Subsidizing Electric Vehicles among Heterogeneous Consumers: Does Vehicle Holding Matter?

Chao Wang*

November 5, 2023

[click here for the latest version]

Abstract

In the United States, most households can access one or more vehicles. I study the impact of household vehicle holding on a household's vehicle purchasing decisions and the benefit of taking into account such an impact on EV subsidy policy. I develop a structural model of household automobile decisions while allowing for the willingness to pay for the new vehicle to vary with the current vehicle holdings. Combining market sales data and survey information in California, I identify and estimate the preference heterogeneity induced by vehicle holding and quantify its welfare implications. Households without vehicles exhibit a higher vehicle purchase propensity, while households with EVs are more likely to acquire additional EVs compared to households with gasoline vehicles. Counterfactual simulations indicate that redistributing subsidy amounts across households with different vehicle holdings could increase EV sales by 8% without augmenting subsidy expenditure, at the cost of a 0.1% reduction in consumer surplus. In contrast, achieving the same level of EV sales under the current subsidy scheme would require an additional $81.6 million in the government's subsidy budget.

Keywords: electric vehicle adoption, household heterogeneity, subsidy policy
JEL Codes: L52, L62, L92

*Department of Economics, Indiana University Bloomington. Email: Wang: cw102@iu.edu. I am grateful to Ruli Xiao, Stefan Weiergraeber, Andrew Butters, Gustavo Torrens, Haizhen Lin, Boyoung Seo, Austin Knies, Cesar Salinas, Alireza Marahel, Yanhao Wang, Brad Larsen, Yunmi Kong and audiences in 2022/2023 Midwest Econometrics Group Conference, 2022 Missouri Valley Economic Association, Indiana University Applied Micro Brownbag, Indiana University Business Economics and Public Policy Brownbag for their valuable comments and feedback. The views expressed here and all errors are my own.

1 Introduction

As of 2020, the US Environmental Protection Agency (EPA) reported that the transportation sector constituted the largest share (27%) of total greenhouse gas emissions in the United States, with light-duty vehicles contributing the majority (57%) of emissions within this sector.[1] Compared to conventional fossil fuels, electricity stands out as a cleaner and more environmentally friendly energy source, emitting fewer greenhouse gases. Thus, replacing conventional energy-powered vehicles with electric vehicles (EVs) emerges as a pivotal strategy for mitigating greenhouse gas emissions in the transportation sector.

Many countries have implemented subsidies and tax rebates to incentivize the adoption of EVs.[2] To best promote EV adoption with a given budget, it is imperative for the government to gain insights into the primary demographic of EV purchasers. In the US, 59% households were multi-car owners in 2020.[3] Household's vehicle holding situation potentially affects their rankings between gasoline vehicles (GVs) and EVs, given other car characteristics the same. For instance, the ratio of EVs to GVs among households with a single vehicle differs from those with multiple vehicles, and a significant proportion of EV consumers own one or more GVs (Davis (2021)). In practice, the current EV subsidy policy does not tailor the rebate to accommodate the composition of a household's current vehicle holdings: whether the household owns a car and, if so, its fuel type. Ignoring such heterogeneity could result in significant misallocation of subsidies, given that different households exhibit varying levels of price sensitivities.

This paper studies the impact of household vehicle holding on a household's vehicle purchasing decisions and the benefit of taking into account such an impact on EV subsidy policy. Firstly, I examine the household vehicle preferences by leveraging consumer survey

[1] https://www.epa.gov/greenvehicles/fast-facts-transportation-greenhouse-gas-emissions

[2] For example, the Plug-in vehicle grant in the UK and the Green vehicle purchasing promotion measures in Japan. In the US, Federal government spent $2.5 billion on EV in 2022. As for state government, California spent about 525 million on state-wide vehicle subsidies in 2022.Other policies to reduce emissions are also considered in the car sector, such as limits on CO_2 emissions, fuel efficiency standards, and green license plates.

[3] https://www.fool.com/the-ascent/research/car-ownership-statistics/

data, revealing significant disparities in these preferences contingent upon the household's existing vehicle holdings. Then, I develop a structural model of oligopoly competition in the automobile industry with differentiated products, allowing the household preferences related to household vehicle holdings. In contrast to conventional demand models that do not consider vehicle holdings, my model introduces a more intricate representation of the diversity in household price sensitivities. Finally, a refined subsidy scheme is proposed with the aim of fostering EV adoption. Counterfactual analysis outcomes emphasize that the redistribution of subsidies among households with differing vehicle holding profiles can efficiently stimulate EV sales without necessitating an increase in subsidy expenditure.

I start by describing household vehicle holdings and choices between GV and EV from the California Vehicle Survey. Households possessing vehicles exhibit a greater inclination to opt for EVs over GVs, and households with pre-existing EVs display a higher inclination to acquire additional EVs in comparison to households owning GVs. Then, I put the households that do not buy vehicles in the analysis. Households with vehicles have a lower tendency to buy vehicles (EV or GV) than households without vehicles. Result changes from adding one more option (not buying a vehicle) underscore the presence of discernible heterogeneity across households with distinct vehicle holdings.

While the descriptions are essential in understanding household heterogeneity regarding vehicle holdings, I develop and estimate a structural model to understand how it enters household vehicle choices across differentiated car model substitution patterns. The demand model estimation is based on Core-based statistical area (CBSA) level market share data for new cars in California between 2014 and 2016. In the model, the decision to purchase a new GV or EV hinges upon the household's existing vehicle holdings. Identifying the parameters governing this preference heterogeneity induced by vehicle holding presents a challenge, as the aggregate market share data do not have enough variation in purchase probability conditional on household vehicle holdings. To solve this problem, I complement the aggregate market share data with the California Vehicle Survey data that give the responding households vehicle information. I construct the micro-moments from the survey data based on vehicle holdings and integrate them with the macro-moments constructed from the market shares. The survey data give variations of purchasse probabilities conditional on vehicle holdings to help identify the respective preference heterogeneity terms in household demand for new vehicles.

