Title
Essays in Public Economics and Urban Economics

Author
Gan, Zhongying

Publication Date
2024

UNIVERSITY OF CALIFORNIA,
IRVINE

Essays in Public Economics and Urban Economics

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Economics

by

Zhongying Gan

Dissertation Committee:
Professor Jan K. Brueckner, Chair
Professor Matthew Freedman
Professor Yingying Dong

2024

TABLE OF CONTENTS

Page

LIST OF FIGURES **v**

LIST OF TABLES **vi**

ACKNOWLEDGMENTS **vii**

VITA **viii**

ABSTRACT OF THE DISSERTATION **ix**

1 Inspecting the Inspectors: Causes and Consequences of Restaurant Inspector Leniency **1**

1.1 Introduction 1
1.2 Background 8
1.2.1 Types of inspections 8
1.2.2 Inspector assignment, compensation and evaluation 11
1.3 Data 12
1.3.1 Routine inspection and OII 12
1.3.2 Complaint investigation 14
1.4 Inspector leniency 15
1.5 Exploring the public health consequences of grade inflation 17
1.5.1 Grade inflation misleads diners about restaurants' hygiene conditions 17
1.5.2 Grade inflation discourages hygiene improvements 20
1.6 Motives behind grade inflation 23
1.6.1 OII avoidance 24
1.6.2 Attachment 31
1.7 Model 34
1.7.1 Model setup 34
1.7.2 Models of restaurants maximizing expected profit under inspector uncertainty 36
1.8 Simulations under the hybrid model 41
1.8.1 Calibration of the hybrid model 41
1.8.2 Evaluating the proposed remedies to inspector leniency 44
1.9 Conclusion 47

2 Do Electric Vehicle Charger Locations Respond to the Potential Charging Demands from Multi-Unit Dwellings? Evidence from Los Angeles County 59
2.1 Introduction 59
2.2 Data 69
2.3 Model 72
2.4 Empirical results 77
2.4.1 The extensive margin of charger supply 77
2.4.2 The intensive margin of charger supply 79
2.4.3 The ZINB model 83
2.4.4 Robustness Checks 85
2.5 Conclusion 86

3 Freeway Congestion and Labor Force Participation: Evidence from Greater Los Angeles 103
3.1 Introduction 103
3.2 Data 106
3.2.1 Labor market outcomes 106
3.2.2 Congestion 108
3.3 Empirical framework and results 111
3.3.1 Congestion and commute flow 111
3.3.2 Congestion and labor market outcomes 116
3.4 Conclusion 119

References 125

Appendix A Chapter 1 138
A.1 Score distributions before and after the policy change 138
A.2 Grade inflation misleads diners about restaurants' hygiene conditions: Discussions 139
A.2.1 Robustness checks 139
A.2.2 Sample selection 139
A.3 Grade inflation discourages hygiene improvements: Details on estimation . . 141
A.3.1 Estimating the score of a routine inspection in the absence of grade inflation 141
A.3.2 Robustness checks 143
A.4 Possible explanations for grade inflation under repeated interactions between an inspector and a restaurant 144
A.5 Threshold model 145
A.6 Restaurants' price level data from Yelp.com 147

Appendix B Chapter 2 157
B.1 Comparing the findings to those of Hsu and Fingerman (2021) and Khan et al. (2022) 157
B.2 The construction of MUD and facility variables 158
B.2.1 MUDs 158

B.2.2 Hotels 158
B.2.3 Office buildings 159
B.2.4 Retail 159
B.2.5 Car dealers 160
B.2.6 Government offices 160
B.2.7 Schools 160
B.2.8 Transportation stations 160
B.3 Accessible distance from chargers 161
B.4 Details on the back-of-the-envelope calculation of MUD charger-to-PEV-ratio 162
B.4.1 Estimating the PEV ownership rate among MUD residents 162
B.4.2 Mapping increase in MUD density onto increase in MUD units 163

Appendix C Chapter 3 **166**
C.1 Details on estimating congestion 166
C.1.1 Cleaning the speed data from stations 166
C.1.2 Availability of freeway speed 167
C.1.3 Free-flow speed 168
C.1.4 Estimating route and travel time 168
C.2 Reliability of the historical travel time estimation 169
C.3 Control variables 172
C.3.1 Percentage of female 172
C.3.2 Percentage of families with children under six 172
C.3.3 Housing price 172
C.3.4 Housing stock 173
C.4 Robustness checks 173

LIST OF FIGURES

Page

1.1 Distribution of routine inspection scores in the full sample 49
1.2 Comparisons of the probability of a subsequent complaint investigation where violations are found among score groups 90 to 100 50
1.3 Understanding OIIs 51
1.4 Distributions of inspector uncertainty 52
1.5 The simulated distributions of optimal efforts and scores under a non-lenient inspector 53
1.6 The simulated distributions of optimal efforts and scores under a lenient inspector 53
1.7 Simulated distributions of scores vs. Empirical distributions of scores 55
1.8 Simulated distributions of optimal efforts under different values of θ 56

2.1 Spatial patterns of the number of L2 chargers and MUD density 88
2.2 MUD charger-to-PEV-ratio across CBGs in LA County 89

LIST OF TABLES

Page

1.1 The effects of grade inflation on restaurant hygiene 57
1.2 Verification of OII avoidance as a motivation behind grade inflation to A . . 58

2.1 Summary statistics 90
2.2 What CBG characterisitics are associated with charger existence in a CBG? 92
2.3 Regression results under the OLS model and the Tobit model 95
2.4 Estimation of the intensive margin in Tobit model 97
2.5 Regression results under the ZINB model 99
2.6 Robustness checks using truncated population-weighted inverse-distance-squared matrix 101

3.1 Summary statistics 121
3.2 Commute flows between tract pairs and lagged congestion along commute routes 122
3.3 Labor force participation rate and lagged congestion 123
3.4 Earnings of residents in tracts facing increasing commute congestion 124

ACKNOWLEDGMENTS

I would like to thank my dissertation chair Jan Brueckner for his incredible mentorship throughout my Ph.D. study, and for always showing excitement for my research. His encouragement always inspired me to preservere through challenges. Specifically, I would like to thank him for encouraging me to reach out to Los Angeles County Department of Public Health to gather institutional knowledge for the first chapter of my dissertation. This chapter is my first independent scholarly project, and I thank Jan for his guidance. I am also thankful to him for helping me navigate through the peer review process for my first publication in *Transport Policy*.

In addition, I would like to thank my dissertation committee members, Matt Freedman and Yingying Dong, and other faculty members at UC Irvine for their guidance and advice. I am also grateful to friends and peers at UC Irvine who have offered me company and help.

I am forever grateful to my partner and friend, Kyle Colanglo, and other family for their love and support.

I also wish to thank Debapriya Chakraborty, Sarah Thomaz, various informants at the Los Angeles County Department of Public Health, an informant from the City of Los Angeles, former Mayor Eric Garcetti's Office of Sustainability, and various informants from a power agency for their kind assistance during this dissertation project.

I would like to thank the Graduate Division, the School of Social Sciences, and the Department of Economics at UC Irvine for financial support.

Chapter 2 of this dissertation is a reprint of the material as it appears in "Do Electric Vehicle Charger Locations Respond to the Potential Charging Demands from Multi-Unit Dwellings? Evidence from Los Angeles County" in *Transport Policy*, Volume 138, Pages 74-93, used with permission from Elsevier Ltd.

VITA

Zhongying Gan

EDUCATION

Doctor of Philosophy in Economics **2024**
University of California, Irvine *Irvine, California*

Master of Arts in Economics **2019**
University of California, Irvine *Irvine, California*

Bachelor of Arts in Mathematical Economics **2018**
Fudan University *Shanghai, China*

RESEARCH EXPERIENCE

Graduate Research Assistant **2019–2020**
University of California, Irvine *Irvine, California*

TEACHING EXPERIENCE

Teaching Assistant **2019–2023**
University of California, Irvine *Irvine, California*

REFEREED JOURNAL PUBLICATIONS

Do Electric Vehicle Charger Locations Respond to the Potential Charging Demands from Multi-Unit Dwellings? Evidence from Los Angeles County **2023**
Transport Policy, 138 (2023): 74-93

AWARDS

Sheen T. Kassouf Endowed Fellowship **2023**
Department of Economics, UCI *Irvine, California*

Ken Small Award for Best Student Paper in Transportation Economics **2022**
Irvine, California
Department of Economics, UCI

ABSTRACT OF THE DISSERTATION

Essays in Public Economics and Urban Economics

By

Zhongying Gan

Doctor of Philosophy in Economics

University of California, Irvine, 2024

Professor Jan K. Brueckner, Chair

This dissertation consists of three chapters, where I use a combination of empirical methods and model simulations to study topics in public economics and urban economics.

Chapter 1 studies inspector leniency as a challenge to regulation enforcement in the context of Los Angeles County restaurant hygiene inspections. Inspectors score restaurants numerically, but only letter grades reflecting wide score intervals are mandatorily disclosed. In this chapter, I show that inspector leniency has compromised the effectiveness of health grades in signaling restaurants' hygiene conditions, as restaurants whose grades are inflated to A have significantly worse hygiene conditions than restaurants who have earned their A. Moreover, I find that inspector leniency discourages hygiene improvements by preventing a re-inspection that a restaurant would likely request if they did not experience grade inflation. This chapter then uncovers a novel motive behind inspector leniency: inspectors avoiding work that does not increase promotion prospects, and proposes changes to the inspector performance evaluation policy as a remedy. Lastly, I evaluate the effects of policy interventions aimed at reducing inspector leniency on inspectors' grading and restaurants' sanitation efforts under a theoretical framework.

Chapter 2 develops an empirical framework to evaluate plug-in electric vehicle (PEV) charger sufficiency. With the transportation sector being the largest contributor of green-

house gas emissions in the United States, states are actively planning the deployment of PEVs. Mass adoption of PEVs requires attracting potential buyers living in multi-unit dwellings (MUDs). Given the current low adoption rate of PEVs in MUDs and MUD residents' reliance on public charging, this chapter studies whether there is a positive correlation between the number of public L2 chargers and MUD density (measured by total square footage of MUD per capita) across census block groups (CBGs) in LA County. The results show that high MUD-density CBGs and low MUD-density CBGs do not differ much in terms of the number of chargers. The charger-to-PEV-ratio range for MUD residents derived in this paper is below the ideal charger-to-PEV-ratio range in the literature. A direct policy implication is that more charging infrastructure should be made available to MUD residents.

Chapter 3 studies how people respond to freeway congestion along their commute to work in Greater Los Angeles. Traffic congestion is a growing problem in major metropolitan areas. This chapter develops a novel approach to measuring historical congestion by using open-sourced road network data augmented with historical speeds collected from freeway traffic detectors. Using annual data on tract-to-tract commute flows within Greater Los Angeles and congestion on the commute from 2010 to 2019, I find that the number of commuters on a given route decreases as congestion along the route increases. In addition, I find suggestive evidence that this decrease more likely results from workers changing jobs for a less-congested commute at the cost of lower pay, than from workers relocating their residences for a less-congested commute, or dropping out of the labor market.

Chapter 1

Inspecting the Inspectors: Causes and Consequences of Restaurant Inspector Leniency

1.1 Introduction

Regulations are ubiquitous in the modern economy: environmental regulations, financial regulations, building regulations, food safety regulations, etc. However, effective enforcement has always been a challenge for policymakers. This paper studies potential sources of enforcement failure and its consequences in the context of restaurant food safety regulations.

Restaurant food safety is a central public health concern. According to Statista, a German online portal for statistics, food and drink sales in the United States restaurant industry reached 745.61 billion U.S. dollars in 2015. Almost 19 million people reported visiting a full service restaurant and over 49 million people reported visiting a quick service restaurant in the United States in the spring of 2016.[1] However, restaurants can generate significant foodborne disease risk. Among the 841 foodborne disease outbreaks[2] reported by 50 states, Washington, D.C., and Puerto Rico in 2017, 489 (58.1%) outbreaks were associated with restaurants (CDC, 2019). In light of the public health risks associated with restaurants, restaurant hygiene inspections are widely adopted by jurisdictions to enforce food safety regulations (Ho, 2012). To induce compliance, inspection results are required to be displayed in restaurants, and are available on the local public health agency's website. Inspection results are also widely available on Yelp.com.[3]

How effective restaurant inspections and inspection-result disclosure are as a way to enforce safety regulations depends highly on the practice of the inspectors. Unfortunately, numerous cases have been documented where inspectors under-report the number or the severity of violations (Jin and Lee, 2018; Ibanez and Toffel, 2020; Kovács et al., 2020; Makofske, 2020b), a practice referred to as inspector leniency. Therefore, understanding inspector leniency is essential in evaluating and potentially improving the restaurant inspection and disclosure policy. However, little is known about the causes and consequences of inspector leniency.

In the context of Los Angeles (LA) County restaurant hygiene inspections, this paper examines the public health consequences of inspector leniency, investigates the inspectors' motives behind grade inflation, proposes corresponding remedies, and evaluates them through a theoretical model. Each restaurant in LA County is assigned a score between 0 to 100 during an unannounced routine inspection, with 90 to 100 points corresponding to grade A, 80 to 89 points corresponding to grade B, and 70 to 79 points corresponding to grade C. Only letter grades are disclosed to the public both through grade cards displayed in the

[1] Retrieved on April 21, 2020 at https://www.statista.com/topics/1957/eating-out-behavior-in-the-us/.

[2] A foodborne disease outbreak is defined by Centers for Disease Control and Prevention (CDC) as an incident in which two or more people experience a similar illness resulting from the ingestion of a common food.

[3] Restaurant inspection results displayed on Yelp.com either come directly from local government agencies that perform health inspections, or from third-party companies that partner with Yelp.

restaurants and on Yelp.com. The inspections are characterized by an unusually large percentage of score 90 and grade A, raising questions about the informativeness of the grade.[4] I find that restaurants whose grades are likely inflated to A have a larger probability of receiving a complaint investigation where violations are found than restaurants that have earned their A. This finding shows that inspector leniency has compromised health grades' effectiveness in signaling restaurants' hygiene conditions. Moreover, inspector leniency discourages hygiene improvements by preventing a re-inspection that motivates hygiene effort, which a restaurant would likely request if they did not experience grade inflation.

The paper then uncovers a novel motive behind inspector leniency — career incentives. If a restaurant owner is dissatisfied with the grade earned in a routine inspection, they can file for an owner-initiated inspection (OII), during which time the same inspector from the previous routine inspection repeats the inspection, and updates the grade. I first demonstrate that a B-grade is a strong predictor of an OII request, as the probability of an OII request increases sharply when the score of the last routine inspection drops below 90. I then present a previously undiscussed, yet critical detail in the inspector promotion policy that creates incentives against performing OIIs: Though routine inspections and OIIs are both full inspections, only the number of routine inspections is counted into performance evaluation. Empirical evidence shows that inspectors are more likely to inflate a restaurant's grade to A when they have a heavier past OII workload. This finding suggests that inspectors avoid OII workload by inflating a restaurant's grade to A, and establishes career incentives as one of the motives behind inspector leniency.

The paper follows by briefly discussing another motive behind inspector leniency that has been documented in the literature (Short et al., 2016; Jin and Lee, 2018; Kovács et al., 2020): attachment to the restaurants formed through repeated inspections. The two motives

[4] Using restaurant inspection data from June 1st, 2017 to December 31st, 2019, this paper finds that more than 13% of the routine inspections end up with a score of 90, causing a large spike at 90 in the score distribution (refer to Figure 1.1), and that more than 94% of the routine inspections end up with an A.

call for drastically different remedies: revisions of inspector performance evaluation policy versus more frequent inspector rotation, which highlights the importance of understanding inspector motives in improving enforcement.

To evaluate how the two proposed remedies to inspector leniency will affect inspectors' grading and restaurants' sanitation efforts, I construct a model where a restaurant maximizes expected profit based on a belief of how an inspector would assess its sanitation effort. A lenient inspector is modelled as always assessing a restaurant's sanitation effort to be larger than its actual effort. A non-lenient inspector is modelled as equally likely to overassess or underassess effort. The model shows that a restaurant exerts less sanitation effort under a lenient inspector than it would have under a non-lenient inspector. I then allow the inspector leniency level to change continuously by modelling an inspector as lenient to a proportion of restaurants, while non-lenient to the remaining restaurants. This hybrid model is calibrated to match empirical data patterns, and then used to study how the remedies to inspector leniency are going to affect inspectors' grading and restaurants' sanitation efforts. Simulations show that under both remedies, the extent of bunching at 90 decreases and restaurants' sanitation efforts increase.

Understanding inspector leniency has important policy implications, as the past effort to increase the informativeness of restaurant health grades in LA County without addressing inspector leniency or even acknowledging its existence was in vain. Stephanie Baer (2015) at the San Gabriel Valley Tribune revealed to the readers the unusually large percentage of score 90 and grade A in LA county restaurant inspections, and claimed that the grades are misleading the public about the actual operating conditions in the kitchen. The article prompted the LA County Department of Public Health (LADPH) to conduct a thorough review of the retail food facility grading system from June 2015 to June 2016. As a result, effective July 2016 with a six-month grace period, if two or more major Critical Risk Violations (each worth four-point deduction) are marked, an additional three points will be deducted, leaving a restaurant with an 89 or below. If a restaurant's permit is suspended due to no water available, improper sewage disposal, or vermin infestation (each worth four-point deduction), an additional seven points will be deducted for each of the three permit-suspension violations cited, also leaving the restaurant with an 89 or below.[5] This policy change, aimed at raising the credibility of the grade[6] and precluding the issuance of an A to restaurants with major health threats, makes it impossible for a restaurant to get a score of 90 or above when it has two or more major violations or when its permit is suspended. However, the percentage of 90 and the percentage of A in the score distribution remain almost unchanged after the policy change, as shown in Appendix A.1, pointing to its failure. This suggests that, to increase the informativeness of the grade, what is needed is not a mechanical change to the grading policy, but a deeper understanding of the inspector motives behind grade inflation, which is one of the main focuses of this paper.

This study has four major contributions. First, it contributes to the broad literature on regulation enforcement, in particular, the ways in which regulation-enforcement agents can compromise or strengthen enforcement efforts (Bennett et al., 2013; Agarwal et al., 2014; Short et al., 2016; Ibanez and Toffel, 2020; Kovács et al., 2020; Makofske, 2020b; Heitz et al., 2021; Kalmenovitz, 2021). This paper is closely related to Kalmenovitz (2021), which finds that when enforcement effort improves promotion chances, promotion incentives drive financial regulators to file more enforcement actions. In contrast, this paper finds that career incentives motivate restaurant inspectors in LA County to cite fewer food safety violations. The opposite effects of career incentives on enforcement stringency implemented by public workers found in the two papers emphasize the importance of properly aligning regulation-enforcement agents' incentives in their performance evaluation scheme. This study is also

[5] For more examples of how this policy change affects scores, refer to http://www.publichealth.lacounty.gov/eh/docs/posts/FAQ-point-deduction-modifications-food-official-inspection-report-en.pdf.

[6] Referring to the grading policy change, Terri Williams, acting director of LADPH's environmental health division, said that it is important for the credibility of the program, and that they want the public to know when they go into a restaurant with an A in the window that the restaurant truly earned that A (Baer, 2016b).

closely related to Makofske (2020b) and Kovács et al. (2020), which document that inspectors exploit point-deduction protocols to produce a more favorable inspection results (i.e., grade inflation) for the restaurants in LA County. This paper extends this specific literature on inspector leniency in two ways. First, it documents two public health consequences of inspector leniency: 1) reducing the effectiveness of health grades in signaling restaurants' hygiene conditions, and 2) discouraging improvements in restaurant hygiene by preventing a re-inspection that a restaurant would likely request if they did not experience grade inflation. Second, motives behind grade inflation are investigated, and it is shown that grade inflation caused by different motives calls for drastically different remedies. To the best of my knowledge, it is also the first paper to propose and provide evidence for owner-initiated re-inspection avoidance as a motive behind inspector leniency in the restaurant inspection setting.

Second, the paper contributes to the literature on the effects of enforcement design on compliance efforts, including enforcement technology (Jin and Lee, 2014; Agarwal et al., 2023), disclosure policy (Jin and Leslie, 2003; Simon et al., 2005; Evans, 2016; Ho et al., 2019; Jin and Leslie, 2019; Makofske, 2020a), perceived probability of being inspected (Makofske, 2019; Makofske, 2021), and the allocation, training, and the enforcement style of the enforcement personnel (May and Wood, 2003; Jin and Lee, 2018). This study is mostly closely related to Jin and Leslie (2003). They find that the mandatory grade card disclosure policy in LA County prompts restaurants to make hygiene quality improvements, though it also gives rise to inspector leniency. By constructing a model where a restaurant chooses sanitation effort to maximize expected profit, this paper shows that a restaurant reduces its sanitation effort when expecting a lenient inspector. The paper also documents that inspector leniency in the form of grade inflation prevents a re-inspection that motivates hygiene effort, which a restaurant would likely request if they did not experience grade inflation. Both findings provide insight on how inspector leniency, an unintended by-product of the disclosure policy, can decrease the effectiveness of disclosure in incentivizing restaurant sanitation efforts.

Third, the research question studied in this paper has implications in other settings where the evaluation of a subject is purposefully inflated at certain cutoffs. It can be applied to restaurant health inspections in other regions where inspectors assign scores while only grades are posted publicly. As far as I know, such regions include but are not limited to San Diego County, Riverside County, San Bernardino County, Boulder County and New York City. It can also be applied in educational contexts, where Diamond and Persson (2016) and Dee et al. (2019) observe scores below grade cutoffs being bumped up by teachers. In the context of home purchase appraisals, appraisal inflation is more common at loan-to-value notches (e.g., 80%), resulting in an unusually high percentage of appraised value equaling the contract price (Agarwal et al., 2020; Conklin et al., 2020; Calem et al., 2021).

Lastly, the paper contributes to the literature on consumption amenities. Starting with Glaeser et al. (2001), the urban literature has been paying more attention to the role of cities as centers of consumption. Restaurants, a critical component of local consumption amenities (Glaeser et al., 2001; Couture and Handbury, 2020),[7] have been a subject of interest in various urban studies.[8] Restaurant inspections and mandatory grade card posting are part of local authorities' efforts to maintain and disclose the quality of local consumption amenities. This paper provides a theoretical argument that such efforts can backfire due to inspector leniency, resulting in a worsening of restaurant quality, in this case, a decrease in restaurants' sanitation efforts. It also documents how grade inflation performed by the inspectors has reduced the health grade to a distorted signal of a restaurant's quality.

[7]Glaeser et al. (2001) show that cities with more restaurants and live performance theaters per capita have grown more quickly than their peers in the last two decades of the 20th century, both in the US and in France. Couture and Handbury (2020) argue that consumption amenities like restaurants and nightlife play a more important role than other commonly-cited factors (eg. job oppurtunites, crime rates and public transit) in explaining the rising number of young college graduates who reside near city centers since 2000.

[8]Kuang (2017) measures the market value of restaurant consumption. Han et al. (2020) examine the neighborhood effects of fast food restaurants on childhood obesity. Davis et al. (2019), Waldfogel (2008), and Schiff (2015) study how geographic and demographic factors affect restaurant consumption and restaurant varieties. Shoag and Veuger (2019) study the effect of land use restrictions on the quality and diversity of restaurants. Kim and Yörük (2015) estimate the impact of restaurant smoking bans on household dining out expenditures.

The remainder of the paper is organized as follows. Section 1.2 introduces the LA County restaurant inspection program, and Section 1.3 describes the data. Section 1.4 discusses the literature that proposes inspector leniency as the dominant contributor to grade inflation. Section 1.5 explores the negative public health consequences of grade inflation. Section 1.6 examines the motives behind inspector leniency and proposes targeted remedies. Section 1.7 develops a model to study the effects of inspector leniency on a restaurant's assigned score and its sanitation effort. Section 1.8 evaluates the proposed remedies using the model. Section 1.9 concludes.

1.2 Background

This section describes different types of inspections under the LA County restaurant inspection program, and provides knowledge on how inspectors are assigned to restaurants and how they are compensated and evaluated. The relevant information is collected from LADPH's websites and by interviewing inspectors and one environmental health department manager.

1.2.1 Types of inspections

There are four types of restaurant inspections: routine inspection, owner-initiated inspection (OII), re-inspection/follow-up inspection and complaint investigation.

1.2.1.1 Routine inspection

The majority of inspections are routine inspections. Each restaurant is subject to 1 to 3 routine inspections every year. The frequency is based on the public health risk associated with the food products served, the methods of food preparation, and the operational history of the food facility. Routine inspections are unannounced. An inspector fills in the Retail Food Official Inspection Report (FOIR) when conducting an inspection. The FOIR lists Critical Risk Factors and Good Retail Practices, and the points to be deducted if violations are found. Critical Risk Factors are marked if high-risk violations are observed.[9] Among these high-risk violations, some are major, with a four-point deduction, some are minor, with a two-point deduction. A major violation poses an imminent health hazard that warrants immediate correction and may require closure of the food facility. A minor violation does not pose an imminent health hazard, but does warrant correction. Some violations are discretionary, as the inspector can decide whether to give out a two-point deduction or a four-point deduction based on the observed severity of the violation. Good Retail Practices are preventive measures that can reduce food borne illness.[10] A violation of good retail practices leads to a one-point deduction. Each food inspection begins with 100 points, and the remaining points after deduction would be the final score.[11]

Both the inspector and the manager/owner of the restaurants are required to sign the inspection report. A grade or score card is issued to each facility at the end of all routine inspections, which must be posted in an area clearly visible to patrons. Restaurants that score at least 70 points will receive a grade card, with 90 to 100 points corresponding to grade A, 80 to 89 points corresponding to grade B, and 70 to 79 points corresponding to grade C. Restaurants that score less than 70 will receive a score card displaying the actual score. Starting from December 2013, LA County also provides restaurant inspection information to Yelp (Brown, 2013).[12]

[9]Violations of Critical Risk Factors include improper cooking time and temperature, no water available, inappropriate disposal of sewage and wastewater, presence of insects, rodents, birds and animals, etc.

[10]Violations of Good Retail Practices include unapproved thawing methods being used, inadequate ventilation, food being stored in unapproved areas, etc.

[11]For more details, refer to the *Reference Guide for the Food Official Inspection Report* compiled by the Department of Public Health at http://publichealth.lacounty.gov/eh/docs/permit/reference-guide-food-official-inspection-report.pdf. It enumerates in great detail the scenarios that are considered violations, with a sample inspection report attached.

[12]On the web version of Yelp on Aug 29th, 2022, the latest inspection grade is displayed under the "Amenities and More—Health Score" section. If a viewer clicks into "Health Score," he/she would be able to see the grades and violation entries for the inspections in the past two and a half years. On the mobile

1.2.1.2 OII

If restaurant owners are dissatisfied with the grade earned in the routine inspection, they can file for an OII within 3 business days. The OII fee needs to be paid within 3 business days after they file for an OII. The OII fee is $330 for a low-risk restaurant, $391 for a moderate-risk restaurant and $440 for a high-risk restaurant in 2019-2020 (LADPH, 2019), and is perceived as affordable to the restaurants by an informant from LADPH. An OII is unannounced, and is performed within 10 calendar days after a restaurant pays the OII fee. The OII inspector is required to be the same inspector as in the previous routine inspection to ensure grading criteria consistency. An OII is a full inspection as a routine inspection, and the grade of the OII will replace the grade of the previous routine inspection. An OII may be requested once in a twelve-month period.

1.2.1.3 Re-inspection and compliant investigation

Both routine inspections and OIIs are full inspections of the food facility, while re-inspections and compliant investigations are brief visits to the food facility and will not yield an inspection score. Re-inspections are scheduled within 14 days to ensure ongoing compliance if the previous routine inspection involves Major Critical Risk Factors, Repeat Minor Critical Risk Factors or Repeat Good Retail Practices. A complaint investigation is required for every diner complaint filed through LADPH before COVID-19 (Kovács et al., 2020). Complaint investigations examine whether a restaurant is subject to the health code violation(s) for which a diner reports the restaurant to LADPH.

---

app version of Yelp on Aug 29th, 2022, the latest inspection grade along with the number of violations is displayed under the "Info" section. If a viewer clicks for more details, he/she would be able to see the violation entries for the most recent inspection, and the grades for the inspections in the past two and a half years. In both cases, the viewers do not know the scores of the inspections, as they are not displayed, and there is no point deduction attached to the violation entries.

1.2.2 Inspector assignment, compensation and evaluation

Restaurant inspections in LA County (except for Pasadena, Long Beach and Vernon) are overseen by the District Surveillance and Enforcement (DSE) branch in LADPH's Environmental Health Division.[13] DSE has 29 offices throughout the county. Each office has eight inspection districts (henceforth districts) and eight inspectors.[14] Generally, a restaurant inspector works at the same office throughout his/her career.[15] A restaurant inspector is assigned to be in charge of one of the eight districts in an office.[16] Inspectors are rotated across districts within an office every two years.

The official job title of a restaurant inspector is level II Environmental Health Specialist (EHS). A restaurant inspector's salary depends on his/her level and step (how long an EHS has worked for LADPH), and does not depend on the number of inspections completed. Restaurant inspectors typically work 40 hours per week.[17] They do receive overtime compensation, but it is not common for them to work overtime. The overtime inspections are usually for the facilities holding the weekend communal events (e.g., amusement parks, stadiums and venues) and need to be approved by the manager.

[13]Each of Pasadena, Long Beach and Vernon has its own city health department that inspects its retail food facilities.

[14]The Lancaster office and the Santa Clarita office are two exceptions. These two offices share the same chief Environmental Health Specialist, and together have eight districts and eight inspectors. There are 224 inspection districts in LA County.

[15]The only way for a restaurant inspector to work at a new office is to either file for a lateral transfer or get promoted. When inspectors are given their first assignment to an office after they complete training, it is possible that they are assigned to offices far away from where they live. Lateral transfer is typically sought by these inspectors to transfer to an office closer to home. For this reason, inspectors usually file for a lateral transfer once or twice during their careers.

[16]An inspector primarily works in the district they are assigned to. However, they may inspect restaurants outside their district or even their office in the following two scenarios. First, an inspector may be asked by their supervisor to inspect restaurants in another district within the same office when the inspector in charge of that district fails to complete the targeted number of routine inspections at the end of each month. Second, there is substantial heterogeneity among the number of restaurants across offices. An inspector from an office with low restaurant inventory may need to take over some of the the inspection duties from another office with high restaurant inventory.

[17]Available work schedules include five days - 40 hours per week, four days - 40 hours per week, and nine days - 80 hours per two weeks.

Restaurant inspectors are required to prioritize food establishments with closer due dates for routine inspections and those that have filed an OII. Apart from that, they have discretion over which establishments to inspect on a work day. Each restaurant inspector has a monthly goal of the number of *routine* inspections he/she needs to complete. Most inspectors are on top of their monthly goals, but if some inspectors fail to meet their goals, the other inspectors in the same office would be appointed by the chief EHS to help with the workload. An inspector's supervisor would step in if he/she has low productivity consistently. The supervisors audit each inspector's report at least once a year by repeating the inspection of a restaurant within one or two days of its last routine inspection.

Promotions to an EHS III are made when there is a need to fill vacancies and the budget allows the positions to be filled.[18] Restaurant inspectors are evaluated based on how they place in the civil service exam and their annual Performance Evaluation (PE) from the last two years.[19] The PE is based on the number of *routine* inspections an inspector has completed.

1.3 Data

1.3.1 Routine inspection and OII

The main dataset used in this paper is the LA County restaurant and market inspections dataset publicly available on the LA County Open Data Portal. This quarterly updated

[18] Job duties for an EHS III include inspector training, program planning, industry engagement, food manufacturer inspection, solid waste inspection, and etc. For more details, refer to http://publichealth.lacounty.gov/eh/about/careers.htm.

[19] The online application for EHS III serves as the civil service exam. The civil service exam is graded by the HR based on the applicant's previous relevant experience and desirable qualifications (e.g., knowing languages other than English). Applicants who achieve a passing score will be added to the hiring list that each public health program can use to fill vacancies. Candidates on the hiring list are grouped into five bands based on their score, and programs move from the top band to the bottom to select candidates. To select from candidates in the same band, the programs rely on the annual PE from the last two years.

dataset presents inspection results for routine inspections and OIIs for restaurants, food markets and caterers in LA County. This paper uses the inspection results for restaurants from June 1st, 2017 to December 31st, 2019.[20] The dataset covers 172,474 inspections of 3,5794 restaurants. Multiple inspections that occur on the same day at a given restaurant are dropped.[21] The remaining sample is referred to as the full sample/main inspection dataset throughout the paper. The full sample covers 172,288 inspections of 35,792 restaurants. Key variables include the date and type (routine inspection/OII) of the inspection, the ID number of the employee who conducted the inspection, the name, address and type (capacity and risk) of the facility being inspected, and the score and grade of the inspection. Information on health codes violated in each inspection and point deductions associated with each violation is publicly available on the LA County Open Data Portal as well, and is linked to the main inspection dataset. In the full sample, routine inspections account for 98.19% of all the inspections, while OIIs account for 1.81%. The dataset does not report the exact capacity of a restaurant. It instead reports four levels of capacity: 0-30 seats, 31-60 seats, 61-150 seats and 151+ seats. Values 1,2,3,4 are assigned to the above categories respectively to get the variable *capacity*. There are three risk levels in the dataset: low, moderate and high.[22] Values 1,2,3 are assigned to the above categories respectively to get the variable *risk*.

[20] There are three reasons why data for this time range are used. First, there was a change in the grading policy that was effective in July 2016 with a six-month grace period until full implementation. The sample period used in this paper starts six months after the full implementation of the policy change, by which time the inspectors or the restaurants are likely to have finished adjusting to the policy change. Therefore, there are unlikely any behavioral changes from the inspectors or the restaurants related to this policy change during the sample period. Second, to ensure no behavioral changes from the inspectors or the restaurants responding to COVID-19, the sample period ends before 2020. Third, this is the set of the data that was available when the project started in January 2020.

[21] There are cases where one restaurant receives two inspections on the same day. Logistically speaking, a routine inspection and an OII are unlikely to happen on the same day, and two routine inspections cannot happen on the same day. Moreover, it is unknown which inspection or inspections are valid. Therefore, such observations are dropped. This results in 186 observations (93 restaurant-days) being dropped from the sample.

[22] Low-risk restaurants handle foods which are generally pre-packaged, ready to eat, or pre-cooked and require heating prior to service. Moderate-risk and high-risk restaurants handle foods such as meat, poultry, seafood, sushi or oysters which are served raw, or require processing, cooling and reheating (Environmental Health, 2011). Examples for low-risk restaurants include coffee shops, juice bars, ice-cream shops and theatre snack bars. Moderate-risk restaurants primarily sell pizzas, pastries, donuts, sandwiches and burgers. High-risk restaurants sell food that requires more processing.

1.3.2 Complaint investigation

The inspection dataset publicly available on the LA County Open Data Portal does not include complaint investigations. Therefore, a secondary inspection dataset is constructed using inspection records from the Environmental Health Inspection Results Page (http://www.publichealth.lacounty.gov/eh/i-want-to/view-inspection-results.htm) on the LADPH website, which contains information on complaint investigations. Each inspection record reports the inspection type (routine inspection/complaint investigation/re-inspection/OII), score, grade, and the number of violations of the inspection, and the name, address and city of the facility. The website holds results of all the environmental health inspections, including housing inspections, food facility inspections, swimming pool inspections, and etc. Ideally, only restaurant inspections are relevant, but the finest level the search tool allows the data to be filtered is to include both restaurants and food markets. In addition, whether a facility is a restaurant or a food market is not labelled in the inspection result. Therefore, data on both restaurant inspections and food market inspections are collected.

The website holds inspection results over the past 5 years for currently active facilities, and is constantly being updated to include the latest inspection results. Inspection records were collected in September 2021, and the sample includes information on inspections performed between September 6th, 2016 and September 20th, 2021. This paper also drops food facilities that have more than one inspection record of the same inspection type on the same day.[23] The resulting sample is henceforth referred to as the secondary inspection dataset.

[23] This outcome could result from entry errors, or it could result from different restaurants sharing the same name and address. They are usually separate food-service establishments housed by the same facility (e.g., hotel, grocery store, food court), and are sometimes inspected on the same day (Makofske, 2019). For inspections at these restaurants, this paper is unable to determine which inspection belongs to which restaurant.

1.4 Inspector leniency

Figure 1.1 shows the score distribution for the routine inspections in the full sample. 13.23% of the routine inspections receive score 90, the lowest score to be qualified for grade A, causing a spike at 90 in the score distribution. Another feature to notice is the high proportion of grade A. Among 169,174 routine inspections, 94.28% of them yield grade A, and 5.09% of them yield grade B.

Existing literature attributes the bunching at 90 to inspector leniency. It documents ways in which inspectors utilize point deduction rules to inflate a restaurant's score to 90 or above when it is just below grade A. It also provides arguments against restaurant optimization as the main cause of the bunching at 90.

Using LA County restaurant inspections from 2014 to 2016, Makofske (2020b) finds that restaurants on the margins of higher letter grades are 28–40% more likely to receive the lesser deduction on discretionary violations,[24] when compared with restaurants exhibiting better hygiene quality. This indicates that inspectors inflate a restaurant's grade by giving out a 2-point deduction (corresponding to a minor violation) instead of a 4-point deduction (corresponding to a major violation) when the larger deduction would have led to a B for a restaurant, and the smaller deduction would have led to an A. This effect is significant across all discretionary violations, including the violation of "hands clean and properly washed; proper glove use." Makofske (2020b) argues that a violation of this health code is evidence against restaurant optimization, as this health code bears little cost to observe since a restaurant could regularly violate this code and still get no deduction during an inspection by simply having its employees behave appropriately in front of the inspector.

Using LA County restaurant inspections from 2000 to 2010, Kovács et al. (2020) argue

[24] In terms of a discretionary violation, an inspector can decide whether to give out a two-point deduction or a four-point deduction based on the observed severity of the violation. More details can be found in Section 1.2.1.1.

that the bunching at 90 is in part due to inspectors coding violations as "Not Observed" (compliance with a health code cannot be visually observed)[25] to avoid point deductions for restaurants they have inspected repeatedly. They also point out that a restaurant staff would need to make simultaneous adjustments in terms of over 70 hygiene rules in order to score just 90, which makes restaurant optimization an implausible explanation for the bunching at 90.

The above literature provides important insights into how inspectors take advantage of the point deduction rules to inflate a restaurant's score to 90 or above when it is just below grade A, and why inspector leniency is a more plausible explanation for the bunching at 90 than restaurant optimization. It also partly explains why the grading policy change implemented by LADPH back in 2016 failed to achieve its intended effects.[26] One of the changes involves mechanical modifications to the point deduction rules such that a restaurant with two or more major Critical Risk Violations (each worth four-point deduction) loses three additional points, ending up with an 89 or below. Such effort to prevent an issuance of A to a problematic restaurant only works when an inspector gives a restaurant two or more major violations. However, the above literature has demonstrated that inspectors are good at taking advantage of the grading protocols to omit violations and mark down the severity of violations. For example, the additional three-point deduction would fail to apply if the inspector changes one of the two major violations a restaurant deserves to a minor violation. A natural question to ask is whether grade inflation incurs social costs, which the paper attempts to answer in the next section.

[25]For example, if an inspection occurs during the lunch hour and the restaurant only cooks and serves hot dishes at dinner, then health code regarding cooking temperatures would be marked as "Not Observed"(Kovács et al., 2020).

[26]Another explanation is simply that routine inspections where two or more major Critical Risk Violations are found, or where a restaurant's permit is suspended due to no water available, improper sewage disposal, or vermin infestation, are rare. Out of 27,757 routine inspections from October 1st, 2015 to June 30th, 2016 (prior to the grading policy change) that have available violation records, the grading policy change would have led to a decrease in score for only 1,092 inspections, and a downgrade for only 824 inspections.

1.5 Exploring the public health consequences of grade inflation

In this section, the paper will discuss two ways in which grade inflation can negatively impact public health: by undermining the effectiveness of health grades in signaling hygiene conditions to diners, and by discouraging restaurants from making hygiene improvements.

1.5.1 Grade inflation misleads diners about restaurants' hygiene conditions

When some restaurants' grades are inflated to A, the A-grade category includes restaurants that have worse hygiene conditions than required to be qualified for A. However, it is unclear whether there are meaningful differences in hygiene conditions between restaurants whose grades are inflated to A and restaurants who have earned their A, such that the informativeness of letter grade on the underlying hygiene of a restaurant is compromised. In the following exercise, the paper compares the average hygiene condition of restaurants in score categories 90 and 91 with the average hygiene condition of restaurants in each of the other score categories above 90. Score categories 90 and 91 are where most restaurants whose grades are inflated to A would end up in, because a utility-maximizing inspector would opt for the methods that allow them to inflate the score by the minimum number of units required for an A.[27] For example, if a restaurant's actual score is 89, it would make sense for an inspector to omit a one-point deduction to inflate the score to 90 or revise a 4-point deduction into a 2-point deduction to inflate the score to 91 if a discretionary violation is present, but it does not make sense for an inspector to inflate the score to be higher than

[27] The benefit of inflating the score to 90 or above is that a restaurant would qualify for an A, which is the same no matter what the inflated score is, as long as it is 90 or above. However, the cost of score inflation increases with the number of units inflated, because 1) it takes effort for inspectors to come up with ways to inflate the score, and 2) a larger score inflation is more likely to be detected by the supervisors.

that. In addition, score category 90 seems to have a larger percentage of restaurants whose grades are inflated to A than score category 91, judging from the spike at 90 in Figure 1.1.

A restaurant's hygiene condition is not directly observed. A natural measure of the hygiene conditions of a restaurant is whether it has caused food-related illnesses. Jin and Leslie (2003) use hospital admission data for digestive disorders from the Office of Statewide Health Planning and Development to measure the incidence of food-related illnesses. Unfortunately, the data are at three-digit zip code level and cannot be linked to individual restaurants. However, complaint investigations and their outcomes offer a possible way to track individual restaurant's hygiene. A complaint investigation is required for every diner complaint filed through LADPH before COVID-19 (Kovács et al., 2020), during which an inspector examines whether a restaurant is subject to the health code violation(s) for which it is reported. This paper uses the probability of a subsequent complaint investigation where violations are found as a measure of the hygiene conditions of a restaurant.

The relationship between the score of a routine inspection and the likelihood of a subsequent complaint investigation where violations are found is estimated using a linear probability model (LPM).[28] The following analysis uses the secondary inspection dataset, and extracts routine inspections from June 1st, 2017 to December 31st, 2019 for the same reason mentioned in Section 1.3.1.[29] The sample includes 155,647 routine inspections. The dependent variable equals 1 if a routine inspection is followed by a complaint investigation where violations are found and 0 otherwise. It is regressed on score dummies while controlling for year fixed effects, month-of-year fixed effects, day-of-week fixed effects and city fixed effects.

[28]Estimates under the logit model are quantitatively similar, and are available upon request. LPM is preferred because 1) it is easier to interpret, and 2) some score and city groups have no restaurants that have experienced complaint investigations where violations are found and these observations are dropped under the logit model.

[29]Besides reasons mentioned in Section 1.3.1, another reason for restricting the sample to include only routine inspections till December 31st, 2019 is so that whether a routine inspection is followed by a complaint investigation is fully observed. Recall from Section 1.3.2 that the original sample includes inspections performed between September 6th, 2016 and September 20th, 2021. Therefore, for each routine inspection, one can follow up for at least nine months to observe whether there is a subsequent complaint investigation.

As discussed above, most of the restaurants whose grades are inflated to A should end up in the score category 90 or 91. Therefore, score 92 is the lowest score under which the restaurants are relatively clean of score inflation. For this reason, the score dummy for score 92 is omitted as the base group.

