 50 meters from the nearest roads that I can reliably measure traffic flow. In (7), estimates are weighted inversely according the distance from transacted property from matched road. Robust standard errors (in parenthesis) are clustered at output area. [c] p<0.10, [b] p<0.05, [a] p<0.01.

Figure 2.12 report estimates from placebo tests associated with fake pre-treatment windows for the WEZ. Like the analysis before in Figure 3.8, I generate rolling 1-year pre-treatment placebo windows for every quarter from 1996Q1 onwards till 2006Q1. Placebo treatment period is between t_{false} and $t_{false}+1year$ and the placebo window is from $t_{false} - 1year$ to $t_{false} + 1year$ where t_{false} represents every quarter 1 year before the implementation of the WEZ. The new key regressor - $WEZ * t_{false}$ - is the interaction of a binary variable of whether the property i in the WEZ is sold during the false treatment period. This falsification test incorporates transactions within 1 kilometre from the WEZ boundary. None of the placebo estimates are bigger than the implementation effects[33] denoted by the dashed line. If anything, home prices in the WEZ dipped before 2000 but these trends should not confound earlier capitalization effects on the WEZ.

[33]This WEZ estimate on house prices is obtained from Column (5) of Panel A from Table 2.5.

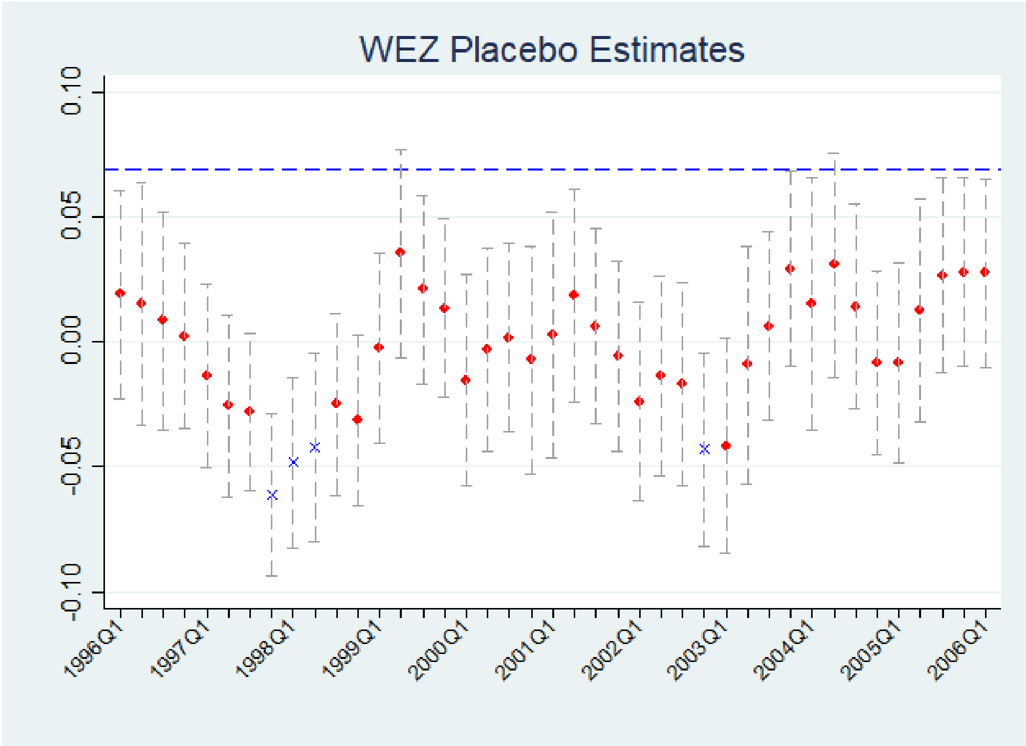


Figure 2.12: The WEZ Placebo Estimates during pre-treatment period. Each point represents a different regression where the treatment period is 1-year rolling window from the corresponding quarter and the pre-treatment period is 1 year before the quarter. The tails represent 95% confidence interval and cross denotes that the estimate is significant at least at 10%, and is insignificant otherwise.

Table 2.14: Reduced form estimates of the Congestion Charge Discount on Traffic

	(1) 5km	(2) 4km	(3) 3km	(4) 2km	(5) 1km
Discount	0.0023	0.0013	0.0088	0.0103	0.0178
	(0.0084)	(0.0087)	(0.0091)	(0.0099)	(0.0120)
WEZ	-0.0407[a]	-0.0415[a]	-0.0332[a]	-0.0315[a]	-0.0258[a]
	(0.0075)	(0.0079)	(0.0080)	(0.0080)	(0.0098)
Obs	46819	39349	31814	24012	15161
R2	0.99	0.99	0.99	0.99	0.99
No.of Postcodes	7576	6284	4979	3699	2232
Discount	0.0053	0.0152	0.0010	0.0220	0.0289
	(0.0158)	(0.0177)	(0.0182)	(0.0193)	(0.0218)
RemWEZ	0.0388[a]	0.0486[a]	0.0465[a]	0.0639[a]	0.0546[a]
	(0.0132)	(0.0142)	(0.0141)	(0.0137)	(0.0151)
Obs	51393	43746	35255	25554	15976
R2	0.99	0.99	0.99	0.99	0.99
No.of Postcodes	7770	6458	5132	3763	2332

Each coefficient is from a different regression. Sample is constrained to sales within 5km (Column 1) to 1km (Column 5). *Discount* is a binary variable equals to one for sales made inside the discount zone after the WEZ is introduced. Dependent variable is the natural logarithm of the traffic flow. Robust standard errors clustered at output area are reported in parenthesis.[c] $p<0.10$, [b] $p<0.05$, [a] $p<0.01$

A5. Effects of the Congestion Charge on Air Quality and Accidents

In this section, I provide some evidence to understand why home buyers are willing to pay for better traffic in the cordoned area after the CC is introduced. In particular, I examine how the effects of the CC - both the CCZ and the WEZ - on traffic collisions and air pollution.

$$A_{rt} = \alpha_r + \delta \mathbf{CC_{rt}} + \omega_t + \varepsilon_{rt}, \tag{2.6}$$

Table 2.15 reports the key estimate (δ) from equation 2.6 for both the CCZ and the WEZ. These regressions include road-level (α_r) and year-quarter (ω_t) fixed effects and the key variable of interest, CC_{rt}, is an indicator variable that takes the value of 1 for roads inside the CCZ/WEZ after the charge is implemented. If the reduction in traffic due to the introduction of CC makes roads safer, I expect δ to be < 0. I constrain the analysis to roads within 3 kilometres from the CCZ/WEZ boundary, with the notion being that this bandwidth covers the entire cordoned area. Here, the dependent variables are the year-quarterly counts[34] of collisions, slight injuries, serious injuries and deaths collected at a particular section of the road (r) at year-quarter (t). Given that the dependent variables are non-discrete count outcomes, I further report Poisson estimates of δ.

As expected, small but significant reductions in collisions are observed and estimates are fairly comparable between OLS and Poisson regressions. Specifically, the counts of accidents and injuries went down by about 4.7 to 4.9% after the CCZ is implemented. Although sizeable reductions in deaths are documented, they are too imprecisely estimated to be statistically significant. Conversely, I do not report significant reductions in collision outcomes after the WEZ is implemented. This is understandable given that the Western Extension of the CC had a negligible effect on reducing traffic flow in the first place.

[34]The reason why this is not conducted at a monthly level is because traffic accidents are fairly rare events and aggregation at a monthly level will lead to a disproportionate number of zeros in the dataset.

Table 2.15: Effect of the CCZ & WEZ on Traffic Accidents

	CCZ				WEZ			
	Acc	*Slight*	*Serious*	*Deaths*	*Acc*	*Slight*	*Serious*	*Deaths*
OLS	-0.0150[a]	-0.0147[b]	-0.0028	-0.0140	0.0019	-0.0007	0.0019	-0.0040
	(0.0045)	(0.0066)	(0.0039)	(0.0107)	(0.0042)	(0.0055)	(0.0061)	(0.0112)
Obs	255789	243192	94271	6428	193431	182007	67578	5019
R2	0.26	0.21	0.05	0.01	0.19	0.16	0.04	0.01
No.of Road Segments	10671	10146	3932	268	8070	7593	2819	210
Poisson	-0.048[b]	-0.050[c]	-0.002	-0.295	0.005	-0.010	0.013	-0.108
	(0.023)	(0.027)	(0.056)	(0.284)	(0.034)	(0.040)	(0.089)	(0.380)
Obs	255789	243192	94271	6428	193431	182007	67578	5019
Absolute	-0.011	-0.012	-0.000	-0.015	0.001	-0.001	0.001	-0.005
% Δ	-4.723	-4.910	-0.182	-25.560	0.544	-0.953	1.268	-10.229
No.of Road Segments	10671	10146	3932	268	8070	7593	2819	210

Dependent variable is the counts of accident outcomes collected at road segment r at quarter t. *Acc* is the total counts of all collision, *Slight* is the number of slight injuries, *Deaths* is the total number of causalities from accidents. Each coefficient (δ) is from different regressions that incorporate roads within 3 kilometres from the CCZ/WEZ boundary estimated with year-quarter and road fixed effects. CC_{rt} is a binary variable that takes the value of 1 for roads in the CCZ/WEZ after the charge is implemented. Robust standard errors, reported in the parenthesis, are clustered at road-level

$$P_{mt} = \alpha_m + \varphi \mathbf{CC_{mt}} + X'_{mt}\phi + \omega_t + \varepsilon_{mt}, \tag{2.7}$$

Table 2.16 reports the key estimate (φ), which captures the percentage change in pollutants after the CCZ and WEZ are implemented from equation 2.7. The dependent variables include the natural logarithm of various pollutants including nitrogen oxide (NOX), nitrogen dioxide (NO2) and particulate matter 10 (PM10) [35] collected at a monitoring station level (m) at month (t). Other than exploiting the monthly variation of air quality within monitoring stations with station (α_m) and year-month (ω_t) fixed effects, I further control for differences in wind speed, wind direction, temperature, relative humidity and barometric pressure.

Panel A presents the effects of the CCZ on air quality within the cordoned area. A surprising and intriguing set of results emerges. The estimate implies that after the CCZ is enforced, the concentration of NOX and NO2 are 4% and 1% higher respectively although the estimates for NO2 are too imprecisely estimated to be statistically significant. This result is consistent with that reported by Green *et al.* (2018). They explain that the implementation of the CC could lead to the substitution of diesel-based vehicles, such as buses and taxis, in the charged zone as they are waived from paying the CC. The combustion of diesel produces more nitrogen oxides and this could explain the higher concentration of these pollutants in the zone.

Moreover, given that I am comparing neighbourhoods very close to one another around the boundary, another possibility is that air pollutants travel across space

[35] There are other pollutants such as sulphur dioxide, PM2.5 and ozone. However, missing observations across the sample period meant that there are insufficient data points for statistical analysis.

into the cordoned area. Findings on PM10 support this hypothesis. In particular, I observe that the concentration of PM10 is about 4% lower after the CCZ is introduced. This pollutant is considerably heavier and less airborne compared to nitrogen oxides. Panel B performs similar analyses for the introduction of the WEZ and reveal smaller effects that are less precisely estimated in the same direction. This is expected given that traffic conditions do not improve significantly after the WEZ is put in place.

All in all, these results suggest that the reduction in traffic from the implementation of the CC leads to safer roads and better air quality in the cordoned area that could explain why homeowners are paying more for homes.

Table 2.16: Effect of the CCZ & WEZ on Air Quality

	Panel A:CCZ			**Panel B:WEZ**		
	NOX	NO2	PM10	NOX	NO2	PM10
CCZ/WEZ	0.039[c]	0.010	-0.038[b]	0.018	0.023	-0.016
	(0.021)	(0.017)	(0.016)	(0.027)	(0.021)	(0.026)
Obs	1422	1412	990	1215	1214	1118
R2	0.94	0.90	0.86	0.94	0.93	0.86
No.of.Stations	22	22	19	27	27	22
Treated	8	8	6	3	3	2
Absolute	5.93	0.65	-1.27	4.13	2.10	-0.46

Dependent variable is the natural log of pollutant collected at monitoring station m at month t. Each coefficient (φ) is from different regressions that incorporate stations within 3 kilometres from the CCZ/WEZ boundary estimated with year-month and monitoring station fixed effects. CC_{mt} is a binary variable that takes the value of 1 for stations in the CCZ/WEZ after the charge is implemented. Control variables include wind speed, wind direction, temperature, relative humidity and barometric pressure. Robust standard errors, reported in the parenthesis, are clustered at district*year-month.

Chapter 3

Do Speed Cameras Save Lives?

3.1 Introduction

Every year, more than 50 million people are injured, with more than 1.2 million people are killed by auto-mobile crashes around the world (Peden *et al.*, 2004). Likewise, across United Kingdom (UK), traffic collisions cause more than 160,000 injuries and 1,730 fatalities in 2015 (DfT, 2016). These crashes have disproportionately affected the younger generation. It is the leading cause of death for those between 5 and 34 years old, accounting for more than 15% of their deaths and inducing many life years lost[1]. In monetary terms, these accidents cost UK a total of £10.3 billion in 2015[2].

Speeding is one of the main reasons[3] why crashes occur. According to Department for Transport (DfT), speeding accounts for more than 24% of the fatal accidents that occurred in UK in 2015. The severity of the crashes is also dependent on the velocity of the colliding vehicles. Studies have shown that the fatality risk at 50 km/h is twice larger than the risk at 40 km/h, and more than five times larger than the risk at 30 km/h (Rosén & Sander, 2009). Although speeding is often considered a menial offence to many, it is immense in determining both the probability and gravity of crashes.

Different interventions, such as traffic police, traffic lights, road humps, speed limits, warning signs, vehicle-activated speed signs and speed enforcement cameras, have been employed to deter speeding. Since the seminal paper by Peltzman (1975),

[1]Based on 2016 figures from Office for National Statistics, there are a total of 3,423 deaths from accidents. This is just slightly more than 10% of the 30,570 who passed on from lung cancer. However, assuming that life expectancy is 79 for males and 82.8 for females, based on the demographics of the victims, the total number of life years lost from accident amounts to 117,285, which is almost half of the 264,424 life years lost from lung cancer. Clearly, one is underestimating how detrimental traffic accidents can be by just looking at casualty figures.

[2]These figures are much larger in United States. A federal study conducted by National Highway Traffic Safety Administration reveals that estimated economic cost from motor crashes is approximately US$242 billion in 2010 (Administration *et al.*, 2014).

[3]Many other factors explain why traffic collisions occur. For instance, intoxication (Dee, 1999; Levitt & Porter, 2001a; Hansen, 2015), distraction from the use of mobile phones (Abouk & Adams, 2013), failure to use seat belts (Levitt & Porter, 2001b; Cohen & Einav, 2003) and visibility (Ho *et al.*, 2017) could increase the risk and severity of crashes.

evaluating these measures has drawn considerable attention from economists. These studies have examined the effect of speed limits (Ashenfelter & Greenstone, 2004; van Benthem, 2015), traffic police (DeAngelo & Hansen, 2014) and red-light cameras (Gallagher & Fisher, 2017) on traffic accidents. Falling back to the economic models of crime (Becker, 1968), these instruments deter reckless driving through punishment, such as fines, driving suspension and incarceration. One widely used strategy that has drawn substantial interest from the transport safety literature, but surprisingly scant attention from economists, is speed camera. These devices are usually deployed at sites prone to collisions (e.g windy, hilly roads) or sites with vulnerable pedestrians (e.g near schools, transportation nodes and petrol stations). They penalize drivers for exceeding speed limits around the cameras.

In this paper, I estimate the effects of fixed speed cameras on reducing the occurrence and severity of collisions. To do so, I put together a rich dataset of more than 2,500 fixed speed cameras across England, Scotland and Wales (Great Britain). I rely on the STATS19 Road Accident Dataset that documents details (location, number of injuries and fatalities etc.) of every reported collision since 1979. This comprehensive dataset allows me to conduct the analysis at a fine spatial scale and capture how enforcement effects change moving away from the camera. In short, I compare accident outcomes before and after the camera is introduced with similar non-camera sites using a quasi-experimental difference-in-difference estimation strategy.

For the estimates to be valid, it requires the mean differences in unobserved characteristics between sites not to be correlated with the installation of enforcement cameras. This assumption, however, is likely to be violated given the endogeneity in site selection. Cameras are often found at areas prone to collisions (e.g more traffic, sharp bends) and this selection process is likely to accentuate the differences between sites with and without cameras. The difficulty in identifying the enforcement effects of cameras is further exacerbated by the fact that installations can happen even when sites do not meet the selection rules (see Section 2). Moreover, the timing of intervention is likely to be endogenous as well. It is more probable for sites that experience a sharp increase in accidents to receive cameras. This means that, regardless of intervention, collisions will probably to revert to lower levels, inducing an over-estimation of enforcement effects.

I adopt several strategies to address these endogeneity concerns. First, to partial out time-invariant differences between sites, I include site fixed effects to exploit the variation in collisions within each site before after the speed camera is installed. Second, to avoid the bias from the obscure selection process, I restrict the analysis to only sites that will ever have enforcement cameras and rely on time variation of installation for identification. Put differently, sites with camera installations in the future (but no cameras now) will be employed as reference groups for sites having installations now. Third, I minimize observable differences by controlling for a rich set of time-variant city level characteristics. The concern is whether there are region-specific shocks that could be correlated with camera installations. Still there is the issue whether "worse" sites are treated first. This means that sites that received cameras far apart in time may be incomparable. Thus, I restrict reference groups to sites that received installations less than six years apart from those treated now.

Finally, I address the endogenous timing of installations with two strategies. First, I plot the pre-treatment collision trends to show that cameras are not strategically introduced after a spike in collisions. In fact, due to bureaucratic red-tapes, cameras are often installed a few years later after the sharp spike, mitigating concerns that the estimates are inflated by mean reversion effects. Second, I exploit the "switching-off" of speed cameras to capture the efficacy of these devices. Due to budget cuts, some local camera partnerships are forced to switch off their cameras. I argue that this decision to cut funding is unlikely be driven by collision trends. The question is whether shutting down of speed cameras attributes to a rebound in traffic collisions.

Other than adopting a more careful identification strategy, this paper improves the existing literature (See Table 3.8 in Data Appendix for details) on several fronts. First, in contrast to previous papers, which are usually city-specific analyses restricted to a small sample of cameras, this paper draw inferences from a representative nationwide dataset to increase the external validity of the research. Second, with fine spatial temporal information on accidents and speed cameras, I can accurately capture how enforcement effects vary across space. Last but not least, I provide a rigorous welfare assessment of speed cameras, after considering a exhaustive list of benefits and costs, to understand whether these cameras improve social welfare.

The headline finding is that speed cameras reduce both the number and severity of collisions. After installing a speed camera, the number of accidents and minor injuries reduce by 17%-39% and 17%-38% respectively, which corresponds to 0.89-2.36 less accidents and 1.19-2.87 less injuries per kilometre per year. As for seriousness of the crashes, the number of fatalities and serious injuries are 0.08-0.19 and 0.25-0.58 lower per kilometre per year, which represents a drop of 58%-68% and 28%-55% respectively. Installing another 1,000 speed cameras reduce around 1130 collisions[4], mitigate 330 serious injuries, and save 190 lives annually[5], generating benefits of

[4] These estimates are taken from the preferred specification in Column (7) of Table 3.2.

[5] The ratio of lives save in my study is much higher than the average national accidents death ratio over the last 10 years from 1995 to 2015 (1.02%). There are several explanations to this finding. First, speed cameras are often found along roads with a much larger proportion of death related accidents. The pre-treatment percentage of deaths from

around £309 million[6]. These findings are robust across a range of specifications that mitigates the risk of potential con-founders from driving the estimates.

I further allow enforcement effects to vary across different speed limits, road types, and over distance. My results show that enforcement effects are larger along roads with higher speed limits. This could be due to the fact that these roads are more dangerous as drivers commute at higher speeds. Enforcement effects also appear highly localised around 500 metres from the camera and dissipate moving away. Beyond 1.5 kilometres from the camera, there are suggestive evidences of rebounds in collisions, injuries and deaths, implying that drivers could have speed up beyond camera surveillance and cause more accidents.

These findings are of interest for at least three reasons. First, the public has always been concerned because of the huge amount of fines that are raked up by these cameras. A total of 166,216 speed tickets was issued in England and Wales in 2015, amounting to more than £31 million[7]. Interests groups[8] have campaigned vehemently against these instruments, believing that alternative strategies, such as vehicle-activated speed limit sign, could be equivalently effective in improving road safety[9]. Second, there are concerns whether these devices could cause more collisions due to "kangaroo" effects (Elvik, 1997). That is when drivers abruptly slow down in proximity to the camera or immediately speed up beyond surveillance. Thirdly, due to budget cuts to the Road Safety Grants, many older obsolete wet-flim cameras are not upgraded and local governments[10] are forced to switch off their cameras. If fixed speed cameras improve road safety, then these devices should be upgraded and switched back on. The objective of this paper is to provide educated answers to these questions through high quality data and rigorous empirical analyses.

My results verify the efficacy of speed cameras in enhancing road safety. However, the limited enforcement effects across space, together with mild rebound of collisions further away, highlight the limitations associated with these fixated de-

collisions around speed camera sites is 2.50% (see Table 3.1), which is more than twice of the national ratio. Second, by reducing speed through deterrence, cameras could have disproportionately mitigated more severe accidents. Another explanation is that speed cameras are less effective in preventing collisions compared to deaths. Possible kangaroo effects, such as sudden braking in front of camera, or speeding up beyond surveillance, could have attributed to more collisions.

[6]This is obtained from multiplying the net benefits from welfare analysis in Table 3.5 by 1,000.

[7]Read more at http://www.bbc.co.uk/news/uk-38724301

[8]Read http://www.safespeed.org.uk/

[9]See https://www.publications.parliament.uk/pa/cm200708/cmhansrd/cm080422/debtext/80422-0003.htm for more information

[10]This include Oxfordshire, West Midlands, Avon and Somerset, Wiltshire, Swindon and Northamptonshire. Recent reports indicate that switching-off of cameras may have been more widespread, raising greater concerns of the efficacy of speed cameras.

vices. My findings also show that switched-off "dummy" cameras can still enforce speed limit and reduce collisions. Local government, therefore, should keep these cameras as a deterrence whether or not they are operating. All in all, with technological advancement, these older devices should be superseded with newer prototypes, such as mobile and variable speed cameras, that can enforce speed limits over a larger area with the flexibility of redeployment .

The remainder of this paper is structured as follows. Section 2 provides a background to speed enforcement cameras in UK. Section 3 describes the identification strategy adopted in this paper. Section 4 outlines the data used in this paper and Section 5 discusses the findings in this paper. Section 6 concludes the study.

3.2 Background

Different enforcement cameras, including fixed, mobile and variable speed, are deployed across the United Kingdom. Fixed speed cameras, the earliest generation of speed detecting devices, are first introduced in 1992. Mobile and Variable[11] speed camera are newer prototypes that only grew in prominence in the last decade. For an illustration of these devices, refer to Figure 3.1. The focus of this paper is on fixed speed cameras as I can reliably determine both the location and installation dates. The minimum penalty for speeding is a fine of £100 and 3 demerit points but offender could be fined up to £2,500 and suspended from driving, depending on how much the speed limit is exceeded.

