Essays on the Effectiveness of Environmental and Urban Transportation Policies

by

Cheng Chen

B.A., Nanjing University, 2011

M.Sc., University of Waterloo, 2013

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Business Administration)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2019

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

Essays on the Effectiveness of Environmental and Urban Transportation Policies

submitted by **Cheng Chen** in partial fulfillment of the requirements for the degree of **DOCTOR OF PHILOSOPHY** in **Business Administration**.

Examining Committee:

Werner Antweiler
Co-supervisor

Sanghoon Lee
Co-supervisor

Robin Lindsey
University Examiner

Kevin Milligan
University Examiner

Additional Supervisory Committee Members:

Sumeet Gulati
Supervisory Committee Member

Abstract

This dissertation is a collection of three essays that study the efficiency of policies targeting environmental and urban transportation issues.

The first essay investigates a government incentive that subsidizes the purchase of electric vehicles(EVs) and asks a question of whether subsidizing public charging facility would be more cost-effective. It estimates a discrete choice model of EVs which relies on both EV characteristics and individual demographic information, using micro-level data from the California Clean Vehicle Rebate Program. Results show: (1) EVs with smaller battery capacity are more reliant on the public charging network; (2) consumers with higher income are less price sensitive. Such results support subsidizing more to public charging facility and reducing the EV purchase subsidy to the more affluent consumers.

The second essay builds on the literature of road congestion and addresses the importance of schedule delay cost occurred due to uncertain traffic time. We propose the notion of a "reliability standard" that commuters use to calculate their schedule time — the buffer time added to a commute so as to ensure being at work on time most of the time. With this tool, we conduct an extensive simulation study to gain further insights into the role of commuter composition, time cost differences, and the degree of inflexibility on optimal road tolls. While the cost of commuting time reliability is economically important, we find that the difference between a full-information road toll and a limited-information road toll is smaller than one would expect. Our study points to how future stated preference studies can be designed to elicit more meaningful information that identifies commuter heterogeneity, and in turn will lead to better designs for mobility pricing.

The third essay examines the relationship between firms' voluntary disclosure in environmental performance information and their level of institutional ownership. Empirical results indicate that US S&P 500 companies with higher institutional ownership ratio are less likely to disclose to the Carbon Disclosure Project. Moreover, disclosure behavior also leads to a lower ratio of institutional ownership. This negative relationship suggests adverse selection in the disclosure decision, hence such voluntary disclosure policy might be of limited value to the public.

Lay Summary

This dissertation evaluates the efficiency of three existing or proposed policies that target on environmental or urban transportation issues. First, in order to promote electric vehicle(EV) adoption, improving the public charging network is no less important than subsidizing individual purchasers. Investing in charging facilities will enhance the attractiveness of EVs with smaller batteries which are also more affordable. Subsidies on luxurious EVs can be reduced because higher-income consumers are less price-sensitive. Second, besides prolonged travel time, the uncertainty of travel time caused by congestion also occurs cost on commuters, therefore this also needs to be accounted for while designing mobility pricing policies. We propose that future surveys should ask about travelers' valuation of travel time, inflexibility criteria, and the correlation between them. Finally, voluntary disclosure of environmental information might be of limited value to the public as empirical results show a negative correlation between disclosure decision and institutional ownership.

Preface

Chapter 2 and 4 are original and independent work by the author, Cheng Chen. Chapter 3 is joint work with my thesis co-supervisor Professor Werner Antweiler. We divided the work primarily along the lines of theory and empirics and we collaborated on all aspects of the paper. My contributions on the empirical side include data analysis, empirical estimation, numerical algorithms, and simulations. Professor Antweiler took the lead on the theoretical parts of the paper.

Table of Contents

List of Tables

List of Figures

Acknowledgments

First and foremost I want to thank my advisors Professor Werner Antweiler and Professor Sanghoon Lee for their continuous support of my Ph.D. study. Both of them have devoted precious time discussing with me research ideas as well as technical details. I learned from them the perseverance, generosity, and enthusiasm of being a researcher. Especially, I owe a lot to Professor Antweiler who collaborates with me in Chapter 3 of the thesis and teaches me the rigor of conducting high-quality research.

I would also like to thank Professor Sumeet Gulati for joining my committee and providing helpful feedbacks on my research. I am also very grateful to Professor Ralph Winter who is willing to offer suggestions regarding all aspects. I also thank Professor Robin Lindsey for his suggestions and feedbacks. Many thanks to other faculty members of the Strategy and Business Economics Division and participants in the Environmental Economics Reading Group, for their comments and acts of kindness.

Special thanks to Xiaonan Sun and Ruoying Wang who have continuously shared their skills and insights with me without reservation. I also thank many other friends from Sauder School of Business and Vancouver School of Economics for our interesting discussions. Finally, I owe my survival in the Ph.D. journey to the unconditional support and ultimate understanding from my parents. Thank you.

Chapter 1

Introduction

This dissertation is a collection of three independent essays that investigate the efficiency of existing or proposed policies targeting on environmental or urban transportation issues. The first essay in Chapter 2 investigates whether providing cash subsidies to individual electric vehicle(EV) purchasers is the most cost-effective policy to promote EV adoption. The second essay in Chapter 3 addresses the importance of volatility in commute time and discusses how to induce an optimal mobility pricing scheme that accounts for such issue. The third essay in Chapter 4 analyzes the effectiveness of a voluntary disclosure policy on firms' environmental performance and suggests that disclosure content is of limited value to the public. In this chapter I will briefly summarize the main methodology and findings of each essay. Because each essay targets an independent topic, I thus leave a more comprehensive discussion of the research question and literature review in each chapter.

The first essay (Chapter 2) is titled "Efficient Government Spending on Promoting Electric Vehicles: a Case of California". Various incentives have been carried out by policy makers in order to promote adoption of electric vehicles(EV), which rely on electricity as an energy source and might generate less greenhouse gas compared to conventional vehicles using fossil fuel. This paper aims to analyze factors that affect consumers' taste on different types of EVs, and discuss whether the current U.S. government incentives are most efficient. Estimating a discrete choice model using micro-level data from California Clean Vehicle Rebate Program (CVRP), this paper discovers that access to both private and public charging facility contributes to decision of EV purchase. Among various types of EVs, pure battery electric vehicles (BEV) with relatively shorter electric range have the most inelastic demand towards public charging facility network. Meanwhile, shorter range BEVs have a modest price compared to longer range BEVs, and are more energy efficient compared to plug-in electric vehicles (PHEVs). Therefore subsidizing charging facility instead of individual EV purchasers might contribute more to energy saving and air pollution reduction, whereas reducing the concern that high-income potential buyers of luxurious longer range BEVs might free ride government subsidy.

The second essay (Chapter 3) is titled "The Cost of Commuting Time Reliability: Theoretical Advances using Reliability Standards and Empirical Findings from Metro Vancouver Traffic Data". The literature on road congestion has long acknowledged the importance of inflexible commuters who incur schedule delay costs in order to arrive on time. This paper advances a number of new insights into this phenomenon. We

introduce the notion of a "reliability standard" that commuters use to calculate their schedule time—the buffer time added to a commute so as to ensure being at work on time most of the time. Using two competing empirical distribution of commuting travel time (Fréchet and Lognormal), we develop a tight linear approximation of the schedule time that neatly separates schedule inflexibility (a commuter characteristic) from the coefficient of variation of travel time (a route characteristic). Our theoretical insights allow us to characterize and quantify the cost of commuting time reliability. Empirically, we use two data sets with traffic flows in the Metro Vancouver area—one with volume and speed of several key congestion points, and one with time distributions across a sample of 2,000 commutes across the entire Metro region—to identify the links between congestion and the cost of commuting time reliability. The empirical results demonstrate that travel time variability is a function of congestion, but interacts with route characteristics. We also conduct an extensive simulation study to gain further insights into the role of commuter composition, time cost differences, and the degree of inflexibility on optimal road tolls. We expand this approach to encompass the correlation structure of time costs, degree of inflexibility, and demand response (cutoff time for commutes). While the cost of commuting time reliability is economically important, we find that the difference between a full-information road toll and a limited-information road toll is smaller than one would expect. Overall, our study points to how future stated preference studies can be designed to elicit more meaningful information that identifies commuter heterogeneity, and in turn will lead to better designs for mobility pricing.

The title of the third essay (Chapter 4) is "Voluntary Carbon Disclosure and Institutional Ownership". This study unveils the relationship between firms' voluntary carbon emission disclosure behavior and the level of institutional ownership. The question of interest is whether disclosure of non-financial information is associated with changes in financial performance. Disclosure improves firms' transparency to shareholders, and also incurs costs such as to maintain good carbon emission record in the future, therefore disclosure can result in financial impact of either direction. Empirical results indicate that among US S&P 500 companies, higher institutional ownership ratio causes less disclosure to Carbon Disclosure Project(CDP), an organization which publishes large companies' self-disclosed information about carbon emission. Moreover the disclosure behavior also leads to lower ratio of institutional ownership. This negative relationship implies that firms' disclosure behavior does not signal carbon strength, yet might be caused by other motivations unfavored by institutional investors.

Chapter 2

Efficient Government Spending on Promoting Electric Vehicles: a Case of California

2.1 Introduction

Among various instruments towards reducing greenhouse gas(GHG) emission, electric vehicles(EVs) have drawn much attention not only from policy makers but also the general public. Different from conventional internal combustion engine vehicles (ICEVs), EVs do not depend on particular sources of energy, which allows them to use cleaner energy that generates less air pollutants and GHG. According to U.S. EPA, EVs convert about 59% - 62% of the electrical energy from the grid to power at the wheels, whereas ICEVs only convert about 17% - 21% of the energy stored in gasoline. Many governments have set ambitious goals on replacing ICEVs with EVs in the future. Britain and France said in 2017 that by 2040 new cars completely relying on petrol or diesel will be illegal. China, with the world's largest EV market, also announced that ICEVs would be banned in the future and required automobile makers with more than 30,000 annual sales to meet a quota of 10% being EVs.

From consumers' point of view, EVs favor potential buyers that they render much lower fuel cost, at about half of gasoline vehicles.[1] Surveys show that EV drivers also appreciate the new technology and environmental benefit. Despite the superiorities EVs embrace, their adoption are hindered by factors including higher prices and concerns on charging. For example, the popular EV Nissan Leaf (2017 model) has a Manufacturer's Suggested Retail Price(MSRP) of $30,680 - $36,790, while a comparable ICE vehicle Toyota Prius (2017 model) has a much lower MSRP of $23,475 - $30,015. According to EPA's estimate on annual fuel cost, Nissan Leaf costs $150 less than Toyota Prius per year, which hardly justifies the price difference of about $6,000. Convenience of charging is another concern to potential EV drivers, as EVs have shorter range with a full battery compared to ICEVs with a full tank of fuel whereas the charging process takes

[1] U.S. EPA website lists estimated annual fuel cost for each vehicle model. The actual fuel cost savings depends on local price of gasoline and electricity.

much more time than refueling at gas station. For example, Nissan Leaf (2017 model) has a range of 107 miles whereas Toyota Prius (2017 model) has a range of 588 miles, and using fastest charging facility, Leaf can be charged to 80% battery capacity in 30 minutes, which takes way more time than people stay at a gasoline pump.

Within the family of EVs, there are plug-in hybrid electric vehicles (PHEVs) and battery electric vehicles (BEVs), their difference being that PHEVs have shorter electricity range and would rely on gasoline after it exhausts its battery whereas BEVs use electricity as their sole energy source. Throughout this paper, I use the word EV to refer to both PHEV and BEV.

Policy makers around the world therefore launched combinations of incentives to encourage EV adoption, examples including: cash rebates or tax credits upon EV purchasing or leasing, unlimited access to high occupancy vehicle(HOV) lanes on highways, utility rate discount for charging EVs, etc. Some local governments also subsidize the installation of EV charging stations.

The current financial incentives given to EV owners are rather generous. EVs purchased in the US after year 2010 may be eligible for a federal income tax credit of up to $7,500, depending on the vehicle's battery capacity. Under such a scheme, BEVs usually qualify for $7,500 tax credit while PHEVs get $2,500 -$4,500. This paper mainly relies on data from California Clean Vehicle Rebate Program (CVRP), which is a state incentive provided by the California Air Resource Board. CVRP provides rebates to various kinds of energy-efficient vehicles including PHEVs, BEVs, and some other alternative fuel vehicles, with PHEVs and BEVs representing about 99% of all the rebated vehicles. Under CVRP, each BEV and PHEV receives a cash rebate of $2,500 and $1,500 respectively during the time period covered by the dataset used in this study.[2] With multiple instruments at hand for the governments to promote EV adoption, a natural question to ask is how to efficiently allocate limited funding so that maximum amount of total gasoline miles could be replaced by electricity miles, leading to most pollution reduction. To answer this question, I propose to first estimate the utility function of PHEVs and BEVs that explains heterogeneous consumers' decision making process in EV choice, and then conduct counterfactual exercises to see how consumers would react under alternative incentive bundles.

The case of California is worth studying due to at least two reasons. First, California is the biggest EV market in the US, accounting for about half of US EV sale. Second, compared with other states, California suffers more serious air pollution from ICEVs whereas embraces a cleaner electricity source, bringing EVs greater environmental benefits (Holland et al., 2016).

The contribution of this paper is two-fold. First, it estimates a discrete choice model of EVs which relies on both EV characteristics and individual demographic information, whereas the literature mainly uses aggregated data. Second, this paper provides implication on how to design cost-effective policies to promote EVs, leveraging the heterogeneous tastes of EV owners towards financial incentives and charging infrastructure.

[2]The CVRP rebate amount adjusted several times since its establishment in 2010, and the details will be presented in a later section of this paper.

2.2 Background

2.2.1 Electric Vehicles

Over the years, global warming has made increasingly visible impact to our planet. For example, growing arctic temperature has melted glaciers dramatically and resulted in sea level rise, which further threats to inundate land areas; places around the world experience enlarging variability in temperature and precipitation, which led to conditions like severe wildfires in California during 2017 fall. Increasing GHG emissions such as carbon dioxide, methane, and nitrous oxide are the dominant cause of climate change.

Car driving accounts for one of the largest sources of GHG emissions as traditional cars rely on fossil fuel as sole energy source. Fortunately the invention of EVs provides an opportunity to reduce GHG emissions, since electricity could be generated from cleaner energy sources that emit less GHG. California, for example, mainly uses natural gas, hydro power, solar power, and nuclear power for electricity generation. Such energy sources not only lowers GHG emissions compared with gasoline, but also generate less air pollutants such as particulate matter, nitrogen oxides, sulfur dioxide, etc.

EVs first came into existence in mid-19th century, yet the series production of highway-capable EVs happened not until recent decade. In 2008, the first EV available to the general public, Tesla Roadster, was launched. It is the first EV using lithium-ion battery cells and has a range greater than 200 miles between charges. The production of Roadster was ceased in 2012, when Tesla introduced Model S, a full-sized all-electric luxury car which receives wide popularity. The earliest generation of Model S with a 60 kWh battery pack has an electric range of 208 miles rated by EPA, and 265 miles for the 85 kWh battery pack model. Despite that the price of Model S starts as high as $66,000, it is the top selling EV in year 2014 and 2015, also the second best selling EV model of all the time. In 2010, the Japanese manufacturer Nissan introduced a compact hatchback electric car — Leaf, which is the world's best selling EV model. First generation of Nissan Leaf comes with a modest electric range at 73 miles (rated by EPA).

Tesla Model S and Nissan Leaf are both BEVs which consume electricity as the sole energy source. Another branch of the EV family is PHEV, which carries both a battery and an internal combustion engine. PHEVs function just like BEVs until their battery is exhausted and then they will turn to the gas tank to power the car. PHEVs began available to the public in 2010, and among the most popular models including the Chevrolet Volt manufactured by General Motors. The first generation of Chevrolet Volt has an electric range of 25-50 miles, and it has a total range of 379 miles with full battery and gasoline.

2.2.2 Government Incentives

Policy makers around the globe have tried various tools to promote EV adoption. Federal government of the United States carried out tax credit for new EV purchasers since 2008, the amount of which increases proportionally with battery size and capped at $7,500. The tax credits will phase out after 200,000 EVs have been sold by one car manufacturer. Until the end of 2017, the 200,000 subsidy limit has not been reached by any manufacturer. Charging equipment installed by homes and businesses also receive tax credit through 2013, the amount ranging from $1,000 to $50,000, depending on installation scale.

In California, the largest EV market of US, additional financial incentives are provided by the Clean

Vehicle Rebate Project (CVRP). The program started in 2010 with $5,000 subsidy to BEVs, and the rebate amount soon decreases to $2,500 after June 2011. PHEVs entered the market in year 2012 with the introduction of Toyota Prius PHEV, and PHEV buyers receive a rebate of $1,500 from CVRP. Starting from Mar 29, 2016, the program gives more subsidy to lower-income buyers whereas sets an income cap to restrict application of rebates from the richer. On Nov 1, 2016, subsidy to lower-income buyers are increased whereas the income cap is further lowered. Details are revealed in Table 2.1. Besides financial incentives, local policies such as allowing solo EV drivers access to HOV lanes, free EV parking, discounted utility rate for EV charging, etc. also contribute to California's relatively high EV sales.

Table 2.1: CVRP Schedule

Vehicle Date of Purchase	Income Cap	Increased Rebate
Prior to Mar 29 2016	Not Applicable	Not Applicable
Mar 29 2016 - Oct 31 2016	$250,000	$1,500
Nov 1 2016 - Present	$150,000	$2,000

Increased rebate are eligible to households with income less than 300 % of the federal poverty level.
Income Cap are for single filers. Household filers have higher income caps.

Other countries leading in EV sales such as Japan, China, United Kingdom, Norway, Netherlands, Sweden, France, Germany, etc. have also carried out various monetary incentives. The policy instruments they used include tax reduction, direct subsidies to purchasers or manufacturers, elimination of car registration fee, etc.

2.2.3 Charging Infrastructure

There are several modes of EV charging. It could be as simple as plug in to a non-dedicated 110-volt household outlet (Level 1 charging), and it takes about 16 hours to add 100 electric miles.

Level 2 charging stations are the most common ones installed by household and public parkades. They require 240-volt outlet and can charge about 6 times faster than Level 1. Smart features are also available for Level 2 charging stations which can schedule charging for off-peak hours. The cost to purchase and install Level 2 charging stations at home is about $1,500, while rebates ranging from $500 to $1,000 are available for some counties in California. Public Level 2 charging stations cost more than home chargers due to more features, and installing chargers at curbside costs more than in a garage. Local incentives for public or workplace charging stations are also available in some areas.

Direct Current(DC) fast charging stations connect EVs to main power grid through an external charger, and can charge most EVs to 80 percent in 20-30 minutes. They are much more expensive to install, costing $50,000 to $100,000 per station.

2.2.4 Related Literature

This study is built on a growing literature that analyze the factors affecting EV adoption empirically. Li et al. (2017) studied how adoption of EV and deployment of EV charging stations affect each other using US vehicle sales data. A feedback loop is found and the authors conclude that policies that subsidize

building charging stations would be twice as effective in promoting EVs compared to the current federal tax incentives for EV purchasers. Studies based on other places around the world also confirmed that EV charging facility plays an important role in pushing forward EV adoption, examples include Sierzchula et al. (2014), Mersky et al. (2016), Wang et al. (2017), and so on. Bailey et al. (2015) found that the availability of home charging has a much more substantial impact than perceived existence of public charging stations on EV demand, using survey of new-vehicle buyers in 2013.

This paper is also closely related to the literature discussing the design of government subsidy on clean vehicles. DeShazo et al. (2017) focuses on designing optimal EV rebate policy that induces more EV sale under a constraint of total government expense. They illustrated that policies which set a cap on EV price or set different rebate schedule for consumers with heterogeneous income would be more efficient. To support this idea, they developed a theory model and tested empirically using data from stated preference surveys. Chandra et al. (2010) estimate the effect of tax rebates offered by Canadian Provinces on the sales of hybrid electric vehicles (HEVs). While confirming the rebate successfully boosted sales of HEVs, they argue that the subsidized consumers would have bought either HEVs or other fuel-efficient vehicles in the absence of the rebate. Therefore they conclude that reduction of carbon emission is too costly under this rebate scheme.

A strand of literature uses stated choice survey data to analyze what factors influence consumers' decision in EV adoption. Hackbarth and Madlener (2013) used discrete choice model on German survey data to analyze consumers' preference to BEVs and PHEVs. They found that due to the longer waiting time of EV charging and higher risk of getting stranded with an empty battery, BEVs are less appealing to consumers than PHEVs, since the latter has a backup gasoline tank that allows the PHEV to be refueled as a conventional ICEV. Lieven (2015) conducted survey analysis regarding potential EV purchasers' preference to combinations of different government incentives. Based on surveys collected from 20 countries, they claimed that the charging infrastructures are must-haves while expediting EV diffusion, whereas monetary incentive and favorable traffic regulations are appealing but not as essential as establishing a charging network.

Several papers apply the random-coefficient discrete choice model which follows the setting of Berry et al. (1995) (referred to as the BLP model) and study the demand for electric vehicles or other new energy vehicles. Zhang et al. (2016) observe the BEV market in Norway and examine how consumers' BEV choice vary with government incentives, in addition to price and car characteristics. They find that government incentives including bus lane access, toll waiver, and charging station expansion promotes BEV adoption, within them increasing density of charging station has the most significant effects for BEV sales. Beresteanu and Li (2011) use BLP style random coefficients model to analyze demand for hybrid vehicles under scenarios of varying gasoline prices and government tax incentives.

2.3 Theoretical Framework

2.3.1 Operational Cost Decomposition

This section decomposes operational costs of different types of vehicles.

Operational cost c_{ij} includes monetary cost and time cost that occur when refueling vehicles. For BEVs,

PHEVs and ICEVs, the general form of their operational costs are similar:

$$c_{ij} = m_j + w_i t_{ij}$$

where m_j stands for dollar fuel cost of driving 1 mile; w_i is valuation of time for individual i; t_{ij} is time spent refueling for 1 mile's driving distance, which includes time searching for refueling stations as well as waiting. Next I decompose c_{ij} in detail.

2.3.1.1 Operational Cost of EVs

A major concern for EV adoption is the reliability and convenience of charging. Assume two scenarios of charging an EV: regular charging with cost $c_{1,ij}$ and backup charging with cost $c_{2,ij}$. Upon purchasing an EV j, consumer i has made a decision on her regular charging mode – installing home charging equipment, or use public chargers located in parkades of workplace or near home, resulting in regular charging cost $c_{1,ij}$. The consumer is also aware that in some non-regular cases when a BEV is about to exhaust all its battery away from her usual charging spot, she would then have to search for public charging station and wait while the BEV is being charged. PHEVs, however, enjoy more flexibility when the battery drains thanks to their backup fuel tanks. In such non-regular cases, charging cost $c_{2,ij}$ is generated. Last but not least, there exists a possibility that a BEV might get stranded due to completely exhausted battery, with an expected cost of stranded being $c_{3,ij}$.

Let q_{ij} denote the probability of "backup charging". The car owners would plan ahead to avoid backup charging, since that either renders more time cost waiting for a BEV to be charged or more monetary cost to add fuel for a PHEV. Assume daily range driven by person i follows log-normal distribution: $r_i \sim$ Lognormal(μ_i, σ_i^2),[3] with its probability density function denoted as $f(r_i)$. With EV j's electric range being R_j, q_{ij} can be expressed as the probability that individual i drives distance longer than its maximum electric range within a day:

$$q_{B,ij} = \int_{\delta R_j}^{+\infty} f(r_i)\mathrm{d}r_i$$

$$q_{P,ij} = \int_{R_j}^{+\infty} f(r_i)\mathrm{d}r_i$$

where $\delta \in (0,1)$ is a ratio which means BEV owners recharge their BEVs once they've driven a distance of δR_j.[4]

Therefore $q_{ij} = q(\mu_i, \sigma_i, R_j)$. For simplicity, I assume $\sigma_i = \sigma$ so that $f(r_i)$ only depends on μ_i, the mean daily driving distance. Then $q_{ij} = q(\mu_i, R_j)$ and $\frac{\partial q_{ij}}{\partial R_j} < 0$, $\frac{\partial q_{ij}}{\partial \mu_i} > 0$. The intuition is that EVs with longer electric ranges will occur less backup charging whereas people who on average drive longer distance per day face higher chance of exhausting battery storage.

[3] To echo the pattern of peoples' daily driving distance from 2009 National Household Transportation Survey, I set $\mu_i = 3.25$, $\sigma_i^2 = 0.75$. The distribution of r_i is skewed to the left with a long right tail.

[4] BEVs suffer from high cost of being stranded therefore BEV owners will not exhaust all of its electric range before recharging, whereas PHEVs do not share such problem.

In the following paragraphs, I discuss the composition of charging costs $c_{1,ij}$ and $c_{2,ij}$, and stranded cost $c_{3,ij}$. The cost of regular charging $c_{1,ij}$ varies with the charging mode selected by the EV owner. Specifically, whether or not the EV owner has access to home charging makes a big difference. Having a home charger guarantees availability and convenience of charging, and does not cause detour to find public charging stations. Assume that EV owners who have chargers at home will not use public charging stations for regular daily charging, then their regular charging cost would be:

$$c_{1|\mathrm{HC},ij} = P_{\mathrm{S}} + \frac{P_{\mathrm{e}}}{\mathrm{E_j}}$$

where P_{S} is the fixed cost to install a Level 2 charging station amortized to each mile driven by the EV[5]; E_j is energy efficiency of vehicle j, i.e., miles/kWh; P_{e} is electricity rate[6]. Note that in the case of home charging, no time cost occurs to individual i.

EV owners without home charging will experience more hassles to regularly charge their EVs due to the physical distance between charger and home, uncertainty of charging spot availability, and more expensive electricity price. The frequency of charging depends on both EV's electric range R_j and EV owner's average daily driving distance r_i. Assume that EV owners go to public charging stations only when they expect the left battery storage would not fully cover the driving distance of the next day, then the days between each time charging would be $\max(\lfloor \delta R_j/r_i \rfloor, 1)$ for BEVs and $\max(\lfloor R_{\mathrm{PE},j}/r_i \rfloor, 1)$ for PHEVs. Then the regular charging cost per mile of EV owners without home charging $c_{1,ij}$ is:

$$c_{1|\mathrm{BEV,NHC},ij} = \frac{w_i t_1(S_{\mathrm{S},ij})}{\min(r_i, R_j)\max(\lfloor \delta R_j/r_i \rfloor, 1)} + \frac{P_{\mathrm{e}}}{E_j}$$

$$c_{1|\mathrm{PHEV,NHC},ij} = \frac{w_i t_1(S_{\mathrm{S},ij})}{\min(r_i, R_j)\max(\lfloor R_{\mathrm{PE},j}/r_i \rfloor, 1)} + \frac{P_{\mathrm{e}}}{E_j}$$

where $S_{\mathrm{S},ij}$ is density of level II public chargers[7]; R_j is the electric range of the BEV; $R_{\mathrm{PE},j}$ is the electric range of PHEV j, w_i is individual i's valuation of time. $t_1(\cdot)$ stands for the time spent detouring to a public charging station and it is a decreasing function of $S_{\mathrm{S},ij}$. With greater density of public charging stations, EV owners spare less efforts finding one and less time walking between charging stations and home or office. Also, an EV with a longer electric range can run for more days once fully charged, therefore it reduces the frequency of charging and time spent on it.

With probability q_{ij}, an EV relies on backup charging channels. Since PHEVs have secondary fuel source as gasoline, PHEVs and BEVs' backup charging costs differ. For BEVs, such costs include time cost of detouring to a charging station and waiting for the EV to be charged, as well as cost of electricity.

$$c_{2|\mathrm{BEV},ij} = q_{B,ij}\left[w_i\left(\frac{t_1(S_{\mathrm{F},ij})}{\delta R_j} + t_{2,j}\right) + \frac{P_{\mathrm{e}}}{E_j}\right]$$

[5]Level 2 charging uses a regular socket with some EV specification protection arrangement. Most EV models can be fully charged in 4-6 hours under Level 2 charging. They are most popular in home charging and public charging stations.

[6]For simplicity reason, assume homogeneous electricity rate in baseline case

[7]Charger density differs by geography therefore has subscript i. Compatibility of chargers with EV models may differ therefore subscript j.

where $S_{\mathrm{F},ij}$ is the density of fast charging stations[8], $t_{2,j}$ is charging time of vehicle j under fast charging mode (minute/mile). Note that in this case, individual i waits for the EV to be charged, because this is not her regular charging spot and she might be in the middle of a trip. DC fast charging stations nowadays usually can add 80 % of an EV battery in 20-40 minutes, whereas level II chargers take more than 3 hours. Therefore I assume that BEV owners search only for fast charging stations in case of backup charging.

For PHEVs, I assume that the owners rely on gasoline as their sole fuel source in cases of backup charging. Similar to BEVs, PHEVs' backup charging cost involves time cost of detouring to a gasoline station and waiting while the gas tank is filled, as well as monetary cost of gasoline.

$$c_{2|\mathrm{PHEV},ij} = q_{P,ij}\left[w_i\left(\frac{t_1(S_{\mathrm{G},i})}{\delta R_{\mathrm{PG},j}} + t_{3,j}\right) + \frac{P_g}{G_j}\right]$$

where $S_{\mathrm{G},i}$ is the density of gasoline stations; $R_{\mathrm{PG},j}$ is PHEV j's gasoline range; $t_{3,j}$ is the time of filling the gas (minute/mile); G_j stands for miles per gallon and P_g is gasoline price considered as a constant.

Furthermore, possibility remains that BEVs will be stranded with an empty battery, which renders undesired inconveniences such as car towing. The expected cost of being stranded $c_{3,ij}$ only occurs to BEV but not PHEV, considering the well developed network of gasoline stations. Probability of failure to find a public charging station, $f_{ij}(S_{\mathrm{S},ij}, S_{\mathrm{F},ij})$ is a decreasing function of density of public charging stations, i.e., $\frac{\partial f}{\partial S_{\mathrm{S},ij}} < 0$ and $\frac{\partial f}{\partial S_{\mathrm{F},ij}} < 0$. BEV owners with accessibility to home charging are only exposed to risk of stranded in case of backup charging, whereas BEV owners without home charging face the stranded risk in all charging modes. Therefore the expected cost of stranded is as below:

$$c_{3|\mathrm{PHEV}} = 0$$

$$c_{3|\mathrm{BEV,HC,ij}} = q_{B,ij} f_{ij}\left(w_i \frac{t_s}{R_j} + m_s\right)$$

$$c_{3|\mathrm{BEV,NHC,ij}} = f_{ij}\left(w_i \frac{t_s}{R_j} + m_s\right)$$

where monetary cost occurred and time wasted after a BEV gets stranded is abstracted to m_s and t_s.

To simplify the analysis, I sum up EV's monetary operational cost and denote it as $M_{B,j}$ for BEVs or $M_{P,j}$ for PHEVs[9]. Then the cost term c_{ij} derived from the previous discussion could be summarized as below, differed by being a BEV or PHEV, and having accessibility to home charging (HC) or not (NHC):

[8] Compatibility of fast charging stations differs across EVs, especially Tesla manages their own fast charging network — Tesla Superchargers, therefore this term has subscript j.

[9] According to EPA, estimated annual fuel cost of a BEV or PHEV varies from \$550 to \$850, which is a relatively small amount compared to vehicle prices. PHEVs bear greater fuel cost in general due to higher price of gasoline than electricity.

$$c_{ij} = \begin{cases} (\text{HC},\text{BEV}):\ P_{\text{S}} + M_{B,j} + q_{ij}\left[w_i\left(\dfrac{t_1(S_{\text{F},ij})}{\delta R_j} + t_{2,j} + f_{ij}\dfrac{t_{\text{s}}}{R_j}\right) + f_{ij}m_{\text{s}}\right] \\ (\text{HC},\text{PHEV}):\ P_{\text{S}} + M_{P,j} + w_i q_{ij}\left(\dfrac{t_1(S_{\text{G},i})}{\delta R_{\text{PG},j}} + t_{3,j}\right) \\ (\text{NHC},\text{BEV}):\ M_{B,j} + \dfrac{w_i t_1(S_{\text{S},ij})}{\min(r_i, R_j)\max(\lfloor \delta R_j / r_i \rfloor, 1)} + w_i\left[q_{ij}\left(\dfrac{t_1(S_{\text{F},ij})}{\delta R_j} + t_{2,j}\right) + f_{ij}\dfrac{t_{\text{s}}}{R_j}\right] + m_{\text{s}} f_{ij} \\ (\text{NHC},\text{PHEV}):\ M_{P,j} + \dfrac{w_i t_1(S_{\text{S},ij})}{\min(r_i, R_{\text{PE},j})\max(\lfloor R_{\text{PE},j} / r_i \rfloor, 1)} + w_i q_{ij}\left(\dfrac{t_1(S_{\text{G},i})}{\delta R_{\text{PG},j}} + t_{3,j}\right) \end{cases} \tag{2.1}$$

2.3.1.2 Operational Cost of ICEVs

Operational cost for conventional ICEVs is listed below for comparison, the functional form of which is similar to the backup charging case of PHEVs. It includes time cost of searching for gasoline station and waiting at the gas pump, and monetary cost of gasoline consumption.

$$c_{\text{ICEV},ij} = w_i\left(\frac{t_1(S_{\text{G},i})}{R_j} + t_{3,j}\right) + \frac{P_g}{G_j}$$

2.3.2 Factors Affecting EV Operational Cost

For consumers without concerns of environmental impacts, their choices between BEVs, PHEVs, and ICEVs with same vehicle attributes depend on trade-off of monetary and time cost of refueling. Therefore efforts to reduce time cost of EV charging is crucial in promoting EV adoption, given that monetary cost of EV charging is relatively fixed.

From equation (2.1), one can discover that driving force of EV charging time cost consist of accessibility to home charging, density of EV charging stations, valuation of time, EV category (BEV or PHEV), and EV range. This section discusses in general how consumer characteristics and EV attributes affect convenience of EV charging and hence consumer utility. Several propositions are listed below, with their proves provided in Section A.1.

Proposition 1 *Accessibility to home charging will lower charging cost in general.*

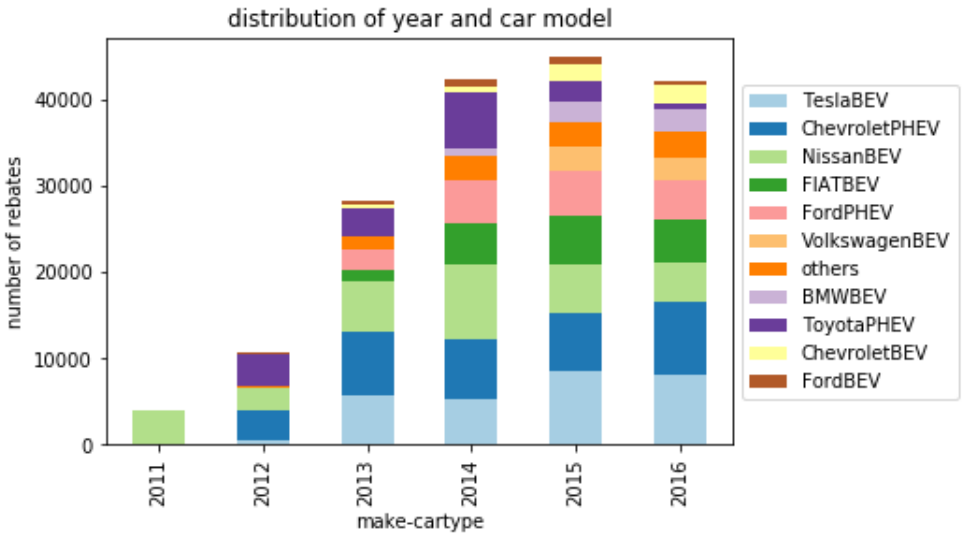
Proposition 2 *Adding public charging stations will lower charging cost in general, and such impact is greater for EV owners without home charging.*

Proposition 3 *Within BEVs, those with shorter electric range benefit more from adding public charging stations.*

Proposition 4 *Greater electric range lowers charging cost, and such impact enlarges with people's valuation of time.*

The propositions in this section will be tested using individual level EV purchase data. Also, I will use aggregated data to show that smaller BEVs (SBEVs, BEVs with shorter electric range) are more sensitive

Figure 2.1: Year Trend of Rebate Applications



to the increasing number of public charging stations than larger BEVs (LBEVs, BEVs with longer electric range). The dataset will be described in Section 2.4, and specification details will be provided in Section 2.5.

A baseline scenario (illustrated in Section A.2) compares the operational costs of several currently available car models which are considered to be close substitutes with each other. It shows how individual characteristics such as accessibility to home charging, valuation towards time affect operational cost hence choice of vehicle within categories of BEVs, PHEVs, and ICEVs.

2.4 Data

2.4.1 Electric Vehicle Adoption in California

This paper focuses on the case of California. California Environmental Protection Agency's Air Resources Board releases information for each EV purchase under CVRP (Center for Sustainable Energy, 2017), including EV make, time of purchase, Census Tract of residence, etc. I will aggregate this whole sample dataset by Census Tract and year, and then examine the impact of public charging stations on EV adoption. Having detailed geographic information — Census Tract code of the vehicle owner is crucial for this study, for it enables observing how many electric vehicle charging stations are present surrounding the drivers' home address. A total number of 8,058 census tracts are involved in this study.

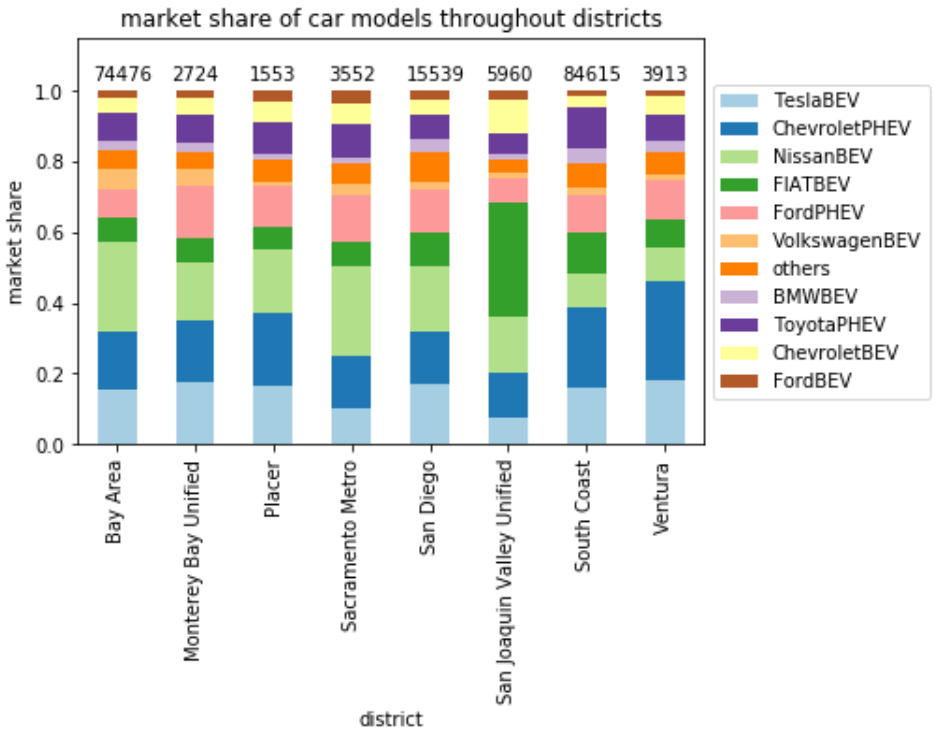
Ever since the CVRP launched in 2010, the number of BEVs and PHEVs[10] applying for rebate kept increasing. Figure 2.1 below exhibits the trend of rebate applications from different vehicle models[11]. Moreover, figure 2.2 shows the market shares of different EV models across different regions, with the number on top of each vertical bar indicating the total number of rebates of the region.

CVRP also conducted a detailed consumer survey that asks about demographic information. 14,442

[10] PHEVs were launched to sell in year 2012. The data in year 2016 only lasts until September 2016.

[11] For simplicity, BEVs under a single automobile maker are considered as same product, so are PHEVs.

Figure 2.2: Market Share across Regions



questionnaires have been answered completely, which represent about 10% of all the rebates in CVRP. The following information is unveiled for each questionnaire: (1) rebate application date, (2) rebate dollar amount, (3) vehicle category (e.g. PHEV, BEV, etc.), (4) vehicle make (e.g. Tesla, Nissan, etc.), (5) residence county[12], (6) income range, (7) housing type, (8) renter/owner of the residence place. With individual EV purchaser's demographic information, this survey data can be used to examine consumers' heterogeneous taste on EV characteristics.

Table (2.2) exhibits vehicles' type and make frequency that are present in the consumer survey. Note that in the survey, some less popular EV makers are documented as "Others", such as Smart and Volkswagen, probably due to privacy reason. Such less popular EV models are excluded from our individual level discrete choice analysis since EV characteristics are unknown. The first column of table (2.2) include all observations; the second and third columns present observations in counties with higher/lower population density; the last two columns present observations in Northern/Southern California counties. Such split of sample is in accordance with the empirical analysis in later part of the paper.

2.4.2 National Household Travel Survey

2017 National Household Travel Survey (NHTS) [13] also provides useful information for this study. The nationwide survey asks households' travel decisions including vehicle purchases. One could observe house-

[12] Unfortunately, more accurate Census Tract information is not asked in the individual questionnaire.

[13] Downloaded on March 20, 2018 from https://nhts.ornl.gov.

Table 2.2: Frequency of Vehicle Make in Surveyed Sample

Type-Make	All	High Pop. Den.	Low Pop. Den.	North CA	South CA
BEV					
BMW	335	186	149	129	206
Chevrolet	499	237	262	242	257
FIAT	1,482	813	669	681	801
Ford	402	164	238	213	189
Nissan	2,974	1,369	1,605	1,789	1,185
Tesla	2,334	1,207	1,127	1,140	1,194
Toyota	352	183	169	173	179
PHEV					
Chevrolet	2,890	1,556	1,334	1,101	1,789
Ford	1,660	789	871	736	924
Toyota	1,514	894	620	618	896
Total	14,442	7,398	7,044	6,822	7,620

holds' vehicle choice including both EVs and ICEVs from NHTS, unlike in CVRP where only EV purchases are observed. Households' demographic information such as income, home ownership, household size, household vehicle count, etc. are also provided.

An important feature of NHTS dataset is that household's residential state is revealed. This helps identify the dollar amount of financial incentive a consumer might receive if she purchases an EV. Since financial incentives differ across states and EV models, such variation makes it possible to estimate consumers' preference towards financial incentives.

For the purpose of research, there are at least two shortcomings of the NHTS dataset. First, the geographical identifiers of households are at state level, which is too broad to learn the pattern of public charging stations deployed near each households. Second, since EVs' market share in the US is about 1%, the sample size of EV purchasers in this dataset is also relatively small (about 400 EV purchases out of 40,000 total vehicle purchases surveyed).

Nevertheless, I use this dataset to determine households' choice between category of EVs and ICEVs given their demographic information, financial incentives, and density of public charging stations at state level.

2.4.3 Charging Stations

Information of all electric vehicle charging stations in California are downloaded from Alternative Fuels Data Center, U.S. Department of Energy[14] on Oct 20, 2015. Charging stations' location and type information are contained in this dataset. Unfortunately the dataset unveils opening dates information for merely about one third of all stations. However, variation in density of stations is crucial in this study, I rescue the situation by using the earliest date one station appears on on-line EV forums[15] to substitute for the actual opening

[14] http://www.afdc.energy.gov/data_download

[15] The first source is Plugshare, which comes with a smartphone application to let EV owners to find nearby charging stations and allow them to leave comments. The date of the first comment is used. The second source is carstations.com, which tells the

date. Yet there are approximately 20% stations still left with unknown opening date, I use Oct 20, 2015 — the day when the list of charging stations is downloaded, to fill in such missing information. The table below casts a light on the expansion of EV charging station construction and EV adoption. During year 2011 to 2015, the ratio of EVs to charging stations in California increased from 9:1 to 57:1.

Table 2.3: Annual Growth of EV Charging Stations and Rebates

Year	2010	2011	2012	2013	2014	2015
Total # Rebates (BEV&PHEV)	61	4,028	14,679	42,996	85,371	130,276
Total # Charging Stations	155	465	960	1,412	1,778	2,296
Δ # Rebates (BEV&PHEV)	61	3,967	10,651	42,375	42465	44,905
Δ # Charging Stations	155	310	495	452	366	518

2.4.4 Demographic Information

Table (2.4) shows the actual EV buyers' information from the surveys, including: income, whether owning a detached house (as a proxy for "having access to home charging"), and number of charging stations nearby. Descriptive statistics are exhibited by buyers of different types of EVs: large BEVs, small BEVs, and PHEVs. Income is given as a range in the original dataset, e.g. $50,000 to $74,999. In our analysis, each individual's income is assumed to be the mean of the income range he/she belongs to[16]. Number of public charging stations nearby is derived from the EV buyer's residence county and purchase year. I first calculate for each Census Tract within a county, how many charging stations are within 30km range; then a weighted average is calculated using these numbers, where the weights are given by purchase frequency of each Census Tract; the weighted average varies by county and year, and is a used as proxy for number of charging stations nearby.