Figure 1.2 plots the coefficients and the 95% confidence intervals of the score dummies for 90 to 100 from the LPM.[30] Score 92 is the base group and its dummy has a coefficient of 0. Overall, the probability of a subsequent complaint investigation where violations are found decreases with score, consistent with the expectation that restaurants with a higher score on average have better hygiene conditions. Restaurants scoring below 96 in routine inspections have comparable hygiene conditions on average, except for restaurants that score 90, which have a 0.35% higher probability of experiencing a subsequent complaint investigation where violations are found than restaurants that score 92, and this difference is significant at a 5% level. The average probability of experiencing a subsequent complaint investigation where violations are found is 2.05% in the sample. Therefore, the difference in the probability of experiencing a subsequent complaint investigation where violations are found between restaurants scoring 90 and restaurants scoring 92 is also economically considerable. Since the score group 90 is most susceptible to grade inflation, this is suggestive evidence that restaurants whose scores are inflated to 90 have significantly worse hygiene conditions than restaurants who achieve an A without score inflation. Therefore, for a restaurant that is subject to grade inflation, its A misleads the diners to believe that it has better hygiene conditions than it actually does. For diners who rely on grades for information on a restaurant's hygiene, they are unable to distinguish poor-hygiene restaurants whose grades are inflated

[30] One may be interested to look at the coefficients of dummies for scores just under 90. However, differences in the probability of a subsequent complaint investigation with violations between a B-restaurant and an A-restaurant can be attributed to factors other than hygiene conditions. For example, since diners are less likely to patronize a B-restaurant, a B-restaurant may even have a lower probability of receiving a diner complaint than an A-restaurant just because it has less foot traffic. Moreover, it is reasonable to assume that diners that patronize a B-restaurant care less about a restaurant's hygiene conditions than diners that patronize an A-restaurant, which can also result in a B-restaurant having a lower probability of receiving a diner complaint.

to A from good-hygiene restaurants who have earned their A.

Moreover, the negative impacts of grade inflation can spread beyond the realm of public health. Since the publicly available health grade is the most accessible and most straightforward measure of restaurant hygiene conditions for diners, it is safe to say that a diner's perceived hygiene condition of a restaurant comes from its health grade. Kuang (2017) finds that the perceived quality of restaurant amenities, measured by Yelp ratings, is capitalized into local real estate values, and that such capitalization is more pronounced after Yelp.com becomes more popular in the area. There is reason to believe that another dimension of perceived restaurant quality, the perceived hygiene condition, may be capitalized into local real estate values as well. Therefore, the information distortion embedded in the assigned grade can ripple through the housing markets and distort local real estate values.

The above analysis uses routine inspections for both restaurants and food markets. To restrict the analysis to restaurant inspections only, a modified approach keeps only the routine inspections that have a match record in the main restaurant inspection dataset. This also allows the paper to further control for inspector fixed effects and restaurant type fixed effects, as the main restaurant inspection dataset contains information on the inspector and the type (ie., risk level and capacity level) of the restaurant. The results under this robustness check are qualitatively similar to the baseline results, and are presented in Appendix A.2. Appendix A.2 also discusses a potential sample selection issue as the secondary inspection dataset used in the analyses includes only restaurants that are still open in September 2021 when the data were collected.

1.5.2 Grade inflation discourages hygiene improvements

Apart from undermining the effectiveness of health grades in signaling hygiene conditions, grade inflation can also discourage hygiene improvements. Using Las Vegas restaurant health inspections, Makofske (2020a) finds that among restaurants with comparable hygiene conditions, those that receive a B and undergo a re-inspection as a result perform better in the next routine inspection than those that receive an A. The same reasoning can be applied to the context of LA County restaurant inspections. I hypothesize that when the grade of a restaurant that should have received a B is bumped up to an A, the restaurant has no incentive to apply for an OII, and therefore misses the opportunity to improve its hygiene conditions by preparing for and going through an OII.

To verify the above hypothesis, the probability of a subsequent complaint investigation where violations are found is compared between restaurants whose grades are inflated to A in the routine inspection (treatment group), and restaurants that receive a B without grade inflation in the routine inspection and update their grades to A[31] through an OII (control group). Limiting to these two groups, the effect of grade inflation on restaurant hygiene is estimated by the following model:

$$complaint_i = \beta_0 + \beta_1 inflated_i + X_i'\gamma + \varepsilon_i, \tag{1.1}$$

where $complaint_i$, a proxy for restaurant hygiene, equals 1 if routine inspection i is followed by a complaint investigation where violations are found and 0 otherwise.[32] $inflated_i$ equals 1 if routine inspection i experiences grade inflation, and equals 0 otherwise. This paper only considers grade inflation that occurs via weak penalties for discretionary violations, as documented in Makofske (2020b): inspectors marking a 2-point deduction instead of a 4-point deduction for discretionary violations.[33] In this case, it is only possible for an inspection

[31]The control group is restricted to restaurants that later update their grade to an A in the OII to ensure that the treatment group and the control group have the same grade (i.e., A) when diners file complaints against them, as differences in probabilities of a subsequent complaint investigation between restaurants with different grades can be due to factors other than hygiene conditions, such as differing tendencies to complain between A-restaurant patrons and B-restaurant patrons (refer to footnote 30 for more details). The restriction of receiving an A in the OII is by no means stringent, as 92.6% to 100% of OIIs lead to an A if the last routine inspection gets a B, as shown by Figure 1.3b.

[32]For the control group, I only consider complaint investigations that take place post-OII.

[33]There are three ways inspectors can achieve grade inflation: First, inspectors can give lesser deduction to discretionary violations (Makofske, 2020b). Second, inspectors can claim that compliance with a health

to be affected by grade inflation if 2-point deductions on all the discretionary violations lead to a higher letter grade than 4-point deductions on all the discretionary violations. Such an inspection is henceforth referred to as an inspection *on the margin.* Appendix A.3.1 estimates the predicted grades in the absence of manipulation for routine inspections on the margin, following Makofske (2020b). A routine inspection on the margin is considered to have an inflated grade if its assigned grade exceeds its predicted grade.

X_i in (1.1) is a vector consisting of inspection-specific controls, including indicators for the month, year, day of the week when the inspection occurs, indicators for the city and type (ie., risk level and capacity level) of the restaurant where the inspection occurs, and indicators for the inspector who conducts the inspection. The analysis uses routine inspections for restaurants which have not previously received complaint investigations.[34]

The parameter of interest is β_1. If grade inflation discourages hygiene improvements, then restaurants that have experienced grade inflation should be more likely to have violations cited in a subsequent complaint investigation than restaurants that have not experienced grade inflation, suggesting a positive β_1. To ensure an accurate estimate of β_1, the treatment and control groups must have comparable hygiene conditions at the time of the routine inspection. As a result, treated units are restricted to routine inspections with a score between 90 and 91, and control units are restricted to routine inspections with a score between 85 and to 89.[35] The estimated β_1 is presented in column (1) of Table 1.1. On average,

code cannot be visually observed by marking a health code as "Not Observed" in the inspection report (Kovács et al., 2020). Third, inspectors can simply choose to not record a violation. It is only possible to identify restaurants that may be affected by the first two grade inflation strategies, through data on discretionary violations and on health codes marked as "Not Observed." Since the inspection dataset does not include information on which health codes have been marked as "Not Observed," only grade inflation achieved through a lesser deduction on discretionary violations is studied in this paper.

[34]A history of complaint investigations suggests that a restaurant has little motivation to exert sanitation effort. Such restaurant is likely to revert to its old ways (i.e., having bad hygiene conditions) even after going through an OII. This means that the hygiene conditions of such restaurant are unlikely to respond to whether it receives treatment (i.e., grade inflation) or not. Therefore, this paper focuses on the local average treatment effects on restaurants that have not previously received complaint investigations, which are more likely to respond to the treatment. Around 77% of the treated inspections, and around 68% of the control inspections belong to restaurants that have not previously received complaint investigations.

[35]Among routine inspections with a score between 90 and 91 and a predicted grade of B, 85.38% of them

restaurants whose grades are inflated to A have a 1.1% higher probability of a subsequent complaint investigation where violations are found than restaurants that receive a B and update their grades to A through an OII. The effect is not precisely estimated (statistically significant at a 10% level), as there is only a small number of control units available (267 routine inspections), due to the infrequent issuance of a B grade. Nevertheless, this finding suggests that grade inflation discourages hygiene improvements.

One concern is that $inflated$ is subject to measurement error, as $inflated$ is unobserved and has to be estimated. Two robustness checks are performed where the treatment group is expanded to include restaurants with *better* hygiene conditions than the control group. Assuming that restaurants with better hygiene conditions have a lower probability of a subsequent diner complaint with violations, the estimated β_1 will be subject to a downward bias, and is therefore a more conservative estimate. The details on the construction of two alternative treatment groups are presented in Appendix A.3.2. The results from the robustness checks are presented in column (2) and column (3) of Table 1.1, and are qualitatively and quantitatively similar to the baseline results in column (1).

Unfortunately, to the best of my knowledge, there is no literature on the relationship between diner complaint and foodborne illness. Therefore, the above estimate on the effect of grade inflation on a subsequent complaint investigation where violations are found cannot be converted to an estimate on the economic cost of foodborne illness caused by grade inflation. However, it could be a possible direction for future research.

1.6 Motives behind grade inflation

Section 1.5 discusses the negative health consequences incurred by grade inflation. To alleviate grade inflation, motives behind inspector leniency need to be studied as inspector

have a predicted score between 85 and 89.

leniency seems to be the dominant cause for grade inflation.

Though this paper is not able to enumerate all the possible motives behind inspector leniency,[36] or pinpoint how much each possible motive contributes to the bunching at 90, the following two subsections will discuss two motives that are verified, either through empirical analyses, institutional knowledge or the existing literature, and will demonstrate that inspector leniency caused by different motives calls for drastically different remedies.

1.6.1 OII avoidance

Figure 1.3a shows the percentage of routine inspections that lead to OIIs in different score categories in the full sample. Few routine inspections lead to an OII when their scores are at or above 90, while the probability of a subsequent OII increases sharply as the routine inspection score drops below 90. When routine inspections yield a B or C, 20.8% to 54.1% of them lead to an OII. This is a considerable proportion, especially considering that restaurants are eligible to apply for an OII only once a year.[37] This pattern suggests that restaurants have a strong tendency to exploit OII as a way to improve grades. The restaurants that file for an OII do successfully end up with an A most of the time, as Figure 1.3b shows. 92.6% to 100% of OIIs lead to an A if the last routine inspection gets a B, and 74.0% to 90.7% of OIIs lead to an A if the last routine inspection gets a C. The high success rate of OII in getting an A combined with the fact that the OII fee is affordable makes a restaurant very likely to file for an OII when it fails to get an A.

Given a restaurant's inclination to file for an OII when it gets a B in a routine inspection, an inspector has an incentive to bump up a restaurant's score to 90 if its score is below but close to 90 for the following two reasons. First, inspectors have no motivation to conduct

[36] See Makofske (2020b) for a discussion of possible motives behind inspector leniency.

[37] Whether a routine inspection leads to an OII cannot be observed if it is the last inspection of a restaurant in the sample. Therefore, this proportion is underestimated.

OIIs. An OII takes a similar amount of time to perform as a routine inspection, but is not counted into the annual PE (Performance Evaluation) or the fulfillment of the monthly goal as a routine inspection. Given that an inspector has a fixed number of working hours per month,[38] the more OIIs an inspector conducts, the less time there is for the inspector to carry out routine inspections, making an inspector less likely to meet monthly goals and get promoted.[39] Since the inspector of the last routine inspection is required to conduct the OII if the restaurant files for one,[40] the inspector is reluctant to give out a B, which would substantially increase the chance of an undesirable OII. Second, score bump-up of this magnitude is hard to be detected by a supervisor. Supervisors audit each inspector's report at least once a year by repeating the inspection of a restaurant within one or two days of its last routine inspection. The audit frequency is low to begin with, so it is unlikely that the restaurant the supervisor audits happens to be among the ones for which the inspector has bumped up the score. If the supervisor finds out that the score given by the inspector is a few points above what its actual hygiene condition qualifies for, an inspector can simply attribute this difference to the restaurant being less vigilant about cleaning since it has just received a routine inspection and will not expect another routine inspection any time soon.[41]

[38] Restaurant inspectors typically work 40 hours per week, and rarely work overtime. For more details, refer to Section 1.2.2.

[39] A second-order motive behind OII avoidance is that OII duties increase the chance of an inspector working beyond his/her typical end-of-shift time. To achieve the monthly goal of routine inspections, an inspector with requested OII inspections would have to complete more inspections in total (routine inspections plus OIIs) than he would have without OII duties. Given the fixed schedule (40 hours per week), it means an inspector has to fit more inspections into the schedule, increasing the chance of prolonging his/her shift. Ibanez and Toffel (2020) find evidence of inspectors' aversion to shift prolonging by showing that restaurant inspectors in Lake County, IL, in Camden County, NJ and in Alaska cite fewer violations when an inspection risks prolonging their workday.

[40] OIIs that appear as a restaurant's first inspection in the sample are dropped, as there is no information on the routine inspection prior to it. OIIs that occur more than 60 days after the previous routine inspections are also dropped. This leaves the paper with 3,070 OII inspections. As many as 83.88% of them are assigned the same inspector as the previous routine inspection, which is consistent with the LADPH's arrangement that the OII inspector is required to be the same inspector as the previous routine inspection. For those scenarios where the OII inspector is not the inspector in charge of the previous routine inspection, it can be because the inspector in charge of the previous routine inspection is on sick leave or vacation leave.

[41] According to an informant from LADPH, the following two conditions would be deemed problematic during the audit: First, the restaurant owner indicates that the inspector conducts the inspection at a different time as the inspector claims in the inspection report. Second, the supervisor finds health code violations that were certainly there when the inspector conducted the inspection, e.g., vermin infestation, but the inspector did not mark the violations. Therefore, as long as the inspector is careful to not omit

In the worst-case scenario where a foodborne illness is confirmed in the restaurant to which the inspector gives an A in the most recent routine inspection, the inspector would NOT be subject to disciplinary action, as DSE recognizes that the inspection is only a snapshot of the hygiene condition at the time of the inspection and therefore the inspector should not be held accountable for a restaurant's violation of health codes at another time.

If OII avoidance is a motivation behind grade inflation, then inspectors would respond to a heavy OII workload by inflating more restaurants' grades to A in future routine inspections. The argument that grade inflation out of OII avoidance should mainly involve grade inflation to A is motivated by the pattern of restaurants' OII requests presented in Figure 1.3a: the probability of an OII request increases sharply as the routine inspection score drops below 90, while it changes continuously around 80. Moreover, restaurants are not always eligible to apply for an OII, as they can only request one OII in a twelve-month period. Since an OII-ineligible restaurant is not allowed to file for an OII, the motivation of OII avoidance should NOT play a role when an inspector decides whether to inflate an OII-ineligible restaurant's grade to A.

The above arguments can be summarized into the following three testable hypotheses:

Hypothesis 1. Inspectors are more likely to inflate a restaurant's grade to A when they have a heavier past OII workload.

Hypothesis 2. Inspectors' propensity to inflate a restaurant's grade to B is not affected by their past OII workload.

Hypothesis 3. Past OII workload only affects an inspector's propensity to inflate a restaurant's grade to A when the restaurant is eligible for an OII.

critical violations as a way to bump up scores, he/she would pass the audit. Moreover, even in the situation where an inspector gave a restaurant an A but the audit one or two days later finds out the restaurant is only qualified for a B, the supervisor would NOT adjust down the restaurant's grade, as the supervisor acknowledges that the grade is only a snapshot of a restaurant's hygiene conditions at the time of the inspection. The only exception is when a facility needs to be closed as a result of the audit, in which case a new inspection would be done, resulting in an updated grade.

One challenge to testing these hypotheses is to track an inspector's propensity for grade inflation. Fortunately, two grade-inflation strategies that are identifiable through the inspection records have been documented: When restaurants are on the margin of a higher grade, an inspector issues a 2-point deduction instead of a 4-point deduction for discretionary violations (Makofske, 2020b), or claims that compliance with a health code cannot be visually observed by marking a health code as "Not Observed" in the inspection report (Kovács et al., 2020). Since the inspection dataset used in this paper does not include information on which health codes have been marked as "Not Observed," I utilize an inspector's deduction decisions on discretionary violations to track an inspector's grade-inflation propensity.

To test Hypothesis 1, the paper estimates the following model, which builds on model (1) presented in Makofske (2020b):

$$y_{di} = \theta_0 + \theta_1 margin\_a_i + \theta_2 no\_OII_i + \theta_3 margin\_a_i \times no\_OII_i + W_d'\delta + Q_i'\lambda + v_{di}, \quad (1.2)$$

where y_{di} equals 1 if two points are deducted for a discretionary violation d found in routine inspection i, and equals 0 if four points are deducted. $margin\_a_i$ is a binary indicator of whether routine inspection i is on the margin of getting an A. Let $score\_max_i$ denote the score when 2 points are deducted for all the discretionary violations detected in inspection i. Let $score\_min_i$ denote the score when 4 points are deducted for all the discretionary violations detected in inspection i. $margin\_a_i = 1$ if $score\_max_i \geq 90$ and $score\_min_i < 90$. $margin\_a_i = 0$ if $score\_min_i \geq 90$. Model (1.2) is estimated for discretionary violations in routine inspections where $margin\_a_i = 1$ or 0. Makofske (2020b)'s original model regresses y_{di} on the margin indicator, and the estimation yields a significantly positive coefficient for the margin indicator. The finding suggests that inspectors are more likely to issue a lesser deduction on discretionary violations when a restaurant is on the margin of a higher grade, providing evidence for grade inflation practiced by inspectors. To study whether inspectors are more likely to inflate a restaurant's grade *when they have higher past OII workload*,

no_OII_i and $margin\_a_i \times no\_OII_i$ are added to the original model. no_OII_i is the number of OIIs the inspector of routine inspection i has completed 1 to 30 days prior to routine inspection i. θ_3 is expected to be positive if Hypothesis 1 is true. A routine inspection is dropped if no full inspections (OII and routine inspection) take place 1 to 30 days prior to it.

W_d contains dummy variables indicating the health code that discretionary violation d is cited for. Q_i is a vector consisting of inspection-specific controls. $no\_routine_i$, the number of routine inspections the inspector of routine inspection i has completed 1 to 30 days prior to routine inspection i, controls for inspector productivity. A productive inspector who has completed enough routine inspections to meet the monthly goal is less concerned about receiving OII requests from restaurants. ln_visit_i, the log of the number of full inspections the inspector of routine inspection i has performed in the restaurant (including routine inspection i), controls for how strong the inspector-restaurant relationship is. $ln(1 + exp_i)$, the log of one plus the number of full inspections the inspector of routine inspection i has completed prior to routine inspection i, controls for inspector experience.[42] Q_i also includes indicators for the month, year, day of the week when the inspection occurs, indicators for the zip code and type (ie., risk level and capacity level) of the restaurant where the inspection occurs, and indicators for the inspector who conducts the inspection. To keep hygiene quality comparable between inspections on the margin of A and grade-A inspections not on the margin, the paper only keeps inspections on the margin of A with $score\_min \in [86, 89]$.[43]

[42] Though the analyses in this section use observations from June 1st, 2017 to December 31st, 2019 for reasons mentioned in footnote 20, data on full inspections and associated violations are available back to Oct 1st, 2015. Since ln_visit_i and $ln(1+exp_i)$ measure inspector-restaurant relationship and inspector experience more accurately when earlier data are available, ln_visit_i and $ln(1 + exp_i)$ are computed using inspections from Oct 1st, 2015.

[43] As pointed out by Makofske (2020b), two inspections with same hygiene quality (having the same number of good retail practice violations and discretionary violations, and the same numbers of major critical violations and minor critical violations excluding discretionary violations) have the same value for $margin\_a_i$, making it impossible to allow for variations in $margin\_a_i$ across routine inspections where restaurants exhibit the same hygiene quality. Therefore, this paper follows Makofske (2020b) to control for inspection's overall hygiene-quality *category* instead. Both the inspections on the margin of A with $score\_min \in [86, 89]$ and grade-A inspections not on the margin fall into the first hygiene-quality group shown in Table 3 in Makofske (2020b).

Estimated results from model (1.2) are presented in column (1) of Table 1.2. The coefficient of $margin \times no\_OII$ is significantly positive, confirming Hypothesis 1 that inspectors are more likely to inflate a restaurant's grade to A when they have a heavier past OII workload. The coefficient of $margin$ is significantly positive, indicating that inspectors with no OII workload in the past 30 days are still more likely to issue a lesser deduction on discretionary violations when a routine inspection is on the margin of getting an A. This result suggests that there are motivations other than OII avoidance that influence an inspector's decision to inflate grades to A, consistent with expectations.

Column (2) presents estimation results from model (1.2) after replacing $margin\_a_i$ with $margin\_b_i$. $margin\_b_i$ is a binary indicator of whether inspection i is on the margin of getting a B. $margin\_b_i = 1$ if $score\_max_i \geq 80$ and $score\_min_i < 80$. $margin\_b_i = 0$ if $score\_min_i \geq 80$ and $score\_max_i < 90$. To keep hygiene quality comparable between inspections on the margin of B and grade-B inspections not on the margin, the paper only keeps inspections on the margin of B with $score\_min \in [76, 79]$ and grade-B inspections not on the margin with $score\_min \in [80, 85]$. The coefficient of $margin \times no\_OII$ is not significant, supporting Hypothesis 2 that inspectors' propensity to inflate a restaurant's grade to B is not affected by their past OII workload. The coefficient of $margin$ is significantly positive, which suggests that though OII avoidance no longer plays a role in inspectors' decision on grade inflation at the margin of B, other motivations are still driving inspectors to bump up restaurants' grades to B. The coefficient of no_OII is significantly positive as well, which indicates that inspectors with a heavier past OII workload are more likely to issue a lesser deduction on discretionary violations when the routine inspection is not on the margin of getting a B. Unfortunately, this paper does not have an explanation for this particular result.

Model (1.2) is estimated separately among a subsample of OII-eligible routine inspections (column (3)) and among a subsample of OII-ineligible routine inspections (column (4)). A routine inspection is OII-eligible if the restaurant's last OII is 365 days or more apart from the routine inspection, and is OII-ineligible otherwise.[44] The coefficient of $margin \times no\_OII$ is significantly positive in column (3), while insignificant in column (4), confirming Hypothesis 3 that past OII workload only affects an inspector's propensity to inflate a restaurant's grade to A when the restaurant is eligible for an OII. It is worth noting that the coefficient of $margin$ is significantly positive in column (4), which suggests that inspectors inflate grades to A for OII-ineligible restaurants for reasons other than OII aversion. Perhaps the inspectors do not want to upset restaurant owners by issuing a B when the restaurant will not be able to improve its grade through an OII.

The above analyses provide evidence for OII avoidance as a motivation for grade inflation practiced by inspectors.[45] To reduce grade inflation motivated by an inspector's aversion to OIIs, changes need to be made in terms of how an inspector's productivity is measured. This paper suggests counting the number of both routine inspections and OIIs into the annual PE and the fulfillment of the monthly goal. One may worry that this may cause inspectors to purposefully depress restaurants' scores to under 90 to increase the number of OIIs they perform. This outcome is not considered plausible for the following two reasons: First, food facility operators have the right to request additional information or clarification regarding decisions made by the inspectors. They may also request that their inspection reports be reviewed by a supervisor at the local DSE office. Therefore, any unreasonably harsh score deductions will elicit questions and even complaints to the higher level of authority from the food facility operators.[46] Second, according to one informant, the reason why the current

[44]Information on the inspections prior to a routine inspection is required to determine if it is OII-eligible. Similar to what the paper has done to compute ln_visit_i and $ln(1 + exp_i)$ discussed in footnote 42, data on OIIs and routine inspections dating back to Oct 1st, 2015 are used to determine a routine inspection's OII-eligibility status. Routine inspections with an ambiguous OII-eligibility status due to dubious data entries are not used for estimation presented in column (3) and (4).

[45]Model (1.2) includes an interaction term between margin and the number of OIIs performed by an inspector in the past 30 days while controlling for the number of routine inspections performed by an inspector in the past 30 days. An alternative model is to include an interaction term between margin and an inspector's OII-routine-inspection ratio in the past 30 days. The results under this alternative model are qualitatively similar.

[46]In fact, Ibanez and Toffel (2020) have observed food-safety inspections conducted in Lake County, IL, in

promotion policy takes into account the placement in the civil serve exam in addition to the number of routine inspections performed is so that inspectors would not boost the number of routine inspections they conduct by rushing through inspections or faking inspections they do not conduct. That being said, the current promotion policy has already been designed to reduce an inspector's incentive to boost the number of inspections through unethical means. Therefore, there is no reason to believe that an inspector would boost the number of OIIs through unprofessional means under the same promotion policy where inspectors are not found to improperly boost the number of routine inspections.

1.6.2 Attachment

Existing literature has established that repeated interactions between the inspector and the subject being inspected lead to fewer citations (Short et al., 2016; Jin and Lee, 2018; Kovács et al., 2020). In the LA County restaurant inspection setting specifically, Kovács et al. (2020) find that inspectors with more repeated visits to a restaurant are more likely to mark a higher number of health codes as "Not Observed", thereby avoiding potential violations and corresponding penalties. One possible explanation is that inspectors have formed attachment to the restaurants through repeated inspections and are reluctant to give them a grade below A.[47]

Currently, each LA County restaurant inspector is in charge of one inspection district for two years, after which he/she is assigned to a new inspection district within the same office. The lack of rotation of inspectors is a likely cause behind inspector's attachment to the

Camden County, NJ and in Alaska, and they witnessed several occasions in which establishment staff were frustrated even when an inspection yielded only one violation.

[47]This paper uses "attachment" (Kovács et al. (2020) use "social relationship") as an umbrella term for an inspector's motives behind grade inflation resulting from repeated interactions with a restaurant. These motives, though different, can all be addressed by more frequent rotation of inspectors. Therefore, this paper does not distinguish between them and uses "attachment" as an umbrella term for them. Refer to Appendix A.4 for a discussion of possible explanations for grade inflation under repeated interactions between an inspector and a restaurant.

restaurants. To reduce grade inflation motivated by an inspector's attachment to restaurants, this paper follows the existing literature (Jin and Lee, 2018; Kovács et al., 2020) and suggests more frequent inspector rotation within each office. Since a restaurant receives at most three inspections per year, randomly rotating inspectors across districts within an office *every four months* would be sufficient to achieve the maximum benefits of rotation within an office: a given restaurant has a probability of 1/8 to be assigned a given inspector in each routine inspection (recall that each office has eight inspectors). The median lifespan of restaurants is about 4.5 years (Luo and Stark, 2015). Under the four-month rotation scheme, a given inspector is expected to inspect a high-risk restaurant (inspected three times a year) with a median lifespan 1.7 times throughout its lifetime, a medium-risk restaurant (inspected twice a year) 1.1 times, and a low-risk restaurant (inspected once a year) 0.6 times. Therefore, it is unlikely that a restaurant establishes a social relationship with any inspector during its lifetime and receives grade inflation motivated by an inspector's attachment to it.[48]

One may be concerned about the increased costs associated with more frequent inspector rotation. There are two main costs: administrative costs and transportation costs.[49] Since the inspector rotation is still kept within each office, the organizational structure where inspectors are managed by the office they belong will remain the same, and thus it is unlikely that administrative costs will increase significantly. A more pressing concern arises from

[48]Using Bureau of Labor Statistics Quarterly Census of Employment and Wages longitudinal database from 1992 to 2011 for eight western states in the US including California, Luo and Stark (2015) estimate the median lifespan to be 4.5 years for full-service restaurants excluding multi-establishment and chain restaurants. The lifespan is not specifically estimated for the restaurants in LA County, but it is the best estimate available in the literature. Chain restaurants are likely to have a median lifespan of more than 4.5 years, as chain restaurants have higher survival rates compared to independent restaurants (Kalnins and Mayer, 2004; Parsa et al., 2011). Therefore, chain restaurants are more likely to develop a social relationship with inspectors, the extent of which this paper is unable to discuss due to data limitations.

[49]One possible advantage of repeated interactions between the inspector and the restaurant, as mentioned in Kovács et al. (2020), is that restaurants may be more willing to share details of their operations and in return, inspectors may be better able to aid restaurants in identifying problems in their food handling routines and learning from previous compliance lapses. More frequent inspector rotation may impede this type of cooperation beneficial to public health that arises from repeated interactions. However, to the best of my knowledge, the prevalence of this type of cooperation or whether it even exists has not yet been documented in the literature. As a result, this paper is unable to evaluate this particular social cost arising from more frequent rotation of inspectors.

the increase in transportation costs. Jin and Lee (2018) expect more frequent rotation of inspectors to increase transportation costs and therefore reduce the number of inspections that can be conducted by an inspector for a given period of time. However, changing from rotating inspectors every two years to rotating inspectors every four months would not substantially increase transportation costs in the context of LA County. On a given day, an inspector minimizes travel time by picking restaurants that are close to each other from those due for an inspection in the list of restaurants in his/her assigned district. Therefore, whatever the rotation schedule is, the travel time *between restaurants* should not change too much because of an inspector's optimization behavior. As an inspector is rotated across districts more often, it is more likely that he/she has to travel to a district farther away from home. However, an office does not serve a large geographical area (141.2 square miles on average).[50] As a result, the travel time from an inspector's home to a district and vice versa should not change too much even when an inspector is rotated to a farther district within an office. Therefore, more frequent inspector rotation within each office can alleviate grade inflation resulting from repeated inspections without incurring high administrative or transportation costs.

This section discusses two possible motives of inspector leniency: OII avoidance and attachment to the restaurants. The remedies to inspector leniency corresponding to the two motives are drastically different. This discussion offers a possible explanation as to why the grading policy change implemented by LADPH back in 2016 failed to achieve its intended effects: the bunching at 90 defies a simple/mechanical fix and requires a deeper understanding of inspectors' possible motives.

[50] According to the county boundary map available on LA County's GIS hub (https://egis-lacounty.hub.arcgis.com/datasets/lacounty::county-boundaries/about), LA County has a total land area of 3953.403 square miles excluding Santa Catalina Island and San Clemente Island (About 400 restaurant are located in the city of Avalon in Santa Catalina Island, but they constitute a negligible proportion of the inspection workload). There are 28 offices with full capacity (29 offices with 2 offices operating with half capacity under the same chief EHS), and therefore, the average geographical area each office serves is 3953.403/28=141.2 square miles.

To evaluate how the two proposed remedies to inspector leniency are going to affect inspectors' grading and restaurants' sanitation efforts, a model is developed in the next section.

1.7 Model

To provide a theoretical framework for understanding the effects of inspector leniency on restaurant behavior and the scores assigned by inspectors, this section constructs a model of a restaurant picking sanitation effort to maximize expected profit under inspector uncertainty. Section 1.7.1 introduces the model and solves the optimization problem when there is no uncertainty. Section 1.7.2.1 and Section 1.7.2.2 model a non-lenient inspector and a lenient inspector respectively by assuming different distributions of inspector uncertainty. Based on these two models, Section 1.7.2.3 develops a hybrid model where a representative inspector is lenient to some restaurants, while non-lenient to the other restaurants.

1.7.1 Model setup

A model is constructed where a restaurant picks effort e to maximize expected profit, which equals expected revenue $\mathbb{E}(R)$ minus the sanitation cost C:

$$\underset{e\in[80,100]}{max}\,\mathbb{E}(R(e+\varepsilon)) \; - \; C(e). \tag{1.3}$$

The optimal effort e^*, is defined as follows:

$$e^* \equiv \underset{e\in[80,100]}{argmax}\,\mathbb{E}(R(e+\varepsilon)) \; - \; C(e).$$

In the optimization problem (1.3), e is the sanitation effort made by the restaurant. Considering that only 0.63% of the inspections have a score less than 80, the lowest score to be qualified for grade B, the lower bound of e is set to be 80. The sanitation cost, $C(\cdot)$, is a function of e, with $C'(\cdot) > 0$, and $C''(\cdot) > 0$. Inspector uncertainty ε is a random variable with support $(\underline{\varepsilon}, \bar{\varepsilon})$ and probability density function $f(\varepsilon)$, with $\underline{\varepsilon} \leq 0$ and $\bar{\varepsilon} \geq 0$. ε is added to e to get the assessed sanitation effort $e+\varepsilon$. The effort is normalized such that the assessed effort $e+\varepsilon$ yields score $e+\varepsilon$ in the inspection. $f(\varepsilon)$ characterizes the restaurant's belief of how the inspector-assessed effort would deviate from its actual sanitation effort. One critical assumption of the model is as follows:

Assumption 1. The restaurant's belief of how an inspector behaves is the same as how that inspector actually behaves.

Therefore, $f(\varepsilon)$ also characterizes how an inspector assesses the sanitation effort to be different from the restaurant's actual effort. The validity of Assumption 1 will be discussed in Section 1.8.2.

The following assumption regarding $f(\varepsilon)$ is imposed:

Assumption 2. $f(\varepsilon)$ is inspector-specific, but not restaurant-specific. In other words, $f(\varepsilon)$ varies across inspectors, while $f(\varepsilon)$ is the same for different restaurants visited by a given inspector. However, the realized value of ε from $f(\varepsilon)$ can vary across restaurants visited by a given inspector.

Recall that customers see only the letter grade and not the underlying score of an inspection, so revenue $R(\cdot)$ is a step function of the score:

$$R(e+\varepsilon) = \begin{cases} \alpha & e+\varepsilon \geq \hat{s}, \\ \beta & e+\varepsilon < \hat{s}, \end{cases} \tag{1.4}$$

where $\hat{s}$ is the lowest possible score to get grade A. In our case, $\hat{s} = 90$. The parameter α is referred to as revenue under grade A, and β is referred to as revenue under grade B. Both

α and β vary with a restaurant's characteristics such as price level and capacity, but $\alpha > \beta$ always holds. The following assumption is required for equation (1.4):

Assumption 3. Customers care about the health grade of a restaurant, so that $\alpha > \beta$.

Using restaurants' quarterly revenue data, Jin and Leslie (2003) show that A-grade restaurants witnessed a 5.7% increase in revenue on average after LA County introduced mandatory grade card disclosure in 1998, while B-grade restaurants and C-grade restaurants witnessed a 0.7% increase and a 1.0% decrease in revenue on average respectively. This is evidence that customers do respond to differences in health grades. Moreover, a survey conducted by LADPH in 2015 shows that more than 85% of respondents consider restaurant grades before eating out (Baer, 2016a). Equation (1.4) requires the customers to only care about the health grade, but not the actual cleanliness of a restaurant. This assumption will be relaxed later in Section 1.7.2.3.

When there is no uncertainty, i.e., $\varepsilon = 0$, then $e^* = 90$ if $\alpha - \beta \geq C(90) - C(80)$. Similarly, $e^* = 80$ if $\alpha - \beta < C(90) - C(80)$. The restaurant would pick the minimum effort to qualify for an A when the revenue gap is sufficiently large compared with the cost. The restaurant would pick the lowest possible effort otherwise.

1.7.2 Models of restaurants maximizing expected profit under inspector uncertainty

When there is uncertainty, the objective function for the maximization problem (1.3) is:

$$\alpha P(\varepsilon \geq 90 - e) \; + \; \beta P(\varepsilon < 90 - e) \; - \; C(e).$$

The objective function can be further written as

$$\alpha \int_{90-e}^{\bar{\varepsilon}} f(\varepsilon) d\varepsilon \ + \ \beta \int_{\underline{\varepsilon}}^{90-e} f(\varepsilon) d\varepsilon \ - \ C(e).$$

Conditional on e, revenue equals α when ε assumes values such that $e+\varepsilon \geq 90$, and equals β otherwise. Using Leibniz's rule, the interior solution of optimal effort e^* satisfies the following first-order and second-order conditions when f is differentiable everywhere:

$$F.O.C. \ \ (\alpha - \beta) f(90 - e^*) \ = \ C'(e^*), \tag{1.5}$$

$$S.O.C. \ \ -(\alpha - \beta) f'(90 - e^*) \ - \ C''(e^*) \ < \ 0. \tag{1.6}$$

Equation (1.5) indicates that at e^*, marginal revenue $(\alpha - \beta) f(90 - e^*)$ from an increase in effort should equal to marginal cost $C'(e^*)$. Note that marginal revenue equals the difference in revenue between grade A and grade B $(\alpha - \beta)$ times the probability $f(90 - e^*)$ that the extra effort shifts the grade from B to A.

The above analysis under unconstrained optimization gives us the economic intuition behind the optimization, assuming that f is differentiable everywhere and that the interior solution e^* happens to fall into the range of [80, 100]. However, complications would arise in the case of corner solutions or when the differentiability assumption does not hold. Therefore, in the following analyses, e^* is solved for computationally, and simulations are used to explore various properties of the models under specific forms of $f(\varepsilon)$, where the differentiability assumption does not hold.

1.7.2.1 Model 1: A non-lenient inspector

Assume that ε follows a symmetric triangular distribution centered around 0. Symmetry implies that the inspector is as likely to overassess as to underassess effort. This is a model where the inspector is non-lenient. The triangular distribution is given by

$$f(\varepsilon) = \begin{cases} \frac{1}{u_1} + \frac{1}{u_1^2}\varepsilon & -u_1 \leq \varepsilon \leq 0, \\ \frac{1}{u_1} - \frac{1}{u_1^2}\varepsilon & 0 < \varepsilon \leq u_1, \end{cases} \tag{1.7}$$

where u_1 can be interpreted as the degree of uncertainty. The larger u_1 is, the larger the probability that the assessed sanitation effort deviates from the real effort. Equation (1.7) can also be written as $\varepsilon \sim tri(-u_1, 0, u_1)$.[51] Figure 1.4a plots the distribution.

Next, a simulation is done where 500 restaurants are inspected by the same non-lenient inspector, whose distribution of inspector uncertainty ε is specified by (1.7). $u_1 = 2$ is chosen, which means an inspector can omit or impose additional violations worth up to two points.[52] The cost function is defined as $C(e) = c(e - 80)^2$. 500 optimization problems defined by (1.3) are solved under different sets of $\{\alpha_i, \beta_i, c_i\}$ to allow restaurant heterogeneity, where i indexes restaurants. $\{\alpha_i\}_{i=1}^{500}$ are independently drawn from $\alpha_i \sim tri(300, 450, 600)$. $\{c_i\}_{i=1}^{500}$ are independently drawn from $c_i \sim tri(0.5, 1.25, 2)$. Each of $\{\beta_i\}_{i=1}^{500}$ is computed as $\beta_i = 0.5\alpha_i - 100$, reflecting the assumption that $\alpha_i - \beta_i$, the revenue gap between grade A and B, monotonically increases with α_i, the revenue under grade A ($\alpha_i - \beta_i = 0.5\alpha_i + 100$). The values of α, β and c are chosen such that the revenue gain from exerting more sanitation effort outweighs the cost for the majority of the restaurants, so that $e^* \neq 80$ holds for the majority of the restaurants. This is because only about 4.26% of inspections yield scores from 75 to 85, which indicates that the corner solution of $e^* = 80$ is not the solution to the optimization problems of the majority of restaurants. The optimal effort picked by the restaurant, e_i^*, is computed for each set of $\{\alpha_i,\ \beta_i,\ c_i,\ u_1\}$. Assessed effort is computed as $s_i = e_i^* + \varepsilon_i$, where ε_i is independently drawn from the distribution $\varepsilon_i \sim tri(-u_1, 0, u_1)$.[53] Therefore, e_i^* and s_i are not restricted to be integers. Assigned score, $score_i$, is computed

[51] $x \sim tri(a, b, c)$ means x follows a triangular distribution with lower limit a, mode b and upper limit c.

[52] Recall that under Assumption 2, ε is not restaurant-specific. Therefore, u_1 takes the same value for all the restaurants under a given inspector, not having subscript i.

[53] s_i could exceed 100 in practice . In this case, s_i is set to be 100.

as $score_i = floor(s_i)$, which means $score_i$ takes the value of the nearest integer less than or equal to s_i.

Figure 1.5a shows the distribution of optimal sanitation efforts. The optimal efforts are above 90, between 91 and 92. Figure 1.5b shows the distribution of simulated scores, which is close to a triangular distribution. The shape of the distribution resembles the empirical score distribution for scores above 90. However, the simulation is not able to produce as many scores equal to or above 95 as in the empirical distribution. Most importantly, the simulated distribution does not have a spike at 90.

1.7.2.2 Model 2: A lenient inspector

Assume ε follows the following triangular distribution:

$$f(\varepsilon) = \frac{2}{u_2^2}\varepsilon, \;\; 0 \le \varepsilon \le u_2, \tag{1.8}$$

which can also be written as $\varepsilon \sim tri(0, u_2, u_2)$. Figure 1.4b plots the distribution. This is a model that incorporates inspector leniency.[54] As $\varepsilon > 0$, restaurants believe that the inspector-assessed sanitation efforts are higher than their real efforts.

Next, a simulation similar to that in Section 1.7.2.1 is done, with the exact same set of $\{\alpha_i, \beta_i, c_i\}_{i=1}^{500}$, but u_2=1 is assumed. In other words, the same sample of restaurants as in Section 1.7.2.1 are now inspected by a lenient inspector, who has a different form of $f(\varepsilon)$ than the non-lenient inspector in Section 1.7.2.1. u_2 is set to 1 so that a restaurant's score

[54] Under this model (henceforth referred to as the baseline model), an inspector bumps up a restaurant's score no matter what the score would have been based on a restaurant's effort level. A more intuitive way of modelling inspector leniency would be to bump up a restaurant's score only when it would have been one or two points under a certain grade based on its effort level (henceforth referred to as the threshold model), which is presented in Appendix A.5. It turns out that the simulated distribution of scores and the simulated distribution of restaurant sanitation efforts are similar under the threshold model as under the baseline model. Appendix A.5 also discusses the pros and cons of both models and explains why the baseline model is used for the main analyses.

is not bumped up to a level too high compared to the score a restaurant's real sanitation effort would have yielded.

Figure 1.6a shows the distribution of optimal sanitation efforts. The optimal efforts are between 89 and 90. Compared with the optimal efforts that are above 90 under a non-lenient inspector, it shows that a restaurant has less incentive to exert sanitation effort when expecting a lenient inspector. Figure 1.6b shows the distribution of simulated scores, all of which equal 90. Figure 1.6b captures one important feature of the empirical score distribution: the high proportion of 90.

1.7.2.3 The hybrid model

Each of the above two models captures one key feature of the empirical score distribution shown in Figure 1.1. Model 1 with a non-lenient inspector captures the shape of the distribution above 90, while model 2 with a lenient inspector captures the high proportion of 90. These two models follow Assumption 2, which restricts an inspector's leniency level to be binary (i.e., lenient or non-lenient). To allow the model to capture both features of the empirical score distribution, Assumption 2 is relaxed such that a given inspector can be lenient to some restaurants, while non-lenient to the others, allowing an inspector's leniency level to change continuously. This relaxed assumption is formally stated as follows:

Assumption 4. $f(\varepsilon)$ is restaurant-specific. For a given inspector, $f(\varepsilon)$ can vary across restaurants.

Specifically, for a given inspector, $f(\varepsilon)$ now takes the form of (1.7) with probability $1-\theta$, and takes the form of (1.8) with probability θ. This means that if an inspector is responsible for N restaurants, he/she is lenient to $N\theta$ restaurants, but not lenient to the remaining $N(1-\theta)$ restaurants. Assumption 1 indicates that the restaurants that the inspector inspects leniently believe that they will be treated leniently, and will solve the optimization problem assuming $f(\varepsilon)$ takes the form of (1.8). The restaurants that the inspector inspects non-leniently believe that they will be treated non-leniently, and will solve the optimization problem assuming $f(\varepsilon)$ takes the form of (1.7).

To account for the issue that the simulation in Section 1.7.2.1 is not able to produce scores equal to or higher than 95, the optimization problem is revised further based on Assumption 5, which is an extension of Assumption 3:

Assumption 5. Customers not only care about the health grade of a restaurant, but also the actual cleanliness of the restaurant, which adds $r(e)$ to the revenue, where $r'(\cdot) > 0$ and $r''(\cdot) < 0$. $r(e)$ quantifies how much customers value the actual cleanliness of a restaurant.

The final model is based on Assumption 1, Assumption 3, Assumption 4 and Assumption 5, and features the following optimization problem for a given restaurant:

$$\underset{e \in [80,100]}{max} \ \mathbb{E}(R(e+\varepsilon)) \ - \ C(e) \ + \ r(e), \tag{1.9}$$

where $f(\varepsilon)$ now takes the form of (1.8) for a restaurant that is inspected leniently, and takes the form of (1.7) for a restaurant that is inspected non-leniently. This model is henceforth referred to as the hybrid model, as it is a hybrid version of the model with a non-lenient inspector and the model with a lenient inspector.