Cameras are managed by a safety camera partnership, which is a joint collaboration of police force, local government, highway agency and health authorities. They work hand-in-hand to identify dangerous sites for enforcement. Sites that chosen for installations must comply with the following national selection rules (DfT, 2004)[12]:

1. Length must be between 0.4 and 1.5 kilometres;

2. At least 4 killed and serious collisions (KSI) & 8 personal injury collisions (PIC) per kilometre in the 3 years before installation[13];

[11] Mobile speed cameras are fixated on auto-mobiles with the flexibility to be deployed in different locations but require manpower to operate. Variable speed cameras enforce speed limit over a stretch by measuring average speed between two points on the road and have the advantage of enforcing speed limit over longer distances.

[12] One other strategy is to utilize a regression discontinuity design over these rules and to obtain some local estimates around these thresholds. This is not adopted due to the following reasons. First, I do not have information on average speed, site length, suitability that affect whether a site receives camera enforcement. Furthermore, these rules are not deterministic for installation. It is possible for sites to have installations without meeting these rules, impeding identification of effects around these thresholds.

[13] One crash can result in multiple causalities. Adding up the number of slight injury collisions and KSI will provide the PIC counts.

(a) Fixed Speed

(b) Average Speed

(c) Mobile Speed

Figure 3.1: Different types of Speed Cameras used in United Kingdom

3. Suitable for the loading and unloading of cameras

4. At least 85% of the traffic is travelling is at or above the Association of Chief Police Officers (ACPO) threshold based on speed surveys;

5. At least 20% of the drivers are exceeding speed limits;

6. No other more cost effective solutions to improve road safety as determined by the road engineers.

The first two guidelines are considered more important for enforcement. While not stated explicitly, I do observe that many of these cameras are near schools, bus stops and petrol stations to ensure pedestrians safety. Even when some of the stated requirements are not met, enforcement could still occur if a large number of non-fatal collisions due to speeding is recorded. These sites are classified as *exceptional sites*. This ambiguity impedes the use of selection rules to identify comparable reference groups.

The local partnerships also decide whether to install mobile, average or fixed speed cameras. Fixed speed cameras are usually deployed when there are many accidents clustered around the sites. To commission new sites for camera installation, partnerships are require to provide full details on these proposed sites for the forthcoming year, subjected to the approval of the national board. They are allowed to recover penalty receipts to cover the cost of camera installations and enforcement. Since 2006, there are several amendments to the guidelines. In particular, the KSI requirements fall from 4 to 3. A risk value is computed for each site and KSI and PIC collisions are given 5 points and 1 point respectively. To qualify for camera installation, sites must have at least 22 points if the speed limit is 40mph or less and have at least 18 points for speed limits beyond 50mph. For more details, refer to DfT (2005).

Once installed, several clear signages must be placed less than 1 kilometre away from the camera. This is to warn drivers about the presence of camera and to inform them about the speed limit. Since 2002, all the cameras are painted in bold yellow and must be visible at least 60 metres away if the speed limit is less than 40 mph and must be visible at least 100 metres away if the speed limits are higher. This is to ensure that drivers do not abruptly reduce speed around the camera to avoid fines.

Most of the cameras across UK are *Gatsometer BV Cameras* that are single direction and rear facing. This means the camera will only take images of the back of a speeding vehicle so as not to blind the offender and impede driving performance. However, some of the newer cameras could be bi-directional[14] or front facing[15]. Majority of the cameras operate though radar technology although there are some that rely on strips on the roads for speed detection (e.g *Truvelo D-Cam*, *SpeedCurb*). If there is a dispute to the fine, the white lines on the roads near the cameras will provide a secondary instrument to determine [16] whether drivers exceed speed limits. For an illustration on how speed cameras operate, refer to Figure 3.2.

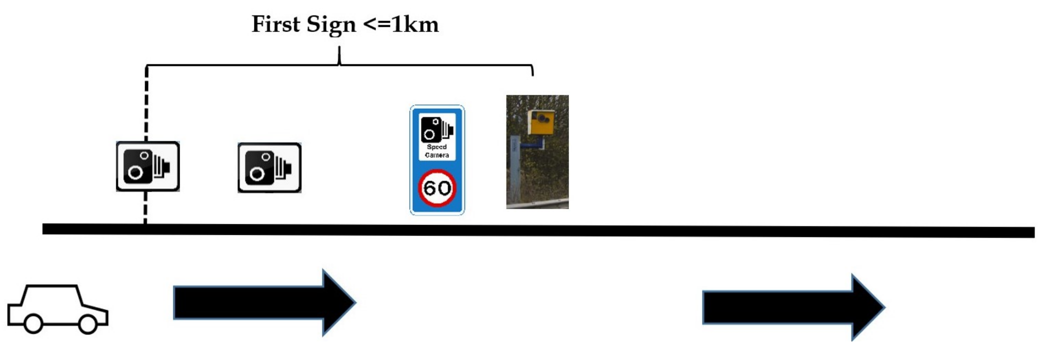


Figure 3.2: Illustration on how fixed speed cameras operate

[14]Cameras installed in the central of the road could be turned periodically to target motorists at either side of the road. Sometimes, multiple cameras could be installed on both sides of the road. Newer devices such as the Truvelo D-Cam can take pictures at both directions.

[15]The second most popular type is *Truvelo Camera* that takes an image of the speeding offender from the front using non visible infra-red flashes. The advantage is that there will be no disputes on who is driving the vehicle.

[16]The distance between each of the white lines represent 5mph. Several images of the moving vehicle over time will illustrate whether driver is speeding.

3.3 Literature Review

Previous literature, largely from transport safety, shows that speed cameras reduce travel speed, accidents, injuries and fatalities near the camera (Gains *et al.*, 2004, 2005; Chen *et al.*, 2002; Shin *et al.*, 2009). These estimates, however, vary substantially across different studies. A survey of existing literature reveals speed cameras reduce travel speed by between 1.7 and 4.4 miles per hour and crashes by between 11% and 51%. For a review of the existing literature, refer to Wilson *et al.* (2010).

Existing empirical work, however, suffers from substantial limitations that questions the validity of the estimates. For one, researches are often limited to a small number of speed cameras constrained in a particular area (Chen *et al.*, 2002; Goldenbeld & van Schagen, 2005; Jones *et al.*, 2008; Shin *et al.*, 2009), raising concerns on the external validity of their findings. This paper overcomes this limitation by analysing a more representative sample of cameras of up to 2,500 fixed speed cameras installed across England, Scotland and Wales.

Secondly, and perhaps most importantly, many studies are restricted to before-and-after analysis with either no or loosely constructed control groups to account for trends in accidents in the absence of camera enforcement (Christie *et al.*, 2003; Jones *et al.*, 2008). Without controlling for the general downward trends of accidents due to technological advancements over time, such as better brake system, more robust car frame and improved road built, these studies are likely to overestimate camera enforcement effects. For studies with control groups, they address the fact that camera location choices are endogenous. Selected camera sites are peculiar accident "black" spots with many drivers exceeding speed limits such that those without cameras are unlikely to be comparable. These differences, if unobserved or imprecisely measured, will enter the specification and confound the estimates. Without due consideration to the endogenous site selection process, studies construct reference groups based on either nearby roads (Newstead & Cameron, 2003; Perez *et al.*, 2007; Shin *et al.*, 2009) or sites with similar observable road and traffic characteristics (Keall *et al.*, 2001; Cunningham *et al.*, 2008).

To create more comparable reference groups, some studies rely on data-generating methods like empirical bayes to identify reference groups with similar trends in accidents and traffic flow (Elvik, 1997; Chen *et al.*, 2002; Gains *et al.*, 2004, 2005). However, these studies fail to show how the reference groups are chosen. To clarify on the matching process, Li *et al.* (2013) uses propensity score matching selection guidelines. This is unlikely to improve identification as sites can receive installations even without meeting all the requirements. Moreover, the surge in accidents surrounding these qualified non-camera sites could be considered transient and is expected to decrease even without intervention. This post-treatment collision reductions could underestimate enforcement effects. In this paper, I adopt the intuitive strategy of using only sites with cameras. That is, sites with cameras in future will be employed as reference groups for sites with installation now.

Another point neglected by the literature is how the effectiveness of enforcement cameras vary over distance. This is important as cameras could attribute to "kangaroo" effects (Elvik, 1997). Several studies, including Newstead & Cameron (2003); Mountain *et al.* (2004); Jones *et al.* (2008), try to break down the impacts across distance but the lack of fine spatial information on collisions mean that results are often uninformative as distance bandwidths are often too big. Relying on fine spatial information on accidents and speed cameras, I can delineate enforcement effects every 100 metres (up to 2 kilometres) to understand whether cameras cause kangaroo effects.

Finally, there is a lack of analysis on how these enforcement cameras fare over time and across different speed limits. One of the few papers that addresses this issue is Christie *et al.* (2003). Their study, however, is limited to an unrepresentative sample of cameras over a short period. Utilising detailed information on speed camera characteristics, and over a longer timespan, I inform how cameras perform over time and across roads with different speed limits. For a succinct summary of the previous literature, refer to Table 3.8 in Data Appendix.

3.4 Data

To examine the effect of speed cameras on accidents, I put together a few data sources. First, I rely on STATS 19 Road Accident Database that provides detailed information for each reported accident to the Police Force in England, Wales and Scotland [17]. Details including location, time, date, road conditions, vehicle type, number of injuries, serious injuries and fatalities (pedestrians and inside the vehicle) are recorded. Shapefiles that delineate the road network and boundaries of local authority districts[18] are provided by Ordinance Survey.

Details of the different speed cameras are hand-collected from websites of various camera partnerships provided by Department for Transport (DfT) [19]. For most of

[17]It is possible that there could be under reporting of non-fatal accidents to the Police Force but this should be less of an issue for more serious crashes. As long as the under reporting of accidents is random across time and is not correlated with camera installations, this should not affect my estimates

[18]Local authorities are responsible of conferring government services within a district. In total, there are 353 different districts in England, 32 in Scotland and 22 in Wales.

[19]For more information on the list of `https://www.gov.uk/government/publications/speed-camera-information`

the partnerships, location of camera, year of installation, speed limits and camera type are provided. For areas that do not provide these data, I request access using Freedom of Information Act (FOI).

Combining various sources of information using Geographic Information System (GIS), I am able to match the location of speed cameras and accidents to the road network. To visualize, refer to Figure 3.3. Imagine the line as a particular stretch of road with a camera installed. With the exact location of each accident, I could sum up the annual accident outcomes along the road that the speed camera i is installed between k and $k-100$ metres interval where $k \in 100, ...1900, 2000$ metres. For instance, within 100 metres around the camera, all the accidents that take place in area "A" in a particular year are taken into account. For my baseline estimates, which examine the effects 500 metres around the camera, I will aggregate all the accidents that took place in areas "A", "B", "C", "D" and "E". Because I know the total number of injuries, serious injuries and deaths associated with each automobile crash, I can construct these collision measures for the different bandwidths as well.

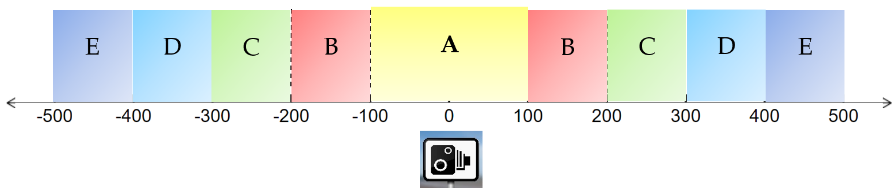


Figure 3.3: Illustration on how accident outcomes are computed across space

To capture the year-on-year variation in region-specific shocks, I rely on several sources. Information on the Annual Average Vehicle Miles Travelled (VMT) is collected from DfT. Details on the average earnings and number of hours worked are complied from Annual Labour Force Survey. Data on population profile are collected from Nomis Population Estimates. For details on how the variables are constructed, refer to Table 3.7 in Data Appendix.

3.5 Identification Strategy and Methodology

The research design adopted in this paper is a fixed effect, quasi-experimental difference-in-difference approach estimated using count regressions models. This is because collision outcomes follow an implicit count process that only takes non-negative integer values. Using Ordinary Least Squares (OLS), which specifies a conditional mean function that takes negative values, one could possibly yield inconsistent estimates (Cameron & Trivedi, 2013). Therefore, I implement two count models, Poisson and Negative Binomial, that is estimated using maximum likelihood estimators (MLE). The latter is adopted because it relaxes the assumption that the conditional mean is equal to the conditional variance, allowing for over-dispersion in the data. To correct for over-dispersion in Poisson regressions, following DeAngelo & Hansen (2014), I report sandwich (robust) standard errors.

To examine the impact of speed cameras on traffic accidents, the following baseline specification is adopted:

$$E(Y_{ijt}) = exp(\alpha_i + \gamma \mathbf{T_{it}} + X'_{jt}\phi + \theta_t + \varepsilon_{ijt}), \tag{3.1}$$

where Y_{ijt} is the counts of Y (**Accidents, Slight Injuries, Serious Injuries, Deaths**[20]) within 0 to 500 metres from camera i in local authority j that is installed in year t. The key variable of interest is $\mathbf{T_{it}}$, a binary variable that equals to unity after the speed camera is installed. If enforcement cameras can deter speeding and improve road safety, I expect γ to be < 0.

α_i represents site fixed effects that captures time invariant unobserved characteristics that influence whether a camera is installed. For instance, sites that are more precarious (e.g on a steep slope, windy roads) or are bypassing areas with more vulnerable pedestrians (e.g schools, petrol stations) are more likely to receive cameras. By including site fixed effects, I am now comparing the change in collision outcomes for each site before and after the cameras become operational with the changes in collision outcomes in some comparable sites.

I further include a vector of time variant city-level controls at local authority j at year t (X'_{jt}). These variables include vehicle miles travelled, population size, percentage of population between 18 to 25 years old, gross annual pay, hours worked and weather conditions. This is to allay concern that there are regional shocks that could be correlated with installation of cameras and influence collision outcomes. For instance, if cameras are installed in areas with an increase in teen drivers that could reduce road safety, γ could be underestimated. θ_t represent year fixed effects to control for any time-specific macro factors that affect traffic collisions across regions. For example, technological advancements on car safety (better car frames, tires, air bags) and roads quality can reduce both the occurrences and severity of collisions over time. For more details on the description of the variables used in this paper, refer to Table 3.7 in Data Appendix.

[20] According to the definition provided by the Department for Transport, slight injury is defined as an injury of minor character that do not require any medical attention. Serious injury is when the injury causes the person to be detained in the hospital for medical treatment and that the injury causes death more than 30 days after the collision. Deaths is defined as a human casualty who sustained injuries from the accident are die less than 30 days from the collision.

ε_{ijt} is the error term and consistent estimation of γ requires $E[\varepsilon_{ijt}|\mathbf{T_{it}} = 0]$. This is unlikely to be plausible even after controlling for camera and year fixed effects, and partial-ling out time variant region specific shocks. Roads with enforcement cameras are peculiar accident-prone sites with many drivers exceeding speed limits and sites without cameras are likely to be very different. The concern is whether these unobserved differences between camera and non-camera sites influence collision outcomes.

Hence, I restrict the sample to only sites with cameras and exploit the variation in the timing of installation. Identification of enforcement effects stems from comparing changes in accident outcomes around camera sites with changes around sites that will have camera installations in the near future. This allow us to attenuate the bias from the "black-box" selection procedure given that these sites will eventually receive cameras in the future. The assumption is that sites having enforcement cameras in the future are not that different from sites having cameras installations now.

However, it is plausible that "worse" sites receive installations first such that later-treated sites are not comparable. Therefore, I remove any observations that are more than 3 years before and after the installation year. To visualize, refer to figure 3.4 that illustrates the timeline for a sample of four cameras (A,B,C & D). Unshaded areas denote the window 3 years before and after the cameras are installed with $T = 0$ representing pre-installation period and $T = 1$ representing post-installation period. Shaded areas denote observations outside the $+3$,-3 window that are not included in the analysis. In this example, CAM B and D are counterfactuals for CAM C. CAM B provides the baseline (collision trends in the absence of camera enforcement) from 1998 to 1999 and CAM D from 2000 to 2001 after CAM C is installed. Conversely, CAM A is not a reference group for CAM C because the treatment dates are too far apart. This also means that only a future recent treated camera will enter as reference group.

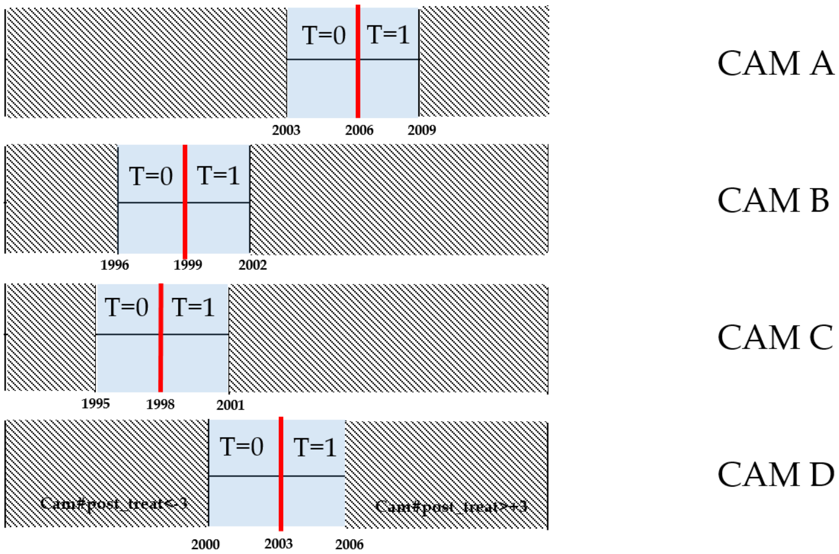


Figure 3.4: Illustration of time-lines for different cameras in sample. Bold lines represent the installation year and unshaded window denotes 3 years before and after the camera is installed. Shaded areas denotes observations more than three years before or after installation and are omitted from the analysis. T=1: Treatment Period; T=0: Pre-Treatment Period

Still, there are concerns that the timing of the installation may be endogenous. Consider the case that the camera is installed only after a sharp increase in collisions. If that is the case, then γ could overstate the enforcement effects as crashes could be reverting to mean (falling) even without cameras. To verify, I plot the conditional mean collision trends at 200 metres from the site 12 years before and after the speed camera is installed in Figure 3.5. I construct this by regressing collision outcomes on site fixed effects and a vector of local authority characteristics and each point represents the respective year-from-installation dummies. Year 0 represents one year before camera installation. Results show that collisions are already falling before the cameras are installed, suggesting that there might have been additional policing around these sites before cameras are put in place. The delay in installation could be due to bureaucratic red-tapes. As explained in by the Department for Transport(DfT, 2004), local camera partnerships are only allowed to request for cameras installations once a year, subjected to the approval of the national safety camera board. Installation could only be scheduled upon approval and the downtime could take up to half a year. These delays in installation mean that my estimates are unlikely to be inflated by any mean reversion effects in collision.

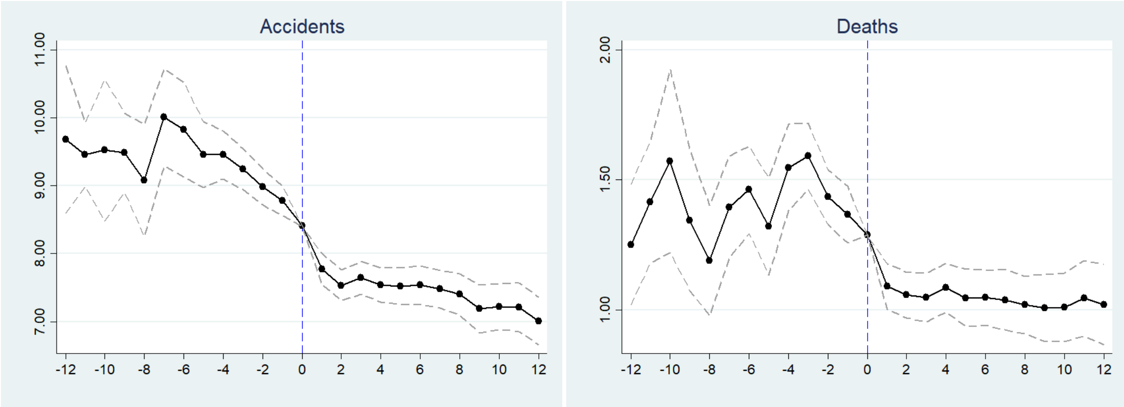


Figure 3.5: Conditional mean collision trends 200 metres from camera site 12 years before and after the installation. Controls include camera fixed effects and a vector of local authority characteristics (Gross Annual Salary, Vehicle Miles Travelled, % of Population from 18 -25, Job Density, Hours Worked). Each dot denotes the coefficients from the number of years from installation dummies. Horizontal axis denotes the number of years from treatment and year 0 is the year before the camera is installed. Vertical axis measures the counts of collision outcomes. Dashed line represents the 95% Confidence Interval.

3.6 Empirical Results

In this section, I estimate the effects of speed enforcement cameras on various accident outcomes. First, I provide some summary statistics before presenting baseline estimates on the effects of speed cameras. I then put these estimates through various robustness and placebo tests that relax identification assumptions. Subsequently, I allow camera enforcement effects to vary across different speed limits, road types, over time and distance. Finally, I compute welfare estimates associated with these devices.

3.6.1 Descriptive Statistics

Figure 3.6 and 3.7 shows the temporal and spatial distribution of fixed speed cameras from 1992 to 2016. 24 cameras are first installed in London in a pilot program in 1992. Soon, other major cities like Manchester, Liverpool and Birmingham begin to adopt these devices. By 2000, there are more than 1,000 cameras distributed across more than half of the local authorities across Great Britain. Fixed speed cameras remain the predominant instrument in enforcing speed limits with another 1,368 devices deployed in the next 8 years. Most local authorities have at least 1 speed camera by 2008. Since then, these devices become less popular as local partnerships rely on newer prototypes, such as variable and mobile speed cameras, for speed enforcement. Only 109 fixed camera sites are added from 2008 to 2016. By

2016, there are approximately 3,500 fixed speed cameras across England, Scotland and Wales. My dataset, which encompasses 2,548 cameras, covers more than 70% of the population. The rest of the 30% are missing either because (1) the local camera partnerships did not respond to data requests[21] or (2) I am not able to accurately determine the location of cameras based on the information provided.