Table 2.4: Demographic Information

Variable	Mean	Std. Dev.	Min.	Max.	N
Large BEV (Tesla) Buyers					
income ($1,000)	334.7	185.3	37.5	600	2334
own house	0.9	0.3	0	1	2334
charging stations in 30km (100)	1.1	0.5	0	2.5	2334
Small BEV (Non Tesla BEV) Buyers					
income ($1,000)	170.7	110.7	37.5	600	6048
own house	0.8	0.4	0	1	6048
charging stations in 30km (100)	1.2	0.6	0	2.5	6044
PHEV Buyers					
income ($1,000)	170.1	110.2	37.5	600	6065
own house	0.8	0.4	0	1	6065
charging stations in 30km (100)	1.1	0.6	0	2.5	6064

concrete date when a station is added on the site. The earliest date of these two data sources will be used as the "opening date".

[16]The highest income group is documented as "income greater than $500,000", no upper limit of income is given. I use $600,000 as a proxy for highest income.

2.5 Empirical Strategy

2.5.1 Discrete Choice Model of EVs

2.5.1.1 Model Setup

Consumer i makes a discrete choice of EVs from J alternatives. She gains utility u_{ij} from the choice denoted as $Y_i = j$. Utility is composed of a systematic part v_{ij} and a random part ε_{ij} which follows generalized extreme value distribution.

$$u_{ij} = v_{ij} + \varepsilon_{ij} \tag{2.2}$$

where i and j stand for consumer and EV model respectively. According to EV's energy source and their electric range, I partition the whole set of EVs into groups: PHEVs, longer-range BEVs, and shorter-range BEVs. For simplicity, I use the term large BEVs(LBEV)/small BEVs(SBEV) interchangeably with longer-range BEVs/shorter-range BEVs. [17] Until year 2017, Tesla stands out as the only BEV model that has a range greater than 200 miles, whereas other BEVs' ranges vary from 80 miles to 100 miles. Therefore the categories large BEVs and small BEVs could also be deemed as Tesla and non-Tesla BEVs respectively. The categorization is demonstrated in Figure 2.3. Then the utility function can be expressed as:

$$u_{i,g,m} = v_{i,g,m} + \varepsilon_{i,g,m}$$

where subscripts g stand for the group of PHEVs, LBEVs and SBEVs, m is the vehicle make.[18] Therefore each EV j could be identified by subscript(g,m).

$v_{i,g,m}$ depends on both vehicle characteristics $\mathbf{z_{g,m}}$ and consumer demographic characteristics $\mathbf{x_i}$, and its functional form closely follows predictions from the theoretical model.

For vehicle characteristics, I include price, annual fuel cost estimated by US EPA, and electric range which determines the convenience of EV charging, and vehicle size (or volume). To capture the effect that consumers with higher income can be less sensitive to EV price change (DeShazo et al. (2017)), I add an interaction term of income and EV price P_j.

As for consumer characteristics, the model estimates the impact of income I_i, owning a detached house H_i, density of public charging stations S_i, and an interaction of H_i and S_i for each category g: PHEVs, longer-range BEVs, and shorter-range BEVs. Therefore the expression of $v_{i,g,m}$ is as below:

$$v_{i,g,m} = \gamma_1 I_i P_j + \boldsymbol{\gamma_2 x_j} + \boldsymbol{\beta_g z_i}$$

where I_i is consumer i's income; P_j is price of EV j; $\boldsymbol{x_j}$ is a vector of EV j's characteristics except price; $\boldsymbol{z_i}$ is a vector of consumer i's characteristics. In order to draw conclusions on how consumer characteristics

[17] In the robustness analysis, another nesting structure is illustrated where choice is between PHEV and BEVs, and under the BEV category is further partitioned into LBEVs and SBEVs.

[18] I do not differentiate vehicle model for a given vehicle make, which is constrained by our data. Fortunately, most EV producers only have one model under their production line until year 2017. Exceptions include Ford and Tesla, for which I take the average of vehicle attributes under each make to represent their EVs.

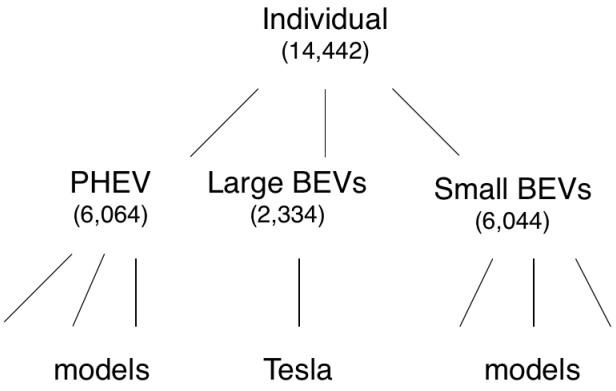
affect preference to different types of EVs, the coefficients of $\boldsymbol{z_i}$ are estimated for each EV group (PHEVs, LBEVs, and SBEVs).

The model here follows the standard nested logit model (McFadden, 1978; Berkovec and Rust, 1985; Goldberg, 1995) and assumes that standard error term ε_{ij} follows generalized extreme value distribution. Consumers first choose within groups of PHEVs, LBEVs, and SBEVs, and then choose an EV model within the group. A tree diagram depicting this decision process is shown in Figure 2.3. The joint probability of choosing an EV (g,m) is:

$$P^i_{g,m} = P^i_g \times P^i_{m|g}$$

where $P^i_{g,m}$ denotes the joint probability of consumer i choosing EV (g,m); P^i_g is the probability of choosing a group from the pool of PHEVs, LBEVs, and SBEVs; $P^i_{m|g}$ represents probability of choosing an EV conditional on the choice of the previous stage.

Figure 2.3: Nested Logit Tree Structure



Following the research cited above, choice probability of (g,m) will be given by nested logit formulas that have the following general form:

$$P^i_{g,m} = P^i_g \times P^i_{m|g} = \frac{\exp(X^i_g\theta_g + I^i_g\lambda_g)}{\sum_{k=1}^{G}\exp(X^i_k\theta_k + I^i_k\lambda_k)} \times \frac{\exp(X^i_{g,m}\theta_{g,m}/\lambda_g)}{\sum_{l=1}^{M_g}\exp(X^i_{g,l}\theta_{g,l}/\lambda_g)}$$

where

$$I^i_g = \log\left[\sum_{l=1}^{M_g}\exp(X^i_{g,l}\theta_g/\lambda_g)\right]$$

The subscript g and m denote for choice on group g and make m respectively. X^i_g represents a vector of explanatory variables of choice g and θ_g is the parameter vector specific to group g to be estimated. M_g is the choice set fo makes given that group g is chosen. I^i_g measures the expected aggregate utility of subset g. Coefficients λ reflect the dissimilarity of alternatives within a nest. As shown by McFadden (1978), if λ's lie within unit interval, the nested structure is consistent with random utility maximization; as λ approaches 1, the distribution of the error term tends towards an iid extreme value distribution and the choice probabilities

are given by multinomial logit model; λ equaling 0 means the error terms are perfectly correlated. In our model, λ for Tesla in the second stage are constrained to equal 1 since they represent the only choice in that nest.

2.5.1.2 Endogeneity

The previous discussion assumes that explanatory variables in the model are independent of the unobserved factors. However, the number of public charging stations could be correlated with the error term, casting doubt on the estimation result from the model previously described. EVs with different electric range have varied demand on public charging station network, especially, we've illustrated that shorter range BEVs would rely on public charging stations more than longer range BEVs and PHEVs. If consumers in a county prefer shorter range BEVs due to reasons unknown to the researcher, and more charging stations are built due to high demand brought by the shorter range BEVs, such reverse causal relationship would make "number of public charging stations" an endogenous explanatory variable.

The following section discusses how to deal with the endogeneity issue. Note that since the choice model is nonlinear, traditional methods such as instrumental variable could not be used. This paper uses the control function approach (Train, 2009; Petrin and Train, 2010), which alleviates the endogeneity problem by including an extra variable that conditions out the variation of unobserved factors that are correlated with the endogenous explanatory variable. To elaborate the control function approach, let's revisit the general form of the utility function (2.2) from previous section, and revise it so that the deterministic part of the utility v_{ij} is expressed as a function of exogenous variables x_{ij}, endogenous variables y_{ij}, and individual specific taste parameters β_i:

$$u_{ij} = v(x_{ij}, y_{ij}, \beta_i) + \varepsilon_{ij} \tag{2.3}$$

Here y_{ij} is the number of public charging stations, which might be correlated with the error term ε_{ij}, and will lead to biased estimation results if the problem left unsolved.

First express the endogenous explanatory variables as a function of observed instruments z_{ij} and unobserved factors μ_{ij}:

$$y_{ij} = w(z_{ij}, \gamma) + \mu_{ij}$$

μ_{ij} and ε_{ij} are independent of z_{ij}, but μ_{ij} is correlated with ε_{ij}. Hence y_{ij} is also correlated with ε_{ij} which is the source of the endogeneity concern. Conditional on μ_{ij}, y_{ij} will be independent of ε_{ij} under such a setting, and this is the key to the control function approach. The following paragraphs discuss choice of z_{ij}.

Regarding to the endogenous deployment of public charging stations, two instrument variables are included. The first one is a dummy variable suggesting whether the EV purchaser's residence county gives financial incentive on installing EV chargers. Such local government policies provides incentives to install more EV chargers, thus bringing an exogenous shock to the chargers stock. The second IV is constructed based on Li et al. (2017). They argue that grocery stores and supermarkets are major providers of public chargers, yet their presence appeared long before the development of EVs, therefore the number of grocery stores and supermarkets brings variation to charger stock and should not be correlated with unobserved vari-

ables in the utility function. For similar reason, other amenities that are potential providers of EV charging stations should be included in IV construction. Here I use the total number of grocery stores, supermarkets, merchandise stores, and restaurants. This instrument satisfies the exclusion restriction because such amenities are established earlier than massive EV adoption. The decision to install EV charging stations involves more than attracting local EV drivers — financial incentives from energy companies and car companies are also a main driver. Moreover, since the number of such places with amenity remains stable overtime whereas charging station network expands rapidly over the years, a year trend variable is also included in the instruments. Suppose consumers' taste towards different types of EVs does not change across year, then year trend should not be correlated with the error terms in the utility function. The first-stage regression results could be found in Table 2.5. The coefficients on both government incentives and number of local amenities are both positive and significant at the 1% level.

Table 2.5: First-Stage Results for Charging Station Equation

Variable	
government incentive on charging stations	0.683***
	(0.126)
number of local amenities	0.0002***
	(0.00003)
year trend	0.0569*
	(0.0298)

Number of observations is 172.
The dependent variable is number of charging stations within 30 km radius of a county in a given year.
* p¡0.1, ** p¡0.05, *** p¡0.01

Regressing the endogenous explanatory variables y_{ij} on the instruments z_{ij} will give us a residual term μ_{ij}, which captures variation of y_{ij} driven by unobserved variables. Further decompose ε_{ij} to the part explainable by μ_{ij} (i.e., the control function, CF) and a residual:

$$\varepsilon_{ij} = \text{CF}(\mu_{ij};\lambda) + \tilde{\varepsilon}_{ij}$$

Plugging in the above expression of ε_{ij} into the utility function (2.3) I will get:

$$u_{ij} = v(x_{ij}, y_{ij}, \beta_i) + \text{CF}(\mu_{ij};\lambda) + \tilde{\varepsilon}_{ij} \tag{2.4}$$

In the utility function (2.4), the error terms are uncorrelated with the explanatory variables by construction, therefore the model can be estimated following standard steps as described in the previous section.

Regarding to the functional form of CF, a straightforward candidate is linear approximation: $\text{CF}(\mu_{ij};\lambda) = \lambda\mu_{ij}$, as is adopted in this paper.

To sum up, the control function approach takes two steps. First, regress the endogenous explanatory variables y_{ij} on the exogenous instruments z_{ij} and obtain the residual terms μ_{ij}. In the second step, add μ_{ij} to the original nested logit model as an extra explanatory variable and estimate the model following the standard approach. Note that in the two-step estimation process, the μ_{ij} used in the second step is an

estimate from the first step rather than the true μ_{ij}, the asymptotic sampling variance of the second step estimator needs to be corrected (Petrin and Train, 2010). Here the bootstrap method is implemented to fix this issue.

2.5.1.3 Results

Results using the control function approach is presented in Table 2.6. The first column shows the estimation result using all observations in the survey (14,420 individuals with 10 EV choices for each person). The results are generally as expected: (1) lower price, lower fuel cost, longer electric range, and larger size will increase the choice probability; (2) interaction term of price and income is positive and statistically significant, suggesting that higher-income consumers are less price sensitive.

The effect of public charging stations is strongest for shorter-range BEVs, which coincides with the theory prediction. The intuition is that shorter-range BEVs rely on their battery as the sole energy source and they have higher probability of getting stranded due to the limited range, hence their reliance on public charging network is the most inelastic. The CF of public chargers is positive and statistically significant, indicating that some factor benefiting EV adoption meanwhile also positively affecting public charger deployment is omitted from the utility function specification.

Owning a detached house gives higher utility for BEVs than PHEVs, which is also intuitive since BEVs rely more on charging facilities than PHEVs and owning a detached house provides the possibility of home charging.

The interaction term of owning a detached house and public charger density is negative, which follows the theory prediction that public charging matters more for drivers without access to home charging. This result is not statistically significant, and it could be because the sample size of EV purchasers not owning a detached house is too small, only about a quarter of those owning a detached house.

The second and third column in Table 2.6 present subsamples containing counties with higher population density (population divided by geographical area) and lower population density respectively. The fourth and fifth column divide the sample to North and South California. The results have similar sign to column (1) using all observations.

Table 2.7 shows the estimation results not using alternative specifications. In Table 2.7, column (1) is the preferred specification same as column (1) in Table 2.6; column (2) drops control function of charging stations; column (3) drops the interaction of income and price; column (4) adds EV model fixed effects thus dropped EV specific characteristics; column (5) adds an additional level of nest such that consumers choose from BEVs and PHEVs first, and the BEV branch is further divided to LBEVs and SBEVs, the rest of the model being the same. Results remain constant with the preferred specification.

2.5.2 Choice between EVs and ICEVs

The above section discusses the question "Conditional on a consumer decides to purchase an EV, which one will she choose". However, one needs to know consumers' decision rule to choose between EVs and conventional ICEVs. Here I use the NHTS dataset and run a binary logit model to determine the factors affecting the choice. Note that ideally one would like to estimate one nested logit choice model with the tree

Table 2.6: Nested Logit Reg. of EV Choice with Control Functions

	Choice Probability				
	All	High Pop. Den.	Low Pop. Den.	North CA	South CA
all EVs					
price(1,000$)	-0.019***	-0.009***	-0.037***	-0.042***	-0.006**
	(0.003)	(0.003)	(0.007)	(0.007)	(0.002)
price(1,000$)	0.083***	0.075***	0.106***	0.123***	0.059***
× income (1,000,000$)	(0.016)	(0.025)	(0.027)	(0.029)	(0.023)
annual fuel cost	-0.983***	-0.891***	-1.284***	-1.898***	-0.469***
	(0.155)	(0.264)	(0.228)	(0.296)	(0.163)
electric range	0.903***	0.827***	1.147***	1.575***	0.463***
	(0.140)	(0.242)	(0.200)	(0.242)	(0.161)
size	0.306***	0.240***	0.464***	0.697***	0.113***
	(0.048)	(0.071)	(0.082)	(0.108)	(0.040)
longer range BEV					
income(1,000$)	0.004***	0.004***	0.003***	0.002**	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
own house	0.660***	0.539*	0.826***	0.595**	0.708***
	(0.184)	(0.325)	(0.274)	(0.279)	(0.248)
charger density	-0.036	-0.259	0.382	0.079	-0.064
	(0.205)	(0.322)	(0.448)	(0.329)	(0.267)
house × charger density	-0.041	0.001	-0.128	-0.034	-0.050
	(0.080)	(0.124)	(0.169)	(0.127)	(0.105)
CF(charger density)	-0.226***	-0.260	0.011	-0.108	-0.418***
	(0.046)	(0.165)	(0.151)	(0.100)	(0.161)
constant	-16.257***	-15.318***	-19.313***	-24.825***	-9.489***
	(2.290)	(4.109)	(3.091)	(3.723)	(2.702)
shorter range BEV					
income(1,000$)	-0.000	-0.000	0.000	0.000	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
own house	0.221**	0.814***	0.071	0.245	0.188
	(0.097)	(0.193)	(0.135)	(0.153)	(0.129)
charger density	0.206**	0.690***	0.659***	0.576***	0.131
	(0.101)	(0.177)	(0.226)	(0.175)	(0.127)
house × charger density	-0.032	-0.208***	0.007	-0.048	-0.004
	(0.042)	(0.068)	(0.088)	(0.070)	(0.053)
CF(charger density)	-0.467***	-0.644***	0.374***	0.307***	-0.883***
	(0.032)	(0.132)	(0.103)	(0.073)	(0.109)
constant	-6.051***	-6.424***	-6.979***	-8.321***	-3.625***
	(0.885)	(1.530)	(1.197)	(1.488)	(1.066)
shorter range BEV λ	0.634***	0.522***	0.891***	1.103***	0.347***
	(0.097)	(0.151)	(0.151)	(0.164)	(0.119)
PHEV λ	2.360***	1.865***	3.774***	6.209***	1.041***
	(0.386)	(0.557)	(0.764)	(1.267)	(0.366)
N	134491	68846	65645	63626	70865

Income is demeaned.

Standard errors in parentheses

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

Table 2.7: Alternative Specifications: Nested Logit Reg. of EV Choice

	Choice Probability				
	(1)	(2)	(3)	(4)	(5)
all EVs					
price(1,000$)	-0.019***	-0.022***	-0.020***		-0.019***
	(0.003)	(0.003)	(0.004)		(0.004)
price(1,000$)	0.083***	0.095***		0.199***	0.083***
× income (1,000,000$)	(0.016)	(0.017)		(0.005)	(0.017)
annual fuel cost	-0.983***	-1.141***	-0.924***		-0.979***
	(0.155)	(0.156)	(0.153)		(0.165)
electric range	0.903***	1.043***	0.854***		0.899***
	(0.140)	(0.141)	(0.140)		(0.150)
size	0.306***	0.354***	0.289***		0.305***
	(0.048)	(0.049)	(0.048)		(0.051)
longer range BEV					
income(1,000$)	0.004***	0.004***	0.007***		0.004***
	(0.001)	(0.001)	(0.000)		(0.001)
own house	0.660***	0.667***	0.660***	0.663***	0.657***
	(0.184)	(0.184)	(0.183)	(0.183)	(0.188)
charger density	-0.036	-0.006	-0.034	-0.037	-0.034
	(0.205)	(0.205)	(0.205)	(0.205)	(0.206)
house × charger density	-0.041	-0.040	-0.041	-0.040	-0.041
	(0.080)	(0.080)	(0.080)	(0.080)	(0.080)
CF(charger density)	-0.226***		-0.226***	-0.225***	-0.228***
	(0.046)		(0.045)	(0.045)	(0.050)
constant	-16.257***	-18.569***	-15.409***		-16.188***
	(2.290)	(2.306)	(2.279)		(2.518)
shorter range BEV					
income(1,000$)	-0.000	0.000	-0.000**		-0.000
	(0.000)	(0.000)	(0.000)		(0.000)
own house	0.221**	0.235**	0.220**	0.232**	0.222**
	(0.097)	(0.097)	(0.097)	(0.097)	(0.098)
charger density	0.206**	0.271***	0.209**	0.208**	0.206**
	(0.101)	(0.100)	(0.101)	(0.101)	(0.101)
house × charger density	-0.032	-0.028	-0.032	-0.029	-0.032
	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)
CF(charger density)	-0.467***		-0.467***	-0.474***	-0.467***
	(0.032)		(0.032)	(0.032)	(0.033)
constant	-6.051***	-6.981***	-5.760***		-6.026***
	(0.885)	(0.891)	(0.886)		(0.960)
EV model dummy	N	N	N	Y	N
Three level nests	N	N	N	N	Y
N	134491	134491	134491	134491	134491

Standard errors in parentheses

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

structure in Figure (2.4). Unfortunately, the dataset of NHTS and CVRP could not be combined, since they contain different variables (mainly granularity of households' geographic location) and the CVRP merely includes EV buyers. A nice feature about the NHTS dataset is that households in different states face various dollar amounts of financial incentives, thus allowing me to identify the impact of financial incentives on choice probability of EVs.

2.5.2.1 Logit Model

Utility of a consumer purchasing an EV is:

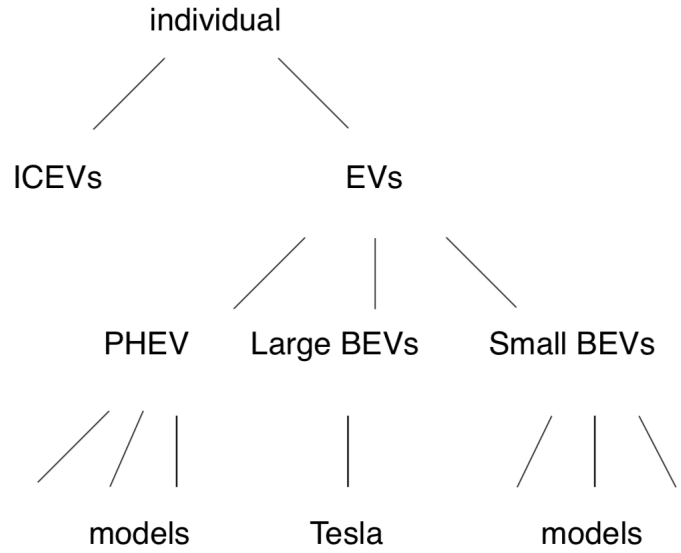
$$u_i = x_i'\beta + \varepsilon_i$$

where x_i' is consumer i specific demographic variable, which includes the dollar amount of financial incentive provided in the residential state, relative density of charging stations (measured by number of charging stations divided by number of gasoline stations in the residential state[19]), the number of vehicles owned by the household, household size, home ownership status, income, etc.

Assuming (1) consumer i will purchase an EV when $u_i > 0$ or an ICEV when $u_i <= 0$ and (2) ε follows logistic distribution, the probability of consumer i purchasing an EV conditional on buying a new car is:

$$p_i = \frac{\exp(x_i'\beta)}{1+\exp(x_i'\beta)} \tag{2.5}$$

Figure 2.4: Ideal Nested Logit Tree Structure



[19]Using state population to normalize the number of charging stations will lead to qualitatively similar results.

2.5.2.2 Results

Results of the logit model described above can be found in Table (2.8), where each column stands for an alternative specification. The implications from various specifications are fairly consistent. First, density of public charging stations have a positive impact on EV adoption, and the result stays after dealing with the problem that charging stations density being endogenous to EV penetration rate. Specification (4) and (5) use logged value of charging station density to capture the effect that marginal benefit of adding public charging stations on EV adoption could be decreasing with the installation base of public charging stations. Second, EV subsidy raises the probability of EV adoption. Specification (4) is chosen as the favorite specification and will be used in counterfactual exercise.

Table 2.8: Logit Regression of Choice between EV/ICEV

	Choice Odds Ratio				
	(1)	(2)	(3)	(4)	(5)
EVs					
charging stations	5.470***	6.101***	7.267***		
	(0.417)	(0.515)	(1.083)		
subsidy ($1,000)	0.176***	0.114*	0.114*	0.113*	0.113*
	(0.048)	(0.059)	(0.059)	(0.060)	(0.060)
home owner	0.522**	0.634***	0.947**	0.638***	1.012**
	(0.232)	(0.230)	(0.398)	(0.229)	(0.419)
income ($1,000)	0.011***	0.010***	0.010***	0.010***	0.010***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
household vehicle count	-0.039	-0.062	-0.061	-0.062	-0.060
	(0.057)	(0.060)	(0.060)	(0.060)	(0.060)
household size	0.080	0.062	0.060	0.061	0.059
	(0.062)	(0.062)	(0.062)	(0.062)	(0.062)
CF(charging stations)		1.012***	1.008***	0.944***	0.934***
		(0.328)	(0.315)	(0.361)	(0.348)
charging stations ×			-1.392		
home owner			(1.237)		
log(charging stations)				7.508***	9.112***
				(0.674)	(1.373)
log(charging stations) ×					-1.899
home owner					(1.550)
constant	-6.863***	-6.417***	-6.690***	-6.501***	-6.830***
	(0.468)	(0.522)	(0.578)	(0.529)	(0.594)
N	29047	29047	29047	29047	29047
log pseudolikelihood	-1.713×10^6	-1.688×10^6	-1.687×10^6	-1.686×10^6	-1.685×10^6

Logit regression is weighted using NHTS sample weight.

Number of charging stations of each state is divided by number of gas stations as normalization.

Standard errors in parentheses

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

2.5.3 Counterfactual Exercise from Discrete Choice Model

With the estimation results from the above sections, we can calculate the choice probability of a specific EV model j by multiplying the probability of purchasing an EV P_e with the probability of purchasing model j conditional on choosing from EVs $P_{j|e}$:

$$P_{e,j} = P_e \times P_{j|e} = P_e \times P_{g|e} \times P_{m|g,e}$$

so that the nested logit structure depicted in Graph (2.4) is estimated separately using two datasets. NHTS dataset gives the choice probability of top level (ICEVs vs EVs) whereas the CVRP dataset pins down the choice probability within the EV class. I conduct several counterfactual exercises by altering the policy provided and see how the choice probability of different group of EVs would vary.

Table 2.9: Counterfactual Exercise from Discrete Choice Model

Panel A: Policy Impact on EV Share				
	status quo	Policy (1)	Policy (2)	Policy(3)
EV share	5.7%	25.4%	5.4%	21.8%
Panel B: Policy Impact on Choice within EVs				
	status quo	Policy (1)	Policy (2)	Policy(3)
share(Large BEVs/EVs)	16.2%	13.9%	16.2%	13.9%
share(Small BEVs/EVs)	41.8%	48.1%	41.8%	48.1%
share(PHEVs/EVs)	42%	42.7%	42%	42.8%
Panel C: Policy Impact on EV Choice Conditional on Buying a New Car				
	status quo	Policy (1)	Policy (2)	Policy(3)
share(Large BEVs)	0.92%	3.53%	0.87%	3.03%
share(Small BEVs)	2.38%	12.2%	2.26%	10.5%
share(PHEVs)	2.39%	10.8%	2.31%	9.31%

Table (2.9) exhibits the counterfactual market share of EVs under alternative scenarios: policy (1): doubling the number of public charging stations; policy (2): eliminate EV subsidies for household with annual income greater than $ 200,000; policy (3): combination of both policy (1) and policy (2). Panel A predicts the probability of buying an EV conditional on buying a new car, using the estimation result from the binary logit model in Section 5.2; Panel B predicts probability of buying a specific group of EV (large BEV, small BEV, or PHEV) conditional on buying an EV, using the estimation result from the nested logit model in Section 5.1; Panel C shows product of Panel A and Panel B, which stands for probability of buying some group of EV conditional on buying a new car. Note that the expression "probability of an individual buying a car" is interchangeable with "predicted market share", under such discrete choice model setting.

Results in Table (2.9) Panel C shows that: (1) doubling the number of public charging stations would increase EV shares (of new cars sold) significantly and such effect is stronger with small EVs than large

EVs; (2) removing the EV subsidy given to the more affluent households would cause a slight decrease in EV shares; (3) combining the two policies together would lead to significant boost in EV shares, relative to the status quo. A back-of-envelope calculation of the cost associated with policy (3) is shown as follows: about 30% of households who purchased EVs have income greater than $200,000 (according to NHTS), eliminating subsidies to such households in California in the year 2015 alone will save about $300 million.[20] Doubling the number of public charging stations in California from the storage in year 2015 will cost about $75 million. [21] Therefore, my estimation result shows that policy (3) which eliminates EV subsidy to higher-income consumers and construct more public charging stations not only leads to higher EV sales, but also costs less than the current subsidy scheme.

2.5.4 Regression with Aggregated Data

The previous section estimates consumer's taste towards different types of EVs, benefiting from the availability of individual level data. Because Census Tract level data are more accurate on EV drivers' geographic location, I test the impact of public charging stations on EV adoption using aggregated Census Tract level data.

2.5.4.1 Model Setup and Endogeneity Issue

For each group of PHEVs, LBEVs (Teslas), and SBEVs (non-Tesla BEVs), I run regressions on Census Tract(CT)-year level observation:

$$\ln(y_{it}) = \beta_0 + \beta_1 \ln(x_{it}) + \mathbf{X}_{it}'\gamma + \delta_t + \sigma_i + \varepsilon_{it}$$

where y_{it} is the number of rebates of CT i in year t, x_{it} is the number of charging stations within 30km distance, $\mathbf{X}_{it}$ are control variables (income, education, household size, travel time to work, etc.), δ_t is year fixed effect, and σ_i is CT fixed effect. The demographic information is averaged at CT-year level, provided by American Community Survey.

The issue of endogenous number of charging stations is similar as described in earlier section. I rescue such concern in two specifications. First, I added CT and year fixed effects to the regression, thus absorbing unobserved Census Tract level characteristics. Second, I add IVs for number of charging stations nearby. The IVs are number of places with amenity (grocery stores, supermarkets, merchandise stores, restaurants) within 30km multiplied by the total number of charging stations outside of 30km radius of CT i in year t.

[20] Federal tax credits amounts to $7,500 and $4,000 for BEVs and PHEVs respectively. California Clean Vehicle Rebate Program provides cash subsidy of $2,500 and $1,500 for BEVs and PHEVs respectively. In 2015, the total number of EV purchase in California is 130,276. Further assume BEVs and PHEVs both occupy half of the EV market share, the total amount of subsidy saved is $130276 \times 30\% \times (0.5 \times (\$7500 + \$2500) + 0.5 \times (\$4000 + \$1500)) = \$302891700$.

[21] In 2015, the stock of public charging stations in California is 2296. Assume 25% of them are DC fast charging stations and 75% of them are level II charging stations. The average cost of constructing a typical DC fast charging station is $50,000, and the cost is $27,000 for a level II charging station Li et al. (2017). The total cost to double the public charging stations in California in 2015 is thus about: $2296 \times (0.25 \times \$50000 + 0.75 \times \$27000) = \$75194000$.

2.5.4.2 Results

From the regressions using aggregated data, one can see that PHEVs and small BEVs benefit most from density of public charging stations. Table 2.10, 2.11, 2.12 exhibit the regression results, in which column (4) is the favorite specification, which includes Census Tract fixed effects, year fixed effects, and IV for charging station density. Note that the geographic information is boiled down to Census Tract level, leaving this analysis the one with most accurate estimate on public charging station density surrounding EV owners.

Table 2.10: Census Tract - Year Panel Regression, PHEVs

	OLS	OLS	OLS	GMM
Variable	(a)	(b)	(c)	(d)
log(no. charging stations within 30km)	0.124***	0.043***	0.146*	0.610***
	(0.014)	(0.009)	(0.073)	(0.141)
log(mean income)	0.135	0.237**	0.381***	0.264***
	(0.098)	(0.111)	(0.106)	(0.077)
log(mean income/median income)	0.245***	0.046	-0.207***	-0.169***
	(0.056)	(0.054)	(0.044)	(0.037)
log(no. detached house)	0.055***	0.016*	0.021***	0.028***
	(0.013)	(0.009)	(0.006)	(0.009)
log(adult population)	0.254***	0.246***	0.151	0.137
	(0.023)	(0.024)	(0.121)	(0.102)
percentage(higher education)	1.057***	1.020***	0.233**	0.161*
	(0.182)	(0.182)	(0.108)	(0.095)
log(time to work)	0.090*	0.171***	0.155*	0.114
	(0.048)	(0.052)	(0.086)	(0.074)
percentage(more than 2 vehicles)	0.426**	0.265	-0.129	-0.193
	(0.207)	(0.241)	(0.121)	(0.136)
year fixed effects	No	Yes	Yes	Yes
census tract fixed effects	No	No	Yes	Yes
N	47448	47448	47448	47442
adj. R^2	0.335	0.501	0.399	0.341

Standard errors in parentheses

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

2.5.4.3 Counterfactual Exercise: Indirect Network Effects and Simulation Results

The following section conducts counterfactual exercise using the aggregated regression results to analyze how the EVs and charging stations will evolve under alternative policy bundles.

2.5.4.3.1 Model Setup The following model is built on Li et al. (2017). We categorize EVs to PHEVs, small BEVs and large BEVs. Assume that EV sales $q_{\theta,t}$ depend on the number of public charging stations in the market N_t, the price of the EV $p_{\theta,t}$, and other product characteristics combined $x_{\theta,t}$; where $\theta \in \{P, L, S\}$ denotes for EV type (PHEV, large BEV, small BEV respectively), and t denotes for time period. The EV stock of type θ in period t is cumulative EV sale less scrappage, denoted as $Q_{\theta,t} = \sum_{h=1}^{t} q_{\theta,h} \cdot s_{t,h}$, where

Table 2.11: Census Tract - Year Panel Regression, small BEVs

	OLS	OLS	OLS	GMM
Variable	(a)	(b)	(c)	(d)
log(no. charging stations within 30km)	0.112***	0.036***	-0.004	0.634***
	(0.019)	(0.012)	(0.070)	(0.168)
log(mean income)	0.127	0.220**	0.616***	0.455***
	(0.083)	(0.096)	(0.143)	(0.090)
log(mean income/median income)	0.146**	-0.049	-0.228***	-0.176***
	(0.061)	(0.063)	(0.038)	(0.031)
log(no. detached house)	0.069***	0.033***	0.004	0.014
	(0.014)	(0.010)	(0.009)	(0.012)
log(adult population)	0.254***	0.244***	0.166	0.145
	(0.027)	(0.029)	(0.121)	(0.094)
percentage(higher education)	1.269***	1.235***	0.401***	0.302**
	(0.136)	(0.145)	(0.129)	(0.120)
log(time to work)	-0.065	-0.003	0.186	0.130
	(0.052)	(0.054)	(0.113)	(0.095)
percentage(more than 2 vehicles)	0.258	0.106	-0.146	-0.234*
	(0.162)	(0.186)	(0.119)	(0.131)
year fixed effects	No	Yes	Yes	Yes
census tract fixed effects	No	No	Yes	Yes
N	47448	47448	47448	47442
adj. R^2	0.344	0.490	0.373	0.186

Standard errors in parentheses

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

$s_{t,h}$ is the survival rate at time t for EVs sold in time h. Assume the survival rate $s_{t,h}$ is δ^{t-h}, where $\delta < 1$. The number of charging stations that have been built $N_t(Q_{P,t}, Q_{L,t}, Q_{S,t}, z_t)$ depends on the stock of EVs and other variables combined z_t that might affect the fixed cost of investment. The EV demand functions and charging stations installment function are specified as below:

$$\ln(q_{\theta,t}) = \beta_{\theta,1}\ln(N_t) + \beta_2 \ln(p_{\theta,t}) + \beta_3 x_{\theta,t}, \text{for } \theta \in \{\text{P,L,S}\} \tag{2.6}$$

$$ln(N_t) = \gamma_1 \ln(Q_{P,t} + Q_{L,t} + Q_{S,t}) + \gamma_2 z_t \tag{2.7}$$

2.5.4.3.2 Feedback Loop The parameters $\beta_{\theta,1}$ and γ_1 capture the magnitude of the indirect network effect on two sides. Feedback loop arises if $\beta_{\theta,1}$ and γ_1 are nonzero. Intuitively, shocks on EV sales in period t will change the installed base of EVs in period $t+1$, causing the number of charging stations to adjust, which in turn affect EV sales in period $t+1$. If both $\beta_{\theta,1}$ and γ_1 are positive, then a policy shock that increases the stock of EVs or charging stations will amplify its effect through the feedback loop.

Table 2.12: Census Tract - Year Panel Regression, large BEVs (Teslas)

	OLS	OLS	OLS	GMM
Variable	(a)	(b)	(c)	(d)
log(no. charging stations within 30km)	0.047***	0.010***	0.118***	0.310***
	(0.007)	(0.003)	(0.025)	(0.079)
log(mean income)	0.212***	0.261***	0.258***	0.267***
	(0.055)	(0.058)	(0.089)	(0.087)
log(mean income/median income)	0.457***	0.365***	-0.112***	-0.111***
	(0.048)	(0.034)	(0.039)	(0.039)
log(no. detached house)	0.028***	0.011*	0.009	0.011
	(0.009)	(0.006)	(0.011)	(0.013)
log(adult population)	0.085***	0.079***	0.001	-0.018
	(0.017)	(0.019)	(0.066)	(0.061)
percentage(higher education)	0.538***	0.512***	0.189**	0.172**
	(0.085)	(0.089)	(0.083)	(0.084)
log(time to work)	-0.022	0.005	0.103**	0.118**
	(0.024)	(0.026)	(0.051)	(0.053)
percentage(more than 2 vehicles)	0.066	-0.012	-0.202**	-0.245**
	(0.107)	(0.117)	(0.101)	(0.123)
year fixed effects	No	Yes	Yes	Yes
census tract fixed effects	No	No	Yes	Yes
N	47448	47448	47448	47442
adj. R^2	0.271	0.352	0.221	0.186

Standard errors in parentheses

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

2.5.4.3.3 Steady State In a longer time span, if a steady state exists where purchases of EVs remain at a certain number cross periods, then $q_{\theta,t} = q_{\theta,t+1} = ... = q^*_\theta$. The equilibrium stock of EV will be $Q^*_\theta = q^*_\theta/(1-\delta)$, given $Q_{\theta,t} = \sum_{h=1}^{t} q_{\theta,h} \cdot s_{t,h}$ and survival rate being $s_{t,h} = \delta^{t-h}$. At the steady state, assume $p_{\theta,t} = p_\theta$, $x_{\theta,t} = x_\theta$ ($\theta \in \{\text{P,L,S}\}$), and $z_t = z$. Combining equations (2.7) and (2.6), equilibrium numbers of EVs (q^*_P, q^*_S, q^*_L) and charging stations (N^*) are defined by the following system of equations:

$$
\begin{aligned}
\ln(q^*_P) &= \beta_{P,1}\ln(N^*) + \beta_2 \ln(p_P) + \beta_3 x_P \\
\ln(q^*_L) &= \beta_{L,1}\ln(N^*) + \beta_2 \ln(p_L) + \beta_3 x_L \\
\ln(q^*_S) &= \beta_{S,1}\ln(N^*) + \beta_2 \ln(p_S) + \beta_3 x_S \\
N^* &= \exp[\gamma_1 \ln(\frac{q^*_P + q^*_L + q^*_S}{1-\delta}) + \gamma_2 z]
\end{aligned}
\tag{2.8}
$$

2.5.4.3.4 Implications on Policy Choices Although difficulties arise when looking for analytical solutions to the system of equations (2.8), we can still draw conclusions from the equilibrium conditions listed in (2.8).

We hereby compare the policy effect of subsidizing EV purchase in several initial periods against building more EV charging stations. Subsidizing EV purchase in initial periods is a common strategy adopted by governments around the globe. Since it is a limited time discount of EVs, studies have shown that such policies greatly led to EV sales increase in the current period. Yet price of EVs will finally return to the original level p_θ once the subsidy budget is exhausted. The equilibrium number of periodical EV sales, cumulative EV stock, number of charging stations would reach the same level as in the case where no subsidy is given. The intuition is that although subsidies in the early periods might speed up EV adoption in the earliest stage, the policy effect will not last as EV stock in the long run is determined by EV characteristics and prices.

Consider an alternative policy that aims at promoting EVs by constructing more charging stations. Through the feedback loop illustrated in the model, this policy will also lead to an increase in EV purchase. The advantage of this alternative policy is that the beneficial effect will last for all periods after. The number of charging stations in the long term will be raised by the policy instead of converging to natural equilibrium level N^* determined in (2.8). The long term EV sales will be higher than the original equilibrium if the number of charging stations with government-enforcing construction surpasses N^*.

By fixing the values of $p_{\theta,t}$, $x_{\theta,t}$, and z_t, and assuming certain values for model parameters as in Li(2017), we conduct simulation to show the effect of feedback loop. The parameters assumed are shown in Table (2.13). The initial EV stock is zero: $Q_{\theta,0} = 0,\ \theta \in \{P,L,S\}$. For each period, we sequentially solve functions (2.6) and (2.7) for $q_{P,t}$, $q_{L,t}$, $q_{S,t}$, and N_t, with condition $Q_{\theta,t} = q_{\theta,t} + \delta Q_{\theta,t-1},\ \theta \in \{P,L,S\}$.[22] Figure (2.5) shows the evolution of three types of EVs and charging stations under this base case.

Table 2.13: Parameters for Simulating Indirect Network Effects

Coefficients	Values	Variables	Values
$\beta_{1,P}$	0.62	p_P	30000
$\beta_{1,L}$	0.32	p_L	70000
$\beta_{1,S}$	0.66	p_S	30000
β_2	-1.2	x_P	16
β_3	1	x_L	18
γ_1	0.6	x_S	16
γ_2	1	z	2
δ	0.9		

To simulate the policy effect of directly subsidizing EV purchase (call it policy A from here), we temporarily deduct the subsidy amount from EV prices for the first 10 periods and solve for $q_{P,t}$, $q_{L,t}$, $q_{S,t}$, and N_t periodically.[23] To mimic the actual policy of Californian EV purchasers, we set the subsidy amount to be \$10,000 and \$5,500 for BEVs and PHEVs respectively.[24]

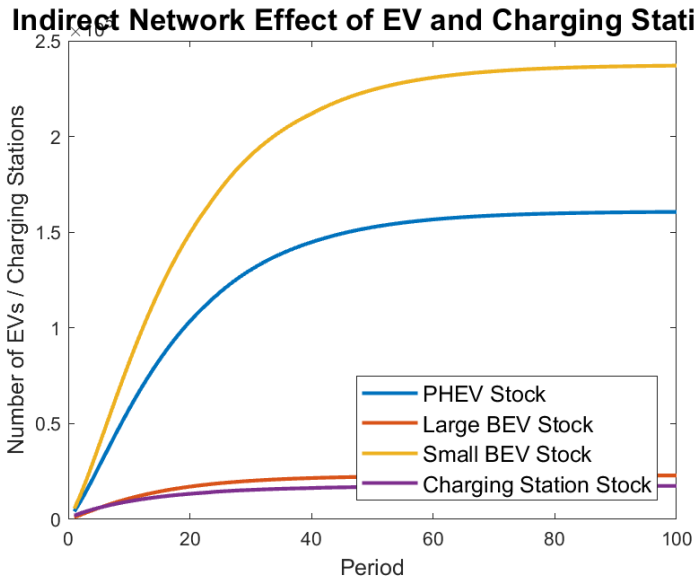
In another simulation, we evaluate policy that uses same amount of money to construct charging stations (call it policy B from here) instead. For each period, we first calculate the subsidy amount which equals

[22] Due to the complex functional form, the simulation results are solved numerically using Matlab.

[23] We assume that charging stock accumulates hence the number of charging stations in each period no less than that in the last period.

[24] US Federal Government provides a tax credit of \$7,500 for BEV buyers and about \$4,000 for PHEV buyers. California Clean Vehicle Rebate Program gives cash rebates of \$ 2,500 for BEVs and \$1,500 for PHEVs.

Figure 2.5: Indirect Network Effect — Base Case Scenario



$5000q_{P,t} + 10000q_{L,t} + 10000q_{S,t}$. Then we calculate number of charging stations that can be built in this period by dividing the subsidy amount by \$27,000.[25] With such policy interfere, number of charging stations $N_{B,t}$ will no longer be endogenously determined by (2.6) and (2.7). When the number of subsidized charging stations surpasses the original equilibrium level, i.e., $N_{s,t} > N_t$, the excess number of charging stations are not supported by the same-period EV stock $Q_{\cdot,t}$ and hence 2.7 does not hold in such case. In the opposite case when $N_{s,t} < N_t$, then we assume that after $N_{s,t}$ charging stations have been constructed by the government, private firms will keep building charging stations to the equilibrium number N_t because of positive profit. Therefore, number of EVs $q_{\cdot,t}$ will always satisfy equation (2.6), whereas number of charging stations is the bigger number between $N_{s,t}$ and N_t.[26]

Figure (2.6) shows the simulated evolution process of EV adoption and charging station construction under three scenarios: a base case, policy A that subsidizes EV purchasing, and policy B that subsidizes charging stations. At least two important conclusions can be drawn from this figure.