1.8 Simulations under the hybrid model

1.8.1 Calibration of the hybrid model

The following functional forms are used: $C(e) = c(e-80)^2$, $r(e) = r(e-80)^{\frac{4}{5}}$. To make sure the parameters α, β, c and r incorporate the attributes of the restaurants in the sample, the values for parameters α, β, c and r are derived as follows. First, 500 observations are randomly picked (without replacement[55]) from the sample of restaurants with nonmissing capacity level, risk level and price level, and these three attributes are extracted: $\{price_i, capacity_i, risk_i\}_{i=1}^{500}$, where i indexes restaurants.[56] Second, a proxy is generated for cost ($cost$) that equals risk multiplied by capacity: $cost_i = risk_i \times capacity_i$, and a proxy is generated for revenue ($revenue$) that equals price multiplied by capacity: $revenue_i = price_i \times capacity_i$. This is based on the assumption that restaurants with a higher risk level (selling food that requires more careful handling) and a larger capacity have higher sanitation costs, while restaurants with a larger capacity and a higher price level have a higher revenue. Third, the two proxies are linearly transformed to get $\{\alpha_i\}_{i=1}^{500}$ ($\alpha_i = 20revenue_i + 400$) and $\{c_i\}_{i=1}^{500}$ ($c_i = 0.2cost_i + 0.5$). For example, for a high-risk restaurant with price level \$\$\$ and 0-30 seats, $price = 3$, $capacity = 1$, and $risk = 3$. Proxies for cost and for revenue are generated as follows: $cost = risk \times capacity = 3$, and $revenue = price \times capacity = 3$. They are then linearly transformed to get the parameter for revenue under grade A ($\alpha = 20revenue + 400 = 460$), and the parameter for sanitation cost ($c = 0.2cost + 0.5 = 1.1$). The above procedure attempts to capture the unobservable distributions of sanitation costs and revenues under grade A across restaurants into the distributions of α and c. As in Section 1.7.2.1 and Section 1.7.2.2, $\beta_i = 0.5\alpha_i - 100$, while r_i is set to be positively correlated with α_i ($r_i = 0.25\alpha_i - 75$), as the additional revenue bonus from the restaurant's cleanliness should be positively correlated with its revenue.

[55]In the model simulation, an inspector never inspects the same restaurant more than once, which deviates from the fact that an inspector can inspect the same restaurant more than once in practice. The reason for this deviation is as follows: In practice, certain factors, such as the time of the day the inspection occurs, would change the model parameters, and therefore a given restaurant can exert different amounts of optimal effort if inspected multiple times. However, the parameters in this model are either assumed to be restaurant-specific or inspector-specific, and thus will not change if a given restaurant is inspected by the same inspector repeatedly. As a result, a given restaurant would pick the same optimal effort if inspected multiple times in this model, making it pointless to repeat restaurants in the simulation sample.

[56]Refer to Section 1.3.1 on how the variables *capacity* and *risk* are constructed. Data used to construct the variable *price* are from Yelp.com. Figure A.4 presents the distributions of the three variables in the sample of restaurants with nonmissing values for these variables. Appendix A.6 describes the procedure to match restaurants' price levels on Yelp.com to the restaurants from the main inspection dataset, and discusses the sample selection issue introduced by using only the attributes from restaurants observable on Yelp.com for model calibration.

The paper runs five simulations similar to the ones in Section 1.7.2.1 and Section 1.7.2.2, with $u_1 = 2$ and $u_2 = 1$. The five simulations involve the same set of $\{\alpha_i, \beta_i, c_i, r_i\}_{i=1}^{500}$, but different values of θ. These simulations can be understood as follows: The same 500 restaurants are inspected by five representative inspectors who differ only in θ. These five inspectors' behavior is characterized by the same $f(\varepsilon)$ as in (1.8) when they are lenient, and is characterized by the same $f(\varepsilon)$ as in (1.7) when they are non-lenient, as u_1 and u_2 take the same values throughout the five simulations. However, inspectors are lenient to different proportions of restaurants, as the value of θ changes across simulations.[57] θ takes values 0, 0.3, 0.5, 0.7 and 1. The larger θ is, the larger proportion of restaurants the inspector is lenient to.

The proportion of 90 in the simulated distributions of scores increases as θ increases, as shown in Figure 1.7a. This result indicates that the more restaurants the inspector is lenient to, the larger the proportion of 90 is in the simulated distribution. Figure 1.7b presents the empirical score distributions of routine inspections for five selected inspectors, where the proportion of 90 varies across inspectors. The hybrid model provides a theoretical framework to understand how the difference in inspector leniency can generate the difference in the extent of bunching at 90 that is seen in the empirical score distributions in Figure 1.7b.[58]

The parameters α, β, c and r are calibrated so that the simulated score distributions produced by the hybrid model resemble the empirical score distributions. First, the linear

[57] It is randomly decided in this model (the baseline model) and the threshold model which restaurants the inspector is lenient to. Another alternative is to introduce a monotonicity assumption: If a restaurant is inspected leniently when the inspector is lenient to $100\theta_1\%$ of the restaurants, then the restaurant would also be inspected leniently when the inspector is lenient to $100\theta_2\%$ of the restaurants, given that $\theta_2 > \theta_1$. The simulation results under this assumption are similar to the simulation results presented in the paper and are available upon request.

[58] Figure 1.7a presents the simulated distributions of scores under specific functional forms for $C(e)$ and $r(e)$, and a specific set of values for α, β, c and r. However, the pattern that the proportion of 90 increases with θ remains as long as 1) $C'(e) > 0$, $C''(e) > 0$, $r'(e) > 0$, and $r''(e) < 0$; 2) The revenue gain $\alpha - \beta$ from getting an A is large enough compared to the increase in the sanitation cost $C(90) - C(80)$ so that getting an A is optimal for enough restaurants; 3) The market value of being clean $r(e)$ is not too large compared to the sanitation cost $C(e)$, so that enough restaurants would change their optimal efforts depending on the type of the inspector (lenient/non-lenient), instead of picking the same large e^* such that $C'(e^*) = r'(e^*)$ holds regardless of the type of the inspector.

transformations from the proxy *revenue* to α, and from the proxy *cost* to c affect the magnitude of c relative to α, which in turn affects the magnitude of sanitation cost relative to the revenue gain from grade-crossing ($\alpha - \beta$ is a function of α) that determines the percentage of restaurants with a score below 90. Therefore, the two linear transformations are picked so that there is only a small percentage of restaurants with a score below 90 across the simulated score distributions. Second, the values of r are then picked so that the market value of being clean ($r(e)$) relative to the sanitation cost ($C(e)$) is large enough to ensure that there are restaurants with scores in the upper 90s, but not too large such that all the restaurants exert huge sanitation efforts regardless of whether the inspector is lenient or not and end up scoring in the upper 90s.

However, the calibration is very preliminary as it utilizes the data on restaurants' capacity level, risk level and price level only. The simulated distributions of scores from the hybrid model also have two deviations from the empirical score distributions: First, they have a lower proportion of scores under 90. Second, they feature a lower proportion of scores above 95. Nonetheless, the calibration process ensures that the parameters are able to capture some key unobserved features of the restaurants (e.g., how high the sanitation cost is compared to the revenue, and how much the customers' preference for cleanliness contributes to a restaurant's effort), and therefore gives credibility to the subsequent evaluation of the two remedies to inspector leniency through the model.

1.8.2 Evaluating the proposed remedies to inspector leniency

To simplify the analysis, each of the two proposed remedies are evaluated assuming its targeted motive is the only motive behind inspector leniency.

1.8.2.1 Evaluating the remedy to OII avoidance

To alleviate an inspector's aversion to OIIs, this paper suggests counting the number of both routine inspections and OIIs into the annual PE and the fulfillment of the monthly goal. To evaluate this proposed policy change, it is assumed that any other policies regarding restaurant inspections remain unchanged, including the inspector rotation schedule.

Under the current rotation schedule, a restaurant is inspected by the same inspector for two years, and would therefore have enough interactions with the inspector to infer his/her type (lenient or non-lenient), validating Assumption 1. When the number of OIIs performed is counted into an inspector's productivity measures together with the number of routine inspections, the inspector would lose the incentive to avoid OIIs, and therefore would stop inflating grades for any restaurant he/she is in charge of. The restaurants the inspector is previously lenient to would update their belief of the inspector's type from lenient to non-lenient through repeated interactions. Therefore, under the revised inspector evaluation policy, the restaurant expects its inspector to be non-lenient when he/she is indeed non-lenient, again validating Assumption 1. In this case, θ, the percentage of restaurants the inspector is lenient to, drops to zero. The score distribution under this revised inspector evaluation policy would resemble the distribution under $\theta = 0$ in Figure 1.7a, with no apparent sign of bunching at 90. Figure 1.8 presents the distributions of restaurants' optimal efforts under different values of θ. As θ decreases to zero, the percentage of optimal efforts above 90 increases, indicating that some restaurants are exerting more sanitation effort under the revised inspector evaluation policy.[59]

[59] In Figure 1.8, the distribution of optimal effort $e^* > 92$ is the same across different values of θ. $C'(e^*) = r'(e^*)$ holds for restaurants with $e^* > 92$ regardless of $f(\varepsilon)$, the distribution of inspector uncertainty, which is why these restaurants exert the same amount of effort whether the inspector is lenient or not. These restaurants have a larger r/c ratio. The parameter r quantifies the value customers put on the actual cleanliness of a restaurant according to Assumption 5, while c quantifies the sanitation cost, which means the value customers put on cleanliness compared with the sanitation cost is larger in restaurants with a larger r/c ratio. Another interpretation of $r(e)$ is that it quantifies a restaurant's level of intrinsic motivation for being clean. In this case, other things being equal, a restaurant with a stronger intrinsic motivation for being clean exerts more sanitation effort in general ($e^* > 92$), and does not change its effort level depending on the

1.8.2.2 Evaluating the remedy to attachment

Under the current two-year rotation schedule, Assumption 1 holds, as discussed in Section 1.8.2.1. When the rotation is implemented every four months within an office, there is not sufficient interaction between an inspector and a restaurant for an inspector to get attached to a restaurant to be lenient during the inspection. When a restaurant has no chance to develop a social relationship with any inspectors, it would expect whichever inspector in charge of the upcoming inspection to be non-lenient. Therefore, under the new rotation scheme, the restaurant expects its inspector to be non-lenient when he/she is indeed non-lenient, validating Assumption 1.

When a given inspector is lenient to none of the restaurants, θ drops to zero. Similar to the effect of the revised inspector evaluation policy discussed above, the score distribution under the new rotation scheme would resemble the distribution under $\theta = 0$ in Figure 1.7a, with no apparent sign of bunching at 90. As θ decreases to zero, the percentage of optimal efforts above 90 increases as shown in Figure 1.8, indicating that some restaurants are exerting more sanitation effort.

1.8.2.3 Grade informativeness under the proposed remedies

One thing worth noticing is that the proportion of A barely changes with θ, as shown in Figure 1.7a.[60] Under the two remedies, restaurants whose grades are previously inflated to A by the inspectors expect the inspectors to be non-lenient, and therefore increase sanitation efforts to qualify themselves for an A, leaving the proportion of A almost unchanged. Though

inspector's type. On the other hand, a restaurant with a weaker intrinsic motivation for being clean exerts less sanitation effort in general ($89 < e^* < 92$), and decreases its effort when expecting a lenient inspector. This pattern is shown in Figure 1.8: As θ increases, the distribution of e^* between 89 and 92 shifts to the left.

[60] When $\theta = 0.3$, all but one inspections are assigned an A. When $\theta = 0.5$, all but two inspections are assigned an A. All inspections are assigned an A when $\theta = 0$, $\theta = 0.7$ and $\theta = 1$.

the proportion of A barely changes, the informativeness of the grade A has improved: The grade A category only includes restaurants who have earned their As under the two remedies, as apposed to including restaurants with worse hygiene conditions whose grades are inflated to A under the status quo.

The hybrid model allows a discussion of how a restaurant's score and sanitation effort are going to change *qualitatively* under the two proposed remedies to inspector leniency. One limitation of the model is that it is not able to *quantify* the change in a restaurant's score and sanitation effort, because 1) the model undergoes very preliminary calibration with data on restaurants' capacity level, risk level and price level only, and 2) the values of θ under the status quo in the two cases, the percentage of restaurants an inspector is lenient to out of OII avoidance and out of attachment respectively, are unknown, as the paper is unable to quantify how much each possible motive accounts for grade inflation.

1.9 Conclusion

This paper studies grade inflation performed by inspectors in the context of LA County restaurant inspections. It first discusses two public health consequences of grade inflation: 1) grade inflation reduces the effectiveness of health grades in signaling restaurants' hygiene conditions; 2) grade inflation discourages improvements in restaurant hygiene by preventing a re-inspection that a restaurant would likely request if they did not experience grade inflation. The paper then investigates two possible motives behind grade inflation and proposes targeted remedies. To reduce grade inflation motivated by an inspector's aversion to OIIs, the number of OIIs should be counted along with the number of routine inspections into an inspector's annual PE and the fulfillment of the monthly goal. Reducing grade inflation motivated by an inspector's attachment to restaurants calls for more frequent inspector rotation within each office. A model is developed to evaluate the two remedies, and both are found to reduce the degree of bunching at 90 and increase restaurants' sanitation efforts.

The study highlights the importance of understanding the motives behind grade inflation in curbing such inflation. The two motives discussed in this paper are drastically different and therefore call for drastically different remedies.

The study also has the following two limitations. First, it is not able to enumerate all the possible motives behind grade inflation, or pinpoint how much each possible motive accounts for grade inflation, which could be possible subjects for future studies. Second, the model developed in this paper allows for a qualitative study of how the proposed remedies for grade inflation would affect inspectors' grading and restaurants' sanitation efforts, but is not able to quantify the magnitude of such effects.

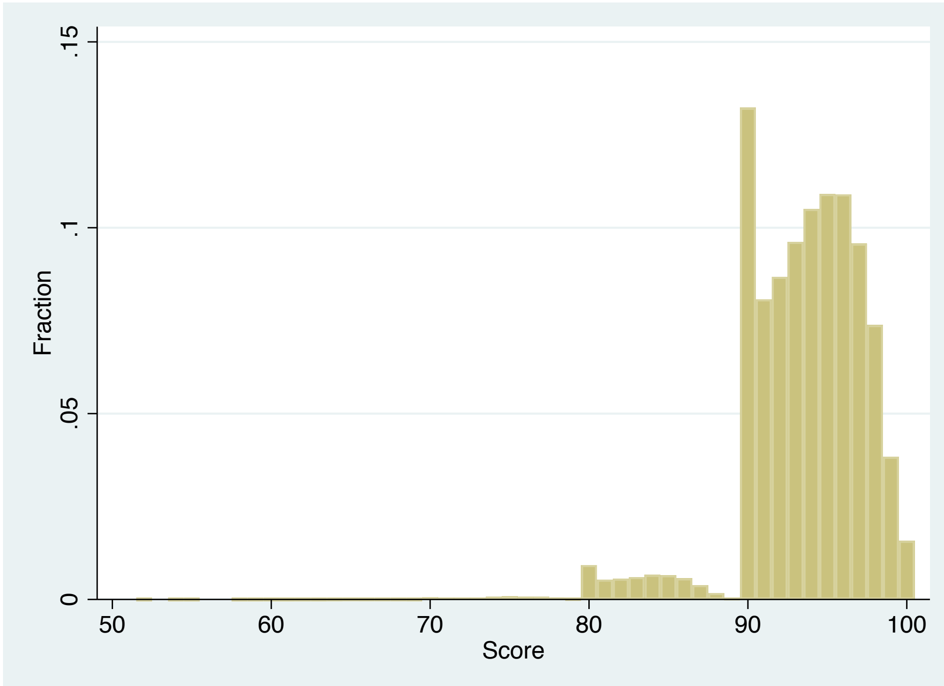


Figure 1.1: Distribution of routine inspection scores in the full sample

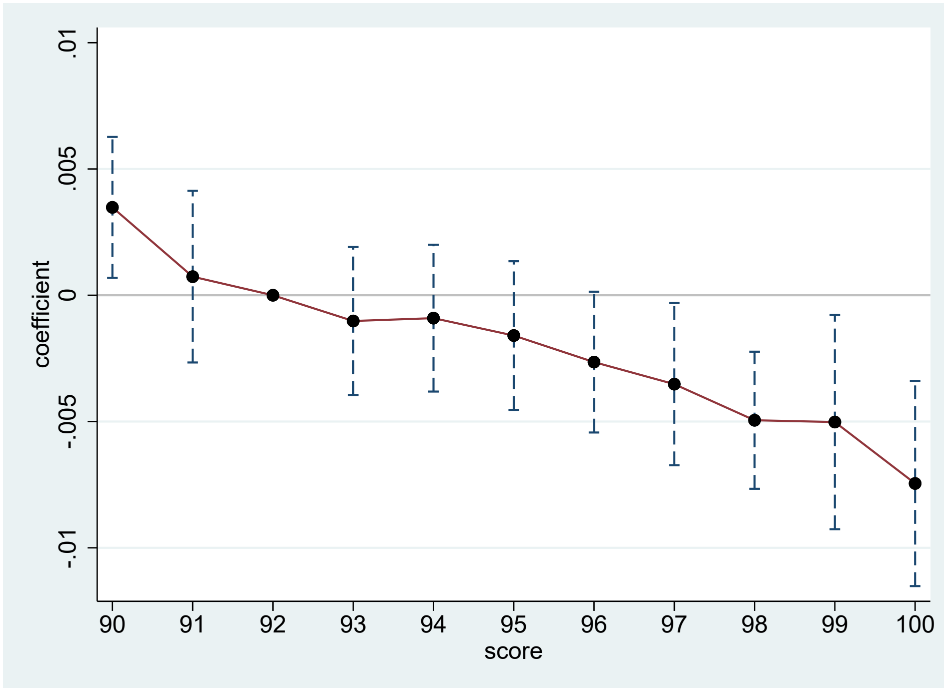
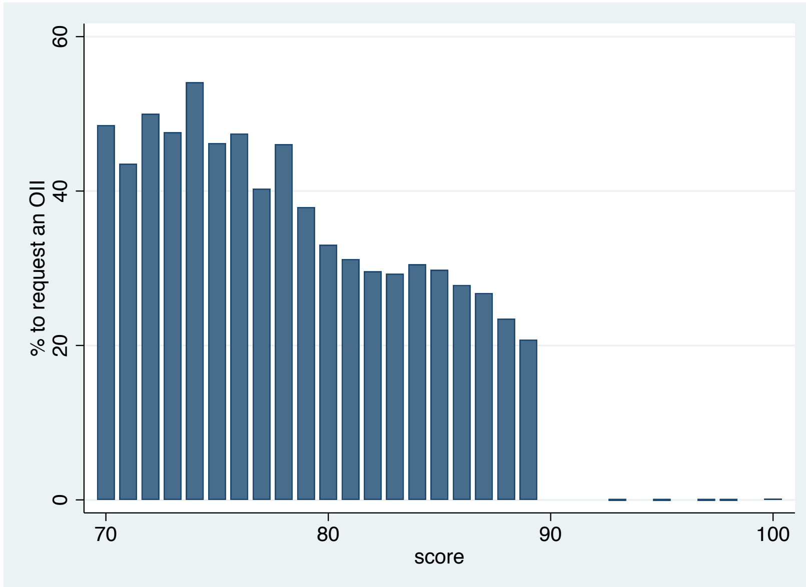
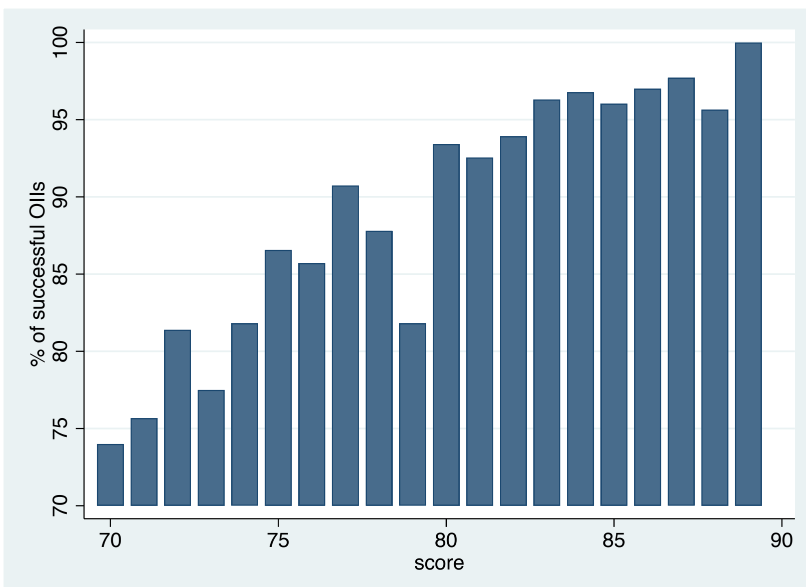


Figure 1.2: Comparisons of the probability of a subsequent complaint investigation where violations are found among score groups 90 to 100

Notes: The dependent variable equals 1 if a routine inspection is followed by a complaint investigation where violations are found and 0 otherwise. The y-axis plots the coefficients of the score dummies from the LPM that regresses the dependent variable on score dummies, while controlling for year fixed effects, month-of-year fixed effects, day-of-week fixed effects and city fixed effects. Score 92 is the base group and its dummy has a coefficient of 0. The dashed lines give the 95% confidence intervals computed using standard errors clustered at the city level.



(a) The percentage of routine inspections that result in an OII in each score level

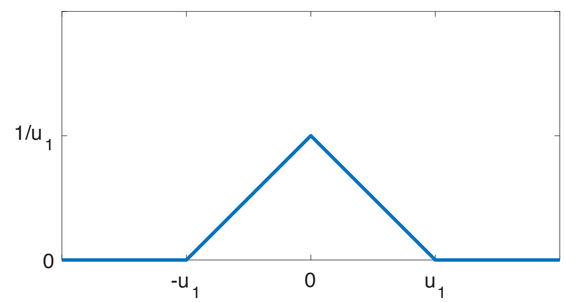


(b) The percentage of OIIs that yield an A

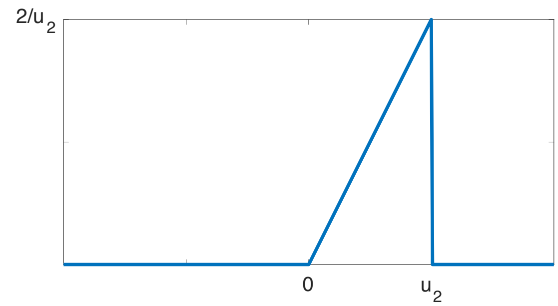
Notes: The sample includes routine inspections that lead to an OII. The x-axis shows the score of the routine inspection prior to the OII. The y-axis shows the percentage of OIIs that yield an A.

Figure 1.3: Understanding OIIs

Notes: Score categories under 70 are not shown as there are not enough inspections whose scores are under 70 in the full sample.

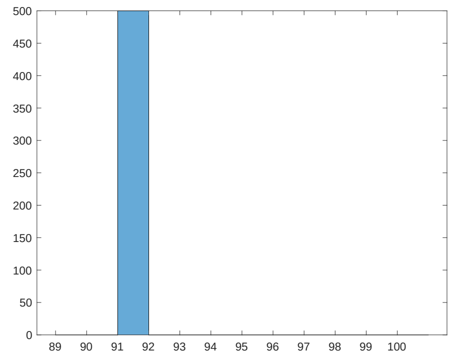


(a) Under a non-lenient inspector: $\varepsilon \sim tri(-u_1, 0, u_1)$

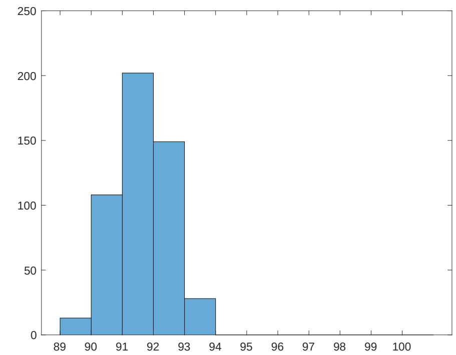


(b) Under a lenient inspector: $\varepsilon \sim tri(0, u_2, u_2)$

Figure 1.4: Distributions of inspector uncertainty



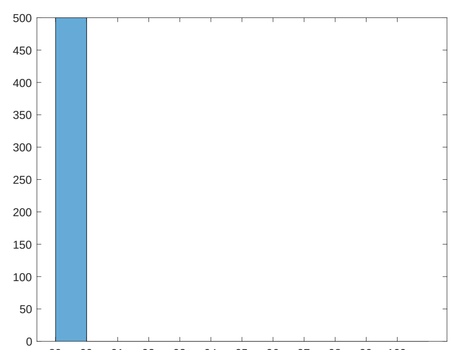
(a) Distribution of optimal efforts



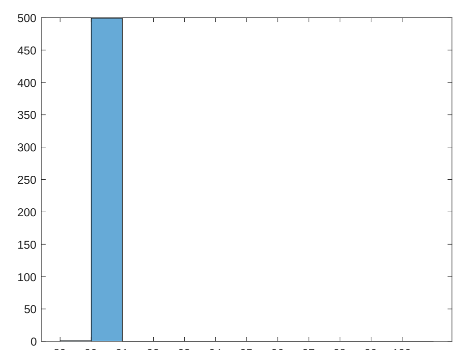
(b) Distribution of scores

Figure 1.5: The simulated distributions of optimal efforts and scores under a non-lenient inspector

Notes: The optimal efforts in Figure 1.5a are continuous, with values falling in the range of each bar. The scores in Figure 1.5b are integers, each equal to the lower integer of the bar it falls in.



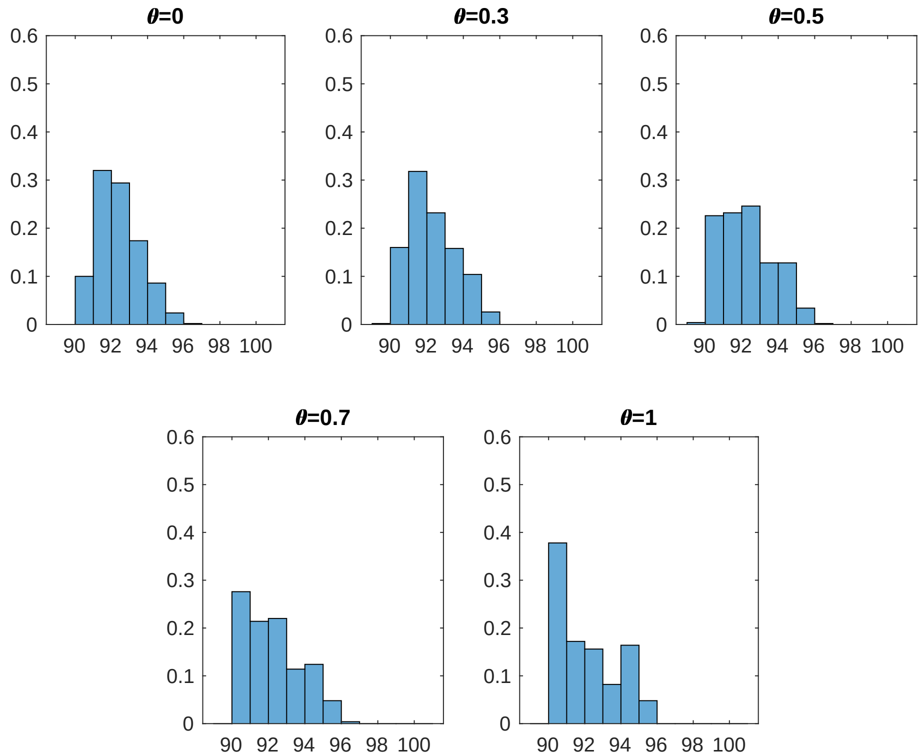
(a) Distribution of optimal efforts



(b) Distribution of scores

Figure 1.6: The simulated distributions of optimal efforts and scores under a lenient inspector

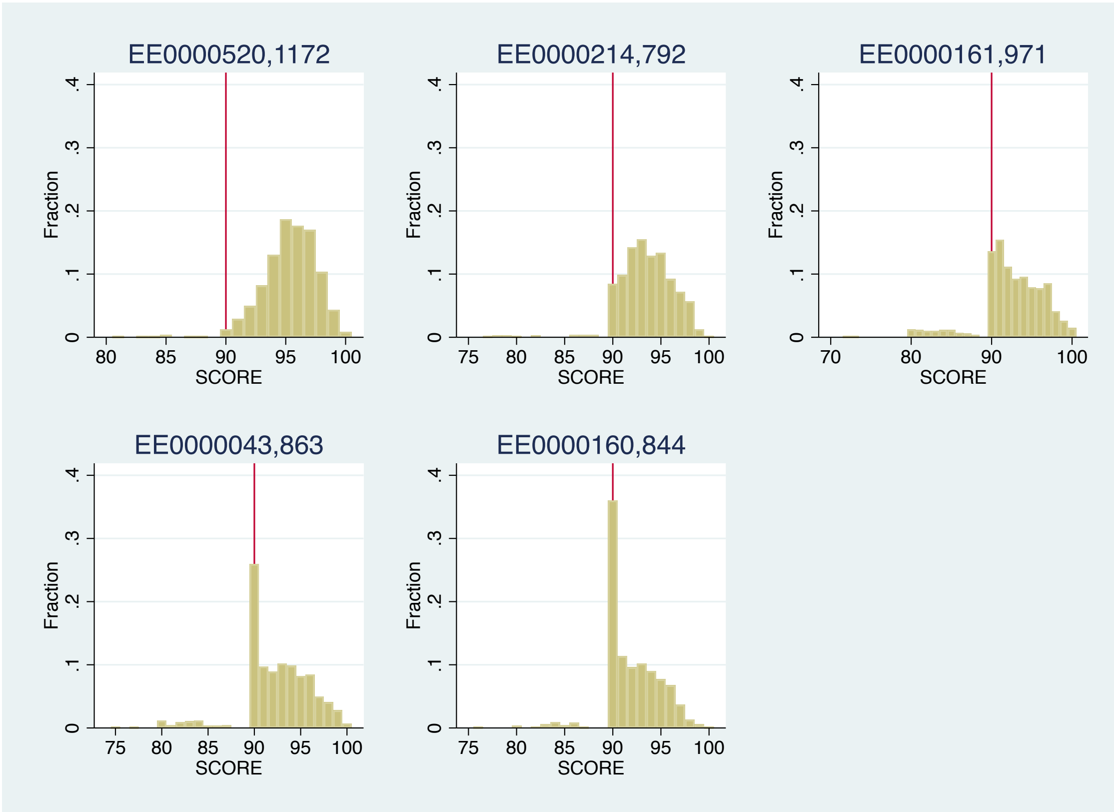
Notes: The optimal efforts in Figure 1.6a are continuous, with values falling in the range of each bar. The scores in Figure 1.6b are integers, each equal to the lower integer of the bar it falls in.



(a) Simulated distributions of scores under different values of θ

Notes: θ is the proportion of restaurants a representative inspector is lenient to. The simulated scores are integers, each equal to the lower integer of the bar it falls in.

(Continued on next page)



(b) Empirical score distributions of routine inspections for select inspectors

Notes: The title of each subfigure is the inspector's ID followed by the total number of routine inspections the inspector has carried out in the full sample. A red line is layered at score 90 to help readers spot it.

Figure 1.7: Simulated distributions of scores vs. Empirical distributions of scores

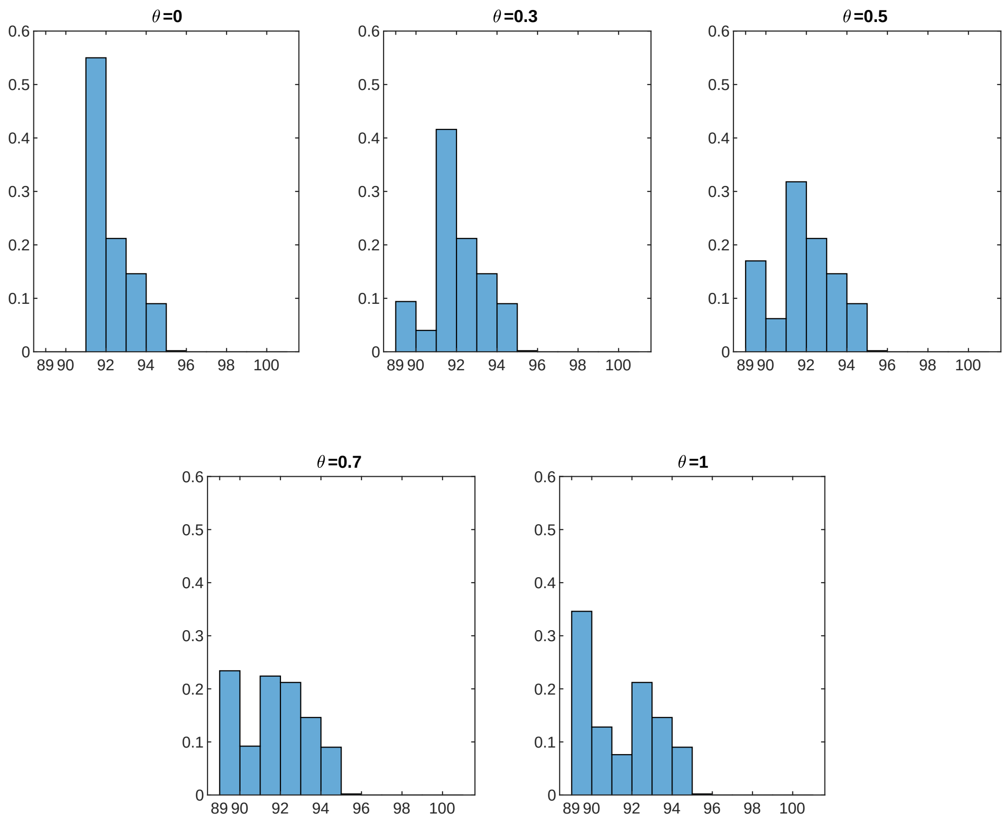


Figure 1.8: Simulated distributions of optimal efforts under different values of θ

Notes: θ is the proportion of restaurants a representative inspector is lenient to. The optimal efforts are continuous, with values falling in the range of each bar.

Table 1.1: The effects of grade inflation on restaurant hygiene

Variable	(1)	(2)	(3)
inflated	0.0110*	0.0111*	0.0109*
	(0.0061)	(0.0061)	(0.0060)
Day-of-Week FE	Yes	Yes	Yes
Month-of-Year FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
City FE	Yes	Yes	Yes
Inspector FE	Yes	Yes	Yes
Restaurant Type FE	Yes	Yes	Yes
No. of routine inspections w/ inflated=1	5,661	10,144	6,321
No. of routine inspections w/ inflated=0	267	267	267
No. of routine inspections	5,928	10,411	6,588

Notes: The dependent variable equals 1 if a routine inspection is followed by a complaint investigation where violations are found and 0 otherwise. In columns (1)-(3), "Routine inspections w/ inflated=0" include any routine inspection for a restaurant that receives a score between 85 and 89 without grade inflation in that routine inspection and later updates its grade to A through an OII. In column (1), "Routine inspections w/ inflated=1" include routine inspections which score 90 or 91 and whose grades are estimated to be inflated from B to A. In column (2), "Routine inspections w/ inflated=1" include routine inspections scoring 90 or 91 that are on the margin of A. In column (3), "Routine inspections w/ inflated=1" include routine inspections whose grades are estimated to be inflated from B to A. Standard errors clustered at the city level are in parentheses. Clustering standard errors at the inspector level yields qualitatively and quantitatively similar results, which are available upon request. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.2: Verification of OII avoidance as a motivation behind grade inflation to A

variable	(1) A-margin	(2) B-margin	(3) A-margin & OII-eligible	(4) A-margin & OII-ineligible
margin	0.2334***	0.2583***	0.2291***	0.3602***
	(0.0102)	(0.0119)	(0.0102)	(0.0266)
no_oii	-0.0029	0.0167***	-0.0031	-0.0041
	(0.0034)	(0.0044)	(0.0032)	(0.0138)
margin×no_oii	0.0168***	-0.0076	0.0167***	0.0085
	(0.0057)	(0.0052)	(0.0054)	(0.0193)
Observations	68,373	10,002	65,272	2,716
Controls	Yes	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes	Yes
Month-of-Year FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Zip Code FE	Yes	Yes	Yes	Yes
Restaurant Type FE	Yes	Yes	Yes	Yes
Inspector FE	Yes	Yes	Yes	Yes

Notes: The dependent variable equals 1 if two points are deducted for a discretionary violation, and equals 0 if four points are deducted. no_OII is the number of OIIs the inspector has completed 1 to 30 days prior. Controls include $no\_routine$, ln_visit and $ln(1 + exp)$. Standard errors clustered at the inspector level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Chapter 2

Do Electric Vehicle Charger Locations Respond to the Potential Charging Demands from Multi-Unit Dwellings? Evidence from Los Angeles County

2.1 Introduction

The transportation sector was the largest contributor of greenhouse gas (GHG) emissions in the United States in 2019, accounting for about 29 percent of nationwide emissions (Environmental Protection Agency, 2022). Between 1990 and 2019, GHG emissions in the transportation sector increased more in absolute terms than in any other sector largely due to increased demand for travel (Environmental Protection Agency, 2020). To move away from reliance on climate-change-causing fossil fuels, many states are actively planning and setting goals for the deployment of zero emission vehicles.[1] A zero emission vehicle (ZEV) is any type of vehicle that has no tailpipe emissions. There are two kinds of ZEV: plug-in electric vehicles (PEVs) and hydrogen fuel cell electric vehicles (FCEVs). Since PEVs have a larger market share than FCEVs and thus more available data, this paper focuses on the analysis of PEVs.[2] PEV is an umbrella term including both 100% battery electric vehicles (BEVs) and plug-in hybrid electric vehicles (PHEVs), which run primarily on batteries but have a back-up tank of gasoline. EV, short for electric vehicle and a term more commonly used in the media, is equivalent to PEV. The term EV and PEV will be used interchangeably in this paper.[3]

Currently, PEV owners mostly consist of people who live in single-family housing, where it is easier to install charging equipment. Overall, 50% to 80% of charging events for PEVs occur at home (California Air Resources Board, 2017; Lee et al., 2020). However, if states' ambitious zero-emission goals are to be achieved, it is important to attract potential buyers living in multi-unit dwellings (MUDs) as well, which consist of more than a quarter[4] of the housing stock in the United States. MUD, also known as multifamily housing, is a type of housing where multiple separate housing units (e.g., condominiums and apartments) for residential inhabitants are contained within one building structure. Unfortunately, the adoption rate of PEVs are low among MUD residents. Taking California as an example, MUDs constitute 32.2% of the housing stock according to 2019 American Community Survey, but fewer than 9% of ZEVs are purchased by MUD residents (Center for Sustainable Energy,

[1]For example, former California Governor Edmund G. Brown Jr.'s Executive Order B-16-2012 sets a goal of 1.5 million ZEVs by 2025. Meanwhile, California Governor Gavin Newsom issued an executive order on Sep 23, 2020 requiring sales of all new passenger vehicles to be zero-emission by 2035. Ten states have joined the Multi-State ZEV Action Plan to accelerate ZEV market growth. For a complete list of state EV action plans, refer to Howard et al. (2021).

[2]For example, according to California Energy Commission, by the end of 2020, PEVs make up 2.193% of the light-duty vehicles in California, while the FCEVs make up only 0.025%. Source: https://www.energy.ca.gov/data-reports/energy-insights/zero-emission-vehicle-and-infrastructure-statistics/vehicle-population.

[3]The paper refers to California Governor's Office of Business and Economic Development (2019) for the definitions of ZEV, PEV, FCEV, PHEV and BEV.

[4]Source: American Community Survey one-year estimates, 2015-2020.

2021).[5]

MUD residents are less likely to have dedicated parking space than residents living in single-family housing, and therefore have limited access to home charging (Alexander, 2022).[6] PEV owners residing in MUDs primarily rely on public charging due to a lack of EV charging infrastructure in MUDs (Chakraborty et al., 2019; Nicholas et al., 2019) and will continue to rely on public charging at least in the near future despite the ongoing effort of promoting charging infrastructure in MUDs.[7] The availability of public charging is an important factor, if not one of the most important factors, when the U.S. consumers decide whether to purchase a PEV (Axsen and Kurani, 2013; Kurani et al., 2016; Li et al., 2017; California

[5] The EV Consumer Survey conducted by the California Clean Vehicle Rebate Project collects data from consumers who purchased or leased an eligible clean vehicle, received a rebate, and responded to a voluntary CVRP Consumer Survey. For those who purchased a PEV between May 1st, 2016 and May 31st, 2017, took the EV Consumer Survey, and did not decline to answer the question, 8.21% of them live in an apartment or a condominium.

[6] Lack of home charging is prevalent among MUD residents, regardless of the income level. Alexander (2022) finds that even among MUD households with an annual income of more than $100,000, fewer than 45% of them have access to Level 1 charging at home, which is the most affordable and most common home charging charging option, as Level 1 charging only requires a residential 120-volt (120V) AC outlet, and most, if not all, PEVs come with a portable Level 1 cordset. Analyses using the 2019 California Vehicle Survey (National Renewable Energy Laboratory, 2019), which are presented in the online appendix, have similar findings: Even among households with an annual income of more than $200,000, fewer than 60% of them have access to electricity at home parking that can be potentially used for PEV charging.

[7] States as well as some local governments are beginning to integrate EV readiness requirements as part of their building codes. These readiness requirements do NOT require placing a charger in the space immediately, but require installation of "raceway" (the enclosed conduit that forms the physical pathway for electrical wiring to protect it from damage) and adequate panel capacity to accommodate future installation of a dedicated branch circuit and charging station(s). However, building code requirements for MUDs in general only apply to newly-constructed MUDs, existing MUDs undergoing substantial improvements, or new parking lots serving MUDs. For the previously existing MUDs or MUD parking lots, landlords have few incentives to install EV charging infrastructure due to high cost of retrofitting electrical infrastructure and uncertainty in whether the investments will pay off. Therefore, it will take some time for the building codes to have a noticable effects on the number of charges in MUDs. For a complete list of local EV infrastructure building codes in the U.S, refer to Southwest Energy Efficiency Project (2022).

Air Resources Board, 2018b),[8] and is valued even more among MUD residents.[9] Given the current low adoption rate of PEV in MUDs and MUD residents' reliance on public charging, it is natural to ask whether the current public charging stations are located in such a way as to encourage MUD residents to adopt PEVs. Moreover, without sufficient infrastructure access to encourage PEV adoption for MUD residents, the federal and state-level incentive programs will continue to disproportionately benefit wealthier households (Borenstein and Davis, 2016; Guo and Kontou, 2021).

Electric vehicle supply equipment (henceforth, charger) consists of all the equipment needed to deliver electrical energy from an electricity source to a PEV battery. There are three types of chargers. Level 1 chargers add 4-6 miles of range per hour, which is the equivalent of plugging into an everyday outlet and is typically used in home charging. Level 2 (L2) chargers add 10-60 miles of range per hour. Direct current fast charging (DCFC) is the fastest charging currently available. DC fast chargers add about 3 to 20 miles per minute, depending on the charger speed and state of charge of the battery. All PEVs can use Level 1 and Level 2 chargers.[10] However, most PHEVs, and some lower-range BEVs are not equipped with DCFC ports (California Governor's Office of Business and Economic Development, 2019). There is no consensus on whether L2 chargers or DC fast chargers are

[8] Li et al. (2017) empirically quantify that a 10% increase in the number of public charging stations would increase PEV sales by about 8%, and that subsidizing charging station deployment would be more than twice as effective as the federal income tax credit with the same expenditure in promoting PEV adoption, pointing to the effectiveness of public charging station deployment in promoting PEV adoption. By eliciting responses from survey respondents who did not design a PEV when asked to design a potential vehicle for the next purchase, Axsen and Kurani (2013) and Kurani et al. (2016) find that lack of public charging infrastructure is one of the most important factors that discourage U.S. households from considering a PEV as their next vehicle. California Air Resources Board (2018b) document that low-income residents, who often live in MUDs, report lack of public charging infrastructure as one of the many factors that held them back from owning/leasing a PEV when interviewed.

[9] The residential PEV owner survey under the 2019 California Vehicle Survey (National Renewable Energy Laboratory, 2019) asks a sample of PEV owners the following question: When deciding to purchase your PEV, how important was the availability of public charging in your decision? Respondents can give one of the five ratings for the importance of public charging, ranging from "not at all important" to "extremely important." Figure B.1 compares the distribution of ratings among PEV owners living in MUDs, and the distribution of ratings among PEV owners living in single-family housing. 49.12% of the MUD residents, compared to 24.87% of single-family housing residents, consider public charging as "very important" and "extremely important."

[10] Source: https://www.epa.gov/greenvehicles/plug-electric-vehicle-charging.

a better candidate to satisfy public charging needs from MUDs. For example, L2 chargers have lower charging fees, making it a more affordable choice for PEV owners living in MUDs. However, it takes DC fast chargers much less time to reach a full charge, making it a more convenient choice for MUD residents. Presumably, both types of chargers are crucial in meeting the charging demands from MUDs.