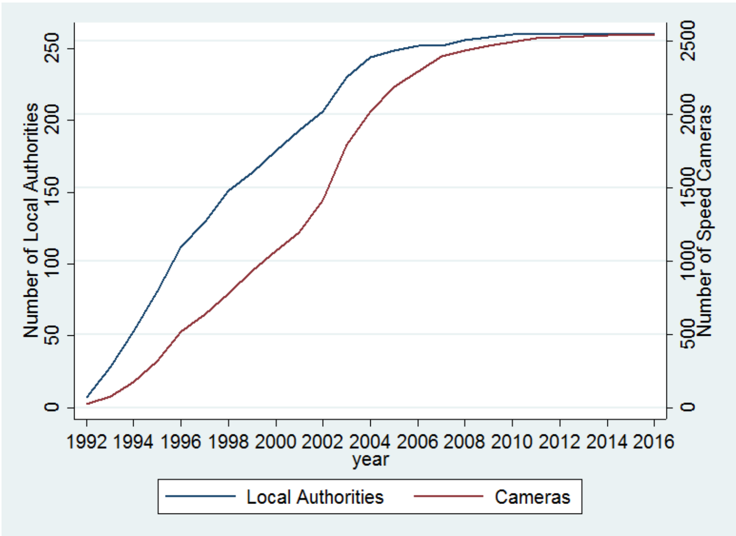


Figure 3.6: Number of Speed Cameras and Local Authorities with speed cameras from 1992 to 2016 across England, Scotland and Wales

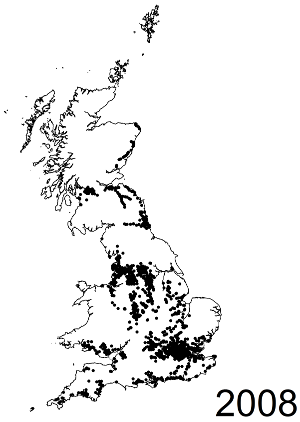
Next, I present some basic summary statistics for pre-treatment accident outcomes, camera, road and local authority characteristics in Table 3.1. Pre-treatment accident outcomes are computed by averaging the number of collisions within 2 kilometres from the site and within five years before the camera is installed. For instance, if a camera is installed in 2000, I will account for the collision outcomes from 1995 to 1999. There are approximately 0.41 counts of accidents every 100 metres annually, resulting in 0.40 counts of slight injuries, 0.08 counts of serious injuries and 0.01 counts of deaths. On average, the limit enforced by speed cameras is around 37mph although bulk of the cameras impose a 30mph limit (more than 70%). Most of the cameras (75%) are installed in A Roads - primary routes that are slightly smaller than motorways (or expressways). The rest are mostly installed in B (11%) and Minor Roads (14%), with less than 2% of the cameras fixed along Motorways and C roads. There are not many fixed cameras on Motorways because variable speed cameras are usually deployed instead to enforce speed limit over a longer distance. Also, approximately 80% of the cameras are located along busier roads in populated urban areas.

As mentioned, one of the major concerns is that earlier camera sites are different

[21]This include Warwickshire, Suffolk, Norfolk, Wiltshire and Swindon.

(a) 24 sites

(b) +1044 sites



(c) +1368 sites

(d) +109 sites

Figure 3.7: Locations of Fixed Speed Cameras across England, Scotland and Wales from 1992 to 2016.

from those receiving installation later. To examine if this is the case, I split the sample into 5 groups (1992 to 1995, 1996 to 2000, 2001 to 2005, 2005 to 2010 and 2010 onwards) according to the year the cameras are installed. I do not find sites that have camera installations first more dangerous than those having installations later. No evident differences are also observed in camera/road characteristics, local authority demographics and labour outcomes. If anything, there seems to be more crashes and injuries for cameras that are installed after 2006. These cameras are often found on roads with higher speed limit. One possible explanation is the change in the guidelines for selecting camera sites. As a precaution, I remove these cameras in my robustness tests but this do not materially affect the results.

Table 3.1: Summary statistics of camera sites across time

	(1) All	(2) 1992 - 1995	(3) 1996 - 2000	(4) 2001 - 2005	(5) 2006 - 2010	(6) 2011 - 2016
Pre-treatment Accident Outcomes						
Accident/100m	0.39	0.36	0.35	0.39	0.50	0.44
	(0.33)	(0.31)	(0.32)	(0.30)	(0.40)	(0.32)
Injuries/100m	0.52	0.48	0.47	0.52	0.64	0.58
	(0.41)	(0.40)	(0.39)	(0.38)	(0.51)	(0.43)
Serious Inj/100m	0.07	0.07	0.07	0.07	0.08	0.07
	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)
Deaths/100m	0.01	0.01	0.01	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Camera/Road Characteristics						
Speed Limit	34.79	37.20	36.53	32.75	34.16	41.05
	(9.33)	(10.92)	(10.73)	(6.72)	(9.39)	(12.35)
A Road	0.73	0.75	0.76	0.69	0.79	0.77
	(0.44)	(0.43)	(0.43)	(0.46)	(0.41)	(0.42)
B Road	0.12	0.11	0.10	0.14	0.08	0.12
	(0.32)	(0.31)	(0.31)	(0.35)	(0.27)	(0.33)
C Road	0.02	0.00	0.01	0.02	0.00	0.00
	(0.12)	(0.06)	(0.12)	(0.15)	(0.06)	(0.00)
Motorway	0.01	0.01	0.01	0.01	0.01	0.04
	(0.09)	(0.08)	(0.11)	(0.07)	(0.08)	(0.19)
Minor Road	0.13	0.14	0.11	0.14	0.13	0.07
	(0.33)	(0.34)	(0.31)	(0.35)	(0.33)	(0.26)
Rural	0.15	0.18	0.21	0.12	0.09	0.14
	(0.36)	(0.38)	(0.41)	(0.32)	(0.28)	(0.35)
Camera/Road Characteristics						
Gross Annual Salary	24245.93	25612.47	24170.51	23423.40	26221.43	23488.91
	(4143.14)	(4317.99)	(3650.61)	(4193.10)	(4143.65)	(2774.21)
Hours Worked	37.86	37.84	37.92	37.87	37.66	37.83
	(0.67)	(0.74)	(0.62)	(0.71)	(0.56)	(0.53)
Job Count	116533.02	118697.66	112070.20	112079.69	143444.02	109272.51
	(95111.37)	(92509.82)	(84887.33)	(93917.90)	(120102.28)	(90747.35)
Job Density	0.88	0.85	0.76	1.01	0.81	0.65
	(2.61)	(0.41)	(0.23)	(3.91)	(0.42)	(0.16)
% of Pop 18 to 25	9.38	9.29	9.11	9.56	9.42	9.73
	(2.59)	(2.44)	(2.35)	(2.92)	(2.00)	(1.82)
Population Size	219312.20	204907.15	221811.27	208318.24	262401.66	254666.19
	(142834.82)	(104695.46)	(135361.03)	(142054.29)	(177959.39)	(178190.47)
Unemployment Rate (%)	6.84	6.33	6.63	6.95	7.25	8.40
	(1.95)	(1.94)	(1.95)	(1.81)	(1.99)	(2.89)
VMT	2797.63	2425.91	2633.17	3271.23	2053.54	1619.41
	(2685.05)	(2370.39)	(2566.26)	(2878.41)	(2368.73)	(1410.15)
Number of Cameras	2548	314	754	1123	301	57

Note: Mean outcomes reported. Standard errors in parenthesis below. Observations further stratified according to the year the camera is installed.

3.6.2 Effects of Speed Cameras on Accidents

Baseline Estimates

Table 3.2 presents a set of baseline estimates from equation (1) that captures the effect of speed enforcement cameras on on various accident outcomes 500 metres left and right of the camera, including number of Accidents, Slight Injuries, Serious Injuries and Deaths. Due to space constraints, I only report results from Poisson regressions. Findings from Negative Binomial regressions in Table 3.9 in Data Appendix and are fairly similar. Only the coefficients (γ) for key estimate T_{it} are reported. To interpret these coefficients, I compute the semi-elasticity (%Δ) by taking the exponential of γ before subtracting by 1. The absolute reductions in collision outcomes (Absolute) are by computed by multiplying %Δ with the pre-treatment mean of collision outcomes. Only ever-treated sites are analysed except of

Column (6). In short, I am comparing changes in collision outcomes for sites after camera installations with sites that have camera installations in the future. The sample is smaller for Serious Injuries and Deaths. This is because there are several sites that experience no fatalities or severe injuries over the sample period and these sites are removed from the analysis.

Moving from left to right, additional covariates are included in the estimation. In column (1), I analyse the entire sample of speed cameras from 1992 to 2016 and limit the analysis to sites that I have a full set of control variables in the column (2). In both specifications, I include site and year fixed effects but do not add any control variables. I observe that enforcement cameras not only reduce the number of crashes, but also abate the severity of collisions. It is also comforting to observe that results are very consistent across the two columns, suggesting that the reduced sample with control variables is fairly representative.

Next, I include a vector of time-variant local authority (LA) characteristics to partial out regional specific shocks that could correlate with the camera installations and affect collision outcomes. This include demographic (population size and % of population between 18 to 25) and labour characteristics (gross annual salary and working hours). Controlling for these differences across LA has an inconsequential effect on the estimates. Subsequently, I control for the annual average vehicle miles travelled (VMT) as more driving could induce more accidents. Estimates remain fairly stable. I further include a number of weather controls including temperature and wind speed. The concern is whether bad weather shocks, which could induce more accidents, are correlated with camera installations. Doing so significantly reduces the sample by more than two-third due to missing data but again this do not materially affect the estimates.

In column (6), I include a sample of non-camera sites[22] despite meeting the selection guidelines (for more information refer to section 2). The rationale is to understand the bias from incorporating non-treated sites based on some matching-on-observables strategy frequently adopted in the previous literature. I observe that estimated enforcement effects are much smaller. This is consistent with the idea that untreated sites experience a large fall in collision outcomes even without camera installation. Because this surge in accidents is deemed to be transient, collision outcomes could fall even in the absence of camera enforcement. This explains why local camera partnerships choose not to install cameras around these sites. This

[22]To create a sample of non-camera sites, I first place random points along major roads (A & B roads) that are at least 2,000 metres from one another and 2,000 metres from the nearest speed camera. Following that I calculate the yearly collision, injuries and death counts within 500 metres from these random points. I only retain sites with more than 4 killed and serious injuries (KSI) and 8 personal injury collisions in a 3 year rolling window. In total, I find 694 sites that meet the selection criterion but are not treated.

result shows that using non-camera sites coul underestimate the enforcement effects of speed cameras.

Furthermore, as mentioned, sites that receive installations later could be different from those earlier treated sites. Thus, in column (7), I restrict the reference groups to just recently treated cameras by excluding any observations more than 3 years before and after the camera is installed. To illustrate, this is equivalent of removing the shaded areas in Figure 3.4 from estimation. Like before, estimates remain comparable, suggesting that the differences between earlier and later treated sites are not confounding the estimates.

Overall, I document that speed cameras not only attribute to significant reductions in the number of collisions, but also abate the severity of the crashes. Results are fairly steady to the addition of controls. I observe substantial decreases for the various accident outcomes significant at 1% level. After an enforcement camera is installed, the number of collisions are, on average, 17% to 39% lower, representing an absolute reduction of 0.89 to 2.36 per kilometre per annum. The counts of Slight injuries also decline by between 1.19 and 2.87 per kilometre per annum, which corresponds to a 17% to 38% decrease. There are between 0.25 and 0.58 less serious injuries surrounding the camera, equivalent to a 28% to 55% fall from pre-treatment levels. The largest effects are documented for traffic fatalities. There are approximately 0.08 to 0.19 less fatalities per kilometre, which represents a substantial 58% to 68% decline[23] from pre installation levels.

Robustness & Alternative Explanations

Table 3.3 summarizes a battery of robustness tests that addresses concerns that earlier estimates could be spuriously driven by other factors.

A1. Traffic Displacement: One issue is whether the installation of cameras induce drivers to switch to non-camera roads. Therefore, the reduction in accidents could be due to less traffic rather than camera enforcement[24]. To mitigate the possibility that traffic displacement is driving the estimates, I limit my analysis to a sub-sample of Motorways and A-Roads. The rationale is that there is less traffic displacement along these major roads because there are less alternative routes available. Results in Columns (1) are fairly similar compared to before, indicating that enforcement effects documented earlier are not driven by lighter traffic.

[23]This is because often there are very little reported deaths on roads, which is why the small estimate could generate significant changes.

[24]The straightforward solution is to include traffic as a control. This, however, will not be advisable as traffic is likely to be a "bad" control. The implementation of speed camera is likely to reduce traffic flow by displacing them to neighbouring unmonitored roads. Moreover, detailed road level traffic data is only available for a small sub-sample of roads.

Table 3.2: Effects of Speed Camera on various accident outcomes within 500 metres from Camera using Poisson Regressions

	(1) All	(2) Baseline	(3) Demo	(4) VMT	(5) Weather	(6) Non-CAM	(7) -3,+3
Accidents	-0.469[a]	-0.488[a]	-0.268[a]	-0.243[a]	-0.184[a]	-0.095[a]	-0.222[a]
	(0.009)	(0.011)	(0.016)	(0.017)	(0.028)	(0.016)	(0.017)
Obs	66868	25720	25720	25720	7383	35929	9841
Absolute	-2.11	-2.36	-1.44	-1.32	-0.89	-0.82	-1.13
% Δ	-37.43	-38.65	-23.54	-21.55	-16.78	-9.09	-19.88
No.of CAM	2481	1555	1555	1555	659	2249	1481
Slight	-0.412[a]	-0.483[a]	-0.278[a]	-0.253[a]	-0.185[a]	-0.057[a]	-0.207[a]
	(0.010)	(0.013)	(0.019)	(0.019)	(0.038)	(0.019)	(0.021)
Obs	57224	21355	21355	21355	5483	31564	8175
Absolute	-2.25	-2.87	-1.82	-1.67	-1.19	-0.64	-1.33
% Δ	-33.74	-38.29	-24.30	-22.34	-16.87	-5.56	-18.70
No.of CAM	2123	1294	1294	1294	518	1988	1223
Serious	-0.788[a]	-0.747[a]	-0.454[a]	-0.414[a]	-0.326[a]	-0.326[a]	-0.373[a]
	(0.015)	(0.022)	(0.033)	(0.034)	(0.074)	(0.028)	(0.042)
Obs	63280	23650	23650	23650	6539	33823	8306
Absolute	-0.58	-0.57	-0.39	-0.37	-0.25	-0.35	-0.33
% Δ	-54.55	-52.63	-36.49	-33.93	-27.81	-27.84	-31.15
No.of CAM	2346	1428	1428	1428	572	2115	1240
Deaths	-0.956[a]	-1.071[a]	-1.029[a]	-1.018[a]	-1.124[a]	-0.761[a]	-0.858[a]
	(0.041)	(0.073)	(0.116)	(0.119)	(0.209)	(0.093)	(0.153)
Obs	42924	11394	11394	11394	2787	18765	2843
Absolute	-0.08	-0.12	-0.12	-0.12	-0.15	-0.09	-0.19
% Δ	-61.57	-65.75	-64.28	-63.85	-67.51	-53.27	-57.59
No.of CAM	1591	683	683	683	220	1155	426
CAM FE	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓
Demographics			✓	✓	✓	✓	✓
VMT				✓	✓	✓	✓
Weather					✓		

Note: Each reported coefficient is the γ from a different Poisson regression estimated using Maximum likelihood. Dependent variable is the annual Y count where Y=accident, injuries, serious injuries and deaths 500m left and right of camera. Absolute is the number of reductions in accident outcomes computed by multiplying the % Δ with the pre-treatment mean of Y. % Δ is the proportional change (semi-elasticity) of collision outcomes after treatment and is computed by taking $exp(\gamma)-1$. In Column (1), I include the entire sample of cameras. In Column (2), I restrict the sample to sites that I have full set of co-variates. In Column (3), I control for population size, % of 18 to 25, Gross Annual Pay & hours worked. In Column (4), I control for the annual average vehicle miles travelled (VMT). In Column (5), weather controls are added into the specification. In Column (6), I include a sample of non-camera sites that are eligible for camera installations. In Column (7), I constraint the analysis to observations just 3 years before and after from the year of installation. Sandwich (robust) standard errors are reported in the parentheses. [c] $p<0.10$, [b] $p<0.05$, [a] $p<0.01$

A2. Change in Selection Rules: In 2006, there are major changes on how sites are selected for camera enforcement. The problem is whether newer guidelines cause these latter sites to be less comparable. Thus, in column (3), I remove sites that receive installations from 2006 onwards. Removing these later treated sites appear to reduce my estimates marginally but inconsequentially. Cameras installed before 2006 appear to reduce more accidents but much lesser slight and serious injuries, suggesting that the newer selection rules are more effective in identifying dangerous roads for camera enforcement.

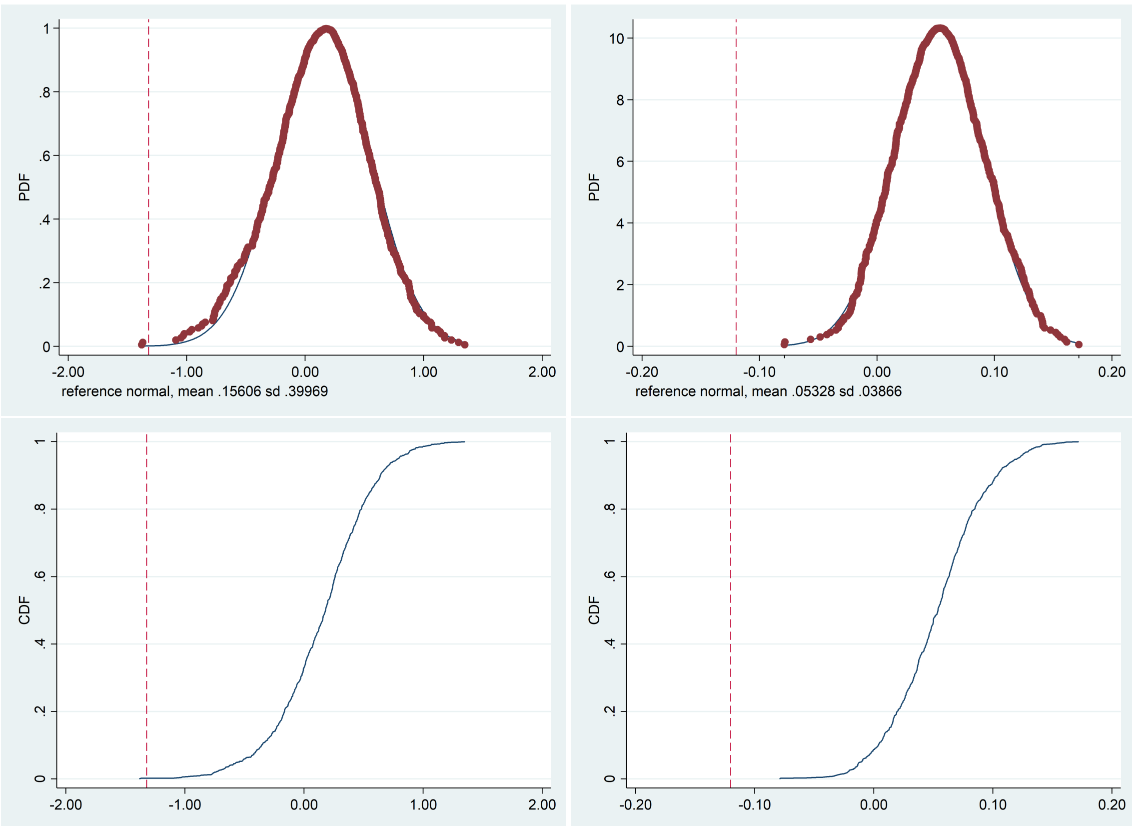
A3. Omitted Variable Bias: Next, I rely on the rich information associated with each camera to further mitigate observable differences. To do so, I match each each site with another site based on the following rules: (1) within 5 kilometres from one another; (2) same rural-urban classification; (3) same road type (A, B, C, Minor or Motorways); (4) similar speed limits; (5) within 5 years from one another in installation dates; (6) within 70% - 130% in pre-treatment collision outcomes. The objective is to benchmark each site with the most similar yet-to-be treated site. This rigorous matching process reduce the sample to 214 sites. I include both site fixed effects and pair-match fixed effects interacted with years in the analysis. In other words, I am now computing changes in collision outcomes before and after the camera is installed, and benchmarking these changes in collision outcomes with the closest reference site matched based on observable characteristics. This procedure does not materially affect the estimates as significant reductions across various collision outcomes are observed.

A4. Endogenous Timing: One of the main concerns is whether a camera is only introduced after a sharp increase in collisions, inflating enforcement effects. While the timing of introduction may be endogenous, the timing for switching cameras off is less affected by collision trends. Due to budget cuts, some local partnerships, including Avon and Somerset and West Midlands, are forced to turned off their cameras [25]. Although these cameras are no longer in operation, local government usually leave them in place as "dummy" cameras to deter speeding. To ensure that these switched-off cameras are benchmarked against comparable reference groups, I further restrict the control group to sites in adjacent local authorities. This reduces the sample to 174 cameras, with 55 cameras in switched-off areas. Results summarized in Column (7) and (8) indicate that enforcement effects are much weaker as compared to before. Although these "dummy" cameras still reduce accidents and slight injuries, the absolute effects are less than half compared to operating cameras.

[25]There are more areas that have shut down their devices. This include Cleveland, Durham, Northamptonshire and North Yorkshire. Avon and Somerset and West Midlands are selected because I can accurately determine the locations of these switched-off devices and the local partnerships publicly announce that these cameras are out of operation.

Moreover, these switched-off devices no longer reduce serious injuries and deaths. These findings suggest that some informed drivers are no longer adhering to the speed limits and their reckless driving behaviors are diluting enforcement effects.

A5. Spurious Timing: Next, I mitigate the issue whether enforcement effects are spuriously documented outside treatment periods. To do so I generate 1,000 random treatment dates at least 5 years before the cameras are installed. For instance, if a camera is installed in 1997, the random generated year will be between 1980 and 1992. These placebo regressions are computed using OLS as it is too computationally intensive for MLE. Cumulative probability and probability density of the estimated γ from 1,000 different placebo regressions for accidents and deaths estimated using OLS are plotted in Figure 3.8. Dash lines denote the estimated effects of speed cameras from the preferred specification from Column (5) of Table 3.2. It is comforting to observe that all but two estimates from these 1,000 placebo regressions are larger (more positive) than the treatment effects, increasing the confidence that earlier findings are not spuriously driven in pre-installation periods.



(a) Accidents (b) Deaths

Figure 3.8: 1,000 Placebo Regressions with random generated treatment dates before camera installation on Accidents & Deaths estimated using OLS: Probability Density Function (PDF) top, Cumulative Density Function (CDF) bottom. Dash line denotes estimates from Column (6) of Table 4.3.

Table 3.3: Robustness Tests

	Major Roads	No 2006-2016	Matched Pairs	Switch Offs
Accidents	-0.248[a]	-0.243[a]	-0.165[a]	-0.224[a]
	(0.018)	(0.019)	(0.028)	(0.082)
Obs	19152	19936	3304	2814
Absolute	-1.58	-1.20	-1.73	-0.54
% Δ	-21.95	-21.61	-15.24	-20.06
No.of CAM	1159	1208	214	174
Slight	-0.245[a]	-0.261[a]	-0.184[a]	-0.269[a]
	(0.020)	(0.022)	(0.033)	(0.104)
Obs	16566	16143	3052	2814
Absolute	-1.89	-1.56	-2.13	-0.84
% Δ	-21.75	-22.98	-16.84	-23.59
No.of CAM	1004	982	196	174
Serious	-0.412[a]	-0.389[a]	-0.174[b]	0.113
	(0.036)	(0.041)	(0.080)	(0.507)
Obs	18382	18405	2348	972
Absolute	-0.40	-0.32	-0.34	0.00
% Δ	-33.75	-32.22	-15.97	12.01
No.of CAM	1110	1114	200	59
Deaths	-0.987[a]	-0.822[a]	-1.064[b]	-0.068
	(0.125)	(0.149)	(0.480)	(0.247)
Obs	9958	7988	230	2427
Absolute	-0.12	-0.11	-0.39	-0.02
% Δ	-62.74	-56.06	-65.50	-6.55
No.of CAM	596	480	64	150
CAM FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
LA Controls	✓	✓	✓	✓
Pair-FE* Year			✓	

Note: Each reported coefficient is the γ from a different poisson regression. Dependent variable is the annual Y count where Y=accident, injuries, serious injuries and deaths 500m left and right of camera. The specification adopted is similar to that of Column (4) in Table 3.2. In columns (1), I restrict the analysis to cameras in A-Roads and Motorway to alleviate the effects of traffic displacement on collisions. In columns (2), I remove cameras installed from 2006 onwards as they could be different from the other cameras. In columns (3), I match each site with another site based on location, pre-treatment accident outcomes and various road characteristics. I exploit the variation now between two speed cameras by including pair-fixed effects interacted with year fixed effects. In columns (4), I restrict the analysis to areas that turn off their cameras and counties that are contiguous to these areas. Sandwich (robust) standard errors are reported in the parenthesis. [c] $p<0.10$, [b] $p<0.05$, [a] $p<0.01$.