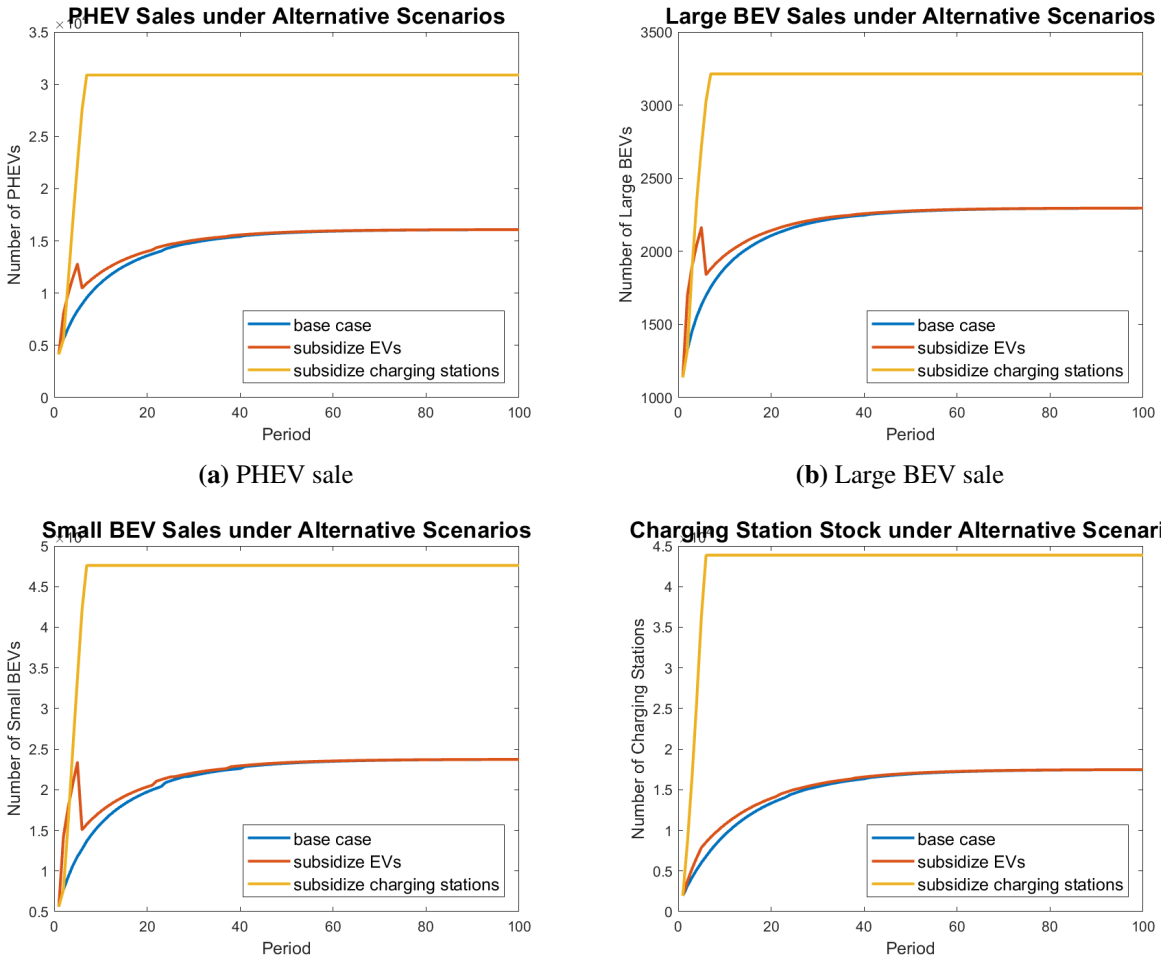
First, policy A has instant effect on EV promotion but does not cause the long run equilibrium of EV sales to increase, whereas policy B's positive impact on EV sales remains after the policy period (10 years). The reason is that factors determining equilibrium q_θ^* and N^* stay the same under policy A; however, under policy B, number of charging stations will be raised exogenously and hence remaining at a higher level than the original equilibrium.

Second, EVs that are more sensitive to public charging stations would experience highest sales increase under policy B.

[25] We follow Li(2017) and assume average cost of building a charging station is at \$27,000. Again we assume that charging station stock accumulates across periods.

[26] The process of calculating EV sales $q_{B,\theta,t}$ and number of charging stations $N_{B,t}$ for each period is: (1) calculate how many charging stations can be built with the government subsidy and add it to the existing charging station stock of last period, denote the resulting number as $N_{s,t}$; (2) calculate $q_{\theta,t}$ and N_t determined by (2.6) and (2.7); (3) if $N_t > N_{s,t}$, then $N_{B,t} = N_t$, $q_{B,\theta,t} = q_{\theta,t}$; (4) if $N_t \leq N_{s,t}$, then $N_{B,t} = N_{s,t}$, $q_{B,\theta,t} = \exp(\beta_{\theta,1}\ln(N_{B,t}) + \beta_2 \ln(p_\theta) + \beta_3 x_\theta)$.

Figure 2.6: $q_{\theta,t}$ and N_t under Different Scenarios



(a) PHEV sale

(b) Large BEV sale

(c) Small BEV sale

(d) Charging Stations

2.6 Conclusion

This paper demonstrates that smaller EVs are more inelastic towards public charging stations than larger EVs, using both individual and aggregated data. The result proves the existence of "range anxiety" of EV drivers — they are afraid of getting stranded on the road due to an exhausted vehicle battery. Such evidence justifies the importance of improving EV charging infrastructure network when the policy maker tries to promote EV adoption.

Since consumers with higher income are less price sensitive, it makes sense to decrease the amount of financial incentive allocated to high income consumer group, so that government subsidies won't be given to "anyway adopters". After estimating EV demand using discrete choice model, I calculated counterfactual EV sales under alternative policies. The result shows that more EV sales could be achieved by eliminating subsidies to high-income consumers and building more public charging stations, meanwhile this policy costs less than the current government subsidy scheme.

Using a feedback model, I showed that constructing more charging stations will lead to more EV sales in

the long run as they can benefit EV drivers through all periods, whereas subsidies only affect current period EV sales.

Moreover, shorter range BEVs have a modest price compared to longer range BEVs, and are more energy efficient compared to plug-in electric vehicles (PHEVs). Therefore subsidizing charging facility instead of individual EV purchasers might contribute more to energy saving and air pollution reduction, whereas reducing the concern that high-income potential buyers of luxurious longer range BEVs might free ride government subsidy.

Chapter 3

The Cost of Commuting Time Reliability: Theoretical Advances using Reliability Standards and Empirical Findings from Metro Vancouver Traffic Data

3.1 Introduction

Despite its compelling economic benefits, congestion pricing remains under-utilized throughout the world. Few major cities have adopted it, but where it is used—Singapore, Stockholm, and London—it is mostly quite effective and motorists have adapted to it well. While resistance to congestion pricing remains widespread in North America, New York is about to introduce congestion pricing in Manhattan (south of Central Park) in 2021.[1] In addition to tolling some bridges and (private) highway sections, there have been a few experiments involving "Lexus lanes": differential pricing for different lanes of the same highway.[2] Hall (2018b) has explored this topic extensively and delivered notable new results that take into account motorists' heterogeneity. This heterogeneity comes in two dimensions: the time cost of commuters, and their schedule inflexibility.

Theoretical and empirical work has highlighted the fact that motorists care not only about the actual driving time, but also about the overall travel time that ensures the reliability of the commute. It matters to commuters that they arrive on time, and thus they choose to arrive early to allow for the dispersion in driving times. Naturally, varying road congestion during different times of the day leads to a distribution of driving time. Motorists may have uncertainty about the road conditions that they encounter and the level of congestion.

Motorists' preferences for commuting time and reliability have been explored widely in the transporta-

[1] New York Times, March 26, 2019.

[2] Washington State's SR 167 high occupancy toll (HOT) lanes; tolls range between 50 cents and $9 depending on congestion level, 5am to 7pm.

tion economics literature, which will be discussed in detail in the following section. We follow a recent strand of literature, in particular Hall (2018b), who separates groups of flexible and inflexible commuters. We also expand on Van den Berg and Verhoef (2011), who take commuter heterogeneity into account and explore distributional outcomes.

This paper contributes to this stream of literature both on the theoretical and empirical level. On the theory side, this paper introduces a novel form of treating the reliability problem. Rather than using a schedule delay[3] cost where the time cost of arriving early or being late needs to be made explicit, our model introduces a *reliability standard* which links a measure of schedule inflexibility to the total travel time that is required to meet the standard. It turns out that the required buffer time (schedule delay) can be approximated well by a linear function of the standard deviation of the road travel (driving) time. Our second innovation is to observe that the standard deviation of the driving time is a function of congestion. Traffic that is slower is also more variable. We believe that this relationship has not been explored rigorously in the prior literature,[4] and we provide a new estimation framework to investigate it. With this observation also comes the question how road traffic is correlated across a road network.

Our empirical work uses data from traffic in the Metro Vancouver area to first identify congestion patterns and estimate speed-volume relationships, followed by identifying how the standard deviation (and coefficient of variation) changes as congestion increases. We find evidence that speed variability increases with congestion levels as well. Difference in reliability may lead to different route choices, which plays a role both for competing routes for motorists as well as for the transportation mode choice. A significant part of our traffic data is obtained from Google's Distance Matrix interface, which provides real-time information about traffic conditions based on data provided voluntarily by users of their map application and Android phones. Our work follows in the growing stream of work that uses geolocation data to infer commuting behaviour (Calabrese et al., 2011; McNeill et al., 2017). In addition to the city-wide analysis, we also focus on a selected number of bottlenecks where we observe speed and volume at hourly intervals. We use a multiplicative heteroskedastic linear model to investigate the effect of congestion on speed and its variance. Lastly, we also explore route choice issues in a simulation exercise that explores issues similar to Hall (2018b) and Van den Berg and Verhoef (2011). We are interested in the sorting effect induced by tolling, especially when competing routes offer a trade-off between time and reliability (one route may be slower but less volatile than another route). Our simulation also investigates the correlation structure of time costs with the degree of schedule inflexibility (for price-inelastic commuters) and with the cutoff time for commutes that defines the extensive margin for price-elastic commuters.

To summarize, our paper makes four novel contributions. First, we introduce a new model using reliability standards to characterize schedule delay costs, and we demonstrate how to turn this into a very effective analytic tool. Second, our empirical work looks at city-wide commuting patterns to identify the importance of commuting time reliability, and we are able to quantify this importance effectively through

[3] Schedule delay refers to the difference between desired and actual arrival time.

[4] Fosgerau and Fukuda (2012) also investigate the variability of travel time, and they express the random travel time as a function of the time of day: $T_t = \mu(t) + \sigma(t)X_t$, where t is time of day, T_t is travel time, $\mu(t)$ is mean travel time, $\sigma(t)$ is interquartile range of travel time, and X_t is a standardized travel time. They find that $\mu(t)$ and $\sigma(t)$ are positively correlated and vary strongly with t. Our paper focuses on traffic volume instead of time of travel, as the root cause of such phenomenon.

our theory-based decomposition into road condition and schedule inflexibility. Third, our bridge traffic analysis demonstrates the effect that the variance of commuting time increases with congestion, and in different ways for different types of bottlenecks. Fourth, our simulation analysis stresses that the magnitude of schedule delay costs can be significant, but that even suboptimal tolls result in strong sorting effects that deliver almost all of the welfare gains from tolling. Even imperfect (limited-information) tolls can be highly effective tolls.

3.2 Travel Time Reliability

Commuters care not only about how quickly they can get to work, but also how reliably. Being late for work may come at a significant penalty if it occurs too frequently. Most employees therefore have a reliability standard in mind when they plan how they commute to work, allowing for sufficient buffer time (known as 'schedule delay' in the literature) to make the journey. Traffic congestion slows down commutes, but it is the variability of traffic congestion that contributes further to the time required for the commute by advancing departure to a time that ensures arrival at the destination at the required time with sufficient likelihood. Our research is predicated on the assumption that commuters have a better understanding of this *reliability standard* than separate and distinct time cost of arriving early or late. This is a crucial point of departure from previous work.

Considering the variability of travel time on different routes, there is a trade-off between average travel time and travel time reliability. Under certain conditions a commuter may prefer the slower but safer route. This argument holds across different route choices as well as across transportation modes. Driving a car to work may be faster on average, but may also come at the greater risk of traffic delays due to congestion or accidents. Some commuters may thus prefer slower but safer public transit options.

Empirical evidence in our paper suggests that increasing congestion not only slows down traffic but also increases travel time variability. Congestion increases the mean travel time as well as the standard deviation of travel time. This link has been previously suggested in the literature, but it has not been explored rigorously in formal models. In this paper we provide empirical evidence about the magnitude of this link, and we explore how it fits with a rigorous micro-economic model of pricing travel time reliability.

Our paper is related to at least three strands of literature. The first strand links travel time reliability to commuters' route choice. [5] There are in general two approaches to model travel time reliability in the transportation literature: the scheduling delay approach and the centrality-dispersion approach.

The scheduling delay approach assumes that commuters leave early and keep a safety margin to avoid being late. The earliest work can be traced to Gaver (1968) and Knight (1974). Small (1982) formulated a model based on a utility maximization framework and defined commuters' utility from travel time and time of arriving early or late. Small's work laid the foundation for the α-β-γ model, where α is the value of travel time, and β and γ are the unit cost of the expected time of arriving early and late, respectively (Arnott et al., 1994). Following this path, Tseng and Verhoef (2008) suggested that value of time varies by time of

[5] In the transportation engineering literature, Iida (1999) is among the earliest to propose the concept of "travel time reliability", which is defined by the probability that traffic can reach a given destination within a stated time. This literature addresses the reliability of a road network whereas we focus more in heterogeneous commuters' behavior.

day, with evidence from stated preference data. Taking the uncertain travel time into account, Asensio and Matas (2008) showed that the value of travel time variability is nonnegligible and highly depends on the time restrictions faced by the individual. Noland and Small (1995) and Noland et al. (1998) extended Small (1982); they further discussed the impact of travel time uncertainty on congestion policies. More recent work in this line includes Bates et al. (2001), Fosgerau and Karlström (2010), and Xiao et al. (2017), which all assume that travel time is a random variable and commuters take an early head start to arrive on time. Our approach is similar, but we specifically assume that commuters have heterogeneous schedule inflexibility. Commuters experience different schedule delays because they adhere to different reliability standards. The traditional α-β-γ model assumes different unit time cost of travelling, and arriving early or late, therefore it is able to account for how much late the commuter is. Our model does not differentiate the unit time cost and assumes a single time cost for each commuter, with the commuter maintaining some probability of arriving on time. Commuters might have better knowledge of their schedule inflexibility — how frequent they are allowed to be late at work, instead of their unit time cost of various activities such as driving, and arriving early or late. This difference allows us to take the dispersion of commuting time into consideration directly. Our approach also provides a straight-forward extension of theoretical modelling of optimal tolls.

Our work has some similarity to Chen and Zhou (2010), who consider both reliability and unreliability aspects of travel time variability for route choices. They assume that with a confidence level α the drivers will arrive on time by leaving early and ensure some buffer time, while they also inspect the worst travel times in the distribution tail of 1–α. Our work shares the focus on explicit probability distribution functions.

The centrality-dispersion approach assumes that commuters have aversion to travel time uncertainty. Related models assume a utility function with a linear combination of mean and variance of travel time. Jackson and Jucker (1982) introduced the mean-variance approach to quantify effects of travel time reliability. They used surveys in which subjects were asked to choose among risky trips with random travel times and discovered that some commuters prefer the more reliable route, even if the expected travel time is higher. Abdel-Aty et al. (1997) followed the above method and used standard deviation as the dispersion measure instead, and they reached similar conclusion. Small et al. (2005) collected traffic data on California State Route 91 in the morning, with both stated and revealed preference data. They estimated a \$21.46/h value of time, and a \$19.56/h value of reliability, where the reliability is defined by the difference between the 80th and 50th percentile travel times.[6]

Recently many researches have investigated the schedule-delay approach together with the centrality-dispersion approach. Carrion and Levinson (2013) showed that reliability of travel time is highly valued through a revealed preference study, using three measures. Engelson and Fosgerau (2016) compared various theoretical measures of the cost of travel time variability. Börjesson et al. (2012) found estimated values of travel time variability differ significantly using different models, and they discussed the assumptions required for each model to be justified. Li et al. (2016) proposed an integrated model of both approaches

[6]While our paper does not follow the centrality-dispersion approach, we mention several important contributions to this strand of literature. Using survey data, de Palma and Picard (2005) find that risk aversion is larger for transit users, blue collars, and for business appointments. Uchida (2014) estimated the value of travel time and reliability based on the risk-averse driver's route choice behavior in a road network. More research work following this path includes Polak (1987); Senna (1994); Devarasetty et al. (2012); Kouwenhoven et al. (2014); Beaud et al. (2016).

and justified its use through stated preference experiments. Travel time reliability is also surveyed elsewhere (Noland and Polak, 2002; Small and Verhoef, 2007; Li et al., 2010; Carrion and Levinson, 2012; Taylor, 2013).

The second strand of related literature focuses on commuter heterogeneity and commuters' valuation of travel time and reliability. Arnott et al. (1994) argued that opposition to congestion pricing has shifted the focus of research from efficiency considerations towards political acceptability (and thus distributional outcomes). Road pricing becomes more viable politically when the tolling scheme benefits almost all urban commuters and does not create intensive losses on the minority. Their research thus focuses on the heterogeneity of agents with respect to time cost, following Vickrey (1969). Key insights are related to the distribution of the drivers with different time costs, notably those with different schedule delay costs. Their work laid the foundations for heterogeneity in commuters' schedule flexibility. Small et al. (2005) estimated motorists' heterogeneous valuation of time and reliability through a discrete choice model, with utility being a function of toll, travel time, and reliability.

Lastly, this paper is also related to the literature on applying tolls to reduce the social cost of congestion. Following the seminal work of Pigou (1920), many have researched how an efficient road tolling system should be constructed. Nevertheless, methods for including travel time reliability in road tolling schemes has not been developed much. Theoretical work on this aspect includes Small and Yan (2001) and Verhoef and Small (2004). Small et al. (2005) and Small et al. (2006) explored toll policy design. Van den Berg and Verhoef (2011) allowed agent preferences to differ across two dimensions: value of time and schedule inflexibility. They demonstrated the possibility to generate a Pareto improvement by pricing a third of the lanes and forgoing revenue by charging a negative toll off-peak. Hall (2018b)'s theoretical work concluded that a carefully designed toll that is applied to some of the lanes of a highway could lead to a Pareto improvement. Hall (2018a) further extended the theory and used survey and traffic data to empirically estimate the joint distribution of road users' value of time, schedule inflexibility, and desired arrival time, and found that the potential social welfare gains from such tolling scheme are substantial. In line with many studies on traffic congestion, Hall's research relies on an important premise that hypercongestion exists (Walters, 1961; Keeler and Small, 1977; Small and Chu, 2003) so that a toll can increase traffic throughput. However, Anderson and Davis (2018) casted doubts on the existence of hypercongestion. Our research approach does not require hypercongestion.[7]

The allocation of welfare gains also remains a central theme in policy discussions about mobility pricing. Classic tolling schemes consider only commuters' value of time, and many researchers concluded that commuters with higher valuation of time would gain by paying efficient tolls, whereas commuters with lower valuation of time could suffer a loss prior to toll revenue redistribution (Cohen, 1987; Arnott et al., 1994).

[7]The notion of *hypercongestion* was introduced by Vickrey to refer to situations where traffic flow is inversely related to traffic density. Network flow follows an inverse U-shape relationship with respect to network density. Network flow first increases as network density rises, reaches a peak, and as network density rises further leads to a fall in network flow and possibly complete gridlock. Travel on the positively-sloped portion of the inverse-U is considered 'congested,' and on the negatively-sloped portion it is considered 'hypercongested.' There is a new class of models that address hypercongestion issues. Arnott (2013) introduced the bathtub model for heavily congested downtown areas where traffic jams cause traffic flow to fall, and in which traffic density (the level of the bathtub) is experiencing either normal (low-level congestion) or hypercongested traffic. Traffic velocity is negatively correlated to traffic density. Time-varying congestion pricing can convert traffic jams into queues for entering into the downtown area.

More recently, Van den Berg and Verhoef (2011) consider bivariate continuous distributions of value of time and schedule delay in their tolling scheme. They found that a majority of users could be left better off even before toll revenue redistribution, with the commuters with intermediate value of schedule delays and lower value of time suffering the greatest loss. Eliasson (2016) discussed the fairness of congestion charges from distinct perspectives. Besides the traditional consumers' perspective which includes monetary and time cost in traveling, the author also discussed the citizens' perspective that consider whether individuals' views of social issues are aligned with policy makers' objective. The conclusions from both perspectives are similar and suggest that middle-income groups win the most. Ke and Gkritza (2018) employed the Suits Index to investigate impact of congestion tolling on urban and rural households. They found that congestion pricing can be progressive on income for urban households, whereas regressive with regards to rural households, with the difference caused by driving distance and income. Our simulation process will also use Suits Index to inspect the equity aspect of the tolling scheme.

To sum up, we provided empirical evidence that increasing traffic volume increases not only the travel time, but also (in most instances) its standard deviation. The estimation of this relationship using real-time traffic data contributes to the literature of travel time reliability. Theoretically, our paper introduces an innovative measure of how travel time reliability affects commuters' cost function. Our model of scheduling delay specifically allows for heterogeneity of commuters' time value and schedule inflexibility, which enables us to assess the welfare impact of tolling on different groups of commuters. The welfare effects from tolling reveal that the inflexible group of commuters benefit more than the flexible group in general, and that more affluent commuters benefit more either by paying an efficient toll or through receiving a per-capita toll refund when they abandon commutes.

3.3 Theory

Developing theory that is useful for empirical work typically requires choosing appropriate functional forms that lead to tractable and estimable algebraic expressions. While much work has been done on the theoretical foundations of optimal road tolling, we employ a set of assumptions that on one hand allow for sufficient flexibility and on the other hand lead to empirically testable implications.

3.3.1 Approaches to Travel Time Reliability

The pertinent literature on travel time reliability refers to the travel time cost as the α-β-γ model (Arnott et al., 1994), where α is the unit cost of expected travel (driving) time, β is the unit cost of the expected time of arriving early and γ is the unit cost of the expected time of arriving late. Following the approach pioneered by Vickrey (1973) for heterogeneous groups, the ratios α/β and α/γ are assumed identical across groups. Van den Berg and Verhoef (2011) provides a recent application of this type of model, which provides for very convenient interpretations of tolls when the heterogeneous commuters follow uniform distributions.

This approach circumvents an important complication: what is the empirical distribution of commuting time and schedule delay costs? There are two parts to this question: how to quantify the cost per unit of time, and how to quantify the probability of being late or arriving early. Essentially, there are different ways on how to approach the γ part for the cost of arriving late.

First, a motorist may incur a per-minute cost for arriving late that is (significantly) higher than the ordinary time cost. For this one needs to model explicitly the tail of the distribution for late arrival. The motorist minimizes total cost, which is the sum of (1) driving time priced at ordinary time cost, (2) early-arrival time priced at a early-arrival cost, and (3) late-arrival time priced at the late-arrival cost. This may be dubbed the "variable-cost lateness model" that is common in the literature.

Second, a motorist may incur a fixed penalty each time when arriving late. This again requires identifying the tail of the distribution for late arrival. The motorist minimizes total travel cost, but the late-arrival component is the product of late-arrival probability and penalty. This is the "fixed-cost lateness model."

Third, a motorist may simply adhere to a reliability standard: plan driving so that arriving late occurs infrequently and meets the motorist's *schedule inflexibility*. This is the "reliability standard model." We choose to take this third route because it (a) does not require explicit knowledge of the late cost (either fixed or variable), and (b) leads to particularly intuitive solutions. Our model shares the important features with Lo et al. (2006) which they called "travel time budget model".[8] While they assume travel time follows a normal distribution, we use log-normal and Fréchet distributions that have long right tails and fit the data better. We arrive at a linearization technique that disaggregates the time mark-up factor into a route-specific coefficient of variation and a multiplier that reflects commuters' inflexibility. Our cleaner functional form lends itself more readily to empirical work.

We begin by defining a motorist's schedule inflexibility $\psi \geq 0$ so that a $\psi = 0$ indicates that a commuter does not care about being late and any positive number indicates increasing schedule inflexibility. In probabilistic terms, a value of $\psi = 0$ will be associated with the median travel time. The odds of being late are given by once in $(\psi + 2)$. We assume that travel time T is distributed with a cumulative distribution function $\Phi(T)$. Thus, the motorist chooses time to depart before desired arrival time T^* (driving time T plus 'buffer time' to ensure timely arrival) so that the probability of arriving on time is

$$\Phi(T^*) \leq \frac{\psi + 1}{\psi + 2} \in [1/2, 1) \tag{3.1}$$

Specifically, we assume that travel times follow the long-tailed Fréchet distribution with minimum travel time $T^\circ > 0$, shape parameter $a \in (0,1)$, and scale parameter $S > 0$ so that

$$\Phi(T; T^\circ, a, S) = \exp\left(- \left[\frac{T - T^\circ}{S} \right]^{-1/a} \right) \tag{3.2}$$

The Fréchet distribution has the mean $\bar{T} = T^\circ + S \cdot \Gamma(1-a)$ and the variance $\sigma^2 = S^2(\Gamma(1-2a) - \Gamma^2(1-a))$. A commuter facing the schedule inflexibility ψ will therefore choose a travel time T^* that obeys

$$T^* = T^\circ + S \cdot \chi^a \quad \text{with} \quad \chi \equiv \left[\ln\left(\frac{2+\psi}{1+\psi} \right) \right]^{-1} \tag{3.3}$$

[8] In Lo et al. (2006), the travel time budget is defined as sum of expected travel time and travel time margin: $b_p = E(T_p) + \lambda \sigma_{T_p}$, where T_p is the random variable of travel timeon route p, $E(T_p)$ and σ_{T_p} are mean and standard deviation of T_p, λ is a parameter that can be related to the probability that a trip arrives within the travel time budget. $P\{T_p \leq b_p = E(T_p) + \lambda \sigma_{T_p}\} = \rho$, where ρ is the probability that the actual trip time is within the travel time budget.

where the buffer time is the difference between T^* and the mean travel time $\bar{T}$. The transformation of ψ into χ is nearly linear for large schedule inflexibility ψ; for example, for $\psi = 250$ the value of $\chi = 251.5$ is just 0.006% larger than ψ.

Consider the following numerical example. A commuter's mean travel time is 45 minutes, with a standard deviation of 6 minutes, and a best feasible travel time of 30 minutes. Then $S = 12.38$ and $a = 0.24$. The parameters a and S can be recovered numerically from solving the two equations for mean and variance of the Fréchet distribution. Assuming that the commuter has a reliability requirement of $\psi = 50$ (the commuter can afford to be late to work five times a year), then $T^* = 30 + 12.38(51.5)^{0.24} = 61.8$ minutes. The commuter should allow for an extra 17 minutes for the typical 45-minute commute in order to arrive on time.

The distributional parameters T°, S and a are all functions of road congestion $\tilde{v} \in [0,1)$ and therefore need to be extracted from the estimated values of mean, variance, and skewness as functions of $\tilde{v}$, where $\tilde{v}$ equals traffic volume divided by road capacity. We note that observing standard deviation and mean of travel time — T_σ and T_μ, given T°, allows backing out $a < 1/2$ directly through

$$1 + T_\rho^2 = \frac{\Gamma(1-2a)}{\Gamma^2(1-a)} \quad \Longrightarrow \quad a(T_\rho) \tag{3.4}$$

where the coefficient of variation is defined as

$$T_\rho \equiv \frac{T_\sigma}{T_\mu - T^\circ} \tag{3.5}$$

The value of a increases with T_ρ. Having solved for a numerically immediately reveals $S = (T_\mu - T^\circ)/\Gamma(1-a)$. (N.B.: $\Gamma(1-a) \in [1, 1.772]$ for $a \in [0, 0.5]$.) The optimal time to depart before the desired arrival time is

$$T^* = T^\circ + (T_\mu - T^\circ)\frac{\chi^a}{\Gamma(1-a)} \tag{3.6}$$

We revisit this equation after introducing the speed-volume relationship.

The Fréchet distribution is only one among many plausible choices. The shifted log-normal distribution can be parameterized similarly with three parameters as

$$\Phi\left(\frac{\ln(T - T^\circ) - \mu}{\sigma}\right) \leq \frac{1+\psi}{2+\psi} \tag{3.7}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function and so that

$$\ln(T^* - T^\circ) = \mu + \sigma \cdot \tilde{\chi} \quad \text{with} \quad \tilde{\chi} \equiv \Phi^{-1}\left(\frac{1+\psi}{2+\psi}\right) \tag{3.8}$$

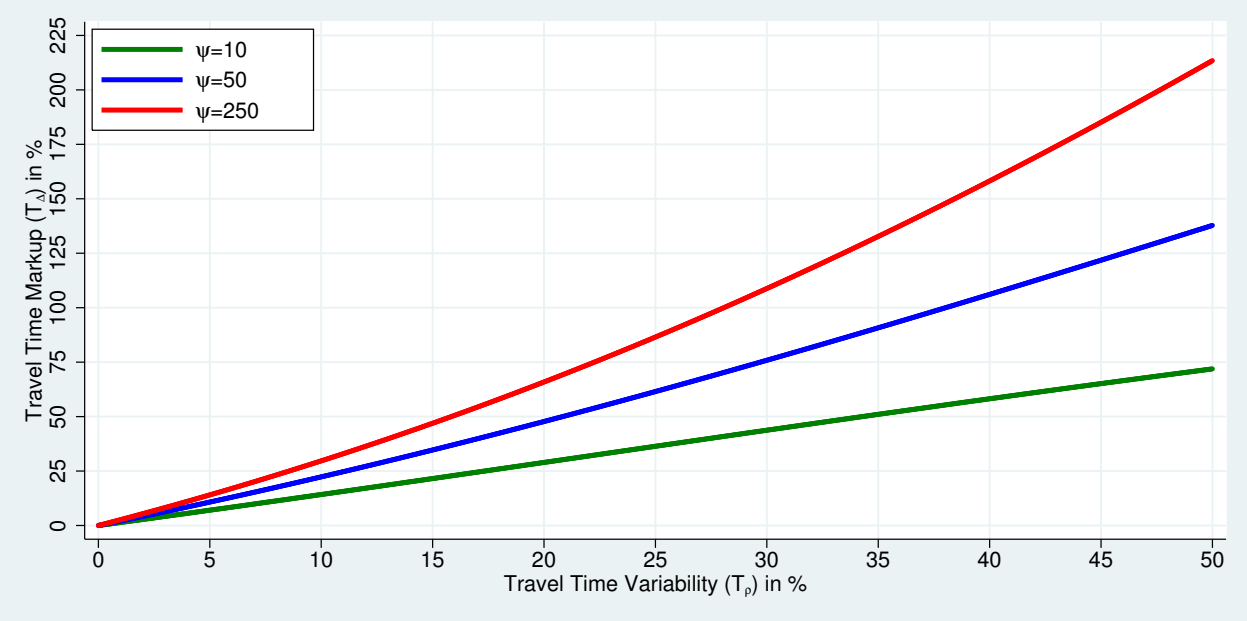
and moments $\mu = \ln(T_\mu) - \sigma^2/2$ and $\sigma^2 = \ln(1 + T_\sigma^2/(T_\mu - T^\circ)^2) = \ln(1 + T_\rho^2)$ that can be obtained from the observational mean T_μ and and standard deviation T_σ. Thus, using the coefficient of variation T_ρ for

simplification,

$$T^* = T^\circ + (T_\mu - T^\circ)\frac{\exp\left(\tilde{\chi}\cdot\sqrt{\ln(1+T_\rho^2)}\right)}{\sqrt{1+T_\rho^2}} \tag{3.9}$$

Consider again the numerical example with a commuter's mean travel time of 45 minutes, with a standard deviation of 6 minutes, and a best feasible travel time of 30 minutes. Then $T_\rho = 0.4$. With $\psi = 50$ the standard normal cumulative distribution function yields $\chi = 2.07$. Then $T^* = 30 + (45-30)(2.220/1.077) = 60.9$ minutes, or about 16 minutes more than the average trip time. The log-normal and Fréchet distributions give very similar results.

Figure 3.1: Travel Time Variability and Travel Time Mark-Up



Note: based on log-normal distribution

It is helpful to illustrate the relationship between congestion variability and trip time. The coefficient of variation T_ρ is a dimensionless number. Similarly, the dimensionless metric

$$T_\Delta \equiv \frac{T^* - T^\circ}{T_\mu - T^\circ} - 1 = \frac{T^* - T_\mu}{T_\mu - T^\circ} \tag{3.10}$$

is a measure of the travel time reliability mark-up: how much extra time is added for reliability relative to the mean-to-minimum time difference. Figure 3.1 shows the mark-up for three values of ψ, 10 (late once every two weeks), 50 (5 times late per year) and 250 (once late per year), for the log-normal distribution. In the numerical example with T_ρ=40%, the markup for $\psi = 50$ is 106%. The fact that T_Δ is a dimensionless number proves to be helpful with the theoretical analysis but doesn't immediately tell how much time a commuter needs to add. The more intuitive and practical measure of the time markup over average travel

time is

$$T_\Lambda \equiv \frac{T^* - T_\mu}{T_\mu} = T_\Delta \left(1 - \frac{T^\circ}{T_\mu}\right) \approx H(\psi) \cdot \frac{T_\sigma}{T_\mu} = H(\psi) \cdot T_\nu \tag{3.11}$$

where $T_\nu \equiv T_\sigma / T_\mu$ is the ordinary coefficient of variation. In the above numerical example with $H(\psi)$=2.65 and T_ν=0.133, the time markup is 35.3%, or 16 minutes on top of the average 45 minutes.

For empirical research as well as further theoretical modelling it is helpful to observe that the relationship between T_ρ and T_Δ is nearly linear for any given ψ. Therefore, one may use an approximation:

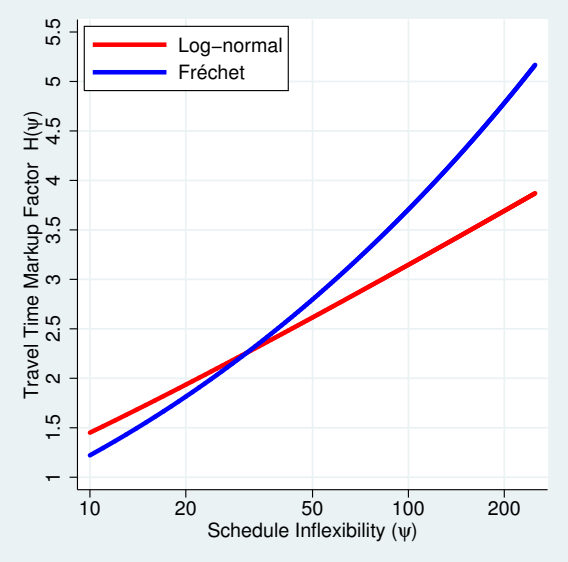
Theorem 1 (Travel Time Reliability Markup) *The travel time markup T_Δ for commuters with schedule inflexibility ψ is related to the parameters T°, T_μ, and T_σ of the full-trip travel time distribution (and implicitly the coefficient of variation T_ρ) approximately through*

$$T_\Delta \approx H(\psi) \cdot T_\rho \quad \Longleftrightarrow \quad T^* \approx T_\mu + H(\psi) \cdot T_\sigma = T_\mu (1 + T_\Lambda) \tag{3.12}$$

where $H(\psi)$ is a positive and increasing function in ψ. The full travel time is larger than the mean driving time by an amount that is proportional to the standard deviation of driving time by the factor $H(\psi)$.

The above theorem suggests a particular schedule delay cost in the notation of the α-β-γ travel cost model. It replaces the late cost with a reliability standard and treats the early arrival period the same as the commuting period.

Figure 3.2: Travel Time Markup Factor



This linearization is extremely useful for practical applications, and at the same time it is highly intuitive. But how good it is this approximation empirically? The R^2 statistics for the underlying regressions are better than 0.99, which means that the linearization is easily justified. However, the different empirical distributions deliver different $H(\psi)$, as is illustrated in figure 3.2, which shows the schedule inflexibility

measure ψ on a logarithmic scale on the horizontal axis and the empirically-estimated travel time markup factor $H(\psi)$ on the vertical axis. For very low levels of schedule inflexibility the markup factor is near one and rises to below 4 for the log-normal distribution and above 5 for the Fréchet distribution.

The link between ψ and $H(\psi)$ is nearly perfectly log-linear for the log-normal distribution, and only slightly less so for the Fréchet distribution.The Fréchet distribution provides a link between ψ and $H(\psi)$ that is slightly steeper than the log-normal distribution and it delivers larger values of $H(\psi)$ than the log-normal distribution when ψ gets large. Table 3.1 provides approximations for the log-linear relationship between ψ and $H(\psi)$ for both distributions and estimation ranges of [10,250] and [50,250] respectively, corresponding to figure 3.2.

Table 3.1: Log-Linear Approximation of $H(\psi)$

Distribution	Approximation	Fit (R^2)	Range (ψ)
Lognormal	$H(\psi) = 0.7656 \cdot \ln(\psi) - 0.3698$	0.9997	[10,250]
Lognormal	$H(\psi) = 0.7824 \cdot \ln(\psi) - 0.4534$	1.0000	[50,250]
Frchet	$H(\psi) = 1.3229 \cdot \ln(\psi) - 2.2870$	0.9894	[10,250]
Fréchet	$H(\psi) = 1.4998 \cdot \ln(\psi) - 3.1711$	0.9975	[50,250]

To summarize, the theoretical model yields insightful links between the statistics T_Δ and T_ρ through the markup factor $H(\psi)$, while practical applications will use the direct link between effective time markup T_Λ and the simple coefficient of variation T_v. Empirical work therefore needs to establish (1) which distributional assumption (Fréchet, Log-normal) provides a better approximation in order to choose the appropriate $H(\psi)$ approximation; and (2) how large the coefficient of variation T_v is in order to quantify the buffer time that commuters add to their journeys. The two coefficients of variation T_v and T_ρ are empirically quite different in nature. Our empirical analysis shows that T_v and T_ρ are virtually uncorrelated because the denominators are quite different in nature. While T_ρ measures the time variation T_σ with respect to the average *delay* $T_\mu - T^\circ$, T_v measures it with respect to the average *time* T_μ. Our theoretical analysis that suggests linearization hinges on the properties of T_ρ rather than T_v, while the practical implications for the extra schedule time hinges on T_v.

3.3.2 Route Choice

Commuters take their entire travel route into account. Their total travel time is the sum of road segments $k \in \mathscr{K}$ in route $\mathscr{K}$ so that $\sum_{k\in\mathscr{K}} T_k$ with variance $\sum_{k\in\mathscr{K}} \sum_{l\in\mathscr{K}} \sigma_k \sigma_l \rho_{kl}$ with standard deviation σ_k of road segment k and ρ_{kl} the correlation coefficient between individual road segments along route $\mathscr{K}$. The nature of this correlation is important. It can be positive when high traffic volume is spread throughout the road system. It can also be negative when congestion at an "upstream" bottleneck is holding back traffic flow and is thus relieving congestion "downstream". For the commuter it is the total variance that matters, not the partial variance at a particular road segment.

The commuter faces a choice between competing routes and will take both speed and reliability into

account. First consider two untolled roads $\mathscr{K} \in \{1,2\}$ connecting A and B. The commuter chooses route

$$\text{argmax}_{\mathscr{K}} \{T^*_{\mathscr{K}}\} \tag{3.13}$$

which may or may not be the same as choosing the route that is faster on average $\text{argmax}_{\mathscr{K}}(T^{\mathscr{K}}_{\mu})$. Using the approximation (3.12), it is possible to describe the choice between two routes 1 and 2 where we assume that route 1 is, on average, faster than route 2: $T^1_\mu < T^2_\mu$. If a commuter chooses routes only based on average travel time, that commuter will always choose route 1. However, when variability comes into play, the commuter may prefer route 2 if it is more reliable than route 1 even though it is slower on average. Then $T^*_2 > T^*_1$ implies $T^2_\mu + H \cdot T^2_\sigma > T^1_\mu + H \cdot T^1_\sigma$. Then the following result holds:

Theorem 2 (Route Choice with Schedule Inflexibility) *If route 1 is on average faster than route 2 so that $T^1_\mu < T^2_\mu$, then a commuter prefers route 2 over route 1 if the ratio of the difference in standard deviation to the difference in mean travel time exceeds the inverse of $H(\psi)$:*

$$\frac{T^1_\sigma - T^2_\sigma}{T^2_\mu - T^1_\mu} \gtrapprox \frac{1}{H(\psi)} \tag{3.14}$$

As schedule inflexibility ψ increases and $1/H$ gets smaller, the commuter will prefer route 2 for ever smaller differences in travel time variability. The numerical values for $1/H$ range from 0.82 (Fréchet distribution and $\psi = 10$) to 0.19 (Fréchet distribution and $\psi = 250$).

Route choice involves multiple road segments. For the commuter only the entire trip matters, and individual road segments influence the overall trip planning. The main empirical complication arises from the fact that studying total travel time requires looking at entire trips rather than road segments individually because of the correlation structure that is captured by the correlation coefficients ρ_{kl}. The implication for empirical work is that it is not sufficient to merely investigate the variability effect of road congestion at individual points in the road network, but also the entire correlation structure.

3.3.3 Speed-Volume Relationship

The starting point for modelling congestion is a relationship between travel time T and traffic volume v on a particular road. Travel time (T) and speed (V) are linked through the relationship $V = L/T$ where L is the length of the road. Thus travel time and speed can be used interchangeably. Empirical work mostly uses speed, as this is easier to measure, while theoretical work prefers time, as this is more meaningful conceptually.

Speed-volume relationships have been studied extensively and there are numerous functional forms that have been explored. For theoretical purposes, an important choice is about whether to account for *hypercongestion*. The presence of hypercongestion leads to a backward-bending speed-volume relationship that allows for two speed points for a given traffic volume: a (normal) congestion speed and a hypercongestion speed. The presence of hypercongestion is a critical element in some models, especially in models that follow the Arnott (2013) "bathtub" model and capture traffic flows in dense urban cores.

In our paper we work with Metro Vancouver traffic data that does not show strong evidence of hyper-

congestion at key bottlenecks. We focus our analysis on situations with moderate levels of congestion, but we acknowledge that our results will not extend to hypercongestion scenarios. We do not require hypercongestion as a condition for deriving our results.

We explore two functional forms for the speed-volume relationship that avoid hypercongestion and merely allow for congestion to approach the capacity limit, at which point travel speed approaches zero.[9]

The first functional form is taken from Pipes (1967). For a given road capacity c, travel time T is a function of uncongested travel time $T^{\odot}$, the road capacity utilization rate $\tilde{v} \equiv v/c \in [0,1]$, and the curvature coefficient $\xi > 0$ so that

$$T(v) = T^{\odot}(1 - v/c)^{-\xi} \tag{3.15}$$

The curvature ξ determines the congestivity of the road. As ξ increases, the roadway is easier and easier to congest. In this version of the equation, values for ξ will be between zero and one, and in practice are quite small. However, while this type of speed-volume equation and its cousins can be fitted relatively easily to the data when it comes to the mean travel time, we are interested in dispersion and thus need to be careful about how individual speed observations are scattered above or below the mean for a particular volume. We will therefore need to adjust this equation to fit our empirical distribution functions and allow for a minimum travel time T°. Conceptually, there is a "safe limit" on speed, which may well be above the posted speed limit as the latter is set to control speed in traffic and not at free flow when the road is nearly empty.

The second functional form we consider is a generalization of the bureau of public roads(BPR) function (Bureau of Public Roads, 1964):

$$T(v) = T^{\odot}\left[1 + \kappa\left(\frac{v}{c}\right)^{\lambda}\right] \tag{3.16}$$

By virtue of having three parameters, this form can provide a better fit for the data. However, it turns out that estimating it in our framework will require non-linear techniques.

Our modelling is predicated on the assumption that there is a minimum travel time $T^{\circ} < T^{\odot}$.

Even when a road is completely uncongested, there is variation around the travel time because of weather conditions, time of day and lighting, or other factors. Our empirical travel time distribution thus requires separating minimum and average travel time, and allowing this to vary under congestion. We therefore reformulate (3.15) as a stochastic equation

$$\frac{T(v_t) - T^{\circ}}{T^{\odot} - T^{\circ}} = \left(1 - \frac{v_t}{c}\right)^{-\xi} \exp(\varepsilon_t) \tag{3.17}$$

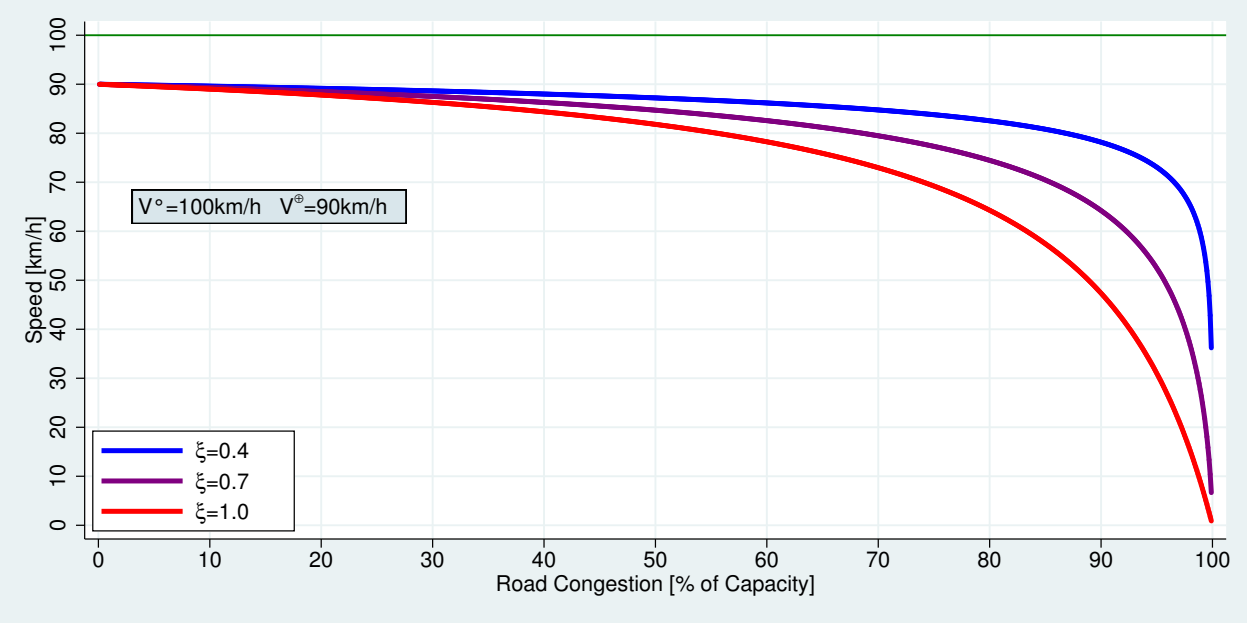
where ε_t is an error term that is distributed normally with mean zero. The variance of the error term is assumed to be independent of congestion, but when we turn to our empirical work we allow it to vary with congestion as well. In equation (3.17), the new variable $T^{\odot}$ (roughly, the average speed without any congestion) becomes an estimable parameter in addition to ξ. This formulation allows us to link up the congestion function directly with our log-normal distribution function. When $\ln(T(v) - T^{\circ})$ is distributed normally, we can estimate a congestion function for the mean commuting time, and we can additionally

[9] We have also tried fitting Vancouver bridge traffic data using other functional forms illustrated in Greenshields et al. (1935); Drake and Schofer (1966), which are outperformed by the two models we finally choose.

allow for the variance of commuting time to depend on congestion.