This paper studies whether the locations of public L2 chargers respond to the potential charging demands from MUDs in Los Angeles (LA) County. Specifically, it examines whether there is a positive correlation between the number of public L2 chargers and MUD density (measured by total square footage of MUD per capita) across census block groups (CBGs) in LA County, and uses the correlation coefficient to estimate the charger-to-PEV-ratio range for MUD residents, which is then compared to the ideal charger-to-PEV-ratio range if the 1.5 million ZEV goal is to be achieved by 2025 in California to evaluate the current progress of charger deployment. The focus is on L2 chargers for the following three reasons: First, L2 chargers are a cheaper choice for daily commute charging than DC fast chargers.[11] Second, MUD charging needs are currently more attended to by the deployment of L2 chargers,[12] than by the deployment of DC fast chargers, which gives more priority to interstate highways to ensure short charging time for long-distance trips.[13] Third and most importantly, a considerable percentage of PEVs owned by MUD residents are PHEVs,[14] most of which do

[11] According to a J.D. Power blog, the premium in cost between DC fast chargers and L2 chargers can be anywhere from 25% to 40% (source: https://www.jdpower.com/cars/shopping-guides/what-is-dc-fast-charging). A ChargePoint blog also mentions that fees are usually higher for DC fast charging than for L2 charging, and that given the extra cost, it is not cost-effective to use fast charging every day (source: https://www.chargepoint.com/blog/when-and-how-use-dc-fast-charging/). California DriveClean website expects drivers in California to pay 30 cents per kWh to charge on L2, and 40 cents per kWh for DC fast charging (source: https://driveclean.ca.gov/electric-car-charging).

[12] Numerous local governments have offered or are planning to offer L2 charging to accommodate the charging needs of residents without home charging. The two most prominent examples are Los Angeles city, with more than 430 public curbside L2 chargers deployed on street lights, and New York City, which is in the process of installing 100 public curbside L2 chargers. Source: https://electrek.co/2019/11/13/la-adds-hundreds-of-ev-chargers-to-streetlights-giving-renters-a-place-to-plug-in/. https://lalights.lacity.org/ourfuture.html. https://www.nyc.gov/html/dot/html/motorist/electric-vehicles.shtml#/curbside.

[13] Source: https://www.wsj.com/articles/ev-charging-network-will-target-interstate-highways-11644487200.

[14] For example, more than half of the PEVs registered to MUD residents in California in 2016 are PHEVs (California Air Resources Board, 2018a).

not work with DC fast chargers. The focus on L2 chargers is by no means an indication that L2 chargers are a superior candidate to DC fast chargers in terms of meeting MUD charging demands. Instead, the paper aims to develop a methodology of evaluating charger sufficiency using public L2 chargers as an example, and the methodology developed in the paper can be applied to evaluate the sufficiency of DC fast chargers in MUDs as well. A study of the charging infrastructure in LA County is appropriate given that LA County is a pioneer in PEV adoption and charging infrastructure investment.[15] The findings will have policy implications for both regions that have already had a relatively mature network of charging infrastructure like LA County and regions that are catching up and hoping to learn lessons from the front-runners. Unless otherwise specified, chargers refer to public L2 chargers throughout the paper.

The study asks whether the existence of chargers and the number of chargers are positively correlated with MUD density across CBGs in LA County in 2020, after controlling for local facilities, socio-economic characteristics and demographic characteristics. The findings show that on the extensive margin, the difference in the predicted probability of charger existence between a CBG with no MUDs and a CBG with an MUD density among the top 5% in LA County is about 20% to 30%. On the intensive margin, there is mixed evidence on whether MUD density is significantly positively correlated with the number of chargers. Even under the model specifications where MUD density is found to be significantly positively correlated with the number of chargers, the implied charger-to-PEV-ratio range for MUD residents is below the ideal charger-to-PEV-ratio range that needs to be realized by 2020 if the 1.5 million ZEV goal is to be achieved by 2025 in California. This suggests that the number of chargers is not responsive enough to the potential charging demands from MUDs in LA County.

[15]The City of Los Angeles is listed as ChargePoint's top 10 cities for electric vehicles. The Los Angeles-Long Beach-Anaheim metropolitan area was ranked fourth among the top 100 most EV-friendly metros by StorageCafé in 2021. Source:https://www.chargepoint.com/about/news/chargepoint-releases-list-top-10-cities-electric-vehicles, https://www.nytimes.com/2021/09/16/realestate/best-places-electric-cars.html.

The study has three major contributions. First, it contributes to the literature on EV charging infrastructure access for MUD residents (Williams and DeShazo, 2015; DeShazo et al., 2017; California Air Resources Board, 2018a; Lopez-Behar et al., 2019; Baldwin et al., 2020; Bryan and Aldridge, 2020). Some of these studies focus on specific areas such as financial viability (Williams and DeShazo, 2015), building codes (California Air Resources Board, 2018a) and installation design (Bryan and Aldridge, 2020) to increase charging access at MUDs. The others offer a comprehensive overview of challenges involved in the installation of charging infrastructure at MUDs and provide targeted policy recommendations (DeShazo et al., 2017; Baldwin et al., 2020). The present paper acknowledges the challenges mentioned in the literature and asks that, given the current lack of chargers at MUDs, whether nearby public chargers can satisfy the charging demand from MUDs. The findings show that the current stock of public L2 chargers is insufficient for the potential charging demand from MUDs, which offers another justification as to why on-site MUD charging is necessary in achieving nationwide zero-emission goals.

Second, the paper complements the literature on EV charging infrastructure projections (Wood et al., 2017, Bedir et al., 2018, Nicholas et al., 2019, Bui et al., 2021). This strand of literature forecasts the amount of EV charging infrastructure necessary to support a certain number of PEVs in the future by utilizing data or surveys on travel patterns, battery ranges, and etc. The scope of forecast can be national-level (Wood et al., 2017, Nicholas et al., 2019), state-level (Bedir et al., 2018) or city-level (Bui et al., 2021). By contrast, this paper estimates the current number of public L2 chargers per PEV at MUDs for LA County. The methodology of the paper can be applied to other types of chargers (e.g., DC Fast chargers), other housing types, and other regions to assess the status-quo prevalence of chargers. The estimated charger-to-PEV-ratio can be compared to the charger-to-PEV-ratio in the charging infrastructure projection literature for the corresponding region to assess the current progress of charger installation.

Third, the paper contributes to the literature studying factors correlated with access to charging infrastructure (Canepa et al., 2019; Brockway et al., 2021; Hsu and Fingerman, 2021; Khan et al., 2022). The paper is closest to Canepa et al. (2019), Hsu and Fingerman (2021) and Khan et al. (2022), which focus on the correlation between socio-demographic features and access to charging infrastructure, while briefly touching upon transportation and housing features.[16] This paper differs from these studies in the following three ways:

First, the three papers, two focusing on California and one focusing on New York City, all claim that there is inequitable access to EV charging infrastructure along socio-demographic dimensions. They compare access to charging infrastructure across geographic units with differing racial/ethnic compositions conditional on no or a few other socio-demographic and amenity variables. Therefore, the disparities in access to charging infrastructure found in the papers can either be caused by racial/ethnic compositions or omitted variables correlated with racial/ethnic compositions. The findings of inequitable access to EV charging infrastructure are interesting and have important policy implications, but identifying the omitted variables that are correlated with both access to EV charging and racial/ethnic compositions can further inform us of the channels through which we can increase charging access in disadvantaged/minority communities. Accordingly, this paper utilizes assessor parcel data and the American Community Survey to include information on land use, local facilities and demographics in each CBG in LA County. The paper thus has the richest set of neighborhood characteristics in this strand of literature, offering the most comprehensive perspective as to what neighborhood characteristics are associated with charging access.[17]

[16] Canepa et al. (2019) look at the difference in the number of charging stations and charging station density between disadvantaged communities (DACs) and non-DACs in California. Hsu and Fingerman (2021) study the relationship between public charger access and a) the percentage of MUDs among total housing units, b) access to freeway, c) income, and d) race and ethnicity across CBGs in California. Khan et al. (2022) explore the correlation between access to charging stations and income, poverty rate, race, ethnicity and access to highway across zip codes in New York City.

[17] Freeway access is not included as a neighborhood characteristic, as in Hsu and Fingerman (2021) and Khan et al. (2022), because of the focus on L2 charger access in this paper (their papers look at both L2 charger and DC fast charger access). Freeway access is more relevant to the discussion of DC fast charger access, as DC fast chargers are mostly located near freeways to ensure short charging time for long-distance trips. Source: http://www.westcoastgreenhighway.com/electrichighway.htm,

There are statistically significant racial and ethnic disparities in access to chargers when there are no other controls or when the only controls are income and MUD density, which is consistent with the findings of Hsu and Fingerman (2021) and Khan et al. (2022).[18] However, the percentage of the Black population and the percentage of the Hispanic population are no longer negatively correlated with access to charging after controlling for the entire set of neighborhood characteristics. The paper does not argue that there are no racial disparities in charging access, but argues that the correlation between other covariates and charging access can tell us the underlying reasons why there are fewer chargers in the minority neighborhoods.

This paper shows that access to charging is positively correlated with MUD density, commercial activities (office density, retail density, hotels and car dealers), amenities (transportation stations), number of government offices, schools and population, while access to charging is negatively correlated with average household size. Perhaps the reason why the minority neighborhoods have fewer chargers is because they have larger household sizes, lower retail densities and fewer amenities, and if so, the corresponding policy implication would be that increasing such amenities in minority neighborhoods would attract new chargers. Further studies are necessary to make formal claims, which could be a possible direction of future research.

The second difference is that the above papers assume that neighborhood characteristics are correlated with access to L2 chargers and access to DC fast chargers in the same way, with the exception of Canepa et al. (2019). In this paper, the focus is solely on L2 chargers, while the discussion suggests possible model modifications for future research on DC fast chargers. Since L2 chargers are more targeted towards daily commuting charging needs while DC fast chargers are more targeted towards charging needs during long-distance trips, access to the two types of chargers is expected to be correlated with neighborhood characteristics in different ways, which justifies separate analyses of the two types of chargers (as in Canepa

https://www.chargepoint.com/blog/california-highways-are-ready-ev-road-trips.

[18]Details are presented in Appendix B.1.

et al. (2019) and in this paper).

The third difference relative to the other papers is the use of more rigorous econometric methods to study the correlation between neighborhood characteristics and the number of chargers, which is the intensive margin of access to charging infrastructure. Canepa et al. (2019) simply compare the aggregate number of charging stations between disadvantaged communities (DACs) and non-DACs, Hsu and Fingerman (2021) do not study the intensive margin, and Khan et al. (2022) present correlations between the number of charging stations and each neighborhood characteristic separately. A prominent feature of the number of chargers across small geographic areas (e.g., census tracts used by Canepa et al. (2019), zip codes used by Khan et al. (2022) and CBGs used by Hsu and Fingerman (2021)) is that a considerable proportion of areas have zero chargers, under which circumstance the simple correlation estimates are biased. This paper utilizes the Tobit model and the zero-inflated negative binomial model to take into account this feature of the data.

One caveat of the study is that the estimated correlation between charger quantity and MUD density does not have a causal interpretation. That is, the estimated correlation cannot be interpreted as the change in the number of chargers *as a result of* an increase in MUD density. This is because MUD density is likely endogenous for the following two reasons. First, there could exist reverse causality, where charger quantity affects MUD density. The correlation between MUD density and the error term in this case will result in a biased estimate of the causal response of charger quantity to MUD density. Second, variables such as neighborhood income levels are correlated with both charger quantity and MUD density. Omitting these variables from the regression will again result in a biased causal estimate. Both sources of endogeneity can be overcome by the use of an instrument/instruments for MUD density. However, variables that affect MUD density typically affect charger quantity as well, making it a fruitless attempt to find a valid instrument that is correlated with MUD density in a CBG, but not correlated with unobserved shocks to charger quantity. It is worth noting that the first source of endogeneity, reverse causality from charger quantity to MUD density, is unlikely. Most MUDs in LA County were built before the emergence of PEVs and the need for chargers, making reverse causality inconsequential. To account for the first source of endogeneity, this paper adds a rich set of controls to the model to reduce the likelihood of omitted variable bias. Nonetheless, the paper aims to draw conclusions on charger sufficiency for MUDs from the correlation between charger quantity and MUD density, while refraining from making any causal interpretations.

The remainder of the paper is organized as follows. Section 2.2 describes the data. Section 2.3 outlines the models. Section 2.4 presents the results, discusses charger sufficiency and describes the various robustness checks that confirm the primary findings. Section 2.5 concludes.

2.2 Data

The data on EV charging stations are drawn from the Alternative Fuels Data Center (AFDC).[19] This dataset includes rich information on the location, the accuracy of the location and the opening date of the charging stations, the number of L2 and DC Fast chargers, respectively, at the charging stations, the types of facility at which the charging stations are located, and whether a charging station is private or publicly accessible. The paper focuses on charging stations that opened before September 1st, 2021,[20] that are located in LA County and publicly accessible, and whose location information is precise to the address. There are 6,948 L2 chargers and 519 DC Fast chargers in the dataset.

[19] Downloaded from https://afdc.energy.gov/data_download/ on Oct 1st, 2021.

[20] The data on MUDs are from LA County Assessor Parcel Data 2020, and therefore using charging stations that opened before September 1st, 2021 allows at least 8 months for charger deployment to respond to new MUD presence in 2020. This sample choice is motivated by the fact that the installment of a charger can take from three months to more than a year, depending on the complexity of site selection, permitting processes and grid connection (California Governor's Office of Business and Economic Development, 2019).

The geographic level at which to perform the analyses must be decided. Ideally, the smallest possible geographic unit (census block in this case) could be used to capture neighborhood characteristics at a local level. However, out of 109,355 census blocks (excluding Santa Catalina and San Clemente islands) in LA County, more than 98% have zero public L2 chargers. The overwhelming percentage of zeros results in little variation in the dependent variable, posing challenges to the estimation. Therefore, the analyses use the second smallest geographic unit, the census block group, or CBG. The study uses the boundaries of LA County CBGs defined by the 2010 Census, yielding 6,419 CBGs after excluding CBGs in the two islands, Santa Catalina and San Clemente.[21]

Figure 2.1a shows the distribution of public L2 chargers across CBGs in LA County. About 85% of the CBGs in LA County have zero chargers. CBGs with the most chargers are located in downtown LA, Long Beach, at and near LAX and in Santa Monica.

The data on MUDs are from LA County Assessor Parcel Data 2020,[22] which is publicly available on the LA County Open Data Portal, and includes information on all parcels built in or before 2020. The following non-vacant parcels are considered MUDs in this paper: condominiums with five or more floors, apartment buildings with five or more units, store and residential combination, and office and residential combination. MUD space is quantified by its total square footage.[23] For mixed-use properties (store and residential combination, and office and residential combination), only the square footage of the MUD

[21] The 2013 TIGER Geodatabase is used. 2013 is the first year that the TIGER Geodatabase is available. CBGs in the 2013 TIGER Geodatabase follows the same geographical boundaries as the 2010 Census. CBGs in Santa Catalina and San Clemente are excluded because they have significantly different topographical and neighborhood characteristics than the inland CBGs in LA County.

[22] Downloaded from https://data.lacounty.gov/Parcel-/Assessor-Parcel-Data-2020/42ne-gwcj on Sep 22nd, 2021.

[23] The other three potential ways to quantify MUDs are: by the number of MUD buildings, by the number of units in MUDs, and by the number of bedrooms in MUDs. Quantifying MUDs by the number of MUD buildings fails to account for differences in the building size and height, while quantifying MUDs by the number of units in MUDs fails to account for differences in the number of occupants in a unit. Quantifying MUDs by the number of bedrooms can potentially solve the above issue. However, the count of bedrooms in MUDs in the assessor records can be inaccurate, as suggested by an appraiser of LA County, leaving quantifying MUDs by their total square footage as the best option. Refer to Appendix B.2 for more details on why MUDs are quantified by their total square footage.

part of the parcel is counted.[24] Total square footage of MUD per capita (henceforth MUD density) is used to quantify the presense of MUDs in CBGs in later analyses. Figure 2.1b shows the distribution of MUD density across CBGs in LA County. In a first glance at Figure 2.1, the relationship between the number of L2 chargers and MUD density is not obvious. Therefore, a closer examination using rigorous econometric methods accounting for other factors affecting charger availability is necessary and will be carried out in Section 2.4.

The types of facilities a CBG has and the demographic composition of a CBG can also affect the number of chargers. AFDC data on the types of facilities where chargers are located in both California as a whole and LA County suggest that hotels, office buildings, government offices, schools, car dealers, transportation stations (Metro, Metrolink and Amtrak stations)[25] and shopping centers/malls are among the places where chargers are most likely to locate. Data on hotels, office buildings, car dealers and shopping centers/malls are from LA County Assessor Parcel Data 2020. Data on government offices, schools and transportation stations are from LA County Points of Interest.[26] Facilities are aggregated within each CBG using ArcGIS Pro. For facilities that are quantified by total square footage, the facility variable in a given CBG is measured by the total facility square footage per capita. For facilities that are quantified by counts, the facility variable in a given CBG is measured by the counts. Refer to Appendix B.2 for details on how the facilities are quantified.

CBG level socio-economic and demographic data are from the 2015-2019 ACS 5-year estimates. Information on the size of the CBG, the percentage of water in the CBG and whether a CBG is on the border of LA County is also collected. All the variables are listed and defined in Table 2.1, along with their summary statistics and data sources.

[24] Detailed information on mixed-use parcels is acquired from https://portal.assessor.lacounty.gov/. The website returns the square footage and design type of each subpart of a parcel. The design type is used to determine which subpart(s) of a mixed-use parcel is MUD. The square footage of the MUD subpart(s) is counted towards the total MUD square footage.

[25] Transportation stations are labeled as parking lots/garages in the AFDC dataset.

[26] Accessed at https://hub.arcgis.com/maps/c493f3d44e97482e90ce9355019b1349_185/about

2.3 Model

There is no consensus on how close a charger has to be for one to conveniently access it. A reasonable range for the maximum distance for a charger to be easily accessible is 0.25 miles to 2 miles.[27] The median span of CBGs in the sample is 0.6 miles, which means that a charger can still be accessible to and therefore be used by residents outside the CBG it is located in.[28] This indicates that whether there exist chargers in a CBG and the number of chargers in a CBG should depend not only on its own characteristics, but also on the characteristics of its neighboring CBGs. A simple OLS model[29] would be as follows:

$$y = X\gamma + WX\lambda + u, \tag{2.1}$$

where $y = (y_1, y_2, ..., y_n)'$ is a vector of the number of chargers in each of the n CBGs, or a vector of charger existence where $y_i = 1$ if there exist chargers in CBG i and $y_i = 0$ otherwise, u is a well-behaved normal error vector, X is the $n \times k$ matrix of k CBG characteristics, and W is the weight matrix, whose rows assign weights to the "neighbors" of a CBG to aggregate each characteristic of the neighboring CBGs into a single variable. W's diagonal elements are set at zero, since a CBG's own characteristics are already captured by the $X\gamma$ term in (2.1). The jth element of row i of W, equal to w_{ij}, is given by one of the following expressions when

[27] For more details, refer to Appendix B.3.

[28] The span is the longest distance between two points on the boundary of a CBG. Strictly speaking, the distance one travels for charging is the routing distance, following street networks, while the span of a CBG is computed in straight line distance. However, they are still comparable when the distance is short.

[29] A more general version of the model is the spatial autoregressive model (Anselin, 1988; Kelejian and Prucha, 1998; Kelejian and Prucha, 1999; Brueckner, 2003; Lee, 2004; Kelejian and Prucha, 2010; Drukker et al., 2013). It is used when there is strategic interaction among neighboring jurisdictions. The econometric representation is as follows:

$$y = \beta Wy + X\gamma + WX\lambda + u,$$

In addition to (2.1), the outcome variable in one jurisdiction depends on the outcome variable of the other neighboring jurisdictions. The reason why the paper does not adopt this more generalized model is because CBGs are created for statistical purposes and therefore are not decision-making entites that interact with each other strategically.

$i \neq j$: $w_{ij} = 1/d_{ij}$ (the inverse-distance weight), $w_{ij} = 1/d_{ij}^2$ (the inverse-distance-squared weight), $w_{ij} = P_j/d_{ij}$ (the population-weighted inverse-distance weight), and $w_{ij} = P_j/d_{ij}^2$ (the population-weighted inverse-distance-squared weight).

These weight patterns are widely used in the spatial econometrics literature (Anselin et al., 1997; Brueckner, 1998; Shafran, 2008; Brueckner, 2022).[30] The inverse-distance weight matrix sets w_{ij} at $1/d_{ij}$, the inverse of the distance between CBG i and CBG j.[31] Under this weight matrix, CBGs near CBG i have a greater influence than those farther away. Since residents do not use chargers that are too far away, the number of chargers of a given CBG should not be affected by the characteristics of CBGs beyond a certain distance. This suggests that we should use a truncated inverse-distance weight matrix, where $w_{ij} = 0$ if d_{ij} is larger than a distance threshold. However, truncated weights are not used in the baseline analyses for the following reasons: First, as discussed in Appendix B.3, the distance threshold is unclear. Second, there exist CBGs whose nearest neighbor is quite far away. Truncating the weight matrix will turn these CBGs into "islands" (without neighbors), which complicates the analysis. Instead, a better candidate would be the inverse-distance-squared weight matrix that gives smaller weights to distant CBGs than under the inverse-distance matrix. Under the inverse-distance-squared matrix, the weight of more distant CBGs are close to zero, yet there is no concern of CBGs turning into "islands." Robustness checks using weight matrices truncated at one-mile distance are reported in Section 2.4.4.

The first two weight matrices listed above represent population-unweighted schemes, while the remaining two weight matrices give population-weighted schemes. Under the population-weighted scheme, weights depend on both the distance to and the population

[30]Another widely used weight matrix is contiguity matrix, where $w_{ij} = 1$ if CBG j shares a border with CBG i, and $w_{ij} = 0$ otherwise. I do not use the contiguity matrix in the paper. In densely-populated areas, the sizes of CBGs are fairly small such that two CBGs are less than 0.5 to 1 mile apart even when they do not share a border. Using the contiguity matrix can unwisely exclude neighboring CBGs that could have nontrivial effects on the number of chargers of the CBG in question.

[31]The distance between CBG i and CBG j is defined as the distance between the centroids of shape polygons of CBG i and CBG j, generated by the Stata command *spshape2dta* (StataCorp, 2019). Distance is measured in miles.

of the neighboring CBGs. Since the demand for chargers comes from local residents, if a CBG has a small population, then its demand for charging would be small, thus playing a negligible role in determining the number of chargers in a nearby CBG. For this reason, a population-weighted scheme is preferable to a population-unweighted scheme. Therefore, the population-weighted inverse-distance-squared matrix is used in the baseline analyses, while results under the other matrices are presented in the online appendix. Each weight matrix is normalized so that its largest eigenvalue is 1, which is referred to as spectral normalization and is commonly used in the spatial econometrics literature.

There are missing values for covariates, as shown in Table 2.1. The majority of missing values occur in either CBGs that have zero population or zero occupied housing units. It is common practice in non-spatial models to exclude observations with missing covariates from the estimation sample. While omitting certain CBGs in this spatial model means that the spillovers from them to the neighbors are no longer being included, it makes sense to drop CBGs with zero population, because CBGs with zero population should theoretically have no spillovers to their neighbors as they have no demand for chargers. 29 CBGs with zero population are dropped from the sample along with 9 CBGs that have zero occupied housing units.[32] After these deletions, 8 additional CBGs having missing values for other variables.[33] After deletion of these CBGs, the final sample has 6,373 CBGs.

When y is a vector of indicator variables of charger existence, (2.1) is a linear probability model (LPM) that estimates the extensive margin of charger supply. LPM has the apparent disadvantage of predicted probabilities outside the $[0, 1]$ interval for values of independent variables that are far away from the averages in the sample. Therefore, a probit model is also used to estimate the correlation between the CBG characteristics and the probability

[32]There are 9 CBGs with zero occupied housing units but positive population. These are group quarters including college residence halls, correctional facilities, and etc. It is reasonable to assume that correctional facilities have no demand for chargers, but chargers often locate at colleges. Luckily, such CBGs occupy a small percentage in the sample.

[33]These CBGs have missing values in either median age, average household size, per capita income or commute time.

of charger existence, which is presented as follows:

$$P(y = 1|X, W) = \Phi(X\gamma + WX\lambda), \tag{2.2}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function.

When y is a vector of the number of chargers, (2.1) is not ideal either. About 85% of the CBGs in LA County have zero chargers. The number of chargers also has a dispersed distribution: 28 CBGs have more than 50 chargers, while 6 CBGs have more than 100 chargers. OLS models do not work well when the outcome variables have non-negative values only and excessive zeros, which calls for alternative models in the estimation.

The first candidate is the Tobit model. The Tobit model expresses the observed response y in terms of an underlying latent variable y^* (Wooldridge, 2013):

$$y^* = X\delta + WX\eta + u, \tag{2.3}$$

$$y = max(0, y^*). \tag{2.4}$$

Since the number of chargers in CBGs has good dispersion where large counts of chargers exist, it is reasonable to treat it as a continuous variable. Tobit model not only takes into account the fact that the outcome variable is left-censored at 0, but it also allows estimation of the effects of covariates on the intensive margin (conditional on a CBG having a charger/chargers, how is a covariate associated with how many chargers a CBG has)(McDonald and Moffitt, 1980).

The second candidate is zero-inflated negative binomial (ZINB) model, which takes into account the fact that the outcome variable is a count variable. Negative binomial is appropriate because the dependent variable is overdispersed, and zero-inflated is appropriate because the outcome variable has about 85% of zeros. Under the ZINB model, zero counts occur in two ways: as a realization of a binary process (a logit model in this paper) and as a realization of the count process when the binary random variable takes value 1 (Cameron and Trivedi, 2005). This paper refers to the zero counts generated by the binary process as *excessive zeros*. The binary process takes value 0 with probability $f_1(0)$, and takes value 1 with probability $f_1(1)$. If the binary process takes value 0, then $y = 0$. If the binary process takes value 1, then y takes count values 0, 1, 2, . . . from a negative binomial density $f_2(\cdot)$. The density function of y, $g(y)$, is as follows:

$$g(y) = \begin{cases} f_1(0) + (1 - f_1(0))f_2(0) & \text{if } y = 0, \\ (1 - f_1(0))f_2(y) & \text{if } y \geq 1, \end{cases} \tag{2.5}$$

To put the model into context, the binary process determines whether a CBG needs chargers, and if a CBG needs chargers, the count process (the negative binomial model) determines how many chargers a CBG installs, which could be 0, 1, 2, ... Zero chargers can be the result of the following two scenarios: First, a CBG does not need chargers, in which case the paper refers to the zero chargers as excessive zeros. Second, a CBG needs chargers, but decides to install zero chargers. The binary process and the negative binomial model can be determined by different covariates. In this paper, the binary process is assumed to be determined by MUD density and the percentage of owner-occupied housing, while the negative binomial model includes all the covariates. The intuition is as follows: MUD density and the percentage of owner-occupied housing determine whether a CBG needs chargers. Since a homeowner can typically charge his/her EV at home, a high percentage of owner-occupied housing means less demand for outside chargers. Similarly, a low MUD density means less charging demand from MUD residents. If there is little charging demand, a CBG will have zero chargers, but for CBGs that do have a demand for chargers, the number of chargers is determined by all the covariates. Note that in the ZINB model, a CBG that has charging demand can still end up with zero chargers, depending on the CBG characteristics. Hypothetically, it could be because the CBG does not have enough facilities where chargers can locate, or because the CBG covers low-income neighborhoods where EV infrastructure companies are unwilling to invest. One may argue that MUD density and the percentage of owner-occupied housing are not the only factors that determine whether a CBG has demand for chargers. Robustness checks that include all the CBG characteristics in the first-stage binary process are presented in the online appendix.

2.4 Empirical results

This section presents the main results. The population-weighted inverse-distance-squared matrix is used across all model specifications, for reasons discussed in Section 2.3. Results under the other three weight matrices are qualitatively and quantitatively similar, and are presented in the online appendix.

2.4.1 The extensive margin of charger supply

Table 2.2 presents results under the LPM and the probit model. Columns (1)-(3) report coefficients from the LPM, while columns (4)-(6) report average marginal effects derived from the probit model. Columns (1) and (4) do not include characteristics of neighboring CBGs, while columns (2) and (5) do. Columns (3) and (6) exclude CBGs that are on the border of LA County, as the number of chargers in these CBGs also depends on the characteristics of other counties, whose data are not included in the sample.[34] Results suggest that a 1,000 square feet per capita increase in MUD density is associated with about 20% to 30% increase in the probability that a CBG has chargers.[35] Note that only 5.16% of the CBGs have MUD

[34] A one-mile buffer zone is drawn at the border of LA County, and any CBGs that intersect this buffer are considered border CBGs. The weighted neighboring characteristics are computed before dropping border CBGs. This way, spillovers from the border CBGs to the rest CBGs are still taken into account.

[35] Another difference of the LPM and the probit model is that in the LPM the effects of MUD density on charger existence are assumed to be constant across levels of MUD density, while in the probit model the effects of MUD density on charger existence vary with MUD density. Columns (4)-(6) present the *overall*

density above 1,000 square feet per capita. Therefore, we can interpret the above results in the following way: the difference in the predicted probability of charger existence between a CBG with no MUDs and a CBG with an MUD density among the top 5% in LA County is about 20% to 30%.

Among own CBG characteristics, facilities (including offices, retail, government offices, car dealers, transportation stations, schools and hotels) and population are positively correlated with charger existence in a CBG, while CBGs with a larger median age (older population) and a larger household size have a smaller probability of having chargers, other things being equal. A larger percentage of owner-occupied housing is associated with a smaller probability of charger existence. This finding is consistent with the fact that homeowners can charge EVs at home and therefore do not rely on public chargers, but the correlation is unfortunately not significant. Among neighboring CBG characteristics, office density, number of transportation stations, population, percentage of population with a bachelor's degree are positively correlated with charger existence, while the percentage of the Hispanic population and the average household size are negatively correlated with charger existence. Unexpectedly, among neighboring CBG characteristics, the percentage of owner-occupied housing and the percentage of population that are not White, Black or Asian are positively correlated with charger existence, while the percentage of the Asian population is negatively correlated with charger existence, which this paper does not have an explanation for. However, whether the coefficient of a certain neighbor characteristic is significantly different from zero or even the sign of the coefficient is not consistent across different weight matrices. This is expected as the form of the weight matrices affects the estimation of the coefficients of

average marginal effects of MUD density on the existence of chargers. When looking at the average marginal effects of MUD density *conditional on* the value of MUD density, the paper finds that the effects of MUD density on existence of chargers increase with MUD density. For example, under the specification of column (5), a one square foot per capita increase in MUD density is associated with 0.02% increase in the probability that a CBG has chargers when a CBG has no MUDs, while a one square foot per capita increase in MUD density is associated with 0.04% increase in the probability that a CBG has chargers for a CBG with an MUD density of 1,000 square feet per capita. Hsu and Fingerman (2021) also allow the effects of socio-demographic characteristics on charger existence to be nonlinear across levels of MUD density.

the neighbor characteristics more than the characteristics of the CBG in question. For this reason, and for the fact that most of the coefficients of neighboring characteristics are insignificant, this paper does not present the coefficients of neighboring characteristics in the other tables.

2.4.2 The intensive margin of charger supply

In this section, the paper studies the correlation between CBG characteristics and the number of chargers in a CBG. Table 2.3 compares estimation results under the OLS model and the Tobit model. Columns (1)-(3) present results under the OLS specification, while columns (4)-(6) present results under the Tobit specification. The Tobit specification is preferred, as it takes into account the fact that a large percentage of CBGs have zero chargers. The OLS estimation underestimates the positive correlation between MUD density and the number of chargers compared to the Tobit estimation, as the excessive zeros in the dependent variable downward bias the estimates. This paper will thus focus on the results from the Tobit model in the future discussions. The coefficients in columns (4)-(6) are estimators of δ in equation (2.3), which measure the partial effects of covariates on the latent outcome variable: $\delta_j = \partial E(y_i^*|X, W)/\partial x_{ij}$. However, it is more interesting to look at the effects of CBG characteristics on the intensive margin ($\partial E(y_i|y_i > 0, X, W)/\partial x_{ij}$) of the outcome variable instead of the latent outcome variable.

Table 2.4 presents the average marginal effects of CBG characteristics on the number of chargers in a CBG. Conditional on the presence of chargers, a 1,000 square feet per capita increase in MUD density is associated with an additional 1.28 chargers. Unsurprisingly, CBG characteristics that are significantly correlated with charger existence in Table 2.2 are also significantly correlated with the number of chargers in a CBG. Among own CBG characteristics, facilities (including offices, retail, government offices, car dealers, transportation stations, schools and hotels) and population are positively correlated with the number of chargers in a CBG, while CBGs with a larger median age (older population), a larger household size and a larger share of high school population have fewer chargers, other things being equal. The percentage of owner-occupied housing is negatively correlated with the number of chargers, though this correlation is not significant. This finding is again consistent with the fact that homeowners can charge EVs at home and therefore do not rely on public chargers.

2.4.2.1 Is this charger response sufficient?

In this section, the paper uses the coefficient of MUD density in Table 2.4 to assess the sufficiency of chargers.[36]

The 2019 California Vehicle Survey (National Renewable Energy Laboratory, 2019) suggests that the PEV ownership rate among MUD residents ranges from 0.04 vehicles per houshold to 0.067 vehicles per household. In addition, a 1,000 square feet per capita increase in MUD density is approximately equivalent to a 645-700 increase in MUD units. With the assumption that one MUD unit holds one household, and that the PEV ownership rate among MUD residents in California is representative of the PEV ownership rate among MUD residents in LA County, a 1,000 square feet per capita increase in MUD density is associated with a 25.8 (0.04×645) to 46.9 (0.067×700) unit increase in PEVs. Using the estimated coefficient of 1.28 from Section 2.4.2, a CBG with 25.8 to 46.9 more PEVs in MUD households has 1.28 more chargers on average. Therefore, the incremental public-L2-charger-to-PEV-ratio for MUD residents (henceforth incremental MUD charger-to-PEV-ratio) ranges from 0.027 (1.28/46.9) to 0.05 (1.28/25.8).

To determine whether the number of chargers is sufficiently responsive to the poten-

[36]According to coefficients of neighbor characteristics that are not presented in Table 2.4, neighbor MUD density is not significantly correlated with the number of chargers conditional on the presence of chargers, which means the number of chargers in a CBG only responds to its own MUD density. Therefore, only the coefficient of "own" MUD density is used in the back-of-the-envelope calculation.

tial charging demands from MUDs, ideally one would compare the above incremental MUD charger-to-PEV-ratio to the minimum acceptable average MUD charger-to-PEV-ratio. Unfortunately, to the best of my knowledge, there has been no such specific criterion for meeting MUD charging needs stated in the literature. However, there exist criteria for overall charger-to-PEV-ratio in the charging infrastructure projection literature. Bedir et al. (2018) forecast the number of PEVs and the number of L2 chargers required to achieve the 1.5 million ZEV target in the former Governor Edmund G. Brown Jr.'s Executive Order B-16-2012 for two time points, by 2020, and by 2025.[37] Since the charger-to-PEV-ratio range for MUD residents is estimated using data on charging stations that opened before September 2021 and data on MUDs built in or before 2020, it will be compared to the target charger-to-PEV-ratio forecasted for 2020.[38] According to the report, to achieve the 1.5 million ZEV goal, by 2020, the target charger-to-PEV-ratio should range from 0.082 to 0.109.[39] The target MUD charger-to-PEV-ratio should be even larger than the above target overall charger-to-PEV-ratio, as a large percentage of PEV owners are assumed to be single-family residents in the report, who rely less on public L2 chargers than MUD residents. The incremental charger-to-PEV-ratio range for MUD residents derived in this paper (0.027 to 0.05) is lower than the ideal charger-to-PEV-ratio range if the 1.5 million ZEV goal is to be achieved (0.082 to 0.109).[40] Therefore, the implied conclusion is that the number of chargers is not responsive

[37] In the report, four major categories of assumptions are made to forecast the number of PEVs and the number of chargers needed to support these PEVs, following Wood et al. (2017). Wood et al. (2017) makes charging infrastructure projections on a national level, while Bedir et al. (2018) focus on charging infrastructure projections in California, which is closer to the context of this study. First, the report makes assumptions on the electric range of BEVs and PHEVs based on California Air Resource Board's technical review. Second, it makes assumptions on charger technology. Third, it utilizes the 2010-2012 California Household Travel Survey to obtain 24-hour daily travel profiles representative of mainstream driving behaviors at the county level. Fourth, it makes assumptions on the shares of BEVs and PHEVs.

[38] As mentioned in footnote 20, since the installment of a charger can take from three months to more than a year, using charging stations that opened before September 1st, 2021 allows at least 8 months for charger deployment to respond to new MUD presence in 2020. The charger-to-PEV-ratio range estimated by using data on charging stations that opened before September 2021 and data on MUDs built in or before 2020 can be considered as the charger-to-PEV-ratio range for MUD residents in 2020.

[39] The range of charger-to-PEV-ratio is computed based on the demand for L2 destination (workplace and public) chargers by 2020 presented in "Table ES.1: Projections for Statewide PEV Charger Demand" in Bedir et al. (2018). Home chargers are not counted towards the charger-to-PEV-ratio.

[40] Technically, the MUD charger-to-PEV-ratio derived in this paper is an incremental ratio, meaning that $\frac{\partial(\text{number of chargers})}{\partial(\text{number of PEVs in MUDs})} \in [0.027, 0.05]$. The charger-to-PEV-ratio in the literature is an average ratio,

enough to the potential charging demands from MUDs in LA County. More details on this back-of-the-envelope calculation are presented in Appendix B.4.

Comparing the current charger-to-PEV-ratio range for MUD residents to the ideal charger-to-PEV-ratio range if a certain PEV deployment goal was met allows the policymakers to gauge the gap between the status-quo charger sufficiency for MUD charging needs and the ideal charger sufficiency for MUDs under a certain policy target, and to adjust the breadth and intensity of charger deployment dedicated to MUDs accordingly. Another equally important and related question is, how to encourage PEV adoption in MUDs so that the adoption level could keep up with the 2025 ZEV goal?[41] This paper only focuses on increasing public charging infrastructure as one way to encourage PEV adoption among MUD residents,[42] but other factors that can encourage PEV adoption in MUDs could be a promising future research subject.[43]

2.4.2.2 The charger sufficiency map of LA County

Figure 2.2 visualizes MUD charger-to-PEV-ratios across CBGs in LA County. CBGs with a darker color have more severe charger insufficiency for MUD residents. The MUD charger-to-PEV-ratio in each CBG is derived by dividing the estimated number of PEVs belonging to MUD households by the number of public L2 chargers. The estimated number of PEVs belonging to MUD households is computed by multiplying the number of MUD units by the PEV ownership rate among MUD residents.[44] 28.2% of the CBGs do not have MUD

meaning that $\frac{\text{number of chargers}}{\text{number of PEVs}} \in [0.082, 0.109]$. However, the incremental MUD charger-to-PEV-ratio can still be compared to the average charger-to-PEV-ratio in the literature to evaluate charger sufficiency.

[41]The author would like to thank one of the anonymous referees for this insight.

[42]Refer to footnote 8 and footnote 9 for evidence on the positive effects of increasing public charging infrastructure on PEV adoption among MUDs.

[43]Existing literature that evaluates policies on promoting PEV adoption has yet to focus on PEV adoption in MUDs specifically (Jenn et al., 2018; Rapson and Muehlegger, 2021; Rapson and Muehlegger, 2018).

[44]In the absence of data on the number of MUD households in each CBG, the paper uses the number of MUD units as a proxy. The PEV ownership rate used is 0.054 vehicles per household, which is the average of the two measures of PEV ownership rates among MUD residents derived in Appendix B.4.1.

units. Among the 4,609 CBGs that have MUDs, 3,825 of them have no public L2 chargers, 142 of them have an MUD charger-to-PEV-ratio below 75% of the lower end of the target charger-to-PEV-ratio (0.06), and 53 of them have an MUD charger-to-PEV-ratio below the lower end of the target charger-to-PEV-ratio (0.08). 589 CBGs have sufficient chargers for MUD residents, with an MUD charger-to-PEV-ratio above the lower end of the target charger-to-PEV-ratio (0.08).

2.4.3 The ZINB model

The Tobit model takes into account that the dependent variable, the number of chargers, is left-truncated at 0, but still treats the variable as continuous. Regression results under the count data models are presented in Table 2.5. Columns (1)-(2) present the results from the negative binomial (NB) regression model. Under the NB model, MUD density is positively correlated with the number of chargers, while the percentage of owner-occupied housing is negatively correlated with the number of chargers. These results are consistent with our expectations, though neither of the coefficients are significant. Though the NB model takes into account the dependent variable's count data nature, it ignores the fact that more than 85% of the dependent variables are zero.

The ZINB model is preferred to the NB model as it explains the occurrence of excessive zeros in the number of chargers. Columns (3)-(6) present the results from the ZINB model. Columns (3)-(4) use MUD density and the percentage of owner-occupied housing in a CBG to predict the occurrence that the number of chargers in that CBG is excessive zero (see the bottom of the table). As mentioned in Section 2.3, one can understand excessive zero as the zero number of chargers resulting from little demand for chargers from the residents in the area. For a one square foot per capita increase in MUD density, the odds ratio (the probability of a CBG not needing chargers over the probability of a CBG needing chargers) decreases by 1.2%.[45] To put this number into perspective, the median MUD density per capita in our sample is 37 square feet. Moving from zero MUD density to median MUD density is associated with a decrease of 36.4% in the odds ratio. For a 1% increase in the percentage of owner-occupied housing, the odds ratio increases by 1%.[46] Both the effects are statistically significant.

Columns (5)-(6) use MUD density and the percentage of owner-occupied housing in both the CBG and the neighboring CBGs to predict the occurrence of excessive zeros (zero number of chargers resulting from little demand for chargers). After allowing MUD density and the percentage of owner-occupied housing in the neighboring CBGs to affect the odds of excessive zeros, a one square foot per capita increase in MUD density is associated with a decrease of 0.7% in the odds ratio.[47] For a 1% increase in the percentage of owner-occupied housing, the odds ratio increases by 1.4%.[48] The neighboring MUD density decreases the odds ratio, but this effect is not significant. The neighboring percentage of owner-occupied housing significantly increases the odds ratio.[49]

For CBGs that do have a demand for chargers, office density, retail density, number of government offices, car dealers, transportation stations, schools, hotels, and population are positively associated with the number of chargers across all specifications, as seen in the upper part of the table. Household size is negatively associated with the number of chargers across all specifications. MUD density and homeownership rate have no significant effect

[45] By taking the more conservative coefficient of MUD density in the first stage logit model, -12.225, dividing it by 1000 (because the MUD density is in 1,000 square feet per capita), and exponentiating it, we get 0.988, which is a 1.2% decrease.

[46] By taking the more conservative coefficient of the percentage of owner-occupied housing in the first stage logit model, 1.039, dividing it by 100, and exponentiating it, we get 1.01, which is a 1% increase.

[47] By taking the more conservative coefficient of MUD density in the first stage logit model, -6.955, dividing it by 1000 (because the MUD density is in 1,000 square feet per capita), and exponentiating it, we get 0.993, which is a 0.7% decrease.

[48] By taking the more conservative coefficient of the percentage of owner-occupied housing in the first stage logit model, 1.388, dividing it by 100, and exponentiating it, we get 1.014, which is a 1.4% increase.

[49] It is hard to interpret the magnitude of the coefficient of neighboring percentage of owner-occupied housing since the variable is a weighted average of homeownership rates in neighboring CBGs using a spectrally normalized weight matrix.

on the number of chargers conditional on their effects on whether a CBG has a demand for chargers.

One drawback of the ZINB model is that the excessive zero element in the model makes the interpretation more abstract. Moreover, it is not possible to compare the effects of the variables contributing to excessive zeros to the effects they have in the other models. However, it is still a nice exercise to understand what factors determine whether a CBG has a demand for chargers, and if a CBG has a demand for chargers, what factors are associated with the number of chargers.