Effects across Road Types, Speed Limits & Time

Next, I allow the effectiveness of speed cameras to vary across different road types, speed limits and over time with the following specification:

$$E(Y_{ijt}) = exp(\alpha_i + \gamma_{\mathbf{w}}(\mathbf{T_{it}} * \mathbf{H'_w}) + X'_{jt}\phi + \theta_t + \varepsilon_{ijt}), \tag{3.2}$$

where the $\mathbf{H'_w}$ represents a vector of binary variables that equals to unity denoting the different:

1. Speed limits (30,40,50,60 & 70)
2. Road Classes (Motorway, A, B & Minor)
3. Years after treatment (1,2...10 years after installation)

In short, I am allowing enforcement effects to vary across these characteristics. Specifications for speed limits and road types are summarized in Panel A and B of Table 3.4 respectively. Only the key estimates γ_w are reported.

From Panel A, although I find significant improvement in road safety across different speed limits, the largest enforcement effects are documented along roads with higher speed limits. Specifically, the number of collisions are, on average, 50% lower along 60mph roads after a speed camera is installed. This is much larger than the 22% reduction along 20mph roads. The number of serious injuries and deaths are 81% and 95% lower after speed cameras are installed along 60mph roads and nuch smaller effects of around 36% and 41% are observed for serious injuries and deaths respectively along 20mph roads. There are several explanations to these findings. First, drivers along the lower speed limit roads are already commuting slowly and reductions in speed achieved by cameras do not matter much in reducing the gravity of collisions. Second, attenuated enforcement effects for more binding speed limits suggest that drivers may be forced to hastily drop speed to avoid fines. This could cause more accidents via kangaroo effects.

Panel B summarizes the results of camera enforcement effects on different road types. Motorways are inter-city major roads for long distance travelling. A-Roads are slightly less important compared to Motorways but can still be considered trunk roads that provide large scale transport links. B-Roads are slightly smaller linkage roads for traffic between A-Roads and Minor roads. Minor Roads are smallest roads that connect local traffic, linking an estate/village with the larger road links. I do not observe stark differences in enforcement effects across the different road types. This is except for Motorway in which no significant reduction in slight injuries and deaths are reported. This is likely due to a sample issue as only 1% of the cameras are found along Motorways.

Table 3.4: Heterogeneous effects of Speed Camera on various accident outcomes across road types and speed limits

	Panel A: Speed Limit			
	Accidents	**Slight**	**Serious**	**Deaths**
Speed Limit 20	-0.247[a]	-0.130[b]	-0.446[b]	-0.514
	(0.066)	(0.066)	(0.211)	(1.007)
	-21.91	-12.23	-35.96	-40.19
	-1.34	-0.92	-0.39	-0.07
Speed Limit 30	-0.227[a]	-0.241[a]	-0.356[a]	-0.891[a]
	(0.017)	(0.020)	(0.035)	(0.124)
	-20.31	-21.43	-29.98	-58.98
	-1.24	-1.61	-0.32	-0.11
Speed Limit 40	-0.371[a]	-0.358[a]	-0.657[a]	-1.277[a]
	(0.036)	(0.040)	(0.067)	(0.238)
	-30.99	-30.10	-48.14	-72.11
	-1.89	-2.25	-0.52	-0.13
Speed Limit 50	-0.213[a]	-0.217[a]	-0.538[a]	-1.188[a]
	(0.057)	(0.063)	(0.138)	(0.285)
	-19.15	-19.49	-41.62	-69.52
	-1.17	-1.46	-0.45	-0.13
Speed Limit 60	-0.697[a]	-0.598[a]	-1.645[a]	-3.060[a]
	(0.162)	(0.203)	(0.275)	(0.755)
	-50.21	-45.01	-80.70	-95.31
	-3.07	-3.37	-0.87	-0.18
Speed Limit 70	-0.314[a]	-0.291[c]	-0.709[b]	-1.505[a]
	(0.115)	(0.163)	(0.286)	(0.307)
	-26.94	-25.21	-50.79	-77.80
	-1.64	-1.89	-0.55	-0.14
Obs	24871	20522	22833	11004
No.of CAM	1503	1243	1378	659
	Panel B: Road Type			
A Road	-0.232[a]	-0.243[a]	-0.393[a]	-0.990[a]
	(0.017)	(0.019)	(0.035)	(0.121)
	-20.70	-21.60	-32.49	-62.84
	-1.26	-1.62	-0.35	-0.12
B Road	-0.325[a]	-0.345[a]	-0.514[a]	-1.289[a]
	(0.040)	(0.053)	(0.077)	(0.301)
	-27.77	-29.19	-40.20	-72.44
	-1.70	-2.19	-0.43	-0.13
Minor Road	-0.317[a]	-0.340[a]	-0.645[a]	-1.346[a]
	(0.055)	(0.094)	(0.096)	(0.344)
	-27.20	-28.82	-47.52	-73.98
	-1.66	-2.16	-0.51	-0.14
Motorway	-0.262[a]	-0.105	-0.496[c]	-0.135
	(0.077)	(0.141)	(0.283)	(0.470)
	-23.05	-9.95	-39.13	-12.62
	-1.41	-0.75	-0.42	-0.02
Obs	25720	21355	23650	11394
No.of CAM	1555	1294	1428	683

Note: Each reported coefficient is the γ_w from a different poisson regression from equation 3.2 estimated using maximum likelihood. Dependent variable is the annual Y counts where Y=accident, injuries, serious injuries and deaths 500m left and right of camera. I allow the effects to vary across different speed limits and road types in Panel A and B respectively. The specification adopted is similar to Column 4 of Table 3.2. Sandwich (robust) standard errors are reported in parentheses. [c] $p<0.10$, [b] $p<0.05$, [a] $p<0.01$.

Next, I examine the effectiveness of speed cameras over time. Results are summarized in Figure 3.11. Results reveal that cameras remain effective and in fact become more potent in reducing collisions and fatalities over time. Weaker effects in the beginning suggest that some drivers could be unfamiliar with the locations of camera and abruptly drop speed to avoid fines. This could cause collisions and dilute enforcement effects. Over time, drivers learn about these locations and are less prone to reckless braking, explaining stronger enforcement effects.

Effects over Distance

A major drawback of fixed speed cameras is that they might cause more accidents further away from the camera. This is know as "kangaroo" effects - when drivers abruptly halt in response to camera signs to avoid being fine, or accelerate beyond surveillance. To precisely capture how the effects change with distance from the camera, the following specification is estimated:

$$E(Y_{ijt}^{k-100,k}) = exp(\alpha_i^{k-100,k} + \gamma \mathbf{T_{it}^{k-100,k}} + X_{jt}^{\prime k-100,k}\phi + \theta_t^{k-100,k} + \varepsilon_{ijt}) \tag{3.3}$$

where k represents the various distance bandwidths (eg. 0 to 100m, 100m to 200m... 1900m to 2000m) up to 2 kilometres left and right of the camera. In brevity, I am estimating the enforcement effects for every 100m bandwidth to identify how enforcement effects vary moving away from the camera. I achieve this by running stratified regressions for every k and $k - 100$ bandwidth for $k \in 100, ...1900, 2000$ metres. If the effects are highly localised, I expect γ_k to be more negative as k is smaller. If there are displacement of accidents, I would expect γ_k to be positive outside camera surveillance.

Figure 3.10 summarizes the estimated effects of speed enforcement cameras at every 100m from the camera using Poisson regressions. Results from Negative Binomial models are summarized in Figure 3.12 in Data Appendix. Like before, results are fairly similar across these two models. Precisely, I am capturing the change in accident outcomes every 100 metres. Every dot denotes γ_k for a different distance bandwidth between k to $k - 100$ from the camera where $k = 100, ...2000$. Dashed lines denote the 95% confidence interval. The coefficients can be interpreted as number of accident outcomes per 100 metre.

Unsurprisingly, I find localised enforcement effects around the camera that dissipate quickly across distance. Reductions are largely around 0 to 500 metres around the camera and strongest effects are reported closest to the camera. This result is fairly consistent across the different accident outcomes. Beyond 700 metres from the device, fixed speed cameras are no longer able to enhance road safety. Moving further away at around 1500 metres from the camera, there are suggestive evidence

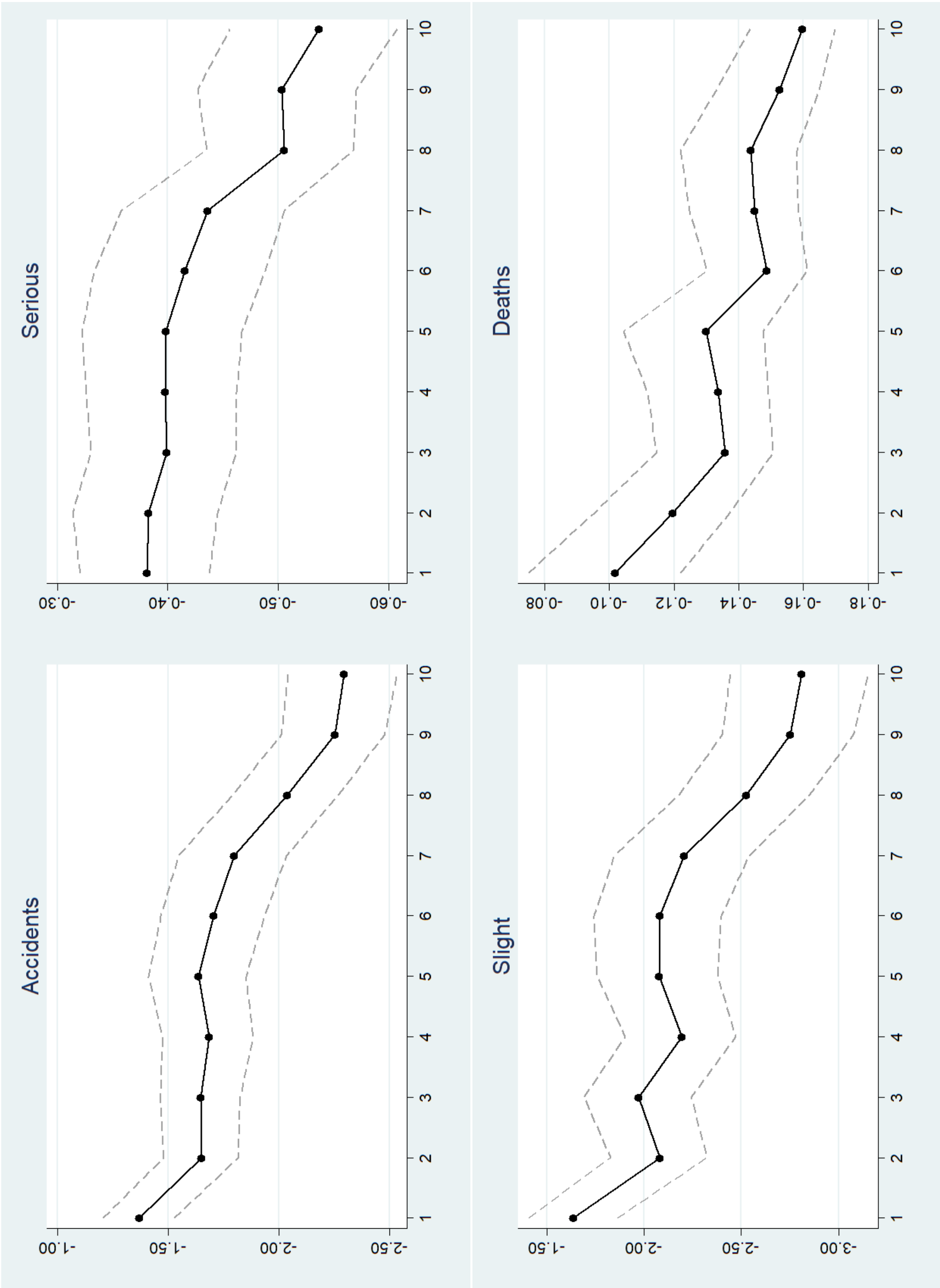


Figure 3.9: Effects of Speed Camera on Accidents 0 to 500m from camera across years from Poisson Regressions. Effects can be interpreted as per km. Dots denote the Coefficients (γ_w) from the estimation of Equation 3.2 where H'_w represents a vector of dummies representing years from installation. 95% Confidence Intervals are denoted by the dashed lines.

of kangaroo effects as I report small rebound in the number of collisions, serious injuries and deaths. A small proportion of drivers could have speed up beyond the surveillance of cameras, inducing more collisions post implementation. However, these effects are quite small compared to the enforcement effects and are too imprecisely estimated to be statistically significant.

3.6.3 Welfare Analysis

This section reports a welfare analysis on speed cameras to understand whether these devices improve social welfare. The costs include the fixed and operating costs of camera and the time delays incurred by bypassing drivers, while the benefits include the savings from less collision, injuries and fatalities. Fines from speeding tickets are not considered as the government could redistribute these revenues through public spending. The parameters considered are summarized in Table 3.5.

For the benefits, I rely on the savings per traffic accidents, injuries and deaths computed by Department of Transport (DfT)[26]. These values account for both (1) casualty-related costs (loss output, medical and ambulance, human costs) and (2) accident-related costs (property damage, insurance and administrative and police costs). Total savings are computed by multiplying earlier estimates on reductions with the savings on per capita or accident basis.

For the costs, I obtain approximated time delays from speed cameras from Gains *et al.* (2005). Speed is around 10kmh slower after the camera is installed. Taking the average speed limit of 58kmh[27] (30mph) and a distance of 1km around the camera, drivers incur a delay of around 0.2 minutes (or around 12 seconds) whenever they bypass a speed camera. According to the average traffic flow along roads provided by DfT, I estimate that there are approximately 3,600 vehicles bypassing each camera every day, corresponding to around 1.3 million vehicles annually. Assuming the average occupancy per car is 1.5, time delays incurred by all bypassing vehicles amount to more than 7,000 hours every year[28].

To compute loss of income from time delays, I rely on the estimates on the value of time savings and the purposes of journey from DfT (2014). I assume that 5.0% of the journeys are made for work, 20% are for commuting towards or from the place

[26]For more information, refer to https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/254720/rrcgb-valuation-methodology.pdf

[27]Since most of the speed cameras impose a 30mph speed limit, time delays will be computed based on the scenario that drivers commute, on average, at a speed limit of 30mph before camera installation. Drivers are assumed to slow down around 500m left and right of the camera. Time delays per driver per trip is therefore approximately equal to $\frac{Distancearoundcamera}{OriginalSpeed-Reductions}$.

[28]This is likely to over-estimate the time delays given that accidents can cause traffic bottlenecks that can increase travel time.

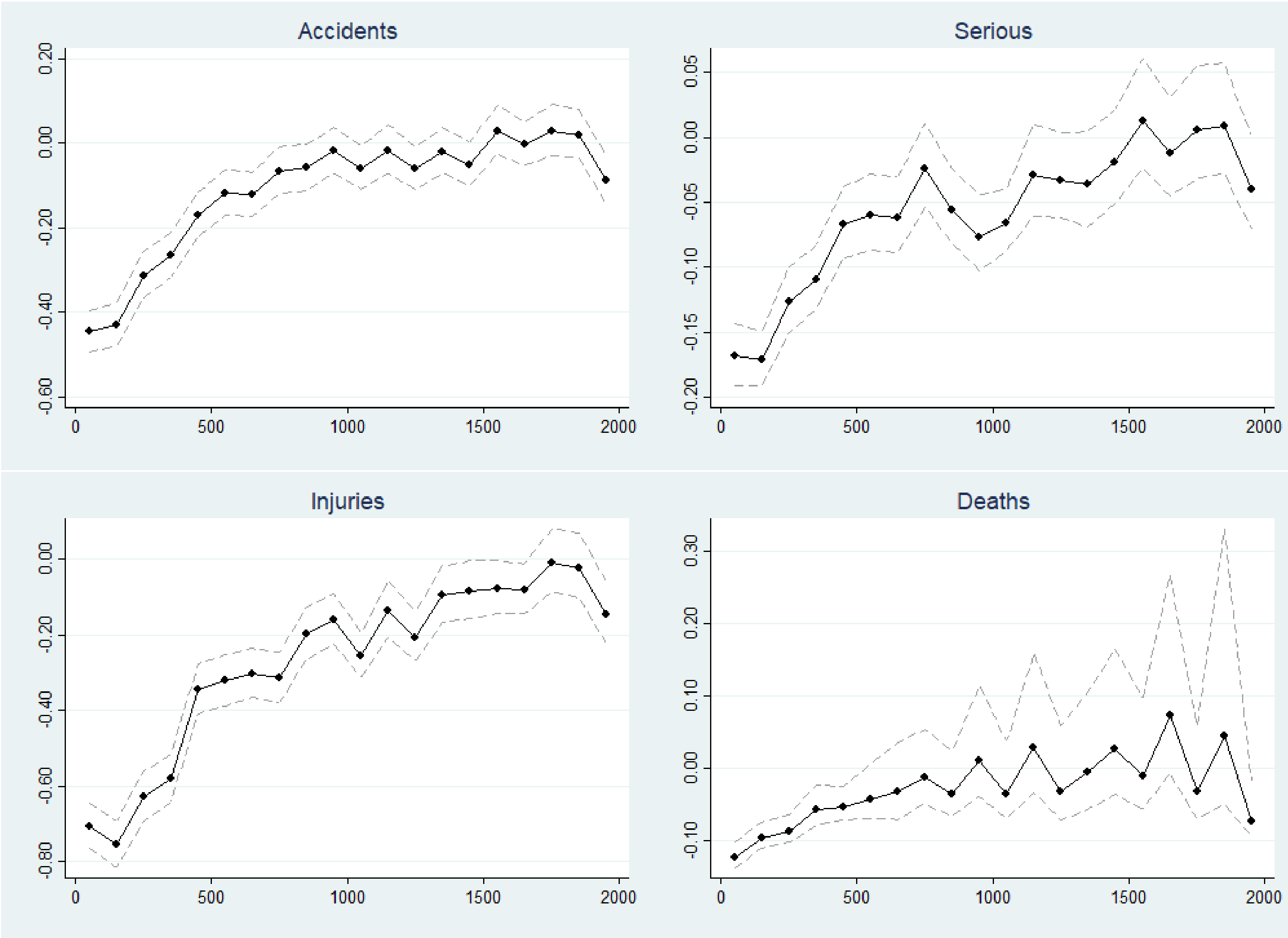


Figure 3.10: Effects of Speed Camera on Accident Outcomes at 100m Intervals (left and right of camera) from Poisson Regressions. Each dot represents the coefficient for every 100 metres distance bandwidth. Effects can be interpreted as per 100 metre. Coefficients (γ_k) are obtained from the estimation of Equation 3.3. Specification is similar to that of Column (4) in Table 3.3. Dot indicates that the coefficient is statistically significant at least at 10% level and cross indicates otherwise. 95% Confidence Intervals are denoted by the tails.

of work, and the rest of the 75% of the journeys are for non-work purposes (e.g leisure trips). The values of travel time (VOT) per hour are much higher at £26.42 for work, and £6.65 and £5.90 for commuting and non-work purposes respectively. Multiplying the total number of hours delayed with the estimated proportion of journeys for various purposes and their respective VOTs, the total loss of income due to delays per speed camera amounts to around £50,000 per annum. Taking the cost of installing a fixed speed camera at £309,000 and the operating and maintenance cost at around £12,500 per annum, the total cost of installing a speed camera per annum is around £121,000.

Table 3.5: Cost-Benefit Analysis per speed camera across Great Britain

Parameter	Source	Value per Unit	Net Cost/Benefit/Year
	Savings from avoiding Accidents		
Damage-only	DfT(2015)	£2,142	$1.13 \times £2,142 = £2,420$
Slight Injuries	DfT(2015)	£15,450	$1.33 \times £15,450 = £20,549$
Serious Injuries	DfT(2015)	£200,422	$0.33 \times £200,422 = £66,139$
Deaths	DfT(2015)	£1,783,556	$0.19 \times £1,783,556 = £338,876$
(A)Total Benefits		≈	**£430,000**
	Time Delays		
Speed Reductions	Gains(2005)	9.65kmh	$\frac{1km}{58kmh-9.65km} - \frac{1km}{58kmh} \approx 0.206mins$
Average No. of Cars	DfT(2016)	3,600 cars/day	$3,600 \times 365 = 1,314,000$
Average Occupancy/car	DfT	1.56/car	
Total Time Loss (h)			$1,314,000 \times \frac{0.206}{60} \times 1.56 \approx 7,038hrs$
Journey By Purpose	DfT (2014)		
Work		5.0% (£26.42/h)	$5.0\% \times 7,038 \times £26.42 = £9,297$
Commuting		20.3% (£6.65/h)	$20.3\% \times 7,038 \times £6.65 = £9,501$
Others		74.7%(£5.90/h)	$74.7\% \times 7,038 \times £5.90 = £31,018$
(B)Loss of Income from delays		≈	£50,000
	Cost of Cameras		
Fixed Cost	Parliament(2008)	£50,000	$£50,000 \times 1.18 = £59,000$
Operating Cost	Hooke(1996)	£8,560	$£8,560 \times 1.45 = £12,441$
(C)Total Camera Costs per year		≈	£71,000
(D)Total Costs		≈	**B+C=£121,000**
Net Costs/Benefits		≈	**A-D=+£309,000**

Note: All the dollar values are adjusted to 2015 price levels. Estimates on savings from avoiding accidents are obtained from Column (7) of Table 3.2. Fixed Costs include planning, signage, installation and procurement, and other fixed costs. Operating costs include operation, administrative, maintenance, publicity and liaison costs that recurs annually. These figures are obtained by averaging across a sample of cameras installed 10 study areas across UK in financial year 1995/96.

It is important to highlight that the estimated benefits from this analysis are likely to underestimate the actual benefits realized as I did not factor in other non-pecuniary perks. These include environmental benefits from slower travelling speed that could save more fuel, reduce emissions and improve health outcomes (van Benthem, 2015). Enforcement cameras could also enhance crime intelligence as images from these devices could help to solve other crimes (Hooke *et al.*, 1996). Even without considering these perks, each speed camera generates net benefits of £309,000 per annum, which is more than twice the cost of implementing a speed camera.