The modification in equation (3.17) is of significance. Effectively, we shift the coordinates of the distribution so that $\forall t : T(v_t) \geq T^{\circ}$. Comparing (3.17) with (3.15), the power term on the right hand side scales up $T^{\odot} - T^{\circ}$ rather than T°, and the former is much smaller in magnitude than the latter. The result is that we can expect much higher estimates of ξ when we estimate (3.17).

Figure 3.3: Speed-Volume Function



Equation (3.17) can also be written in terms of speed rather than time, as showing speed in a diagram is often more insightful than showing time for a given distance. Ignoring the effect of the stochastic term ε_t, the velocity $V(v)$ as a function of traffic volume v is given by

$$V(v) = V^{\circ}\left[1 + \left(\frac{V^{\circ}}{V^{\odot}} - 1\right)\left(1 - \frac{v}{c}\right)^{-\xi}\exp(\varepsilon_t)\right]^{-1} \tag{3.18}$$

The speed-volume chart in figure 3.3 illustrates the relationship between $\tilde{v}$ and speed for V°=100km/h and $V^{\odot}$=90km/h and three different ξ. When we estimate equation (3.17), our dependent variable will become $\ln(T_t/T^{\circ} - 1)$, and as we observe speed, we replace this (identical in value) with $\ln(V^{\circ}/V_i - 1)$.

3.3.4 Speed-Volume Dispersion

A central theme in this paper is the empirical fact that congestion not only increases average travel time but also increases the variability of travel time. The key metric for dispersion is the standard deviation of travel time T_{σ}, or its normalized cousin, the coefficient of variation T_{ρ} as defined in (3.5).

The analysis in the preceding section introduces functional forms (3.15) and (3.16) that capture the effect of congestion on travel time. We have modified (3.15) to capture variation in commuting time and be consistent with log-normal distributions through (3.17). This new formulation allows us to draw a link

between variance and mean by assuming that T_ρ, the coefficient of variation for commuting time mark-up, is approximately constant. Then $T_\sigma = T_\rho(T_\mu - T^\circ)$. Using the linearization insight expressed in theorem 1 and equation (3.12) with $H(\psi)$ as a function of the schedule inflexibility parameter, we are able to derive a theorem about the commuting time under congestion

Theorem 3 (Commuting Time with Congestion and Schedule Infexibility) *The commuting time for commuters facing schedule inflexibility ψ and travel time coefficient of variation T_ρ is approximated by*

$$T^*(\tilde{v}) = T^\circ + (T^\odot - T^\circ) \cdot G(T_\rho, v) \cdot (1 + T_\rho H(\psi)) \tag{3.19}$$

if the traffic congestion follows (3.17) and the coefficient of variation of travel time T_ρ is approximately independent of the level of congestion.

As we will show in the empirical discussion of the paper further below, the function G is derived from the properties of the log-normal distribution function and is given by

$$G(T_\rho, v) = \frac{\sqrt{1 + T_\rho^2}}{(1 - v/c)^\xi} \tag{3.20}$$

and follows from the assumption of log-linearity of $T - T^\circ$. Equation (3.19) compares directly to the commuter who is flexible. The log-linear distribution implies that the mean commuting time (without scheduling inflexibility) is

$$T_\mu(\tilde{v}) = T^\circ + (T^\odot - T^\circ) \cdot G(T_\rho, v) \tag{3.21}$$

The difference of time that is added to the commute as buffer time is therefore given by $(T^\odot - T^\circ) \cdot G(T_\rho, v) \cdot T_\rho \cdot H(\psi)$. Quantifying this will be important to gauge the effect that scheduling inflexibility has on the total cost of commuting.

The assumption that T_ρ is independent of the level of congestion is perhaps strong. We find empirical evidence that supports this notion for across-city commutes. However, we find some traffic bottlenecks where road congestion also increases variance more than proportionally and thus T_ρ may increase with congestion.

The discussion above only accounts for one dimension of dispersion: namely that for a given level of congestion there are varying speed outcomes. There are several sources for this type of dispersion. The composition of traffic can be different in terms of trucks and cars. There can be differences in weather conditions that interact with the level of congestion. These physical reasons can all contribute to dispersion that varies with the level of congestion.

There is also a second type of dispersion, which could be described as "congestion uncertainty." Assume that the speed-volume relationship does not exhibit dispersion in the time dimension. Rather, the commuter does not know which volume of traffic she will encounter on a given day due to random demand shifters. This means that a commuter may expect a certain amount of traffic on average, but with random dispersion. Observationally, this does not make any difference for the planning by the commuter. Whether the dispersion is caused by speed dispersion or volume dispersion, ultimately it is the commuting time that varies.

3.3.5 Optimal Road Toll and Agent Heterogeneity

The optimal road toll problem has been described extensively, and what is added here is a more involved treatment of agent heterogeneity in the presence of time cost dispersion and schedule inflexibility dispersion. There are several ways of modelling the commuter's choice problem: as a welfare maximization problem, as a travel cost minimization problem, or a utility maximization discrete choice problem. Our starting point is a travel cost minimization problem to derive the optimal tolls.

The planner's objective is to choose a vector $\boldsymbol{\tau}$ of optimal link tolls to minimize total travel cost for the $i \in \{1,..,n\}$ individuals, each choosing an individual route R_i with path elements $\delta_{ik} = 1(k \in R_i)$, and choosing to travel $\pi_i(\boldsymbol{\tau}) \in \{0,1\}$ or not, so that the total social cost reflects individuals' time cost ω_i, schedule inflexibility ψ_i, and cost z_i of not travelling. The non-travel cost can also be viewed as the maximum duration $T_i^{\varnothing}$ that a commuter is willing to spend on this trip, which implies $T_i^{\varnothing} = z_i/\omega_i$. All toll revenue is returned to commuters, so from the point of view of the planner the toll revenue does enter into the cost problem. This is an important consideration as otherwise the toll revenue could influence the commuting decision.

It is practical to group commuters into those with a flexible schedule ($\psi = 0$) and those with an inflexible schedule (and a common $\psi \gg 0$) so that the first group has an extensive margin and respond to prices, while the second group will always commute. Let α denote the fraction of inflexible commuters and $1-\alpha$ the fraction of flexible commuters. Further denote as $\pi(\boldsymbol{\tau}) \in [0,1]$ the fraction of flexible commuters that choose to drive after seeing posted tolls.

Aggregating these groups and averaging them yields the social cost function

$$\mathscr{C} = \min_{\boldsymbol{\tau}} \left[\alpha \bar{\omega}^* \bar{T}^*(\boldsymbol{\tau}) + (1-\alpha)\bar{\omega}\left(\pi(\boldsymbol{\tau})\bar{T}(\boldsymbol{\tau}) + (1-\pi(\boldsymbol{\tau}))\bar{T}^{\varnothing}\right)\right] n \tag{3.22}$$

where $\bar{\omega}$ is the average time cost of flexible commuters and $\bar{\omega}^*$ is the average time cost of inflexible commuters, and where $v_k = \sum_i \delta_{ik}$ is the number of commuters who take road k. Furthermore, $\bar{T}^*$ and $\bar{T}$ are the average travel times of the inflexible commuters (including their schedule delay) and the flexible commuters (who choose to travel), respectively. We also introduce $\bar{T}^{\varnothing}$ under the assumption that z_i and ω_i are uncorrelated, an assumption we will relax in our empirical work. However, for expositional purposes we use this assumption to derive analytically tractable results.

Total traffic volume is $v = n[\alpha + (1-\alpha)\pi(\tau)]$. Therefore, the share of flexible commuters that commutes is

$$\pi(\tau) = \frac{(v/n) - \alpha}{1-\alpha} \tag{3.23}$$

Inflexible commuters choose route R_i

$$R_i = \text{argmin}_{\delta_{ik}} \left\{ \omega_i^* T^*\left(\boldsymbol{v}(\boldsymbol{\tau}), \psi\right) + \sum_k \delta_{ik}\tau_k \right\} \tag{3.24}$$

while flexible (inframarginal) commuters choose route R_i

$$R_i = \text{argmin}_{\delta_{ik}} \left\{ \omega_i T\left(\boldsymbol{v}(\boldsymbol{\tau})\right) + \sum_k \delta_{ik}\tau_k \right\} \tag{3.25}$$

conditional on making the trip

$$\pi_i = 1\left(\omega_i T\left(\boldsymbol{v}(\boldsymbol{\tau})\right) + \sum_k \delta_{ik}\tau_k < z_i + \varepsilon_i\right) \tag{3.26}$$

The above implies that a commuter makes the trip when the total cost of making the trip (travel time cost and tolls) is smaller than the cost of avoiding the trip. This is simply another way of putting the discrete choice problem and could be expressed similarly in utility terms.

If ε_i is assumed to be distributed as an independent and identical type I extreme value (i.e., Gumbel), the conditional logit model emerges with an aggregate probability of making trips given by

$$\pi = \left[1 + \exp(u(\bar{\omega}(\bar{T}^{\varnothing} - \bar{T}(\pi)) - \bar{\tau}))\right]^{-1} \tag{3.27}$$

This provides a solution to the optimal traffic flow v. Inverting the above demand function yields the optimal toll:

$$\bar{\tau} = \bar{\omega}(\bar{T}^{\varnothing} - \bar{T}(\pi)) - \frac{1}{u}\ln\left(\frac{\pi}{1-\pi}\right) \tag{3.28}$$

Unambiguously, the optimal toll increases when a lower travel volume (lower π) is required in optimum.

The cost minimum for $\mathscr{C}$ is found from the first-order condition for the optimal traffic volume v. In turn, the travel demand function $v(\tau)$ reveals the optimal toll given optimal traffic volume v. It is important to acknowledge the economic intuition underlying this principle of minimizing cost with respect to traffic volume rather than optimal toll. The size of the toll does not matter for welfare because all revenue is returned to economic agents. As long as the returned funds do not influence commuting behaviour, the two-stage process of minimizing costs with respect to traffic volume and then finding the optimal toll for the required volume is economically sound.

Before returning to the general case of a road network, consider the simple case of a single road. To obtain a closed-form solution for the first-order condition, it is useful to introduce the variable travel time $(T - T^{\circ})$ elasticity

$$\eta \equiv \frac{\mathrm{d}(T - T^{\circ})}{\mathrm{d}v}\frac{v}{T - T^{\circ}} = \xi\left[\frac{\tilde{v}}{1-\tilde{v}}\right] \tag{3.29}$$

with respect to traffic volume v. We also introduce the time cost ratio $\varphi \equiv \bar{\omega}^{*}/\bar{\omega}$ to capture asymmetric time costs. Note that $\varphi > 1$ when inflexible commuters have a higher time cost than flexible commuters.

We also make use of our congestion function (3.17), and we introduce a function

$$U(v) \equiv \left[\frac{T^{\varnothing} - T^{\circ}}{T^{\odot} - T^{\circ}}\right]\frac{(1 - v/c)^{\xi}}{\sqrt{1 + T_{\rho}^{2}}} - 1 \tag{3.30}$$

which decreases with congestion v. The term in square brackets is a time ratio: the difference between non-commute threshold and minimum time relative to the difference between average free-flow time and minimum time. Further recall that $\eta(v)$ increases in v, and thus $U(v)/\eta(v)$ decreases with v. Thus the first-order condition for a cost minimum reveals the following theorem:

Theorem 4 (Optimal Single-Road Traffic Volume) *For a single road travelled by α inflexible commuters (with average time cost of $\bar{\omega}^*$ and average travel time $\bar{T}^*$) and $1-\alpha$ flexible commuters (with average time cost of $\bar{\omega}$, average travel time $\bar{T}$, and maximum trip time $\bar{T}^{\varnothing}$), the optimal traffic volume is given by*

$$\frac{v}{n} = \alpha \left[\frac{\varphi(1+H(\psi)T_\rho)-1}{U(v)/\eta(v)-1} \right] \tag{3.31}$$

which can be solved for v numerically.

We are interested in understanding how changes in the share α of inflexible commuters, as well as changes in their degree of inflexibility ψ, affect the optimal traffic volume. We are also interested to know how differences in travel cost between the two types of commuter matter. Is it reasonable to assume that time costs are the same ($\varphi=1$), or will they differ substantially? It is clear that an increase in φ has a similar effect as an increase in α, as both enter multiplicatively in the numerator.

The expression in square brackets in (3.31) must be greater than one because the share α of inflexible commuters always commutes. For $v/n > \alpha$, it must hold that $\varphi(1+T_\rho H(\psi)) > U(v)/\eta(v)$. This says that either the schedule inflexibility or the time cost ratio for the inflexible commuters must be large enough to generate a solution that reduces the commutes of flexible commuters. Furthermore, we require that numerator and denominator are both positive. For the numerator, this means that the schedule inflexibility ψ must be large enough to compensate for a time cost ratio $\varphi < 1$ where the flexible commuters have a higher cost than the inflexible commuters.[10]

Equation (3.31) is nonlinear in v and thus the effect of key parameters on optimal traffic volume can only be identified through implicit function calculus. Defining (3.31) as $F(v,\alpha)=0$, it follows that

$$\frac{\mathrm{d}v}{\mathrm{d}\alpha} = -\frac{\partial F/\partial \alpha}{\partial F/\partial v} < 0 \quad \text{and} \quad \frac{\mathrm{d}v}{\mathrm{d}\psi} = -\frac{\partial F/\partial \psi}{\partial F/\partial v} < 0 \tag{3.32}$$

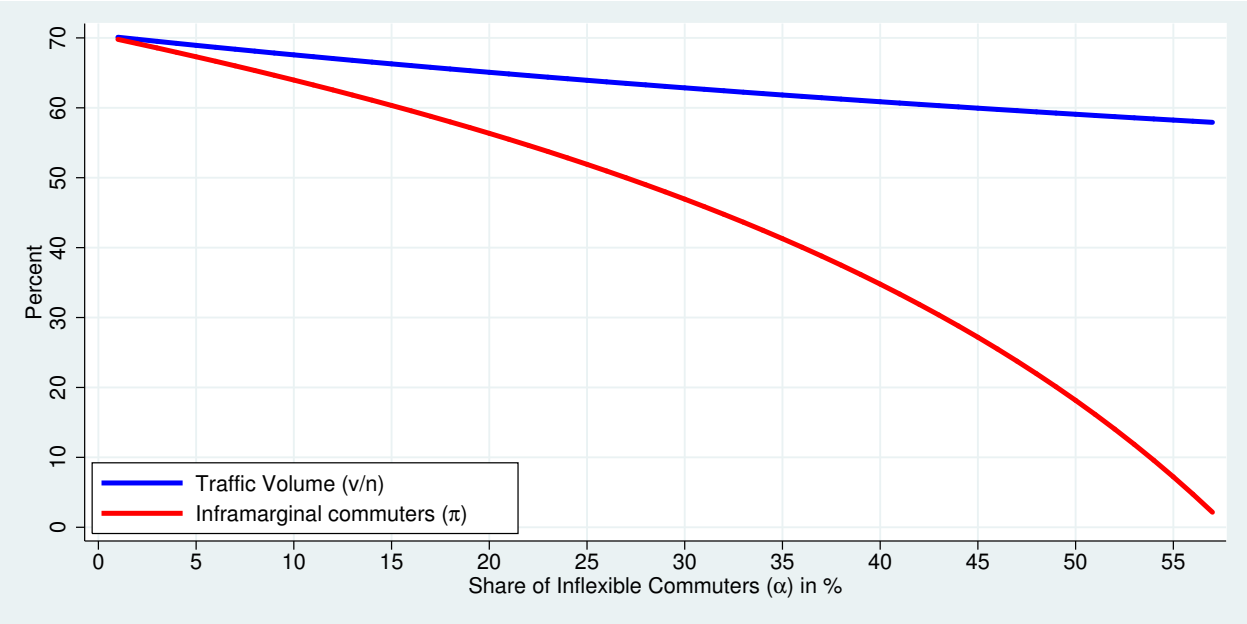
Figure 3.4 illustrates the effect of inflexible commuter share on optimal traffic volume. It shows that as the share of inflexible commuters increases, the optimal traffic volume decreases gently. However, as the proportion of flexible commuters $1-\alpha$ gets smaller and smaller, a larger and larger number of them have to refrain from commuting to enable the optimal traffic volume to be attained. When α reaches about 58% in this numerical simulation, it is optimal to induce all flexible commuters to stop commuting.

Figure 3.5 depicts the effect of an increase in schedule inflexibility on the optimal traffic volume. It uses the same parameters as before except that α is fixed at one-half, and the factor $H(\psi)T_\rho$ is varied from 0 to the point where all inframarginal commuters stop commuting. As schedule inflexibility increases, the relative importance of the inflexible commuters increases because these commuters add more and more extra time to ensure timely arrival. As a result, the number of inframarginal commuters is pushed towards zero, while the total share of commuters approaches the lower bound of α.

When there are no inflexible commuters ($\alpha = 0$), there is a particularly simple solution for the optimal

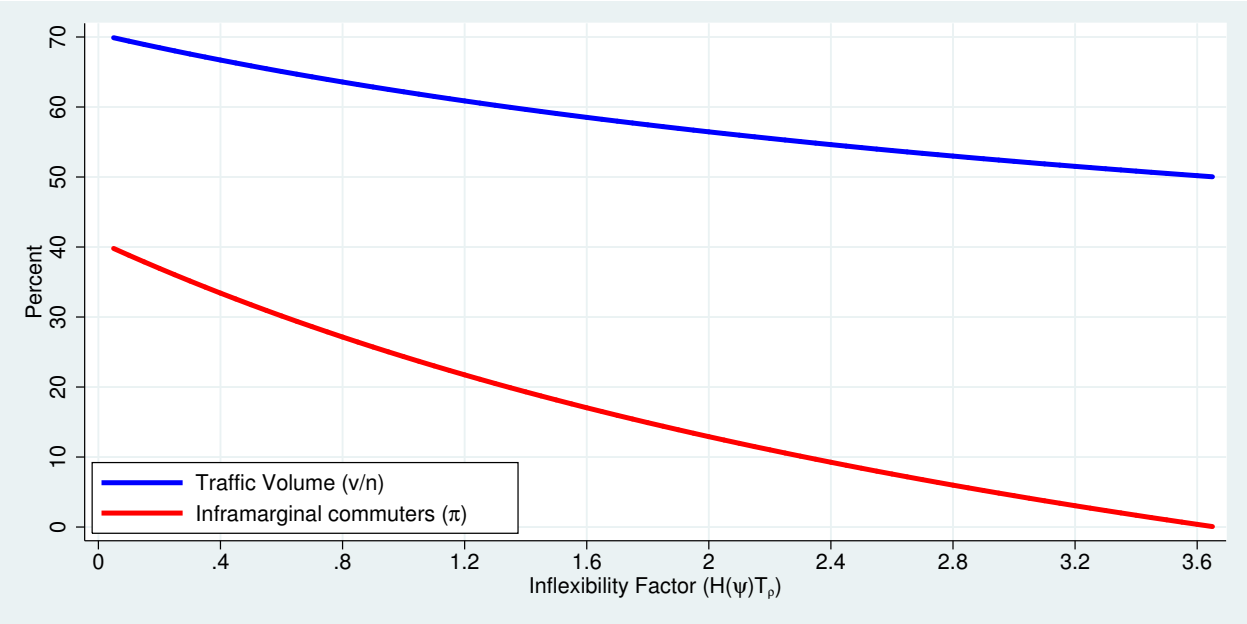
[10] If the time cost of flexible commuters is high and the schedule inflexibility of the other commuters is low, then it makes no economic sense to curtail the commutes of flexible drivers.

Figure 3.4: Effect of Inflexible Commuter Share on Optimal Traffic Volume



Parameters: n=1900, c=2000, κ=2, λ=4, $\bar{\omega}$=$\bar{\omega}^*$=25, $H(\psi)$=6, T_ρ=0.25, T°=30, $\bar{T}^\varnothing$=90.

Figure 3.5: Effect of Schedule Inflexibility on Optimal Traffic Volume



Parameters: n=1900, c=2000, κ=2, λ=4, $\bar{\omega}$=$\bar{\omega}^*$=25, α=0.5, T°=30, $\bar{T}^\varnothing$=90.

vehicle flow:

$$\frac{v}{c} = \left[\frac{1}{1+\xi/U(v)}\right] = \frac{\pi(\tau)n}{c} \tag{3.33}$$

This non-linear equation can be solved for v.

3.3.6 Commuters' Buffer Time

A planner who fails to take into account the buffer time (the time mark-up) of inflexible commuters will solve a cost minimization problem where the planner falsely assumes that $H = 0$. Because of (3.31), the linearization result that is implied there suggests that the numerator is treating changes in the time cost ratio φ and changes in the mark-up factor $1 + T_\rho H(\psi)$ exactly the same. This means that there is a fundamental equivalence in the way a researcher can treat the time mark-up for inflexible drivers and their time cost:

Theorem 5 (Reliability Pricing Equivalence) *The optimal traffic volume (and thus optimal toll) depends directly on*

$$\varphi^* \equiv (\bar{\omega}^*/\bar{\omega}) \cdot [1 + T_\rho H(\psi)] = \varphi(1 + T_\Delta) \tag{3.34}$$

and therefore a higher time cost of inflexible drivers is observationally equivalent with a higher time-markup T_Δ.

The implication is that a road toll that only allows for the differential time cost of inflexible drivers, but not their time mark-up, will result in a suboptimal outcome as if the time cost of inflexible drivers had been underestimated by a factor of $1 + T_\Delta$. This result has important implications for empirical research. Researchers need to estimate the "effective time cost" of inflexible commuters by estimating their nominal time cost $\bar{\omega}^*$ and the degree of their schedule inflexibility ψ.

We can use theorem 1 about travel time reliability markup to characterize the total cost, noting that $\bar{T}^* = \bar{T}[1 + H(\psi)T_v]$, where $T_v = \bar{T}_\sigma/\bar{T}$ is the ordinary coefficient of variation of commuting time. Then

$$\mathscr{C} = n\bar{\omega}\bar{T}\left[\alpha\varphi(1 + H(\psi)\bar{T}_v) + (1-\alpha)(\pi + (1-\pi)\Omega)\right] \tag{3.35}$$

where $\Omega \equiv \bar{T}^\varnothing/\bar{T} > 1$ is the ratio of the threshold time for non-commuting to average commuting time. Road tolls determine the share of flexible commuters π entering the road and decrease average travel time $\bar{T}$ and the coefficient of variation $\bar{T}_v$. In the absence of tolling, $\pi = 1$. Then we can compare the cost with reliability costs against the case where reliability costs are ignored (and thus $H(\psi) = 0$). Dividing the two cost measures yields a measure of the schedule time cost:

Theorem 6 (Schedule Time Cost) *Given the share of inflexible commuters α, schedule inflexibility ψ, commuting time coefficient of variation $\bar{T}_v$, and time cost ratio φ of inflexible to flexible commuters, the extra cost (in percent) induced by the presence of schedule time is given by*

$$\mathscr{A} \equiv 100\% \cdot H(\psi) \cdot \bar{T}_v \left[1 + \frac{1-\alpha}{\alpha\varphi}\right]^{-1} \tag{3.36}$$

in the absence of road tolls.

Effectively, equation (3.36) allows us to put an upper limit on the extra cost due to schedule inflexibility; it is $H(\psi)\bar{T}_V$. Furthermore, when the time cost of flexible and inflexible commuters is the same, then the extra cost due to schedule time is simply proportional to the share of inflexible commuters so that $A = 100\% \cdot \alpha H(\psi)\bar{T}_V$.

3.3.7 Reliability Pricing in a Road Network

The results obtained so far paint a very simple picture about how reliability pricing can be brought into a framework for optimal road tolling in a road network, which itself has been well established in the literature already. The key modifications derived from the analysis in this paper is that researchers need to identify:
(a) the composition of commuters (the share α of inflexible commuters);
(b) the relative time cost (φ) of inflexible and flexible commuters;
(c) the schedule inflexibility (ψ) of commuters; and
(d) the coefficient of variation (T_V) of commuting trip time.

It is the latter that is specific on individual routes, and thus empirical research will need to identify the reliability of each road segment and identify how the variance of individual road segments adds up across the multitude of trip patterns in a region. This is where the empirical research in this paper attempts to provide some novel insights by studying commuting patterns in the Metro Vancouver region. Just how large are these coefficients of variation across commuting trips? While our study does not identify the three other key parameters (α, φ, and ψ), which require other microeconomic data sets from surveys, we aim to identify T_ρ and T_V from observed traffic flows.

3.4 Data

We utilize data from a variety of sources. From Statistics Canada's National Household Survey 2016 we observe commuting patterns among pairs of census tracts (CTs). A CT is a small areas that captures between 2,500 and 8,000 persons in census metropolitan areas (CMA). We employ data for the Metropolitan Vancouver area (CMA code 933) area that encompasses the cities listed in table B.2 with a total of 482 census tracts. The commuting patterns in the household survey only identifies 43,763 CT pairs with a minimum number of commutes. We restrict our analysis to commutes with origin and destination CT both in Metro Vancouver and exclude those trips that start or end outside Metro Vancouver.

We obtain travel time information from the Google Maps Distance Matrix API. This Application Programmer Interface (API) provides information on distance, normal travel time, and time in traffic for origin-destination pairs, which we identify by longitude and latitude as the mid-points of census tracts. The API employs a specific traffic model for calculating the time in traffic. Google employs live traffic data, which it acquires from users of their navigation applications that share location information in real time. We use their "best guess" model, whose reported duration in traffic is their best estimate of travel time given what is known about both historical traffic conditions and live traffic.

In addition to the region-wide commuting pairs we also focus on several bottlenecks of traffic flows—bridges and tunnels—which are well known to experience high levels of congestion during rush hour. Only one of these, the Port Mann Bridge, was tolled between 2012 and 2017. We employ post-toll data, but note

that the removal of the toll on September 1, 2017 led to a significant increase (about 30%) in traffic volume. The main feature of our bottleneck data is our ability to match travel time through these bottlenecks in both directions with volume measurements from the BC Ministry of Transportation Traffic Data Program.

3.4.1 Commutes among Census Tract Pairs

We sample 2,000 "typical" commuting trips and ensure that the sampling frequency is matched to the volume of commutes reported in the Household Survey. Thus CT pairs with many commuters will be sampled proportionally more than those with fewer commuters. Figure 3.6 shows the cumulative distribution of CT pairs in descending order of traffic volume reported by the household survey. We sample by taking 2,000 routes so that the frequency with which we sample routes is weighted by the traffic volume.[11]

Figure 3.7, 3.8 and 3.9 describe the overall distribution of the sampled routes of census-tract pairs. The distributions for distance and commuting time are noticeably right-skewed and thus have means that are higher than their medians. In our sample the mean commuting distance is 13.5km (8.4mi) and takes 21 minutes. The average speed that we observe is about 35.6km/h (22.2mi/h); however some commutes can be as slow as 10km/h and as fast as 70km/h.

Our commuting data lends itself to summarizing and analyzing region-wide patterns. We first look at average commuting speed and commuting delay and try to explain how it varies during rush hour. Our sample of 2000 trips across Metro Vancouver during seven half-hour intervals between 06:30 and 10:00 provides estimates of how the log of average speed and the average delay (measured as the log of the ratio of observed travel time to normal travel time as reported by the Google API) depends on distance, time of day, and day of week. Table 3.2 summarizes the results.

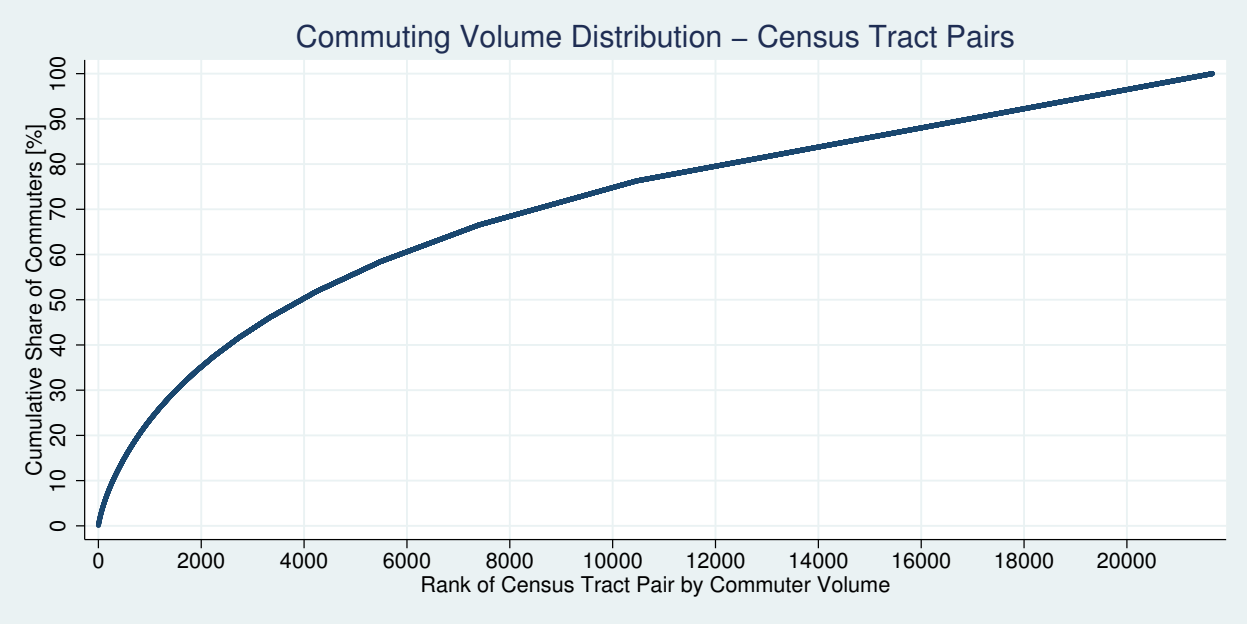
Column (A) in table 3.2 reveals that average speed increases with distance so that a 10% increase in distance increases average speed by 2.4%. This effect is a result of longer commutes containing longer sections of highway driving. The time of day effect is captured by a linear and square time effect that suggests that speed is lowest (and delay is highest) around 08:30. Delays increase very little with distance. A 10% increase in distance is associated only with an 0.5% increase in delay. The best days for commuting are Mondays and Fridays, whereas other days of the week are slower. Weekends are excluded from the analysis. We cluster standard errors in our analysis.

Column (B) in table 3.2 captures the delay, the log ratio of travel time in traffic to normal travel time. Both figures are provided directly by the Google API. This means that under ideal conditions the observed travel time may be higher or lower than the normal travel time. We again find a link to distance, but much weaker than for speed. A 10% increase in distance increases delays by about half a percent. Delays are smaller on Mondays and Fridays compared to the three other weekdays.

We use the available time distribution data in several ways. For each commuting trip we calculate minimum travel times, and for each CT pair and time slot (2,000 times 7) we calculated mean and standard deviation in order to obtain empirical estimates of T_ν and T_ρ that we analyze further. We also use the CT-pair time-slot data set to estimate the parameters of the underlying Fréchet and Lognormal distributions in order

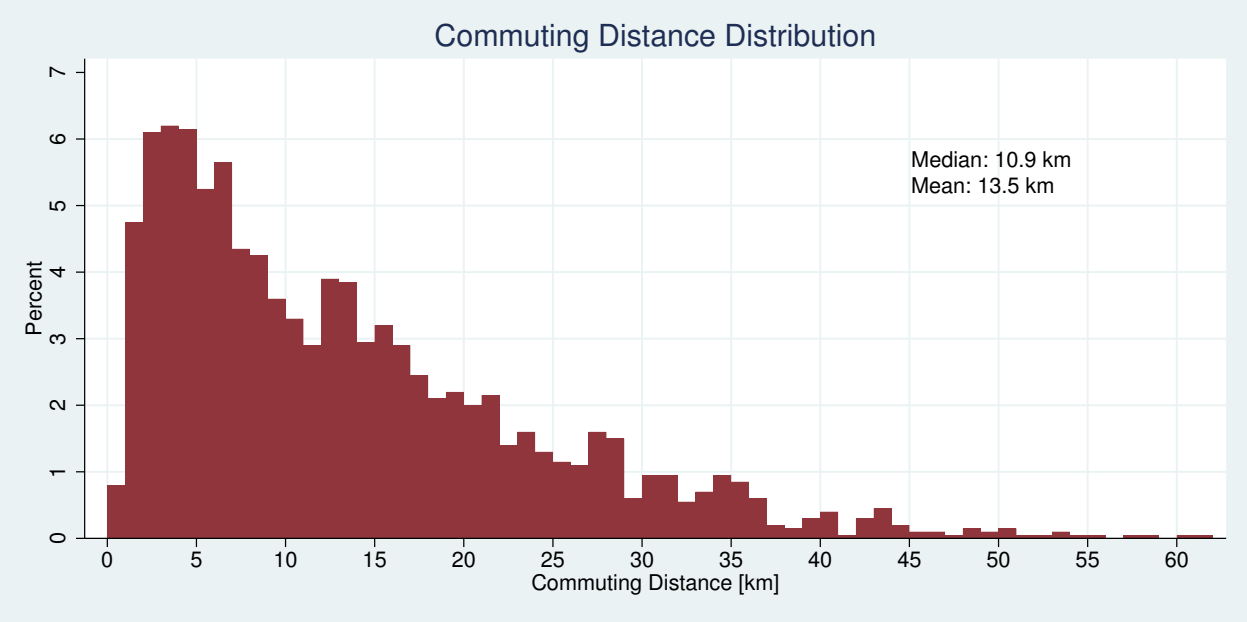
[11] We split the cumulative distribution into 2,000 segments by taking a route sample each time we cross the next 0.05% threshold of cumulative volume.

Figure 3.6: Census Tract Commuting Trip Sample



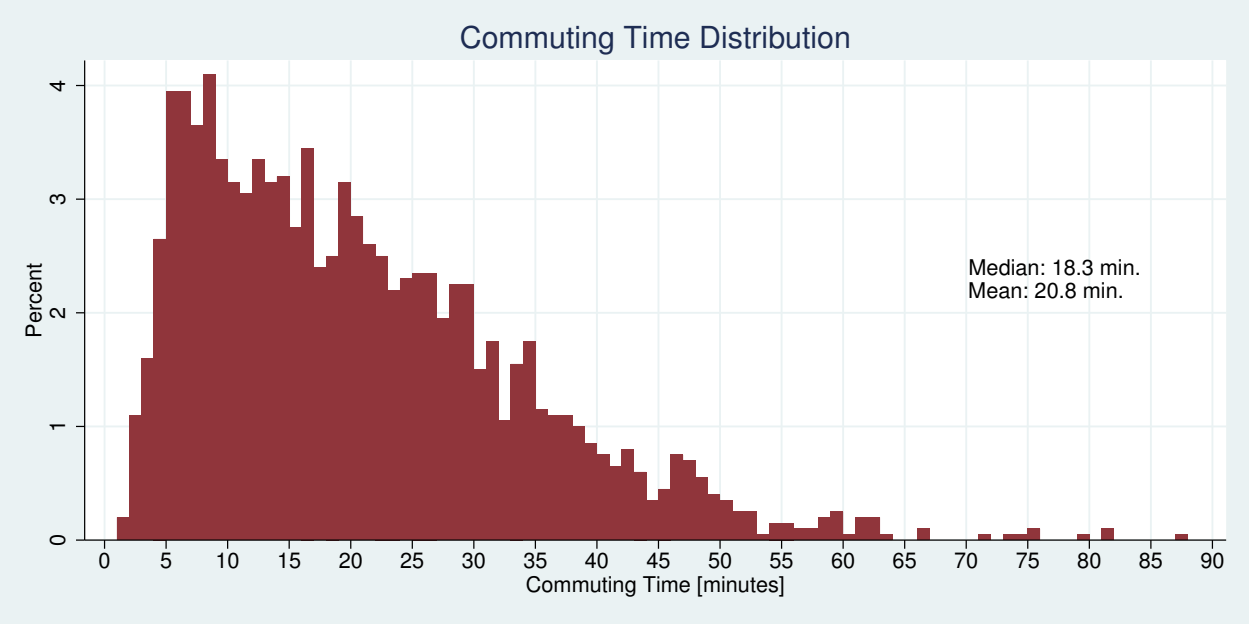
Note: census tract pairs are sorted in descending order of traffic volume. Source: Statistics Canada, National Household Survey (2016).

Figure 3.7: Commuting Distance



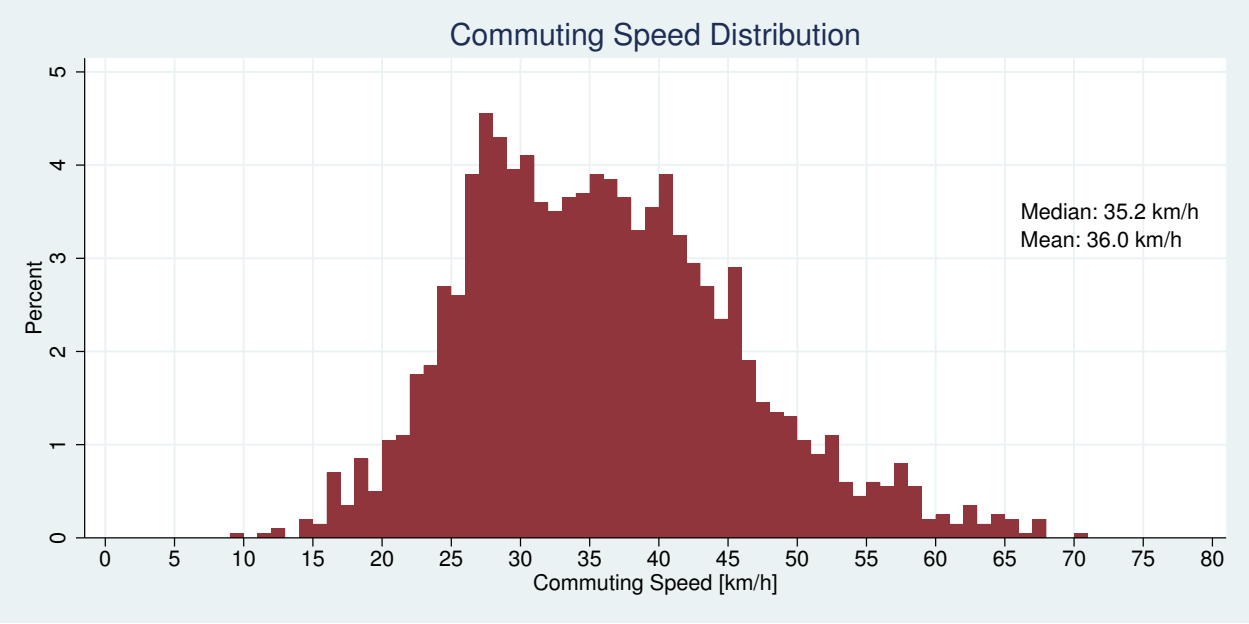
Note: representative sample of 2000 trips across census tract pairs.

Figure 3.8: Commuting Time Distribution



Note: representative sample of 2000 trips across census tract pairs.

Figure 3.9: Commuting Speed Distribution



Note: representative sample of 2000 trips across census tract pairs.

Table 3.2: Determinants of Commuting Speed and Delay

		(A)	(B)
Dependent Variable	Speed	Delay	
Intercept	2.951^c (246.)	-0.042^c (8.85)	
Distance	$\ln(d)$	0.242^c (50.3)	0.045^c (20.7)
Time of Day (linear)	$h-8$	-0.053^c (46.1)	0.056^c (53.0)
Time of Day (squared)	$(h-8)^2$	0.055^c (75.6)	-0.054^c (76.4)
Implied Peak Time	[hour]	8.478^c (828.)	8.524^c (844.)
Monday	0.021^c (38.4)	-0.020^c (40.0)	
Tuesday	-0.013^c (27.0)	0.008^c (22.2)	
Thursday	-0.018^c (30.4)	0.012^c (26.1)	
Friday	0.026^c (41.5)	-0.023^c (42.0)	
Regression fit (R^2)	0.559	0.290	
Observations	685,993	685,993	

Note: Statistical significance at the 95%, 99%, and 99.9% level of confidence is indicated by superscripts a, b and c, respectively. Standard errors are clustered for the 2,000 commutes. Numbers in brackets are t-statistics. The dependent variables are the log of observed speed and the log of the ratio of travel time in traffic to normal travel time. The implied peak is the hour of minimum speed or maximum delay. Wednesday is the base day for the regressions and thus is omitted. Weekends are excluded.

to determine which distribution describes the data more accurately. The result from this analysis allows us to decide whether $H(\psi)$ is more suitable based on either of the two distributions.

3.4.2 Bridge Data: Volume and Time

The commuting pattern data faces one important limitation: we do not observe traffic volumes directly. The BC Ministry of Transportation collects data on traffic volumes at a number of measurement stations across the region, including key bottlenecks (Opus International Consultants, 2014). Table 3.3 identifies the locations we study. We obtain hourly data for one full year from October 2017 through September 2018.[12]

Two of our routes have alternating counterflow lanes. The Lions Gate Bridge has three lanes in total and switches back and forth between 2+1 and 1+2 lanes in each direction. The George Massey Tunnel has four lanes in total and switches between 1+3, 2+2, and 3+1 configurations. Our study of these routes tries to control carefully for the times of day in which they are operated in a particular configuration as they follow a set schedule during the day.

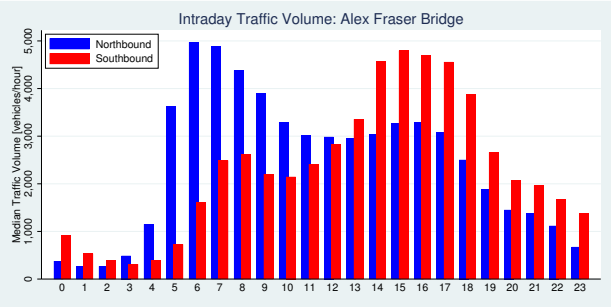
We have also obtained speed information through the Google Distance Matrix API in real time for road segments of a few kilometers in length. We have sampled speeds in 5-minute intervals between 06:00 and 21:55, and at 15-minute intervals during the night. We have averaged the data at the hourly level to match our bridge volume data.

Figure 3.10 shows the diurnal traffic volume for four of our bridges. The northbound and southbound traffic can show significant asymmetries. While the traffic on the Second Narrows Bridge (connecting

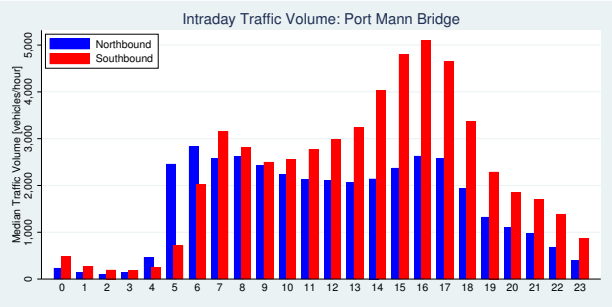
[12]The publication of volume data for some bridges is lagging or has suffered from technical outages. We are missing data for the Second Narrows Bridge in November 2017 and parts of July 2018. Our data for the Alex Fraser Bridge is missing from March-September 2018.