To sum up, MUD density is found to be positively associated with the likelihood that a CBG has chargers. The difference in the predicted probability of charger existence between a CBG with no MUDs and a CBG with an MUD density among the top 5% in LA County is about 20% to 30%. On the intensive margin, there is mixed evidence on whether MUD density is significantly positively correlated with the number of chargers. The findings under the ZINB model suggest that MUD density's positive correlation with the number of chargers is not significant given that a CBG has demand for chargers. Under the Tobit model, a CBG with an MUD density among the top 5% in LA County is estimated to have 1.3 more chargers than a CBG with no MUDs, conditional on the presence of chargers. This estimate is statistically significant. However, the implied charger-to-PEV-ratio range for MUD residents is below the ideal charger-to-PEV-ratio range if the 1.5 million ZEV goal is to be achieved by 2025 in California. This suggests that the number of chargers is not responsive enough to the potential charging demands from MUDs in LA County.

2.4.4 Robustness Checks

One may worry that measuring neighbor characteristics by aggregating the characteristics of all the CBGs in LA County other than the CBG in question brings noise into the estimation, as the number of chargers in a CBG should only depend on the characteristics of the CBGs that are close enough. To address this concern, a truncated population-weighted inverse-distance-squared matrix is used, where w_{ij}, the jth element of row i, is given by the following expression:

$$w_{ij} = \begin{cases} P_j/d_{ij}^2 & \text{if } d_{ij} < 1 \text{ and } i \neq j, \\ 0 & \text{otherwise ,} \end{cases} \tag{2.6}$$

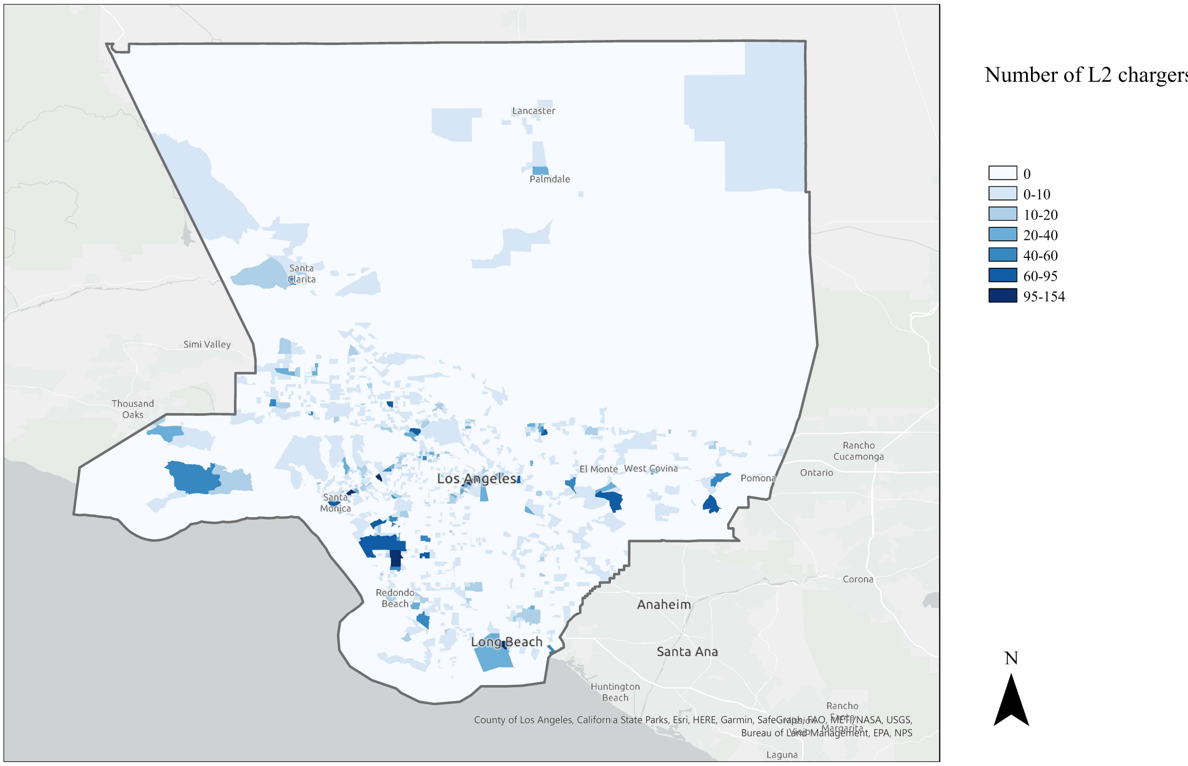
Since distance is measured in miles, the influence of CBG j on chargers in CBG i is zero when CBG i and CBG j are one mile or more apart. After computing the aggregate neighbor characteristics, the CBGs that are "islands" (CBGs with no neighbors after distance truncation) are dropped from the sample. The "islands" dropped have a larger size than the rest of the CBGs, which is why they have fewer neighbors within one mile. Since CBG size is negatively correlated with CBG population density, this subsample consists of more-populated CBGs, which are of more interest from a policy point of view. Border CBGs are also dropped. The results are presented in Table 2.6. The coefficients of MUD density are qualitatively and quantitatively similar to the coefficients in the previous specifications.

To ensure that the main findings are not driven by the choice of the population-weighted inverse-distance-squared matrix, the main analyses are repeated under the other three weight matrices in the online appendix. The results are qualitatively and quantitatively similar.

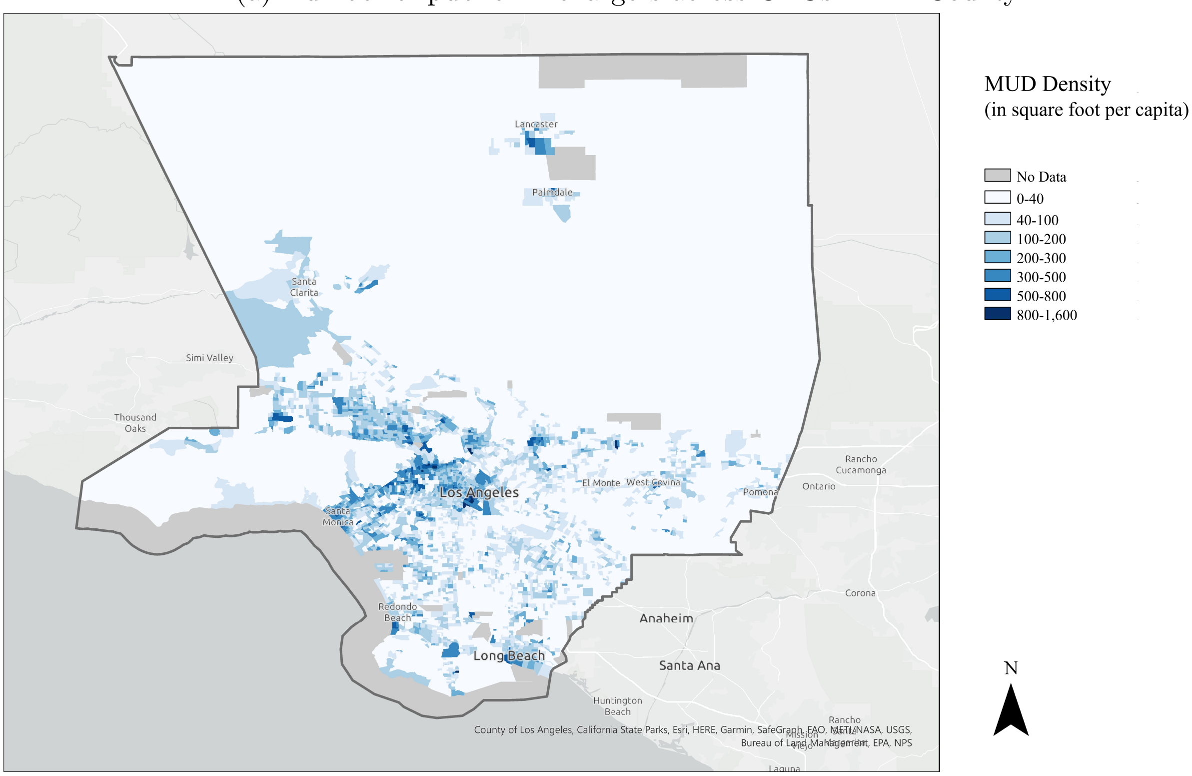
2.5 Conclusion

This paper studies whether the locations of public L2 chargers respond to the potential charging demands from MUDs in Los Angeles (LA) County. Specifically, it asks whether the existence of chargers and the number of chargers are positively correlated with MUD density across CBGs in LA County in 2020, after controlling for local facilities, socio-economic characteristics and demographic characteristics. The results show that high MUD-density CBGs and low MUD-density CBGs differ in terms of the probability of charger existence, but do not differ much in terms of the number of chargers. The charger-to-PEV-ratio range for MUD residents in LA County derived in this paper is below the ideal charger-to-PEV-ratio range in 2020 if the 1.5 million ZEV goal is to be achieved by 2025 in California. A direct policy implication is that more charging infrastructure should be made available to MUD residents. This includes public L2 chargers near MUDs, which are discussed in this paper, and onsite MUD charging, which is still at an early stage.

This paper focuses entirely on L2 chargers, but the methodology can be applied to other types of chargers, other housing types, and other regions to evaluate the adequacy of the charging infrastructure. An immediate extension of the paper would be to look at the correlation between neighborhood characteristics and the number/availability of DC fast chargers. Some changes in the model are necessary to account for the fact that DC fast chargers are mainly targeted for the fast charging needs of long-distance travelers. For example, one may include freeway access and long distance passenger travel volume as neighborhood characteristics as well.



(a) Number of public L2 chargers across CBGs in LA County



(b) MUD density across CBGs in LA County

Figure 2.1: Spatial patterns of the number of L2 chargers and MUD density

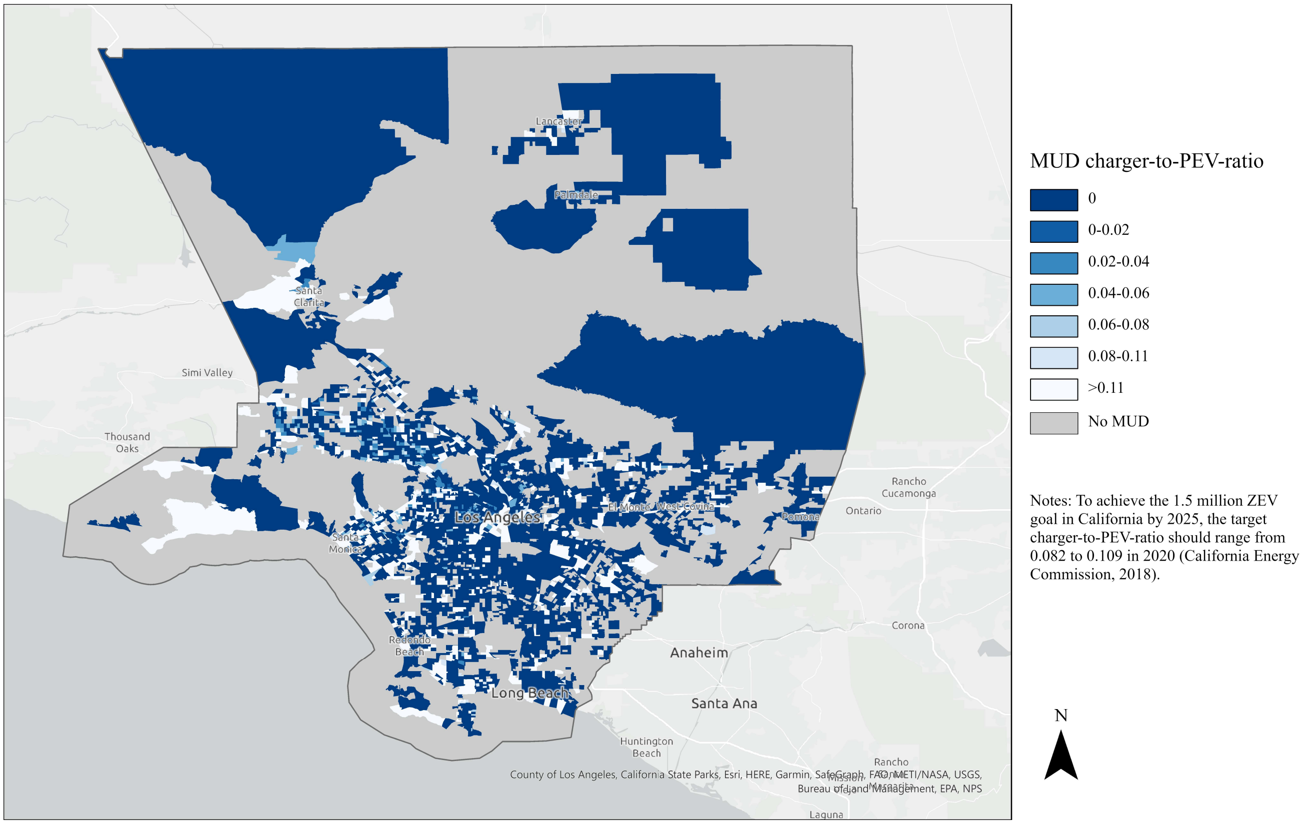


Figure 2.2: MUD charger-to-PEV-ratio across CBGs in LA County

Table 2.1: Summary statistics

Variable	N	Mean	SD	Min	25^{th} pctl	Median	75^{th} pctl	Max	Definition	Sources
L2	6,419	1.082	6.263	0	0	0	0	154	Number of L2 chargers	A
MUD	6,390	0.087	0.124	0	0	0.037	0.126	1.568	MUD density*	B
Office	6,390	0.033	0.222	0	0	0	0.008	8.601	Office density*	B
Retail	6,390	0.021	0.123	0	0	0	0.004	5.373	Retail density*	B
Government	6,419	0.039	0.346	0	0	0	0	20	Number of government offices	B
Car Dealer	6,419	0.119	0.629	0	0	0	0	19	Number of car dealers	B
Transportation	6,419	0.024	0.172	0	0	0	0	3	Number of transportation stations	B
School	6,419	0.108	0.401	0	0	0	0	4	Number of schools	B
Hotel	6,419	0.062	0.445	0	0	0	0	16	Number of hotels	B
Population†	6,419	1.570	0.841	0	0.996	1.403	1.977	12.1	Population	C
White‡	6,390	0.524	0.224	0	0.354	0.526	0.707	1	White	C
Black‡	6,390	0.084	0.144	0	0.002	0.029	0.094	1	Black	C
Asian‡	6,390	0.141	0.169	0	0.023	0.084	0.188	0.959	Asian	C
Other Races‡	6,390	0.251	0.177	0	0.103	0.213	0.376	1	Other races	C
Hispanic‡	6,390	0.467	0.302	0	0.19	0.441	0.736	1	Hispanic or Latino	C
Female‡	6,390	0.508	0.061	0	0.474	0.508	0.542	1	Female	C
Age	6,387	38.31	7.941	10	32.5	37.4	43.5	75.7	Median age	C
Owner-Occupied	6,381	0.499	0.287	0	0.256	0.496	0.756	1	Share of owner-occupied housing	C
Household Size	6,377	3.137	0.851	1.05	2.53	3.06	3.69	8.18	Average household size	C
Below High‡	6,388	0.210	0.173	0	0.062	0.167	0.335	0.885	No high school diploma	C
High School‡	6,388	0.206	0.099	0	0.132	0.206	0.273	1	With high school diploma	C
Some College‡	6,388	0.261	0.097	0	0.192	0.253	0.323	1	Some college but no Bachelor's degree	C
Bachelor‡	6,388	0.209	0.133	0	0.097	0.192	0.308	0.658	Bachelor's degree	C
Above Bachelor‡	6,388	0.114	0.111	0	0.027	0.078	0.172	0.642	Degrees above Bachelor's degree	C
Commute Time	6,381	31.20	4.519	5	28.23	31.13	34.09	60	Commute time§	C

(Continued)

Variable	N	Mean	SD	Min	25^{th} pctl	Median	75^{th} pctl	Max	Definition	Sources
Income	6,385	36.54	25.08	1.6	19.43	28.99	45.47	288.96	Per capita income in the past 12 months¶	C
CBG Area‖	6,419	18.22	199.4	0.271	2.216	3.442	5.845	8676.8	CBG Area	D
Water	6,419	0.009	0.062	0	0	0	0	1	Share of water area	D
Border	6,419	0.034	0.180	0	0	0	0	1	Located at the border of LA County**	D

Notes: The "N" column reports the number of CBGs that have nonmissing values for each variable. Under the "Sources" column, "A" stands for Alternative Fuels Data Center, "B" stands for LA County Assessor Parcel Data 2020, "C" stands for 2015-2019 ACS 5-year estimates, and "D" stands for 2013 TIGER Geodatabase.

*In 1,000 square feet per capita.

†In 1,000 persons.

‡The variables are the population shares of a certain demographic or socio-economic group.

§The average commute time of each CBG is computed as follows: For each bin of travel time to work by car, truck, or van, take the mean of the bin (for travel time over 60 minutes, use 60 minutes), and weigh the means by the number of population in the bin.

¶In 1,000 2019 inflation-adjusted dollars.

‖In 1,000,000 square feet.

**A one-mile buffer zone is drawn at the border of LA County, and any CBGs that intersect this buffer are considered to be located at the border.

Table 2.2: What CBG characterisitics are associated with charger existence in a CBG?

Dependent Variable:	LPM			Probit		
Charger existence	(1)	(2)	(3)	(4)	(5)	(6)
MUD	0.308***	0.286***	0.297***	0.216***	0.218***	0.223***
	(0.070)	(0.080)	(0.080)	(0.051)	(0.056)	(0.057)
Office	0.173***	0.160***	0.151**	0.202***	0.186***	0.179***
	(0.065)	(0.061)	(0.059)	(0.062)	(0.059)	(0.058)
Retail	0.485***	0.469***	0.464***	0.510***	0.486***	0.483***
	(0.056)	(0.057)	(0.057)	(0.064)	(0.062)	(0.063)
Government	0.058**	0.053**	0.050**	0.066***	0.062***	0.058***
	(0.023)	(0.024)	(0.023)	(0.021)	(0.020)	(0.021)
Car Dealer	0.044***	0.043***	0.043***	0.034***	0.033***	0.033***
	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)
Transportation	0.157***	0.152***	0.143***	0.108***	0.101***	0.097***
	(0.036)	(0.036)	(0.036)	(0.023)	(0.022)	(0.023)
School	0.056***	0.051***	0.050***	0.035***	0.031***	0.030***
	(0.014)	(0.014)	(0.014)	(0.010)	(0.010)	(0.010)
Hotel	0.048*	0.046	0.045	0.056***	0.056***	0.056***
	(0.029)	(0.030)	(0.029)	(0.015)	(0.015)	(0.015)
Population	0.042***	0.033***	0.033***	0.039***	0.030***	0.030***
	(0.006)	(0.006)	(0.007)	(0.005)	(0.005)	(0.005)
Black	-0.038	0.029	0.033	-0.037	0.025	0.030
	(0.033)	(0.050)	(0.051)	(0.036)	(0.051)	(0.052)
Asian	-0.053*	0.071	0.095**	-0.044	0.072*	0.092**
	(0.032)	(0.044)	(0.048)	(0.030)	(0.040)	(0.042)
Other Races	-0.004	-0.031	-0.032	-0.003	-0.045	-0.047
	(0.032)	(0.038)	(0.039)	(0.037)	(0.043)	(0.044)
Hispanic	-0.043	0.019	0.026	-0.045	0.030	0.036
	(0.037)	(0.045)	(0.047)	(0.038)	(0.046)	(0.047)
Female	0.025	0.032	0.015	0.063	0.074	0.056
	(0.078)	(0.077)	(0.078)	(0.071)	(0.071)	(0.072)
Age	-0.002**	-0.002**	-0.002**	-0.002**	-0.002*	-0.001*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Owner-Occupied	-0.001	-0.042	-0.042	0.001	-0.033	-0.034
	(0.030)	(0.033)	(0.033)	(0.028)	(0.030)	(0.031)
Household Size	-0.030***	-0.032***	-0.031***	-0.032***	-0.029***	-0.028**
	(0.009)	(0.010)	(0.010)	(0.010)	(0.011)	(0.011)
High School	-0.073	-0.068	-0.068	-0.096	-0.095	-0.092
	(0.060)	(0.064)	(0.065)	(0.068)	(0.073)	(0.074)
						(Continued)

Some College	-0.102*	-0.093	-0.095	-0.127**	-0.130*	-0.131*
	(0.057)	(0.064)	(0.065)	(0.060)	(0.068)	(0.069)
Bachelor	0.126*	0.095	0.092	0.050	0.012	0.013
	(0.071)	(0.078)	(0.079)	(0.067)	(0.074)	(0.075)
Above Bachelor	0.035	-0.015	-0.018	-0.020	-0.053	-0.049
	(0.084)	(0.090)	(0.093)	(0.076)	(0.085)	(0.087)
Commute Time	-0.0005	0.0001	0.0004	0.0000	0.0005	0.0007
	(0.0009)	(0.0010)	(0.0010)	(0.0009)	(0.0010)	(0.0010)
Income	0.0001	0.0000	0.0001	0.0002	0.0002	0.0003
	(0.0004)	(0.0004)	(0.0004)	(0.0003)	(0.0003)	(0.0003)
Neighbor Variables						
MUD		-0.214	-0.150		-0.512	-0.465
		(0.386)	(0.393)		(0.393)	(0.397)
Office		1.739**	1.819**		0.889*	0.964**
		(0.709)	(0.708)		(0.484)	(0.488)
Retail		-0.673	-0.744		-0.365	-0.415
		(1.570)	(1.580)		(1.297)	(1.308)
Government		-0.421**	-0.423**		-0.150	-0.154
		(0.204)	(0.204)		(0.174)	(0.176)
Car Dealer		0.074	0.075		0.032	0.033
		(0.086)	(0.086)		(0.076)	(0.077)
Transportation		1.023**	1.008**		0.873***	0.863***
		(0.444)	(0.446)		(0.324)	(0.326)
School		0.267	0.259		0.117	0.113
		(0.221)	(0.223)		(0.185)	(0.187)
Hotel		-0.138	-0.143		-0.084	-0.090
		(0.147)	(0.147)		(0.112)	(0.113)
Population		0.186***	0.204***		0.168***	0.186***
		(0.063)	(0.064)		(0.053)	(0.054)
Black		-0.177	-0.165		-0.178	-0.166
		(0.262)	(0.264)		(0.261)	(0.264)
Asian		-0.958***	-1.039***		-1.006***	-1.082***
		(0.254)	(0.260)		(0.254)	(0.260)
Other Races		0.508**	0.526***		0.814***	0.826***
		(0.201)	(0.202)		(0.258)	(0.259)
Hispanic		-0.703**	-0.703**		-0.838***	-0.839***
		(0.276)	(0.278)		(0.316)	(0.320)
Female		0.211	0.255		0.750	0.766
		(0.934)	(0.940)		(0.888)	(0.898)
						(Continued)

Age		-0.013	-0.013		-0.009	-0.009
		(0.008)	(0.008)		(0.009)	(0.009)
Owner-Occupied		0.755**	0.747**		0.652*	0.631*
		(0.305)	(0.306)		(0.340)	(0.344)
Household Size		-0.152*	-0.145*		-0.186*	-0.177*
		(0.083)	(0.084)		(0.101)	(0.102)
Below High		1.062	0.947		0.983	0.871
		(0.872)	(0.880)		(0.930)	(0.939)
High School		1.728*	1.620		1.042	0.954
		(1.002)	(1.012)		(1.044)	(1.054)
Some College		-1.190*	-1.333*		-1.110	-1.253*
		(0.722)	(0.732)		(0.725)	(0.733)
Bachelor		1.716*	1.736*		1.671*	1.726*
		(0.901)	(0.911)		(0.938)	(0.948)
Above Bachelor		-0.430	-0.494		-0.548	-0.607
		(1.025)	(1.037)		(0.938)	(0.945)
Commute Time		-0.005	-0.004		-0.004	-0.003
		(0.009)	(0.009)		(0.009)	(0.009)
Income		-0.000	-0.002		-0.003	-0.004
		(0.005)	(0.005)		(0.004)	(0.004)
Observations	6,373	6,373	6,167	6,373	6,373	6,167
Neighbor Characteristics	No	Yes	Yes	No	Yes	Yes
Exclude Border CBGs	No	No	Yes	No	No	Yes

Notes: The dependent variable is an indicator variable that equals to one if there exist chargers in a CBG and equals to zero otherwise. Columns (1)-(3) report coefficients from the LPM. Columns (4)-(6) report average marginal effects derived from the probit model. The population-weighted inverse-distance-squared matrix is used to aggregate neighbor characteristics. Variables *White* (percentage of the White population), *Below High* (percentage of the population without a high school diploma) and *Neighbor White* (percentage of the White population in neighboring CBGs) are omitted as base groups. *Water*, *CBG Area* and the corresponding neighbor variables are also controlled but omitted from the regression outputs as they are not variables of interest. Robust standard errors in parentheses in columns (1)-(3). Delta-method standard errors in parentheses in columns (4)-(6). *** p<0.01, ** p<0.05, * p<0.1.

Table 2.3: Regression results under the OLS model and the Tobit model

Dependent Variable:	OLS			Tobit		
Number of Chargers	(1)	(2)	(3)	(4)	(5)	(6)
MUD	3.071*	4.239**	4.268**	6.929**	6.996*	6.977*
	(1.666)	(1.907)	(1.936)	(3.469)	(3.847)	(3.898)
Office	9.763***	9.708***	9.642***	13.379***	12.888***	12.636***
	(1.507)	(1.511)	(1.529)	(2.010)	(1.928)	(1.893)
Retail	7.182***	7.108***	7.077***	18.907***	17.890***	17.513***
	(1.690)	(1.667)	(1.679)	(2.597)	(2.558)	(2.529)
Government	4.931***	5.036***	5.055***	6.444***	6.496***	6.420***
	(1.340)	(1.340)	(1.349)	(0.971)	(0.917)	(0.946)
Car Dealer	-0.125	-0.155	-0.151	0.979***	0.845***	0.829***
	(0.133)	(0.137)	(0.138)	(0.293)	(0.291)	(0.292)
Transportation	3.032***	2.886***	2.918**	7.538***	6.865***	6.711***
	(1.122)	(1.116)	(1.149)	(1.681)	(1.683)	(1.737)
School	0.627*	0.591*	0.594*	3.193***	2.800***	2.810***
	(0.321)	(0.319)	(0.336)	(0.735)	(0.730)	(0.761)
Hotel	1.840***	2.000***	2.045***	2.370***	2.653***	2.645***
	(0.589)	(0.599)	(0.607)	(0.692)	(0.720)	(0.720)
Population	0.574***	0.593**	0.600**	2.796***	2.272***	2.254***
	(0.207)	(0.232)	(0.243)	(0.449)	(0.499)	(0.513)
Black	0.554	0.664	0.718	-1.097	4.638	4.923
	(0.535)	(0.857)	(0.870)	(2.609)	(3.786)	(3.801)
Asian	0.117	0.831	1.037	-2.003	6.531**	8.034**
	(0.542)	(0.764)	(0.841)	(2.220)	(3.011)	(3.189)
Other Races	-0.429	-0.320	-0.284	-0.707	-1.952	-1.905
	(0.447)	(0.548)	(0.555)	(2.644)	(3.097)	(3.113)
Hispanic	-0.088	-0.694	-0.704	-2.378	1.506	1.585
	(0.462)	(0.546)	(0.561)	(2.526)	(2.995)	(3.026)
Female	-0.527	-0.932	-1.015	1.645	1.556	0.359
	(1.112)	(1.113)	(1.118)	(4.621)	(4.705)	(4.714)
Age	-0.039**	-0.039*	-0.038*	-0.131*	-0.130*	-0.129*
	(0.019)	(0.020)	(0.021)	(0.069)	(0.071)	(0.073)
Owner-Occupied	1.499***	1.506**	1.433**	-0.534	-2.325	-2.633
	(0.542)	(0.584)	(0.598)	(2.022)	(2.199)	(2.235)
Household Size	-0.272	-0.328*	-0.286	-2.807***	-2.796***	-2.625***
	(0.172)	(0.186)	(0.190)	(0.825)	(0.876)	(0.878)

(Continued)

High School	-2.938***	-3.045***	-3.005***	-9.930**	-11.847**	-11.902**
	(0.692)	(0.785)	(0.800)	(4.558)	(5.002)	(5.043)
Some College	-1.061	-1.296	-1.362	-6.323	-8.322*	-8.361*
	(0.864)	(0.992)	(1.017)	(4.137)	(4.777)	(4.831)
Bachelor	-1.767*	-1.951*	-1.917*	2.674	-1.040	-1.350
	(0.921)	(1.003)	(1.028)	(4.446)	(4.944)	(5.004)
Above Bachelor	2.394*	1.513	1.761	4.289	-0.419	0.343
	(1.248)	(1.260)	(1.330)	(5.058)	(5.588)	(5.729)
Commute Time	0.021*	0.023*	0.024*	0.025	0.068	0.081
	(0.012)	(0.013)	(0.013)	(0.059)	(0.063)	(0.065)
Income	-0.006	-0.006	-0.005	-0.009	-0.009	-0.007
	(0.005)	(0.006)	(0.006)	(0.020)	(0.021)	(0.022)
Observations	6,373	6,373	6,167	6,373	6,373	6,167
Neighbor Characteristics	No	Yes	Yes	No	Yes	Yes
Exclude Border CBGs	No	No	Yes	No	No	Yes
R^2/Pseudo R^2	0.407	0.414	0.418	0.118	0.126	0.127

Notes: The population-weighted inverse-distance-squared matrix is used to aggregate neighbor characteristics. Variables *White* (percentage of the White population), *Below High* (percentage of the population without a high school diploma) and *Neighbor White* (percentage of the White population in neighboring CBGs) are omitted as base groups. *Water*, *CBG Area* and the corresponding neighbor variables are also controlled but omitted from the regression outputs as they are not variables of interest. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.4: Estimation of the intensive margin in Tobit model

Dependent Variable: No. of chargers	(1)	(2)	(3)
MUD	1.275**	1.282*	1.280*
	(0.639)	(0.705)	(0.716)
Office	2.463***	2.361***	2.319***
	(0.370)	(0.353)	(0.347)
Retail	3.480***	3.278***	3.214***
	(0.477)	(0.467)	(0.463)
Government	1.186***	1.190***	1.178***
	(0.179)	(0.168)	(0.174)
Car Dealer	0.180***	0.155***	0.152***
	(0.054)	(0.053)	(0.053)
Transportation	1.387***	1.258***	1.232***
	(0.308)	(0.307)	(0.317)
School	0.588***	0.513***	0.516***
	(0.135)	(0.134)	(0.140)
Hotel	0.436***	0.486***	0.485***
	(0.127)	(0.132)	(0.132)
Population	0.515***	0.416***	0.414***
	(0.082)	(0.091)	(0.094)
Black	-0.202	0.850	0.904
	(0.480)	(0.693)	(0.697)
Asian	-0.369	1.196**	1.474**
	(0.408)	(0.551)	(0.585)
Other Races	-0.130	-0.358	-0.350
	(0.487)	(0.567)	(0.571)
Hispanic	-0.438	0.276	0.291
	(0.465)	(0.549)	(0.555)
Female	0.303	0.285	0.066
	(0.850)	(0.862)	(0.865)
Age	-0.024*	-0.024*	-0.024*
	(0.013)	(0.013)	(0.013)
Owner-Occupied	-0.098	-0.426	-0.483
	(0.372)	(0.403)	(0.410)
Household Size	-0.517***	-0.512***	-0.482***
	(0.151)	(0.160)	(0.161)
			(Continued)

High School	-1.828**	-2.170**	-2.184**
	(0.838)	(0.914)	(0.923)
Some College	-1.164	-1.525*	-1.534*
	(0.761)	(0.874)	(0.885)
Bachelor	0.492	-0.191	-0.248
	(0.818)	(0.906)	(0.918)
Above Bachelor	0.789	-0.077	0.063
	(0.931)	(1.024)	(1.051)
Commute Time	0.005	0.012	0.015
	(0.011)	(0.012)	(0.012)
Income	-0.002	-0.002	-0.001
	(0.004)	(0.004)	(0.004)
Observations	6,373	6,373	6,167
Neighbor Characteristics	No	Yes	Yes
Exclude Border CBGs	No	No	Yes

Notes: The population-weighted inverse-distance-squared matrix is used to aggregate neighbor characteristics. Variables *White* (percentage of the White population), *Below High* (percentage of the population without a high school diploma) and *Neighbor White* (percentage of the White population in neighboring CBGs) are omitted as base groups. *Water*, *CBG Area* and the corresponding neighbor variables are also controlled but omitted from the regression outputs as they are not variables of interest. Delta-method standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.5: Regression results under the ZINB model

Dependent Variable:	NB		ZINB			
No. of Chargers	(1)	(2)	(3)	(4)	(5)	(6)
MUD	0.846	0.565	0.505	0.306	0.762	0.704
	(0.532)	(0.561)	(0.503)	(0.527)	(0.532)	(0.531)
Office	3.370***	3.086***	2.783***	2.463***	2.515***	2.481***
	(0.555)	(0.546)	(0.538)	(0.533)	(0.539)	(0.551)
Retail	3.796***	3.487***	2.809***	2.499***	2.510***	2.336***
	(0.683)	(0.585)	(0.542)	(0.484)	(0.499)	(0.472)
Government	1.425***	1.281***	1.212***	1.059***	0.999***	0.993***
	(0.279)	(0.270)	(0.280)	(0.278)	(0.271)	(0.274)
Car Dealer	0.197***	0.129**	0.177***	0.107*	0.108**	0.110**
	(0.063)	(0.060)	(0.060)	(0.055)	(0.052)	(0.053)
Transportation	1.738***	1.329***	1.606***	1.203***	1.200***	1.179***
	(0.487)	(0.364)	(0.480)	(0.344)	(0.343)	(0.340)
School	0.852***	0.826***	0.771***	0.709***	0.737***	0.734***
	(0.155)	(0.161)	(0.126)	(0.131)	(0.139)	(0.139)
Hotel	0.446***	0.407***	0.328***	0.267***	0.300***	0.302***
	(0.165)	(0.156)	(0.122)	(0.102)	(0.111)	(0.108)
Population	0.414***	0.284***	0.357***	0.228***	0.227***	0.222***
	(0.066)	(0.062)	(0.063)	(0.056)	(0.055)	(0.056)
Black	0.375	1.555**	0.408	1.708**	1.838**	1.868**
	(0.539)	(0.736)	(0.539)	(0.728)	(0.738)	(0.734)
Asian	0.062	0.869	0.055	0.922*	0.860*	1.002*
	(0.398)	(0.536)	(0.386)	(0.514)	(0.510)	(0.516)
Other Races	0.002	-0.135	0.061	-0.003	0.136	0.223
	(0.581)	(0.651)	(0.563)	(0.622)	(0.607)	(0.610)
Hispanic	0.276	0.683	0.181	0.677	0.544	0.532
	(0.545)	(0.589)	(0.525)	(0.566)	(0.569)	(0.570)
Female	-0.086	0.498	-0.439	0.185	0.144	-0.018
	(1.073)	(0.973)	(0.963)	(0.896)	(0.882)	(0.880)
Age	-0.013	-0.015	-0.013	-0.014	-0.015	-0.014
	(0.011)	(0.011)	(0.010)	(0.010)	(0.010)	(0.010)
Owner-Occupied	0.006	-0.167	0.828**	0.672*	0.691	0.640
	(0.358)	(0.374)	(0.595)	(0.391)	(0.454)	(0.429)
Household Size	-0.520***	-0.491***	-0.448***	-0.373***	-0.375***	-0.340**
	(0.130)	(0.132)	(0.129)	(0.131)	(0.132)	(0.134)

(Continued)

High School	-0.334	-0.712	-0.504	-0.713	-0.647	-0.636
	(0.890)	(0.987)	(0.832)	(0.923)	(0.905)	(0.911)
Some College	0.134	-0.452	0.040	-0.392	-0.103	-0.106
	(0.720)	(0.835)	(0.683)	(0.791)	(0.796)	(0.801)
Bachelor	0.769	-0.165	0.874	0.189	0.324	0.341
	(0.844)	(0.938)	(0.804)	(0.863)	(0.857)	(0.854)
Above Bachelor	1.195	0.046	1.563	0.784	0.713	0.955
	(0.977)	(1.059)	(0.952)	(1.049)	(1.058)	(1.062)
Commute Time	0.005	0.016	0.010	0.020	0.021*	0.022*
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.012)
Income	0.003	-0.001	0.001	-0.002	-0.002	-0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
First-Stage Logit						
MUD			-12.550***	-12.225***	-6.955***	-7.462***
			(2.575)	(2.771)	(2.133)	(2.242)
Owner-Occupied			1.039*	1.290**	1.418**	1.388**
			(0.595)	(0.567)	(0.689)	(0.677)
Neighbor MUD					-35.470	-30.630
					(28.704)	(19.603)
Neighbor Owner-Occupied					10.927***	10.325***
					(3.118)	(3.099)
Observations	6,373	6,373	6,373	6,373	6,373	6,167
Neighbor Characteristics	No	Yes	No	Yes	Yes	Yes
Exclude Border CBGs	No	No	No	No	No	Yes

Notes: The population-weighted inverse-distance-squared matrix is used to aggregate neighbor characteristics. Variables *White* (percentage of the White population), *Below High* (percentage of the population without a high school diploma) and *Neighbor White* (percentage of the White population in neighboring CBGs) are omitted as base groups. *Water*, *CBG Area* and the corresponding neighbor variables are also controlled but omitted from the regression outputs as they are not variables of interest. Unreported likelihood-ratio tests comparing the ZINB model with the zero-inflated Poisson model suggest that the zero-inflated negative binomial model is preferred to the zero-inflated poisson model. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.6: Robustness checks using truncated population-weighted inverse-distance-squared matrix

Model	LPM	Probit	Tobit Int. Margin	ZINB	ZINB
	(1)	(2)	(3)	(4)	(5)
MUD	0.346***	0.251***	1.494**	0.742	0.708
	(0.080)	(0.056)	(0.691)	(0.512)	(0.503)
Office	0.145**	0.176***	2.277***	1.999***	1.969***
	(0.057)	(0.055)	(0.333)	(0.487)	(0.486)
Retail	0.440***	0.455***	3.023***	2.068***	2.034***
	(0.056)	(0.060)	(0.448)	(0.467)	(0.461)
Government	0.046**	0.055***	1.122***	0.945***	0.906***
	(0.021)	(0.021)	(0.184)	(0.255)	(0.255)
Car Dealer	0.042***	0.034***	0.140***	0.103*	0.099*
	(0.008)	(0.007)	(0.053)	(0.059)	(0.057)
Transportation	0.136***	0.092***	1.181***	0.867***	0.905***
	(0.036)	(0.023)	(0.314)	(0.246)	(0.254)
School	0.044***	0.025**	0.430***	0.473***	0.502***
	(0.014)	(0.010)	(0.137)	(0.119)	(0.122)
Hotel	0.045	0.054***	0.483***	0.243***	0.238***
	(0.030)	(0.015)	(0.131)	(0.091)	(0.090)
Population	0.027***	0.029***	0.377***	0.129**	0.129**
	(0.007)	(0.005)	(0.099)	(0.060)	(0.057)
Black	0.016	0.011	0.820	1.258*	1.392**
	(0.049)	(0.050)	(0.688)	(0.647)	(0.655)
Asian	0.014	0.014	0.614	0.414	0.519
	(0.044)	(0.039)	(0.539)	(0.468)	(0.474)
Other Races	-0.031	-0.041	-0.339	-0.016	0.043
	(0.038)	(0.043)	(0.549)	(0.630)	(0.618)
Hispanic	-0.008	-0.002	-0.082	0.454	0.466
	(0.045)	(0.045)	(0.536)	(0.570)	(0.564)
Female	0.003	0.042	-0.072	0.055	-0.053
	(0.079)	(0.073)	(0.867)	(0.896)	(0.888)
Age	-0.002*	-0.001*	-0.020	-0.014	-0.014
	(0.001)	(0.001)	(0.013)	(0.010)	(0.010)
Owner-Occupied	-0.057*	-0.044	-0.734*	0.364	0.501
	(0.034)	(0.031)	(0.410)	(0.395)	(0.395)
Household Size	-0.026***	-0.024**	-0.398**	-0.277**	-0.273**
	(0.010)	(0.011)	(0.156)	(0.132)	(0.132)
High School	-0.097	-0.131*	-2.615***	-1.375	-1.399
	(0.064)	(0.072)	(0.905)	(0.938)	(0.921)

(Continued)

Some College	-0.145**	-0.180***	-2.037**	-0.902	-0.890
	(0.063)	(0.067)	(0.858)	(0.777)	(0.765)
Bachelor	0.093	0.004	-0.365	-0.194	-0.172
	(0.079)	(0.074)	(0.888)	(0.879)	(0.868)
Above Bachelor	-0.047	-0.085	-0.285	0.446	0.591
	(0.091)	(0.085)	(1.038)	(1.024)	(1.030)
Commute Time	0.000	0.001	0.017	0.016	0.016
	(0.001)	(0.001)	(0.012)	(0.012)	(0.012)
Income	-0.000	0.000	-0.002	-0.002	-0.002
	(0.000)	(0.000)	(0.004)	(0.004)	(0.004)
First-Stage Logit					
MUD				-11.238***	-9.710***
				(2.682)	(2.306)
Owner-Occupied				1.644***	1.869***
				(0.633)	(0.629)
Neighbor MUD					1.695
					(2.049)
Neighbor Owner-Occupied					9.458***
					(3.523)
Observations	6,073	6,073	6,073	6,073	6,073
Neighbor Characteristics	Yes	Yes	Yes	Yes	Yes
Exclude Border CBGs	Yes	Yes	Yes	Yes	Yes

Notes: In columns (1)-(2), the dependent variable is an indicator variable that equals to one if there exist chargers in a CBG and equals to zero otherwise. Columns (3)-(5) use the number of chargers as the dependent variable. Column (1) presents the estimation results under the LPM. Column (2) presents estimation results under the probit model. Column (3) presents the average marginal effects of CBG characteristics on the number of chargers in a CBG conditional on the presence of chargers, derived from the Tobit model. Columns (4)-(5) present the estimation results under the ZINB model. The population-weighted inverse-distance-squared matrix truncated at distance equal to one mile is used to aggregate neighbor characteristics. The sample consists of CBGs that have neighbor CBGs within one mile and are not border CBGs. Variables *White* (percentage of the White population), *Below High* (percentage of the population without a high school diploma) and *Neighbor White* (percentage of the White population in neighboring CBGs) are omitted as base groups. *Water*, *CBG Area* and the corresponding neighbor variables are also controlled but omitted from the regression outputs as they are not variables of interest. Delta-method standard errors in parentheses in column (3). Robust standard errors in parentheses in the rest columns. *** p<0.01, ** p<0.05, * p<0.1.

Chapter 3

Freeway Congestion and Labor Force Participation: Evidence from Greater Los Angeles

3.1 Introduction

Traffic congestion is a growing problem in major metropolitan areas. According to the Texas A&M Transportation Institute (TTI), the annual delay per auto commuter caused by congestion increased from 38 hours in 2000 to 54 hours in 2019.[1] Congestion increases commute time, which has been found to affect both the extensive (Black et al., 2014) and intensive margin of labor supply (Gutiérrez-i-Puigarnau and van Ommeren, 2010; Gershenson, 2013; Fu and Viard, 2018; Farré et al., 2023), job search behavior (Le Barbanchon et al., 2021),

[1] Accessed from https://mobility.tamu.edu/umr/report/ on Mar 29, 2024. The annual delay per commuter is a yearly sum of all the per-trip delays for people who travel in private vehicles during peak periods (6 to 10 am and 3 to 7 pm). Note that this definition of auto commuter in the TTI report is different from the one adopted in this paper, which refers to people who travel *to/from work* in private vehicles during peak periods.

and can potentially affect workers' choice of residential location and workplace location as well.

The existing literature that examines the effects of congestion on the labor market focuses on its broader economic impact on large urban areas. Using measures of congestion from TTI's *Urban Mobility Reports* at urban-area[2] level, Hymel (2009) and Sweet (2014) find that congestion reduces employment growth. There is little research on the individual commuter's response to congestion, however, mainly due to difficulty in measuring historical congestion at the commute-route level, which requires data on route-specific historical commute/travel time. Unfortunately, web mapping services commonly used in the urban and transportation literature, such as Google Maps, only support queries of *real-time* travel time. Therefore, using these web mapping services to obtain historical travel time either requires advance planning or strict assumptions on traffic conditions.[3] Other studies utilize historical speeds collected from freeway traffic detectors to estimate historical travel time (Beland and Brent, 2018). However, due to complications in routing, Beland and Brent (2018) restrict their focus to commute flows that travel on select freeways. This paper builds and improves on Beland and Brent (2018), in that it is able to estimate historical commute time for *any commute route* with the help of OSMnx (Boeing, 2017), a Python package that makes it possible to incorporate historical freeway speeds into open-sourced road networks (i.e., OpenStreetMap), and to find the optimal route between locations.