Next, I conduct a battery of sensitivity analyses to the welfare estimates in Table

Table 3.6: Sensitivity Analysis of Welfare Estimates

Net Benefits	(1) £200,000	(2) £281,000	(3) £280,000	(4) £270,000	(5) £222,000	(6) £35,000
25% ↓ in effectiveness	✓					✓
50% ↑ in traffic		✓				✓
25% of journey work trips			✓			✓
50% ↑ in cost				✓		✓
25% ↓ in VSL					✓	✓

Note: All the dollar values are adjusted to 2015 price levels. In Column (1), I reduce the effectiveness of speed cameras by 25%. In Column (2), I increase the traffic that bypasses speed cameras by 50%. In Column (3), I change the composition of trips such that 25% of the journeys made now are for work purposes. In Column (4), the cost of installing and operating fixed speed cameras are now 50% higher. In Column (5), I reduce the cost of death from collisions by 25%.

3.6. Different assumptions on the effectiveness, time delay and cost associated with speed cameras do not affect the main findings. If I assume that the cameras are 25% less effective as before, I still observe positive benefits fall to around £200,000. If the amount of traffic bypassing speed cameras goes up by 50%, the net benefits dipped to around £281,000. If the percentage of trips made for work purposes increases to 25%, the benefits reduce to £280,000. If the cost of installing and operating a speed camera increase by 50%, the benefits fall to around £270,000. Finally, cutting the value of statistical life (VSL) from traffic collisions by 25% reduce the benefits considerably to around £222,000. This is understandable considering the bulk of the welfare from speed cameras is from the reduction of fatalities from accidents. In the worst case scenario when I consider that all the following conditions happen, I still observe substantial benefits of around £35,000, which represents around 30% of the cost. All these results suggest that installing speed cameras improve social welfare.

3.7 Conclusion

This paper utilizes micro geo-coded dataset on traffic accidents to evaluate the effectiveness of speed enforcement cameras. These devices deter reckless driving on roads particularly prone to collisions by imposing fines when drivers exceed speed limits. In contrast to earlier literature, this paper addresses the selection bias by analyzing only sites that will ever have a speed camera installed. The empirical strategy is a quasi-experimental difference-in-difference framework that relies on comparing accident outcomes before and after a speed camera is installed with other sites that will experience installation in the near future.

Assuming a linear relationship between cameras and collisions, putting another 1,000 speed cameras on roads could reduce approximately 1130 crashes, preventing around 330 serious injuries and in turn, saving 190 lives every year and generating benefits up to £309 million. These results remain robust across a range of specifications that relaxes the identification strategies. Dwelling further, however, reveal that these effects are largely localised within 0 to 500 metres from the camera and there are suggestive evidence of a rebound in collisions further away from the camera. This illustrates the possibility of drivers speeding up beyond the surveillance of cameras and inducing more accidents. Nevertheless, simple cost-benefit analysis reveals that the perks from installing a camera are much larger than the cost of cameras, suggesting that these devices improve social welfare. But with technology advancement, newer prototypes, such as mobile and variable speed cameras, should be considered to circumvent the weaknesses associated with fixed speed cameras to more effectively deter speeding.

3.8 Data Appendix

Table 3.7: List of Variables

Variable	Source	Description
Dependent Variable (Y_{ijt})		
Accident	STATS19	Number of Accidents at site i in LA j in year t
Slight Injuries	STATS19	Number of Slight Injuries at site i in LA j in year t
Serious Injuries	STATS19	Number of Serious Injuries at site i in LA j in year t
Deaths	STATS19	Number of Deaths at site i in LA j in year t
Local Authority Characteristics(X'_{jt})		
Gross Annual Salary	Annual Labour Force Survey	Average Gross annual salary at LA j
Hours worked	Annual Labour Force Survey	Average number of hours worked in LA j
Job Density	Nomis	Number of Jobs per unit area of LA j (hectare)
% of 18 to 25	Nomis Population Estimates	Percentage of population aged 18 to 25 in LA j
VMT	DfT	Annual average vehicles miles travelled in LA j
Max Temperature	MIDAS	Annual average max air temperature in LA j
Min Temperature	MIDAS	Annual average min air temperature in LA j
Wind Speed	MIDAS	Annual average wind speed in LA j
Camera/Road Characteristics		
Speed Limit	-	Binary variable denoting whether speed camera in site i has a speed limit of l where l=30,40,50,60 or 70
Road Type	-	Binary variable denoting whether speed camera in site i in road type r where r=Motorway, A, B, C or Minor
Rural	ONS Rural Urban 2011 classification	Binary variable denoting whether speed camera in site i is in rural area, otherwise it is located in urban area

Table 3.8: Review of Existing Literature on Speed Camera Evaluation

Authors	Dataset	Methodology	Results
Chen *et al.* (2002)	12 Photo Radar Programs (PRP) over 22km along a highway in British Columbia, Canada, 2 years pre post	EB	2.8 km/h (3%) ↓ in speed; 7% ↑ in traffic; overall 16% ↓ in collisions across entire corridor with positive spillover effects at non PRP locations. Unlike cameras, drivers are unsure of PRP deployment
Christie *et al.* (2003)	101 Mobile Speed cameras in South Wales, UK, 3 years pre, 1 year post	BA with circle and route based measures	50% (1.8) ↓ in injury crashes; effects are within 300 to 500m and no longer significant beyond; effects are stable across time and similar for 30mph and 60-70mph roads
Cunningham *et al.* (2008)	Mobile Speed cameras in Charlotte, North Carolina, US, 5 years	BA with comparable reference groups constructed based on characteristics	10% ↓ in collisions, decrease in travelling speed
Elvik (1997)	64 Speed cams in Norway,	EB	20% ↓ in injury crashes; 12% ↓ in property crashes; effects are largely driven by road sections with warrants - a certain level of crash and speed limit for the use of speed limit.
Gains *et al.* (2004, 2005)	2,300 speed cameras across 23 areas across UK, 3 years pre post	BA & EB	6% ↓ in speed; 91% ↓ in excessive speeding (>15mph); 22% ↓ in collisions; 42% ↓ in casualties
Goldenbeld & van Schagen (2005)	28 Rural Roads in Friesland, Netherlands, 5 years pre and 8 years post	BA with other rural roads as comparables	4 km/h ↓ in speed; overall 21% ↓ in collisions and casualties
Hess & Polak (2003)	43 fixed speed cams in Cambridgeshire, England, over 11 years	ARIMA, BA with comparable reference sites, long pre-treatment period to mitigate RTM	18% ↓ in collisions & 32% ↓ in injury crashes
Jones *et al.* (2008)	29 mobile cams in Norfolk, England, for 4 years	BA with 48 fixed speed cam sites as comparables	18% ↓ in collisions & 35% ↓ in fatal crashes; no evidence of migration of accidents
Li *et al.* (2013)	771 fixed speed cam sites across England, 9 years	DID-PSM, EB; reference groups by matching on observables	23-31% (0.9-1.4) ↓ in collisions; 0.12 - 0.34 ↓ in fatal crashes; effects smaller with PSM & localised within 200m ; no spillovers of accidents
Li & Graham (2016)	771 fixed speed cam sites across England, 9 years	DID-PSM, EB; reference groups by matching on observables	Cameras are more effective in reducing collisions on riskier sites, measured by higher historical collision counts.
Keall *et al.* (2001)	Visible and Hidden cameras in 4 regions in New Zealand, 1 year pre and post	BA with matching on road characteristics for comparables	0.7 km/h ↓ in speed; overall 11% ↓ in collisions, 19% ↓ casualties; hidden cameras has a more general effect across road
Mountain *et al.* (2004)	62 fixed speed cams across Great Britain, 3 years pre post	EB	35% ↓ in speeding, 26% (1.36) ↓ in collisions, 34% (0.31) ↓ in fatal crashes 500m from cam; effects ↓ moving away from cam
Mountain *et al.* (2005)	79 enforcement schemes (17 mobile, 62 fixed) across Great Britain, 3 years pre post	EB	4% ↓ for every 1mph ↓ in speed ;Larger ↓ reported for lower speed roads; Vertical deflections (speed humps) more effective in reducing speed and accidents
Newstead & Cameron (2003)	Speed cameras in Queensland, Australia, over a 5 year span (2006 to 2007)	Poisson BA with reference sites more than 6km away	21% ↓ in non-injury crashes, 31% ↓ in injury crashes, largest effects localised within 2km
Perez *et al.* (2007)	8 mobile cams in Barcelona, Spain	BA Poisson regressions with nearby reference sites	9 mph ↓ in speed; 27% ↓ in collisions and injuries, greater effects on weekends
Shin *et al.* (2009)	6 speed cameras in Scottsdale, Arizona US, over a 2 year span (2006 to 2007)	BA, EB with nearby reference sites	9 mph ↓ in speed; overall 44-55% ↓ in all collisions, 28-48% ↓ in injury crashes, but no effect on rear-end crashes; no discernable spillovers

EB - Empirical Bayes, BA - Before and after analysis, DID - Difference-in-Difference

Table 3.9: Effects of Speed Camera on various accident outcomes within 500 metres from Camera using Negative Binomial Regressions

	(1) All	(2) Baseline	(3) Demo	(4) VMT	(5) Weather	(6) Non-CAM	(7) -3,+3
Accidents	-0.489[a]	-0.520[a]	-0.280[a]	-0.268[a]	-0.212[a]	-0.168[a]	-0.232[a]
	(0.009)	(0.013)	(0.016)	(0.016)	(0.028)	(0.013)	(0.017)
Obs	66868	25720	25720	25720	7383	35929	9841
Absolute	-2.18	-2.47	-1.49	-1.43	-1.01	-1.40	-1.18
% Δ	-38.68	-40.52	-24.40	-23.47	-19.07	-15.48	-20.70
No.of CAM	2481	1555	1555	1555	659	2249	1481
Slight	-0.433[a]	-0.527[a]	-0.423[a]	-0.379[a]	-0.332[a]	-0.221[a]	-0.310[a]
	(0.010)	(0.013)	(0.021)	(0.019)	(0.036)	(0.013)	(0.021)
Obs	57224	21355	21355	21355	5483	31564	8175
Absolute	-2.35	-3.07	-2.59	-2.36	-1.99	-2.28	-1.89
% Δ	-35.16	-40.97	-34.52	-31.52	-28.27	-19.84	-26.67
No.of CAM	2123	1294	1294	1294	518	1988	1223
Serious	-0.785[a]	-0.758[a]	-0.533[a]	-0.499[a]	-0.447[a]	-0.403[a]	-0.443[a]
	(0.016)	(0.018)	(0.034)	(0.039)	(0.079)	(0.025)	(0.044)
Obs	63280	23650	23650	23650	6539	33823	8306
Absolute	-0.58	-0.57	-0.45	-0.42	-0.32	-0.41	-0.38
% Δ	-54.40	-53.15	-41.30	-39.31	-36.07	-33.15	-35.81
No.of CAM	2346	1428	1428	1428	572	2115	1240
Deaths	-0.934[a]	-1.006[a]	-1.048[a]	-1.018[a]	-0.950[a]	-0.756[a]	-0.883[a]
	(0.043)	(0.071)	(0.108)	(0.109)	(0.185)	(0.077)	(0.131)
Obs	42924	11394	11394	11394	2787	18765	2843
Absolute	-0.08	-0.12	-0.12	-0.12	-0.14	-0.09	-0.19
% Δ	-60.69	-63.41	-64.95	-63.85	-61.32	-53.04	-58.63
No.of CAM	1591	683	683	683	220	1155	426
CAM FE	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓
Demographics			✓	✓	✓	✓	✓
VMT				✓	✓	✓	✓
Weather					✓		

Note: Each reported coefficient is the γ from a different Negative Binomial regression estimated using Maximum likelihood. Dependent variable is the annual Y count where Y=accident, injuries, serious injuries and deaths 500m left and right of camera. Absolute is the number of reductions in accident outcomes computed by multiplying the % Δ with the pre-treatment mean of Y. % Δ is the proportional change (semi-elasticity) of collision outcomes after treatment and is computed by taking $exp(\gamma) - 1$. In Column (1), I include the entire sample of cameras. In Column (2), I restrict the sample to sites that I have full set of co-variates. In Column (3), I control for population size, % of 18 to 25, Gross Annual Pay & hours worked. In Column (4), I control for the annual average vehicle miles travelled (VMT). In Column (5), weather controls are added into the specification. In Column (6), I include a sample of non-camera sites that are eligible for camera installations. In Column (7), I constraint the analysis to observations just 3 years before and after from the year of installation. Bootstrapped standard errors are reported in the parentheses. [c] $p<0.10$, [b] $p<0.05$, [a] $p<0.01$

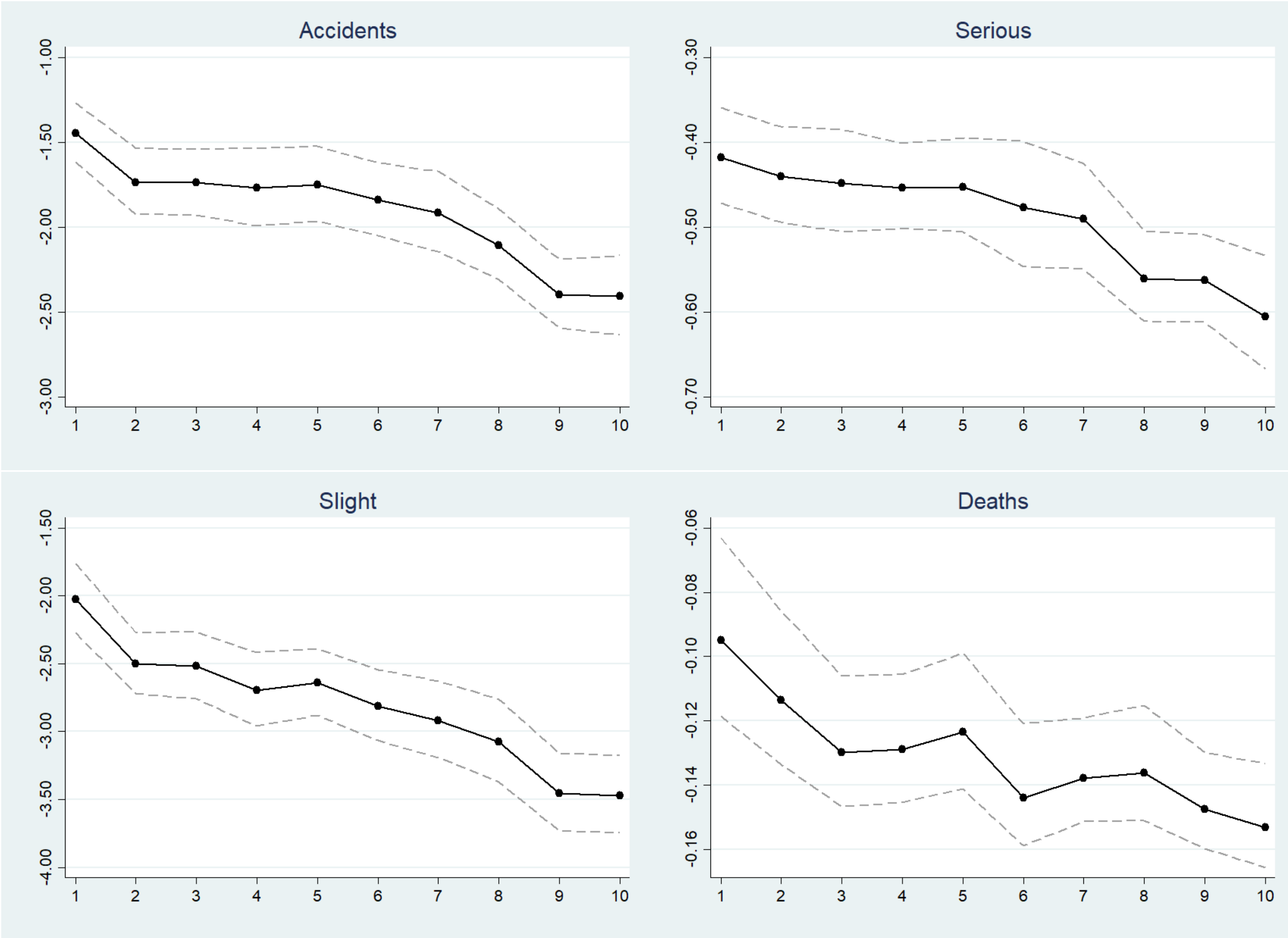


Figure 3.11: Effects of Speed Camera on Accidents 0 to 500m from camera across years from Negative Binomial regressions. Effects can be interpreted as per km. Dots denote the Coefficients (γ_w) from the estimation of Equation 3.2 where H'_w represents a vector of dummies representing years from installation. 95% Confidence Intervals are denoted by the dashed lines.

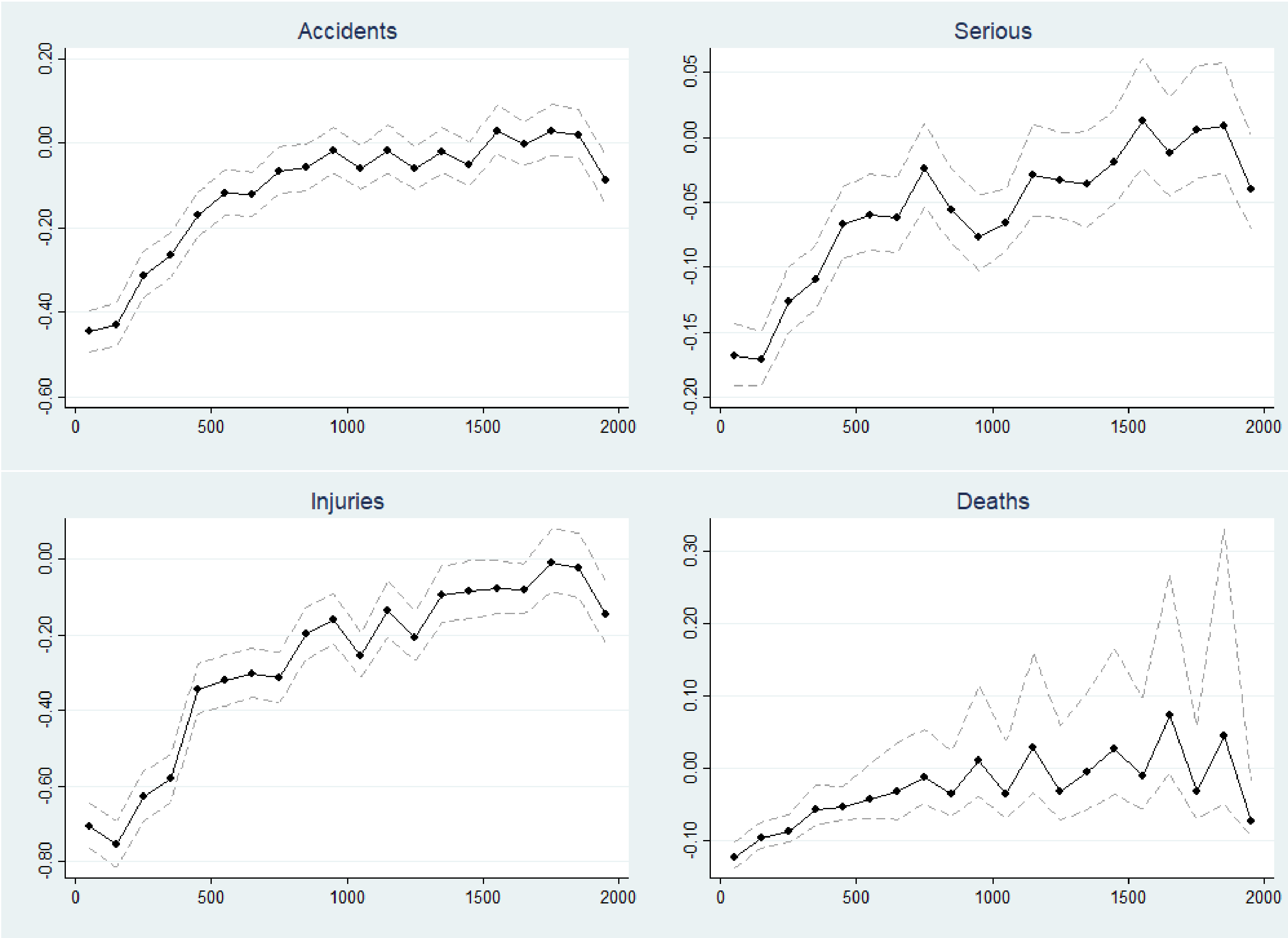


Figure 3.12: Effects of Speed Camera on Accident Outcomes at 100m Intervals (left and right of camera) from Negative Binomial Regressions. Each dot represents the coefficient for every 100 metres distance bandwidth. Effects can be interpreted as per 100 metre. Coefficients (γ_k) are obtained from the estimation of Equation 3.3. Specification is similar to that of Column (4) in Table 3.3. Dot indicates that the coefficient is statistically significant at least at 10% level and cross indicates otherwise. 95% Confidence Intervals are denoted by the tails.

Chapter 4

Beauty Pays for Crime? Evidence from Sentencing Outcomes

4.1 Introduction

According to the Fourteenth Amendment to the U.S Constitution, all citizens must be guaranteed equal protection from laws. This implicitly means that when facing trial, all defendants should not be treated differently based on unimportant factors, such as physical appearance, race, gender etc. Judges are entrusted with the task of upholding this principle. Although for more serious offences, the jury will agree upon whether to convict a suspect, judges decide on the harshness of the punishment. They are required to be impartial and not be influenced by personal experiences, emotions or other unimportant factors. Their verdict should ultimately be based on laws and evidences presented in the courts. Yet, time and time again, judges have been in the spot light for the inconsistency in the punishment for considerably similar offences.

Different extraneous factors have been identified to influence judicial outcomes. Numerous studies have shown that outcome for football games (Eren & Mocan, 2016), duration from food breaks (Danziger *et al.*, 2011), media attention (Lim *et al.*, 2015), race of jury (Anwar *et al.*, 2012), judge (Shayo & Zussman, 2011) and defendants (Abrams *et al.*, 2012; Alesina & La Ferrara, 2014), and gender of defendants (Mustard, 2001) can impact judicial rulings. These studies seem to suggest that these "expert" decision makers are easily swayed by unimportant factors.

In this paper, I contribute to the existing literature by evaluating whether facial attractiveness of criminals impact sentencing outcomes. This is motivated by the prevalence of the discrimination against the physically unattractive in multiple contexts. It has been long documented that labour markets discriminate based on appearance as attractive individuals earn higher wages while plain looking people are penalized with poor labour market outcomes (Hamermesh & Biddle, 1994; Biddle & Hamermesh, 1998; Graham *et al.*, 2016). Researches suggest that better looking people earn more because they appear to be more confident with superior communication skills (Mobius & Rosenblat, 2006). Being attractive also reduces the propensity for young adult to engage in criminal activities (Mocan & Tekin, 2010). This bias towards more attractive individuals is also documented widely outside labour markets. While better-looking politicians appear to win more votes in elections (Berggren *et al.*, 2010), lenders also trust better looking borrowers more than unattractive counterparts (Duarte *et al.*, 2012; Ravina *et al.*, 2008; Jenq *et al.*, 2015). Looks could also impede human capital accumulation. Hatfield & Sprecher (1986) suggest that better-looking adolescents are showered with more attention from teachers that can improve their confidence and social skills. Supporting this, Mocan & Tekin (2010) show that less physically attractive teenagers have poorer test scores.