Figure 3.10: Diurnal Direction Traffic Volume

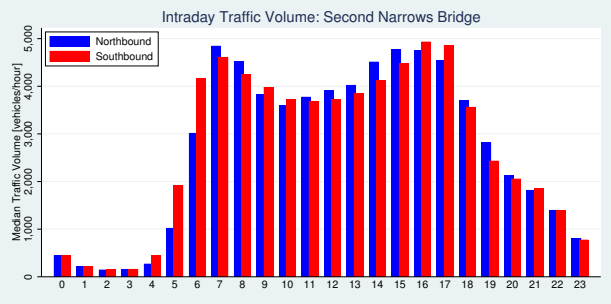
PANEL A



PANEL B



PANEL C



PANEL D

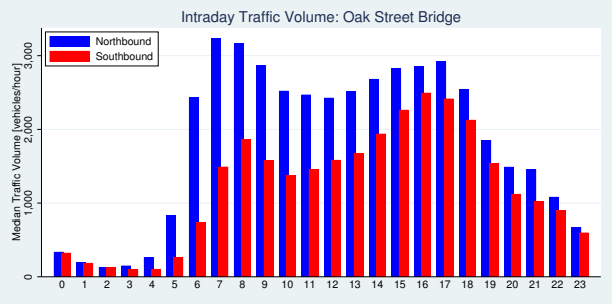
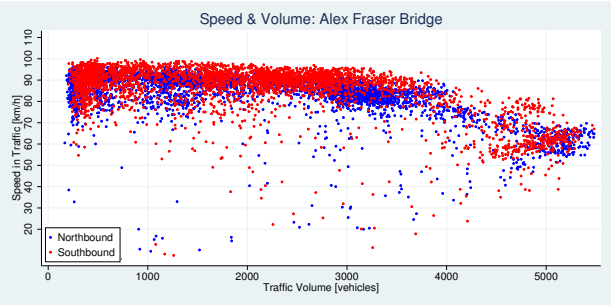
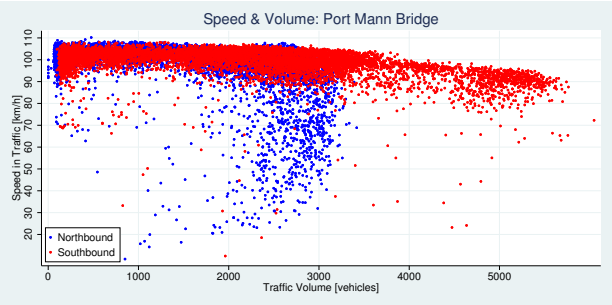


Figure 3.11: Speed Volume Charts

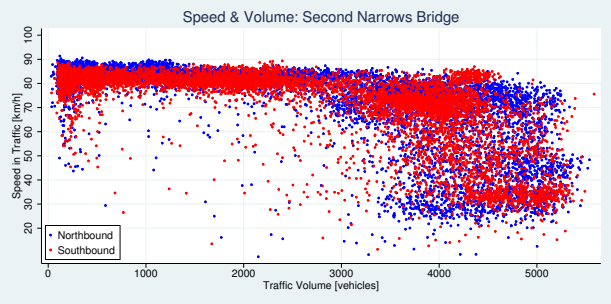
PANEL A



PANEL B



PANEL C



PANEL D

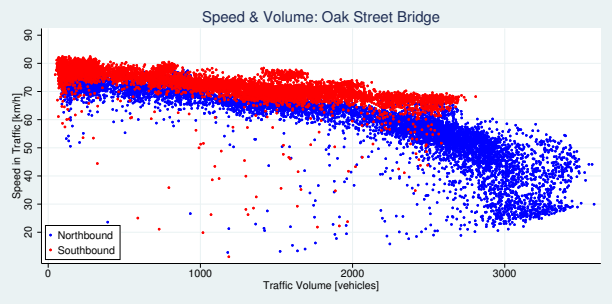


Table 3.3: Traffic Bottlenecks in Metro Vancouver

Code	Name	Direction	Lanes*	Limit
PMB-N	Port Mann Bridge	Northbound (west)	2+3	90 km/h
PMB-S	Port Mann Bridge	Southbound (east)	2+3	90 km/h
GMT-N	George Massey Tunnel	Northbound	1/3	80 km/h
GMT-S	George Massey Tunnel	Southbound	1/3	80 km/h
AFB-N	Alex Fraser Bridge	Northbound	3	90 km/h
AFB-S	Alex Fraser Bridge	Southbound	3	90 km/h
LGB-N	Lions Gate Bridge	Northbound	1/2	60 km/h
LGB-S	Lions Gate Bridge	Southbound	1/2	60 km/h
SNB-N	Second Narrows Bridge	Northbound	2	70 km/h
SNB-S	Second Narrows Bridge	Southbound	2	70 km/h

* a/b indicates counterflow lanes minimum/maximum; a+b indicates separated lanes

Vancouver to North Vancouver) is quite balanced in both directions, the Alex Fraser Bridge has strong northbound traffic in the morning and southbound traffic in the afternoon. The Oak Street Bridge connecting Richmond to Vancouver has stronger northbound than southbound traffic in the morning.

Important for our research is the relation between traffic volume and commuting time. In figure 3.11 we show scatter plots of hourly volume and speed for the same routes as in figure 3.10.[13] They illustrate the typical speed-volume link, although there is no clear *prima facie* evidence of hypercongestion. The new Port Mann Bridge show few signs of significant congestion. Given its significant capacity, only traffic accidents seem to slow down traffic noticeably.

3.5 Empirics

3.5.1 Travel Time Coefficient of Variation

Figures 3.12 and 3.13 provide histograms of the distribution of the two coefficients of variation $T_\rho = T_\sigma/(T_\mu - T^\circ)$ and $T_\nu = T_\sigma/T_\mu$. The first coefficient of variation is measured with respect to the average *delay* $T_\mu - T^\circ$ and the second coefficient is measured with respect to the average time T_μ. We construct the "best" travel time T° for commuting trip i from observations across different times t, and then we allow for a 10% faster drive under optimal conditions in order to allow for individual variation in driving behaviour.

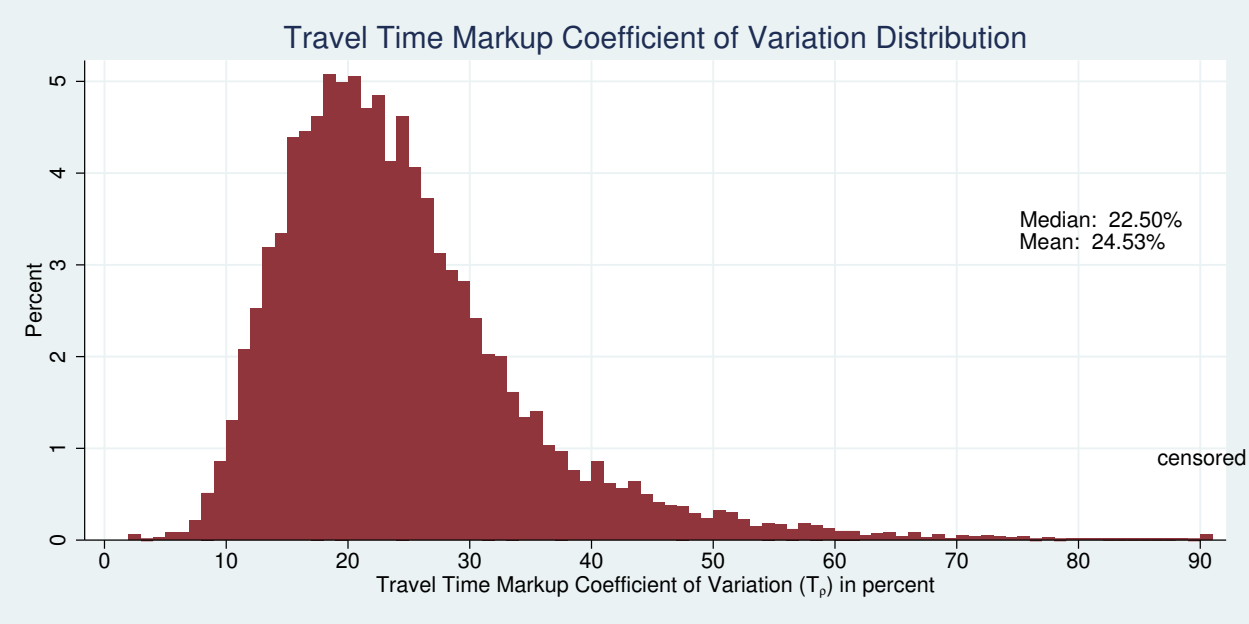
$$T_i^\circ = 0.9 \min_t \{T_{it}\} \tag{3.37}$$

This procedure ensures that T_i° is truly providing a floor for the travel time.

Empirically, we find that both measures are strongly positively correlated (about 0.88), as one would expect when both measures share the same numerator. Both distributions have a long right tail. The travel time markup coefficient of variation T_ρ has a mean of 24%, whereas the commuting time coefficient of variation has a mean of about 6.1%. Recall that each data point was sampled to represent roughly equal

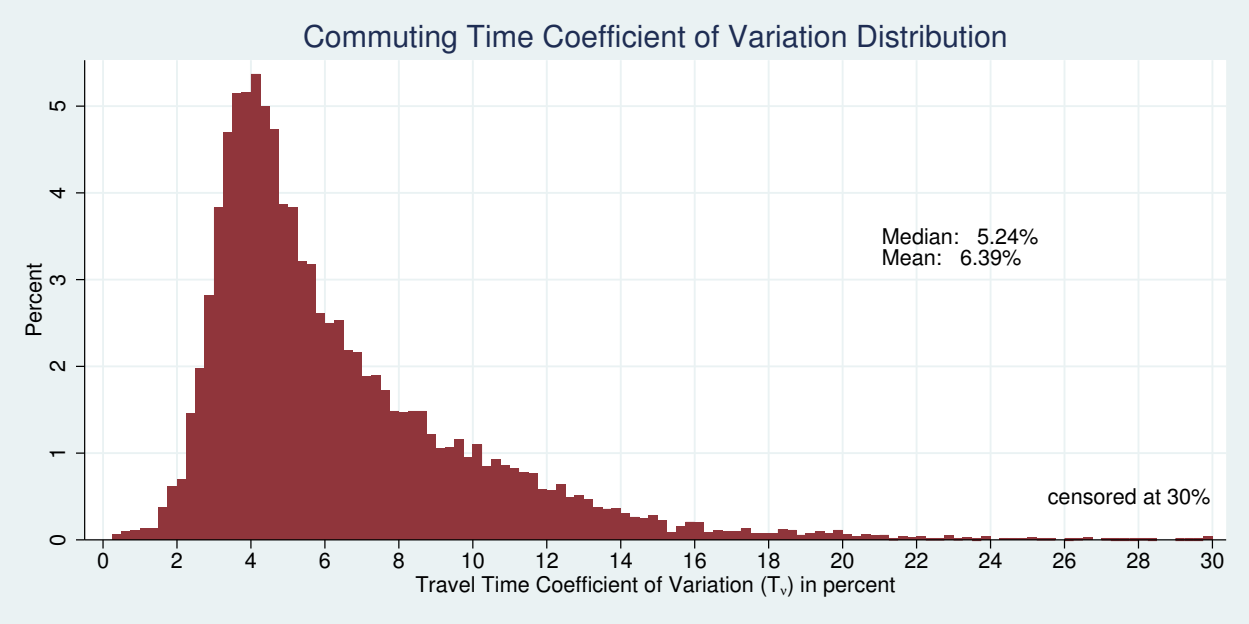
[13] We show speed rather than time because of convention in the transportation literature.

Figure 3.12: Coefficient of Variation Distribution: Time Markup T_ρ



Note: representative sample of 2000 trips across census tract pairs.

Figure 3.13: Coefficient of Variation Distribution: Travel Time T_v



Note: representative sample of 2000 trips across census tract pairs.

number of commuters.

We observe seven half-hour time blocks of morning commutes in Metro Vancouver, but we have only time but no volume data. Therefore we construct a road congestion index (RCI) based on the average delay that commuters incur:

$$\text{RCI}_{kt} = 100\% \cdot \left(\frac{\bar{T}_{kt}}{T_k^\circ} - 1 \right) \tag{3.38}$$

This proxy for actual traffic volume is zero when average travel time does not exceed the minimum (best) travel time. It has a mean of 21.8% and median of 17.8%, and the 5th and 95th percentiles are 3% and 53%, respectively.

Table 3.4: Congestion and Travel Time Coefficients of Variation

	(A)	(B)	(C)
Dependent Variable	$100\%T_v$	$100\%T_p$	$100\%T_p$
Road Congestion Index	0.093^c (35.8)	-0.148^c (17.9)	
Timeslot 07:00			-0.861^c (5.08)
Timeslot 07:30			-3.659^c (19.3)
Timeslot 08:00			-3.026^c (13.2)
Timeslot 08:30			-1.579^c (6.14)
Timeslot 09:00			0.754^b (2.58)
Timeslot 09:30			1.029^b (3.09)
Regression fit (R^2)	0.266	0.065	0.065
Observations	14,000	14,000	14,000
CT Pair Fixed Effects	yes	yes	yes

Note: Statistical significance at the 95%, 99%, and 99.9% level of confidence is indicated by superscripts *a*, *b* and *c*, respectively. Standard errors are clustered for the 2,000 commutes. Numbers in brackets are t-statistics.

We are interested in identifying the extent to which the coefficients of variation increase or decrease with the level of congestion. In Table 3.4 we report three regression specifications that link our observed coefficients of variation to the RCI measure defined in (3.38). Column (A) finds that the ordinary coefficient of variation T_v increases slightly with our road congestion index measure defined in (3.38). The empirical fit (the R^2 statistic) is somewhat weak at 0.266, but the intuition that variance increases with congestion more than the average is intuitive. The estimated coefficient is relatively small, moving the RCI by 10%-points shifts T_v by less than 1%-point.

The effect of congestion on T_p is somewhat weaker and actually negative, reported in column (B). We conclude from this that there is no case to be made that T_p increases with congestion. The negative point estimate rules that out convincingly. The estimation in column (B) is slightly sensitive to the choice of T° in (3.37).[14] The larger magnitude of the estimate is not surprising given that the mean T_p is several times larger than T_v. However, the fit in column (B) is much weaker than in column (A), and we conclude that it is not unreasonable to assume that T_p is not strongly influenced by congestion.

In column (C) we also use dummy variables to proxy for the effect of road congestion during different

[14] It is possible to find a T° that would bring the estimate close to zero.

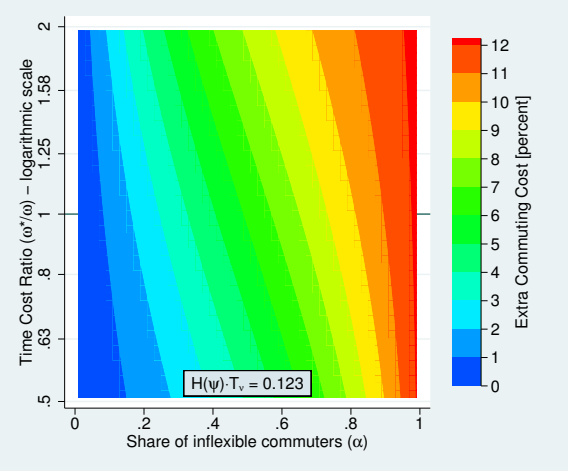
time slots, with 06:30 as the baseline time slot. T_ρ, with an average of about 21%, is slightly lower during peak congestion. The overall explanatory power of road congestion with respect to T_ρ is not very high. There is a weak effect of congestion on T_ρ, but the magnitude is not large enough to negate the theoretical assumption to hold T_ρ constant in our theoretical modelling.

Table 3.5: Schedule Time Cost [%]

	Q1	Median	Q3
T_v	3.89	5.24	7.93
$\psi = 16$	6.82	9.19	13.90
$\psi = 40$	9.55	12.86	19.46
$\psi = 100$	12.26	16.51	24.97
$\psi = 250$	15.05	20.27	30.66

Our analysis is also able to provide an evaluation of the overall cost of scheduling time as we are able to estimate T_v for the Metro Vancouver region and match it with assumptions about schedule inflexibility of commuters, the share of inflexible commuters, and the time cost ratio. Table 3.5 reports the median, as well as first and third quartile (Q1 and Q3) of our empirical findings of T_v. We combine this information with four assumptions of ψ, spaced apart at a factor of 2.5 from each other. A schedule inflexibility of $\psi = 16$ is rather low and can be considered the lowest plausible point. At $\psi = 40$, the commuter is aiming to be late no more than once every 8 weeks. At $\psi = 100$ and $\psi = 250$, the commuter does not want to be late late more than once every 20 weeks and once a year, respectively. Our cautious "best estimate" is the median of T_v at $\psi = 40$, which suggests a schedule time cost of 12.3%. Individual commuters may experience more or less schedule time. A very inflexible commuter ($\psi = 250$) at the third quartile point may incur time costs close to 30%.

Figure 3.14: Schedule Time Variation



The schedule time factor T_Λ that we calculate in table 3.5 is the maximum attainable individually, but the overall time cost increase across all commuters—flexible and inflexible—depends also on the share of inflexible commuters α and the time cost ratio φ of inflexible to flexible commuters. The contour plot in

figure 3.14 illustrates how our key estimate of schedule time (for $\psi=40$ and median T_v) varies with these two parameters, based on our Schedule Time Theorem expressed in equation (3.36). The effect of differences in φ are relatively modest as we explore a range from 1:2 to 2:1 in our diagram. A higher φ leads to a slightly higher extra commuting cost. The stronger effect comes from the composition of commuters. When $\varphi=1$, then $\mathscr{A}=\alpha T_\Lambda$; the time cost increase is directly proportional to the share α of inflexible commuters.

3.5.2 Fréchet v. Lognormal: which provides a better fit?

Our discussion of the reliability standard hinges empirically on the choice of distribution for commuting time. Our analysis employs both Fréchet and log-normal distributions as these are obvious choices for non-negative distributions. However, there are important differences in the shape of these distributions: for identical moments of mean and standard deviation, the Fréchet distribution has a more compressed and focused shape than the log-normal distribution. (Figure B.1 in the appendix illustrates it.) The empirical consequence of the choice of distribution function is reflected in our linear approximation $H(\psi)$ of schedule inflexibility as illustrated in figure 3.2. The Fréchet distribution has a higher slope of $H(\psi)$, and has higher values of $H(\psi)$ when schedule inflexibility is high (roughly when $\psi>40$). We are therefore interested in providing empirical evidence about the relative merits of the distributional choice. Our commuting pattern data enables us to do so as we observe distributions of commuting time across census tract pairs and time slot.

Because our time dimension is limited (we only have 50 days), we exploit the fact that T_ρ is not dependent on the level of road congestion, and pool our data across the 7 different time slots. This means that for every of our 2,000 census tract pairs (except the 50 that show no commuting time variation) we have $7\times50=350$ observations at our disposal. Specifically, we estimate the following two log-likelihood functions for the two distributions

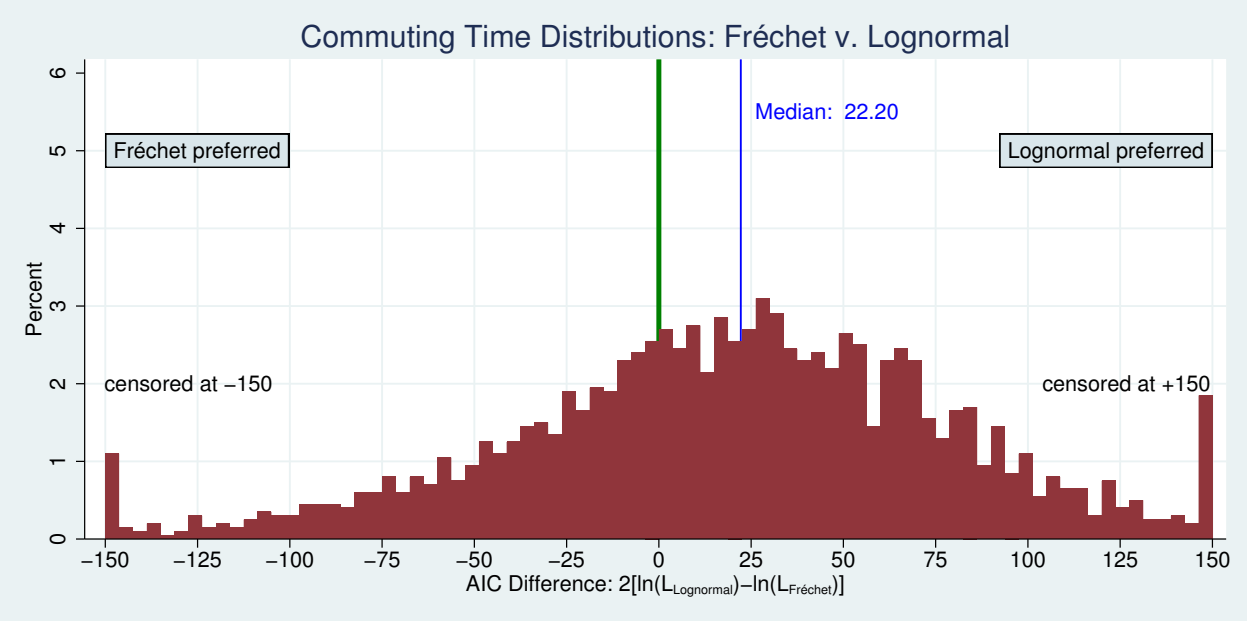
$$\mathrm{LL}_{\mathrm{Fr}}=-\sum_{it}\left\{\ln\left(a\sum_k s_k\delta_{kt}\right)+\left[\frac{\sum_k s_k\delta_{kt}}{T_{it}-T_i^\circ}\right]^{1/a}-\left(1+\frac{1}{a}\right)\ln\left(\frac{\sum_k s_k\delta_{kt}}{T_{it}-T_i^\circ}\right)\right\} \tag{3.39}$$

and

$$\mathrm{LL}_{\mathrm{Ln}}=-\sum_{it}\left\{\ln\left(\sigma(T_{it}-T_i^\circ)\sqrt{2\pi}\right)+\frac{\left[\ln(T_{it}-T_i^\circ)-\sum_k\mu_k\delta_{kt}\right]^2}{2\sigma^2}\right\} \tag{3.40}$$

where $\delta_{kt}\equiv 1(k=t)$ is an indicator variable that is one when time period t matches the variable number k. There are $K+1$ parameters: one for each time slot and a "shape" parameter for the distribution. Our identification strategy, which is based on T_ρ being independent of road congestion, allows us to estimate a single shape parameter for each CT pair while estimating one scale or locational parameter for each of the 7 time slots. Recall that in the Fréchet distribution the shape parameter a is a function of T_ρ; see equation (3.4).

Figure 3.15: Log-Likelihood Comparison of Commuting Time Distribution



Note: representative sample of 2000 trips across census tract pairs.

It is not possible to directly test one distribution against the other using a test statistic. We employ the Akaike information criterion $\text{AIC} = 2K - 2\ln L$, a measure of the relative quality of statistical models for a given set of data. When comparing two models, the lower AIC indicates the better fit. Because both models have the same number of parameters K, the measure we use to characterize the relative performance of the two distributions is the metric

$$\Delta\text{AIC} \equiv (-2\text{LL}_{\text{Fr}}) - (-2\text{LL}_{\text{Ln}}) \tag{3.41}$$

Fréchet is preferred over lognormal if ΔAIC is negative, and lognormal is preferred over Fréchet if ΔAIC is positive. We report the results of our analysis from running 2,000 pairs of regressions in figure 3.15, a histogram of each of the ΔAIC we compute from our regression pairs. As is visible, the bulk of the estimation results favour the lognormal distribution over the Fréchet distribution. We conclude, therefore, that generally the lognormal distribution provides a better fit for commuting time distributions than the Fréchet distribution.

The important implication of our analysis is therefore that our $H(\psi)$ linear approximation of commuting time mark-up as a function of schedule inflexibility (ψ) should be based on the lognormal distribution, and thus the lower $H(\psi)$ that are implied by this choice for larger ψ.

3.5.3 Route Choice: is there a speed-reliability trade-off?

We sample 2,000 commutes seven times a day across Metro Vancouver. How often do commuters take the same route, and how many times will they find a better alternative on a different route? Our data set is able to answer these questions. Table 3.6 shows the distribution of the number of routes commuters may take. Even though the Google API will not provide the exact route, it provides the distance of the route recommended.

We granularize the data by considering a route as different if the integer number of kilometers is different. So routes that are 3.3km and 4.2km would be considered different, while routes that are 3.8km and 4.1km would be considered roughly the same. This granularization is meant to suppress minute differences in routes and focus instead on major route choice differences.

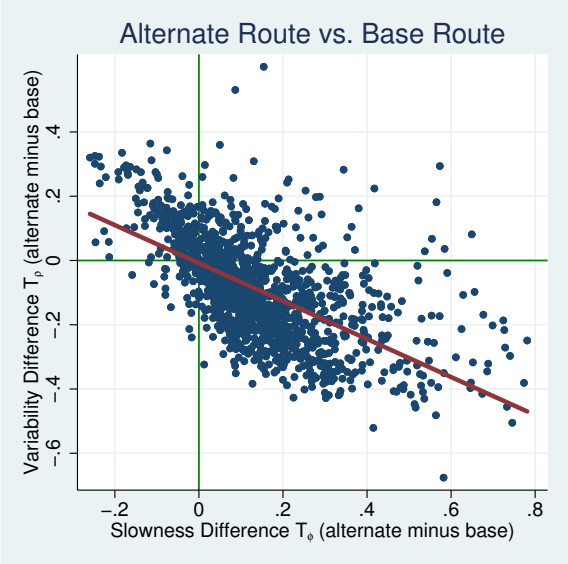
Table 3.6: Distribution of Route Choices among 2,000 Commutes

Routes	1	2	3	4	5	6	7	8	9	10+
Frequency	27.9%	27.3%	15.7%	9.7%	6.2%	4.5%	2.8%	1.8%	1.7%	2.6%

Roughly one third of commuters will only take one route, while about 60% take no more than two routes. As the data in the table shows, there are some commutes that are much more variable, and some commutes even observe more than ten routes. On average, commuters take 2.7 different routes. As commutes get longer, obviously there are also more permutations of route segments. The table above does not reveal the intensity of alternative route choices: when commuters take alternative routes, they may do so more sparingly. We see that 70% of commutes are on the first route choice, and another 17% on the second route choice, and another 6% on the third route choice. About 93% of commuters thus occur on the top three route choices.

We also regress the number of routes on the log of distance. We employ a count regression using the Poisson distribution and find a positive coefficient (0.64) that is highly statistically significant. It confirms that route choices become more complex with greater distance.

Figure 3.16: Alternate Route Trade-Offs



The focus of our paper is on commuter time reliability, and our line of inquiry thus takes us to the question how a commuter's usual route (defined as the route most often taken on a given commute) compares with that commuter's alternate routes (defined as all the routes taken less frequently). Our theoretical work suggests that the coefficient of variation T_ρ is the right measure to compare route performance in terms of

variability, and thus we construct the difference $\Delta T_\rho \equiv T_\rho^A - T_\rho^B$ as the metric that compares the alternate (A) routes against the base (B) route. We construct a similar measure for the slowness of each of our commuting routes: $T_\phi \equiv (\bar{T} - T^\circ)/T^\circ$. This measures captures the markup of average travel time over fastest travel time. To compare alternate (A) routes against base (B) routes, we construct $\Delta T_\phi \equiv T_\phi^A - T_\phi^B$ as our measure of difference in slowness. Some of our 2,000 commutes never use alternate routes or have no time variation, which leaves us with 1,215 observations. In our diagram 3.16 we ignore eight outliers with unusually high or low ΔT_ϕ and show a simple linear fit. The green lines divide the diagram into four quadrants.

On average, it is unsurprising that alternate routes are slower than preferred routes. What is less obvious is how the route choice is related to the time variability of the route, our T_ρ measure. The correlation in figure 3.16 is striking: slower routes are associated with a reduction in variability. This means, as a stylized fact, that *alternate routes are slower but less risky*. The mean ΔT_ϕ is about 0.12 (an increase from 0.37 for base route to 0.49 for alternate route) and the mean ΔT_ρ is about –0.11 (a drop from 0.40 for base route to 0.29 for alternate route). The conclusion from this exercise is that route choice is not just about speed, but also about reliability. If it wasn't, there should be no correlation in figure 3.16.

3.5.4 Bridge Traffic Analysis

In our citywide analysis we only have a rough proxy for congestion as we do not have flow data for all segments of our sampled commutes. However, the BC Ministry of Transportation publishes hourly volume data for a number of select bridges and tunnels that we have described in our data preview section. This data enables us to link volume and speed data to answer the question how increasing congestion affects the dispersion of commuting time. While this analysis allows us to observe the dispersion characteristics at these bottlenecks, these findings are not representative of the entire commute. Bottlenecks are obviously highly important for the overall commuting time and its dispersion, but queueing in front a bottleneck and other spillovers are influencing the effect on the total commuting time. Commuters' route choice and travel decision depends on the total commuting time, including buffer time, but not on the travel time through a bottleneck alone. Our empirical analysis reveals the magnitude of two key parameters: how congestion increases travel time through the bottleneck, and how congestion increases (or decreases) the variance of travel time. It is the latter that is a novel contribution of our paper.

In estimating the effect of congestion on commuting time and its dispersion, we employ a log-linear version of (3.17). However, this approach requires us to introduce the minimum travel time T° and road capacity c as constants. The minimum travel, or analogously maximum speed V°, constrains the empirical distribution. We define the maximum speed as

$$V^\circ \equiv \lfloor \max_t \{V_t\}/0.9 \rfloor \tag{3.42}$$

which mirrors (3.37). We report this speed as an integer number and consider this 10% increase over the highest observed speed as a plausible limit on what is still feasible under completely ideal conditions. Naturally, this maximum speed is much higher than the average speed and the speed limit ($\bar{V}$).[15]

[15]Nevertheless, there are always drivers who may exceed even what we consider a "generous" upper limit. In May 2018 it was

We also need to define the capacity of each of our bridges and tunnels. There is an extensive literature on rating road capacity. We rely on Hall (1975) and Federal Highway Administration (2016, 2017) to determine the capacity of our multilane highways as

$$c = \max\{2200, 1000 + 12.427 \cdot \text{FFS}\} \cdot \frac{\text{Lanes}}{1 + \text{HV}/200} \tag{3.43}$$

where HV is the percentage of heavy traffic on the route, and with the free-flow speed (FFS) given by

$$\text{FFS} = 121.32 - f_{\text{LW}} - f_{\text{LC}} \quad [\text{km/h}] \tag{3.44}$$

Here, f_{LW} and f_{LC} are adjustment factors for lane width and right shoulder later clearance, respectively. The adjustment factor for lane width are 3.6m (0.00 km/h); 3.3m (3.06 km/h); 3.0m or less (10.62 km/h). The adjustment factor for right shoulder lateral clearance is zero for grater than 1.8m, and a maximum of 5.8 km/h and 3.9 km/h for 2-lane and 3-lane roads, respectively, when no right shoulder is present. We have rather incomplete information on the adjustment factors that are needed for our bridges and tunnels because of widely different design characteristics. However, the posted speed limit $\bar{V}$ reflects many of these characteristics such as narrow lane width, absent shoulders, or concrete barriers on the side. We fall back on a classic calculation by Johannesson (1931)

$$c = \left\lfloor \frac{2400}{1 + 18.3/\bar{V}} \right\rfloor \cdot \text{Lanes} \tag{3.45}$$

that provides a remarkably sturdy approximation of lane capacity. We show the results of this formula in table 3.7.

Table 3.7: Speed Limit-Capacity Relationship

Speed Limit [km/h]	50	60	70	80	90	100	110	120
Lane Capacity [vehicles]	1756	1839	1902	1953	1994	2028	2057	2082

We use multiplicative heteroskedastic linear regression (MHLR) to analyze our bridge traffic data and determine how dispersion *and* travel time vary with the level of congestion. We estimate a linear model of the form $y_i = \boldsymbol{x}_i\boldsymbol{\beta} + \varepsilon_i$ with dependent variable y_i and a vector of independent variables $\boldsymbol{x}_i$ and i.i.d. errors with mean zero and variance σ_i^2. The variance function $\sigma_i^2 = \exp(\boldsymbol{z}_i\boldsymbol{\gamma})$ is parametric and is explained by variable vector $\boldsymbol{z}_i$. Equivalently, $\ln(\sigma_i^2) = \boldsymbol{z}_i\boldsymbol{\gamma}$ is linear. Thus we can estimate the MHLR log-likelihood function

$$\text{LL} = -\frac{1}{2}\left[n\ln(2\pi) + \sum_{i=1}^{n} \boldsymbol{z}_i\boldsymbol{\gamma} - \frac{(y_i - \boldsymbol{x}_i\boldsymbol{\beta})^2}{\boldsymbol{z}_i\boldsymbol{\gamma}}\right] \tag{3.46}$$

Stata, the statistical software that we use, has implemented this type of regression through the command `hetregress` for the linear case. It is straight-forward to generalize this to the non-linear case.

reported that a Ferrari driver was fined $750 for clocking 210 km/h on Lions Gate Bridge. The posted speed limit is 60 km/h. There is no limit to recklessness.

We start out by noting that $\ln(T_i - T^\circ)$ is normally distributed, and because T° is a constant, $y_i \equiv \ln(T_i/T^\circ - 1)$ is also normally distributed. Empirically we normalize distance for our bridge and tunnel analysis and thus we observe velocities V_i and infer V°. Therefore, for any distance, it holds that $T_i/T^\circ = V^\circ/V_i$. Thus we calculate $y_i = \ln(V^\circ/V_i - 1)$. The mean is expected to follow a congestion function of the types proposed in (3.15) and (3.16),[16] and implemented in (3.17) as a stochastic version that fits with our log-normality assumption for $T_i - T^\circ$. The error term ε_i in (3.17) is distributed normally with mean zero and variance σ^2, which we allow to grow analogously to the mean. Thus we can estimate the linear system

$$y_i = \ln(T^\odot/T^\circ - 1) - \xi \ln(1 - v_i/c) + \varepsilon_i \tag{3.47}$$

$$\ln(\sigma_i^2) = \ln(\Sigma^\odot) - \zeta \ln(1 - v_i/c) \tag{3.48}$$

which provides estimates of $T^\odot$ and ξ in the mean equation, and $\Sigma^\odot$ and ζ in the variance equation. Hence, we calculate the observational mean as

$$T_\mu(\tilde{v}) = T^\circ + \left[\frac{T^\odot - T^\circ}{(1-\tilde{v})^\xi}\right] \cdot \exp\left(\frac{\Sigma^\odot/2}{(1-\tilde{v})^\zeta}\right) \tag{3.49}$$

and the standard deviation as

$$T_\sigma(\tilde{v}) = (T_\mu(\tilde{v}) - T^\circ)\sqrt{\exp\left(\frac{\Sigma^\odot}{(1-\tilde{v})^\zeta}\right) - 1} \tag{3.50}$$

and therefore

$$T_\rho = \frac{T_\sigma}{T_\mu - T^\circ} = \sqrt{\exp\left(\frac{\Sigma^\odot}{(1-\tilde{v})^\zeta}\right) - 1} \tag{3.51}$$

Note that when $\zeta = 0$, the expression on the right turns into a constant $\sqrt{\exp(\Sigma^\odot) - 1}$ and thus becomes independent of congestion. This means that T_σ increases proportional to $T_\mu - T^\circ$ as congestion increases. In the $\zeta = 0$ case,

$$T_\mu(\tilde{v}) = T^\circ + \frac{T^\odot - T^\circ}{(1-\tilde{v})^\xi}\sqrt{1 + T_\rho^2} \tag{3.52}$$

When we interpret the results, it is useful to express $T^\odot$ as a speed equivalent, knowing that $T^\odot/T^\circ = V^\circ/V^\odot$. Thus, using the estimate of the intercept in the mean equation $\hat{a} = \ln(T^\odot/T^\circ - 1)$, it follows that $V^\odot = V^\circ/(1 + \exp(\hat{a}))$. Similarly, when we interpret the constant $\hat{a} = \ln(\Sigma^\odot)$ in the variance equation, we report the corresponding initial $T_\rho^\odot$ at $\tilde{v} = 0$ as $\sqrt{\exp(\exp(-\hat{a})) - 1}$.

If ζ is positive, then the coefficient of variation will decrease with congestion. However, if ζ is negative, then the coefficient of variation will in fact increase with congestion. Both are possible outcomes. Both ξ and ζ determine the travel time in (3.49). When ζ is positive, it reinforces the effect of ξ; however, when ζ is negative, it dampens the effect of ξ. The intuition is that a positive ζ increases the variance σ^2 of the log-normal distribution when congestion increases, while a negative ζ decreases that variance. Because of

[16] There is no log-linearization of the Pipes-Munjal congestion function because $\ln(1 + \kappa(1-\tilde{v})^\lambda)$ does not allow separating κ and λ. It is therefore necessary to estimate this version non-linearly. We leave this exercise for future study as the small improvements in fit do not change our qualitative results in this section.

the right-skewed shape of the log-normal distribution, increasing variance pulls up the arithmetic mean.

Table 3.8 reports the results of our estimation for several of our bridges, opposing traffic directions, and different lane numbers due to counterflow traffic: the Alex Fraser Bridge northbound and southbound (AFB-N and AFB-S); the Oak Street Bridge (OSB-N and OSB-S); the Second Narrows Bridge (SNB-N and SNB-S); the Lions Gate Bridge (LGB-N and LGB-S) in two-lane mode; and the George Massey Tunnel (GMT-N and GMT-S) in 2-lane normal mode and 3-lane counterflow mode. We are not presenting results for the Massey Tunnel or the Lions Gate Bridge when only a single lane is in operation because these phases are heavily driven by lane merging and queuing. Similarly, we do not show results for the Port Mann Bridge because the bridge has distinct express and local lanes, while the volume data is aggregate. For this bridge we are not able to discern separate volumes for the long-distance and local traffic, and the northbound direction also has a complicated lane merger that results in capacity curtailment.

Table 3.8 first shows the relevant fixed parameters for our bridges and tunnels: the number of lanes, posted speed limit, maximum speed inferred by (3.42), and road capacity inferred by (3.45). There are two blocks of regression estimates in each column: the first for the mean equation (3.47), and the second for the variance equation (3.48). The congestion effect estimates identify ξ and ζ, respectively. The two constants are difficult to interpret without transformation. Therefore, we report the initial speed $V^{\odot}$ for the mean equation and the initial T_{ρ} for the variance equation. "Initial" means that these are the speeds (or coefficients of variation) with zero congestion. We finally report the number of observations and the Aikaike Information Criterion (AIC) for model selection purposes (noting that smaller AICs are better). Our result for the Lions Gate Bridge northbound has a negative ξ estimate, which we attribute to a rather high maximum speed compared to the initial speed. We discard this set of estimates as unreliable.

Overall, we find strong evidence that the congestion parameter ξ for the mean equation varies between about 0.6 and 1.0. This means that, as can be expected, congestion strongly increases travel times. The effect is somewhat more pronounced in some places than others. The Oak Street Bridge northbound has much stronger congestion behaviour than southbound. This observation is consistent with the fact that the northbound traffic transitions freeway traffic to urban (signal) traffic, while southbound a string of traffic lights regulate the traffic volume onto the bridge.

Our key result concerns the variance equation and fully justifies our choice of using an MHLR model: variance is strongly influenced by congestion. Our estimates of ζ generally tend to be mostly highly positive (in seven of twelve cases), and sometimes negative (in three of twelve cases). In some cases the results are insignificant. The case of the Oak Street Bridge southbound is particularly interesting because it characterizes transition from urban (signal) traffic to freeway traffic, so inflow to the bridge is throttled by the signal phases. This results in the coefficient of variation being essentially constant. Where we see relative unimpeded traffic flow, such as on the Alex Fraser Bridge, the variance also does not seem to increase with congestion. The variances effect kicks in much stronger in the cases where some amount of queuing takes places (northbound traffic on Oak Street, the Massey Tunnel, but also on the Second Narrows Bridge that has strong on-ramp traffic merging onto it.

There are three key takeaways from this section. First, variance matters—a lot. Second, variance appears to increase with congestion when traffic has strong "bottleneck" features, while the coefficient of variation

Table 3.8: Estimation Results Bridge & Tunnel Traffic

Route	AFB-N	AFB-S	OSB-N	OSB-S
Lanes	3	3	2	2
Speed Limit [km/h]	90	90	60	60
Max. Speed [km/h]	108	111	88	91
Road Capacity	5982	5982	3678	3678
Constant (Mean Eq.)	-1.633^{c} (158.)	-1.648^{c} (147.)	-1.650^{c} (361.)	-1.719^{c} (384.)
Congestion Effect	0.669^{c} (57.1)	0.692^{c} (44.7)	1.010^{c} (171.)	0.777^{c} (89.3)
Initial Speed [km/h]	90.355^{c} (593.)	93.080^{c} (553.)	73.820^{c} (1359)	77.166^{c} (1470)
Constant (Var. Eq.)	-1.607^{c} (43.7)	-1.644^{c} (43.0)	-2.933^{c} (124.)	-2.958^{c} (107.)
Congestion Effect	-0.450^{c} (8.96)	-0.078 (1.39)	0.718^{c} (32.4)	-0.095 (1.74)
Initial T_p	0.471^{c} (49.3)	0.462^{c} (47.6)	0.234^{c} (82.3)	0.231^{c} (70.2)
Observations	3,550	3,550	8,635	8,635
AIC	3487.46	4098.71	4256.52	-1382.5

Route	GMT-N	GMT-S	SNB-N	SNB-S
Lanes	2	2	3	3
Speed Limit [km/h]	80	80	70	70
Max. Speed [km/h]	103	102	101	99
Road Capacity	3906	3906	5706	5706
Constant (Mean Eq.)	-1.515^{c} (265.)	-1.707^{c} (320.)	-1.683^{c} (254.)	-1.646^{c} (247.)
Congestion Effect	0.581^{c} (51.0)	0.710^{c} (73.0)	0.841^{c} (76.6)	0.778^{c} (74.1)
Initial Speed [km/h]	84.437^{c} (971.)	86.339^{c} (1221)	85.176^{c} (963.)	82.995^{c} (928.)
Constant (Var. Eq.)	-2.657^{c} (98.8)	-2.823^{c} (103.)	-2.321^{c} (83.1)	-2.381^{c} (85.3)
Congestion Effect	0.918^{c} (24.5)	0.815^{c} (22.5)	1.227^{c} (43.2)	1.090^{c} (37.6)
Initial T_p	0.270^{c} (71.8)	0.247^{c} (70.8)	0.321^{c} (68.2)	0.311^{c} (68.5)
Observations	5,821	5,821	7,560	7,560
AIC	3839.87	2741.97	11339.3	9904.03

Route	GMT-N	GMT-S	LGB-N	LGB-S
Lanes	3	3	2	2
Speed Limit [km/h]	80	80	60	60
Max. Speed [km/h]	102	95	67	50
Road Capacity	5859	5859	3678	3678
Constant (Mean Eq.)	-1.759^{c} (86.0)	-1.731^{c} (44.7)	-0.541^{c} (18.8)	-1.361^{c} (146.)
Congestion Effect	0.896^{c} (48.1)	0.764^{c} (26.8)	-0.352^{c} (14.9)	0.344^{c} (49.6)
Initial Speed [km/h]	87.014^{c} (333.)	80.712^{c} (172.)	42.340^{c} (94.7)	39.798^{c} (526.)
Constant (Var. Eq.)	-2.289^{c} (23.7)	-1.072^{c} (9.06)	-1.478^{c} (16.4)	-3.221^{c} (57.1)
Congestion Effect	0.730^{c} (10.6)	-0.480^{c} (4.95)	-0.386^{c} (5.02)	0.250^{c} (8.35)
Initial T_p	0.327^{c} (19.7)	0.639^{c} (14.3)	0.506^{c} (19.8)	0.202^{c} (34.8)
Observations	1,090	1,456	1,456	1,074
AIC	1606.55	1769.55	1387.28	-78.424

Note: Statistical significance at the 95%, 99%, and 99.9% level of confidence is indicated by superscripts *a*, *b* and *c*, respectively. Standard errors are clustered for the 2,000 commutes. Numbers in brackets are t-statistics.

remains relatively flat when congestion is less driven by merging traffic and queueing effects. Third, when variance and mean both increase with congestion, the variance effect intensifies the mean effect because of the right-skewed shape of the log-normal distribution. This in turn matters for reliability calculations: schedule-sensitive drivers will avoid the roads that are getting excessively unpredictable and in tendency favour alternative routes that trade-off more reliability against slower speed. This is one of the two topics that we explore more deeply in our next section, which simulates route choice.

3.6 Simulation

The purpose of our simulation exercise is to explore competing routes where commuters can choose among two options that differ in terms of volume, speed, and dispersion characteristics. As commuters can choose routes, this problem is inherently more difficult to describe algebraically and instead is best addressed through simulation. We have at the minimum two distributions to draw from randomly, and this makes analytic solutions very difficult if not impossible to obtain. To focus on the essence of our model we simulate a simple two-road system. The framework we use can be used easily for more complex networks, but by focusing on two competing routes we are able to better identify the sorting effect between flexible and inflexible commuters. Our two routes can be envisioned either as two separated lanes of the same highway (the "Lexus Lane" model), but they can also be envisioned as the choice between two transportation modes (public transit and cars). Our analysis is applicable to either type of interpretation.

The interpretation of our analysis as a transportation mode choice requires some modification, however. A car-owning commuter only decides between competing routes. A commuter choosing between competing transportation modes will take the difference in transportation cost into account, which includes the variable cost of an individual trip as well as the levelized fixed cost of the transportation mode. If ownership of a car is treated as a sunk cost, the decision at the margin compares fuel cost against transit fare. In Metro Vancouver, transit fares are divided into three zones ($2.95, $4.20, $5.70, respectively, for a single trip). A typical 13.5km commute by car in Vancouver would cost about $2.00-$2.50, depending on the fuel efficiency of the vehicle and the price of gasoline. In most instances, the car will be the cheaper option at the margin. However, the difference in cost is negligible in comparison to the time cost difference, and we thus focus on this alone in our simulation for route choice. We capture the cost trade-off by focusing on the asymmetry in the route choice—slower but more reliable versus faster but more volatile—and using identical tolls on both routes: identical per-trip prices regardless of commuting mode.

3.6.1 Simulation Configuration

Our simulation explores two scenarios that characterize the effect of reliability on commuting.