This paper studies how people respond to freeway congestion along their commute to work in Greater Los Angeles (LA), one of the most congested metropolitan areas in the United States.[4] Greater LA, an area that includes LA County, Orange County, San Bernardino

[2]Urban areas are defined by the Census Bureau. For more details, refer to https://www.census.gov/programs-surveys/geography/about/faq/2010-urban-area-faq.html.

[3]Kim and Long (2024) use historical travel time that was collected real-time from Google Maps for a past project. Akbar and Duranton (2017) use trip durations queried real-time in 2015 as a measure for trip durations in 2011 by assuming similar traffic conditions in the two periods.

[4]According to TTI, the Los Angeles-Long Beach-Anaheim urban area has the highest Travel Time Index (ratio of peak-period travel time to free-flow travel time) in the United States, ranging from 1.45 to 1.52 between 2010 and 2019. The Riverside-San Bernardino urban area is among the 15 most congested urban

County, Riverside County, and Ventura County, is also a commuting zone,[5] making it an appropriate context to study changes in residential location and labor market outcomes in response to congestion. Utilizing open-sourced OpenStreetMap and historical freeway speed data maintained by the California Department of Transportation (Caltrans), I estimate the average delay caused by freeway congestion during morning rush hours from 2010 to 2018 for more than 1.3 million commute routes (defined by residence-workplace census-tract pair) within Greater LA. Pairing data on congestion-induced delay with data on tract-to-tract commute flows constructed from Longitudinal Employer-Household Dynamics Origin-Destination Employment Statistics (LODES), this study finds that the number of commuters on a given route decreases as congestion along the route increases. In addition, there is suggestive evidence that the decrease in commuters on congested routes more likely results from workers changing jobs for a less-congested commute, than from workers relocating their residences for a less-congested commute, or dropping out of the labor market. Lastly, I find that switching jobs for a less-congested commute comes with the cost of lower earnings.

This study has two major contributions. First, it contributes to the literature on the economic costs of congestion. Existing literature has documented that congestion causes negative health effects (Currie and Walker, 2011; Knittel et al., 2016; Brent and Beland, 2020; Bencsik et al., 2023), increases domestic violence (Beland and Brent, 2018), and discourages employment growth (Hymel, 2009; Sweet, 2014). This paper adds to this strand of literature by exploring the effects of congestion on individual labor market outcomes. In addition, it develops an innovative approach to measuring historical congestion by using open-sourced road network data augmented with historical speeds collected from freeway traffic detectors. As speeds recorded by traffic detectors have fine temporal (e.g., speed data from Caltrans are precise to 30-second intervals) and spatial granularity, this approach can be used in other research that requires estimation of congestion in varying degrees of granularity.

areas in terms of its Travel Time Index, which ranges from 1.31 to 1.34 between 2010 and 2019.

[5]From 2010 to 2019, more than 93% of people who work in Greater LA live in Greater LA, and more than 93% of residents in Greater LA work in Greater LA, according to LODES data.

This study also complements the literature that studies the effects of commute time on labor market decisions (Gutiérrez-i-Puigarnau and van Ommeren, 2010; Gershenson, 2013; Black et al., 2014; Fu and Viard, 2018; Le Barbanchon et al., 2021; Farré et al., 2023). This paper focuses on the role that congestion, an increasingly important determinant of commute time, plays in an individual's labor market decisions.

The remainder of the paper is organized as follows. Section 3.2 introduces the data. Section 3.3 describes the empirical models and discusses results. Section 3.4 concludes.

3.2 Data

3.2.1 Labor market outcomes

3.2.1.1 Commute flow

Data on commute flows within Greater LA are from LODES.[6] The LODES data cover approximately 95 percent of wage and salary jobs. Omitted workers include informal workers, self-employed workers, workers affiliated with the military and other security-related federal agencies, etc. (Graham et al., 2014). To construct commute flows, I use the Origin-Destination (OD) data, which include the count of workers for each pair of workplace census block and residence census block.[7] Given its focus on commute trips, this paper uses "origin" interchangeably with "residence," and "destination" interchangeably with "workplace." Since noise infusion in LODES for confidentiality purposes makes block-level analyses unreliable, the counts of jobs are aggregated to the census-tract-pair level, and are referred to

[6]LODES data are publicly available at https://lehd.ces.census.gov/data/.

[7]The OD data provide the count of *jobs* for each pair of workplace census block and residence census block. I restrict the job type to be "primary," meaning that if a person holds two or more jobs, only the one with the most earnings (primary job) is counted. Under this restriction, the count of jobs is equivalent to the count of workers.

as tract-to-tract commute flows.[8] There are 7,322,550 tract pairs within Greater LA with commute flows in at least one year from 2010 to 2019. For computational ease, the following steps are taken to construct a set of tract pairs that represent typical commuting patterns, similar to Beland and Brent (2018). First, I compute the total commute flows from 2010 to 2019 between each tract pair. Second, for a given origin tract, the set of destination tracts are ordered from the most commuted to the least commuted. Last, I keep the top destination tracts that account for 75% of the total commute flows from the origin tract,[9] resulting in a sample of 1,372,819 tract pairs for future analyses. The rationale behind dropping destination tracts that are less commuted to is as follows. For destination tracts that consistently have small commute flows, congestion is likely to have a smaller effect on commute flows, compared to factors such as lack of job opportunities and remoteness. Therefore, the effects of congestion along these routes are of less policy importance, justifying their exclusion from the sample.

There are two caveats about constructing commute flows using the LODES data. First, a workplace is defined as the physical or mailing address reported by employers in the administrative data, and may not be the actual location that a worker reports to. To alleviate this concern, I perform robustness checks in Section 3.3.1 by excluding tracts where a considerable number of residents work from home. Still, the paper is unable to identify or address the situation where an employer's work location is different from the physical or mailing address reported by employers. Despite this limitation, LODES data provide the most spatially and temporally granular data on linked residence and workplace job flows among publicly available datasets. Second, the LODES data do not include information on transport mode, making it impossible to construct auto-specific commute flows. To

[8] Refer to Graham et al. (2014) for details regarding the noise infusion practices. Aggregation of data to the census-tract level is commonly adopted when using LODES data due to noise infusion (Couture and Handbury, 2020; Owens et al., 2020; Shoag and Veuger, 2021; Tyndall, 2021).

[9] In the majority of cases, multiple destination tracts have the same number of workers commuting from a given origin tract, resulting in a tie between destination tracts at the 75% cutoff. Instead of randomly selecting a subset from these destination tracts, this paper includes all of them in the sample.

address this limitation, I perform robustness checks in Section 3.3.1 by excluding tracts where a considerable number of residents commute to work by walking, biking or taking public transportation.

3.2.1.2 Labor force participation rate

The tract-level civilian labor force participation rate for population 16 to 64 years old (henceforth, labor force participation rate) from 2011 to 2019 is computed using American Community Survey (ACS) 5-year estimates.[10]

3.2.1.3 Percentage of working residents with high earnings

The OD data tabulate the count of workers by three categories of monthly earnings: \$1,250/month or less, \$1,251/month to \$3,333/month, and greater than \$3,333/month. The percentage of working residents with high-paying jobs in a given residence tract is its percentage of residents working in Greater LA with monthly earnings over \$3,333/month.

Summary statistics for the above three variables are presented in Table 3.1a.

3.2.2 Congestion

Annual congestion between residence-workplace census-tract pairs is estimated from 2010 to 2018. Quantifying congestion requires the estimation of the free-flow travel time ($travel\_time^0_{ij}$) and average travel time during morning rush hours ($travel\_time_{ijt}$) from census tract i to census tract j in year t. Morning rush hours are defined to be between 5 am and 9 am from Monday to Friday, following Beland and Brent (2018). Congestion is measured by the

[10]Table B23001 of ACS tabulates the estimates of civilians in labor force and population not in labor force for different age groups.

increase in travel time with congested traffic compared to travel time without congestion, referred to as congestion-induced delay[11]:

$$delay_{ijt} = travel\_time_{ijt} - travel\_time^0_{ij}. \tag{3.1}$$

Measuring historical travel time has always been a challenge in the urban and transportation literature, mainly because commonly-used web mapping services, such as Google Maps, only support queries of real-time travel time.[12] This paper develops a novel approach to measuring historical travel time by using open-sourced road network data augmented with historical speeds collected from freeway traffic detectors. The details are as follows.

The road network of Greater LA is retrieved from OpenStreetMap and processed through the Python package OSMnx (Boeing, 2017).[13] The road network consists of smaller road segments (henceforth edges), with an average length of 0.54 miles for freeway edges. Figure C.1 gives an illustration of these edges. Using the road network to estimate travel time requires knowledge on the driving speed on edges. Unfortunately, OpenStreetMap only comes with speed limits for edges. Therefore, I set the default speed for each edge to be its speed limit,[14] and revise the speed of the freeway edges based on historical freeway speed data to estimate

[11] This paper focuses on congestion during morning rush hours. This is because morning congestion is more disruptive to commuters as they are usually required to arrive in the office by a certain time, while they are more flexible as to when to arrive at their residence after work. Moreover, for a given commuter, morning congestion is highly correlated with afternoon/evening congestion.

[12] To my best knowledge, using Google Maps to obtain historical travel time is rare in the literature, with two exceptions, where either advance planning (Kim and Long, 2024) or strict assumptions on traffic conditions (Akbar and Duranton, 2017) is required. Kim and Long (2024) use historical travel time that was collected real-time from Google Maps for a past project. Akbar and Duranton (2017) use trip durations queried real-time in 2015 as a measure for trip durations in 2011 by assuming similar traffic conditions in the two periods.

[13] OpenStreetMap is a widely-used open-sourced road network in the urban and transportation literature (Heilmann, 2018; Trajkovski et al., 2021; Akbar et al., 2023; Mo, 2023), and has been used for estimating driving time and distance between locations (Haller and Heuermann, 2016; Heuermann et al., 2017; Holmes et al., 2019; Luco, 2019). OSMnx is a Python package used to easily download, model, analyze, and visualize street networks and other geospatial features from OpenStreetMap. This paper mainly uses OSMnx to download street networks from OpenStreetMap, revise speed on the road segments, and solve routing problems. For more details, refer to https://osmnx.readthedocs.io/en/stable/index.html.

[14] For edges with missing speed limit, its speed is set to be the average speed limit of the edges of the same highway type. For highway types, see https://wiki.openstreetmap.org/wiki/Key:highway.

$travel\_time^0_{ij}$ and $travel\_time_{ijt}$. Because edges not on the freeway are given the same (free-flow) speed over the years, the congestion estimated in this paper only measures congestion *on the freeway* along the commute route. In addition, the paper assumes that commuters do not travel on toll roads or HOV lanes.[15] Accordingly, the speeds of freeway edges that are tolled (e.g., part of SR73) are set to zero. A freeway that has HOV lanes or express lanes (e.g., SR91) is mapped as two sets of edges in the OpenStreetMap road networks: one representing the regular lanes, and the other representing the HOV lanes or express lanes. The speeds of freeway edges representing HOV lanes or express lanes are set to zero as well. The above design ensures that the commute routes to be estimated in the next steps will not involve toll roads or HOV lanes.[16] Henceforth, unless otherwise specified, freeway edges refer to those representing regular freeway lanes.

The freeway speed data are publicly available from the Performance Measurement System (PeMS) maintained by Caltrans. PeMS collects traffic data from vehicle detector stations (henceforth stations) located on freeways. I access data on hourly average speed and coordinates for mainline[17] stations from 2010 to 2019 in District 7 (LA County and Ventura County), District 8 (Riverside County and San Bernadino County), and District 12 (Orange County) using the "Data Clearinghouse" tool. There is one station every 0.59 to 0.69 miles, and the distance between stations decreases over time as more stations are added. For each edge, I find a matching station, which is the nearest station with the same freeway number and direction (e.g., I5-S). This process is illustrated in Figure C.2.

To compute $travel\_time_{ijt}$, the speed of a freeway edge is set to be the average morning rush-hour speed in year t of its matching station. The free-flow travel time, $travel\_time^0_{ij}$,

[15] According to the 2010–2012 California Household Travel Survey, among respondents who commuted within Greater LA, 2.25% and 3.77% of them use toll roads and HOV lanes respectively. Therefore, the assumption that commuters do not travel by toll roads or HOV lanes is consistent with the actual travel patterns of commuters.

[16] Another option is to remove freeway edges that are toll roads and HOV lanes from the road network. This option is not preferred as it runs the risk of damaging the connectivity of the network.

[17] Stations are located on mainlines, on-ramps, off-ramps, HOV lanes, etc. Only mainline and HOV stations report speed.

is year-invariant. To compute $travel\_time^0_{ij}$, the speed of a freeway edge is set to be the median of average speed between 10 pm and 11 pm of its matching station during the sample period. $travel\_time_{ijt}$ and $travel\_time^0_{ij}$ are estimated by finding the route with the minimum travel time between the centroid[18] of census tract i and the centroid of census tract j through OSMnx, using the road network with corresponding speeds. Estimating year-specific optimal routes for a given origin-destination pair acknowledges the possibility that drivers may change commute routes in response to changes in traffic conditions over time. Table 3.1b presents the change in morning commute time and delay caused by freeway congestion by commute distance for 2010 and 2018. Delay caused by freeway congestion has increased by 32% to 39% from 2010 to 2018.

To demonstrate the reliability of the above method in measuring historical travel time, I compare the self-reported travel time from the 2010–2012 California Household Travel Survey (CHTS) (National Renewable Energy Laboratory, 2013) with the estimated travel time using this method in Appendix C.2.

3.3 Empirical framework and results

3.3.1 Congestion and commute flow

To explore whether commuters respond to changes in congestion along their commute routes, I estimate the following regression model:

$$flow_{ijt} = \alpha_0 + \alpha_1 delay_{ij,t-1} + \lambda_{ij} + \eta_{it} + \phi_{jt} + \varepsilon_{ijt}, \tag{3.2}$$

[18]The coordinates of tract internal points available in the 2019 Census Bureau TIGER Geodatabase are used as the coordinates of tract centroids. According to https://www.census.gov/programs-surveys/geography/about/glossary.html, the internal point is usually the centroid, when the centroid is located inside the boundaries of a geographic entity. When the centroid is located outside the boundaries of a geographic entity (e.g., crescent-shaped areas), the internal point is identified as a point inside the entity boundaries nearest to the centroid.

where $flow_{ijt}$ is the number of workers that commute from residence tract i to workplace tract j in year t = 2011, 2012, ..., 2019. The variable of interest, $delay_{ij,t-1}$, is the one-year lag of the average congestion-induced delay during the morning commute from residence tract i to workplace tract j. Tract-pair fixed effects, λ_{ij}, control for tract-pair characteristics that do not vary over time, including the distance between residence tract i and workplace tract j. By controlling for tract-pair fixed effects, the model looks at how the commute flow along a given route changes in response to *temporal* changes in congestion. Residence-tract-by-year fixed effects, η_{it}, controls for residence tract characteristics that change over time, including housing price, demographic composition, etc. Workplace-tract-by-year fixed effects, ϕ_{jt}, controls for workplace tract characteristics that change over time, including job opportunities, etc.

α_1 is expected to be negative: The number of residents commuting along a given route is expected to decrease following an increase in congestion on that route. One may be concerned that congestion-induced delay is not exogenous, as there may exist reverse causality: More commute flows between a given tract-pair can lead to more congestion along the commute route between that tract-pair. Fortunately, the presence of reverse causality is expected to cause an upward bias in the estimate of α_1, making it a conservative estimate of the negative effect of congestion on the size of the commute flow. Second, I try to mitigate reverse causality by using lagged congestion-induced delay in the model. Last and most importantly, the paper's focus on tract-pair level commute flows and freeway congestion minimizes the possibility of reverse causality. This is because the commute flow between a given tract-pair constitutes a negligible share of total traffic on the freeway segments along which they travel, and therefore is unlikely to affect congestion on these freeway segments. Indeed, between 2010 and 2018, among the 1,152,699 tract-pair commute routes that pass through freeways, the commute flows along more than 99.7% of the routes contribute to less than 1% of the commute flows on the corresponding freeways, and the commute flows along more than 93.4% of the routes contribute to less than 0.1% of the commute flows on the corresponding freeways.

At first sight, delay caused by freeway congestion as a measure of congestion on a commute has its obvious limitation: it does not take into account congestion on roads other than freeways. However, including congestion on surface/local streets may raise the concern of reverse causality, as the commute flow between a given tract-pair can constitute a sufficiently large share of traffic along local streets on the commute route, such that an increase in commute flow can result in an increase in congestion.

Among the initial sample of 1,372,819 tract pairs introduced in Section 3.2.1.1, 6,202 tract pairs that involve tracts whose land area is 0 or island tracts are dropped, as the congestion estimated using road network for driving can be inaccurate for these tract pairs. An additional 1,427 tract pairs are excluded where I fail to find a route that does not involve toll roads or HOV lanes, resulting in the final sample of 1,365,190 tract pairs.

Column (1) of Table 3.2 reports estimates of model (3.2). Consistent with the hypothesis that congestion reduces commute flows, I find that a one-minute increase in congestion-induced delay reduces the commute flow by 0.086 persons on average. This change in the commute flow can be caused by people changing their residence tract (moving), changing their workplace tract (changing job, moving to another office), or disappearing from the LODES data altogether (dropping out of the labor market).

However, the likelihood of people moving/changing jobs not only depends on the amount of congestion they have been experiencing on their commute route, but also depends on the commute distance and the amount of congestion they would experience if they were to move/change job. As commute time on alternative routes decreases, it is expected that commuters are more likely to switch to these routes by moving/changing job, causing a decrease in the number of commuters on the current commute route. This dynamic suggests a positive correlation between commute flows and expected travel time on alternative commutes.

Meanwhile, congestion on the current route is also expected to be positively correlated with the expected congestion on alternative commutes. The current route and alternative routes tend to overlap in some part, as they either share the same origin (if commuters change job) or the same destination (if commuters move). Therefore, omitting expected travel time if commuters move/change workplace from the estimation can cause an upward bias in the estimate of α_1. Model (3.3) includes additional variables measuring expected travel time along alternative routes, and their interactions with $delay_{ij,t-1}$:

$$flow_{ijt} = \alpha_0 + \alpha_1 delay_{ij,t-1} + \alpha_2 travel\_time^{exclu\_W}_{ij,t-1} + \alpha_3 delay_{ij,t-1} \times travel\_time^{exclu\_W}_{ij,t-1} + \alpha_4 travel\_time^{exclu\_R}_{ij,t-1} + \alpha_5 delay_{ij,t-1} \times travel\_time^{exclu\_R}_{ij,t-1} + \lambda_{ij} + \eta_{it} + \phi_{jt} + \varepsilon_{ijt}. \tag{3.3}$$

For a given tract-pair ij, $travel\_time^{exclu\_W}_{ij,t-1}$, the expected travel time under workplace change for workers commuting between tracts i and j, is the average travel time in year $t-1$ on commute routes out of residence tract i weighted by commute flows in year $t-1$, *excluding* the one going to workplace tract j.[19] $travel\_time^{exclu\_W}_{ij,t-1}$ can also be interpreted as the average commute time faced by commuters who live in the same residence tract, tract i, as the focal commuters, but work in tracts other than j. The expected travel time under residence change, $travel\_time^{exclu\_R}_{ij,t-1}$, is the average travel time on commute routes to workplace tract j weighted by commute flows, *excluding* the one from residence tract i.[20]

The estimation results are reported in column (2) of Table 3.2. As expected, the coefficient of lagged delay decreases with the inclusion of additional variables that measure

[19] Mathematically, $travel\_time^{exclu\_W}_{ij,t-1} = \dfrac{\sum_{k\neq j} job\_flow_{ik,t-1} \times travel\_time_{ik,t-1}}{\sum_{k\neq j} job\_flow_{ik,t-1}}$.

[20] Mathematically, $travel\_time^{exclu\_R}_{ij,t-1} = \dfrac{\sum_{k\neq i} job\_flow_{kj,t-1} \times travel\_time_{kj,t-1}}{\sum_{k\neq i} job\_flow_{kj,t-1}}$.

expected travel time along alternative routes. The estimated coefficient of $travel\_time_{t-1}^{exclu\_W}$ is 2.81: Among commute routes out of the same residence tract, the commute flow to a given workplace tract increases by 2.81 persons on average as the average commute time to the other workplace tracts increases by one minute. This finding suggests that some commuters change workplaces for a less-congested commute, following an increase in delay on their commute route. In contrast, the estimated coefficient of $travel\_time_{t-1}^{exclu\_R}$ is economically small and statistically indistinguishable from zero. This result indicates that there is no evidence suggesting that commuters change residence for a less-congested commute, following an increase in delay on their commute route. The above findings imply that commuters tend to choose changing workplace over changing residence in response to a more congested commute. The estimated coefficients on both interaction terms are positive and statistically significant, indicating that the negative effects of congestion on the number of commuters on a given route decreases as workers' alternative commute routes become more congested.

Column (3) presents results after dropping commute routes (represented by tract-pairs) that may experience endogenous congestion. Specifically, I drop routes whose commute flows contribute to more than 0.1% of the total commute flows on the freeways they pass through. The estimates are quantitatively similar.

In addition, column (4) drops tract-pairs with residence tracts where at least 50% of workers work from home or commute to work via public transportation, biking or walking, as the commute time of these workers are not affected by freeway congestion.[21] As expected, the effects of congestion-induced delay on commute flow increase in magnitude after excluding commute routes less susceptible to congestion.

[21] Table B08006 of ACS tabulates the number of workers 16 years and over who went to work using public transportation, biked to work, walked to work, or worked from home the week prior to being surveyed. I compute the percentage of these workers using ACS 2010-2014 data, and ACS 2015-2019 data, respectively, and take the maximum. In column (3), tract-pairs with residence tracts where the maximum percentage of workers who work from home or commute to work via public transportation, biking or walking is at least 50% are dropped from the sample. Dropping tract-pairs based on this time-invariant threshold makes sure that the sample is still a balanced panel.

3.3.2 Congestion and labor market outcomes

Section 3.3.1 provides evidence that the number of workers on a given commute route decreases in response to increasing congestion on the route. This change in the commute flow can be caused by people moving, changing job, or dropping out of the labor market altogether. In this section, the paper takes a closer look at these different possible responses to congestion.

3.3.2.1 Congestion and labor force participation

First, the analysis examines whether congestion induces workers to drop out of the labor market. Specifically, I look at whether the labor force participation rate in a given census tract decreases, as its residents face more congestion along their commutes. The following model is estimated:

$$lfpr_{icp} = \beta_0 + \beta_1 avg\_delay_{ic,t-1} + X'_{icp}\delta + \xi_i + \psi_{cp} + u_{icp}, \tag{3.4}$$

where $lfpr_{icp}$ is the labor force participation rate in residence tract i in county c over a five-year period $p = [t, t+4]$, where t = 2011, 2012, ..., 2015. $avg\_delay_{ic,t-1}$ is the average delay in year t-1 on commute routes out of residence tract i weighted by commute flows in year t-1. The vector X_{icp} includes controls for residence tract i's percentage of female, percentage of families with children under six years old, housing price, and housing stock over period p. Appendix C.3 details the construction of the above control variables. Model (3.4) also controls for residence-tract fixed effects, ξ_i, and county-by-period fixed effects, ψ_{cp}.

The estimates are presented in columns (1) and (2) in Table 3.3. Column (1) includes controls, while column (2) does not. The coefficients of $avg\_delay_{ic,t-1}$ in both columns are small in magnitude and statistically indistinguishable from zero. Based on these results, there is no evidence that workers drop out of the labor market in response to increasing congestion along their commute. Since lagged average delay in (1) and (2) is estimated using a time-varying weight (lagged commute flow), change in average delay can be a result of factors other than changes in delay along individual commute routes. Therefore, in columns (3) and (4), lagged average delay is estimated using a time-invariant weight: Delay on a given route out of residence tract i is weighted by the average commute flows on that route between 2010 and 2018. This specification yields quantitatively similar results. The estimated coefficients of $avg\_delay_{ic,t-1}$ are economically small and statistically indistinguishable from zero.

One may be concerned that controls in X_{icp} are bad controls, as the demographic and housing characteristics in a given census tract over a five-year period $p = [t,\ t{+}4]$ can be *affected* by congestion in year t-1. For example, Ossokina and Verweij (2015) find that a decrease in traffic density results in an increase in housing prices in the Netherlands. To mitigate this concern, in Appendix C.4, I control for demographic and housing characteristics from a period before average delay is determined, and the results are quantitatively similar.

3.3.2.2 Earnings of residents in tracts facing increasing commute congestion

Section 3.3.1 finds that workers change workplaces for a less-congested commute as they experience a longer delay along their commutes. This section explores whether changing workplaces for a less-congested commute comes at the cost of less earnings. Specifically, I examine the relationship between the percentage of working residents with high earnings in a given census tract, and the average delay on commute routes out of the tract. The following regression model is estimated:

$$high\_pay_{ict} = \gamma_0 + \gamma_1 avg\_delay_{ic,t-1} + X'_{ic,t-2}\theta + \sigma_i + \zeta_{ct} + v_{ict}, \tag{3.5}$$

where $high\_pay_{ict}$ is the percentage of working residents (with a job in Greater LA) who earn more than \$3,333 per month in residence tract i and county c during year t. $avg\_delay_{ic,t-1}$ is the average delay on commute routes out of residence tract i in year $t-1$, estimated in the same way as in Model (3.4). The vector $X_{ic,t-2}$ includes two sets of controls: 1) five-year estimates of residence tract i's percentage of female, percentage of families with children under six years old, and housing stock over the period [t-6, t-2], and 2) housing price in tract i in year t-2. Model (3.5) also controls for residence-tract fixed effects, σ_i, and county-by-year fixed effects, ζ_{ct}.

The estimates are presented in Table 3.4. Similar to Table 3.3, lagged average delay in (1) and (2) is estimated using time-varying weights, while lagged average delay in (3) and (4) is estimated using time-invariant weights. Columns (1) and (3) report estimates under t=[2011, 2019] without controls, while columns (2) and (4) report estimates under t=[2012, 2019] with controls in the regression. The inclusion of controls in columns (2) and (4) results in the exclusion of t=2011, as it requires mapping five-year estimates of tract level characteristics over 2005-2009 from 2000 tract geography to 2010 tract geography, introducing additional measurement errors in the estimation.

The results suggest that a one-minute longer delay in commutes out of a given tract reduces the percentage of working residents with high earnings in that tract by 0.25% to 0.49%. There are two possible explanations for this finding: First, high-earning residents in high-congestion tracts switch to jobs/offices with a lower pay, in exchange for a less congested commute. This shift does not change the number of working residents in the tract, but reduces the number of residents with high earnings. Second, high earning residents move out of high-congestion tracts for a less congested commute, as they have a high value of time and are therefore less tolerant of congestion. In this case, both the number of working residents and the number of residents with high earnings decrease, leading to a decline in the percentage of working residents with high earnings. Recall the previous finding in

Section 3.3.1 that commuters tend to choose changing workplace over changing residence in response to a more congested commute, which provides suggestive evidence that workers switching to jobs with a lower pay for a less-congested commute might be a more likely explanation.

3.4 Conclusion

This paper studies how workers respond to freeway congestion along their commute to work in Greater LA, one of the most congested metropolitan areas in the United States. First, I develop a novel approach to measuring historical congestion by using open-sourced road network data augmented with historical speeds collected from freeway traffic detectors. Second, pairing congestion data with data on tract-to-tract commute flows constructed from LODES, this study finds that a one-minute increase in delay caused by freeway congestion on a given commute route reduces the commute flow on the route by 0.086 persons on average. In addition, the negative effect of congestion on commute flows decreases as workers' alternative commute routes, achievable through moving or changing workplaces, become more congested. Third, there is suggestive evidence that the decrease in commuters on congested routes more likely results from workers changing workplaces for a less-congested commute, than from workers relocating their residences for a less-congested commute, or dropping out of the labor market. Lastly, the paper uncovers a negative labor market outcome imposed by congestion: workers switching to jobs/offices with a lower pay in exchange for a less congested commute.

This study deepens the understanding of the economic costs of congestion, and provides important policy implications. The finding that commuters are willing to take pay reduction in order to experience less congestion lends support to the adoption of congestion pricing policies. Future research that quantifies this "earnings effects" of congestion avoidance may help gauge commuters' willingness to pay to reduce travel time, and contribute to the design of congestion pricing.

This paper can be extended in a few ways. First, recent studies suggest that women's labor market outcomes might be more susceptible to congestion than men (Le Barbanchon et al., 2021; Farré et al., 2023). With restricted access census data, one can construct tract-to-tract commute flows by gender and study the effects of congestion on *women*'s labor market outcomes specifically. This extension may help shed light on the contribution of traffic congestion to the gender gap in labor force participation and the gender wage gap. Second, recent findings suggest that the volatility of commute times due to congestion might have a more important role in one's labor market outcomes than the commute time itself (Bento et al., 2020; Maloney, 2022). One may extend this paper by exploring the effects of unreliable commute time caused by congestion.

Table 3.1: Summary statistics

(a) Summary statistics: Labor market outcomes

Variable	Unit of Obs.	No. of Tracts/ Tract-Pairs	Sample period	Total Obs.	Mean	SD	Min	Max	Definition	Sources
flow	Tract-Pair × Year	1,365,190	2011-2019	12,286,710	3.54	8.56	0	1456	No. of commuters	A
lfpr	Tract × 5-Year-Period	3,916	[2011-2014] to [2015-2019]	19,480	0.72	0.09	0	1	Civilian labor force participation rate for population 16 to 64 years old	B
high_pay	Tract × Year	3,916	2011-2019	35,244	0.45	0.12	0.12	0.79	% of working residents with earnings > $3,333/month	A

Notes: In terms of data on commute flows (*flow*), among the initial sample of 1,372,819 tract pairs introduced in Section 3.2.1.1, 6,202 tract pairs that involve tracts whose land area is 0 or island tracts are dropped. An additional 1,427 tract pairs are excluded where the algorithm fails to find a route that does not involve toll roads or HOV lanes, resulting in the final sample of 1,365,190 tract pairs. Among 3,925 tracts in Greater LA, 6 tracts whose land area is 0 and island tracts are dropped, resulting in the final sample of 3,916 tracts. Under the "Sources" column, "A" stands for LODES, and "B" stands for ACS 5-year estimates.

(b) Summary statistics: Congestion

Distance (mile)	Avg. travel time (min)		Avg. delay (min)		Δ Avg. delay (min)	% increase in delay
	2010	2018	2010	2018		
≤5	5.46	5.48	0.04	0.05	0.01	32.00%
5-10	12.82	12.96	0.44	0.58	0.14	32.38%
10-20	20.10	20.62	1.47	1.99	0.52	35.62%
20-30	30.17	31.20	2.92	3.95	1.03	35.45%
30-40	40.73	42.19	4.44	5.89	1.46	32.86%
40-50	50.55	52.35	5.26	7.07	1.81	34.37%
>50	78.20	81.03	7.29	10.12	2.83	38.88%

Notes: Table 3.1b presents summary statistics for average delay caused by freeway congestion from residence to workplace between 5 am and 9 am on workdays in 2010 and 2018 for 1,365,190 residence-workplace tract pairs. Summary statistics are presented by the distance of the commute and year. Among the initial sample of 1,372,819 tract pairs introduced in Section 3.2.1.1, 6,202 tract pairs that involve tracts whose land area is 0 or island tracts are dropped. An additional 1,427 tract pairs are excluded where I fail to find a route that does not involve toll roads or HOV lanes, resulting in this final sample of 1,365,190 tract pairs. Both average travel time and average delay are weighted by the average commute flow on each route between 2010 and 2018.

Table 3.2: Commute flows between tract pairs and lagged congestion along commute routes

Dependent Variable: Commute flow	(1)	(2)	(3)	(4)
$delay_{t-1}$	-0.0860***	-0.1732***	-0.1632***	-0.1694***
	(0.0016)	(0.0057)	(0.0054)	(0.0050)
$travel\_time^{exclu\_W}_{t-1}$		2.8091***	2.8193***	3.1235***
		(0.3257)	(0.3848)	(0.3117)
$delay_{t-1} \times travel\_time^{exclu\_W}_{t-1}$		0.0019***	0.0021***	0.0021***
		(0.0001)	(0.0001)	(0.0001)
$travel\_time^{exclu\_R}_{t-1}$		-0.0024	0.0271	0.0143
		(0.0266)	(0.0291)	(0.0284)
$delay_{t-1} \times travel\_time^{exclu\_R}_{t-1}$		0.0013***	0.0015***	0.0017***
		(0.0002)	(0.0002)	(0.0002)
Observations	12,286,629	12,283,164	10,688,921	10,347,607
Tract-Pair FE	Yes	Yes	Yes	Yes
R-Tract-By-Year FE	Yes	Yes	Yes	Yes
W-Tract-By-Year FE	Yes	Yes	Yes	Yes
Excl. Pairs w/ Endogenous Delay	No	No	Yes	Yes
Excl. Public Transit & WFH	No	No	No	Yes

Notes: All columns control for tract-pair fixed effects, residence-tract-by-year fixed effects, and workplace-tract-by-year fixed effects. Column (3) drops commute routes that may experience endogenous congestion by dropping tract-pairs whose commute flows contribute to more than 0.1% of the total commute flows on the freeways they pass through. Column (4) additionally drops tract-pairs with residence tracts where at least 50% of workers work from home or commute to work via public transportation, biking or walking (refer to footnote 21 for details). The regressions are performed using the Stata command "reghdfe," by Correia (2014), which iteratively drops singleton groups (groups with only one observation) to avoid biasing standard errors (Correia, 2015). The number of observations reported is the number after dropping singleton groups. Standard errors clustered at the tract-pair level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.3: Labor force participation rate and lagged congestion

Dependent Variable: Labor force participation rate	(1)	(2)	(3)	(4)
avg_delay_{t-1}	-0.0001	0.00004	0.0006	0.0006
	(0.0008)	(0.0007)	(0.0012)	(0.0011)
Observations	19,083	19,064	19,084	19,065
Controls	No	Yes	No	Yes
Weights	Year-varying	Year-varying	Fixed	Fixed
Residence-Tract FE	Yes	Yes	Yes	Yes
County-By-Period FE	Yes	Yes	Yes	Yes
Excl. Pub. Transit & WFH	Yes	Yes	Yes	Yes
Excl. Pairs w/ Endogenous Delay	Yes	Yes	Yes	Yes

Notes: All columns control for residence-tract fixed effects and county-by-period fixed effects. In all columns, average delay on routes out of a given residence tract is computed after dropping commute routes that may experience endogenous congestion, which are tract-pairs whose commute flows contribute to more than 0.1% of the total commute flows on the freeways they pass through. All columns drop tract-pairs with residence tracts where at least 50% of workers work from home or commute to work via public transportation, biking or walking (refer to footnote 21 for details). In columns (1) and (2), $avg\_delay_{ic,t-1}$ is the average delay in year t-1 on commute routes out of residence tract i weighted by commute flows in year t-1. In columns (3) and (4), $avg\_delay_{ic,t-1}$ is the average delay in year t-1 on commute routes out of residence tract i weighted by average commute flows between 2010 and 2018. "Controls" include percentage of female, percentage of families with children under six years old, ln(housing price), and ln(housing stock). Standard errors clustered at the residence-tract level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.4: Earnings of residents in tracts facing increasing commute congestion

Dependent Variable: % of working residents with high earnings	(1)	(2)	(3)	(4)
avg_delay_{t-1}	-0.0025***	-0.0025***	-0.0039***	-0.0049***
	(0.0004)	(0.0004)	(0.0006)	(0.0006)
Observations	34,361	30,497	34,362	30,497
Controls	No	Yes	No	Yes
Weights	Year-varying	Year-varying	Fixed	Fixed
Residence-Tract FE	Yes	Yes	Yes	Yes
County-By-Period FE	Yes	Yes	Yes	Yes
Excl. Pub. Transit & WFH	Yes	Yes	Yes	Yes
Excl. Pairs w/ Endogenous Delay	Yes	Yes	Yes	Yes

Notes: All columns control for residence-tract fixed effects and county-by-period fixed effects. In all columns, average delay on routes out of a given residence tract is computed after dropping commute routes that may experience endogenous congestion, which are tract-pairs whose commute flows contribute to more than 0.1% of the total commute flows on the freeways they pass through. All columns drop tract-pairs with residence tracts where at least 50% of workers work from home or commute to work via public transportation, biking or walking (refer to footnote 21 for details). In columns (1) and (2), $avg\_delay_{ic,t-1}$ is the average delay in year t-1 on commute routes out of residence tract i weighted by commute flows in year t-1. In columns (3) and (4), $avg\_delay_{ic,t-1}$ is the average delay in year t-1 on commute routes out of residence tract i weighted by average commute flows between 2010 and 2018. "Controls" include percentage of female, percentage of families with children under six years old, ln(housing price), and ln(housing stock). Standard errors clustered at the residence-tract level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

References

Agarwal, S., Han, Y., Qin, Y., & Zhu, H. (2023). Disguised pollution: Industrial activities in the dark. *Journal of Public Economics*, *223*, Article 104904.

Agarwal, S., Lucca, D., Seru, A., & Trebbi, F. (2014). Inconsistent regulators: Evidence from banking. *The Quarterly Journal of Economics*, *129*(2), 889–938.

Agarwal, S., Song, C., & Yao, V. (2020). *Relational contracts in the housing market* (Research Paper No. 3076944). Georgetown McDonough School of Business. Retrieved May 31, 2022, from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3076944

Akbar, P., Couture, V., Duranton, G., & Storeygard, A. (2023). Mobility and congestion in urban india. *American Economic Review*, *113*.

Akbar, P., & Duranton, G. (2017). Measuring the cost of congestion in a highly congested city: Bogotá. *Development Bank of Latin America Working Paper*.

Alexander, M. (2022). *Home charging access in california* (Publication Number: CEC-600-2022-021). California Energy Commission. Retrieved January 15, 2023, from https://www.energy.ca.gov/sites/default/files/2022-01/CEC-600-2022-021.pdf

Anselin, L. (1988). *Spatial econometrics: Methods and models*. New York: Springer.

Anselin, L., Varga, A., & Acs, Z. (1997). Local geographic spillovers between university research and high technology innovations. *Journal of urban economics*, *42*(3), 422–448.

Axsen, J., & Kurani, K. S. (2013). Hybrid, plug-in hybrid, or electric—what do car buyers want? *Energy policy*, *61*, 532–543.

Baer, S. K. (2016a). Public wants more information about conditions at la county restaurants, survey finds. *San Gabriel Valley Tribune*. Retrieved May 27, 2022, from https://www.sgvtribune.com/2016/01/05/public-wants-more-information-about-conditions-at-la-county-restaurants-survey-finds/

Baer, S. K. (2016b). That 'a' grade at your favorite la restaurant will be more meaningful soon. here's why. *San Gabriel Valley Tribune*. Retrieved June 9, 2022, from https://www.sgvtribune.com/2016/06/10/that-a-grade-at-your-favorite-la-restaurant-will-be-more-meaningful-soon-heres-why/

Baer, S. K. (2015). What that restaurant letter grade isn't telling you about health and cleanliness. *San Gabriel Valley Tribune*. Retrieved June 8, 2022, from https://www.sgvtribune.com/2015/06/25/what-that-restaurant-letter-grade-isnt-telling-you-about-health-and-cleanliness/

Baldwin, S., Myers, A., & O'Boyle, M. (2020). *Increasing electric vehicle charging access at multi-unit dwellings: Workshop summary report* (tech. rep.). San Francisco: Energy Innovation: Policy, Technology LLC.

Bedir, A., Crisostomo, N., Allen, J., Wood, E., & Rames, C. (2018). *California plug-in electric vehicle infrastructure projections: 2017-2025* (Publication Number: CEC-600-2018-001). California Energy Commission. Retrieved March 8, 2022, from https://www.nrel.gov/docs/fy18osti/70893.pdf

Beland, L. P., & Brent, D. A. (2018). Traffic and crime. *Journal of Public Economics*, *160*.

Bencsik, P., Lusher, L., & Taylor, R. L. (2023). Slow traffic, fast food: The effects of time lost on food store choice (no. 16036). *Institute of Labor Economics (IZA) Discussion Paper*.

Bennett, V. M., Pierce, L., Snyder, J. A., & Toffel, M. W. (2013). Customer-driven misconduct: How competition corrupts business practices. *Management Science*, *59*(8), 1725–1742.

Bento, A., Roth, K., & Waxman, A. (2020). The value of urgency: Evidence from real-time congestion pricing (no. w26956). *National Bureau of Economic Research*.

Black, D. A., Kolesnikova, N., & Taylor, L. J. (2014). Why do so few women work in new york (and so many in minneapolis)? labor supply of married women across us cities. *Journal of Urban Economics*, *79*.

Boeing, G. (2017). Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, environment and urban systems*, *65*.

Borenstein, S., & Davis, L. W. (2016). The distributional effects of us clean energy tax credits. *Tax Policy and the Economy*, *30*(1), 191–234.

Brent, D., & Beland, L. P. (2020). Traffic congestion, transportation policies, and the performance of first responders. *Journal of Environmental Economics and Management*, *103*, Article 102339.

Brockway, A. M., Conde, J., & Callaway, D. (2021). Inequitable access to distributed energy resources due to grid infrastructure limits in california. *Nature Energy*, *6*(9), 892–903.

Brown, E. (2013). Now on yelp: Los angeles county restaurant health grades. *The Los Angeles Times*. Retrieved September 2, 2020, from https://www.latimes.com/local/lanow/la-me-ln-restaurant-grades-yelp-20131219-story.html

Brueckner, J. K. (2022). Is strategic interaction among governments just a modern phenomenon? evidence on welfare competition under britain's 19th-century poor law. *International Tax and Public Finance*.

Brueckner, J. K. (2003). Strategic interaction among governments: An overview of empirical studies. *International regional science review*, *26*(2), 175–188.

Brueckner, J. K. (1998). Testing for strategic interaction among local governments: The case of growth controls. *Journal of urban economics*, *44*(3), 438–467.

Bryan, S. L., & Aldridge, M. (2020). *Innovation in electric vehicle charging for multi-unit dwellings: Community innovation on-ramp grant final project report* (Prepared for East Bay Community Energy). Santa Cruz: Ecology Action.

Bui, A., Slowi, P., & Lutsey, N. (2021). *Los angeles electric vehicle charging infrastructure needs and implications for zero-emission area planning* (White Paper). Washington, DC: International Council on Clean Transportation.

Calem, P., Kenney, J., Lambie-Hanson, L., & Nakamura, L. (2021). Appraising home purchase appraisals. *Real Estate Economics*, *49*(S1), 134–168.

California Air Resources Board. (2017). *California's advanced clean cars midterm review appendix g: Plug-in electric vehicle in-use and charging data analysis*. Retrieved May 25, 2021, from https://ww2.arb.ca.gov/resources/documents/2017-midterm-review-report

California Air Resources Board. (2018a). *Electric vehicle (ev) charging infrastructure: Multifamily building standards* (tech. rep.).

California Air Resources Board. (2018b). *Low-income barriers study, part b: Overcoming barriers to clean transportation access for low-income residents*. Retrieved January 13, 2023, from https://ww2.arb.ca.gov/sites/default/files/2018-08/sb350_final_guidance_document_022118.pdf

California Governor's Office of Business and Economic Development. (2019). *Electric vehicle charging station permitting guidebook*. Retrieved March 7, 2022, from https://businessportal.ca.gov/wp-content/uploads/2019/07/GoBIZ-EVCharging-Guidebook.pdf

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and applications*. New York: Cambridge University Press.

Canepa, K., Hardman, S., & Tal, G. (2019). An early look at plug-in electric vehicle adoption in disadvantaged communities in california. *Transport Policy*, *78*, 19–30.

Center for Sustainable Energy. (2021). *California air resources board clean vehicle rebate project, ev consumer survey dataset*. Retrieved May 24, 2021, from http://cleanvehiclerebate.org/eng/survey-dashboard/ev

Centers for Disease Control and Prevention. (2019). *Surveillance for foodborne disease outbreaks united states, 2017: Annual report*. Retrieved September 2, 2020, from https://www.cdc.gov/fdoss/pdf/2017_FoodBorneOutbreaks_508.pdf

Chakraborty, D., Bunch, D. S., Lee, J. H., & Tal, G. (2019). Demand drivers for charging infrastructure-charging behavior of plug-in electric vehicle commuters. *Transportation Research Part D: Transport and Environment*, *76*, 255–272.