Despite the burgeoning evidences on the disparity in judicial sentencing due to extraneous factors, and the pervasiveness of bias against the unattractive in different settings, the literature linking appearance of felons and sentencing outcomes is remarkably limited. This is probably because the empirical analysis is beset with multiple challenges. First, measuring attractiveness is challenging as it is hard to clearly define what constitutes beauty. Measures widely used in the psychology literature to define attractiveness[1] include facial symmetry, averageness and sexual dimorphism. Furthermore, as the saying goes, *beauty lies in the eyes of the beholder*. The definition of attractiveness could vary tremendously across cultures and time. The argument made by many academics is that there seems to be consensus among individuals on the standards of beauty within culture and time. To circumvent this problem, studies often require different respondents[2] to rate the same subjects to obtain an objective measure of attractiveness. This resource-intensive research set up, however, meant that the sample sizes are usually small, questioning the external validity of these studies.

Even if appearance is accurately measured, there could be unobserved confounders that correlate with facial attractiveness and influence judicial sentencing. For instance, if labour market outcomes are aversely affected by appearance (Hamermesh & Biddle, 1994; Biddle & Hamermesh, 1998), labour market penalties could incentivise unattractive individuals to engage in illicit activities. Higher socio-economic status could allow attractive felons to engage in more reputable attorneys who can reduce the severity of their sentences. All these factors could exacerbate the sentencing gap between the better and worse looking criminals.

Bearing these challenges in mind, this research contributes to the existing lit-

[1]This list is by no means exhaustive. Other measures used include grooming, youthfulness, babyfaced-ness and expression. Moreover, there is a lack of consistency in the definition of attractiveness.

[2]For instance, in the seminal study by Hamermesh & Biddle (1994), assessors of between 7 and 50 years old are asked to rank the same images and their responses appear to be highly correlated. Similarly, Mocan & Tekin (2010) rely on different evaluators to assess beauty to obtain a representative measure of attractiveness.

erature on several fronts: I conduct the analysis to a universe of almost 300,000 convicted criminals from 1998 to 2015 in Florida to increase the representativeness of my findings. This is possible because I rely on a facial recognition algorithm on the mugshots to accurately delineate various facial features (eyes, nose, mouth, forehead and jawline) and compute facial symmetry to approximate attractiveness. Third, I exploit the random assignment of cases to judges in the courts in Florida. This means that cases are randomly allocated to judges not based on the appearance of the felon. Finally, to mitigate the risk of observed differences between better and worse looking criminals from biasing the estimates, I control for a rich set of covariates that includes (1) physical attributes of felons (age, height, weight, race, tattoo, eye and hair color), (2) case facts (crime type, number of concurrent charges, whether is the principal criminal) and (3) criminal history (counts and types of previous offences).

The main finding is that judges hand out lenient sentences to criminals with more symmetrical faces. The disparity in punishment between criminals with more symmetric faces, at the 25th percentile, and criminals with less symmetric faces, at the 75th percentile, is around 1.0% to 1.9% of the mean sentence length, which amounts between 17 and 32 days. These results hold across a battery of sensitivity analyses. This bias against less attractive felons appears to vary across race, gender and type of crimes. The effects are smaller associated with black inmates, suggesting that cross-race judgement of facial symmetry could be weaker than within race as most of the judges in Florida are white. Conversely, female felons with more symmetrical features are given harsher sentences. This reversal in relationship could be explained by paternalism towards female felons as asymmetric facial features could be correlated with sympathetic life circumstances. I further observe the effects of facial symmetry on sentences are weaker for serious offences (e.g murder, manslaughter), suggesting that judges have less flexibility to depart from sentencing guidelines for high profile cases.

The remainder of the paper is organized as follows. Next, I review several streams of literature related to this research, before describing the data and how I measure attractiveness using facial symmetry. Subsequently, I will illustrate the institutional background of judicial sentencing before highlighting the identification strategy adopted in this paper. Finally, I will present the findings before concluding the research.

4.2 Literature Review

There are several streams of literature particularly related to this research. The first line of research is understanding what defines attractiveness. The three main components prominent in psychology are averageness (Langlois & Roggman, 1990), symmetry (Scheib *et al.*, 1999), and sexual dimorphism (Perrett *et al.*, 1998). An average face has mathematically mean trait values of the population and is low in distinctiveness. A symmetrical face has physiognomy features that are highly proportional. Sexual dimorphism refers to facial traits that emerge during puberty, signalling sexual maturity and reproductive potential.

These measures of facial attractiveness each has their pros and cons. For instance, an average face is usually created out of compositing many digitized faces that confounds averageness with symmetry and smoothness of the skin (Alley & Cunningham, 1991). It is also challenging to define the standards of "attractive" sexual dimorphism. For instance, it is subjective to consider what is a distinctive jawline. Conversely, among the different measures, it is the most straightforward to measure facial symmetry as it is based on the proportionality of various facial features. Numerous studies have also shown that facial symmetry is highly correlated with attractiveness Grammer & Thornhill (1994); Perrett *et al.* (1999); Jones & Hill (1993). Hence, this paper will focus on facial symmetry to measure attractiveness.

Why do attractive faces matter? They are desirable because they reflect functional optimality, disease resistance, development stability and signal better mate quality (Rhodes, 2006). An attractive face could also draw positive first impressions and elicit desirable traits - a popular stereotype of *what is beautiful is good* (Dion *et al.*, 1972). Previous literature has shown that attributes such as trustworthiness, competency, sociability and confidence etc. are highly correlated with ratings of facial attractiveness. For a summary of the existing literature, refer to Feingold (1992) and Langlois *et al.* (2000). All these results suggest that facial attractiveness could matter in the courtroom.

Despite the copious amount of research indicating that judicial decisions are swayed by extraneous factors, existing literature on the presence of discrimination on appearance in the courtroom is largely limited to the field of psychology. The bulk of these studies conclude that appearance matters as judges tend to hand out favourable sentences to better looking felons. Stewart (1980) conducted an observational analysis for 74 defendants from Pennsylvania during trials, requiring observers to rate them on attractiveness, grooming and cleanliness etc. Although he finds that attractive defendants receive lighter sentences, appearance do not matter for conviction rates. Zebrowitz & McDonald (1991) examine the impact of babyfacedness and attractiveness for a sample of 504 cases from the small claims court in Massachusetts. They report that the appearance of defendants and plaintiffs matter as judges side youthful-looking individuals. Whether attractiveness matters also depends on the type of crimes committed. Attractiveness could work against felons who committing crimes that rely on their appearances (Sigall & Ostrove, 1975). The bias against less attractive felons also appears to be implicit. Subjects seem unaware of their bias against the uglier suspects despite recommending lighter sentences and lower conviction rates for better looking felons (Efran, 1974).

A review of the existing literature reveals several shortcomings. Many of these studies mentioned above, especially earlier ones, are laboratory experiments that are not only constrained to a limited sample, but also require non-judge subjects to make judicial decisions outside courtroom settings. Hence, it is unlikely that the results from these field experiments are going to be representative or reflective of actual sentencing outcomes. Even for observational studies, most of them are fairly dated and lack information on cases and felons. These limitations restrict the analyses to parsimonious regression models plagued with omitted variable bias.

Economics researches on the discrimination in judicial sentencing have made more headway in detecting prejudice. Most of these papers focus on understanding racial and gender bias in judicial outcomes. These studies rely on random assignment of cases to judges and a richer set of control variables to improve the identification of discrimination. Researches have provided overwhelmingly support for racial or gender bias in capital punishment (Alesina & La Ferrara, 2014), bail (Ayres & Waldfogel, 1994), probability of being convicted (Mustard, 2001; Abrams *et al.*, 2012) and sentence length (Abrams *et al.*, 2012; Butcher *et al.*, 2017). Most of these studies report favourable treatment towards females and whites. Several studies seek to understand the underlying motivations for discrimination in policing with some innovative strategies. For instance, Knowles *et al.* (2001) examine the guilty rates for motor vehicle searches across different races. The idea is that if race is indicative of the propensity to carry contraband substance - consistent with statistical discrimination - then even if the number of searches are higher for a particular group, the guilty rates should be very similar across races. Indeed, they find evidence supporting statistical discrimination. Rank-order tests[3] are conducted to detect whether judges are consistent in the judgement towards felons belonging in different groups (Park, 2017; Butcher *et al.*, 2017).

Overall, it is surprising to find no studies on appearance bias in judicial outcomes in the economics literature despite the sheer volume of work illustrating the discrimination against the less attractive in multiple settings. This is a gap that this paper aims to fill.

[3]This strategy is first adopted by Anwar & Fang (2006) on motor vehicle searches. The idea is that taste-based discrimination prevails if white police officers search black motorists more than black officers and vice versa. However, if both black and white officers are more prone to search black motorists, then race might be signalling some latent criminality and statistical discrimination prevails.

4.3 Data

The main source of data is from the OBIS database provided by the Florida Department of Corrections (FDOC). This comprehensive database documents information for a universe of convicted felons, including those released, those currently under supervision and incarceration from 1997 onwards. For each case record, details on the sentencing outcomes and case facts, which include length of sentence, date of sentencing and offence committed, number of concurrent charges, type of crime and location of court, are reported. Other details recorded for each felon include criminal history, gender, race, age, height, weight, body tattoos and residential address.

To assess facial attractiveness of the convicts, I first scrape inmate mugshots from FDOC. These pictures are usually taken when inmates are initially incarcerated in the facility. Using a facial recognition package, I extracted 69 facial landmarks from each of the mugshots[4]. These features delineate the top of the forehead, eyebrows, eyes, nose, lips and jaw line. An illustration is shown in the Figure 4.1.

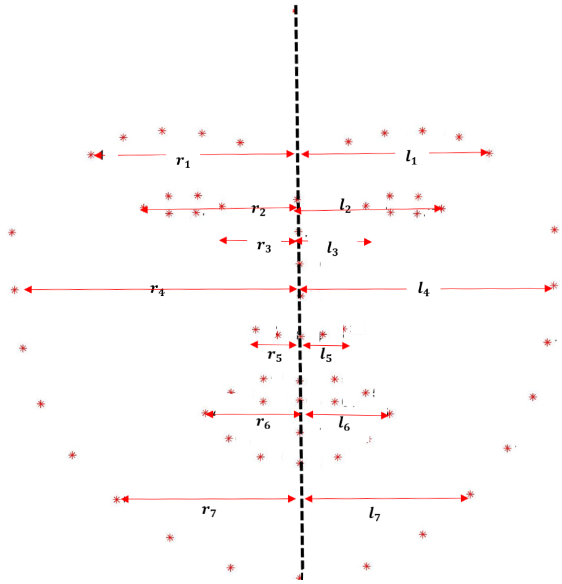


Figure 4.1: An illustration of the 69 facial features extracted & how facial symmetry is computed

Using the location of these facial landmarks, I measure the distance of the left and right features from the middle of the face denoted by the dashed line. A total of seven features, which include eye brows (r_1 & l_1), eyes (r_2, l_2), interocular (r_3, l_3), middle face (r_4 & l_4), nose (r_5 & l_5), mouth (r_6 & l_6) and lower face (r_7 & l_7), are accounted for. I first measure the absolute difference between the distance of the

[4]The algorithm is written by Vahid Kazemi and Josephine Sullivan to locate facial landmarks rapidly using machine learning. For more details, refer to Kazemi & Sullivan (2014).

left (l_n) and right (r_n) for each of the 7 features, before averaging them to compute facial symmetry for each inmate (FS_i) as shown in equation 4.1. Details of these features are summarized in Table 4.1.

$$FS_i = \frac{\sum_{n=1}^{7} |r_n - l_n|}{7} \tag{4.1}$$

Table 4.1: Facial Symmetry Ratios

Feature	Computation	Description
Eyebrows	$r_1 - l_1$	Right Brow - Left Brow
Eyes	$r_2 - l_2$	Right Eye - Left Eye
Interocular	$r_3 - l_3$	Right Interocular - Left Interocular
Middle Face	$r_4 - l_4$	Right Face Width - Left Face Width
Nose	$r_5 - l_5$	Right Nose Width - Left Nose Width
Mouth	$r_6 - l_6$	Right Mouth Width - Left Mouth Width
Lower Face	$r_7 - l_7$	Right Bottom Face Width - Left Bottom Face Width

While the pros of using facial symmetry to proxy for attractiveness are that it is straightforward to measure and it provides a continuous measure of attractiveness, there are some cons with using facial symmetry. First, FS_i only measures horizontal symmetry and will not account for inmates with disproportionately wide or long faces. Therefore, I include face-width-height ratio as an additional control in the robustness tests (See Table 4.5).

Another concern is that facial symmetry captured from mugshots could be inaccurately measured from badly taken pictures. For instance, if the inmates' head is tilted to the left or right when the mugshot is taken (See Figure 4.2), it is possible for faces to be asymmetrical even though this might not be true. The concern is whether defiant inmates, who receive harsher sentences, take poorer mugshots. To correct for measurement error in facial symmetry from poorly taken mugshots, I compute various adjustment angles that account for how slanted the face is. First, I calculate eye-adjustment angle (k) - the angle between the actual eye-line and the horizontal eye-line. Then, I account for the nose-adjustment angle (d) - the angle between the actual nose-line and the vertical nose-line. For illustration of tilted and well-taken mugshots, and how I compute these adjustment angles, refer to Figure 4.2. A well-taken mugshot should have a flat horizontal eye-line and a vertical nose-line as observed in Figure 4.2c.

When taking mugshots, inmates could also be facing left or right, or could be lifting their heads up or down, affecting facial symmetry. Hence, I further compute the absolute of the difference between right eye-to-face-edge (D_r) and left eye-to-face-edge distance (D_l), and the absolute difference between forehead-eyebrow distance

(a) This image has been removed as the copyright is owned by another organisation

(b) This image has been removed as the copyright is owned by another organisation

(c) This image has been removed as the copyright is owned by another organisation

Figure 4.2: Examples of side-way tilted mugshots that could affect facial attractiveness. To find d and k, one can do it using Pythagoras Theorem where $d = Cos^{-1}(\frac{VerticalNoseLength}{ActualNoseLength})$ and $k = Cos^{-1}(\frac{HorizonalEyeLength}{ActualEyeLength})$. Source: Florida Department of Corrections

(D_u) and mouth-chin distance (D_d). These measures are illustrated in Figure 4.3.

(a) This image has been removed as the copyright is owned by another organisation

(b) This image has been removed as the copyright is owned by another organisation

Figure 4.3: Examples of left or right facing and up or downward facing mugshots that could affect facial attractiveness. To account for this I calculate the $|D_r - D_l|$ and $|D_u - D_d|$ and include these measures as controls. Source: Florida Department of Corrections

4.4 Institutional Background

Sentencing guidelines are adopted in Florida since 1983 to eradicate disparity in punishments between similar offences to ensure fairness, and to make sure that felons are adequately punished for their transgressions. A scoresheet is prepared for each defendant by the state attorney, before being presented to the defense counsel and signed off by the judge. Offences are categorized into 10 levels, depending on the severity of the crime that is based on the purported level of harm inflicted on the community. The more severe the crime, the more points will be allocated. Additional points could be given depending on the way the crime is committed[5] and the criminal history of the felon. The total points will determine the severity the punishment[6]. Although the points determine the recommended sentencing, judges are allowed to depart from the recommendations if proper reasons[7] could be provided.

[5]Additional points could be given if the felon inflicts victim injury or made sexual contact, the felon has serious offence history, there is a use of firearms or the crime is gang related.

[6]Offenders receiving 44 points or less could receive a non-state prison sanction, while those exceeding will be given a minimum sentence in months by subtracting the points received by 28 before decreasing the value by 25%.

[7]Departure of sentences could occur depending on whether the defendant is juvenile, whether the defendant is an accomplice, and whether the defendant is able to appreciate the criminal nature of the conduct. For more information, refer to http://www.dc.state.fl.us/pub/sen_cpcm/cpc_manual.pdf.

Sentencing guidelines went through major changes[8] over the years. The Criminal Punishment Code (CPC) became effective on the 1st October 1998. The new code not only lowers the threshold for incarceration, but also increases the maximum sentences permitted. Under CPC, the maximum sentence is determined by the maximum provided in statute 775.082 rather than being capped by the 25% upward discretion under previous guidelines. Adhering to the statutory maximum means that potentially all felony offenders could be incarcerated, and that they are likely to receive longer sentence lengths[9]. With more flexibility given to judges to depart from the guidelines, it is more likely for extraneous factors, such as appearance, to have a consequential effect on judicial outcomes.

4.5 Identification Strategy and Methodology

To estimate the impact of facial symmetry on sentencing length, I specify the following equation:

$$D_{ijct} = \alpha_j * \delta_c + \gamma \mathbf{FS_i} + X'_{ij}\phi + M'_i\omega + \theta_t + \varepsilon_{ijct}, \tag{4.2}$$

where D_{ijct} is the natural logarithm of the sentence length measured in days for defendant i committing crime type j and trialled in court c at time t. The key variable of interest, $\mathbf{FS_i}$, is the absolute deviation from the perfect facial symmetry. This is measured based on the average symmetry from 7 different facial features as described earlier in Table 4.1. Hence, γ captures the percentage change in sentencing length from a one unit increase in deviation from a perfectly symmetrical face. If judges are favoring felons with symmetric facial features, I would expect γ to be $>$ 0.

X'_{ij} represents a vector of personal characteristics, case facts and criminal history associated with defendant i committing crime type j. Personal details include race, gender, age, tattoos (counts of visible tattoos), height, weight, hair and eye colour. By partial-ing out the effect of height and weight, this paper focuses on the impact

[8]The first reforms took place in 1994. New guidelines were created in response to the epidemic of certain offences, such as crack-related crimes, the passage of unfunded mandatory minimum sentence legislation and the population boom to reduce the strain on correction facilities. The objective is to ensure that state incarceration will be enforced on repeated offenders who are threatening the society committing serious or violent crimes. These reforms include better categorization of the offences, and allowing additional points to be given for repeated offenders and for the manner the crime is committed. In 1995, the guidelines were amended again. Point values were increased in a variety of areas and additional policy levers were created to permit tougher punishments. These guidelines went through more modifications in 1997 and 1998 that further exacerbates sanctions.

[9]Based on the new code, a felon charged with life felony could receive life imprisonment, while 1st, 2nd and 3rd degree could receive jail terms of up to 30, 15 and 5 years respectively.

of physiognomy on judicial outcomes. As the majority of the inmates are either black or white males, I restrict the analysis to these groups, which makes up more than 88% of the sample, to reduce heterogeneity in the sample. I will examine the differential effects of facial symmetry on sentencing outcomes across different races and gender in subsequent sections.

Case facts accounted for include the number of concurrent charges and whether the felon is the principal criminal. Those playing lesser roles in the transgression are less culpable and could be given lighter sentences. I further control for the criminal history of each defendant by including the total number of times the felon has been previously incarcerated as a regressor. The idea is that serial offenders are subjected to heftier punishment to deter re-offending. M_i' represents a vector of mugshot related characteristics. These variables include various adjustment angles and distances that correct for badly taken mugshots that could affect the measurement of facial symmetry. Refer to Figure 4.2 and 4.3 for more information. I also control for the image data size as the fuzziness of the mugshots could exacerbate measurement error.

δ_c represents court fixed effects that are indicators for the court responsible for handling the case. α_j are indicators denoting the type of crime committed. Relying on the detailed information provided for each sentence, I classify the offences into 1294 unique categories[10]. I include the interaction of crime (α_j) and court (δ_c) fixed effects [11]. In other words, I am exploiting the variation of facial symmetry and sentences between felons committing the same crime type j and trial in the same court c. Adding crime-court fixed effects not only mitigates the risk of time-invariant unobserved differences between offences from confounding the estimates, but also ensure that effects are not induced by the variation of harshness in sentencing across courts.

θ_t denotes month and year fixed effects for both the offence and sentencing dates that controls for general trends in the sentencing across counties over time. Due to major changes in sentencing guidelines over the sample period, I restrict the sample to offences committed after the passing of the Criminal Punishment Code on the 1st of October 1998. ε_{ijct} represents the standard errors that are clustered at court-by-year levels. The assumption is that $E[\varepsilon_{ijct}|FS_i = 0]$.

[10]For more information on how the different transgression are classified, refer to http://www.dc.state.fl.us/appcommon/offctgy.asp

[11]This specification is chosen also because I do not have information on the judges. Ideally, I would like to include judge fixed effects and exploit the variation in facial symmetry and sentencing within judge and crime type. Given the lack of judicial details, interacting the crime and court fixed effects ensures that I am closer to exploiting within judge variation given that there seems to be some specialization in the types of crimes that are handled by judges in a court.

4.6 Empirical Results

4.6.1 Descriptive Statistics

Table 4.2: Summary Statistics

	All		Above Median (More Symmetric)		Below Median (Less Symmetric)		Difference	
	Mean	SD	Mean	SD	Mean	SD	Diff	T-stats
Sentencing Outcomes								
Sentence Length (Years)	4.65	5.56	4.61	5.56	4.88	5.76	0.27	23.97
Personal Traits								
Age	34.63	10.84	34.32	10.52	34.61	10.89	0.29	13.32
Height(feet)	5.27	0.40	5.27	0.39	5.28	0.40	0.01	17.32
Weight(lbs)	182.82	35.79	182.07	35.05	183.13	36.08	1.06	14.90
Black	0.41	0.49	0.39	0.49	0.41	0.49	0.02	20.50
Hispanic	0.04	0.19	0.04	0.20	0.04	0.19	-0.00	-6.76
White	0.55	0.50	0.57	0.50	0.55	0.50	-0.02	-17.46
Female	0.13	0.34	0.14	0.35	0.12	0.32	-0.03	-37.62
Counts of Tattoos	3.30	3.89	3.36	3.94	3.33	3.90	-0.03	-3.39
Have Face Tattoos	0.15	0.35	0.15	0.36	0.15	0.36	-0.00	-0.46
Case Related Characteristics								
Principal Criminal	0.99	0.08	0.99	0.08	0.99	0.08	0.00	1.80
No. of Charges per Case	4.83	12.39	4.72	11.89	5.05	12.84	0.33	13.16
No. of previous offences	1.18	3.98	1.17	3.95	1.19	4.12	0.02	2.72
Assault	0.09	0.28	0.08	0.28	0.09	0.28	0.00	6.86
Drug	0.24	0.43	0.24	0.43	0.23	0.42	-0.01	-9.67
Manslaughter	0.00	0.07	0.00	0.07	0.00	0.07	-0.00	-1.06
Murder	0.01	0.09	0.01	0.09	0.01	0.10	0.00	5.92
Other Crimes	0.20	0.40	0.20	0.40	0.20	0.40	-0.00	-5.96
Property Crime	0.01	0.09	0.01	0.09	0.01	0.09	0.00	4.16
Robbery	0.05	0.22	0.05	0.22	0.06	0.23	0.00	7.16
Sex	0.06	0.24	0.06	0.23	0.07	0.25	0.01	19.65
Theft	0.31	0.46	0.31	0.46	0.31	0.46	-0.01	-6.69
Weapon	0.03	0.17	0.03	0.17	0.03	0.17	0.00	2.79

All represents the entire sample of 807,013 sentences from 278,240 inmates . Above Median represents a group of inmates with facial features more symmetrical that the median face, while Below Median represents a group of inmates with facial features less symmetrical than the median face. Difference captures the absolute difference in the mean of observable covariates between the Above and Below Median and T-statistics denote the statistical significance of such differences.