- **Two symmetric roads with one road tolled**. This configuration mirrors the 'Lexus lane' approach studied by Hall (2018b). We are exploring how introducing a toll on one of two routes will improve the allocation efficiency by sorting inflexible and flexible commuters.
- **Two asymmetric roads with both roads tolled**. With this configuration we are interested in the trade-off between a slower but more predictable route (or mode) and a faster but more volatile route

(or mode). One of the routes has a significantly higher T_ρ than the other (double). Even without a toll, commuters sort themselves so that inflexible commuters prefer the slower but safer route (or mode) over the alternative. However, how much more sorting is induced by tolling?

The two scenarios allow us to determine how route choice is affected by policy (congestion pricing) and by self-sorting based on route characteristics. The two scenarios are parameterized as shown in table 3.9. We refer to the two competing routes as 'A' and 'B', where route 'A' is the tolled road in the 'Lexus Lane' scenario, and the slow-but-certain route in the 'Asymmetry' scenario.

Table 3.9: Simulation Scenario Parametrization

			'Lexus Lane' (Toll: A)		'Asymmetry' (Toll: Both)	
Parameter	Symbol	Unit	A	B	A	B
Lane Capacity	c	[vehicles/hr]	2200	2200	2200	2200
Minimum Travel Time	T°	[min]	15	15	18	12
Initial Average Travel Time	$T^\odot$	[min]	20	20	23	17
Congestivity	ξ	[–]	0.50	0.50	0.50	0.50
Coefficient of variation	T_p	[–]	0.25	0.25	0.20	0.40

Table 3.10: Simulation Scope Parametrization

Parameter	Symbol	High (H)	Medium (M)	Low (L)
Share of Inflexible Commuters	α	0.8	0.6	0.4
Schedule Inflexibility	ψ	250	100	40
Time Cost Ratio	φ	1.25	1.00	0.80

For the time cost distribution we employ Canadian census data that identifies the individual income in the Metro Vancouver region as CAD 44,700 and CAD 33,600 for mean and median, respectively, in the year 2016. Following Small and Verhoef (2007), we assume that time cost is 50% of income. Based on OECD estimates[17] of 1,706 average annual hours actually worked per worker for Canada in 2016 we determine a mean and median time cost of \$13.10/h and \$9.85, which we transform through the moments of the lognormal distribution into $\mu = \ln(\text{median}) = 2.0894$ and $\sigma = \sqrt{2\ln(\text{mean}/\text{median})} = 0.75567$. In accordance with the theoretical consideration, the cutoff time $T_i^\varnothing$ that determines the cutoff cost through the mean time cost $\bar{\omega}$ so that $z_i = \bar{\omega} T_i^\varnothing$ is drawn from an extreme value type I (Gumbel) distribution. We calibrate this distribution by choosing a mean and median of 50 and 47 minutes, respectively.[18]

In table 3.10 we define the base parameterizations for our simulation. We explore our three key parameters– the share of inflexible commuters, the schedule inflexibility, and the time cost ratio—at three levels that characterize the low end, mid range, and high end of plausible parameterizations.

For the share of inflexible commuters (α) we argue that during the morning rush hour it likely exceeds 50% and think that 60% is a conservative estimate, but allow for a 40% and 80% alternatives. We thus have

[17] OECD.Stat time data table ANHRS.

[18] See details in the mathematical appendix.

a bracketing of 20% between steps. For the degree of inflexibility we assume that a ψ of 100 (being late to work 2-3 times per work year) is a reasonable standard, but also allow for a stricter level (once per year) and a weaker level (once every 8 weeks). There is a ratio of 2.5:1 between each of the two steps for ψ. For the time cost difference between the two commuter groups we start with a neutral assumption of 1:1, and allow for a 25% deviation in either direction.

We thus have $3\times3\times3=27$ cases that we explore. For each of these 27 scenarios we begin by dispatching all commuters to fill up both routes, which provides our baseline against which we compare the results from an optimal toll. We explore two optimal tolls: the full-information optimal toll that takes schedule delay cost of the inflexible commuters into account, and the limited-information (sub-)optimal toll that only takes observable driving time into consideration. For ease of reference, we call them 'optimal' and 'classic' tolls. Where we need to distinguish between the two, we denote these outcomes with an asterisk (*) or a bullet ($^\bullet$) superscript. We describe the algorithm for computing the toll optimum below. Once we have computed the two toll optima, we are interested in a number of different outcomes:

(a) the schedule delay premium $\mathscr{A}$, expressed as a percentage, when everyone commutes (no toll);
(b) the optimal toll τ;
(c) the percentage share π of flexible commuters who decide to commute (with $1-\pi$ the share of non-commuters;
(d) the utilization rates of both routes, $\tilde{v}_A$ and $\tilde{v}_B$, as a percentage;
(e) the percentage share of inflexible commuters who prefer the tolled or slow-but-certain route ϕ_A^I and the percentage share flexible commuters who prefer the untolled or fast-but-volatile route ϕ_B^F;
(f) the cost savings $\Delta\mathscr{C}$ achieved through the toll as a percentage reduction over the no-toll case.

Additionally, we will also explore distributional outcomes that we define further below.

3.6.2 Optimal Toll Algorithm

We have implemented our simulations both in C++ and Python code in order to cross-verify our simulation methodology. This dual implementation strategy has reassured us that our algorithm is reliable. We use the more-efficient C++ code for our high-replication results. Our optimal toll algorithm consists of five steps:

1. assign random time cost ω_i and cutoff cost z_i to each commuter (and in the expanded version, also draw a random schedule inflexibility ψ_i);
2. dispatch all commuters sequentially, letting them choose optimal routes conditional on current levels;
3. reoptimize route choices until the choices of all commuters are stable;
4. find the optimal toll through golden section search (GSS)[19] using a cost function that includes commuting time including schedule delay, but excludes tolls (as they are rebated uniformly per capita). Each search step consists of assigning a new toll level, reoptimizing the route choices, and calculating the cost function; and
5. find the 'classic' toll in the same way using a cost function that only includes travel time.

We are able to speed up the simulation runs by computing in advance the travel times, or the key parts thereof, that correspond to each of the discrete capacity utilization points for each route. This means that

[19] We choose this method because it is robust even when the cost function is not smooth.

we can rely on inexpensive additions and multiplications to compute final travel and commute times, rather than computing power functions and square roots. This simple trick allows us to run between 30 and 70 simulations per second in our C++ code.

3.6.3 Results: Single and Dual Toll Scenarios

The first aim of our simulation exercise is the characterization of optimal tolls in the presence of distinct commuter groups (flexible and inflexible) with different time costs, and where the degree of inflexibility (the reliability standard) of the inflexible commuters may vary. Tables 3.11 and 3.12 present the results of our simulation for the 'Lexus Lane' and 'Asymmetry' scenarios, respectively, where we interpret the second scenario also as as a potential mode choice. Both tables show the results of the average of 1,000 simulation runs.

Table 3.11: Simulation Results: 'Lexus Lanes' Scenario

Parameters			Optimal Toll								Classic Toll		
α	ψ	φ	$\mathscr{A}$	τ	π	$\tilde{v}_A$	$\tilde{v}_B$	ϕ_A^I	ϕ_B^F	$\Delta\mathscr{C}$	τ	π	$\Delta\mathscr{C}$
L	L	L	11.21	1.19	81.3	77.9	90.8	60.0	65.2	−2.69	1.00	81.7	−2.64
L	L	M	12.78	1.32	80.7	76.7	91.3	66.3	71.4	−3.36	1.05	81.3	−3.26
L	L	H	14.37	1.46	79.8	75.2	91.8	72.1	77.6	−4.23	1.12	80.8	−4.06
L	M	L	14.38	1.31	81.0	77.4	91.0	63.0	68.1	−3.02	1.03	81.6	−2.90
L	M	M	16.40	1.45	80.3	76.1	91.5	69.1	74.3	−3.81	1.07	81.2	−3.61
L	M	H	18.46	1.61	79.4	74.5	92.0	74.6	80.4	−4.82	1.15	80.7	−4.52
L	H	L	17.66	1.44	80.7	76.8	91.2	65.6	70.7	−3.40	1.05	81.5	−3.20
L	H	M	20.13	1.59	79.9	75.4	91.7	71.5	76.9	−4.30	1.10	81.1	−3.97
L	H	H	22.64	1.77	78.9	73.7	92.3	76.7	82.9	−5.43	1.18	80.6	−4.96
M	L	L	18.00	1.61	78.8	80.9	92.9	56.1	71.6	−3.71	1.36	79.3	−3.64
M	L	M	19.61	1.84	77.6	79.8	93.2	60.0	80.8	−4.59	1.52	78.5	−4.48
M	L	H	21.11	2.12	76.4	78.5	93.5	63.2	88.8	−5.61	1.72	77.6	−5.47
M	M	L	23.09	1.79	78.2	80.4	93.0	58.0	75.9	−4.18	1.44	79.0	−4.05
M	M	M	25.17	2.06	77.1	79.2	93.3	61.6	84.7	−5.18	1.60	78.3	−4.99
M	M	H	27.08	2.37	75.8	78.0	93.6	64.3	91.9	−6.29	1.83	77.3	−6.06
M	H	L	28.33	2.01	77.7	79.9	93.2	59.6	79.8	−4.68	1.51	78.9	−4.47
M	H	M	30.87	2.31	76.5	78.6	93.5	62.8	88.0	−5.78	1.68	78.1	−5.46
M	H	H	33.26	2.65	75.4	77.5	93.8	65.2	94.4	−6.94	1.94	77.2	−6.60
H	L	L	26.11	2.11	74.6	85.5	94.9	52.3	78.9	−4.67	1.91	75.2	−4.64
H	L	M	27.20	2.45	73.2	84.8	95.0	53.8	89.1	−5.44	2.19	74.0	−5.39
H	L	H	28.15	2.78	72.3	84.4	95.1	54.8	95.7	−6.11	2.53	73.0	−6.07
H	M	L	33.51	2.38	74.0	85.2	94.9	53.1	83.9	−5.17	2.07	74.8	−5.11
H	M	M	34.91	2.76	72.7	84.6	95.1	54.3	92.7	−5.92	2.39	73.7	−5.85
H	M	H	36.12	3.13	71.9	84.2	95.1	55.0	97.6	−6.57	2.78	72.8	−6.51
H	H	L	41.12	2.67	73.4	84.9	95.0	53.7	88.1	−5.65	2.23	74.5	−5.54
H	H	M	42.85	3.05	72.4	84.4	95.1	54.7	95.2	−6.41	2.58	73.5	−6.30
H	H	H	44.34	3.49	71.7	84.1	95.2	55.1	98.6	−6.99	3.04	72.7	−6.91

Note: Initial schedule delay cost markup $\mathscr{A}$ in percent. Optimal toll τ in dollars. Flexible commuters commuting π in percent. Route utilization rates $\tilde{v}$ in percent. Share of commuter types I and F taking routes A and B in percent. Cost savings over baseline $\Delta\mathscr{C}$ in percent. Average of 1,000 simulation runs.

Table 3.12: Simulation Results: 'Asymmetry' Scenario

Parameters			Optimal Toll								Classic Toll		
α	ψ	φ	$\mathscr{A}$	τ	π	$\tilde{v}_A$	$\tilde{v}_B$	ϕ_A^I	ϕ_B^F	$\Delta\mathscr{C}$	τ	π	$\Delta\mathscr{C}$
L	L	L	7.33	3.92	70.1	76.0	79.9	100.	100.	−5.35	3.92	70.1	−5.35
L	L	M	8.35	3.51	70.1	76.0	79.9	100.	100.	−5.16	3.51	70.1	−5.16
L	L	H	9.39	3.14	70.1	76.0	79.9	100.	100.	−4.96	3.14	70.1	−4.96
L	M	L	9.39	3.92	70.1	76.0	79.9	100.	100.	−5.39	3.92	70.1	−5.39
L	M	M	10.70	3.51	70.1	76.0	79.9	100.	100.	−5.18	3.51	70.1	−5.18
L	M	H	12.04	3.14	70.1	76.0	79.9	100.	100.	−5.01	3.14	70.1	−5.01
L	H	L	11.54	3.92	70.1	76.0	79.9	100.	100.	−5.40	3.92	70.1	−5.40
L	H	M	13.15	3.51	70.1	76.0	79.9	100.	100.	−5.22	3.51	70.1	−5.22
L	H	H	14.79	3.14	70.1	76.0	79.9	100.	100.	−5.05	3.14	70.1	−5.05
M	L	L	15.98	5.84	57.1	79.4	78.0	69.7	100.	−9.73	5.23	61.3	−9.57
M	L	M	17.28	5.70	53.3	77.8	76.8	68.2	100.	−11.04	5.05	58.5	−10.83
M	L	H	20.04	5.57	49.1	75.9	75.4	66.6	100.	−12.41	4.85	55.5	−12.13
M	M	L	22.35	6.17	55.3	80.3	75.8	70.4	100.	−10.99	5.32	61.2	−10.69
M	M	M	25.17	6.02	51.2	78.6	74.3	69.0	100.	−12.41	5.09	58.6	−11.99
M	M	H	27.77	5.89	46.6	76.8	72.6	67.4	100.	−13.96	4.92	55.5	−13.45
M	H	L	29.54	6.48	53.4	80.9	73.6	71.0	100.	−12.17	5.33	61.4	−11.62
M	H	M	32.99	6.34	49.0	79.3	71.9	69.6	100.	−13.73	5.13	58.7	−13.05
M	H	H	35.96	6.19	44.2	77.5	70.1	68.0	100.	−15.44	4.95	55.6	−14.64
H	L	L	28.36	9.02	31.8	83.1	81.0	54.6	100.	−13.23	7.37	42.9	−12.76
H	L	M	29.96	9.03	25.4	81.7	79.9	53.8	100.	−14.75	7.26	37.8	−14.23
H	L	H	31.35	9.11	18.9	80.4	78.8	52.9	100.	−16.31	7.19	32.0	−15.78
H	M	L	37.42	9.63	28.4	83.7	79.1	55.1	100.	−14.80	7.47	42.7	−14.04
H	M	M	39.47	9.69	21.9	82.5	77.9	54.2	100.	−16.43	7.35	37.5	−15.63
H	M	H	41.22	9.81	15.5	81.2	76.7	53.4	100.	−18.03	7.28	31.7	−17.22
H	H	L	47.08	10.23	25.2	84.2	77.3	55.4	100.	−16.20	7.53	42.6	−15.11
H	H	M	49.57	10.31	18.6	83.1	76.0	54.7	100.	−17.94	7.41	37.4	−16.80
H	H	H	51.74	10.54	12.5	82.0	74.8	53.9	100.	−19.62	7.32	31.7	−18.49

Note: Initial schedule delay cost markup $\mathscr{A}$ in percent. Optimal toll τ in dollars. Flexible commuters commuting π in percent. Route utilization rates $\tilde{v}$ in percent. Share of commuter types I and F taking routes A and B in percent. Cost savings over baseline $\Delta\mathscr{C}$ in percent. Average of 1,000 simulation runs.

Both tables characterize 27 scenarios of combinations of the α, ψ, and φ parameters defined in table 3.10 as low (L), medium (M) and high (H). We consider the MMM scenario as our baseline—but we do not have data to reliably identify any of the three parameters.

The column presenting the $\mathscr{A}$ measure reveals the empirical importance of the schedule delay cost. When all vehicles are dispatched (no toll), the extra time due to schedule delay ranges from about 13% in the LLL scenario to abut 46% in the HHH scenario in table 3.11 and about 7% to 52% in table 3.12. On average, our model suggests that schedule delays add about 27% cost to inflexible commuters. For a 20-minutes driving time, schedule delay adds another 5–6 minutes. It is immediately apparent that $\mathscr{A}$ increases with all three parameters: a larger share of inflexible commuters, higher degree of inflexibility, and higher time cost of inflexible commuters all increase the economic importance of schedule delay.

The full-information optimal toll ('Optimal' in both tables) is characterized in both tables with dollar figures. In the single-toll case they range from \$1.19 to \$3.49, a significant range in relative terms (2.9:1).

The toll induces a number of flexible drivers to cease commuting, and thus the share of flexible commuters who continue to commute (π) varies between about 72% and 81%. A drop of traffic volume of up to 30% due to tolls is perfectly within the range of what has been found in practice. As the cost $\mathscr{A}$ of schedule delay increases, so does the toll, which in turn reduces π. The tolls that our model finds look quite reasonable in terms of magnitude.

It may surprise at first glance that the utilization rate of roads is increasing when $\mathscr{A}$ goes up, but note that as α goes up, the share π applies to a smaller cohort of commuters. So in the HHL scenario route A will get utilized much more than in scenario LLH. What is clearly apparent, though, is that route A always has a lower utilization rate of route B, a direct result of the toll. This wedge in utilization makes route A faster than route B, allowing sorting of high-value commuters and low-value commuters. This result agrees with Small and Yan (2001) and Verhoef and Small (2004) who address that commuter heterogeneity plays an important role for such "Lexus Lane tolling" to work. The beneficial effect of tolling comes both from reducing demand (a scale effect) and from sorting commuters (a composition effect). This becomes even clearer when we turn to the question which commuter takes which route. Without tolling, commuters of both types are perfectly indifferent between taking route A and B and we start with a 50:50 mix. Tolling ensures that a higher fraction of inflexible commuters takes the tolled route A, and more flexible commuters take the untolled route B. This sorting effort is strongest, unsurprisingly, win the HHH scenario.

The column $\Delta\mathscr{C}$ shows the cost savings of the tolled equilibrium over the untolled equilibrium. The overall cost savings that we find from tolling range from about 3% to 7%. They are highest in our MHH and HHH scenarios: inflexibility is pronounced and inflexible commuters have higher time cost than flexible commuters. Our cost measure assumes that all toll revenue is returned in full to commuters.

The limited-information optimal toll ('classic' in both tables) where the social planner sets the toll based on observable driving time, but where commuters choose route based on their full information about schedule delay, shows some rather important features that have strong implications for policy design. The demand response from the 'classic' toll (identified by the magnitude of π) is rather similar to the full-information optimal toll case. So even though the classic toll is always a bit too small and thus the induced scale (demand) effect is a little less, the cost savings ($\Delta\mathscr{C}$) are quite comparable to the full-information case. The conclusion is thus that even the suboptimal toll induces a sufficiently strong sorting effect that generates most of the gains for commuters. In other words, even when schedule delay costs are large, incomplete information about them does not result in hugely inefficient tolls.

We now turn to the dual-toll case where the routes are highly asymmetric. Recall that route A is slow but predictable, and route B is fast but unpredictable. We make route A 6 minutes slower than route B (23 versus 17 minutes), but also make route A twice as reliable as route B by choosing appropriate T_ρ values. Table 3.12 shows the results. We now find that much higher tolls are needed to get the desired sorting effect. For our low-α scenarios we find perfect sorting: inflexible commuters take route A, and flexible commuters take route B. All the gains come from the demand (scale) effect, as the dual toll has a neutral sorting effect. The self-sorting for the flexible types remains in place throughout our range of parameters. Flexible commuters always prefer the faster route. However, as the toll increases, an increasing number of low-time-cost inflexible commuters chooses to switch from A to B because route B is now less busy.

The tolls that we compute cover again a wide range, and because they tend to be significantly higher than in our first scenario, the demand response is much stronger. When schedule delay costs are high, the toll strongly favours policies that benefit the group of inflexible commuters and starves off commutes from flexible commuters.

We note again a key result with respect to cost savings ($\Delta\mathscr{C}$): there is hardly any difference between the 'optimal' and 'classic' tolls, which suggests again that the beneficial effect from tolls with respect to sorting is induced very effectively even when the toll level is suboptimal. While perfect sorting occurs for the low-α scenario, classic and optimal toll are actually the same because the tolls do not induce any sorting. Only when the share of inflexible commuters is large enough to allow additional sorting, do we see a gap opening between classic and optimal tolls.

Our Asymmetry scenario generates welfare gains from a demand effect, rather than through sorting. The commuters are already self-sorting because of the asymmetry in road conditions, and thus there is not much that can be gained from differential tolls. One key insight therefore is that asymmetric tolling is useful when road conditions are similar, while symmetric tolling is useful when road conditions are dissimilar.

The results in table 3.12 can also be interpreted as mode choice: flexible commuters commute by car, and most inflexible commuters prefer the (slower but more reliable) train. As the toll rises, some of the inflexible commuters also commute by car because now the traffic density on route B decreases sufficiently to make this route fast enough to remain competitive for inflexible commuters with low time cost. Tolling increases the use of the reliable but slower route A and at the same time reduces congestion on the faster but volatile route B.

What we have gained from comparing our two scenarios is one key insight: optimal tolling induces sorting (composition) and demand (scale) effects. When conditions are similar across routes, the benefits from tolling arrive more from the sorting effect and required tolls tend to be low. When conditions are dissimilar across routes with a strong negative correlation of reliability and speed, then larger tolls are needed to induce a stronger demand respond to relieve congestion because the heterogeneous commuter groups self-select into the optimal routes already—it isn't necessary to create the selection effect artificially.

3.6.4 Distributional Outcomes and Fairness

The distributional outcome of mobility pricing have received much attention because perceived unfairness of such a pricing scheme is perhaps the most formidable obstacle to its implementation (Eliasson, 2016).

We pursue two strategies to characterize the distributional outcome. The first method looks at the net gains of different commuter groups relative to our baseline no-toll scenario for the commuters in each of five groups:

1. Inflexible commuters taking (tolled or slow-but certain) route A, W_A^I.
2. Inflexible commuters taking (untolled or fast-but-volatile) route B, W_B^I.
3. Flexible commuters taking (tolled or slow-but certain) route A, W_A^F.
4. Flexible commuters taking (untolled or fast-but-volatile) route B, W_B^F.
5. Flexible commuters who are not commuting because the commuting cost exceeds their cutoff cost, $W_\varnothing^F$.

Table 3.13: Distributional Outcomes: 'Lexus Lanes' Scenario

Parameters				Optimal Toll							Classic Toll						
α	ψ	φ	SI$^{\circ}$	R	W_A^I	W_B^I	W_A^F	W_B^F	$W_{\varnothing}^F$	SI	R	W_A^I	W_B^I	W_A^F	W_B^F	$W_{\varnothing}^F$	SI
L	L	L	45.9	122	93	10	1	18	116	44.2	104	89	12	9	19	101	44.5
L	L	M	46.9	133	121	5	−9	21	123	45.0	108	114	9	4	22	103	45.4
L	L	H	48.1	144	156	−2	−21	25	130	45.8	114	145	6	−2	26	107	46.3
L	M	L	45.8	133	110	8	−3	20	126	44.1	107	103	12	9	21	103	44.4
L	M	M	47.0	145	143	2	−15	24	133	44.9	110	131	9	4	25	105	45.4
L	M	H	48.3	158	183	−6	−30	29	140	45.8	117	167	6	−2	29	109	46.5
L	H	L	45.8	145	129	6	−8	23	135	43.9	109	118	12	9	24	105	44.4
L	H	M	47.1	158	168	−1	−22	27	143	44.9	113	149	9	4	27	107	45.5
L	H	H	48.6	172	213	−11	−39	32	150	45.9	120	189	6	−2	32	112	46.7
M	L	L	46.9	171	114	22	−5	34	159	44.8	147	111	24	6	33	139	45.1
M	L	M	48.0	193	146	20	−19	46	173	45.6	162	143	24	−2	44	150	46.0
M	L	H	49.1	219	183	17	−38	64	190	46.3	181	180	23	−13	59	163	46.7
M	M	L	46.9	189	134	22	−12	41	173	44.6	155	130	25	5	39	146	45.0
M	M	M	48.1	214	171	19	−29	56	189	45.5	171	166	25	−3	52	157	46.0
M	M	H	49.2	243	212	16	−50	78	209	46.2	193	207	25	−15	69	173	46.8
M	H	L	46.9	211	156	22	−20	49	189	44.4	162	149	27	5	45	152	45.0
M	H	M	48.2	238	197	19	−40	68	208	45.4	179	188	27	−3	60	164	46.1
M	H	H	49.4	270	240	15	−64	95	230	46.2	205	234	28	−16	81	184	47.0
H	L	L	48.3	238	139	44	−12	62	214	45.7	216	139	44	−1	60	197	45.9
H	L	M	49.0	274	173	49	−30	91	240	46.2	246	174	50	−15	86	220	46.5
H	L	H	49.7	309	207	54	−48	129	269	46.7	283	208	55	−32	121	250	46.9
H	M	L	48.3	267	162	48	−21	78	237	45.5	234	162	49	−4	73	212	45.8
H	M	M	49.1	307	198	54	−42	113	267	46.1	268	199	55	−20	104	239	46.5
H	M	H	49.8	347	233	59	−62	160	302	46.7	310	235	61	−39	146	275	47.0
H	H	L	48.2	298	186	53	−31	96	261	45.3	252	186	54	−7	88	227	45.8
H	H	M	49.2	338	225	59	−51	138	294	46.0	290	225	61	−23	124	258	46.5
H	H	H	49.9	386	259	66	−76	194	338	46.6	339	262	68	−46	174	303	47.0

Note: Suits Index (SI), with superscript $^{\circ}$ for baseline; average annual rebate (R); welfare gain (W) for inflexible (I) and flexible (F) commuters taking routes A and B, or not commuting ($\varnothing$). Optimal toll: full information; classic toll: limited information. Average of 1,000 simulation runs. See appendix table B.3 for corresponding measures of average time cost by commuter group.

These five groups are strongly sorted by time costs. The commuters with the highest time cost prefer to take the tolled route, while the commuters with the lowest time cost prefer to take the untolled route. The W metrics are scaled up to 250 commutes per year in order to capture the annualized dollar value of the changes in welfare. A positive W represents a gain over the untolled scenario, and a negative W represents a loss over the untolled scenario.

In order to calculate distributional effects, we need to make a specific assumption about how the revenue from the tolls are recycled. We assume that all commuters receive an equal share of the toll revenue.[20]

Our second approach follows a familiar approach for investigating the regressivity or progressivity of tax policies, the Suits index. This method is also used in Ke and Gkritza (2018), who apply this measure to

[20]It would be easy to explore alternative assumptions such as not returning revenue to commuters, or only returning a fraction of revenue to commuters.

Table 3.14: Distributional Outcomes: 'Asymmetric' Scenario

Parameters				Optimal Toll							Classic Toll						
α	ψ	φ	SI°	R	W_A^I	W_B^I	W_A^F	W_B^F	$W_\varnothing^F$	SI	R	W_A^I	W_B^I	W_A^F	W_B^F	$W_\varnothing^F$	SI
L	L	L	45.8	805	−119	0	0	35	651	38.2	805	−119	0	0	35	651	38.2
L	L	M	46.7	721	−94	0	0	32	583	39.7	721	−94	0	0	32	583	39.7
L	L	H	47.8	644	−70	0	0	29	521	41.2	644	−70	0	0	29	521	41.2
L	M	L	45.8	805	−114	0	0	36	651	38.3	805	−114	0	0	36	651	38.3
L	M	M	46.8	720	−89	0	0	32	583	39.8	720	−89	0	0	32	583	39.8
L	M	H	47.9	644	−63	0	0	29	521	41.4	644	−63	0	0	29	521	41.4
L	H	L	45.8	805	−109	0	0	35	652	38.4	805	−109	0	0	35	652	38.4
L	H	M	46.9	720	−82	0	0	32	582	40.0	720	−82	0	0	32	582	40.0
L	H	H	48.1	644	−57	0	0	29	521	41.7	644	−57	0	0	29	521	41.7
M	L	L	47.1	1,209	178	−19	0	−103	749	37.5	1,104	180	3	0	−68	724	38.2
M	L	M	48.4	1,159	268	−8	0	−125	683	38.8	1,052	261	21	0	−82	664	39.5
M	L	H	49.6	1,109	311	138	0	−150	619	40.0	997	301	172	0	−95	604	40.8
M	M	L	47.1	1,266	192	109	0	−121	758	37.4	1,123	198	128	0	−69	729	38.3
M	M	M	48.6	1,212	250	220	0	−146	689	38.8	1,063	250	245	0	−81	666	39.8
M	M	H	49.9	1,159	325	333	0	−174	622	40.0	1,010	317	368	0	−96	607	41.1
M	H	L	47.0	1,318	192	287	0	−141	764	37.3	1,127	204	297	0	−68	729	38.5
M	H	M	48.6	1,261	251	426	0	−170	693	38.7	1,071	258	441	0	−81	667	40.0
M	H	H	50.1	1,203	341	550	0	−200	622	40.0	1,018	337	582	0	−95	606	41.4
H	L	L	48.4	1,948	214	120	0	−134	1,103	37.0	1,632	234	160	0	−58	986	38.1
H	L	M	49.3	1,922	279	183	0	−173	1,062	38.1	1,589	301	232	0	−80	929	39.3
H	L	H	50.0	1,908	358	252	0	−216	1,042	39.0	1,554	383	314	0	−106	881	40.3
H	M	L	48.3	2,063	240	207	0	−163	1,143	36.9	1,653	270	250	0	−58	989	38.4
H	M	M	49.3	2,044	311	286	0	−208	1,113	38.0	1,608	344	341	0	−82	932	39.6
H	M	H	50.1	2,038	392	372	0	−255	1,106	39.0	1,570	430	445	0	−108	882	40.6
H	H	L	48.3	2,174	259	306	0	−196	1,187	36.8	1,667	300	349	0	−60	991	38.6
H	H	M	49.4	2,158	334	405	0	−245	1,164	38.0	1,620	380	463	0	−83	931	39.8
H	H	H	50.2	2,174	416	509	0	−299	1,185	39.0	1,580	474	592	0	−109	883	40.9

Note: Suits Index (SI), with superscript ° for baseline; average annual rebate (R); welfare gain (W) for inflexible (I) and flexible (F) commuters taking routes A and B, or not commuting ($\varnothing$). Average of 1,000 simulation runs. See appendix table B.4 for corresponding measures of average time cost by commuter group.

data from Oregon. The Suits index (scaled up by 100 for easier use) is calculated as

$$\text{SI} = 100 \left[1 - \frac{1}{2} \sum_i [T(y_i) - T(y_{i-1})](y_i - y_{i-1}) \right] \quad \in [-100, +100] \tag{3.53}$$

where y_i is the cumulative share of income, which we proxy for by the time cost ω_i, and the corresponding cumulative time cost $T(y_i)$. Also, $y_0 = 0$ and $T(0) = 0$. The Suits Index is negative when the policy is regressive, and positive when the policy is progressive. We find that in all our simulations the Suits Index is positive: the cost of commuting is progressive—time is more valuable to high-income earners.

The results of our analysis are reported in tables 3.13 and 3.14. We compare the Suits Index in the baseline scenario S° against the corresponding outcomes for the 'optimal' and 'classic' tolls, and we report the five W statistics for both types of tolls.

So who gains, and who loses? As a reference point, it is useful to report the average rebate (R in the tables) from the tolls, which is necessarily considerably more in the 'asymmetry' case than the 'Lexus lane' case because both routes are tolled. This amount is redistributed among the various commuter groups so that on average nobody gains or loses because our setup mandates that all toll revenue goes back to all commuters. Consider the baseline case in table 3.13. The average commuter pays $215 in tolls and gets just as much back in rebates, for a net zero gain. However, the different commuter groups fare differently. The affluent inflexible commuters who pay the toll and take route A tend to gain much more than the less affluent inflexible commuters who end up taking the untolled route. The toll-induced sorting effect benefits them the most. In some cases, especially when the overall proportion of inflexible commuters is small, the less affluent group of inflexible commuters ends up with net losses in some cases. The mid-affluent flexible commuters who take the tolled route A tends to lose as well, while the less affluent commuters who stay on route B gain slightly—but this group is large. The most affluent flexible commuters will cease commuting, and thus their gains tend to be the highest because they benefit from the rebate the most as their losses are constrained by their cutoff time (reservation cost). The picture for the 'classic' toll case is qualitatively similar, although the magnitudes are somewhat smaller.

The Suits Index summarizes the total distributional effect: there is a small drop in progressivity of the cost of commuting, typically not more than a few percentage points. In other words, tolling does not radically change the distributional outcomes in table 3.13. However, the picture in table 3.14 for the 'asymmetry' case reveals a different picture as the Suits Index drops about up to 10 percentage points. Tolling reduces the progressivity of commuting costs much more noticeably.

The results in the Asymmetry scenario in table 3.14 differ in some important ways from those in table 3.13. When the share of inflexible commuters is low, this group stands to lose, but when their share is high, they gain significantly regardless of which route they take. Tolling clearly benefits the inflexible commuters, and it also benefits the flexible commuters with high time cost who cease commuting. Who pays for it are the large number of flexible commuters who all take route B and pay the full toll, but increasingly share their route with the less affluent inflexible commuters. Note that in table 3.14 there are many zeros, which is because there are no travellers under such category, for example, no flexible travellers will ever take the slower but more reliable route A.

If we interpret the Asymmetry scenario as a mode choice problem, then the inflexible commuters who need to get to work are the ones that gain as long as they are the majority of commuters. Then the tolling system works in their favour. The flexible commuters who can choose to commute or not will gain if they are induced to stay home and lose if they still commute. That does not seem such an unfair outcome at first because it shifts the burden from the inflexible commuters (who need to get to work) to the flexible commuters (who could perhaps delay their trips). The only caveat is that the burden falls disproportionately on the less affluent flexible commuters with the lowest time value. This is what shows in the large drop of the Suits Index.

3.6.5 Correlation Structure of Time Costs with Inflexibility and Non-Commutes

Our baseline simulation explored variations in the time cost and inflexibility in a limited fashion by considering high, medium, and low scenarios for the time cost ratio φ of inflexible to flexible commuters, and the degree of inflexibility of the former group. We kept the decision about non-commutes "neutral" by keeping the distribution of the cutoff time for non-commutes, $T_i^{\varnothing}$, uncorrelated with the time cost ω_i.

The correlation between time cost and cutoff cost has not been explored systematically. We use our simulation to vary the correlation of time cost with inflexibility, $\rho_{\omega,\psi}$ and the correlation of time cost with cutoff cost, $\rho_{\omega,z}$ systematically between −1 and +1. We use 51 steps of 0.04, and thus 2,601 points in the space spanned by the two correlation coefficients. We use 100 replications for each correlation point, for a total of 260,100 simulations, and average over these replications. We illustrate our results in contour plots for four key variables: the optimal toll, the share of flexible commuters commuting, the cost savings, and the change in the Suits Index.

Figure 3.17: Simulation Results: Correlation Structure

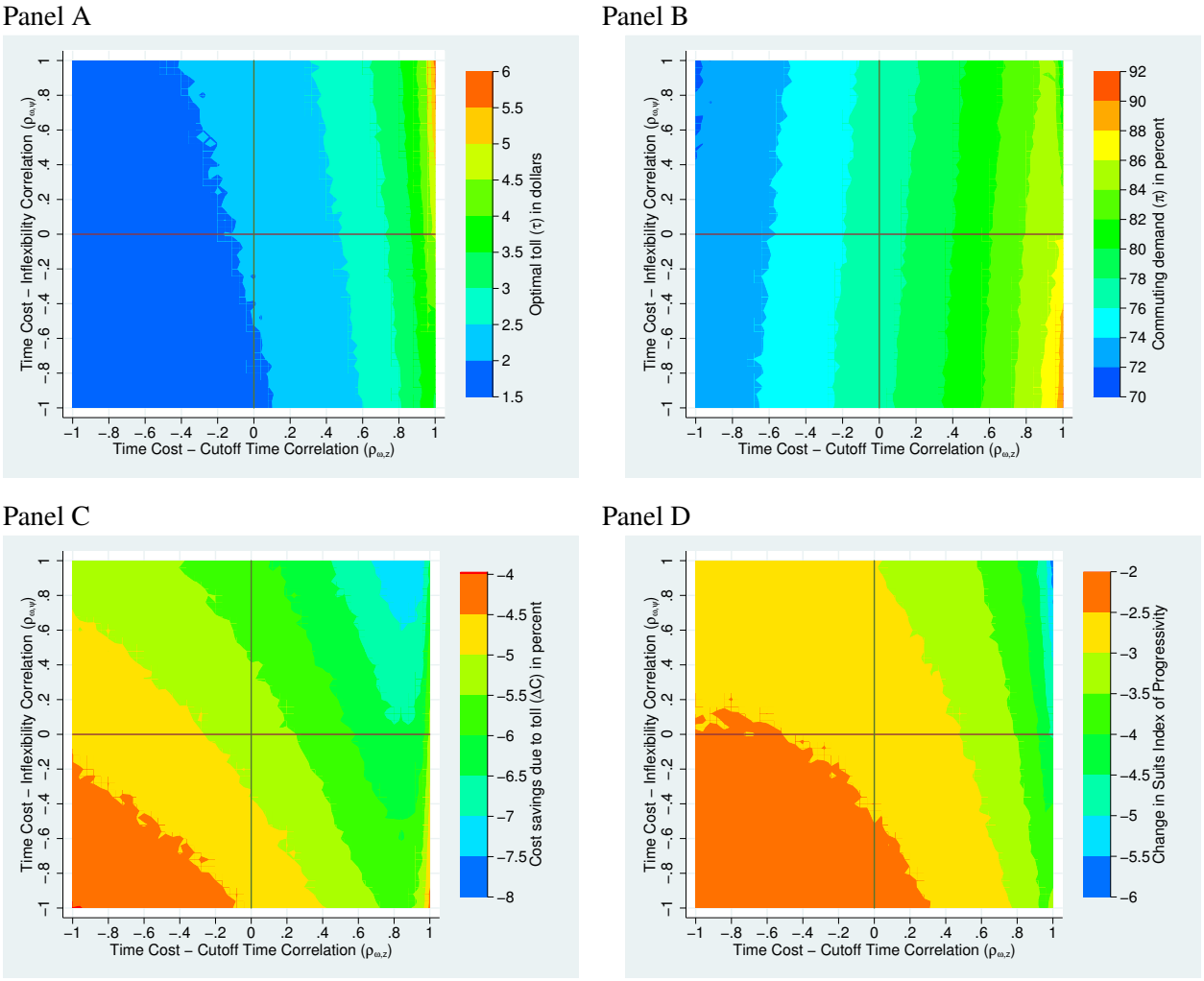


Figure 3.17 summarizes our results. Panel A shows the optimal toll τ; panel B shows the commuting demand π of flexible commuters; panel C shows the cost savings $\Delta\mathscr{C}$ from baseline (non-tolled) to optimal (tolled) scenario; and panel D shows the change in the Suits Index between baseline (non-tolled) and optimal (tolled) scenario. The four quadrants in each chart identify the four combinations of positive and negative correlations. If we had data on the distributions for our individual commuters we could pin down a single point in each diagram. By exploring the entire feasible range we hope to gain a better understanding how much these correlations matter.

Overall, the correlation $\rho_{\omega,z}$ of time cost with reservation price (cutoff time) for the flexible commuters is the dominant influence, which is visible in the strong vertical patterns in panels A and B for the optimal toll and demand response. The correlation $\rho_{\omega,\psi}$ of time cost with the degree of inflexibility of inflexible commuters matters much less in general.

For the optimal toll, the influence of a strong positive correlation between time cost and cutoff time is visible in Panel A in figure 3.17. When affluent flexible commuters also have high cutoff costs, tolls need to be higher. Consider the extreme case of perfect positive correlation when $z_i = \omega_i \bar{T}^{\varnothing}$ and all flexible commuters have the same cutoff time for their trips. To induce commuters to stay at home, ever higher tolls are needed to get a demand respond. If there is no (i.e., zero) correlation, on the other hand, this means that commuters simply have a reservation price for each trip that is independent of their income and time cost. In that case, tolls are just about half of what is needed as when $\rho_{\omega,z}$ approaches perfect positive correlation. Panel B corresponds to this and shows clearly that higher $\rho_{\omega,z}$ correlation generates less and less of a demand response.

Panels C and D in figure 3.17 explore the welfare implications of the correlation structure. Panel C shows the cost reduction $\Delta\mathscr{C}$ when going from no toll to optimal toll. The cost savings range from about 4 percent in the negative/negative quadrant to to over 7 percent in the positive/positive quadrant. Here, the correlation structure of $\rho_{\omega,\psi}$ between time cost and schedule inflexibility comes into play much stronger, as the cost function is strongly influenced by the size of ψ.

Panel D shows the change in progressivity (ΔSI) as captured by the change in the Suits Index when going from no toll to optimal toll. The pattern in panel D is quite similar to the pattern in panel C. The smallest changes to the progressivity of congestion occur again in the negative/negative quadrant, and the largest changes in the positive/positive quadrant. The economic logic of what we find in panels C and D is closely tied to the need to have higher tolls when the correlation coefficients increase to get the desired outcome.

The main conclusion of our correlation structure analysis is that it matters significantly. In order to design effective policy, we need to know more than just averages of the time cost, inflexibility, and reservation prices. Effective mobility pricing requires knowledge of the *distribution* of time cost, inflexibility, and reservation prices, as well as the two *correlations*. Based on what we found in panel A for the optimal toll, if we had incorrectly assumed zero correlations when in fact we had strong positive correlations, tolls would be much too low.

3.7 Caveats and Extensions

This empirical work in this paper explores commuting patterns in the Metro Vancouver region and demonstrated the importance of schedule delay costs. We have been able to quantify this importance, and our results also emphasize the importance to allow for congestion effects on the variance of commuting time. But there are important limitations on our results, and these limitations lay out an agenda for future empirical work in this area.

Our traffic data focuses on vehicle commutes. While this is a significant part of the overall commutes, it does not account for public transit and other forms of transportation. Work on road tolls has often suffered from an exclusive focus on road traffic without being able to quantify the substitution effects. In order to address commuting reliability, we need to know the entire choice set of commuters. An obvious extension of the work in this paper would be to investigate the reliability of public transit commuters where they are available. Where are there margins of substitutions?

The biggest caveat of our empirical work is that we know little about the commuters making their morning commutes. We do not know which part is truly inflexible, having to get to work, and which part is flexible, being able to delay or cancel trips. We also do not know the degree of inflexibility—the reliability standard that inflexible commuters adhere to. While we are able to make reasonable and plausible assumptions about this standard (a big improvement over guessing parameters in the α-β-γ model), we do not know the distribution of this measure ψ and how it is correlated with income or time cost ω. We also do not know how the reservation price (cutoff time) for flexible commuters is distributed and could be correlated with time cost ω. Our research shines a bright light on these gaps in our knowledge. Future empirical work, especially stated preference work, needs to explore these distributions and their correlation structure rigorously. Our work yields a better understanding of what to ask individual commuters in surveys,[21] gaining deeper insights into the crucial distributions of ω, ψ, and z, and the composition α of commuters.

With respect to designing road tolls and their distributional outcomes, our work highlights the importance of pinning down the aforementioned distributions. Almost inevitably, road tolls generate benefits to some commuters and losses to other commuters. Our work shows clearly that these gains and losses depend not only on the heterogeneity of the commuters, but also on the heterogeneity of road conditions (and competing mode choices). If anything, our work suggests that the frontier of transportation research lies in microeconomic studies of individual commuting behaviour. The literature on mobility pricing has come far in demonstrating that such pricing is hugely beneficial—and London, Singapore, and Stockholm show that this is not just a theoretical point. However, to overcome resistance to adopting mobility pricing especially in North America, we need to gain a better understanding of the distributional outcomes region-wide and in the presence of greater limitations on mode choice.

A major caveat in some of our theoretical and empirical work is the lack of allowing for variation across departure times. We focus on morning commutes, but as our data from Metro Vancouver shows, congestion follows a well-established pattern around a peak. Tolls that vary by time of day are the obvious answer to this

[21] Lo et al. (2006) has surveyed for the risk averse parameter λ in their paper, which is a similar measure to our reliability standard ψ. We would like to conduct similar surveys and ask individuals their expected frequency of being late, which would be even more intuitive.

issue. We do not explore the mix of commuters that are at different points of their journey and the resulting dynamics. However, we duly note that variation around departure times will also change the *composition* of commuters: our key variables α, ψ, φ, and the various correlations associated with ω, ψ, and z are all time dependent. An obvious extension of our work would involve exploring the dynamics effects of composition changes. If the composition changes, the tolls need to change accordingly. Even if (untolled) traffic volume were constant over time, the optimal toll could still change over time if the composition of commuters changes.

We acknowledge other important caveats of our approach. We assume in our optimal-toll simulation that motorists and toll-setters have full information on traffic conditions *and* commuter preferences (time cost, schedule inflexibility). Research work such as Lindsey (2009) and Boyles et al. (2010) explores how information asymmetry between motorists and toll-setters can affect the traffic outcome. Lindsey (2009) showed in theory that with uncertain capacity, tolls could be justified for optimal capacity only if tolls are set using information no less than drivers have. Boyles et al. (2010) discussed how optimal tolling scheme can be designed for cases where travelers have either full information on traffic network or only know the probability distributions, and tolls being static or could response to realized traffic. We acknowledge the existence of information asymmetry by exploring a classic (limited information) toll scheme in comparison to the full information optimal toll. We still allow for motorists to gain full information about the *distribution* moments of traffic flow.

3.8 Conclusions

When commuters plan their work-daily trips, a significant share of them needs to allow for additional buffer time to ensure that they arrive on time for work. This means that they need to take the dispersion of their commuting time into account. Just how large is the cost that is added to commutes as a result of this schedule delay? How much does the existence of this additional schedule delay affect optimal road tolling? If policy makers do not take this extra cost into consideration, will policies based on incomplete information about the true commuting time turn out to be significantly suboptimal?