Conklin, J., Coulson, N. E., Diop, M., & Le, T. (2020). Competition and appraisal inflation. *The Journal of Real Estate Finance and Economics*, *61*(1), 1–38.

Correia, S. (2014). *Reghdfe: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects* (Statistical Software Components S457874) (Revised 21 Aug 2023). Boston College Department of Economics. Retrieved April 1, 2024, from https://ideas.repec.org/c/boc/bocode/s457874.html

Correia, S. (2015). *Singletons, cluster-robust standard errors and fixed effects: A bad mix* (Technical Note). Duke University.

Couture, V., & Handbury, J. (2020). Urban revival in america. *Journal of Urban Economics*, *119*, Article 103267.

Currie, J., & Walker, R. (2011). Traffic congestion and infant health: Evidence from e-zpass. *American Economic Journal: Applied Economics*, *3*.

Davis, D. R., Dingel, J. I., Monras, J., & Morales, E. (2019). How segregated is urban consumption? *Journal of Political Economy*, *127*(4), 1684–1738.

Dee, T. S., Dobbie, W., Jacob, B. A., & Rockoff, J. (2019). The causes and consequences of test score manipulation: Evidence from the new york regents examinations. *American Economic Journal: Applied Economics*, *11*(3), 382–423.

DeShazo, J., Wong, N., & Karpman, J. (2017). *Overcoming barriers to electric vehicle charging in multi-unit dwellings: A westside cities case study* (tech. rep.). Los Angeles: UCLA Luskin Center for Innovation.

Diamond, R., & Persson, P. (2016). The long-term consequences of teacher discretion in grading of high-stakes tests (no. w22207). *National Bureau of Economic Research*.

Drukker, D. M., Egger, P. H., & Prucha, I. R. (2013). On two-step estimation of a spatial autoregressive model with autoregressive disturbances and endogenous regressors. *Econometric Reviews*, *32*(1), 5–6.

Environmental Health, Los Angeles County Department of Public Health. (2011). *Proposed adjustment of public health fees for fy 2011-2012*. Retrieved August 23, 2020, from http://www.lapublichealth.org/eh/docs/FeeStudyPresentationFood.pdf

Environmental Protection Agency, U. S. (2020). *Fast facts: U.s. transportation sector ghg emissions 1990-2019* (EPA-420-F-21-076). EPA: Office of Transportation, Air Quality. Retrieved March 17, 2022, from https://nepis.epa.gov/Exe/ZyPDF.cgi?Dockey=P1013NR3.pdf

Environmental Protection Agency, U. S. (2022). *Sources of greenhouse gas emissions: Transportation sector emissions*. Retrieved March 17, 2022, from https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions#t1fn2

Evans, M. F. (2016). The clean air act watch list: An enforcement and compliance natural experiment. *Journal of the Association of Environmental and Resource Economists*, *3*(3), 627–665.

Farré, L., Jofre-Monseny, J., & Torrecillas, J. (2023). Commuting time and the gender gap in labor market participation. *Journal of Economic Geography*, *23*.

Fu, S., & Viard, V. B. (2018). Commute costs and labor supply: evidence from a satellite campus. *Journal of Economic Geography*, *19*(3), 723–752.

Gershenson, S. (2013). The causal effect of commute time on labor supply: Evidence from a natural experiment involving substitute teachers. *Transportation Research Part A: Policy and Practice*, *54*, 127–140.

Glaeser, E. L., Kolko, J., & Saiz, A. (2001). Consumer city. *Journal of Economic Geography*, *1*(1), 27–50.

Graham, M. R., Kutzbach, M. J., & McKenzie, B. (2014). Design comparison of lodes and acs commuting data products. *Working Papers 14-38, Center for Economic Studies, U.S. Census Bureau.*

Guo, S., & Kontou, E. (2021). Disparities and equity issues in electric vehicles rebate allocation. *Energy Policy*, *154*, Article 112291.

Gutiérrez-i-Puigarnau, E., & van Ommeren, J. N. (2010). Labour supply and commuting. *Journal of urban economics*, *68*.

Haller, P., & Heuermann, D. F. (2016). Job search and hiring in local labor markets: Spillovers in regional matching functions. *Regional Science and Urban Economics*, *60*.

Han, J., Schwartz, A. E., & Elbel, B. (2020). Does proximity to fast food cause childhood obesity? evidence from public housing. *Regional Science and Urban Economics*, *84*, Article 103565.

Heilmann, K. (2018). Transit access and neighborhood segregation. evidence from the dallas light rail system. *Regional Science and Urban Economics*, *73*.

Heitz, A., Wang, Y., & Wang, Z. (2021). Corporate political connections and favorable environmental regulatory enforcement. *Management Science*.

Heuermann, D. F., Assmann, F., vom Berge, P., & Freund, F. (2017). The distributional effect of commuting subsidies-evidence from geo-referenced data and a large-scale policy reform. *Regional Science and Urban Economics*, *67*.

Ho, D. E. (2012). Fudging the nudge: Information disclosure and restaurant grading. *Yale Law Journal*, *122*(3), 574–688.

Ho, D. E., Ashwood, Z. C., & Handan-Nader, C. (2019). New evidence on information disclosure through restaurant hygiene grading. *American Economic Journal: Economic Policy*, *11*(4), 404–28.

Holmes, M. J., Otero, J., & Panagiotidis, T. (2019). Property heterogeneity and convergence club formation among local house prices. *Journal of Housing Economics*, *43*.

Howard, B., Vaidyanathan, S., Cohn, C., Henner, N., & Jennings, B. (2021). *The state transportation electrification scorecard* (Research Report). Washington, DC: American Council for an Energy-Efficient Economy.

Hsu, C. W., & Fingerman, K. (2021). Public electric vehicle charger access disparities across race and income in california. *Transport Policy*, *100*, 59–67.

Hymel, K. (2009). Does traffic congestion reduce employment growth? *Journal of Urban Economics*, *65*.

Ibanez, M. R., & Toffel, M. W. (2020). How scheduling can bias quality assessment: Evidence from food-safety inspections. *Management Science*, *66*(6), 2396–2416.

Jenn, A., Springel, K., & Gopal, A. R. (2018). Effectiveness of electric vehicle incentives in the united states. *Energy policy*, *119*, 349–356.

Jin, G. Z., & Lee, J. (2014). Inspection technology, detection, and compliance: Evidence from florida restaurant inspections. *The RAND Journal of Economics*, *45*(4), 885–917.

Jin, G. Z., & Lee, J. (2018). A tale of repetition: Lessons from florida restaurant inspections. *The Journal of Law and Economics*, *61*(1), 159–188.

Jin, G. Z., & Leslie, P. (2003). The effect of information on product quality: Evidence from restaurant hygiene grade cards. *The Quarterly Journal of Economics*, *118*(2), 409–451.

Jin, G. Z., & Leslie, P. (2019). New evidence on information disclosure through restaurant hygiene grading: Reply. *American Economic Journal: Economic Policy*, *11*(4), 429–43.

Kalmenovitz, J. (2021). Incentivizing financial regulators. *The Review of Financial Studies*, *34*(10), 4745–4784.

Kalnins, A., & Mayer, K. J. (2004). Franchising, ownership, and experience: A study of pizza restaurant survival. *Management Science*, *50*(12), 1716–1728.

Kelejian, H. H., & Prucha, I. R. (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review*, *40*(2), 509–533.

Kelejian, H. H., & Prucha, I. R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics*, *17*(1), 99–121.

Kelejian, H. H., & Prucha, I. R. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*, *157*(1), 53–67.

Khan, H. A. U., Price, S., Avraam, C., & Dvorkin, Y. (2022). Inequitable access to ev charging infrastructure. *The Electricity Journal*, *35*(3), Article 107096.

Kim, D., & Yörük, B. K. (2015). The impact of restaurant smoking bans on dining out expenditures: Evidence from panel data. *Journal of Urban Economics*, *88*, 38–49.

Kim, J. (2022). Does roadwork improve road speed? evidence from urban freeways in california. *Regional Science and Urban Economics*, *93*, Article 103773.

Kim, J., & Long, D. (2024). Working from home, commuting time, and intracity house-price gradients. *Journal of Regional Science*.

Knittel, C. R., Miller, D. L., & Sanders, N. J. (2016). Caution, drivers! children present: Traffic, pollution, and infant health. *Review of Economics and Statistics*, *98*.

Kovács, B., Lehman, D. W., & Carroll, G. R. (2020). Grade inflation in restaurant hygiene inspections: Repeated interactions between inspectors and restaurateurs. *Food Policy*, *97*, Article 101960.

Kuang, C. (2017). Does quality matter in local consumption amenities? an empirical investigation with yelp. *Journal of Urban Economics*, *100*, 1–18.

Kurani, K. S., Caperello, N., & TyreeHageman, J. (2016). *New car buyers' valuation of zero-emission vehicles: California*. Retrieved January 12, 2023, from https://ww2.arb.ca.gov/sites/default/files/classic/research/apr/past/12-332.pdf

Langer, A., & McRae, S. (2014). *Fueling alternatives: Evidence from real-world driving data* (Working Paper). University of Arizona.

Le Barbanchon, T., Rathelot, R., & Roulet, A. (2021). Gender differences in job search: Trading off commute against wage. *The Quarterly Journal of Economics*, *136*.

Lee, J. H., Chakraborty, D., Hardman, S. J., & Tal, G. (2020). Exploring electric vehicle charging patterns: Mixed usage of charging infrastructure. *Transportation Research Part D: Transport and Environment*, *79*, Article 102249.

Lee, L. F. (2004). Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica*, *72*(6), 1899–1925.

Li, S., Tong, L., Xing, J., & Zhou, Y. (2017). The market for electric vehicles: Indirect network effects and policy design. *Journal of the Association of Environmental and Resource Economists*, *4*(1), 89–133.

Lopez-Behar, D., Tran, M., Froese, T., Mayaud, J. R., Herrera, O. E., & Merida, W. (2019). Charging infrastructure for electric vehicles in multi-unit residential buildings: Mapping feedbacks and policy recommendations. *Energy policy*, *126*, 444–451.

Los Angeles County Department of Public Health. (2019). *Fy 2019-2020 public health and financial management fee schedule*. Retrieved August 25, 2020, from http://publichealth.lacounty.gov/eh/docs/LicensePermitFees.pdf

Luco, F. (2019). Who benefits from information disclosure? the case of retail gasoline. *American Economic Journal: Microeconomics*, *11*.

Luo, T., & Stark, P. B. (2015). Nine out of 10 restaurants fail? check, please. *Significance*, *12*(2), 25–29.

Makofske, M. P. (2020a). Disclosure policies in inspection programs: The role of specific deterrence. *Economics Letters*, *196*, Article 109533.

Makofske, M. P. (2019). Inspection regimes and regulatory compliance: How important is the element of surprise? *Economics Letters*, *177*, 30–34.

Makofske, M. P. (2020b). Mandatory disclosure, letter-grade systems, and corruption: The case of los angeles county restaurant inspections. *Journal of Economic Behavior & Organization*, *172*, 292–313.

Makofske, M. P. (2021). Spoiled food and spoiled surprises: Inspection anticipation and regulatory compliance. *Journal of Economic Behavior & Organization*, *190*, 348–365.

Maloney, M. (2022). Why are the labor outcomes of married women better in detroit? the role of long and variable commutes. Retrieved April 24, 2024, from https://drive.google.com/file/d/1WcNRtqDq32MQjzj5OdPTGIV5OGjI_mIB/view

May, P. J., & Wood, R. S. (2003). At the regulatory front lines: Inspectors' enforcement styles and regulatory compliance. *Journal of public administration research and theory*, *13*(2), 117–139.

McDonald, J. F., & Moffitt, R. A. (1980). The uses of tobit analysis. *The review of economics and statistics*, *62*(2), 318–321.

Mo, J. (2023). Heterogeneous effects of urban transport infrastructure on population distribution: The role of educational access. *Regional Science and Urban Economics*, *101*, Article 103902.

National Renewable Energy Laboratory. (2013). *Transportation secure data center*. Retrieved November 3, 2022, from www.nrel.gov/tsdc

National Renewable Energy Laboratory. (2019). *Transportation secure data center*. Retrieved January 23, 2022, from www.nrel.gov/tsdc

Nicholas, M., Hall, D., & Lutsey, N. (2019). *Quantifying the electric vehicle charging infrastructure gap across us markets* (White Paper). Washington, DC: International Council on Clean Transportation.

NuStats, L. L. C. (2013). *2010–2012 california household travel survey final report*. California Department of Transportation. Retrieved March 24, 2024, from https://www.nrel.gov/transportation/secure-transportation-data/assets/pdfs/calif_household_travel_survey.pdf

Ossokina, I. V., & Verweij, G. (2015). Urban traffic externalities: Quasi-experimental evidence from housing prices. *Regional Science and Urban Economics*, *55*, 1–13.

Owens, R. I., Rossi-Hansberg, E., & Sarte, P. D. (2020). Rethinking detroit. *American Economic Journal: Economic Policy*, *12*.

Parsa, H. G., Self, J., Sydnor-Busso, S., & Yoon, H. J. (2011). Why restaurants fail? part ii - the impact of affiliation, location, and size on restaurant failures: Results from a survival analysis. *Journal of Foodservice Business Research*, *14*(4), 360–379.

Peer, S., Knockaert, J., Koster, P., & Verhoef, E. T. (2014). Over-reporting vs. overreacting: Commuters' perceptions of travel times. *Transportation Research Part A: Policy and Practice*, *69*.

Rapson, D. S., & Muehlegger, E. (2021). *The economics of electric vehicles* (Working Paper No. w29093). National Bureau of Economic Research.

Rapson, D. S., & Muehlegger, E. (2018). *Subsidizing mass adoption of electric vehicles: Quasi-experimental evidence from california* (Working Paper No. w25359). National Bureau of Economic Research.

Schiff, N. (2015). Cities and product variety: Evidence from restaurants. *Journal of Economic Geography*, *15*(6), 1085–1123.

Shafran, A. P. (2008). Risk externalities and the problem of wildfire risk. *Journal of urban economics*, *64*(2), 488–495.

Shoag, D., & Veuger, S. (2021). Ban-the-box measures help high-crime neighborhoods. *The Journal of Law and Economics*, *64*.

Shoag, D., & Veuger, S. (2019). Do land use restrictions increase restaurant quality and diversity? *Journal of Regional Science*, *59*(3), 435–451.

Short, J. L., Toffel, M. W., & Hugill, A. R. (2016). Monitoring global supply chains. *Strategic Management Journal*, *37*(9), 1878–1897.

Simon, P. A., Leslie, P., Run, G., Jin, G. Z., Reporter, R., Aguirre, A., & Fielding, J. E. (2005). Impact of restaurant hygiene grade cards on foodborne-disease hospitalizations in los angeles county. *Journal of Environmental Health*, *67*(7), 32–6.

Southwest Energy Efficiency Project. (2022). *Ev infrastructure building codes: Adoption toolkit*. Retrieved March 17, 2022, from https://www.swenergy.org/transportation/electric-vehicles/building-codes

StataCorp. (2019). *Stata spatial autoregressive models reference manual release 16*. Stata Press.

Sweet, M. (2014). Traffic congestion's economic impacts: Evidence from us metropolitan regions. *Urban Studies*, *51*.

Trajkovski, S., Zabel, J., & Schwartz, A. E. (2021). Do school buses make school choice work? *Regional Science and Urban Economics*, *86*, Article 103607.

Tyndall, J. (2021). The local labour market effects of light rail transit. *Journal of Urban Economics*, *124*, Article 103350.

Waldfogel, J. (2008). The median voter and the median consumer: Local private goods and population composition. *Journal of Urban Economics*, *63*(2), 567–582.

Williams, B., & DeShazo, J. R. (2015). Pricing plug-in electric vehicle recharging in multi-unit dwellings: Financial viability and fueling costs. In D. Beeton & G. Meyer (Eds.), *Electric vehicle business models: Global perspectives* (pp. 89–107). Springer International Publishing Switzerland.

Wood, E., Rames, C., Muratori, M., Raghavan, S., & Melaina, M. (2017). *National plug-in electric vehicle infrastructure analysis*. Retrieved March 9, 2022, from https://www.nrel.gov/docs/fy17osti/69031.pdf

Wooldridge, J. M. (2013). *Introductory econometrics: A modern approach* (5th ed.). Cengage learning.

Appendix A

Chapter 1

A.1 Score distributions before and after the policy change

Figure A.1 shows a comparison between score distributions for routine restaurant inspections in LA County before and after the policy change implemented by LADPH in response to the media questioning the credibility of restaurant inspection grades. Figure A.1a is taken from Makofske (2020b). It presents the score distribution before the full implementation of the policy change, using 140,163 routine inspections involving violations conducted from October 1, 2014 to September 30, 2016. Therefore, routine inspections scoring 100 are not included in the dataset. Given this fact, the actual percentage of each score should be smaller than what is plotted in Figure A.1a. Figure A.1b presents the score distribution after the full implementation of the policy change, using 169,174 routine inspections from June 1, 2017 to December 31, 2019. The score distribution has a slightly larger percentage of 90 after the policy change than before the policy change. It is not possible to compare the percentages of A in the two graphs, without knowing the precise percentage of each score at or above 90 in Figure A.1a, but they seem to be similar. This pattern indicates that the grade policy change implemented by LADPH that is targeted to increase grade informativeness failed to have its intended effects.

A.2 Grade inflation misleads diners about restaurants' hygiene conditions: Discussions

A.2.1 Robustness checks

The baseline analysis in Section 1.5.1 uses routine inspections for both restaurants and food markets. To restrict the analysis to restaurant inspections only, the paper keeps only the routine inspections that have a match record in the main restaurant inspection dataset, leaving the analysis with 121,660 inspections (78.16% of the original sample).[1] This also allows the analysis to further control for inspector fixed effects and restaurant type fixed effects, as the main restaurant inspection dataset contains information on the inspector and the type of the restaurant. Figure A.2 plots the coefficients and the 95% confidence intervals of the score dummies for 90 to 100 from the LPM. The economic magnitude and statistical significance of the coefficients are consistent with the ones in Figure 1.2.

A.2.2 Sample selection

The baseline analysis in Section 1.5.1 uses the secondary inspection dataset collected from the Environmental Health Inspection Results Page, which only includes inspection results for the restaurants that are open at the time being. Since the data were collected during September 2021, the secondary inspection dataset suffers from the sample selection issue that

[1] Routine inspections that are dropped are supposed to be food market inspections. This indicates that the majority of the sample used in the baseline analysis are restaurant inspections, which are the focus of this paper.

only restaurants that were still active during September 2021 are included. The parameters of interest are the coefficients of the score dummies, which estimate the difference in the probability of a diner compliant between score category 92 and any other score categories under grade A. The estimates will suffer from the sample selection bias if such differences are different for restaurants that were still open during September 2021 (incumbents) than for restaurants that have closed (exiters). One possible scenario is that incumbents are better at establishing a favorable relationship with the inspectors, and are therefore more likely to have their scores inflated to 90 than exiters. In this case, the difference in the probability of a diner compliant between score category 92 and score category 90 would be larger for incumbents than for exiters, which means the estimated coefficient of the score dummy for 90 in Section 1.5.1 is biased upward. Another possible scenario would downwardly bias the estimated coefficient of the score dummy for 90: Exiters have a larger percentage of independent restaurants than incumbents. Since a low grade would hurt independent restaurants more than chain restaurants, inspectors may be more likely to inflate scores to 90 for independent restaurants than for chain restaurants, causing the difference in the probability of a diner compliant between score category 92 and score category 90 to be smaller for incumbents than for exiters. Unfortunately, this paper is not able to pin down the direction of the sample selection bias or correct it.

However, the estimated coefficients for the incumbents are more relevant to this paper. After all, incumbents are the ones that will continue being visited by diners and the ones whose hygiene conditions will continue having public health implications. Moreover, to the best of my knowledge, this dataset is the only one among publicly available datasets that contains information on complaint investigations.

A.3 Grade inflation discourages hygiene improvements: Details on estimation

A.3.1 Estimating the score of a routine inspection in the absence of grade inflation

This section describes the process to estimate the score that a routine inspection would have received without grade inflation (referred to as the predicted score). This paper focuses on a specific type of grade inflation documented in Makofske (2020b): inspectors marking a 2-point deduction (corresponding to a minor violation) instead of a 4-point deduction (corresponding to a major violation) for discretionary violations when an inspection is on the margin of a higher letter grade. An inspection is considered on the margin if 2-point deductions on all the discretionary violations lead to a higher grade than 4-point deductions on all the discretionary violations.[2] For example, consider an inspection where eight violations of Good Retail Practices (one-point deduction each) and one discretionary violation are found. This inspection is on the margin, as a 4-point deduction on the discretionary violation would lead to a score of 88 and a grade of B, while a 2-point deduction on the discretionary violation would lead to a score of 90 and a grade of A. Since a less-than-warranted deduction on a discretionary violation is assumed to be the only form of score manipulation, an inspection that does not involve discretionary violations is considered free of score manipulation and has the same predicted score as assigned score. If an inspection involves discretionary violations but is not on the margin, it is also considered free of score manipulation, because there is little incentive for an inspector to manipulate scores when manipulation will not lead to a

[2] A grading change implemented in January 2017 complicates this process: If two or more four-point deductions are marked, an additional three points will be deducted. Consider an inspection with one major critical risk violation (four-point deduction) and one discretionary violation. This inspection is on the margin, as its lowest possible score would be 89 (one four-point deduction from a major critical risk violation, one four-point deduction from a discretionary violation, plus a three-point deduction as a result of two four-point deductions), leading to a B grade, while its highest possible score would be 94, leading to an A grade.

better grade. In this case, whether a discretionary violation leads to a 2-point deduction or a 4-point deduction is a truthful representation of a restaurant's hygiene conditions. Exploiting this feature, this paper estimates how violation-level and inspection-level characteristics are correlated with the likelihood of a lesser deduction on a discretionary violation among routine inspections *not on the margin*, and uses the estimated parameters to extrapolate predicted scores for routine inspections *on the margin*. The details are as follows: First, the following equation is estimated for the sample of discretionary violations that belong to routine inspections not on the margin:

$$y_{di} = \alpha_0 + W_d'\eta + Z_i'\phi + u_{di}, \tag{A.1}$$

where y_{di} equals 1 if two points are deducted for a discretionary violation d found in routine inspection i, and equals 0 if four points are deducted. W_d contains dummy variables indicating the health code that discretionary violation d is cited for. Z_i is a vector consisting of the following inspection-specific and restaurant-specific controls: indicators for the month, year, day of the week when the inspection occurs, indicators for the zip code[3] and the type of the restaurant, indicators for the inspector who conducts the inspection, the number of violations of good retail practices, the number of discretionary violations, the numbers of major critical violations and minor critical violations excluding discretionary violations respectively, and the number of permit-suspension violations.

Second, estimated parameters from (A.1) are used to predict y_{di} for the sample of discretionary violations that belong to the routine inspections on the margin. The predicted deduction for the discretionary violation d found in inspection i is 2 points if $\hat{y}_{di} \geq 0.5$, and 4 points if $\hat{y}_{di} < 0.5$, which is then used to derive the predicted score and the predicted grade.

[3] For a restaurant with a zip code that is shared by twenty or fewer restaurants in the main inspection dataset, the paper replaces its zip code with the zip code returned by Google Map.

A.3.2 Robustness checks

In Section 1.5.2, model (1.1) is estimated to study the effect of grade inflation on restaurant hygiene. One caveat to the estimation is that the variable of interest, $inflated$, is subject to measurement error, as whether a routine inspection's grade is inflated to A is unobserved and has to be estimated.

The baseline estimation in Section 1.5.2 compares restaurants estimated to have experienced grade inflation and restaurants estimated to not have experienced grade inflation with *comparable hygiene conditions*. The baseline treatment group includes routine inspections with a score between 90 and 91 whose grades are estimated to be inflated to A. The following two robustness checks are performed where the treatment group is expanded to include restaurants with *better* hygiene conditions than the control group. Assuming that restaurants with better hygiene conditions have a lower probability of a subsequent complaint investigation with violations, the estimated β_1 will be subject to a downward bias, and is therefore a more conservative estimate.

First, the treatment group is expanded to include all routine inspections with a score of 90 or 91 that are on the margin of A, *whether their grades are estimated to be inflated or not*. Compared with the baseline treatment group, this alternative treatment group includes additional restaurants that can be subject to grade inflation (i.e., on the margin of A) but have an estimated grade of A. These additional restaurants are assumed to have better hygiene conditions as their estimated grade is A instead of B. The results are presented in column (2) of Table 1.1, and are quantitatively and qualitatively similar to the results in column (1).

Second, routine inspections that have an assigned grade of A and a predicted grade of B are used as the treatment group. Compared with the baseline treatment group, this alternative treatment group does not restrict the score of the routine inspections to be between 90 and 91. It therefore includes restaurants with higher assigned scores, which may have better hygiene conditions. The estimated β_1 is presented in column (3) of Table 1.1, which is quantitatively and qualitatively similar to the estimated β_1 presented in column (1).

A.4 Possible explanations for grade inflation under repeated interactions between an inspector and a restaurant

Section 1.6.2 discusses one possible motive behind the grade inflation performed by an inspector: inspectors have formed attachment to the restaurants through repeated inspections and are reluctant to give them a grade below A. However, there are other possible explanations for grade inflation under repeated interactions that are discussed in the literature. Such motives may include: inspector becoming increasingly inclined to give the benefit of the doubt to the restaurants in ambiguous situations (Kovács et al., 2020) and inspectors taking bribes from the restaurants. Jin and Lee (2018) study repeated interactions between the inspector and the inspectee in the context of Florida restaurant inspections, where restaurants are neither graded nor scored and only violations are recorded and posted online, but the inspection results are not required to be posted inside an establishment. They argue that the following two mechanisms can best explain why repeated interactions lead to fewer citations: First, a restaurant learns about an inspector's stringency and preferences and targets compliance. Second, an inspector pays less attention the more he/she has inspected a given restaurant. The first mechanism does not really fall into the realm of inspector bias that this paper focuses on, and is not likely to be the dominant cause for the bunching at 90 discussed in this paper, as Section 1.4 has discussed arguments from the existing literature against restaurant targeting 90 as a plausible cause for the bunching at 90. The second mechanism indeed sheds light on another possible inspector bias: diminishing attention. The reason behind diminishing attention can be as benign as inspectors trying to reduce detection costs by skipping areas where no violations were found previously. However, it is unlikely that score inflation under this benign motive happens to cause the bunching at 90. Diminishing attention can also be driven by an inspector's increased unwillingness to "look for" violations as they have formed attachment to the restaurants, which is the motive emphasized in Section 1.6.2.

However different these possible underlying causes for grade inflation may seem, they can all be addressed by more frequent inspector rotation proposed in Section 1.6.2. Therefore, this paper does not distinguish between them and uses "attachment" as an umbrella term for them.

A.5 Threshold model

In Section 1.7, a model is constructed where a restaurant picks sanitation effort to maximize expected profit under inspector uncertainty. In that baseline model, a lenient inspector inflates a restaurant's score no matter what the score would have been based on a restaurant's effort level. However, it would be more intuitive to model a lenient inspector as inflating a restaurant's score only if it is on the margin between two letter grades, as documented in Makofske (2020b). This section introduces a threshold model that models a lenient inspector in this alternative way.

Under the threshold model, for a lenient inspector, inspector uncertainty ε follows different distributions depending on a restaurant's sanitation effort e:

$$\begin{cases} \varepsilon = 90 - e & 88 \leq e \leq 90, \\ \varepsilon \sim tri(-2, 0, 2) & e < 88 \text{ or } e > 92, \\ \varepsilon \sim tri(90 - e, 0, e - 90) & 90 < e \leq 92, \end{cases} \tag{A.2}$$

A lenient inspector would inflate a restaurant's score to 90 as long as the restaurant's sanitation effort qualifies for a score of at least 88 (when $88 \leq e \leq 90$, assessed effort $s = e + \varepsilon = 90$, yielding a score of 90). When a restaurant's effort is above 90 or below 88, a lenient inspector grades no differently than a non-lenient inspector, and is as likely to overassess as to underassess effort. As discussed in Section 1.7.2.1, inspector uncertainty under a non-lenient inspector in the baseline model is given by $\varepsilon \sim tri(-2, 0, 2)$, meaning that a non-lenient inspector can omit or impose additional violations worth up to two points. To ensure consistency with the baseline model, ε should follow $\varepsilon \sim tri(-2, 0, 2)$ as well in the threshold model when a restaurant's effort is above 90 or below 88. However, a restaurant with effort between 90 and 92 can end up with assessed effort below 90 as a result. This violates the implicit monotonicity assumption that if an inspector is lenient enough to inflate the score to 90 for a restaurant with effort between 88 and 90, he should give a score of at least 90 if a restaurant's effort is above 90. Therefore, the following adjustment is made: If $90 < e \leq 92$, $\varepsilon \sim tri(90 - e, 0, e - 90)$. In this case, assessed effort $s = e + \varepsilon \sim tri(90, e, 2e - 90)$, ensuring that a restaurant with effort larger than 90 would end up with assessed effort of at least 90.

Figure A.3a plots the simulated distributions of scores, and Figure A.3b plots the simulated distributions of optimal sanitation efforts under the threshold model. The simulations use the same set of parameters ($\{\alpha_i, \beta_i, c_i, r_i\}_{i=1}^{500}$) as in the simulations under the baseline model, to allow a comparison of the two models. The simulated distributions of both scores and optimal sanitation efforts display similar patterns as the simulations under the baseline model.

In the threshold model, an inspector inflates a restaurant's score only if it is at most two points short of an A, while in the baseline model, an inspector bumps up a restaurant's score no matter what its original score would have been. In this sense, the inspector behavior modeled in the threshold model is closer to the inspector behavior in practice. However, the threshold model is ad hoc and is therefore more complicated than the baseline model.

In the baseline model, though score inflation happens on the entire score spectrum, score inflation only affects a restaurant's revenue (and thus its optimal effort) if it results in grade-crossing, as a restaurant's revenue is assumed to depend on its grade, not score, given the fact that diners can only see the grade of a restaurant. Therefore, the baseline model is still appropriate for studying a restaurant's optimal effort. However, it might not be an ideal model for simulating a restaurant's score, as a lenient inspector in the baseline model inflates scores for every restaurant, as opposed to inflating scores only for restaurants at the grade margin in practice.

Nonetheless, since the simulations from the baseline model are similar to the ones from the threshold model, and given that the baseline model is simpler, it is used in the main text.

A.6 Restaurants' price level data from Yelp.com

The paper uses the name and the address of a restaurant to get its price level from Yelp.com. In some cases, Yelp.com would return a different restaurant than the restaurant being searched. The following algorithm is developed to determine whether the restaurant that Yelp search returns is the same restaurant as the one in the inspection dataset.

The paper utilizes the Python package *fuzzywuzzy* to measure the similarity between the name and address of a given restaurant in the inspection dataset and the name and address Yelp search returns based on the Levenshtein Distance used in string matching, yielding two similarity indices: *name_similarity* and *address_similarity*. The paper also takes the difference between a restaurant's latitude recorded in the inspection dataset and the latitude Yelp search returns (*diff_lat*), and the difference between a restaurant's longitude recorded in the inspection dataset and the longitude Yelp search returns (*diff_long*). The restaurants are sorted by their unique IDs. It is manually checked whether Yelp returns the correct restaurant for the first 450 restaurants, and a value for the match indicator is given, which equals one if the restaurant Yelp returns is correct and equals 0 otherwise. This sample of 450 restaurants is used to estimate a logit model with *name_similarity*, *address_similarity*, *diff_lat* and *diff_long* as regressors, and the estimated coefficients are used to predict the probability of a successful match for restaurants with nonmissing latitude and longitude values. This sample of 450 restaurants is also used to estimate a logit model with only *name_similarity*, *address_similarity* as regressors, so that the estimated coefficients can be used to predict the probability of a successful match for restaurants with missing latitude and longitude values. The estimated match indicator equals 1 if the predicted probability of a successful match is larger than 0.6, and equals 0 otherwise. This algorithm is applied to the sample of next 50 restaurants to test its performance. Out of the 50 observations, the algorithm returns the same estimated match indicator as the match indicator determined manually for 49 observations.

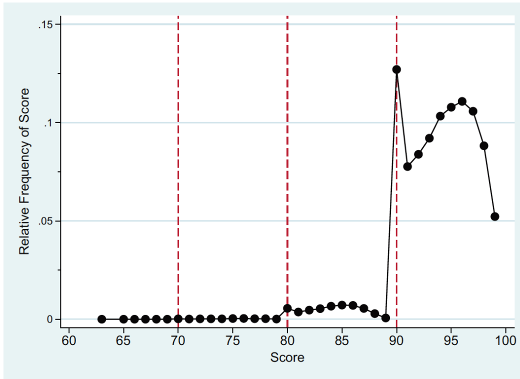
There are 35,792 restaurants to begin with from the full sample, and the paper ends up with 27,262 (76.2%) restaurants for which Yelp.com returns the correct counterparts. Then, the analysis deletes restaurants for which Yelp.com returns a category other than restaurant or food categories,[4] cutting the sample to 21,363 restaurants. These restaurants are primarily deli departments of grocery stores, restaurants that belong to a hotel or resort, restaurants or snack bars for stadiums or theatres. The price level of these restaurants on Yelp focuses mostly on the overall price level of the hotels/grocery stores/stadiums, and thus reflect less

[4]This paper refers to https://www.yelp.com/developers/documentation/v3/all_category_list to determine if a Yelp category belongs to the broader restaurant or food category.

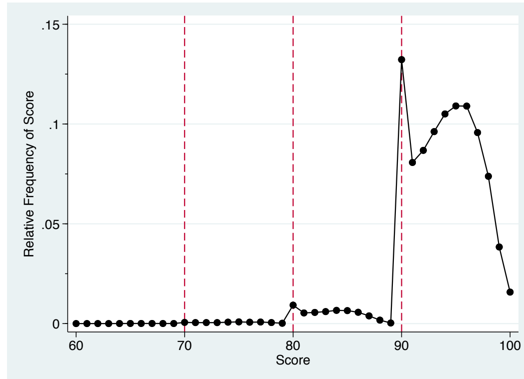
about the restaurants affiliated with them. Lastly, restaurants with missing price levels are dropped. The final sample of 19,968 restaurants is henceforth referred to as the attribute sample. Yelp has four price levels: $, $$, $$$ and $$$$, and values 1,2,3,4 are assigned to the price levels respectively to get the variable *price*.

The attribute sample is used for the extraction of data on capacity level, risk level and price level for the model calibration in Section 1.8.1. Figure A.4 presents the distributions of these attributes. Since the data on price level are only available for restaurants with a matching Yelp profile, one may be concerned that the paper is selecting on restaurants that were still open at the time of the Yelp search (between Feb 15th and 23rd, 2020), as Yelp's search engine tends to filter out establishments that are closed. In this case, calibration of the model uses the sample that overrepresents the attributes of the incumbents and underrepresents the attributes of the exiters. To address this concern, the paper first does a simple comparison between the attribute sample and the full sample in terms of the distributions of capacity levels and risk levels, as the data on these two variables are available for both samples. It turns out that the distributions of capacity level and risk level are similar in the two samples: The proportions of low-risk, moderate-risk and high-risk restaurants in the full sample are 9.56%, 30.62% and 59.82% respectively, compared to 5.02%, 33.02% and 61.96% in the attribute sample. The proportions of 0-30 seats, 31-60 seats, 61-150 seats and 151+ seats restaurants are 52.79%, 23.65%, 17.39% and 6.17% respectively in the full sample, compared to 50.98%, 26.84% and 18.07% and 4.12% in the attribute sample.

A more sophisticated analysis identifies the restaurants that had a matching Yelp profile at the time of the Yelp search but were eventually closed, and compares the distributions of attributes between these restaurants and the restaurants that have not closed. The LA County restaurant and market inspections dataset updated on April 5, 2022 covers inspections conducted from April 1st, 2017 to March 31st, 2022. Therefore, inspection results of the restaurants can be observed for more than two years after the sample period. 4,132 restaurants from the attribute sample do not have inspection records in the updated inspection dataset after the sample period. Since every restaurant is subject to 1 to 3 routine inspections every year, it is safe to deduct that restaurants with no inspection records for more than two years have closed, and therefore their attributes should be representative of the attributes of the exiters. These 4,132 restaurants are referred to as exiters and the remaining restaurants in the attribute sample are referred to as incumbents. Table A.1 compares the distributions of attributes between incumbents and exiters, which are very similar except for the fact that the percentage of low-capacity (0-30 seats) restaurants among the exiters is 6% higher than that among the incumbents. Therefore, using the attribute sample for model calibration is unlikely to introduce sample selection bias into the procedure.



(a) Before: October 1, 2014 to September 30, 2016 (Makofske, 2020b)



(b) After: June 1, 2017 to December 31, 2019 (this paper)

Figure A.1: Score distributions in Makofske (2020b) vs. this paper

Notes: Figure A.1a is Fig. 1 from Makofske (2020b). It plots the score distribution for 140,163 routine inspections involving violations conducted from October 1, 2014 to September 30, 2016, which is before the full implementation of the grade policy change. The dataset does not include inspections that score 100. Figure A.1b plots the score distribution from 169,174 routine inspections from June 1, 2017 to December 31, 2019, which is after the full implementation of the grade policy change. It omits scores under 60 to be consistent with the x-axis of the figure in Makofske (2020b). Six routine inspections end with with scores under 60, accounting for 0.004% of the sample. In both Figure A.1a and Figure A.1b, black dots indicate the relative frequency of each score within that sample.

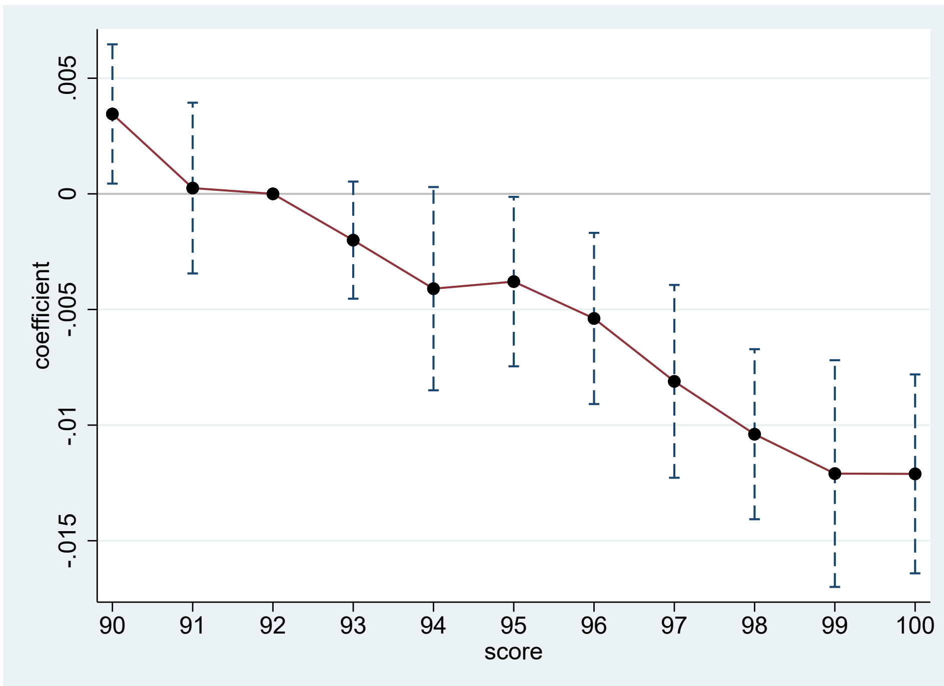
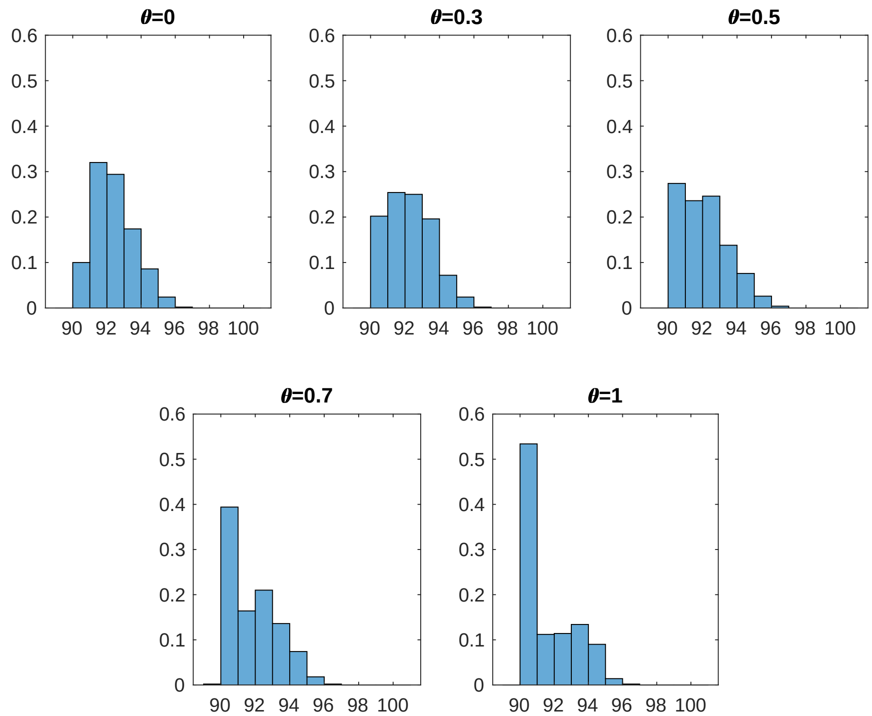


Figure A.2: Comparisons of the probability of a subsequent complaint investigation where violations are found among score groups 90 to 100 using restaurants only

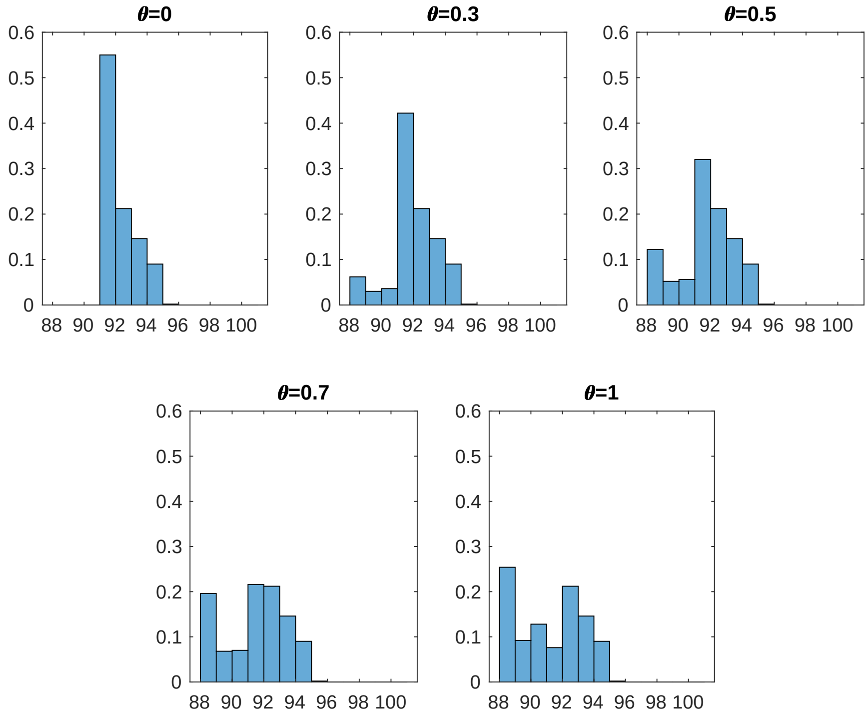
Notes: The dependent variable equals 1 if a routine inspection is followed by a complaint investigation where violations are found, and 0 otherwise. The sample includes 121,660 routine inspections that have a match record in the main restaurant inspection dataset. The y-axis plots the coefficients of the score dummies from the LPM that regresses the dependent variable on score dummies, while controlling for year fixed effects, month-of-year fixed effects, day-of-week fixed effects, city fixed effects, inspector fixed effects and restaurant type fixed effects. Score 92 is the base group and its dummy has a coefficient of 0. The dashed lines give the 95% confidence intervals computed using standard errors clustered at the city level. Clustering standard errors at the inspector level yields qualitatively and quantitatively similar results, which are available upon request.



(a) Simulated distributions of scores under the threshold model

Notes: θ is the proportion of restaurants a representative inspector is lenient to. The simulated scores are integers, each equal to the lower integer of the bar it falls in.