Table 4.2 presents the summary statistics of sentencing outcomes, observable personal and case related characteristics surrounding the felons in the analysis. I first report statistics for the entire sample, before stratifying the sample into above and below median facial symmetry. I further report the differences-in-mean for various observable characteristics. These preliminary results show whether felons with more symmetric facial features are different from others in observable characteristics.

There are significant differences in sentencing outcomes. On average, criminals with more proportional facial features serve sentences that are 0.27 years (about 3 months) shorter, suggesting that judges are more lenient to those felons with more symmetrical faces. In terms of personal traits, I observe that inmates with less symmetrical facial features are, on average, 0.3 years older, around 1 pound heavier, and have fewer tattoos. There are, on average, a slightly bigger proportion of black inmates and smaller proportion of white inmates with facial symmetry below median.

Surrounding case related characteristics, more facially symmetrical criminals are charged with 0.34 fewer counts of concurrent offences compared to those below median symmetry. There are no stark differences in terms of the types of crime committed although less facially symmetry criminals are more prone to be convicted for sex related offences, and less likely to be remanded for theft and drug crimes.

Overall, although there are disparities associated with personal traits and case-related facts between criminals with above and below median facial symmetry, these differences are unlikely to bias the estimates as they are controlled for in the regression analyses as shown in the subsequent sections.

4.6.2 Baseline Results

Table 4.3 presents the baseline estimation of γ from equation 4.2. On top of restricting the sample to white and black male felons, I further truncate the top and bottom 1% of facial symmetry to alleviate the possibility of outliers from driving the results. Column (1) shows results from a parsimonious model with only crime*court fixed effects and indicators for the months and years of sentencing and offence dates. Put differently, I am comparing the effect of facial symmetry of criminals committing the same crime type and being sentenced in the same court on their sentencing outcomes. This amounts to more than 14,000 fixed effects (from 20 Circuit Courts) associated with almost 280,000 different felons. Results suggest that a unit increase in deviation from a perfectly symmetrical face increases sentencing length by 0.2%. Putting these estimates into perspective, the sentencing gap between the felon at the 25th percentile and the felon at the 75th percentile in facial attractiveness is approximately 17 days, which corresponds to around 1% of the mean sentence length. I extend the baseline model by including observable inmate characteristics in Column (2). Indicator variables for age, race, ethnicity and the presence of face tattoos, the counts of tattoos, height and weight of the inmate are controlled for. These personal traits do not appear to matter much as the estimates remain fairly similar to before.

Next, I add a vector of photo-related characteristics into the analysis. These controls include (1) the image data size of the photo, which accounts for the resolution of the picture, (2) the nose and eye adjustment angles that measure how tilted the face is, and (3) the absolute face to eye ratio that measures whether the inmate is facing sideways, and (4) the forehead to mouth-chin ratio that captures whether the inmate is facing up or downwards. This specification accounts for measurement error associated with using facial symmetry to measure attractiveness because of poorly

Table 4.3: Baseline Results

	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{FS_i}$	0.002[a]	0.002[a]	0.004[a]	0.003[a]	0.005[b]	
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	
$\mathbf{FS_i^2}$					-0.000	
					(0.000)	
Quintile1						-0.031[a]
						(0.008)
Quintile2						-0.029[a]
						(0.008)
Quintile3						-0.026[a]
						(0.007)
Quintile4						-0.012[c]
						(0.006)
Obs	807037	802952	795944	795944	795944	795944
R2	0.54	0.55	0.56	0.57	0.57	0.57
No. of Inmates	278483	277166	274972	274972	274972	274972
Crime*Court FEs	14118	14089	14021	14021	14021	14021
Absolute Effects (25% $\rightarrow$ 75%)	17.32	15.80	31.31	23.81	32.62	.
%Δ (25% $\rightarrow$ 75%)	0.99	0.90	1.81	1.37	1.88	.
Crime*Court Fixed Effects	✓	✓	✓	✓	✓	✓
Year Month Fixed Effects	✓	✓	✓	✓	✓	✓
Inmate Details		✓	✓	✓	✓	✓
Photo Details			✓	✓	✓	✓
Case Facts				✓	✓	✓

Dependent variable is logarithm of total number of days of sentence i by felon j. Key Independent variable, $\mathbf{FS_i}$, measures the deviation from a perfectly symmetrical face. Absolute effects report the sentencing gap in the number of days moving from the 25th to 75th percentile in facial attractiveness, Crime*Court fixed effects is the interaction of j different categories of crime with c court fixed effects. Year Month Fixed effects include both year and month dummies for date of sentencing. Inmate characteristics include race, age, height, weight, number of tattoos and whether the tattoos are visible. Photo Characteristics entail adjustment angle, distances and picture size. Case facts include the number of charges included in the same case, whether felon i is the principal criminal, and detailed criminal history capturing the number of times the felon i has been incarcerated, and the number of times felon i commit the same crime j. Standard errors are clustered in court*year level. Columns (5) and (6) report the non-linear relationship between facial attractiveness and sentencing. In Column (5), $\mathbf{FS_i^2}$ is the squared term of $\mathbf{FS_i}$. In Column (6), continuous measure of facial attractiveness is divided into quintiles where Quintile1 takes the value of 1 for inmates classified in the first quintile (most attractive) or otherwise zero while Q5, the base (excluded) group, takes the value of 1 for inmates in the last quintile (least attractive) or otherwise zero. Standard errors clustered at court-by-year levels are reported in the parenthesis. [c] $p<0.10$, [b] $p<0.05$, [a] $p<0.01$.

taken mugshots. Indeed, controlling for these differences increases sentencing gaps considerably to 31 days, suggesting that measurement error in facial symmetry could have attributed to attenuation bias.

In Columns (4), I control for details of the case that include the number of concurrent charges, whether the felon is the principal criminal and the criminal history that is measured by the number of times the felon has been convicted for the same crime type before. These details should affect the harshness of sentencing and controlling for these differences should mitigate the concern that unattractive felons could be serial criminals prone to committing more severe transgressions.

Accounting for these differences marginally reduces sentencing gap to 24 days but the effects remain statistically significant at 1%.

Next, I examine whether the relationship between facial attractiveness and harshness of sentencing could be non-linear. That is, whether judges show stronger favouritism to the most attractive felons. I first include the quadratic term of facial symmetry ($\mathbf{FS_i^2}$) in Column (7). Although the squared term is not statistically significant, the coefficient is negative and the inclusion appears to accentuate the estimates. This suggests that the most attractive felons could enjoy the sharpest reductions in sentences. Now, the sentencing gap is around 25 days moving from the 25th to the 75th percentile in facial symmetry.

I further divide facial symmetry into quintiles, with Quintile1 representing the top quintile of felons with the most symmetrical faces and Quintile5, the omitted reference group, representing felons with the least symmetrical faces. The difference in sentence length between the most (Quintile1) and least attractive felons (Quintile5) is 3.1%, which corresponds to a gap of 44 days. This magnitude of favouritism decreases as faces become less proportional as evidenced by the smaller estimated effects moving down the quantiles. The difference in sentencing between felons in the fourth quintile (Quintile4) and the least attractive faces are smaller at 1.2%, which works out to be around 21 days. This sharp drop in effects moving into the fourth quintile suggests that judges are siding more on criminals with distinctively more symmetrical facial features.

In summary, my baseline results reveal that the facial symmetry affects the sentencing outcomes. In particular, criminals with more symmetrical faces are given shorter sentences for the crimes they committed compared to criminals with less symmetrical faces. Further analysis suggests that the bias towards criminals with more proportional faces could be non-linear. Judges appear to hand out the most lenient punishments to the most proportional faces and this preferential treatment is smaller as faces become less symmetric.

4.6.3 Heterogeneous Effects

Next, I explore whether the effects of facial symmetry on sentencing vary between different groups. I analyse the relationship separately for black, white, male and female felons[12]. Results are summarized in Table 4.4.

Columns (1) and (2) report the effect of facial symmetry on sentence length for black and white felons respectively[13] for the entire population of inmates and

[12] Result are similar when I combine the analysis with interaction variables. They are available upon request.

[13] In other unreported results, I include other minority races including Asian and Hispanics. This takes up about less than 4% of the sample. Similar to the blacks, I do not

for those released. It appears that effects of facial symmetry on sentencing are fairly consistent between black and white felons although the estimates are slightly larger for white inmates. To put the estimates into perspective, compared to the 75th percentile, a white inmate at the 25th percentile in facial symmetry serve, on average, 27 fewer days. Conversely, black inmates with more proportional facial features serve, on average, 20 fewer days. Why are the effects stronger for white inmates? One possible explanation is that more than 91% of the judges in Florida are white. Cross-race judgement for attractiveness could be slightly weaker than within race assessment (Malpass & Kravitz, 1969; Bernstein *et al.*, 1982). White judges could be more able to discern facial symmetry for white felons, explaining slightly larger effects associated with white felons.

Columns (3) and (4) further breakdown the analysis to male and female felons. Results in Column (3) are similar to that reported in Column (4) of Table 4.3. It is very interesting to observe a reverse in the relationship for female felons. Now, comparing the sentencing outcomes for female inmates at the 25th percentile with the 75th percentile in facial symmetry, holding all other factors constant, I observe a 32 day reduction in sentencing length. While an in-depth analysis of this intriguing result is out of the purview of this paper, the divergence in effects for female felons could stem from the fact that judges are paternalistic towards women believing that they are physically weaker (Moulds, 1978; Spohn, 1999). Unobserved sympathetic life circumstances, such as mental illness, poverty and addiction, could be correlated with facial symmetry and these factors could explain why less attractive female defendants are less harshly punished. Refer to (Starr, 2014) for more detailed discussion of possible explanations to male-female sentencing gaps.

Next, I explore how the effect of facial symmetry on sentencing outcomes to manifest across different crime types. This is estimated by allowing bias to vary across the 9 major crime types. I interact the indicator variable of each crime with $\mathbf{FS_i}$ and coefficients of this interaction term are plotted in Figure 4.4. A few notable observations can be made. It is interesting to observe that criminals with more proportional facial features are more likely to get away with lighter sentences for sex-related transgression although the estimates are too imprecisely estimated to be statistically significant. This is consistent with the findings reported by Jacobson (1981) that the outcome of rape cases could be influenced by the appearance of defendants and plaintiffs. Furthermore, more serious transgressions, such as murder and manslaughter, are more equitably sentenced regardless of the facial symmetry of the felons. This suggests that judges may have less flexibility to depart from the sentencing guidelines for more higher profile cases.

find that judges favoured more proportional faces for races other than white

Table 4.4: Effects of Gender & Race on Discrimination of Facial Attractiveness

	(1) Black	(2) White	(3) Male	(4) Female
FS_i	0.003^b	0.004^a	0.003^a	-0.008^a
	(0.001)	(0.001)	(0.001)	(0.002)
Obs	336513	456281	795944	117757
R2	0.59	0.57	0.57	0.52
No. of Inmates	124364	149826	274972	40079
Absolute Effects (25% → 75%)	20.25	26.80	25.36	-31.50
%Δ (25% → 75%)	0.86	0.95	0.91	-1.53

Refer to the notes from earlier Table 4.3 for more information. In Columns (1) & (2), I stratify the analysis to Black and White inmates, allow the discrimination to vary between black and white felons. In Columns (3) & (4), I stratify the analysis to Male and Female inmates. This group of 40,079 female inmates are previously omitted from the analysis . ^{c} p<0.10, ^{b} p<0.05, ^{a} p<0.01.

4.6.4 Robustness Tests & Alternative Interpretations

Table 4.5 summarizes the findings from a battery of robustness tests that examine the validity of the results.

Table 4.5: Robustness Tests

	(1) FWHR	(2) Zipcode FE	(3) Age<=70	(4) <=5 Crime	(5) After2005	(6) <=90%	(7) High Res	(8) Crime#Time
FS_i	0.003^a	0.004^a	0.003^a	0.004^a	0.002^b	0.005^a	0.004^a	0.003^a
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Obs	795944	469951	794298	640159	582299	639113	626907	758672
R2	0.57	0.36	0.57	0.58	0.60	0.57	0.59	0.64
No. of Inmates	274972	176864	274382	260270	200996	221208	206671	264247
Crime*Court FEs	14021	8157	13999	13426	12529	13051	13193	66291
Absolute Effects (25% → 75%)	23.08	18.88	25.19	28.16	15.04	26.92	32.52	23.51
%Δ (25% → 75%)	1.33	1.69	1.45	1.72	0.89	1.56	1.74	1.37

Refer to notes in Table 4.3 for more information. The specification adopted is similar to Column (6) in Table 4.3. In Column (1), I account for face-width-height ratio of each inmate. In Column (2), I include zipcode fixed effects that control for the zipcode of the released inmate. In Column (3), I remove any inmates that are older than 70 years old. In Column (4), any felons that are charged with 5 or more crimes are omitted from the analysis. In column (5), I only examine inmates sentenced after 2005. In column (6), I exclude the top 10% least attractive faces. In column (7), any low resolution images less than 12,000 bytes are removed from the sample. In column (8), I include crime*year*court fixed effects. Standard errors clustered at court-by-year levels are reported in the parenthesis. ^{c} p<0.10, ^{b} p<0.05, ^{a} p<0.01.

1. *Vertical Symmetry:* As mentioned before, face symmetry only accounts for horizontal symmetry and do not account for faces that could be disproportionately wide or long. Therefore, I include face-width-height ratio (FWHR) as an additional control in Column (1). FWHR is computed by taking the ratio of the face height and the face width. This do no matter much as documented sentencing gaps remain fairly similar to before.

2. *Zipcode Fixed Effects:* If less attractive felons have poorer labour market outcomes (Hamermesh & Biddle, 1994; Biddle & Hamermesh, 1998) and prevent them from

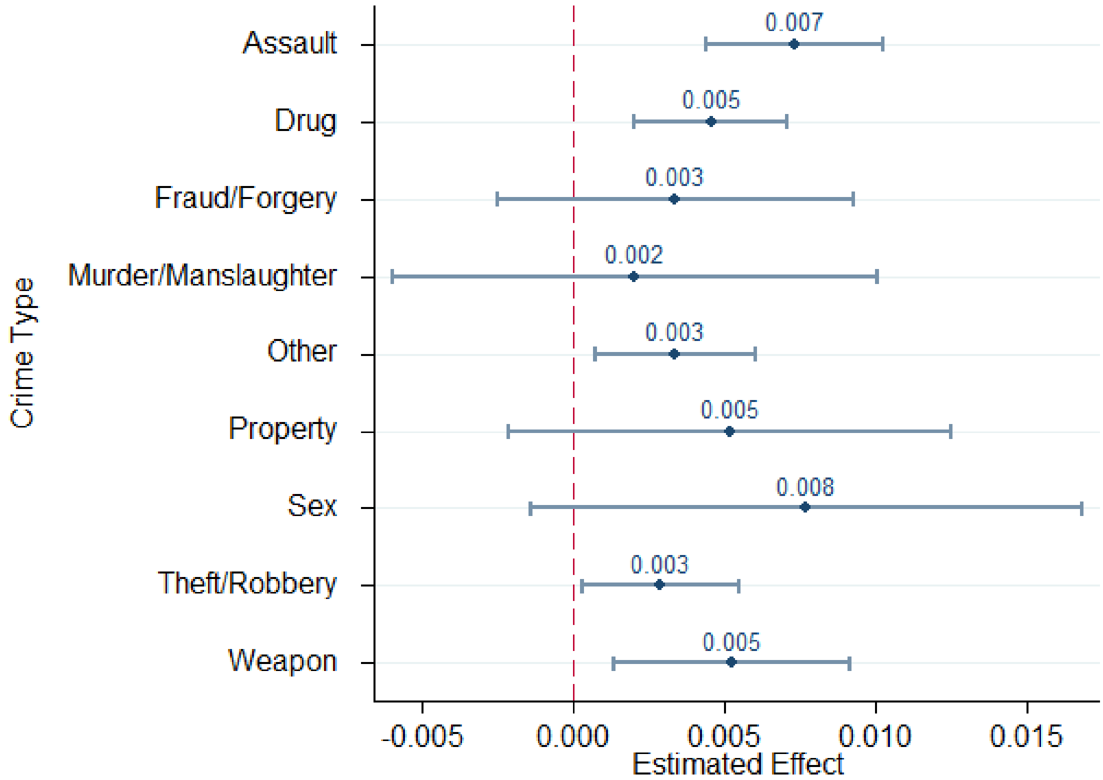


Figure 4.4: Effects of facial attractiveness on sentencing outcomes across different crimes. Each dot denotes the estimated effect (γ) of facial symmetry on the sentencing outcomes for each crime type. Tails indicate the 95% confidence intervals. The specification is similar to Column (5) of table 4.3 but I allow the discrimination towards facial attractiveness to vary across the 9 major crime types. Other crimes include escape from correction facility, kidnapping, racketeering, crime on elderly and other violent crimes.

hiring better attorney to defend their cases, criminals with less symmetric faces could be more harshly sentenced due to their lower socio-economic status (SES). Thus, I proxy for SES based on the neighbourhood of residence with the inclusion of zipcode fixed effects. The assumption is that inmates living in same zipcode should have fairly similar economic background. Doing so reduces the sample considerably by more than 40% as I only have residential address for released inmates. After accounting for disparity in SES, as shown in Column (2), the sentencing gap remains significant at around 19 days moving from the 25th to 75th percentile in facial symmetry. The smaller absolute sentence gaps despite larger estimated effects could be because sentencing lengths are on average shorter for those who are already released at the point of analysis.

3. *Age Effects:* If ageing felons have less symmetrical facial features, and judges are sympathetic towards elderly criminals, it is plausible that earlier estimates are underestimating the effect of facial symmetry on sentencing. Thus, in Column (1),

I remove any inmates who are more than 70 years old from the analysis[14]. Results remain similar to before in terms of both magnitude and statistical significance, ameliorating the concerns that earlier effects could be driven by age.

4. *Serial Felons:* Given that less attractive individuals could be forced to engage in illicit activities due to poorer labour outcomes (Mocan & Tekin, 2010), the concern is whether criminals with less proportional facial features are more likely to be repeated offenders. Since the harshness of punishment is influenced by the number of concurrent and historical charges, less attractive criminals could be more harshly sentenced even when there is no discrimination against appearance. Therefore, I limit the analysis to felons with 5 or less concurrent and historical charges in Column (2). Removing repeated offenders from the sample has a negligible impact on the estimates, suggesting that the discrimination against criminals with less symmetric faces are not driven by recalcitrant offenders.

5. *Measurement Error:* As mentioned, it is plausible that facial symmetry could be subjected to measurement error that could lead to attenuation bias. I consider the following cases that increase the risk that facial symmetry could be wrongly measured:

(1) Dated Pictures: The mugshots taken may not reflective of the appearance of inmates when facing trial. This might be the case as mugshots are updated periodically in the correction facilities. Although previous researches have argued that appearance is highly correlated over time (Hamermesh & Biddle, 1994), for inmates who are incarcerated for an extended period, the outlook in these pictures could still be different from how inmates appear in courts. Hence, I limit the analysis to more recent cases post 2005 to ensure that the mugshots are reflective of the felons' appearance during trial. Now, moving from the 25th to the 75th percentile in facial symmetry increases sentence length by almost 15 days, which represents a gap of 0.9%.

(2) Resolution of mugshots: Measurement error could also be intensified by the fuzziness of the images. Hence, I remove the bottom 25% of the images measured by the file size. Results summarized in Column (5) indicate that sentencing gaps between the 25th and 75th percentile are slightly higher now at 26 days when constrained to a sample of higher quality mugshots. This suggests that the measurement error associated with blurred photos are not driving the results.

[14]Results remain very stable even when I constraint the analysis to felons not more than 40 years old.

4.6.5 Discussion

There are still several issues yet to be addressed in this paper. Most of them are driven by the lack of data. One concern is the potential bias that arises from sample selection. In this analysis, I only observe cases when the felon is found guilty. If judges are lenient towards felons with more symmetric facial features, then my sample comprises of felons with asymmetric faces and felons with symmetric faces but some unobserved factors that could increase their conviction rates. This means that convicted felons with asymmetric faces are unlikely to be comparable with those convicted felons with symmetric faces.

Ideally, I would like to address this issue using a two-stage Heckman correction model but this would require me to observe all the cases, including suspects who are charged but not found guilty. Unfortunately, this data is not available. But how will this selection bias affect my estimates? Presumably, better-looking individuals found guilty are more likely to have stronger evidence presented against them. Conversely, bad-looking felons are more likely to be involved in borderline cases with weak evidence against them. Hence, based on evidences presented in courts, criminals with more symmetric faces are more likely to receive heftier punishment, which will attribute to an underestimation of discrimination on facial symmetry by judges. It is comforting to observe significant effects of facial symmetry on sentencing even in the presence of selection issues.

Another issue is that I do not have judge identifiers for each case. However, this should not be a major concern given that cases are randomly assigned to judges in Florida. Judges allocated to cases with better-looking felons should not be different, on average, from judges assigned to cases with worse-looking felons. Second, there are no details on the sentencing cells that each felon is allocated to based on the points given for the crime(s). This could exacerbate unobserved severity between cases. I address this limitation by including indicator variables for micro-crime categories, which are much finer than sentencing cells. This is permissable given that I have detailed information on the crimes committed by each felon that allows me to exploit variation of facial symmetry and sentencing outcomes within these precisely defined crime types. In addition, the enhanced flexibility of judges to depart from recommended sentences under the Criminal Punishment Code also meant that sentencing cells could be less effective in controlling for the severity of crimes.

4.7 Conclusion

In this paper, I examine whether the appearance of an inmate, measured by facial symmetry, affects sentencing outcomes. This research contributes to the burgeoning literature that documents the pervasiveness of appearance-based discrimination in different contexts, and the malpractices of judges allowing extraneous factors such as race, gender and emotions to influence decision-making.

This study employs facial recognition algorithms on mugshots of convicted felons to locate 69 different facial landmarks, including forehead, eyes, nose, mouth, ears and jawline to compute facial symmetry associated with each inmate. This automated way of detecting facial features allows this study to be conducted at much larger scale, incorporating the entire universe of more than 200,000 inmates in Florida. This is an improvement from previous studies that are usually restricted to small sample sizes as researchers are required to painstakingly collect multiple attractiveness ratings from respondents on subjects to objectively measure appearance. To address the concern of omitted con-founders from biasing the estimates, I include a rich set of controls on personal characteristics and case-related facts associated with each inmate, and crime-by-court fixed effects. Put differently, I am examining the effects of facial symmetry of inmates conducting the same crime (1294 categories) and being trialled at the same court (20 courts) on their sentencing outcomes.

Using a universe of sentencing outcomes associated with criminals from Florida, I observe that judges hand out harsher sentences to criminals with asymmetric facial features. The disparity in punishment between criminals with more symmetric faces, at the 25th percentile, and criminals with less symmetric faces, at the 75th percentile, is around 1.0% to 1.9% of the mean sentence length, which corresponds to between 17 and 32 days. Additional analyses reveal that this bias against felons with less proportional faces could vary across race, gender and type of crimes.