The literature on congestion pricing has long acknowledged the importance of commuter heterogeneity. This paper deviates from the conventional assumption of schedule delay costs and replaces this with a travel time reliability standard that introduces an explicit metric of schedule inflexibility. This metric provides for an economically intuitive relationship between the standard deviation of travel time and the total travel time that allows sufficient buffer time upon arrival to meet the reliability standard. The notion of a reliability standard has several advantages. It allows us to employ two empirical distributions of travel time—Fréchet and log-normal—and link a measure of schedule inflexibility with the time mark-up on driving time. This setup also lends itself very effectively to empirical work we conduct in our paper. The reliability standard leads to a linearization technique that disaggregates the time mark-up factor into two separate components: a route-specific coefficient of variation, and a multiplier that reflects the schedule inflexibility. This decomposition technique allows us to model commuting time reliability very succinctly and derive new theoretical insights about optimal road tolls.

Empirically, our paper explores three separate avenues. First, we have assembled data from real-time

Google traffic data to determine commuting time across a sample of 2,000 commutes across the Metro Vancouver region over a 10-week period. Second, we have also used such data to measure speeds at numerous congestion points in the region: high-traffic bridges and tunnels. The latter we are able to link to traffic volume measurements that allow us to infer speed-volume relationships and dispersions. Third, we use simulation techniques to infer the effect of route characteristics on optimal tolling, and the effect of tolls on distributional outcomes.

Our analysis of region-wide commuting patterns provides an estimate of the economic importance of schedule delay cost, which we put at about 13% as our preferred estimate and an upper bound of 30%. The overall importance of schedule delay costs depends significantly on the share of inflexible commuters, which we are unable to ascertain. We also find evidence for a speed-reliability tradeoff: alternate routes are slower but less risky.

Analysis of our bridge traffic data through multiplicative heteroskedastic linear regression reveals that the *variance* of crossing time is related to the level of congestion, and not just the mean crossing time. However, while we find a strong positive congestion effect on variance for several bridges, there are some for which this pattern does not hold because of idiosyncracies in traffic management or differences in network topology. Overall, we believe that our results reinforce the notion that the second-order effect of congestion on the variance of travel time is important and in previous contributions to the literature has almost always been ignored.

Our simulation explores route choice under two different scenarios: one where two roads are symmetric but only one is tolled, and another one where two roads are asymmetric (one road is slower but more reliable) but both are tolled. Optimal tolls induce both a scale (demand) effect and a composition (sorting) effect. When route conditions are similar, even low-ish tolls can generate significant benefits from inducing sorting. When route conditions are dissimilar and are negatively correlated in terms of speed and reliability (which has similarities to transportation mode choice), we find that most of the traction comes from the demand response and little additional sorting.

We also investigate the welfare effects by studying who would gain and lose from a toll. Commuting costs are strongly progressive, and we find that tolls have a benign effect and tend to reduce progressivity as captured by the Suits Index. Overall, inflexible commuters stand to gain the most along with the flexible commuters induced to stay home, based on the notion that toll revenue is distributed equally to all commuters.

Our simulation study also sheds light on the use of a limited-information optimal toll in comparison to a full-information optimal toll. The limited-information case involves information asymmetry between commuters and social planner, where the social planner only observes driving time and determines an optimal toll without taking schedule delay into consideration, while commuters use their full private information to determine route choice. We find, surprisingly, that an imperfect toll is not as bad as feared. In our 'Lexus Lane' scenario even a toll that is somewhat too low will induce a strong sorting effect that will benefit commuters almost as much as the perfect full-information toll. In the asymmetric route case, the full-information toll does generate benefits, but they are small in magnitude. These results are encouraging news: mobility pricing does not have to be perfect as long as it induces a fair bit of sorting.

Our paper makes several novel contributions. We introduce the notion of a reliability standard as a viable and practical alternative to conventional α-β-γ cost models. Our modelling allows us to take the dispersion of commuting time into consideration directly, and it provides for a straight-forward extension of theoretical modelling of optimal tolls. The simplicity of our linearization approach makes this tool very useful for future work in this area. Our empirical work based on Metro Vancouver commutes shows clearly that dispersion matters, and we can put forward plausible estimates of the extra cost added by reliability considerations. Our congestion points analysis allows us to estimate key characteristics of the speed-volume relationship in the presence of dispersion. Simulation results also demonstrate that reliability considerations can lead to strong self sorting of commuters across routes, and that optimal congestion prices don't need to be perfect to achieve the desired outcome. Even limited-information tolls can generate beneficial sorting effectively.

Our research also suggests how future stated preference studies can better identify commuting behaviour for empirical work: we need to ask commuters whether (or how many) trips are flexible or inflexible, and how important it is to arrive on time. Asking people about how often they can afford to be late is much more intuitive than asking them about the cost of being late or arriving early, which most commuters will be unable to quantify. In order to gauge demand, it is also easier to ask flexible commuters at what duration they would rather not make a trip in order to gauge the distribution of cutoff times. Importantly, our research demonstrates the importance of the correlation of time cost with the measure of inflexibility (for inflexible commuters) and cutoff cost (for flexible commuters).

We conclude our paper with two policy recommendations. First, to better understand the welfare implications of mobility pricing, we need to study commuter heterogeneity better. Our paper has concrete suggestions about the questions we need to ask commuters in order to determine their type, inflexibility, and demand response. Second, policy makers need to address commuting reliability by providing options (e.g., rapid public transit that is immune to road congestion) or by allowing commuters to self-sort through prices. In addition to tackling congestion in urban cores (e.g., as is proposed for Manhattan), turning High-Occupancy Vehicle (HOV) lanes into High-Occupancy Toll Lanes (HOT) lanes is probably among the most practical options in North America.

Chapter 4

Carbon Emission Disclosure and Institutional Ownership

4.1 Introduction

With a growing attention on the global warming issue, legislation gradually emphasizes limiting the emission amount of greenhouse gas (GHG). Carbon pricing mechanisms have been well established and implemented, including cap-and-trade programs and carbon tax systems. Besides these legislation that explicitly sets a price on carbon emission, non-pricing mechanisms also came into being to motivate companies to reduce GHG emission, among which carbon reporting mechanism is one good example. Mandatory emission reporting mechanisms such as US EPA's greenhouse gas inventory track emission from individual facilities and suppliers of certain fossil fuels and industrial gases. This study focuses on a voluntary carbon emission disclosure program initiated by a global non-for-profit organization — Carbon Disclosure Project (CDP).

The objective of CDP is to raise more public attention towards big companies' environmental impact and to motivate companies to reduce their emission. CDP sends its annual report to more than 767 institutional investors such as mutual funds, banks, insurance companies, etc., and it hopes that institutional investors would take into account companies' carbon footprints when they make investment decisions. If institutional investors, upon receiving CDP reports, actually modify their portfolio to invest more on carbon efficient companies, firms might take actions to gradually reduce their emission. In CDP's news reports, it claimed that companies that did a better job in disclosing carbon emission information to CDP experienced a superior financial return, compared to the companies that do not disclose, or those disclose limited information. The causal relationship between voluntary carbon emission disclosure and firms' financial success, however, is dubious, because such correlation could be driven by many other forces. For example, if cautious managers are more likely to disclose their carbon emission and also more capable to bring profit to their firms, then the financial success would not be caused merely by the carbon emission disclosure behavior.

The questions asked are: (1) whether higher institutional ownership ratio induce stronger incentive to participate in such voluntary disclosure program, and (2) whether such disclosure cause a greater share of institutional ownership. The reasons to focus on institutional ownership are as follows. First, institutional

investors have more expertise in analyzing stock performance and hold a relatively stable number of shares, as a result, it is generally better for a company to have higher institutional ownership ratio. Therefore institutional ownership is an indicator of firms' financial success, and provides implication whether disclosing is a signal of better future financial performance. Second, since CDP devotes to inform institutional investors about large companies' carbon emission information, it makes sense to analyze whether the disclosed information affects companies' institutional ownership. Third, Kim and Lyon (2011) studied the relationship between CDP disclosure and stock price, and found no evidence suggesting participation in CDP would increase share prices. They further inferred that participation in CDP was not entirely voluntary but due to pressure from shareholders, regulators, and the institutional investors involved in the CDP. Their work provided a motivation to analyze institutional investors' responses to CDP disclosure reports.

In order to answer the two questions, this study builds a simultaneous equations model which captures the feature that both institutional ownership and CDP participation are endogenously determined and could affect each other. The results show that within US S&P 500 companies, (1) higher institutional ownership ratio lead to less disclosure to CDP, and (2) disclosure to CDP cause lower institutional ownership ratio.

4.2 Theoretical Framework

4.2.1 Motivation of Voluntary Carbon Disclosure

Voluntary disclosure of carbon emission information can be costly. First, by disclosing such information, the company might lose its green consumers or green investors if its performance does not meet their expectations. Second, the company might be giving away its production information to its competitors, since carbon emission is highly correlated with production activities. Third, the company seems to commit itself to maintain its carbon emission level near its currently disclosed emission amount, by making the gesture to participate in the CDP program. Why do some companies still disclose to CDP voluntarily, considering such unnecessary costs? Do the disclosing companies really have superiority in carbon efficiency, or do they do so mainly due to shareholder pressure?

Lyon and Maxwell (2004) provided several candidate models to unveil the true motivation for firms to conduct voluntary environmental activities. First, this effort might be taken to attract green consumers or green investors. If this motivation dominates, firms' altruistic environmental activities should be followed by their financial success either due to more profit raised from green consumers or more favored capital investment from green investors. Second, this could be a strategic behavior to preempt the regulation makers about the future environmental standard to be set for the industry. After firms indicate their current emission level, it will become a benchmark when future regulations are designed. If firms claim that their efforts to voluntarily reduce emission involve major capital investment, this sunk cost would become firms' bargaining power to persuade the regulation makers to loosen the environmental standard that was planned to be set. In this case, such voluntary environmental activities might not be associated with superior environmental performance or financial success. In one of their later studies, Lyon and Maxwell (2011) indicate that firms engage in more self-regulation when they perceive a greater threat of government regulation. They focused in firms' disclosure behavior in US Department of Energy's Voluntary Greenhouse Gas Registry and their

actual emission. This program allows firms to choose to report emission amount at aggregate level or at project level. Empirical results show that participants in this program engage in highly selective reporting — the participants actually increase total emission during the reporting period while they report reduction per project, however non-participants decrease aggregate emission over time.

Reid and Toffel (2009) verified the hypothesis that firms voluntarily disclose environmental information due to shareholder pressure using CDP data. More specifically, firms will be more likely to participate in CDP if it is targeted by a shareholder resolution on related topics such as environmental disclosure or climate change. They also found that firms subject to greater regulatory threats will be more likely to participate in CDP.

4.2.2 Impact of Voluntary Carbon Disclosure

Another question to ask is what impact will follow after a company discloses its carbon emission information to the public. The positive effect of carbon disclosure is to increase the company's transparency to the public, therefore reducing its long term risk in case of future regulation towards carbon emission changes. Meanwhile, the actual investment to reduce carbon emission might be against the company's objective of maximizing shareholders' value. Therefore the financial impact of voluntary carbon disclosure can be ambiguous.

On one hand, a related literature investigates the relationship between firms' social responsibility (CSR) and their financial performance, including stock return and institutional ownership, and suggests that such voluntary environmental activities benefit companies' financial performance. CSR counts for companies' responsibility to the stakeholders in many aspects, including environment protection. Heiner (1983) suggested that institutional investors are more capable than individual investors to assimilate and act upon information about socially responsible corporate practices. Graves and Waddock (1994) found a positive relationship between a company's CSR and the number of institutions holding the company's shares, yet the relationship between CSR and the percentage of shares owned by institutions is not statistically significant. Taking one step ahead, Cox et al. (2004) discovered that long-term institutional investment (defined by the shares held by pension funds, life insurance companies, etc.) is positively related to CSR. They concluded that CSR can be viewed as a significant determinant of firm exposure to long-term risks, therefore CSR can affect long-term investors' portfolio choice decision.

On the other hand, some studies suggest that voluntary carbon reduction activity leads to negative financial impact. Fisher-Vanden and Thorburn (2011) assert that corporations that commit to reduce GHG emission appear to be more likely to experience negative stock return, based on an empirical test regarding companies announcing membership in EPA's Climate Leaders. Jacobs et al. (2010) also found that announcements of voluntary emission reductions are associated with significant negative market reaction.

4.2.3 Relationship between Voluntary Carbon Disclosure and Institutional Ownership

Responding to CDP is not mandatory, therefore firms will evaluate their benefits and potential risks when deciding to disclose their carbon emission information. As explained above, firms' decision to voluntarily disclose their carbon emission, as a managerial strategy, and firms' institutional ownership, as one aspect of

financial performance, can affect each other simultaneously.

Firms' disclosure in carbon emission might cause change in its institutional investors' holding volume and hence its institutional ownership. Such voluntary environmental activity could affect companies' future financial performance because it reveals information to its investors. If investors take such voluntary disclosure activity as a good signal that the company is confident about its carbon emission amount and thus facing less financial risk in case of future regulation change, they would be more likely to appreciate such disclosure behavior. However, the reverse consequence might arise if either of the following situations happen. First, heavy emitters in the past could easily reduce its emission intensity and pick the low-hanging fruit, however the companies that have always been making investment to maintain low level of carbon emission face more difficulties to further reduce their emission intensity. Second, heavy emitters might have more incentive to preempt policy makers to loosen emission restriction level before any legislation is finalized, therefore they could be more likely to disclose to CDP. If such adverse selection phenomenon indeed exists and is recognized by the investors, the financial impact of voluntary disclosure could be negative. Besides, if disclosing firms are exposed to higher regulation risk, the disclosure behavior might also imply future financial burden from investors' perspective. Since CDP sends its annual reports to more than 767 institutional investors which manage assets of more than US$ 92 trillion, a natural question to ask is whether institutional investors of the S&P 500 companies react to whether they disclose to CDP. The reason that this study limits within S&P 500 companies is that CDP currently only sends out questionnaires to the biggest companies around the world, therefore within all US companies only S&P 500 companies are surveyed.

Meanwhile, institutional ownership could also affect firms' disclosure decision. If the institutional investors demand more transparent information about the companies' environmental performance which may or may not cause the firm future legislation risk, then we would expect higher institutional ownership leads to greater probability of disclosure. If instead, the institutional investors communicate directly with the companies, requesting and analyzing their carbon emission amount, then the companies with high institutional ownership would be less willing to disclose such information.

4.3 Model Specification

As described above, this study is mainly concerned with the relationship between companies' decision in voluntary disclosure of carbon emission information and their ratio of institutional ownership. To account for the endogeneity between the disclosure decision and institutional ownership, I formulate a simultaneous equations model and treat both of the variables as endogenous. Two equations are contained in this model, one explaining the impact of disclosure on institutional ownership ratio, the other one explaining the causal relationship from the opposite direction. Instrumental variables are added in both equations to identify the true impact of the two endogenous variables. The model setup is as below.

$$\begin{aligned} InstOwn_{it} = &\beta_0 + \beta_1 Discl_{it} + \beta_2 IV_{1it} + \beta_3 controls + \\ &year\,dummies + industry\,dummies + u_{it} \end{aligned} \tag{4.1}$$

$$Discl_{it} = \gamma_0 + \gamma_1 InstOwn_{it} + \gamma_2 IV_{2it} + \gamma_3 controls + year\,dummies + industry\,dummies + v_{it} \tag{4.2}$$

The measure of the variable *InstOwn* (institutional ownership) and *Discl* (disclosure dummy) will be defined in the next section. Equation (4.1) measures how the disclosure dummy variable influences the company's institutional ownership, and equation (4.2) measures how different levels of institutional ownership affect the likelihood that a company would participate in CDP's voluntary disclosure program. This simultaneous equations model takes into account the endogeneity generated from reverse causal relationship, with both equations containing the endogenous variables on both sides. The variables IV_1 and IV_2 serve as IVs for *InstOwn* and *Discl* correspondingly. In equation (4.1), the model assumes that *Discl* affects *InstOwn*, however *InstOwn* also influences *Discl* reversely, if firms with different levels of institutional ownership have various level of motivation to disclose to CDP. Adding an exogenous variable IV_1 that only correlates with *InstOwn* to the equation will minimize the reverse causality bias and helps identify the direct impact of *Discl* on *InstOwn*. The explanation for equation (4.2) is similar. The objective of equation (4.2) is to find out how institutional ownership ratio affects the decision to disclose to CDP, yet causality could also arise from the opposite direction. The exogenous variable IV_2 only correlates with *Discl*, and adding it to equation (4.2) eliminates the reverse causality bias. Candidates for IVs will be discussed in the following section.

4.3.1 Carbon Disclosure and Institutional Ownership Measures

The variable $InstOwn_{it}$ stands for stock i's institutional ownership in year t, and institutional ownership is defined by number of shares held by institutional investors divided by the total number of outstanding shares. The dummy variable $Discl_{it}$ equals 1 if company i chose to answer the questionnaire from CDP in year t, and equals 0 otherwise. The data source will be discussed in Section 4.4.

4.3.2 Instrumental and Control Variables

Instrumental variables need to be highly correlated with the regressors they serve as instruments for, but unrelated to the error terms in the equation those regressors are used.

To select for an IV for Institutional Ownership, I use the conclusion from Falkenstein (1996) and Gompers and Metrick (2001) that institutional investors prefer stocks with better liquidity and higher stock prices. Therefore the IV chosen here is the share turnover (trading volume within a year/number of outstanding shares at the end of the year) and stock price at the end of each year. These two variables should be exogenous to firm's disclosure decision, since there is no clear link between firm's stock liquidity and environmental performance or disclosure tradition.

The IV for Disclosure is "state greenness", which will be defined and justified in the following. State greenness is supposed to measure the environmental consciousness of the state that the company's headquarter locates in. The more one state cares about environmental sustainability, the more likely firms locat-

ing in that state would participate in such voluntary disclosure programs. However, where the company's headquarter locates should be independent of the company's institutional ownership. To measure the state greenness, I use the online searching frequency of "global warming" of each state from Google. Google publishes the regional searching frequency of popular phrases on Gtrend, which is publicly available. Another IV used is firm's clean energy strength which is a dummy variable taken from the KLD dataset. I assume that firms known of using clean energy devote more to environment protection, therefore they could be more likely to disclose to the public about their carbon emission.

The control variables that appear in both equations include accounting records such as firm size (measured by company market value and number of employees), revenue (measured by annual sales), and financial risk (leverage ratio, defined as total debt divided by total equity). Firms' past environmental performance and management skills could affect both the disclosure decision and institutional ownership, and therefore, should be controlled. As a candidate, corporate social responsibility (CSR) characterizes firms' true environmental performance and management skills, however the data for some firms are missing. Therefore I run regression models both with and without CSR data. In regression model 1, I include CSR data of: the number of environmental strengths, the number of environmental concerns, the number of corporate governance strengths, and the number of corporate governance concerns. Regression model 2 does not include CSR ratings in the control variables, and keeps all other variables the same as model 1.

4.4 Data

4.4.1 Carbon Disclosure Project

Carbon Disclosure Project(CDP) is an international organization that sends out annual surveys to the biggest companies around the world to ask about their carbon emission information. Each year more than half companies respond to the questionnaires and provide details about their carbon emission generated from production process, energy use, or supply chain. From year 2006, CDP has gained attention from an increasing number of institutional investors who subscribe for CDPs annual report at no charge or pay an annual fee to achieve enhanced access to the dataset. The main difference of CDP compared to other disclosure programs are that (1) it includes global carbon emission of a company, not just within a single country, and (2) this emission data is at company level instead of facility level, and involves companies from all industries instead of merely the emission intensive ones.

This study uses annual CDP reports containing US S&P 500 companies' self-disclosed information from year 2006 to 2013. In each year, every US S&P 500 company receives a questionnaire from CDP, and some of them choose to respond to CDP voluntarily. CDP then compiles the annual report and makes the data available to the public. Not only does CDP report the information provided by those companies who answered the questionnaire, it also lists out which companies failed to participate in this voluntary disclosure program. I use a disclosure dummy variable to indicate whether a company has replied to the questionnaire in a specific year. The dummy variable equals 1 if the company disclosed its carbon emission information to CDP in the referred year and it equals 0 otherwise.

In addition to studying firms' disclosure decision, I also investigated the disclosure quality and the ac-

tual carbon emission amount of those companies that gave back the questionnaire. In order to encourage the companies to deliver high quality questionnaires, CDP assigns a disclosure score to each completed questionnaire and highlights the companies with the highest scores. The more questions the company has answered in the questionnaire, the higher its disclosure score is. Moreover, companies can gain credit in disclosure score if they manage to seek for external verification institution to verify their disclosed information. I use this disclosure score to proxy for companies' disclosure quality. The actual emission amount data is used to calculate emission intensity which is defined by emission per employee. The emission accounted in this study only includes emission generated from the actual production process but not that generated from supply chain, because the latter information is disclosed by only a few companies.

4.4.2 Institutional Ownership

Data of stock holding volume and stock price information all comes from Thompson Reuter database. For each stock, the database provides quarterly data of institutional ownership, share turnover, etc. Institutional ownership of a stock is defined by the total amount of stock held by institutional investors (such as banks, pension funds, charity funds, etc.) divided by the total number of outstanding shares. To create annual data from quarterly data, I use the fourth quarter's data as a representation. Stock turnover is the total number of transactions of a stock within one year divided by the number of outstanding shares at year end.

4.4.3 Corporate Social Responsibility

Kinder, Lydenberg, Domini Research & Analytics (KLD) dataset provides annual rating of companies' performance in environment, employee relationships, governance, etc., which are known as corporate social responsibility (CSR). The ratings come from a comprehensive research of firms' financial statements, popular press, government reports, etc., and the database is widely used by institutional investors and scholars to evaluate firms' impact on the environment and stakeholders such as consumers, employees, communities, etc.

This study mainly concerns with two aspects of CSR, environment and corporate governance. When firms are making the decision of whether to voluntarily disclose their carbon emission information, which is an environmental concern, their decisions are clearly depending on their true environmental performance. Also, management skills, which can be captured by KLD dataset's corporate governance rating, influences a company's disclosure behavior in general.

Each year, the KLD dataset gives a 1/0 score for firms strengths and concerns in some specific aspects. The involved categories include environmental strengths, environmental concerns, corporate governance strengths, and corporate governance concerns.[1] The data is available from year 1995 to 2012.

The variables used in the regression model include: clean energy strength, transparency strength, number

[1] Environmental strengths include: a. environmental beneficial products and services; b. pollution prevention; c. recycling; d. clean energy; e. communications; f. property, plant, and equipment; g. others. Environmental concerns include: a. hazardous waste; b. regulatory problems; c. ozone depleting chemicals; d. substantial emissions; e. agricultural chemicals; f. climate change; g. other concerns. Corporate governance strengths include: a. limited compensation; b. ownership strength; c. transparency strength; d. political accountability strength; e. public policy strength; f. others. Corporate governance concerns include: a. high compensation; b. ownership concern; c. accounting concern; d. transparency concern; e. political accountability concern; f. public policy concern; g. governance structures controversies; h. others.

of environmental strengths other than clean energy, number of environmental concerns, number of corporate governance strength other than transparency strength, and number of corporate governance concerns.

4.4.4 Firms' Accounting Records

The regression model controls for some of the firms' financial and accounting characteristics such as firm size, profitability, industry. This data comes from Compustat North America database. The variables used in this study include number of employees, equity, market value, financial leverage ratio, global industrial code (GIC), revenue, etc.

Table 4.1 shows summary statistics and Table 4.2 shows cross correlations of the variables.

Table 4.1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
disclosure	0.656	0.475	0	1	3300
institutional ownership	0.787	0.178	0	1	4871
state greenness	4.297	0.126	3.951	4.533	4363
log share turnover	0.562	0.811	-6.977	3.334	4863
clean energy strength	0.293	0.455	0	1	3517
transparency strength	0.21	0.408	0	1	2918
# environmental strengths	0.579	0.905	0	4	3517
# environmental concerns	0.522	1.013	0	5	3574
# corporate governance strengths	0.092	0.305	0	2	2918
# corporate governance concerns	0.679	0.746	0	4	3574
log employee number	2.827	1.449	-3.297	7.696	4467
log market value	9.315	1.123	2.724	13.348	4108
log financial leverage ratio	1.096	0.695	0.044	6.899	4394
log sales	8.926	1.248	0.443	13.07	4517
log stock price	3.551	0.967	-4.605	7.058	4286

Table 4.2: Cross-correlation table

Variables	(1) disclosure	(2) Institutional Ownership	(3) state greeness	(4) log share turnover	(5) clean energy	(6) transparency strength
(1) disclosure	1.000					
(2) Inst. Ownership	-0.113	1.000				
	(0.000)					
(3) state greenness	0.063	-0.031	1.000			
	(0.000)	(0.041)				
(4) log share turnover	-0.106	0.324	-0.014	1.000		
	(0.000)	(0.000)	(0.376)			
(5) clean energy strength	0.282	-0.204	0.051	-0.162	1.000	
	(0.000)	(0.000)	(0.004)	(0.000)		
(6) transparency strength	0.246	-0.208	0.030	-0.181	0.474	1.000
	(0.000)	(0.000)	(0.115)	(0.000)	(0.000)	

4.5 Regression Results and Discussion

Table 4.3 exhibits the regression results of the model defined by Equation (4.1) and Equation (4.2).

From Table 4.3, one can observe that both coefficients of disclosure dummy and institutional ownership are negative and statistically significant, in both models. The results of model 1 directly reveals that: (1) controlling the other characteristics of a firm, it would have 6.82 percent points less institutional ownership if it changed from not disclosing to disclosing to CDP, and (2) controlling the other characteristics, the probability that a firm voluntarily discloses its carbon emission information would decrease by 8.83 percent points if its institutional ownership increased by 1 percent point. The coefficient given by model 2 is very close to that of model 1.

An advantage of using simultaneous equations model is that one can learn the feedback effects of endogenous variables. When the institutional ownership percentage changes, the probability of disclosure will also modify, which again leads to variations in institutional ownership. Taking this feedback effect into account, the marginal effects of the two variables have on each other are different from the coefficients directly read from the regression result table.

Considering the feedback effects, the results indicates: (1) controlling the other characteristics of a firm, it would have 7.26 percent points less institutional ownership if it changed from not disclosing to disclosing to CDP, and (2) controlling the other characteristics, the probability that a firm voluntarily discloses its carbon emission information would decrease by 9.40 percent points if its institutional ownership increased by 1 percent point.

Results show that higher institutional ownership does not lead to greater probability of disclosing. This finding somehow contradicts the conjecture in the literature that firms might be voluntarily disclosing their carbon emission information due to pressure from institutional investors (Kim and Lyon, 2011), although CDP is designed to share firms' carbon emission information with institutional investors. One possible reason that firms with higher institutional ownership tend to hide their carbon emission information from

Table 4.3: Simultaneous Equations Model

VARIABLES	Model 1		Model 2	
	Discl	InstOwn	Discl	InstOwn
institutional ownership	-0.883***		-0.882**	
	(0.334)		(0.310)	
state greenness	0.154**		0.227***	
	(0.0759)		(0.0694)	
clean energy strength	0.129***		0.158***	
	(0.0267)		(0.0224)	
number of environmental concerns	0.00216	-0.00948***		
	(0.0128)	(0.00296)		
number of environmental strengths	0.0168	0.00219		
	(0.0144)	(0.00363)		
transparency strength	0.0858***			
	(0.0268)			
number of corporate governance strengths	-0.0380	0.00619		
	(0.0331)	(0.00773)		
number of corporate governance concern	0.0208	0.00530		
	(0.0145)	(0.00346)		
log employees	0.0107	0.00294	0.0124	0.00311
	(0.0151)	(0.00361)	(0.0130)	(0.00314)
log market value	0.0310	-0.0346***	0.0199	-0.0335***
	(0.0195)	(0.00480)	(0.0182)	(0.00476)
log leverage ratio	0.0160	-0.00149	0.00589	-0.00162
	(0.0190)	(0.00459)	(0.0182)	(0.00435)
log sale	0.00512	-0.00438	-0.0073	-0.0069
	(0.0197)	(0.00477)	(0.0198)	(0.0047)
disclosure		-0.0682**		-0.0684**
		(0.0346)		(0.0286)
log share turnover		0.0526***		0.0487***
		(0.00588)		(0.0052)
log stock price		0.0360***		0.0363***
		(0.00434)		(0.00399)
Constant	0.0565	1.052***	-0.267	1.073***
	(0.578)	(0.0371)	(0.545)	(0.0326)
Observations	2,013	2,013	2487	2487
R-squared	0.121	0.333	0.0987	0.3356

Standard errors in parentheses
2-digit industry FE, and year FE are included.
*** p<0.01, ** p<0.05, * p<0.1

the public could be that firms communicate directly with their institutional investors about every aspect of their business, therefore disclosing to the rest of the investors becomes less necessary.

Another result demonstrates that more disclosure leads to lower institutional ownership, which means the disclosure behavior is not appreciated by institutional investors. If firms' disclosure behavior are not driven by philanthropic environmentalism, but by the motivation to manipulate the standard of future regulation on GHG emission, then disclosure does not serve as a signal of greater possibility of financial success in this case, therefore would not attract more institutional investment. This type of motivation amplifies in states where environmental regulations are generally stricter and are more likely to charge for GHG emissions in the future. That being the case, such regression result favors the perspective in existing literature that companies who disclose more details face stronger regulatory threat. The IV "state greenness" brings variation across different states, therefore distinct levels of state regulatory threat. If institutional investors believe that disclosing firms will likely encounter unfavorable GHG regulation changes and incur financial burden, the phenomenon that voluntary carbon disclosure results in lower institutional ownership could be easily understood.

The relationship between institutional ownership with (1) disclosure quality (measured by the disclosure score evaluated by CDP) and (2) emission intensity (emission from production process and energy use divided by number of employees) is also examined using the same method — simultaneous equations model. The regression results are in Table 4.4 and Table 4.5. The results indicate that these relationships are not statistically significant. Therefore no conclusion can be drawn to say that the level of institutional ownership would affect disclosure quality or actual emission intensity. This again strengthens the argument that such voluntary disclosure behavior might not be due to pressure coming from institutional investors, since higher institutional ownership ratio does not lead to better disclosure quality or lower emission intensity.

Nevertheless, comparing Table 4.3, Table 4.4 and Table 4.5, one can find that firms with greater number of environmental strengths tend to have better disclosure quality, and that firms with greater number of environmental concerns have higher level of emission intensity. The observations in these two regression only include the firms that have responded to CDP. Therefore, although the voluntary disclosure activity involves self-selection, the disclosed information can still reveal the firms' true environmental performance.

4.6 Conclusion

This study looks at the causal relationship between firms' decision in voluntarily disclosing its carbon emission information to CDP, a global non-for-profit organization that collects biggest firms' carbon emission data and publishes annual reports to make the data publicly available, and firms' level of institutional ownership. A simultaneous equations model is used to capture the endogenous characteristic of this question. The results show that the impact is negative and statistically significant, from both directions — the impact of disclosure on institutional ownership and the impact of institutional ownership on the probability of disclosure. This suggests that disclosing firms might not be the most carbon efficient ones, but rather those disclosing with other motivations such as preempting the regulators, and they could be subject to higher regulation risk. Therefore voluntary carbon disclosure mechanism such as CDP is sub-optimal in the attempt to unveil firms' environmental footprints to the public. Considering this point, firms' voluntary disclosure

to CDP can not serve as a positive signal of their greenness status or financial success potential. Mandatory mechanisms should be considered in the future to realize the objective of disclosing firms' carbon footprints and motivating them to take action to reduce emission. If disclosure is mandatory, the information released would directly inform investors of firms' environmental performance, and there would be no selection problem in this case. Nevertheless, if investors want to learn some knowledge about firms' true environmental performance from CDP reports, they can specifically inspect the firms' self-disclosed emission amount and their disclosure quality, which is available only from the companies that chose to participate in CDP.

This study still faces a lot of challenges, and the most important one being finding a stronger instrumental variable for the disclosure dummy variable. In cases where the companies' headquarter location state affects its institutional ownership, the IV would not be valid. This might happen if stocks of companies from some specific states consistently show superior financial performance and are favored by institutional investors, or if more institutional investors are located in some states and they tend to invest in companies that are physically close to them because of lower information costs.

Another potential improvement of this research is to decompose institutional holding by various types of institutional investors. For example, public investors might value more in corporate social responsibility of the stocks they invest in, therefore more likely to invest in companies that are willing to disclose their environmental performance. Moreover, some institutional investors, such as pension funds, may have longer investment horizon than the others that favor short-term appreciation in stocks. If better environmental performance requires higher operational cost in the short-term but brings cost-saving in the long-term, then institutional investors with longer investment horizon would value more for firms that disclose to CDP.

As legislation changes, UK has launched a mandatory carbon reporting project and the legislation came into effect in October 2013. All quoted companies in the UK are subject to this new regulation. An interesting follow-up research could be conducted to discover this mandatory carbon reporting project and see whether the companies that already participated in CDP emit more or less compared to the companies that failed to respond to CDP. A better idea of whether there is adverse selection problem in participation in CDP can be elicited with the new data from the UK's mandatory reports.

Table 4.4: Disclosure Quality and Institutional Ownership

VARIABLES	disclosure quality	institutional ownership
Institutional Ownership	-15.44	
	(19.33)	
state greenness	2.305	
	(5.116)	
clean energy strength	6.499***	
	(1.589)	
number of environmental concerns	0.234	-0.0118***
	(0.788)	(0.00393)
number of environmental strengths	2.606***	0.00377
	(0.853)	(0.00604)
transparency strength	3.866**	
	(1.610)	
number of corporate governance strengths	0.494	0.0174*
	(2.025)	(0.0105)
number of corporate governance concerns	-0.825	-0.00437
	(0.915)	(0.00478)
log employees	0.775	0.00404
	(0.984)	(0.00513)
log market value	2.180*	-0.0287***
	(1.150)	(0.00595)
log leverage ratio	0.982	0.00103
	(1.135)	(0.00610)
log sale	-2.017	-0.0105
	(1.292)	(0.00677)
disclosure score		-0.00143
		(0.00100)
log share turnover		0.0709***
		(0.00872)
log stock price		0.0334***
		(0.00593)
Constant	47.31	1.080***
	(34.65)	(0.0619)
Observations	994	994
R-squared	0.197	0.384

Standard errors in parentheses
2-digit industry FE, and year FE are included.
*** p<0.01, ** p<0.05, * p<0.1

Table 4.5: Emission Intensity and Institutional Ownership

VARIABLES	emission intensity	institutional ownership
institutional ownership	-2.232	
	(1.875)	
state greenness	-0.202	
	(0.505)	
clean energy strength	0.255*	
	(0.149)	
number of environmental concerns	0.380***	-0.00193
	(0.0693)	(0.0145)
number of environmental strengths	0.0295	-0.00141
	(0.0769)	(0.00515)
transparency strength	0.0273	
	(0.138)	
number of corporate governance strengths	-0.191	0.00676
	(0.183)	(0.0134)
number of corporate governance concerns	0.0944	-0.00492
	(0.0888)	(0.00657)
log employees	-0.790***	-0.0107
	(0.0956)	(0.0285)
log market value	-0.0194	-0.0317***
	(0.114)	(0.00786)
log leverage ratio	0.0709	-0.00403
	(0.113)	(0.00782)
log sale	0.193	-0.00815
	(0.119)	(0.0108)
emission intensity		-0.0160
		(0.0348)
log share turnover		0.0726***
		(0.0107)
log stock price		0.0303***
		(0.00892)
Constant	5.953	1.107***
	(3.709)	(0.103)
Observations	717	717
R-squared	0.536	0.453

Standard errors in parentheses
2-digit industry FE, and year FE are included.
*** p<0.01, ** p<0.05, * p<0.1

Chapter 5

Conclusion

In this dissertation, we investigate the efficiency of three existing or potential policies targeting on environmental or urban transportation issues.

Chapter 2 suggests that in order to promote the adoption of electric vehicles (EVs), the government should invest on building more public charging stations, in addition to providing cash incentives to individual EV purchasers. Meanwhile, subsidies giving to luxurious EVs or affluent consumers could be cut down. Empirical results demonstrate that smaller EVs are more inelastic towards public charging stations than larger EVs, because of a greater possibility of exhausting their batteries during a trip. Larger EVs, on the other hand, are much more expensive and would be affordable by higher-income consumers. Empirical results show that more affluent consumers are less price sensitive, meaning that some of them would purchase luxurious EVs with or without subsidies. Therefore, sizing down subsidies to higher-income EV purchasers or to luxurious EVs while building more charging stations could contribute more to induce EV purchasing, meanwhile reducing the concern that potential buyers of luxurious EVs might free ride government subsidy.

Chapter 3 characterizes the cost of travel time reliability and discusses the optimal tolling scheme that considers for such cost. When commuters plan their work-daily trips, a significant share of them needs to allow for additional buffer time to ensure that they arrive on time for work. Deviating from the conventional assumption of schedule delay costs, we propose a travel time reliability standard that introduces an explicit metric of schedule inflexibility. This setup allows linearization technique that disaggregates the time mark-up factor into two separate components: a route-specific coefficient of variation, and a multiplier that reflects commuter's schedule inflexibility. This decomposition technique allows us to model commuting time reliability very succinctly and derive new theoretical insights about optimal road tolls. Empirical results indicate that schedule delay cost is economically important and that both mean and variation of travel time increase with congestion level. Our simulation explores route choice of a two-road system under two different scenarios: (1) symmetric roads with one road being tolled, and (2) asymmetric roads with both being tolled. Optimal tolls induce both a scale (demand) effect and a composition (sorting) effect. When route conditions are similar, even low-ish tolls can generate significant benefits from sorting. When route conditions are dissimilar and negatively correlated in speed and reliability, the most of the traction comes from the demand response and little additional sorting. Through a welfare analysis we find that tolling reduces progressivity of commuting cost.

Chapter 4 looks at the causal relationship between firms' decision in voluntarily disclosing its carbon emission information and their level of institutional ownership. Although environmental performance does not directly indicate financial success, it might put companies at risk in case of policy change, such as carbon tax. Institutional investors are more sophisticated and might account for such risk in their investing decisions. Moreover, due to the larger share of stocks held by institutional investors, they might put pressure on companies to improve their environmental performance. To capture the endogenous characteristics of this question (that comes from reverse causality), the study employs a simultaneous equations model. The results show negative influence from both directions — the impact of disclosure on institutional ownership and the impact of institutional ownership on the probability of disclosure. This suggests that disclosing firms might not be the most carbon efficient ones, but rather those disclosing with other motivations such as preempting the regulators, and they could be subject to higher regulation risk. Therefore voluntary carbon disclosure mechanism is sub-optimal in the attempt to unveil firms' environmental footprints to the public. Mandatory mechanisms should be considered in the future to realize the objective of disclosing firms' carbon footprints and motivating them to take action to reduce emission.

Bibliography

Abdel-Aty, M. A., R. Kitamura, and P. P. Jovanis (1997). Using stated preference data for studying the effect of advanced traffic information on drivers' route choice. *Transportation Research Part C: Emerging Technologies 5*(1), 39–50.

Anderson, M. L. and L. W. Davis (2018). Does hypercongestion exist? new evidence suggests not. Technical report, National Bureau of Economic Research.

Arnott, R. (2013). A bathtub model of downtown traffic congestion. *Journal of Urban Economics 76*, 110–121.

Arnott, R., A. de Palma, and R. Lindsey (1994, May). The welfare effects of congestion tolls with heterogeneous commuters. *Journal of Transport Economics and Policy 28*(2), 139–161.

Asensio, J. and A. Matas (2008, November). Commuters' valuation of travel time variability. *Transportation Research Part E: Logistics and Transportation Review 44*(6), 1074–1085.

Bailey, J., A. Miele, and J. Axsen (2015). Is awareness of public charging associated with consumer interest in plug-in electric vehicles? *Transportation Research Part D: Transport and Environment 36*, 1–9.

Bates, J., J. Polak, P. Jones, and A. Cook (2001). The valuation of reliability for personal travel. *Transportation Research Part E: Logistics and Transportation Review 37*(2-3), 191–229.

Beaud, M., T. Blayac, and M. Stéphan (2016). The impact of travel time variability and travelers' risk attitudes on the values of time and reliability. *Transportation Research Part B: Methodological 93*, 207–224.

Beresteanu, A. and S. Li (2011). Gasoline prices, government support, and the demand for hybrid vehicles in the united states. *International Economic Review 52*(1), 161–182.

Berkovec, J. and J. Rust (1985). A nested logit model of automobile holdings for one vehicle households. *Transportation Research Part B: Methodological 19*(4), 275–285.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, 841–890.

Börjesson, M., J. Eliasson, and J. P. Franklin (2012). Valuations of travel time variability in scheduling versus mean–variance models. *Transportation Research Part B: Methodological 46*(7), 855–873.

Boyles, S. D., K. M. Kockelman, and S. T. Waller (2010). Congestion pricing under operational, supply-side uncertainty. *Transportation Research Part C: Emerging Technologies 18*(4), 519–535.

Bureau of Public Roads (1964). Traffic assignment manual. *US Department of Commerce*.

Calabrese, F., G. Di Lorenzo, L. Liu, and C. Ratti (2011, April). Estimating origin-destination flows using mobile phone location data. *IEEE Pervasive Computing 10*(4), 36–44.

Carrion, C. and D. Levinson (2012). Value of travel time reliability: A review of current evidence. *Transportation Research Part A: Policy and Practice 46*(4), 720 – 741.

Carrion, C. and D. Levinson (2013). Valuation of travel time reliability from a gps-based experimental design. *Transportation Research Part C: Emerging Technologies 35*, 305–323.

Center for Sustainable Energy (2017). California air resources board clean vehicle rebate project, rebate statistics. data last updated [mar 1, 2017]. retrieved [mar 28, 2017] from https://cleanvehiclerebate.org/eng/rebate-statistics.

Chandra, A., S. Gulati, and M. Kandlikar (2010). Green drivers or free riders? an analysis of tax rebates for hybrid vehicles. *Journal of Environmental Economics and management 60*(2), 78–93.

Chen, A. and Z. Zhou (2010). The α-reliable mean-excess traffic equilibrium model with stochastic travel times. *Transportation Research Part B: Methodological 44*(4), 493–513.

Cohen, Y. (1987). Commuter welfare under peak-period congestion tolls: who gains and who loses? *International Journal of Transport Economics/Rivista internazionale di economia dei trasporti*, 239–266.

Cox, P., S. Brammer, and A. Millington (2004). An empirical examination of institutional investor preferences for corporate social performance. *Journal of Business Ethics 52*(1), 27–43.

de Palma, A. and N. Picard (2005). Route choice decision under travel time uncertainty. *Transportation Research Part A: Policy and Practice 39*(4), 295–324.

DeShazo, J., T. L. Sheldon, and R. T. Carson (2017). Designing policy incentives for cleaner technologies: Lessons from california's plug-in electric vehicle rebate program. *Journal of Environmental Economics and Management 84*, 18–43.

Devarasetty, P. C., M. Burris, and W. D. Shaw (2012). The value of travel time and reliability-evidence from a stated preference survey and actual usage. *Transportation research part A: policy and practice 46*(8), 1227–1240.

Drake, J. and J. L. Schofer (1966). A statistical analysis of speed-density hypotheses. *Highway Research Record 154*, 53–87.

Eliasson, J. (2016, November). Is congestion pricing fair? consumer and citizen perspectives on equity effects. *Transport Policy 52*, 1–15.

Engelson, L. and M. Fosgerau (2016). The cost of travel time variability: Three measures with properties. *Transportation Research Part B: Methodological 91*, 555–564.

Falkenstein, E. G. (1996). Preferences for stock characteristics as revealed by mutual fund portfolio holdings. *The Journal of Finance 51*(1), 111–135.

Federal Highway Administration (2016, December). Highway performance monitoring system - field manual. Technical report, U.S. Department of Transportation, Office of Highway Policy Information.

Federal Highway Administration (2017, October). Simplified highway capacity calculation method for the highway performance monitoring system. Technical report, U.S. Department of Transportation, Office of Highway Policy Information.

Fisher-Vanden, K. and K. S. Thorburn (2011). Voluntary corporate environmental initiatives and shareholder wealth. *Journal of Environmental Economics and management 62*(3), 430–445.

Fosgerau, M. and D. Fukuda (2012). Valuing travel time variability: Characteristics of the travel time distribution on an urban road. *Transportation Research Part C: Emerging Technologies 24*, 83–101.

Fosgerau, M. and A. Karlström (2010). The value of reliability. *Transportation Research Part B: Methodological 44*(1), 38–49.

Gaver, Jr., D. P. (1968). Headstart strategies for combating congestion. *Transportation Science 2*(2), 172–181.

Goldberg, P. K. (1995). Product differentiation and oligopoly in international markets: The case of the us automobile industry. *Econometrica: Journal of the Econometric Society*, 891–951.

Gompers, P. A. and A. Metrick (2001). Institutional investors and equity prices. *The quarterly journal of Economics 116*(1), 229–259.

Graves, S. B. and S. A. Waddock (1994). Institutional owners and corporate social performance. *Academy of Management journal 37*(4), 1034–1046.

Greenshields, B. D., W. S. Channing, and H. H. Miller (1935). A study of traffic capacity. In *Highway research board proceedings*, Volume 1935, pp. 448–477. National Research Council (USA), Highway Research Board.

Hackbarth, A. and R. Madlener (2013). Consumer preferences for alternative fuel vehicles: A discrete choice analysis. *Transportation Research Part D: Transport and Environment 25*, 5–17.

Hall, F. L. (1975). Traffic stream characteristics. In D. L. Gerlough and M. J. Huber (Eds.), *Traffic Flow Theory*, Chapter 2, pp. 2/1–2/36. Washington, DC: Federal Highway Administration, US Department of Transportation. Transportation Research Board (TRB) Special Report 165.