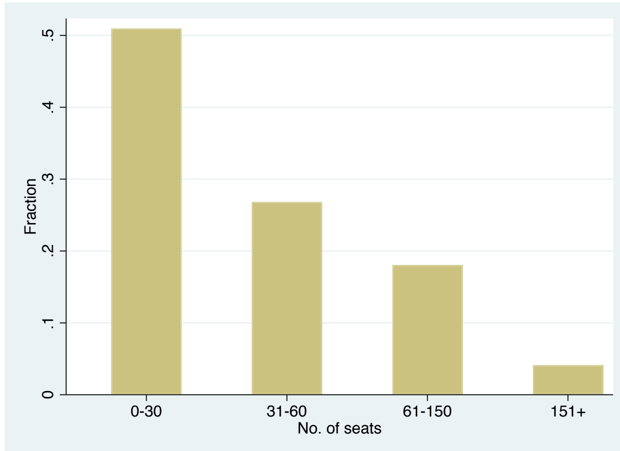
(Continued on next page)



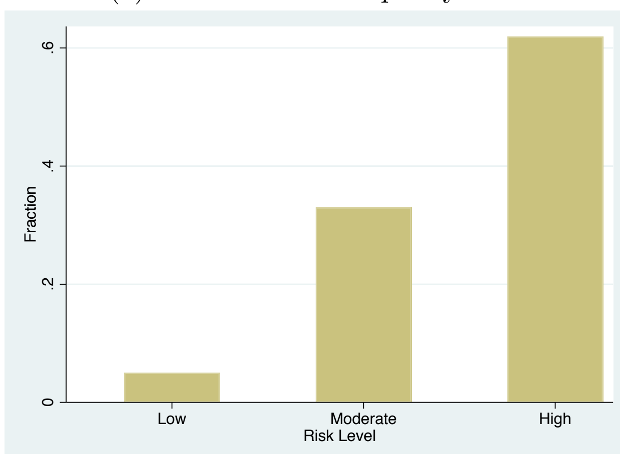
(b) Simulated distributions of optimal efforts under the threshold model

Notes: θ is the proportion of restaurants a representative inspector is lenient to. The optimal efforts are continuous, with values falling in the range of each bar.

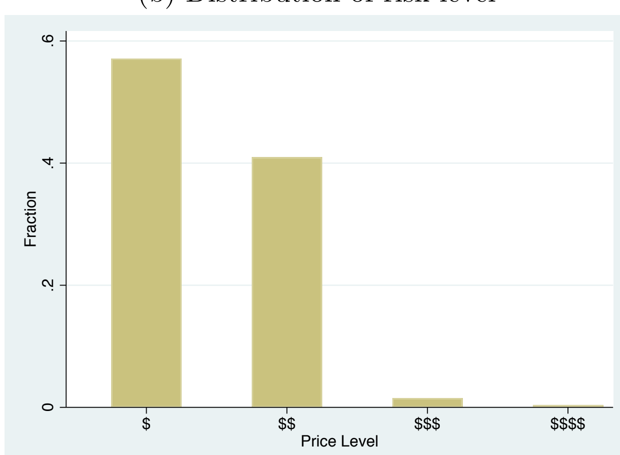
Figure A.3: Simulated distributions of scores and optimal efforts under the threshold model



(a) Distribution of capacity level



(b) Distribution of risk level



(c) Distribution of price level

Figure A.4: Distributions of restaurant attributes

Notes: The sample consists of 19,968 restaurants. The data on a restaurant's capacity level and risk level are from the LA County restaurant and market inspections dataset introduced in Section 1.3.1. The data on a restaurant's price level are from Yelp.com.

Table A.1: Distributions of restaurant attributes among incumbents vs. exiters

risk	incumbent	exiter
low risk	4.94	5.30
moderate risk	33.02	33.03
high risk	62.04	61.67

(a) Distribution of risk level among incumbents vs. exiters

price	incumbent	exiter
$	57.33	56.36
$$	40.65	42.11
$$$	1.58	1.21
$$$$	0.43	0.31

(b) Distribution of price level among incumbents vs. exiters

capacity	incumbent	exiter
0-30 seats	49.61	56.22
31-60 seats	27.15	25.65
61-150 seats	19.09	14.16
151+ seats	4.16	3.97

(c) Distribution of capacity level among incumbents vs. exiters

Appendix B

Chapter 2

B.1 Comparing the findings to those of Hsu and Fingerman (2021) and Khan et al. (2022)

Table B.1 studies the relationship between charger access and race and ethnicity. Table B.1a uses a linear probability model, while Table B.1b uses a Probit model, with the 0-1 dependent variable indicating whether the CBG has a charger. Column (1) has no controls. Column (2) controls for only income and MUD density. The results in column (1) and column (2) are comparable to the findings in Hsu and Fingerman (2021), showing significantly negative correlations between the percentage of Black population and charger existence, and between the percentage of Hispanic population and charger existence, conditional on MUD and income only. However, if we compare CBGs that are different in racial and ethnic composition but are same in other characteristics by adding controls as in columns (3)-(5), racial and ethnic composition are not significantly correlated with the probability that a CBG has chargers. This suggests that the negative correlation between racial and ethnic composition and charger access in columns (1)-(2) can be explained by the other characteristics that are controlled in columns (3)-(5).

B.2 The construction of MUD and facility variables

B.2.1 MUDs

There are four potential ways to quantify MUDs: by the number of MUD buildings, by the number of units in MUDs, by the number of bedrooms in MUDs, and by the total square footage of MUDs. First, it is not possible to count MUD buildings as some MUD buildings appear as one parcel, while for other MUD buildings, each of its units is counted as one parcel. Parcels in the same MUD buildings can have different coordinates and addresses, making it impossible to identify which group of parcels should be counted as one MUD building. Moreover, quantifying MUDs by the number of MUD buildings fails to account for differences in the building size and height. Second, a unit can have different numbers of bedrooms, and can thus accommodate different numbers of people. Therefore, quantifying MUDs by the number of units can mask the variations in the number of people a unit can hold. Quantifying MUDs by the number of bedrooms can potentially solve this problem. However, an appraiser of LA County indicated that for MUDs, the count of bedrooms in the assessor records can be inaccurate. Therefore, the total square footage ends up as the most accurate way to quantify MUDs and is thus used in this paper.

B.2.2 Hotels

The assessor dataset divides hotels into two types: hotels with under 50 rooms and hotels with 50 rooms and over. Based on the types of hotels that have EV chargers, it appears that EV chargers are more likely to be located at higher-end hotels. Since the number of rooms is a proxy for how high-end a hotel is, hotels are quantified by counting the number of hotels with 50 rooms and over with unique coordinates in the assessor dataset.[1]

B.2.3 Office buildings

The "office building" property use category in the assessor dataset includes standalone office buildings and office and residential combinations. Judging from the types of office buildings that have EV chargers, it appears that larger companies are more likely to install EV chargers near their buildings. This is consistent with the fact that larger companies try to go green by offering EV charging to their employees. Therefore, only standalone office buildings are included in the sample, as offices in the office and residential combinations usually belong to small businesses. Office buildings are quantified by aggregating the total square footage instead of counting the number of buildings. This is because an office building in downtown typically has more floors than the one in the suburbs, and the total square footage can account for that difference.

B.2.4 Retail

Parcels in the assessor dataset whose property use is "department store," "shopping center" or "supermarket with 12,000 square feet or more" are included, with the total square footage aggregated. These parcels cover shopping centers/malls and larger grocery stores. Though not as popular as shopping centers/malls as a location for EV chargers, grocery stores are the 11th most popular EV charger location in LA County. Therefore, in this paper, shopping

[1]In the assessor dataset, a single property (e.g., a hotel) can be split into multiple parcels. According to an appraiser, this may be requested by taxpayers, cities or for mapping reasons. Therefore, counting parcels with unique coordinates ensures that the same property is not counted multiple times. On the other hand, dropping parcels with duplicate coordinates is unnecessary when quantifying a type of property by aggregating the square footage. This is because when a property is divided into multiple parcels, its total square footage is divided too. Therefore, aggregation of the square footage of all the sub-parcels is needed to get the total square footage of the property.

centers/malls and larger grocery stores are grouped together and called retail.

B.2.5 Car dealers

Parcels are counted whose property use codes are “New Car Sales and Service” and “Used Car Sales.” The count covers the number of parcels with unique coordinates.

B.2.6 Government offices

Data on government offices include city halls, county offices and government offices. Since some government offices can appear in more than one dataset, and some government offices may have the same locations, government offices with unique coordinates are counted.

B.2.7 Schools

Data on schools include public high schools, school district offices and colleges and universities. Elementary schools and middle schools are not counted as EV chargers are rarely located there. Schools with unique coordinates are counted.

B.2.8 Transportation stations

Transportation stations include Metro stations, Metrolink stations and Amtrak stations. The dataset does not include certain metro stations as it has not been updated since 2013. Therefore, the dataset is supplemented using information from the Metro website.[2]

[2] E Line, L Line and G Line have missing stations in the dataset. The supplement data on the stations are from https://www.metro.net/riding/guide/E-line/, https://www.metro.net/riding/guide/l-line/ and https://www.metro.net/riding/guide/g-line/.

B.3 Accessible distance from chargers

So far, there is no existing literature on how far people are willing to travel for EV charging.[3] Relevant information comes from three sources. First, according to an informant from the City of Los Angeles, the accessible distance to an L2 charger is about 5 to 10 minutes of walking distance, which is about 0.25 to 0.5 miles. Second, the City of Torrance in California launched the "One Mile, One Charger Project" to facilitate the expansion of EV infrastructure throughout the city so that an EV driver is never more than one mile from a charging station within the city.[4] This suggests that policymakers consider one mile an accessible distance to the chargers. Third, 2019 California Vehicle Survey (National Renewable Energy Laboratory, 2019) includes a survey for PEV owners. The survey asks one how far the closest public charger is from one's home, how many times one has used a public charger and the types of public chargers one has used in the month preceding the survey. The paper looks at respondents who have exclusively used L2 chargers for public charging and have used these chargers at least three times in the month preceding the survey. There are 55 of them. 30 users report the distance of the closet public charger in miles: one mile for 13 of them, two miles for 7 of them, with the maximum distance being 11 miles. 25 of them report the distance in minutes of driving time: less than 5 minutes for 13 of them, 6 to 10 minutes for 7 of them, with the maximum driving time being 20 minutes. EV owners are thus willing to travel a longer distance for charging in the survey than according to the informant. However, we should be cautious about the results from the survey because of the small sample size. Moreover, how far one is willing to travel for a charger also depends on one's charging style. If one charges one's car overnight at the charging station, one needs to walk home and walk back to the car the next day. In this case, the accessible distance to a charger would be much smaller than if one drives to do grocery shopping and leaves the car

[3] The literature on willingness to travel for fuels has focused on gas (e.g., Langer and McRae (2014)), and has yet to expand to alternative fuels.

[4] Source: https://www.torranceca.gov/our-city/community-development/sustainability/one-mile-one-charger-project.

for charging when one is shopping.

B.4 Details on the back-of-the-envelope calculation of MUD charger-to-PEV-ratio

B.4.1 Estimating the PEV ownership rate among MUD residents

The PEV ownership rate among MUD residents is estimated using data from the 2019 California Vehicle Survey (National Renewable Energy Laboratory, 2019).[5] The sample is restricted to include only households living in the following three types of housing: "building with 2-4 apartments/condos/studios/rooms," "building with 5-19 apartments/condos/studios/rooms," and "building with 20 or more apartments/condos/studios/rooms." There are 924 households in this sample, and the following two methods are used to measure the PEV ownership rate among MUD residents. The first method uses the following two survey questions: (1) "Has your household ever owned or leased a plug-in hybrid electric vehicle (PHEV)?", and (2) "Has your household ever owned or leased a fully electric vehicle (also called a battery electric vehicle, or BEV)?". If the household answers "yes" to question (1), then that household is counted as having one PHEV. If the household answers "yes" to question (2), then that household is counted as having one BEV. 19 households answered "yes" to question (1), and 18 households answered "yes" to question (2). Therefore, the PEV ownership rate among MUD residents is 37 vehicles per 924 households, which is 0.04 vehicles per household. This measure can both underestimate and overestimate the PEV ownership rate among MUD residents. It can underestimate the PEV ownership rate because if a household has ever owned or leased a PEV, it may have owned or leased multiple PEVs, whereas it is counted as owning or leasing one PEV. It can overestimate the PEV ownership

[5]The term PEV *ownership* rate also counts PEVs that are *leased* by MUD residents.

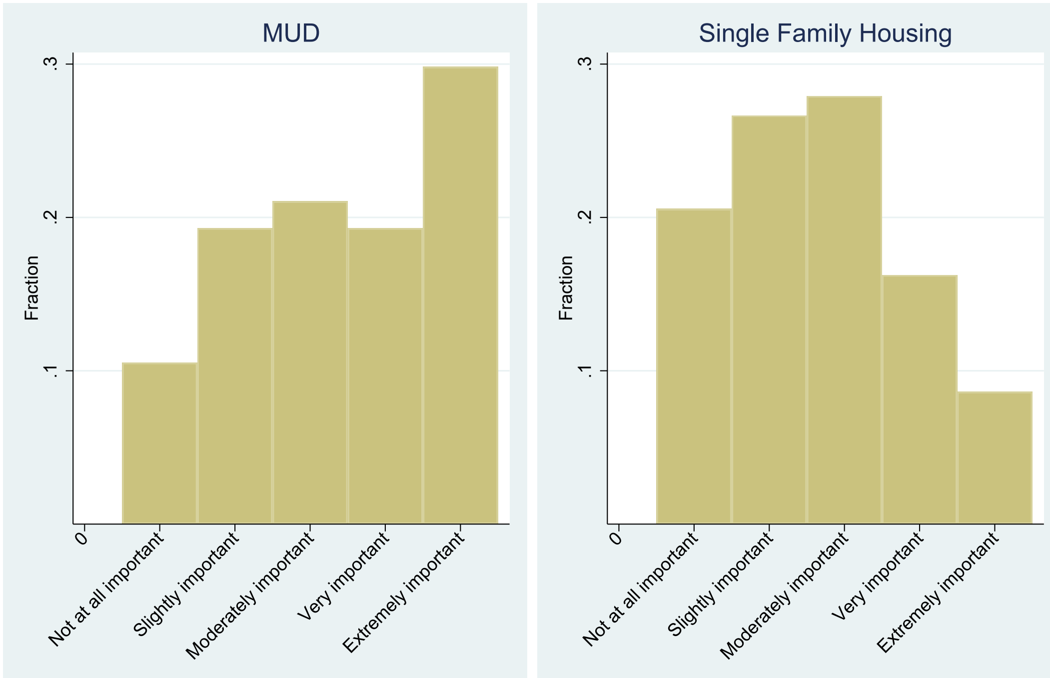
rate because we are counting a household as owning or leasing a PEV even when they used to own or lease a PEV and did not have a PEV when they answered the survey.

The second method uses the survey question that asks the household respondent to list the fuel type of every vehicle in the household. Based on this response, there are 25 PHEVs and 37 BEVs in total. Therefore, the PEV ownership rate among MUD residents is 62 vehicles per 924 households, which is 0.067 vehicles per household.

This is how the range of PEV ownership rate among MUD residents in Section 2.4.2.1 is derived: 0.04 to 0.067 vehicles per household.

B.4.2 Mapping increase in MUD density onto increase in MUD units

Assume that there is monotonic mapping from MUD density to the number of MUD units. This means that if CBG i has a larger MUD density than CBG j, then CBG i also has more MUD units than CBG j. Only 5.16% of the CBGs in the sample have MUD density above 1,000 square feet per capita. The 94th percentile of the number of MUD units is 645, and the 95th percentile of the number of MUD units is 693. That is why Section 2.4.2.1 states that a 1,000 square feet per capita increase in MUD density is approximately equivalent to 645-700 increase in MUD units.



(a) (b)

Figure B.1: How important is the availability of public charging? MUD residents vs. Single family housing residents

Notes: Figure B.1 plots the distributions of answers among MUD residents (Figure B.1a) and single family housing residents (Figure B.1b) respectively to the following question: When deciding to purchase your PEV, how important was the availability of public charging in your decision? Figure B.1a consists of a sample of 57 PEV owners who live in the following three types of housing: "building with 2-4 apartments/condos/studios/rooms," "building with 5-19 apartments/condos/studios/rooms," and "building with 20 or more apartments/condos/studios/rooms." Figure B.1b consists of a sample of 394 PEV owners who live in the following two types of housing: "single family house not attached to any other house," and "single family house attached to one or more houses (townhouse, duplex, triplex) each with separate entry."

Table B.1: Charger access disparities in terms of race and ethnicity

(a) Linear Probability Model

	(1)	(2)	(3)	(4)	(5)
Black	-0.110***	-0.079**	-0.038	0.029	0.033
	(0.030)	(0.033)	(0.033)	(0.050)	(0.051)
Asian	-0.080***	-0.020	-0.053*	0.071	0.095**
	(0.031)	(0.032)	(0.032)	(0.044)	(0.048)
Other Races	0.044	-0.011	-0.004	-0.031	-0.032
	(0.034)	(0.033)	(0.032)	(0.038)	(0.039)
Hispanic	-0.207***	-0.079***	-0.043	0.019	0.026
	(0.023)	(0.029)	(0.037)	(0.045)	(0.047)
Observations	6,373	6,373	6,373	6,373	6,167
Income and MUD	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes
Neighborhood Controls	No	No	No	Yes	Yes
Exclude Border CBGs	No	No	No	No	Yes

(b) Probit Model

	(1)	(2)	(3)	(4)	(5)
Black	-0.096***	-0.078**	-0.037	0.025	0.030
	(0.030)	(0.035)	(0.036)	(0.051)	(0.052)
Asian	-0.062**	-0.016	-0.044	0.072*	0.092**
	(0.027)	(0.029)	(0.030)	(0.040)	(0.042)
Other Races	0.054	-0.003	-0.003	-0.045	-0.047
	(0.040)	(0.039)	(0.037)	(0.043)	(0.044)
Hispanic	-0.209***	-0.091***	-0.045	0.030	0.036
	(0.025)	(0.029)	(0.038)	(0.046)	(0.047)
Observations	6,373	6,373	6,373	6,373	6,167
Income and MUD	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes
Neighborhood Controls	No	No	No	Yes	Yes
Exclude Border CBGs	No	No	No	No	Yes

Notes: The dependent variable is an indicator variable of whether a CBG has a charger. It equals one if a CBG has chargers and equals zero otherwise. Table B.1a reports coefficients from the LPM. Table B.1b reports average marginal effects derived from the probit model. Column (2) controls for only income and MUD density. Column (3) controls full CBG characteristics. Column (4) controls for full CBG characteristics and full neighboring characteristics. The coefficients of column (4) in Table B.1a are the same as those of column (2) in Table 2.2. The coefficients of column (4) in Table B.1b are the same as those of column (5) in Table 2.2. Column (5) controls for full CBG characteristics and full neighboring characteristics, and also excludes CBGs at the border of LA County. The coefficients of column (5) in Table B.1a are the same as those of column (3) in Table 2.2. The coefficients of column (5) in Table B.1b are the same as those of column (6) in Table 2.2. *White* (percentage of the White population), *Below High* (percentage of the population without a high school diploma) and *Neighbor White* (percentage of the White population in neighboring CBGs) are omitted as base groups. The population-weighted inverse-distance-squared matrix is used to aggregate neighbor characteristics. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Appendix C

Chapter 3

C.1 Details on estimating congestion

C.1.1 Cleaning the speed data from stations

I access data on hourly average speed (from "Station Hour" datasets) from 2010 to 2019 for stations in District 7 (LA County and Ventura County), District 8 (Riverside County and San Bernadino County), and District 12 (Orange County) using the "Data Clearinghouse" tool. The dataset includes each station's ID, type (e.g., mainline, HOV), its reported hourly average speed, and the quality of the speed data. I keep only stations on the mainline. PeMS receives data points from each station every 30 seconds, which are then aggregated into hourly average speed. However, there can often be gaps among these 30-second data points when detectors malfunction, stop working, or cease sending data. In this case, PeMS estimates the missing data through imputation. I only keep hourly average speed constructed with 80% of imputed data points or less, following Kim (2022). Since this paper focuses on congestion during morning rush hours, I keep hourly speed between 5 am and 9 am on weekdays, and compute their average for each year.

Information on the location of traffic stations is required to match freeway edges to traffic stations. Longitudes and latitudes of traffic stations are included in separate "Station Metadata" datasets, which I use to construct a yearly station-coordinate dataset. "Station Metadata" datasets are updated whenever there are changes to station-specific variables, and coordinates of a station can change in the middle of the year. In the case where a station has multiple coordinates throughout the year, I let the coordinate of that year be the coordinate with the longest duration. The yearly station-coordinate dataset is then merged with the station-speed dataset.

C.1.2 Availability of freeway speed

Column "Sample" in Table C.1 presents the total length of freeway edges in the OpenStreetMap for which historical speeds are utilized (henceforth, coverage length) across freeways. The coverage length from 2010 to 2019 stays the same for a given freeway-direction, except that there are no historical speed data for the following freeway-year pairs: SR74 (2010-2011), SR90 (2010-2011), SR142 (2010-2012), SR142-E (2017), SR170 (2019). The unavailability of historical speed data on these freeway-year pairs is either due to a lack of traffic stations, or poor speed data quality from the stations. For each freeway-direction, its coverage length is compared with its full length (column "Full") within Greater LA.[1] For the majority of the freeways, historical speeds from traffic sensors are incorporated into all the freeway segments, referred to as the freeway "having a full coverage" in column "Coverage."[2]

[1]The full length of a freeway within Greater LA is estimated using the California Enhanced National Highway System (NHS) line feature class available at https://gisdata-caltrans.opendata.arcgis.com/datasets/1f71fa512e824ff09d4b9c3f48b6d602_0/about, except for SR74 and SR90. Caltrans has relinquished parts of SR74 and SR90 to local jurisdictions, which do not show up on the NHS line feature class provided by Caltrans. Therefore, the full length of SR74 and SR90 within Greater LA is computed using the 2019 road shapefiles available from the Census Bureau website at https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.2019.html.

[2]There is a slight discrepancy between the coverage length and the full length of a freeway, even when the freeway has a full coverage. This is because the two lengths are estimated using different data sources.

For freeways that are partially incorporated with historical speeds, it is either because vehicle detector stations are not available for parts of the freeway (partial censor coverage), or because the parts of freeway with tolls are restricted to have zero speed (not utilized by commuters), as mentioned in Section 3.2.2.

C.1.3 Free-flow speed

For a given year, I take the median of average speed between 10 pm and 11 pm for each station. I then take the median of the median speeds from 2010 to 2019 for each station, which is an estimate of the free-flow speed for that station. To compute the free-flow travel time, $travel\_time^0_{ij}$, the speed of a freeway edge is set to be the free-flow speed of its matching station. The free-flow speed is smaller than the maximum of annual average morning rush-hour speeds from 2010 to 2019 for 1,041 (20.6%) of the freeway edges. However, $delay_{ijt}$ is supposed to be nonnegative, as it should take longer to travel under traffic than without traffic. Therefore, the free-flow speed on these freeway edges is revised to be the maximum of annual average morning rush-hour speeds from 2010 to 2019.[3]

C.1.4 Estimating route and travel time

For each edge, I compute edge travel time by dividing edge length by edge speed. 20 seconds are added to the edge travel time if there is a traffic light on the edge, and 5 seconds are added if there is a stop sign on the edge. I then use the Python package "NetworkX" to find the minimum travel time between the centroid of census tract i and the centroid of census tract j, which is the estimated travel time between census tract i and census tract j. Since

The former is estimated using the OpenStreetMap, and the latter is estimated using the NHS line feature class from Caltrans.

[3]For 60.13% of the freeway edges whose free-flow speeds are revised, the difference between the free-flow speed and the maximum of annual average morning rush-hour speeds is below 1 mph. The difference is below 2 mph and 5 mph for 82.61% and 99.42% of the freeway edges respectively.

the estimation of travel time is required for 1,372,819 tract pairs from 2010 and 2018, and for the baseline free-flow scenario, the process can be computationally intensive. "NetworkX" reduces the computation time, as it computes the minimum travel time without returning the detailed route data.

The construction of Table 3.1b and the estimation of the share of total traffic each tract-pair commute flow contributes to in Section 3.3.1 require the detailed commute route for 1,372,819 tract pairs. The Python package "OSMnx" is used to find the route that minimizes travel time for each tract pair in 2015, which returns all the edges on the route. I only estimate the route for one year, and assume that the routes in other years are the same, as route estimation is computationally intensive.

C.2 Reliability of the historical travel time estimation

The CHTS collected one day's travel information from a sample of household residents in California from February 1, 2012 through January 31, 2013. The members of sampled households used a travel diary to record travel information for a pre-assigned 24-hour period, including trip purpose, places visited during a trip, arrival time, departure time, etc. (NuStats, 2013). In addition to the above information, the survey dataset also includes trip duration (henceforth self-reported travel time) calculated based on arrival time and departure time. Though correspondents report the precise addresses of trip origins and destinations, the smallest geographic unit of the publicly accessible survey data is the census tract. However, the dataset does include trip distances estimated through Google Maps (henceforth CHTS travel distance) using the *precise* addresses of trip origins and destinations. I construct a sample of commute trips within Greater LA that are recorded in a travel diary, with the mode of transport being auto, van or truck. I further limit the sample to morning commutes whose arrival times are between 5:00 am and 11:15 am. Commute trips using toll roads or

HOV lanes are excluded from the sample.

To estimate the travel time for these commute trips, I first find the halfway timestamp between the trip start time and the trip end time. The speed of a freeway edge on OpenStreetMap is then set to be the average speed recorded by its matching station in the hour where the timestamp falls in. For example, for a trip that starts at 8:50 am and ends at 10:10 am (halfway timestamp being 9:30 am) on August 2nd, 2012, I use the average hourly speed recorded by traffic stations between 9 am and 10 am on the same day to construct freeway edge speeds. The travel time and travel distance are estimated by finding the route with the minimum travel time between the centroid of the residence tract and the centroid of the workplace tract using the OpenStreetMap.

Before comparing the estimated distance and travel time with those recorded in the survey, I exclude one trip with a reported distance over 350 miles, and trips with no reported distance. An additional trip is excluded where I fail to find a route that does not involve toll roads or HOV lanes. The final sample consists of 3,374 commute trips.

Figure C.3a compares the estimated travel distance using the OpenStreetMap augmented with historical freeway speeds with the travel distance provided by CHTS. In addition, column (1) of Table C.2 presents the coefficient (1.003) when regressing estimated travel distance on the CHTS travel distance. Both results indicate that the estimated distance is very similar to the distance provided by CHTS, which supports the use of augmented OpenStreetMap as a routing device and indicates that centroids of residence and workplace tracts are good proxies of the actual locations of the residence and workplace.

Column (2) of Table C.2 presents the results when regressing estimated travel time on self-reported travel time. Though the two variables are significantly correlated, the coefficient (0.542) indicates that estimated travel time is only about half of the self-reported travel time. This result could be driven by respondents' misreporting of travel time. To identify self-reported travel time that are likely subject to reporting errors, I predict travel time based on distance using the following model:

$$reported\_time_i = \beta_0 + \beta_1 CHTS\_distance_i + \beta_2 CHTS\_distance_i^2 + \varepsilon_i, \qquad \text{(C.1)}$$

where i denotes a commute trip, $reported\_time_i$ is self-reported travel time, and $CHTS\_distance_i$ is trip distance provided by CHTS that is estimated using the *precise* addresses of trip origins and destinations

Figure C.3b compares the estimated travel time with the self-reported travel time after excluding self-reported travel time that are out of the 95% confidence interval of the predicted travel time. Column (3) of Table C.2 presents the result of regressing estimated travel time on self-reported travel time using the same sample, while column (4) of Table C.2 excludes self-reported travel time that are out of the 90% confidence interval of the predicted travel time. The estimated travel time is closer to the self-reported travel time after excluding trips likely subject to misreporting. However, the self-reported travel time is still larger than the estimated travel time by a factor of 1.45 (1/0.6887), which at first sight may challenge the validity of this paper's methodology of estimating historical travel time. However, Peer et al. (2014) find that reported travel times of morning commutes are overstated by a factor of 1.5 when compared to actual travel time in an experiment in Netherlands.[4] Their finding suggests that the gap between estimated travel time and self-reported travel time found in this paper can simply be interpreted as a result of commuters' over-reporting of travel times,

[4]There are two main differences between Peer et al. (2014) and CHTS in terms of study subject and design. First, Peer et al. (2014) focus on Dutch drivers, while CHTS focuses on Californian drivers. Second, Peer et al. (2014) ask the study participants to report the *average* travel time of their 20 most recent morning commute trips, while CHTS asks the study participants to report the departure time and arrival time of their trips *on a given day*, from which their travel time is computed. Peer et al. (2014) propose numerous causes for over-reporting. Given the difference in the context and design between Peer et al. (2014) and CHTS, the following explanations can be applied to over-reporting of commute time in CHTS: First, respondents may have recorded their departure and arrival time in such a way that the travel time also have included the time it took them to walk from their home to their car, as well as the time it took them to walk from their car to their workplace. Second, the respondents may have recorded their departure and arrival time in such a way as to overstate their commute time in hope of motivating the implementation of policies targeted at decreasing travel times.

instead of as evidence against the validity of the methodology of estimating historical travel time.

C.3 Control variables

C.3.1 Percentage of female

Data on tract-level gender composition are from Table B23001 of ACS 5-year estimates. It is measured by the percentage of female among civilians 16 to 64 years old.

C.3.2 Percentage of families with children under six

Data on tract-level percentage of families with children under six are from Table B11003 of ACS 5-year estimates.

C.3.3 Housing price

Data on housing price are constructed using Zillow Home Value Index (ZHVI).[5] The ZHVI, available at the zip-code level, are mapped into 2010 tract geography using the HUD-USPS ZIP Code Crosswalk[6] supplied by the U.S. Department of Housing and Urban Development. The ZHVI data are monthly time series, and the yearly ZHVI is computed as the monthly average. Five-year estimates of ZHVI are computed by taking the average of yearly ZHVIs over five years, which are used in regression (3.4).

[5] Accessed from https://www.zillow.com/research/data/ on May 1, 2024.
[6] Accessed at https://www.huduser.gov/portal/datasets/usps_crosswalk.html on May 26, 2024.

C.3.4 Housing stock

Data on tract-level total owner-occupied and renter-occupied units are from Table B25003 of ACS 5-year estimates.

C.4 Robustness checks

In Model (3.4) in Section 3.3.2.1, one concern is that the control variables in X_{icp} are themselves outcomes of the regressor of interest, congestion. In the following revised model, I control for demographic and housing characteristics from a period before average delay is determined:

$$lfpr_{ic,[t,t+4]} = \beta_0 + \beta_1 avg\_delay_{ic,t-1} + X'_{ic,[t-6,t-2]}\delta_1 + \delta_2 h_{ic,t-2} + \xi_i + \psi_{c,[t,t+4]} + u_{ic,[t,t+4]}, \quad \text{(C.2)}$$

where t = 2012, 2013, 2014, 2015. The vector $X_{ic,[t-6,t-2]}$ includes controls for residence tract i's percentage of female, percentage of families with children under six years old, and housing stock over period $[t-6, t-2]$. Variable $h_{ic,t-2}$ is the housing price in residence tract i in year t-2.[7]

The estimates are presented in columns (1) and (2) of Table C.3. Similar to results presented in Table 3.3, the estimated coefficients of $avg\_delay_{ic,t-1}$ are economically small and statistically indistinguishable from zero.

[7]Tract-level data on percentage of female, percentage of families with children under six years old, and housing stock are only available as 5-year estimates, while tract-level data on housing price are available yearly. Refer to Appendix C.3 for more details.

Figure C.1: Freeway edges in OpenStreetMap

Notes: This figure displays select freeway edges in OpenStreetMap on I-5 South against the navigation base map provided by ArcGIS Pro. Each black line with arrow represents a freeway edge, with the arrow pointing to the direction of the freeway. Though freeway segments have curvature, they are illustrated with straight lines for simplicity.

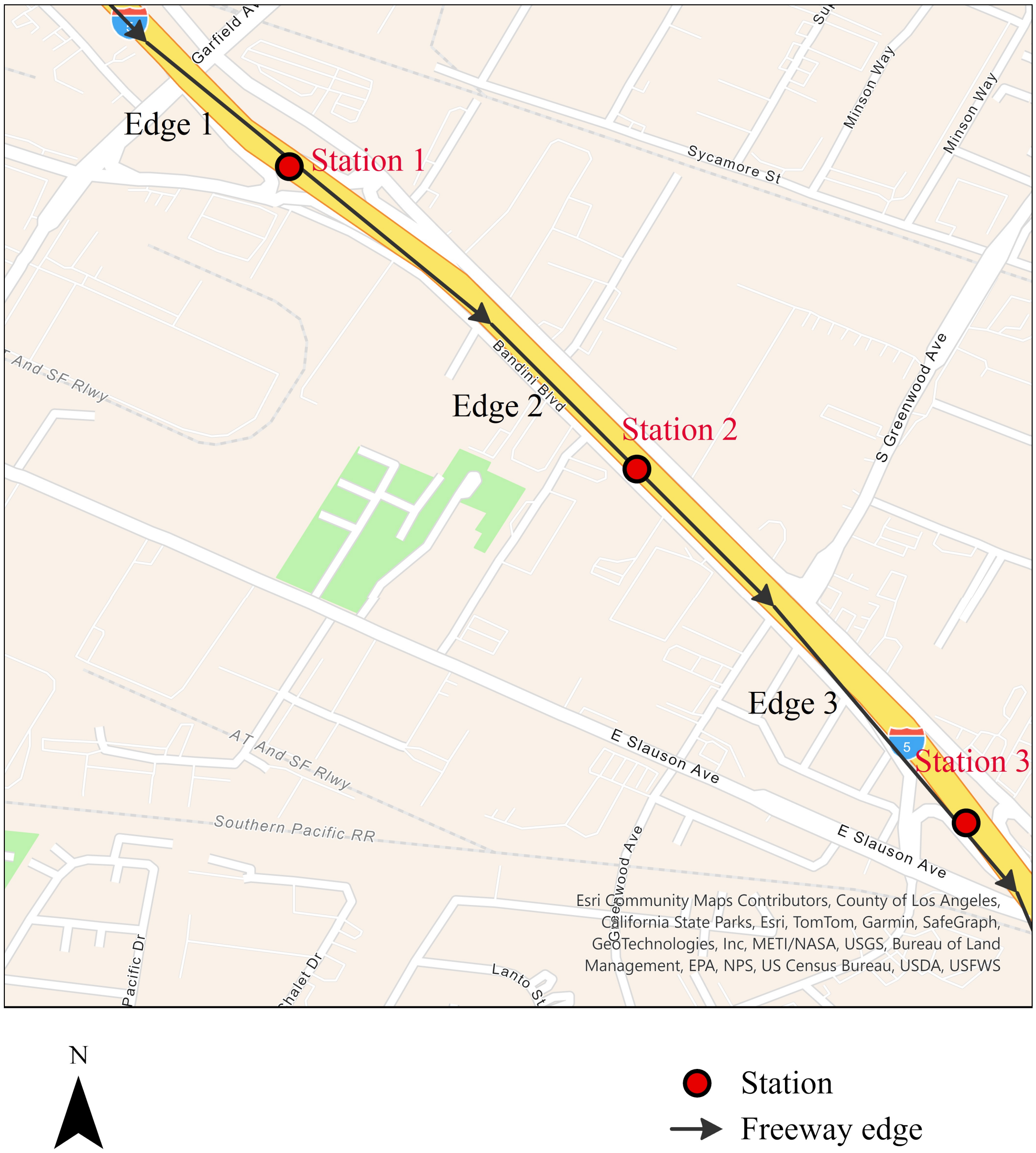
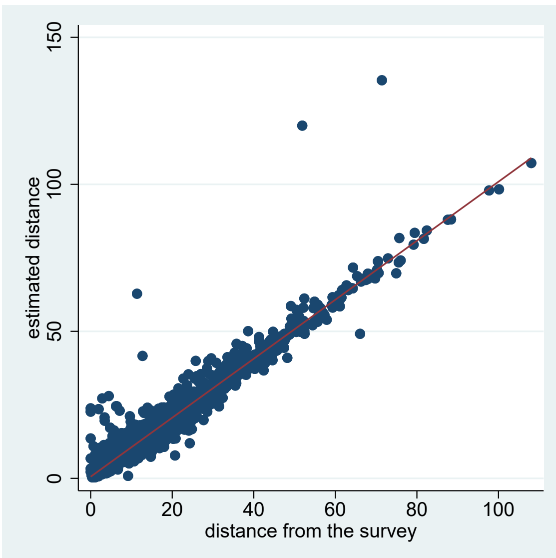
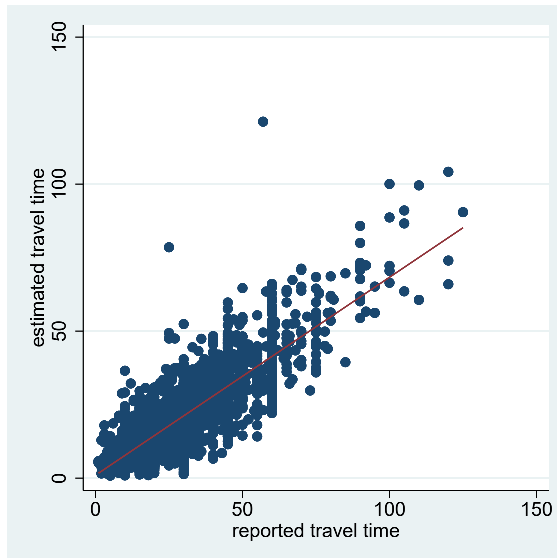


Figure C.2: Finding a matching station for a given freeway edge

Notes: This figure displays how the paper finds a matching station for a given freeway edge for select freeway edges on I-5 South. Each black line with arrow represents a freeway edge on I-5 South, with the arrow pointing to the direction of the freeway. Though freeway segments have curvature, they are illustrated with straight lines for simplicity. Each red circle represents a vehicle detector station on I-5 South. In this example, Station 1 is the matching station for Edge 1, Station 2 is the matching station for Edge 2, and Station 3 is the matching station for Edge 3.



(a) Estimated distance vs. Actual distance



(b) Estimated travel time vs. Self-reported travel time

Figure C.3: Comparing estimated commute distance and time with commute distance and time accessed from CHTS

Notes: Each observation point represents a morning commute trip from residence to workplace in CHTS. A linear fit is superimposed. Figure C.3a represents the full sample of 3,374 commute trips. The x-axis plots the trip distance reported in CHTS, which is estimated through Google Maps using the precise addresses of trip origins and destinations. The y-axis plots the trip distance estimated by finding the route with the minimum travel time between the centroid of the origin tract and the centroid of the destination tract using the OpenStreetMap augmented with historical speeds on freeways. Figure C.3b excludes trips with self-reported travel time that is out of the 95% confidence interval of the predicted travel time based on travel distance. The x-axis plots the travel time reported by survey respondents of CHTS. The y-axis plots the travel time estimated by finding the route with the minimum travel time between the centroid of the origin tract and the centroid of the destination tract using the OpenStreetMap augmented with historical speeds on freeways.

Table C.1: Availability of historical speed

Freeway	Sample	Full	Sample	Full	Coverage
	Eastbound length (mile)		Westbound length (mile)		
I10	245.25	241.58	245.49	241.65	Full coverage
I105	17.92	18.15	17.96	18.15	Full coverage
I210	86.16	85.64	86.14	85.62	Full coverage
SR118	29.11	46.32	29.27	46.42	Full coverage
SR134	15.35	13.34	15.09	13.39	Full coverage
SR142	5.84	11.38	5.75	11.36	Partial censor coverage
SR2	9.22	18.73	9.66	18.52	Partial censor coverage
SR22	15.34	14.73	15.09	14.66	Full coverage
SR60	73.51	70.97	73.15	71.25	Full coverage
SR74	12.52	114.60	12.52	114.60	Partial censor coverage
SR90	3.36	16.12	3.36	16.12	Partial censor coverage
SR91	59.70	59.19	62.42	59.52	Full coverage
	Northbound length (mile)		Southbound length (mile)		
I110	31.13	31.06	31.11	31.05	Full coverage
I15	234.45	239.07	235.47	239.46	Full coverage
I215	55.36	55.09	57.45	56.08	Full coverage
I405	74.89	72.41	72.82	72.44	Full coverage
I5	132.34	131.65	132.41	131.97	Full coverage
I605	28.70	27.97	28.50	28.07	Full coverage
I710	24.44	23.92	24.77	23.86	Full coverage
SR133	5.07	13.65	5.03	13.65	Partial toll road
SR14	52.10	52.19	52.62	52.17	Full coverage
SR170	7.82	6.05	7.25	6.19	Full coverage
SR23	8.71	15.29	8.60	14.99	Partial censor coverage
SR47	2.05	3.43	2.03	3.25	Partial censor coverage
SR55	15.90	17.65	15.99	17.66	Partial censor coverage
SR57	25.47	24.14	24.60	24.12	Full coverage
SR71	16.79	16.53	17.48	16.48	Full coverage
SR73	5.88	18.02	5.66	17.72	Partial toll road
US101	80.42	83.02	80.66	82.98	Partial censor coverage

Notes: Column "Sample" presents coverage length (the total length of freeway edges in the OpenStreetMap for which historical speeds are utilized) for each freeway-direction. The coverage length from 2010 to 2019 stays the same for a given freeway-direction, except that there are no historical speed data for the following freeway-year pairs: SR74 (2010-2011), SR90 (2010-2011), SR142 (2010-2012), SR142-E (2017), SR170 (2019). Column "Full" presents the full length for each freeway-direction within Greater LA, estimated using the California Enhanced National Highway System (NHS) line feature class, and the 2019 road shapefiles from the Census Bureau. Column "Coverage" indicates whether historical speeds from sensors are incorporated into all the segments for a given freeway on the OpenStreetMap.

Table C.2: Comparing estimated commute distance and time with commute distance and time accessed from CHTS

	Dependent variable			
	(1)	(2)	(3)	(4)
	Est. distance	Est. travel time	Est. travel time	Est. travel time
CHTS distance	1.0033***			
	(0.0089)			
Self-reported travel time		0.5416***	0.6733***	0.6887***
		(0.0242)	(0.0075)	(0.0074)
Observations	3,374	3,374	3,226	3,145
Sample	Full Sample	Full Sample	Outliers Removed	Outliers Removed

Notes: The unit of observation is a morning commute trip from residence to workplace in CHTS. The dependent variable, estimated distance/travel time, is the distance/travel time estimated by finding the route with the minimum travel time between the centroid of the origin tract and the centroid of the destination tract of a commute trip, using the OpenStreetMap augmented with historical speeds on freeways. "CHTS distance" is the trip distance reported in CHTS, which is estimated through Google Maps using the precise addresses of trip origins and destinations. "Self-reported travel time" is the commute time reported in CHTS, which is derived from arrival time and departure time recorded in survey respondent's travel diaries. Column (3) excludes trips with self-reported travel time that is out of the 95% confidence interval of the predicted travel time based on travel distance. Column (4) excludes trips with self-reported travel time that is out of the 90% confidence interval of the predicted travel time based on travel distance. Robust standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table C.3: Robustness checks: Labor force participation rate and lagged congestion

Dependent Variable: Labor force participation rate	(1)	(2)
avg_delay_{t-1}	0.0004	0.0007
	(0.0007)	(0.0010)
Observations	15,245	15,245
Additional Controls	Yes	Yes
Weights	Year-varying	Fixed
Residence-Tract FE	Yes	Yes
County-By-Period FE	Yes	Yes
Excl. Pub. Transit & WFH	Yes	Yes
Excl. Pairs w/ Endogenous Delay	Yes	Yes

Notes: All columns control for residence-tract fixed effects and county-by-period fixed effects. In all columns, average delay on routes out of a given residence tract is computed after dropping commute routes that may experience endogenous congestion, which are tract-pairs whose commute flows contribute to more than 0.1% of the total commute flows on the freeways they pass through. All columns drop tract-pairs with residence tracts where at least 50% of workers work from home or commute to work via public transportation, biking or walking (refer to footnote 21 for details). In column (1), $avg\_delay_{ic,t-1}$ is the average delay in year t-1 on commute routes out of residence tract i weighted by commute flows in year t-1. In column (2), $avg\_delay_{ic,t-1}$ is the average delay in year t-1 on commute routes out of residence tract i weighted by average commute flows between 2010 and 2018. "Controls" include percentage of female, percentage of families with children under six years old, ln(housing price), and ln(housing stock) that are *from a period before average delay is determined*. Standard errors clustered at the residence-tract level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.