4.8 Data Appendix

Table 4.6: List of Variables

Variable	Source	Description
Dependent Variable (D_{ijct})		
Log(Sentence)	OBIS	The natural logarithm of the sentence length (in days) of inmate i for crime type j
Facial Attractiveness(FS_i)		
Facial Symmetry	-	Absolute deviation from a perfectly symmetrical face measured based on the locations of eyebrows, eyes, mid face, nose, mouth and lower face for inmate i
Personal Characteristics(X'_{ij})		
Gender	OBIS	Binary variable denoting whether inmate i is male or female
Ethnicity	OBIS	Binary variable denoting whether inmate i is black, white, hispanic or others
Address	OBIS	Zipcode of the residential address of inmate i upon release
Age	OBIS	Age of inmate i at sentencing date
Height	OBIS	Height of inmate i
Weight	OBIS	Weight of inmate i
Eye Color	OBIS	Binary variable denoting eye color of inmate i
Hair Color	OBIS	Binary variable denoting hair color of inmate i
Total Tattoo Counts	OBIS	Total counts of tattoo for inmate i
Facial Tattoo Counts	OBIS	Counts of facial tattoo for inmate i
Criminal History(X'_{ij})		
Concurrent Charges	OBIS	Number of concurrent charges for inmate i during sentence
Criminal History	OBIS	Total counts of historical charges for inmate i similar to the sentence charged
Mugshot Characteristics(M'_i)		
Adjustment Angle	-	Average of eyeline angle and noseline angle of inmate i
Face Edge to Eye Ratio	-	ratio of the distance from left face edge to left eye corner and the distance from right face edge to right eye corner for inmate i
Forehead-Bottom Ratio	-	Ratio of the distance between forehead and eyebrow to the distance between chin and mouth of inmate i
Picture Size	-	The picture size for inmate i in kilobytes

Bibliography

Abouk, Rahi, & Adams, Scott. 2013. Texting bans and fatal accidents on roadways: Do they work? Or do drivers just react to announcements of bans? *American Economic Journal: Applied Economics*, **5**(2), 179–199.

Abrams, David S, Bertrand, Marianne, & Mullainathan, Sendhil. 2012. Do Judges Vary in Their Treatment of Race? *The Journal of Legal Studies*, **41**(2), 347–383.

Administration, National Highway Traffic Safety, *et al.* 2014. The economic and societal impact of motor vehicle crashes, 2010. *Report DOT HS*, **812**, 013.

Agarwal, Sumit, Koo, Kang Mo, & Sing, Tien Foo. 2015. Impact of electronic road pricing on real estate prices in Singapore. *Journal of Urban Economics*, **90**, 50–59.

Alesina, Alberto, & La Ferrara, Eliana. 2014. A test of racial bias in capital sentencing. *The American Economic Review*, **104**(11), 3397–3433.

Alley, Thomas R, & Cunningham, Michael R. 1991. Article Commentary: Averaged Faces Are Attractive, but Very Attractive Faces Are Not Average. *Psychological science*, **2**(2), 123–125.

Anas, Alex, & Lindsey, Robin. 2011. Reducing urban road transportation externalities: Road pricing in theory and in practice. *Review of Environmental Economics and Policy*, 66–88.

Andersson, Henrik, Jonsson, Lina, & Ögren, Mikael. 2010. Property prices and exposure to multiple noise sources: Hedonic regression with road and railway noise. *Environmental and Resource Economics*, **45**(1), 73–89.

Anwar, Shamena, & Fang, Hanming. 2006. An alternative test of racial prejudice in motor vehicle searches: Theory and evidence. *American Economic Review*, **96**(1), 127–151.

Anwar, Shamena, Bayer, Patrick, & Hjalmarsson, Randi. 2012. The impact of jury race in criminal trials. *The Quarterly Journal of Economics*, **127**(2), 1017–1055.

Ashenfelter, Orley, & Greenstone, Michael. 2004. Using mandated speed limits to measure the value of a statistical life. *Journal of Political Economy*, **112**(S1), S226–S267.

Ayres, Ian, & Waldfogel, Joel. 1994. A market test for race discrimination in bail setting. *Stanford Law Review*, 987–1047.

Bayer, Patrick, Ferreira, Fernando, & McMillan, Robert. 2007. A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy*, **115**(4), 588–638.

Becker, Gary S. 1968. Crime and punishment: An economic approach. *Pages 13–68 of: The economic dimensions of crime*. Springer.

Beevers, Sean D, & Carslaw, David C. 2005. The impact of congestion charging on vehicle emissions in London. *Atmospheric Environment*, **39**(1), 1–5.

Berggren, Niclas, Jordahl, Henrik, & Poutvaara, Panu. 2010. The looks of a winner: Beauty and electoral success. *Journal of Public Economics*, **94**(1), 8–15.

Bernstein, Ira H, Lin, Tsai-Ding, & McClellan, Pamela. 1982. Cross-vs. within-racial judgments of attractiveness. *Perception & Psychophysics*, **32**(6), 495–503.

Biddle, Jeff E, & Hamermesh, Daniel S. 1998. Beauty, productivity, and discrimination: Lawyers' looks and lucre. *Journal of Labor Economics*, **16**(1), 172–201.

Black, Sandra E. 1999. Do better schools matter? Parental valuation of elementary education. *Quarterly Journal of Economics*, **114**(2), 577–599.

Boarnet, Marlon G. 1997. Infrastructure services and the productivity of public capital: the case of streets and highways. *National Tax Journal*, **50**(1), 39–57.

Butcher, Kristin F, Park, Kyung H, & Piehl, Anne Morrison. 2017. Comparing apples to oranges: Differences in women and men incarceration and sentencing outcomes. *Journal of Labor Economics*, **35**(S1), S201–S234.

Cameron, A Colin, & Trivedi, Pravin K. 2013. *Regression analysis of count data*. Vol. 53. Cambridge university press.

Chay, Kenneth Y, Greenstone, Michael, *et al.* 2005. Does Air Quality Matter? Evidence from the Housing Market. *Journal of Political Economy*, **113**(2), 376–424.

Chen, Greg, Meckle, Wayne, & Wilson, Jean. 2002. Speed and safety effect of photo radar enforcement on a highway corridor in British Columbia. *Accident Analysis & Prevention*, **34**(2), 129–138.

Christie, SM, Lyons, Ronan Anthony, Dunstan, Frank D, & Jones, Sarah J. 2003. Are mobile speed cameras effective? A controlled before and after study. *Injury Prevention*, **9**(4), 302–306.

Cohen, Alma, & Einav, Liran. 2003. The effects of mandatory seat belt laws on driving behavior and traffic fatalities. *Review of Economics and Statistics*, **85**(4), 828–843.

Cunningham, Christopher, Hummer, Joseph, & Moon, Jae-Pil. 2008. Analysis of automated speed enforcement cameras in Charlotte, North Carolina. *Transportation Research Record: Journal of the Transportation Research Board*, 127–134.

Currie, Janet, & Walker, Reed. 2011. Traffic congestion and infant health: Evidence from E-ZPass. *American Economic Journal: Applied Economics*, **3**(1), 65–90.

Currie, Janet, Davis, Lucas, Greenstone, Michael, & Walker, Reed. 2015. Environmental health risks and housing values: evidence from 1,600 toxic plant openings and closings. *The American economic review*, **105**(2), 678–709.

Danziger, Shai, Levav, Jonathan, & Avnaim-Pesso, Liora. 2011. Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences*, **108**(17), 6889–6892.

Davis, Lucas W. 2004. The effect of health risk on housing values: Evidence from a cancer cluster. *The American Economic Review*, **94**(5), 1693–1704.

De Borger, Bruno. 2009. Commuting, congestion tolls and the structure of the labour market: Optimal congestion pricing in a wage bargaining model. *Regional Science and Urban Economics*, **39**(4), 434–448.

DeAngelo, Gregory, & Hansen, Benjamin. 2014. Life and death in the fast lane: Police enforcement and traffic fatalities. *American Economic Journal: Economic Policy*, **6**(2), 231–257.

Dee, Thomas S. 1999. State alcohol policies, teen drinking and traffic fatalities. *Journal of Public Economics*, **72**(2), 289–315.

Department of Environment, Transport, & the Regions. 1998. Traffic Speeds in Inner London. **98:22**.

DfT. 2004. Handbook of Rules and Guidance for the National Safety Camera Programme for England and Wales: 2005/2006.

DfT. 2005. Handbook of Rules and Guidance for the National Safety Camera Programme for England and Wales: 2006/2007.

DfT. 2014. Values of Time and Vehicle Operating Costs.

DfT. 2016. Reported Road Casualties in Great Britain: Main Results 2015.

Dion, Karen, Berscheid, Ellen, & Walster, Elaine. 1972. What is beautiful is good. *Journal of personality and social psychology*, **24**(3), 285.

Duarte, Jefferson, Siegel, Stephan, & Young, Lance. 2012. Trust and credit: The role of appearance in peer-to-peer lending. *The Review of Financial Studies*, **25**(8), 2455–2484.

Duranton, Gilles, & Turner, Matthew A. 2011. The fundamental law of road congestion: Evidence from US cities. *American Economic Review*, **101**(6), 2616–2652.

Efran, Michael G. 1974. The effect of physical appearance on the judgment of guilt, interpersonal attraction, and severity of recommended punishment in a simulated jury task. *Journal of Research in Personality*, **8**(1), 45–54.

Elvik, Rune. 1997. Effects on accidents of automatic speed enforcement in Norway. *Transportation Research Record: Journal of the Transportation Research Board*, 14–19.

Eren, Ozkan, & Mocan, Naci. 2016. Emotional judges and unlucky juveniles.

Feingold, Alan. 1992. Good-looking people are not what we think. *Psychological bulletin*, **111**(2), 304.

Fernald, John G. 1999. Roads to prosperity? Assessing the link between public capital and productivity. *American Economic Review*, 619–638.

Gains, Adrian, Heydecker, Benjamin, Shrewsbury, John, & Robertson, Sandy. 2004. The national safety camera programme: Three-year evaluation report.

Gains, Adrian, Nordstrom, Michael, Heydecker, BG, & Shrewsbury, John. 2005. The national safety camera programme: Four-year evaluation report.

Gallagher, Justin, & Fisher, Paul J. 2017. Criminal Deterrence when there are Offsetting Risks: Traffic Cameras, Vehicular Accidents, and Public Safety.

Gayer, Ted, Hamilton, James T, & Viscusi, W Kip. 2000. Private values of risk tradeoffs at superfund sites: housing market evidence on learning about risk. *Review of Economics and Statistics*, **82**(3), 439–451.

Gibbons, Stephen, & Machin, Stephen. 2005. Valuing rail access using transport innovations. *Journal of Urban Economics*, **57**(1), 148–169.

Gibbons, Stephen, Machin, Stephen, & Silva, Olmo. 2013. Valuing school quality using boundary discontinuities. *Journal of Urban Economics*, **75**, 15–28.

Gibbons, Steve. 2004. The costs of urban property crime. *The Economic Journal*, **114**(499), F441–F463.

Goldenbeld, Charles, & van Schagen, Ingrid. 2005. The effects of speed enforcement with mobile radar on speed and accidents: An evaluation study on rural roads in the Dutch province Friesland. *Accident Analysis & Prevention*, **37**(6), 1135–1144.

Graham, Daniel J. 2007. Variable returns to agglomeration and the effect of road traffic congestion. *Journal of Urban Economics*, **62**(1), 103–120.

Graham, John R, Harvey, Campbell R, & Puri, Manju. 2016. A corporate beauty contest. *Management Science*.

Grammer, Karl, & Thornhill, Randy. 1994. Human (Homo sapiens) facial attractiveness and sexual selection: the role of symmetry and averageness. *Journal of comparative psychology*, **108**(3), 233.

Green, Colin, Heywood, John Spencer, & Navarro Paniagua, Maria. 2018. Did the London Congestion Charge Reduce Pollution?

Green, Colin P, Heywood, John S, & Navarro, Maria. 2016. Traffic accidents and the London congestion charge. *Journal of Public Economics*, **133**, 11–22.

Hamermesh, Daniel S, & Biddle, Jeff E. 1994. Beauty and the Labor Market. *The American Economic Review*, **84**(5), 1174–1194.

Hansen, Benjamin. 2015. Punishment and deterrence: Evidence from drunk driving. *The American Economic Review*, **105**(4), 1581–1617.

Hatfield, Elaine, & Sprecher, Susan. 1986. *Mirror, mirror: The importance of looks in everyday life*. Suny Press.

Hess, Stephane, & Polak, John. 2003. Effects of speed limit enforcement cameras on accident rates. *Transportation Research Record: Journal of the Transportation Research Board*, 25–33.

Ho, Teck-Hua, Chong, Juin Kuan, & Xia, Xiaoyu. 2017. Yellow taxis have fewer accidents than blue taxis because yellow is more visible than blue. *Proceedings of the National Academy of Sciences*, **114**(12), 3074–3078.

Hooke, Andrew, Knox, Jim, & Portas, David. 1996. *Cost benefit analysis of traffic light & speed cameras*. Home Office, Police Research Group.

Hughes, William T, & Sirmans, CF. 1992. Traffic externalities and single-family house prices. *Journal of Regional Science*, **32**(4), 487–500.

Hymel, Kent. 2009. Does traffic congestion reduce employment growth? *Journal of Urban Economics*, **65**(2), 127–135.

Jacobson, Marsha B. 1981. Effects of victim's and defendant's physical attractiveness on subjects' judgments in a rape case. *Sex Roles*, **7**(3), 247–255.

Jenq, Christina, Pan, Jessica, & Theseira, Walter. 2015. Beauty, weight, and skin color in charitable giving. *Journal of Economic Behavior & Organization*, **119**, 234–253.

Jones, Andrew P, Sauerzapf, Violet, & Haynes, Robin. 2008. The effects of mobile speed camera introduction on road traffic crashes and casualties in a rural county of England. *Journal of safety research*, **39**(1), 101–110.

Jones, Doug, & Hill, Kim. 1993. Criteria of facial attractiveness in five populations. *Human Nature*, **4**(3), 271–296.

Kazemi, Vahid, & Sullivan, Josephine. 2014. One millisecond face alignment with an ensemble of regression trees. *Pages 1867–1874 of: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.

Keall, Michael D, Povey, Lynley J, & Frith, William J. 2001. The relative effectiveness of a hidden versus a visible speed camera programme. *Accident Analysis & Prevention*, **33**(2), 277–284.

Knittel, Christopher R, Miller, Douglas L, & Sanders, Nicholas J. 2016. Caution, drivers! Children present: Traffic, pollution, and infant health. *Review of Economics and Statistics*, **98**(2), 350–366.

Knowles, John, Persico, Nicola, & Todd, Petra. 2001. Racial bias in motor vehicle searches: Theory and evidence. *Journal of Political Economy*, **109**(1), 203–229.

Kuminoff, Nicolai V, Parmeter, Christopher F, & Pope, Jaren C. 2010. Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities? *Journal of Environmental Economics and Management*, **60**(3), 145–160.

Langlois, Judith H, & Roggman, Lori A. 1990. Attractive faces are only average. *Psychological science*, **1**(2), 115–121.

Langlois, Judith H, Kalakanis, Lisa, Rubenstein, Adam J, Larson, Andrea, Hallam, Monica, & Smoot, Monica. 2000. Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological bulletin*, **126**(3), 390.

Leape, Jonathan. 2006. The London congestion charge. *The Journal of Economic Perspectives*, **20**(4), 157–176.

Levitt, Steven D, & Porter, Jack. 2001a. How dangerous are drinking drivers? *Journal of political Economy*, **109**(6), 1198–1237.

Levitt, Steven D, & Porter, Jack. 2001b. Sample selection in the estimation of air bag and seat belt effectiveness. *The review of economics and statistics*, **83**(4), 603–615.

Li, Haojie, & Graham, Daniel J. 2016. Heterogeneous treatment effects of speed cameras on road safety. *Accident Analysis & Prevention*, **97**, 153–161.

Li, Haojie, Graham, Daniel J, & Majumdar, Arnab. 2012. The effects of congestion charging on road traffic casualties: A causal analysis using difference-in-difference estimation. *Accident Analysis & Prevention*, **49**, 366–377.

Li, Haojie, Graham, Daniel J, & Majumdar, Arnab. 2013. The impacts of speed cameras on road accidents: An application of propensity score matching methods. *Accident Analysis & Prevention*, **60**, 148–157.

Lim, Claire SH, Snyder Jr, James M, & Strömberg, David. 2015. The judge, the politician, and the press: newspaper coverage and criminal sentencing across electoral systems. *American Economic Journal: Applied Economics*, **7**(4), 103–135.

Malpass, Roy S, & Kravitz, Jerome. 1969. Recognition for faces of own and other race. *Journal of personality and social psychology*, **13**(4), 330.

Mobius, Markus M, & Rosenblat, Tanya S. 2006. Why beauty matters. *The American Economic Review*, **96**(1), 222–235.

Mocan, Naci, & Tekin, Erdal. 2010. Ugly criminals. *The review of economics and statistics*, **92**(1), 15–30.

Moulds, Elizabeth F. 1978. Chivalry and paternalism: Disparities of treatment in the criminal justice system. *Western Political Quarterly*, **31**(3), 416–430.

Mountain, LJ, Hirst, WM, & Maher, MJ. 2004. Costing lives or saving lives: a detailed evaluation of the impact of speed cameras. *Traffic, Engineering and Control*, **45**(8), 280–287.

Mountain, LJ, Hirst, WM, & Maher, MJ. 2005. Are speed enforcement cameras more effective than other speed management measures?: The impact of speed management schemes on 30mph roads. *Accident Analysis & Prevention*, **37**(4), 742–754.

Mustard, David B. 2001. Racial, ethnic, and gender disparities in sentencing: Evidence from the US federal courts. *The Journal of Law and Economics*, **44**(1), 285–314.

Newbery, David M. 1990. Pricing and congestion: economic principles relevant to pricing roads. *Oxford Review of Economic Policy*, **6**(2), 22–38.

Newstead, Stuart, & Cameron, Maxwell H. 2003. *Evaluation of the crash effects of the Queensland speed camera program.* Monash University Accident Research Centre Victoria, Australia.

Palmquist, Raymond B. 1992. Valuing localized externalities. *Journal of Urban Economics*, **31**(1), 59–68.

Park, Kyung H. 2017. Do judges have tastes for discrimination? evidence from criminal courts. *Review of Economics and Statistics*, **99**(5), 810–823.

Peden, Margie, Scurfield, Richard, Sleet, David, Mohan, Dinesh, Hyder, Adnan A, Jarawan, Eva, Mathers, Colin D, *et al.* 2004. *World report on road traffic injury prevention.*

Peltzman, Sam. 1975. The effects of automobile safety regulation. *Journal of political Economy*, **83**(4), 677–725.

Percoco, Marco. 2014. The impact of road pricing on housing prices: preliminary evidence from Milan. *Transportation Research Part A: Policy and Practice*, **67**, 188–194.

Perez, Katherine, Mari Dell Olmo, Marc, Tobias, Aurelio, & Borrell, Carme. 2007. Reducing road traffic injuries: effectiveness of speed cameras in an urban setting. *American journal of public health*, **97**(9), 1632–1637.

Perrett, David I, Lee, Kieran J, Penton-Voak, Ian, Rowland, D, Yoshikawa, Sakiko, Burt, D Michael, Henzi, SP, Castles, Duncan L, & Akamatsu, Shigeru. 1998. Effects of sexual dimorphism on facial attractiveness. *Nature*, **394**(6696), 884.

Perrett, David I, Burt, D Michael, Penton-Voak, Ian S, Lee, Kieran J, Rowland, Duncan A, & Edwards, Rachel. 1999. Symmetry and human facial attractiveness. *Evolution and human behavior*, **20**(5), 295–307.

Pigou, Arthur Cecil. 1924. *The Economics of Welfare*. Palgrave Macmillan.

Ravina, Enrichetta, Gabriel, Suggestions Paul, Galak, Jeff, Gokli, Ami, Munro, Amelia, Patel, Harshi, & Qian, Dawei. 2008. Love & loans: the effect of beauty and personal characteristics in credit markets,SSRN Working Paper 1101647.

Rhodes, Gillian. 2006. The evolutionary psychology of facial beauty. *Annu. Rev. Psychol.*, **57**, 199–226.

Richards, Martin Gomm. 2006. *Congestion charging in London-The policy and the politics.*

Rosén, Erik, & Sander, Ulrich. 2009. Pedestrian fatality risk as a function of car impact speed. *Accident Analysis & Prevention*, **41**(3), 536–542.

Rosen, Sherwin. 1974. Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy*, **82**(1), 34–55.

Scheib, Joanna E, Gangestad, Steven W, & Thornhill, Randy. 1999. Facial attractiveness, symmetry and cues of good genes. *Proceedings of the Royal Society of London B: Biological Sciences*, **266**(1431), 1913–1917.

Shayo, Moses, & Zussman, Asaf. 2011. Judicial ingroup bias in the shadow of terrorism. *The Quarterly Journal of Economics*, **126**(3), 1447–1484.

Shin, Kangwon, Washington, Simon P, & Van Schalkwyk, Ida. 2009. Evaluation of the Scottsdale Loop 101 automated speed enforcement demonstration program. *Accident Analysis & Prevention*, **41**(3), 393–403.

Sigall, Harold, & Ostrove, Nancy. 1975. Beautiful but dangerous: effects of offender attractiveness and nature of the crime on juridic judgment. *Journal of Personality and Social Psychology*, **31**(3), 410.

Small, Kenneth A, Winston, Clifford, & Yan, Jia. 2005. Uncovering the distribution of motorists' preferences for travel time and reliability. *Econometrica*, **73**(4), 1367–1382.

Spohn, Cassia. 1999. Gender and sentencing of drug offenders: Is chivalry dead? *Criminal Justice Policy Review*, **9**(3-4), 365–399.

Starr, Sonja B. 2014. Estimating gender disparities in federal criminal cases. *American Law and Economics Review*, **17**(1), 127–159.

Stewart, John E. 1980. Defendant's attractiveness as a factor in the outcome of criminal trials: An observational study. *Journal of Applied Social Psychology*, **10**(4), 348–361.

TfL. 2003a. Congestion Charging: Six months on.

TfL. 2003b. Impacts Monitoring Programme: First Annual Report.

TfL. 2004. Impacts Monitoring Programme: Second Annual Report.

TfL. 2005. Impacts Monitoring Programme: Third Annual Report.

TfL. 2009. Proposed Western Extension of the Central London Congestion Charging Scheme.

Thaler, Richard. 1978. A note on the value of crime control: evidence from the property market. *Journal of Urban Economics*, **5**(1), 137–145.

van Benthem, Arthur. 2015. What is the optimal speed limit on freeways? *Journal of Public Economics*, **124**, 44–62.

Vickrey, William S. 1963. Pricing in urban and suburban transport. *The American Economic Review*, **53**(2), 452–465.

Wilson, Cecilia, Willis, Charlene, Hendrikz, Joan K, Le Brocque, Robyne, & Bellamy, Nicholas. 2010. Speed cameras for the prevention of road traffic injuries and deaths. *The Cochrane Library*.

Zebrowitz, Leslie A, & McDonald, Susan M. 1991. The impact of litigants' baby-facedness and attractiveness on adjudications in small claims courts. *Law and human behavior*, **15**(6), 603.

Zhang, Yi, & Shing, Hui-Fai. 2006. The London congestion charge and property prices: an evaluation of the impact on property prices inside and outside the zone. *MPRA Paper*, **4050**.

List of Figures

List of Tables