Hall, J. D. (2018a, July). Can tolling help everyone? estimating the aggregate and distributional consequences of congestion pricing. University of Toronto Working Paper.

Hall, J. D. (2018b, February). Pareto improvements from lexus lanes: The effects of pricing a portion of the lanes on congested highways. *Journal of Public Economics 158*, 113–125.

Heiner, R. A. (1983). The origin of predictable behavior. *The American economic review 73*(4), 560–595.

Holland, S. P., E. T. Mansur, N. Z. Muller, A. J. Yates, et al. (2016). Are there environmental benefits from driving electric vehicles? the importance of local factors. *American Economic Review 106*(12), 3700–3729.

Iida, Y. (1999). Basic concepts and future directions of road network reliability analysis. *Journal of advanced transportation 33*(2), 125–134.

Jackson, W. B. and J. V. Jucker (1982). Empirical study of travel time variability and travel choice behavior. *Transportation Science 16*(4), 460–475.

Jacobs, B. W., V. R. Singhal, and R. Subramanian (2010). An empirical investigation of environmental performance and the market value of the firm. *Journal of Operations Management 28*(5), 430–441.

Johannesson, S. (1931). *Highway Economics*. McGraw Hill.

Ke, Y. and K. Gkritza (2018, November). Income and spatial distributional effects of a congestion tax: A hypothetical case of oregon. *Transport Policy 71*, 28–35.

Keeler, T. E. and K. A. Small (1977). Optimal peak-load pricing, investment, and service levels on urban expressways. *Journal of Political Economy 85*(1), 1–25.

Kim, E.-H. and T. P. Lyon (2011). Strategic environmental disclosure: Evidence from the doe's voluntary greenhouse gas registry. *Journal of Environmental Economics and Management 61*(3), 311–326.

Knight, T. E. (1974). An approach to the evaluation of changes in travel unreliability: a "safety margin" hypothesis. *Transportation 3*(4), 393–408.

Kouwenhoven, M., G. C. de Jong, P. Koster, V. A. van den Berg, E. T. Verhoef, J. Bates, and P. M. Warffemius (2014). New values of time and reliability in passenger transport in the netherlands. *Research in Transportation Economics 47*, 37–49.

Li, H., H. Tu, and D. A. Hensher (2016). Integrating the mean–variance and scheduling approaches to allow for schedule delay and trip time variability under uncertainty. *Transportation Research Part A: Policy and Practice 89*, 151–163.

Li, S., L. Tong, J. Xing, and Y. Zhou (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.

Li, Z., D. A. Hensher, and J. M. Rose (2010, May). Willingness to pay for travel time reliability in passenger transport: A review and some new empirical evidence. *Transportation Research Part E: Logistics and Transportation Review 46*(3), 384–403.

Lieven, T. (2015). Policy measures to promote electric mobility–a global perspective. *Transportation Research Part A: Policy and Practice 82*, 78–93.

Lindsey, R. (2009). Cost recovery from congestion tolls with random capacity and demand. *Journal of Urban Economics 66*(1), 16–24.

Lo, H. K., X. Luo, and B. W. Siu (2006). Degradable transport network: travel time budget of travelers with heterogeneous risk aversion. *Transportation Research Part B: Methodological 40*(9), 792–806.

Lyon, T. P. and J. W. Maxwell (2004). *Corporate environmentalism and public policy*. Cambridge University Press.

Lyon, T. P. and J. W. Maxwell (2011). Greenwash: Corporate environmental disclosure under threat of audit. *Journal of Economics & Management Strategy 20*(1), 3–41.

McFadden, D. (1978). Modeling the choice of residential location. *Transportation Research Record* (673).

McNeill, G., J. Bright, and S. A. Hale (2017, Oct). Estimating local commuting patterns from geolocated twitter data. *EPJ Data Science 6*(1), 24.

Mersky, A. C., F. Sprei, C. Samaras, and Z. S. Qian (2016). Effectiveness of incentives on electric vehicle adoption in norway. *Transportation Research Part D: Transport and Environment 46*, 56–68.

Noland, R. B. and J. W. Polak (2002). Travel time variability: a review of theoretical and empirical issues. *Transport reviews 22*(1), 39–54.

Noland, R. B. and K. A. Small (1995, July). Travel-time uncertainty, departure time choice, and the cost of morning commutes. *Transportation Research Record 1493*, 150–158.

Noland, R. B., K. A. Small, P. M. Koskenoja, and X. Chu (1998). Simulating travel reliability. *Regional science and urban economics 28*(5), 535–564.

Opus International Consultants (2014, April). Traffic reports user documentation. Technical report, B.C. Ministry of Transportation and Infrastructure.

Petrin, A. and K. Train (2010). A control function approach to endogeneity in consumer choice models. *Journal of marketing research 47*(1), 3–13.

Pigou, A. (1920). *The Economics of welfare*. London: Macmillan.

Pipes, L. A. (1967). Car following models and the fundamental diagram of road traffic. *Transportation Research 1*(1), 21 – 29.

Polak, J. (1987). *Travel time variability and departure time choice: a utility theoretic approach*. Polytechnic of Central London Transport Studies Group.

Reid, E. M. and M. W. Toffel (2009). Responding to public and private politics: Corporate disclosure of climate change strategies. *Strategic Management Journal 30*(11), 1157–1178.

Senna, L. A. D. S. (1994, May). The influence of travel time variability on the value of time. *Transportation 21*(2), 203–228.

Sierzchula, W., S. Bakker, K. Maat, and B. van Wee (2014). The influence of financial incentives and other socio-economic factors on electric vehicle adoption. *Energy Policy 68*, 183–194.

Small, K. A. (1982). The scheduling of consumer activities: work trips. *The American Economic Review 72*(3), 467–479.

Small, K. A. and X. Chu (2003). Hypercongestion. *Journal of Transport Economics and Policy (JTEP) 37*(3), 319–352.

Small, K. A. and E. T. Verhoef (2007). *The Economics of Urban Transportation*. Routledge.

Small, K. A., C. Winston, and J. Yan (2005, July). Uncovering the distribution of motorists' preferences for travel time and reliability. *Econometrica 73*(4), 1367–1382.

Small, K. A., C. Winston, J. Yan, N. Baum-Snow, and J. A. Gómez-Ibáñez (2006). Differentiated road pricing, express lanes, and carpools: Exploiting heterogeneous preferences in policy design [with comments]. *Brookings-Wharton Papers on Urban Affairs*, 53–96.

Small, K. A. and J. Yan (2001). The value of "value pricing" of roads: Second-best pricing and product differentiation. *Journal of Urban Economics 49*(2), 310–336.

Taylor, M. A. (2013). Travel through time: the story of research on travel time reliability. *Transportmetrica B: transport dynamics 1*(3), 174–194.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

Tseng, Y.-Y. and E. T. Verhoef (2008). Value of time by time of day: A stated-preference study. *Transportation Research Part B: Methodological 42*(7), 607 – 618.

Uchida, K. (2014). Estimating the value of travel time and of travel time reliability in road networks. *Transportation Research Part B: Methodological 66*, 129 – 147.

Van den Berg, V. and E. T. Verhoef (2011). Winning or losing from dynamic bottleneck congestion pricing?: The distributional effects of road pricing with heterogeneity in values of time and schedule delay. *Journal of Public Economics 95*(7-8), 983–992.

Verhoef, E. T. and K. A. Small (2004). Product differentiation on roads. *Journal of Transport Economics and Policy (JTEP) 38*(1), 127–156.

Vickrey, W. S. (1969, May). Congestion theory and transport investment. *American Economic Review 59*(2), 251–260.

Vickrey, W. S. (1973). Pricing, metering, and efficiently using urban transportation facilities. *Highway Resarch Record 476*, 36–48.

Walters, A. A. (1961). The theory and measurement of private and social cost of highway congestion. *Econometrica: Journal of the Econometric Society*, 676–699.

Wang, N., H. Pan, and W. Zheng (2017). Assessment of the incentives on electric vehicle promotion in china. *Transportation Research Part A: Policy and Practice 101*, 177–189.

Xiao, Y., N. Coulombel, and A. De Palma (2017). The valuation of travel time reliability: does congestion matter? *Transportation Research Part B: Methodological 97*, 113–141.

Zhang, Y., Z. S. Qian, F. Sprei, and B. Li (2016). The impact of car specifications, prices and incentives for battery electric vehicles in norway: Choices of heterogeneous consumers. *Transportation Research Part C: Emerging Technologies 69*, 386–401.

Appendix A

Appendix for Chapter 2

A.1 Proof of Propositions

Proof of Proposition 1:

From cost function (2.1) it is easy to show that $c_{\text{BEV,NHC,ij}} - c_{\text{BEV,HC,ij}} > 0$ and $c_{\text{PHEV,NHC,ij}} - c_{\text{PHEV,HC,ij}} > 0$ under reasonable assumption that cost of installing home charger amortized to each mile P_{S} is trivial compared to the other cost components. The intuition is that having a reliable charging spot saves the trouble looking for public chargers meanwhile largely reducing the risk of getting stranded.

Proof of Proposition 2:

To simplify the analysis, assume that time to find a slow charging station (with level I&II chargers) $t_1(S_{\text{S},ij})$ and time to find a DC fast charging station $t_1(S_{\text{F},ij})$ are both decreasing functions of total number of charging stations S_i in this region, i.e., $\frac{\partial t_1(S_{\text{S},ij})}{\partial S_i} < 0$ and $\frac{\partial t_1(S_{\text{F},ij})}{\partial S_i} < 0$. The probability of failure to find a public charging station is also assumed to be a decreasing function of S_i, i.e. $\frac{\partial f_{ij}(S_{\text{S},ij}, S_{\text{F},ij})}{\partial S_i} < 0$. The influence of number of charging station having on EV operational cost is as below:

$$\frac{\partial c_{ij}}{\partial S_i} = \begin{cases} (\text{HC,BEV}): \ q_{ij}\left[\frac{w_i}{\delta R_j}\frac{\partial t_{1,F}}{\partial S_i} + (\frac{w_i t_{\text{s}}}{R_j} + m_{\text{s}})\frac{\partial f_{ij}}{\partial S_i}\right] \\ (\text{HC,PHEV}): \ 0 \\ (\text{NHC,BEV}): \ \frac{w_i}{\min(r_i, R_j)\max(\lfloor \delta R_j / r_i \rfloor, 1)}\frac{\partial t_{1,S}}{\partial S_i} + \frac{w_i q_{ij}}{\delta R_j}\frac{\partial t_{1,F}}{\partial S_i} + (\frac{w_i t_{\text{s}}}{R_j} + m_{\text{s}})\frac{\partial f_{ij}}{\partial S_i} \\ (\text{NHC,PHEV}): \ \frac{w_i}{\min(r_i, R_{\text{PE},j})\max(\lfloor R_{\text{PE},j} / r_i \rfloor, 1)}\frac{\partial t_{1,S}}{\partial S_i} \end{cases} \tag{A.1}$$

From equation (A.1) it is easy to see that $\frac{\partial c_{ij}}{\partial S_i}$ is non-positive in every case, and that $\frac{\partial c_{\text{NHC,BEV},ij}}{\partial S_i} < \frac{\partial c_{\text{HC,BEV},ij}}{\partial S_i}$, $\frac{\partial c_{\text{NHC,PHEV},ij}}{\partial S_i} < \frac{\partial c_{\text{HC,PHEV},ij}}{\partial S_i}$, i.e., $\frac{\partial c_{ij}}{\partial S_i}$ has greater absolute value for non home owners in both cases of BEV and PHEV. The intuition is that EV owners without home chargers rely more on public

chargers, whereas EV owners with access to home charging seldomly need to seek public chargers. Therefore adding public charging stations bring most benefit to those without opportunity of home charging.

Proof of Proposition 3:

To prove Proposition 3, one needs to calculate derivative of $\frac{\partial c_{ij}}{\partial S_i}$ regarding to R_j. For home owners, one can see:

$$\frac{\partial^2 c_{ij}}{\partial S_i \partial R_j} = \frac{\partial q_{ij}}{\partial R_j}[\frac{w_i}{\delta R_j}\frac{\partial t_{1,F}}{\partial S_i} + (\frac{w_i t_s}{R_j} + m_s)\frac{\partial f_{ij}}{\partial S_i}] - q_{ij}\frac{w_i}{\delta R_j^2}\frac{\partial t_{1,F}}{\partial S_i} - q_{ij}\frac{w_i t_s}{R_j^2}\frac{\partial f_{ij}}{\partial S_i}$$

The above term is positive because $\frac{\partial q_{ij}}{\partial R_j} < 0$, $\frac{\partial t_{1,F}}{\partial S_i} < 0$, $\frac{\partial f_{ij}}{\partial S_i} < 0$, and all other parameters are positive. Similarly, for non home owners, this term would be:

$$\frac{\partial^2 c_{ij}}{\partial S_i \partial R_j} = -\frac{w_i}{g(R_j)^2}g'(R_j)\frac{\partial t_{1,S}}{\partial S_i} + \frac{\partial q_{ij}}{\partial R_j}(\frac{w_i}{\delta R_j}\frac{\partial t_{1,F}}{\partial S_i}) - q_{ij}\frac{w_i}{\delta R_j^2}\frac{\partial t_{1,F}}{\partial S_i} - \frac{w_i t_s}{R_j^2}\frac{\partial f_{ij}}{\partial S_i}$$

where $g(R_j) = \min(r_i, R_j)\max(\lfloor \delta R_j / r_i \rfloor, 1)$ and $g'(R_j) \geqslant 0$, assuming fixed r_i. Together with condition $\frac{\partial q_{ij}}{\partial R_j} < 0$, $\frac{\partial t_{1,S}}{\partial S_i} < 0$, $\frac{\partial t_{1,F}}{\partial S_i} < 0$, $\frac{\partial f_{ij}}{\partial S_i} < 0$, and other parameters positive, one would reach conclusion that the term above is positive.

Therefore, with or without access to home charging, a greater electric range would increase $\frac{\partial c_{ij}}{\partial S_i}$, while decreasing its absolute value, i.e., the cost reduction effect brought by adding public charging stations decrease with BEV's electric range.

Proof of Proposition 4:

For BEV owners with home charging:

$$\frac{\partial c_{ij}}{\partial R_j} = \frac{\partial q_{ij}}{\partial R_j}[w_i(\frac{t_1(S_F)}{\delta R_j} + t_{2,j} + r_{ij}\frac{t_s}{R_j}) + r_{ij}m_s] - q_{ij}w_i(\frac{t_1(S_F)}{\delta R_j^2} + \frac{r_{ij}t_s}{R_j^2})$$

$$\frac{\partial^2 c_{ij}}{\partial R_j \partial w_i} = \frac{\partial q_{ij}}{\partial R_j}(\frac{t_1(S_F)}{\delta R_j} + t_{2,j} + r_{ij}\frac{t_s}{R_j}) - q_{ij}(\frac{t_1(S_F)}{\delta R_j^2} + \frac{r_{ij}t_s}{R_j^2})$$

Since $\frac{\partial q_{ij}}{\partial R_j} < 0$ and other parameters are all positive, it follows that $\frac{\partial c_{ij}}{\partial R_j} < 0$ and $\frac{\partial^2 c_{ij}}{\partial R_j \partial w_i} < 0$.

For PHEV owners with home charging:

$$\frac{\partial c_{ij}}{\partial R_{PE,j}} = w_i\frac{\partial q_{ij}}{\partial R_{PE,j}}(\frac{t_1(S_G)}{\delta R_{PG,j}} + t_{3,j})$$

$$\frac{\partial^2 c_{ij}}{\partial R_{\mathrm{PE},j}\partial w_i} = \frac{\partial q_{ij}}{\partial R_{\mathrm{PE},j}}\left(\frac{t_1(S_\mathrm{G})}{\delta R_{\mathrm{PG},j}} + t_{3,j}\right)$$

With $\frac{\partial q_{ij}}{\partial R_{\mathrm{PE},j}} < 0$, one can see $\frac{\partial c_{ij}}{\partial R_{\mathrm{PE},j}} < 0$ and $\frac{\partial^2 c_{ij}}{\partial R_j \partial w_i} < 0$.

For BEV owners without home charging:

$$\frac{\partial c_{ij}}{\partial R_j} = w_i\left[-\frac{t_1(S_\mathrm{S})}{g(R_j)^2}g'(R_j) + \frac{\partial q_{ij}}{\partial R_j}\left(\frac{t_1(S_\mathrm{F})}{\delta R_j} + t_{2,j}\right) - \left(q_{ij}\frac{t_1(S_\mathrm{F})}{\delta R_j^2} + r_{ij}\frac{t_\mathrm{s}}{R_j^2}\right)\right]$$

$$\frac{\partial^2 c_{ij}}{\partial R_j \partial w_i} = -\frac{t_1(S_\mathrm{S})}{g(R_j)^2}g'(R_j) + \frac{\partial q_{ij}}{\partial R_j}\left(\frac{t_1(S_\mathrm{F})}{\delta R_j} + t_{2,j}\right) - \left(q_{ij}\frac{t_1(S_\mathrm{F})}{\delta R_j^2} + r_{ij}\frac{t_\mathrm{s}}{R_j^2}\right)$$

where $g(R_j) = \min(r_i, R_j)\max(\lfloor \delta R_j / r_i \rfloor, 1)$. With $\frac{\partial q_{ij}}{\partial R_j} < 0$ and $g'(R_j) \geqslant 0$, one can see $\frac{\partial c_{ij}}{\partial R_{\mathrm{PE},j}} < 0$ and $\frac{\partial^2 c_{ij}}{\partial R_j \partial w_i} < 0$.

For PHEV owners without home charging:

$$\frac{\partial c_{ij}}{\partial R_{\mathrm{PE},j}} = w_i\left[-\frac{t_1(S_\mathrm{S})}{g_P(R_{\mathrm{PE},j})^2}g'_P(R_{\mathrm{PE},j}) + \frac{\partial q_{ij}}{\partial R_{\mathrm{PE},j}}\left(\frac{t_1(S_\mathrm{G})}{\delta R_{\mathrm{PG},j}} + t_{3,j}\right)\right]$$

$$\frac{\partial^2 c_{ij}}{\partial R_{\mathrm{PE},j}\partial w_i} = -\frac{t_1(S_\mathrm{S})}{g_P(R_{\mathrm{PE},j})^2}g'_P(R_{\mathrm{PE},j}) + \frac{\partial q_{ij}}{\partial R_{\mathrm{PE},j}}\left(\frac{t_1(S_\mathrm{G})}{\delta R_{\mathrm{PG},j}} + t_{3,j}\right)$$

where $g_P(R_{\mathrm{PE},j}) = \min(r_i, R_{\mathrm{PE},j})\max(\lfloor R_{\mathrm{PE},j}/r_i \rfloor, 1)$. With $\frac{\partial q_{ij}}{\partial R_{\mathrm{PE},j}} < 0$ and $g'_P(R_{\mathrm{PE},j}) \geqslant 0$, one can see $\frac{\partial c_{ij}}{\partial R_{\mathrm{PE},j}} < 0$ and $\frac{\partial^2 c_{ij}}{\partial R_{\mathrm{PE},j}\partial w_i} < 0$.

Therefore, operational cost decreases with electric range, and such effect enlarges with w_i — valuation of time. The intuition is that EVs with longer electric range need to be charged less frequently and have lower probability of getting stranded, thus reducing time cost of driving, and such characteristic would be appreciated more by people with higher valuation of time.

A.2 Choice of Vehicles under Baseline Scenario

The choices within BEVs, PHEVs, and ICE vehicles rely on trade-off between time cost and monetary cost, when price and vehicle attributes are the same. Using currently available car model attributes data listed in Table 1 and assuming distribution of peoples' daily driving distance r_i follows log normal distribution with $\mu_i = 3.25$, $\sigma_i^2 = 0.75$, this section presents how the valuation of time can affect the operational cost of BEVs, PHEVs, and ICEVs, thus flipping preference towards them.

The vehicles in Table A.1 are divided to two groups, within each group the vehicles are deemed as substitutes with similar price and amenity. The vehicles in the same group are assumed to differ only in the way they are refueled. The first group represents for energy efficient small vehicles which has one model of ICEV, PHEV, BEV respectively, whereas the second group represents for luxurious vehicles which only

includes an ICEV and a BEV[1].

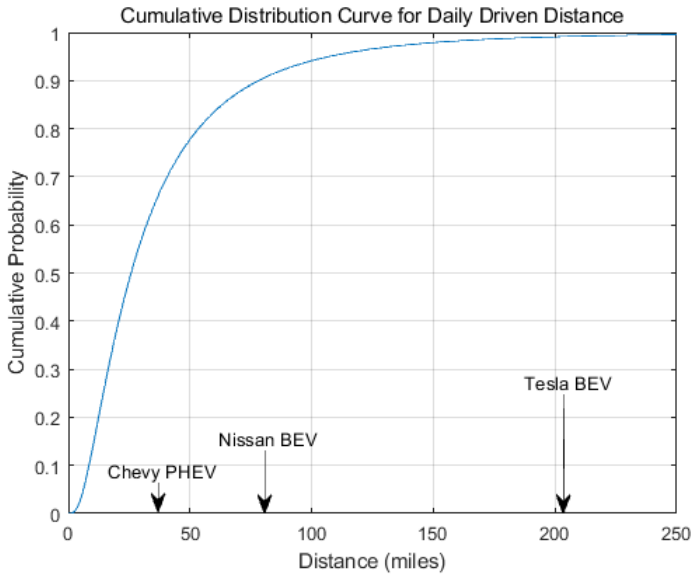
Table A.1: Characteristics of Representative EVs

Make-Type	price	electric range	gas range	mi/kWh	mi/gallon	refuel time
Toyota Prius	$24,200	-	571 mi	-	47.6	5 min
Chevrolet PHEV	$25,185	38 mi	342 mi	2.9	37.1	5 min
Nissan BEV	$22,065	84 mi	-	3.3	-	30 min
BMW 740i	$74,000		464 mi		22.2	5 min
Tesla BEV	$59,900	208 mi	-	2.9	-	40 min

Prices are MSRP deducting government tax incentives and tax rebates.
Refuel time is time to fill up the gas tank (PHEVs); for BEV it is to charge 80% of its battery capacity under DC fast charging. The time is a coarse estimate with unit being minutes.

The cumulative distribution function of r_i is shown in Figure A.1. With the currently popular EV fleets, Chevy PHEV's electric range can satisfy about 60% of the daily trips. BEV models Nissan and Tesla Model S's electric range cover 90% and 99% of daily trips respectively.

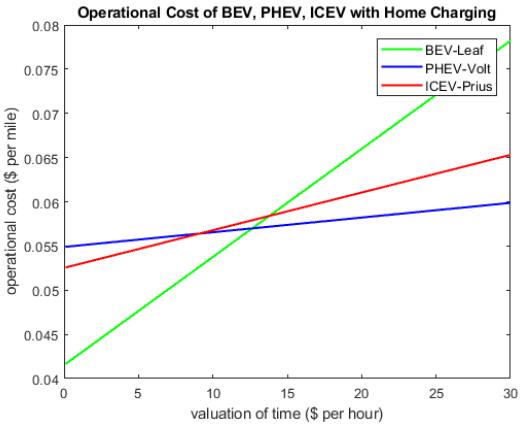
Figure A.1: Distribution of Daily Driven Distance



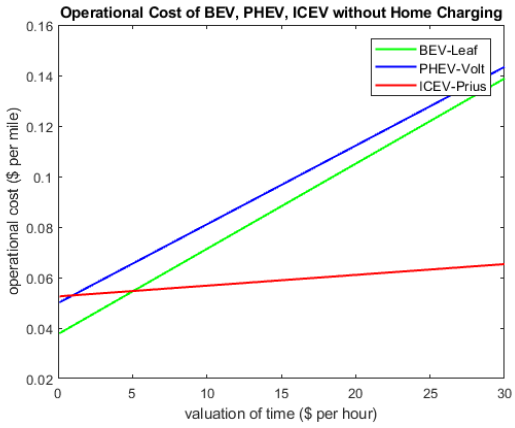
Using baseline assumptions on density of public chargers, I draw graphs (as shown in Figure A.2) that show the impact of evaluation of time on vehicle choice, with cases differentiated by whether or not home charging is available.

[1] PHEVs are currently absent in the luxurious vehicle class

Figure A.2: Operational Cost of Small Vehicles



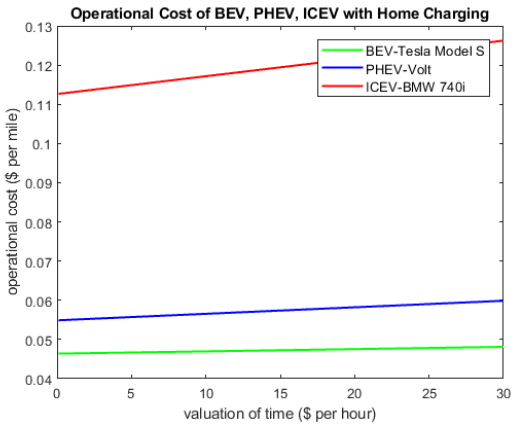
(a) with home charging



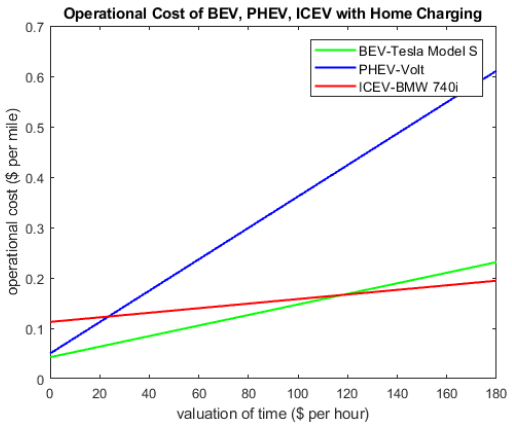
(b) without home charging

In Figure A.3, I compare the operational cost of three popular small vehicle models: Nissan Leaf, Chevrolet Volt, and Toyota Prius, each representing categories of BEV, PHEV, and ICEV. With home charging available, EV owners save time detouring for refueling, therefore EVs' operational cost cuts lower than ICEVs. PHEVs superior in quickly refuel with gasoline after its battery drains therefore they are more favored than BEVs for people with higher valuation of time. Without home charging, however, the time spent on looking for public charging stations hinders EV adoption, leaving ICEVs the dominant choice for people with valuation of an hour higher than 5$.

Figure A.3: Operational Cost of Large Vehicles



(a) with home charging



(b) without home charging

This pair compares popular large vehicle models: Tesla Model S (BEV) and BMW 740i (ICEV). Lacking for a large PHEV model, I include the Chevrolet Volt (PHEV) as a comparison. As Tesla's luxury cars have a long electric range that greatly reduce possibility of running out of battery, people with home charging will always find Tesla's operational cost lowest. For people without access to home charging, those who value an hour of time greater than 120$ will find BMW 740i, which serves as Tesla's luxury counterpart,

occur lower operational cost.

This section considers a baseline scenario where density of public charging stations is fixed and representative vehicle models from category of BEV, PHEV, and ICEV are compared. One can intuitively find that EV range and accessibility to home charging can make a big difference on EV's operational cost.

Appendix B

Appendix for Chapter 3

B.1 Glossary of Variables in Chapter 3

B.2 Probability Distributions and Random Variables

Our work makes extensive use of probability distributions. At times we calibrate our distributions to real-world data such as mean and median income. For a random variable x that is distributed log-normally with location and shape parameter μ and σ, it is straight forward to back out these parameters from the observational mean ($\bar{x}$) and median ($\tilde{x}$) with $\tilde{x} < \bar{x}$:

$$\mu = \ln(\tilde{x}) \quad \text{and} \quad \sigma = \sqrt{2\ln(\bar{x}/\tilde{x})} \tag{B.1}$$

and therefore the standard deviation s_x and coefficient of variation cv_x of x is given by

$$s_x = \bar{x}\sqrt{(\bar{x}/\tilde{x})^2 - 1} \quad \text{and} \quad \text{cv}_x = s_x/\bar{x} = \sqrt{(\bar{x}/\tilde{x})^2 - 1} \tag{B.2}$$

We use a similar method to calibrate the Gumbel (extreme value type I) distribution. Using the Euler-Mascheroni constant $E = 0.57721566490153286$ and the constant $\ln(\ln(2)) = -0.366513$, it follows that

$$\sigma = (\bar{x} - \tilde{x})/(E + \ln(\ln(2)) \quad \text{and} \quad \mu = \bar{x} - \sigma E \tag{B.3}$$

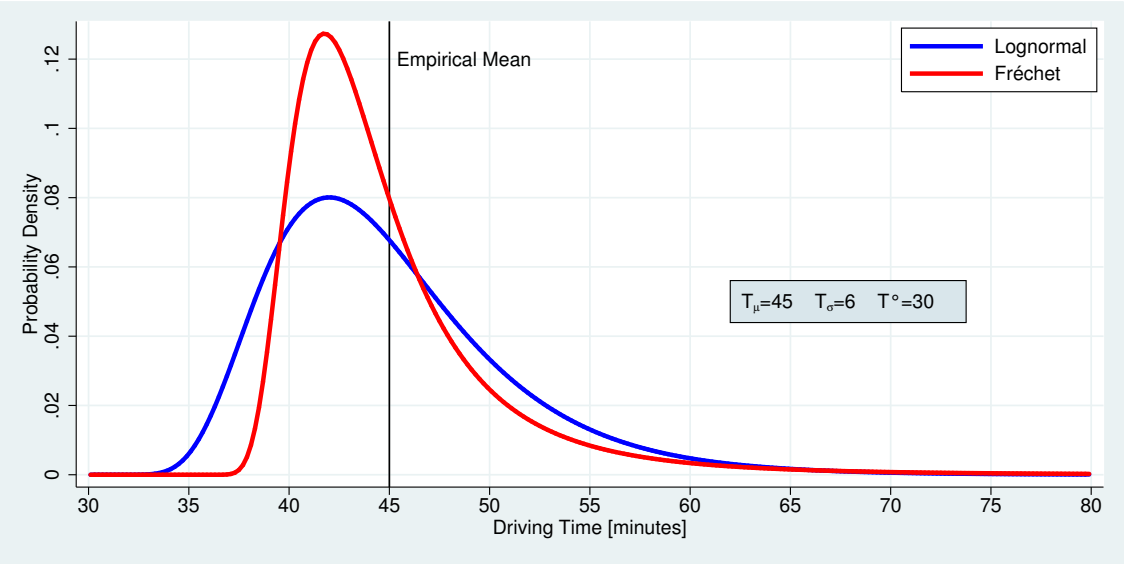
and the standard deviation as $s_x = \sigma \cdot \pi/\sqrt{6} = 1.28255 \cdot \sigma$.

Figure B.1 below shows the shapes of the Fréchet and Log-normal distributions for comparison for identical empirical mean T_μ and standard deviation T_σ, as well as left boundary T°. The Fréchet distribution has a more compact shape than the lognormal distribution for the same moments.

Table B.1: Glossary of Variables in Chapter 3

Variable	Description
ψ	Motorist's schedule inflexibility
T	Travel time
T^*	Optimal time to depart before desired arrival time
T°	Distributional parameter: minimum travel time
S	Distributional parameter: scale parameter
a	Distributional parameter: shape parameter
$\bar{T}$	Mean of travel time
σ^2	Variance of travel time
$\tilde{v}$	Road congestion: traffic volume divided by road capacity
T_μ	Observed mean of travel time
T_σ	Observed standard deviation of travel time
T_ρ	Coefficient of variation of travel time, $T_\rho \equiv \frac{T_\sigma}{T_\mu - T^\circ}$
μ	Parameter in the shifted log-normal distribution
σ	Parameter in the shifted log-normal distribution
T_Δ	Measure of travel time mark-up, $T_\Delta \equiv \frac{T^* - T_\mu}{T_\mu - T^\circ}$
T_Λ	Measure of travel time mark-up, $T_\Lambda \equiv \frac{T^* - T_\mu}{T_\mu}$
T_ν	Ordinary coefficient of variation, $T_\nu \equiv T_\sigma / T_\mu$
V	Travel velocity
L	Length of the road
v	Traffic volume
c	Road capacity
$T^\odot$	Uncongested travel time
ξ	Curvature coefficient
V°	Maximum possible velocity
$V^\odot$	Uncongested traffic velocity
$\boldsymbol{\tau}$	Optimal link tolls
ω	Individual time cost
z	Cost of not travelling
$T^\varnothing$	Maximum duration that a commuter is willing to spend on a trip
$\pi(\boldsymbol{\tau})$	Fraction of flexible commuters that choose to drive after seeing posted tolls
α	Fraction of inflexible commuters
n	Number of potential commuters
η	Travel time elasticity, $\eta \equiv \frac{\mathrm{d}(T - T^\circ)}{\mathrm{d}v} \frac{v}{T - T^\circ}$
$\bar{\omega}$	Average time cost of flexible commuters
ω^*	Average time cost of inflexible commuters
φ	Time cost ratio

Figure B.1: Comparison of Fréchet and Lognormal Distribution



B.3 Data Description

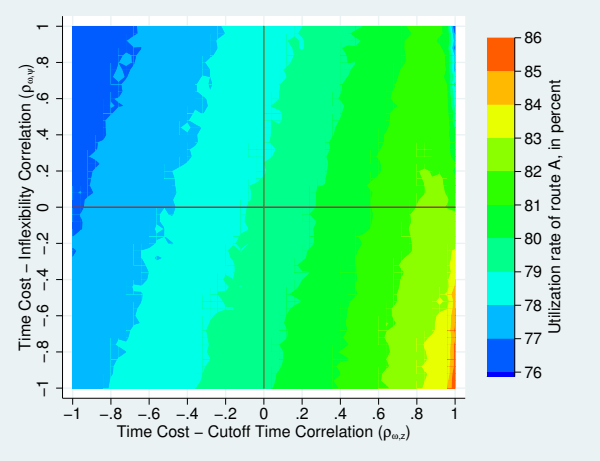
Table B.2: Census Tracts in Vancouver Census Metropolitan Area (2016)

Municipality	#
Burnaby	42
Coquitlam	25
Delta	19
Greater Vancouver A	5
Langley	30
Maple Ridge	15
New Westminster	13
North Vancouver	27
Pitt Meadows	4
Port Coquitlam	9
Port Moody	8
Richmond	39
Surrey	95
Vancouver	117
West Vancouver	8
White Rock	5
All Other	21
Total	482

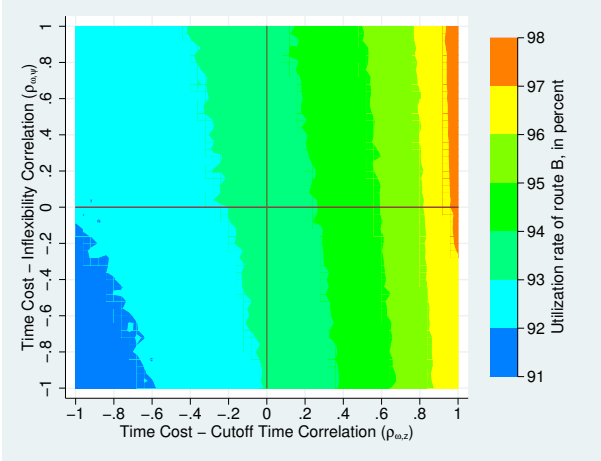
B.4 Additional Figures and Tables

Figure B.2: Simulation Results: Correlation Structure

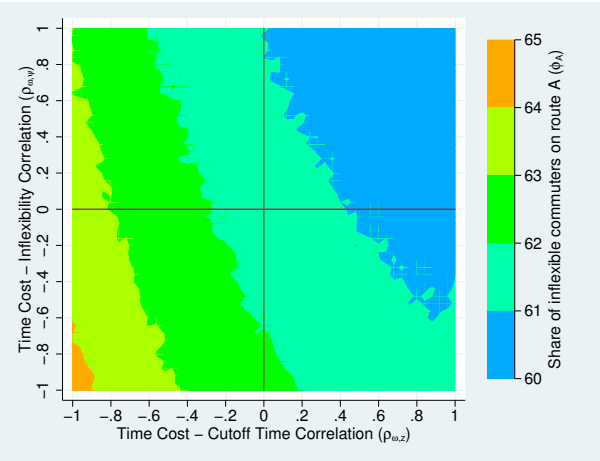
Panel A



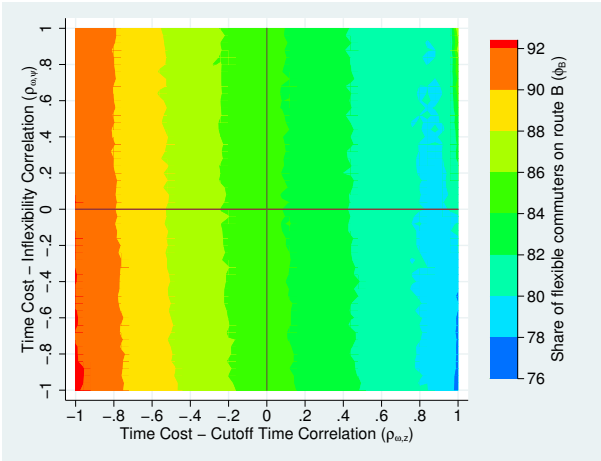
Panel B



Panel C



Panel D



The four figures above show the results of the $51 \times 51 \times 100$ simulation runs for the correlation space for $\rho_{\omega,\psi}$ and $\rho_{\omega,z}$, the correlation between time cost and inflexibility for the inflexible commuters and the correlation between time cost and cutoff time (reservation price) for the flexible commuters. Panels and A and B show the utilization rates of routes A and B (tolled and untolled, respectively), and panels C and D show the share of inflexible commuters taking route A and the share of flexible commuters taking route B, respectively.

Table B.3: Distributional Outcomes: Average Time Costs by Group in the 'Lexus Lane' single-toll scenario with symmetric roads

Parameters			Groups		Optimal Toll					Classic Toll				
α	ψ	φ	$\bar{\omega}^I$	$\bar{\omega}^F$	ω_A^I	ω_B^I	ω_A^F	ω_B^F	$\omega_\varnothing^F$	ω_A^I	ω_B^I	ω_A^F	ω_B^F	$\omega_\varnothing^F$
L	L	L	11.72	14.65	16.48	4.58	17.56	6.87	32.30	16.42	4.56	17.49	6.84	32.60
L	L	M	13.10	13.10	17.40	4.66	16.62	6.54	28.51	17.31	4.61	16.50	6.50	28.93
L	L	H	14.65	11.72	18.46	4.71	15.78	6.20	25.06	18.34	4.65	15.62	6.15	25.59
L	M	L	11.72	14.65	16.02	4.39	18.01	7.07	32.12	15.94	4.35	17.89	7.02	32.57
L	M	M	13.10	13.10	16.97	4.45	17.06	6.72	28.30	16.85	4.39	16.89	6.66	28.89
L	M	H	14.65	11.72	18.11	4.51	16.28	6.36	24.89	17.95	4.42	16.03	6.29	25.58
L	H	L	11.72	14.65	15.66	4.21	18.46	7.26	31.88	15.55	4.16	18.30	7.20	32.47
L	H	M	13.10	13.10	16.63	4.27	17.53	6.89	28.07	16.49	4.19	17.29	6.81	28.83
L	H	H	14.65	11.72	17.76	4.32	16.78	6.51	24.62	17.57	4.21	16.42	6.43	25.48
M	L	L	11.72	14.65	17.09	4.85	18.11	7.20	30.73	17.03	4.82	18.03	7.17	31.10
M	L	M	13.10	13.10	18.40	5.12	17.80	7.01	27.02	18.31	5.08	17.68	6.99	27.46
M	L	H	14.65	11.72	20.00	5.46	17.82	6.77	23.73	19.87	5.40	17.61	6.76	24.22
M	M	L	11.72	14.65	16.80	4.72	18.86	7.50	30.53	16.72	4.69	18.75	7.46	31.03
M	M	M	13.10	13.10	18.15	5.01	18.69	7.27	26.81	18.03	4.95	18.50	7.25	27.41
M	M	H	14.65	11.72	19.79	5.37	18.89	6.98	23.62	19.60	5.28	18.57	6.98	24.21
M	H	L	11.72	14.65	16.53	4.61	19.67	7.77	30.31	16.41	4.56	19.49	7.73	31.01
M	H	M	13.10	13.10	17.93	4.91	19.63	7.51	26.58	17.76	4.83	19.32	7.48	27.37
M	H	H	14.65	11.72	19.64	5.30	20.08	7.15	23.49	19.39	5.18	19.61	7.16	24.20
H	L	L	11.72	14.65	17.75	5.11	18.57	7.47	28.83	17.71	5.10	18.53	7.46	29.11
H	L	M	13.10	13.10	19.53	5.60	19.06	7.35	25.36	19.48	5.57	18.99	7.35	25.65
H	L	H	14.65	11.72	21.66	6.18	19.95	7.03	22.54	21.60	6.15	19.85	7.06	22.76
H	M	L	11.72	14.65	17.60	5.06	19.73	7.82	28.55	17.54	5.04	19.65	7.82	28.94
H	M	M	13.10	13.10	19.45	5.56	20.42	7.59	25.26	19.38	5.53	20.30	7.62	25.63
H	M	H	14.65	11.72	21.62	6.16	21.59	7.17	22.45	21.54	6.12	21.45	7.22	22.73
H	H	L	11.72	14.65	17.50	5.02	20.95	8.12	28.40	17.42	4.99	20.84	8.13	28.89
H	H	M	13.10	13.10	19.38	5.53	21.90	7.81	25.14	19.28	5.49	21.73	7.85	25.56
H	H	H	14.65	11.72	21.60	6.15	23.41	7.27	22.44	21.50	6.11	23.21	7.33	22.75

Note: This table shows the average time cost by commuter groups corresponding to the welfare analysis in table 3.13, the 'Lexus Lane' scenario. The group means are shown first, followed by the average time cost for the five groups of commuters under an 'optimal' and 'classic' toll regime.

Table B.4: Distributional Outcomes: Average Time Costs by Group in the dual-toll scenario with asymmetric roads

Parameters			Groups		Optimal Toll					Classic Toll				
α	ψ	φ	$\bar{\omega}^I$	$\bar{\omega}^F$	ω_A^I	ω_B^I	ω_A^F	ω_B^F	$\omega_\varnothing^F$	ω_A^I	ω_B^I	ω_A^F	ω_B^F	$\omega_\varnothing^F$
L	L	L	11.72	14.65	11.72			9.84	25.90	11.72			9.84	25.90
L	L	M	13.10	13.10	13.10			8.81	23.15	13.10			8.81	23.15
L	L	H	14.65	11.72	14.66			7.88	20.73	14.66			7.88	20.73
L	M	L	11.72	14.65	11.70			9.85	25.89	11.70			9.85	25.89
L	M	M	13.10	13.10	13.09			8.81	23.18	13.09			8.81	23.18
L	M	H	14.65	11.72	14.63			7.88	20.70	14.63			7.88	20.70
L	H	L	11.72	14.65	11.72			9.84	25.89	11.72			9.84	25.89
L	H	M	13.10	13.10	13.11			8.80	23.19	13.11			8.80	23.19
L	H	H	14.65	11.72	14.66			7.88	20.71	14.66			7.88	20.71
M	L	L	11.72	14.65	13.63	7.36		9.26	21.80	13.53	7.28		9.40	22.93
M	L	M	13.10	13.10	15.66	7.58		8.19	18.70	15.44	7.58		8.32	19.82
M	L	H	14.65	11.72	16.22	11.52		7.24	16.02	15.87	11.93		7.36	17.13
M	M	L	11.72	14.65	12.35	10.18		9.26	21.33	12.29	10.19		9.44	22.86
M	M	M	13.10	13.10	13.33	12.61		8.18	18.27	13.15	13.01		8.35	19.83
M	M	H	14.65	11.72	14.59	14.79		7.23	15.63	14.25	15.62		7.40	17.09
M	H	L	11.72	14.65	11.09	13.24		9.24	20.85	11.05	13.56		9.48	22.90
M	H	M	13.10	13.10	11.98	15.64		8.14	17.87	11.83	16.44		8.37	19.82
M	H	H	14.65	11.72	13.33	17.41		7.21	15.28	13.01	18.75		7.43	17.08
H	L	L	11.72	14.65	12.76	10.47		8.42	17.55	12.65	10.54		8.63	19.17
H	L	M	13.10	13.10	14.12	11.93		7.47	15.03	13.92	12.10		7.64	16.43
H	L	H	14.65	11.72	15.72	13.45		6.60	12.90	15.41	13.73		6.74	14.05
H	M	L	11.72	14.65	12.03	11.34		8.42	17.10	11.93	11.44		8.67	19.08
H	M	M	13.10	13.10	13.32	12.84		7.44	14.66	13.12	13.07		7.66	16.34
H	M	H	14.65	11.72	14.81	14.45		6.61	12.67	14.50	14.81		6.78	14.03
H	H	L	11.72	14.65	11.34	12.17		8.41	16.77	11.25	12.34		8.70	19.09
H	H	M	13.10	13.10	12.56	13.75		7.42	14.41	12.36	14.07		7.69	16.33
H	H	H	14.65	11.72	13.97	15.44		6.59	12.44	13.67	15.90		6.78	13.99

Note: This table shows the average time cost by commuter groups corresponding to the welfare analysis in table 3.14, the 'Asymmetry' scenario. The group means are shown first, followed by the average time cost for the five groups of commuters under an 'optimal' and 'classic' toll regime.