I take these model estimates to study the effectiveness of counterfactual interventions aimed at promoting EV sales. This evaluation entails two simulations. I first compare the market outcome and welfare under the models with and without incorporation of preference heterogeneity induced by vehicle holding, in a scenario where subsidies for EV are repealed. While the general performance remains consistent between the two models, discernible variations emerge in the predictions for different household segments. In particular, households without vehicles purchase much more (including GV and EV) when the model accounts for preference heterogeneity induced by vehicle holding, a trend that aligns with the demand parameters. This highlights the critical impact of incorporating preference heterogeneity induced by vehicle holding in accurately predicting household responses to subsidy incentives.

In the second simulation, I restructured the EV subsidy subsidy framework to align with household vehicle ownership, ensuring that households with distinct vehicle holdings receive varying subsidy levels. While maintaining a comparable subsidy expenditure, I observed that this adjustment resulted in incremental improvements in both total welfare and EV sales, albeit to a limited extent. On the other hand, reallocating subsidies from households with GV to subsidies with no vehicle can yield an 8% increase in EV sales, albeit with a corresponding 0.1% reduction in consumer surplus. This outcome mirrors the scenario akin to third-degree price discrimination, where different segments of households are subject to distinct pricing structures, determined by government subsidies in this context. These segments remain isolated from one another, causing a shift of consumer surplus to producer surplus and, correspondingly, an impact on subsidy expenditure. If the government keeps the current subsidy scheme, an additional $81.6 million would be required to achieve the same 8% increment in sales.

This paper makes several contributions to the literature. First, I contribute to the empirical literature on substitution between products within households. Much of the work focuses on other products bought in combinations within the household (see Gentzkow (2007), Wakamori (2015), Archsmith et al. (2020)). I contribute to the literature by analyzing the context of choices between EV and GV. Moreover, I do not assume households choosing products simultaneously, approaching to the real decisions in automobile choices. The few recent studies analyze the total cost of owning an EV, including driving habit changes (Jakobsson et al. (2016), Karlsson (2017), Abotalebi, Scott, and Ferguson (2019)). Their calculations reveal that multi-car households adapt to EVs faster than single-car households. I complement to the research by explicitly estimating a structural model incorporating the household preference heterogeneity induced by vehicle holdings and analyzing the welfare implications.

Second, my work relates to recent literature studying the demand for EVs and the role of government policies. The research focuses on estimating the effectiveness of policies in EV markets from different aspects, including income heterogeneity (Muehlegger and Rapson (2018), Ku and Graham (2022), Hardman et al. (2021)), network effects (S. Li et al. (2017), Zhou and S. Li (2018), J. Li (2019), Springel (2021)), and non-financial incentives (Langbroek, Franklin, and Susilo (2016), Ma, Xu, and Fan (2019), Wang, Pan, and Zheng (2017), Hao (2022)). The results support that the policies play significant roles in promoting EV sales. I complement the literature by proposing a more effective subsidy scheme from the household preference heterogeneity induced by vehicle holdings, and this scheme can be incorporated into other policy schemes mentioned above.

The rest of the paper is organized as follows. Section 2 reviews the related literature. Section 3 describes the data and provides summary statistics. Section 4 provides some reduced-form evidence on household vehicle choice. Section 5 presents the structural model. Section 6 reports the structural estimation results. Section 7 provides two counterfactual experiments on policy design, and section 8 concludes.

2 Literature Review

The body of literature pertaining to the estimation of demand for differentiated automobiles spans the 1990s. S. Berry, Levinsohn, and Pakes (1995) (henceforth BLP) provides a seminal tool allowing the substitution patterns to reflect the consumer heterogeneity in tastes for observed product characteristics using aggregate level sales data. Petrin (2002) improves the BLP model with micro-moments derived from consumer-level data, effectively estimating impact of introducing minivans. Beresteanu and S. Li (2011) also employs both aggregate-level sales data and household-level data to estimate demand for hybrid vehicles using the US automobile market, focusing on the impact of gasoline prices and income tax incentives. J. Li (2019) estimates demand for electric vehicles and the impact of uniforming charging station modes. Following Petrin (2002) method, I complement the existing literature by combining aggregate market sales data and household data to estimate demand for electric vehicles, focusing on the household preference heterogeneity induced by vehicle holding rarely addressed in the literature.

The second strand of literature concentrates on examining the substitution pattern between products within households. Gentzkow (2007) analyzes the impact of online newspapers on print newspapers and employs a structure model that allows households to choose products jointly. Wakamori (2015) takes a model akin to Gentzkow (2007) to investigate the consequences of introducing Kei-car (one kind of small-sized car) in Japan, focusing on the size complementarity between cars within a household. There are also studies providing reduced-form evidence. For example, Archsmith et al. (2020) provide insights into households' preferences for combinations of high and low gasoline-per-mile vehicles. I contribute to this strand of research by analyzing the relationship between types of vehicle engine (gasoline or electric) within a household and comparing the behavior of households with and without vehicles. There are two reasons why I do not adopt the joint estimation method proposed by Gentzkow (2007) on automobile decisions in the US market. First, unlike the simultaneous consumption of online and printed newspapers, households rarely buy two or more new vehicles simultaneously, as new vehicle models typically have a two-year lifecycle before newer versions replace them. Second, extending the joint decision period to, for instance, five years, would introduce non-independent and identically distributed (*i.i.d.*) unobserved shocks among different bundles of vehicles. In the market with a five-year decision period, where options like the Mazda 3 (2015) and Mazda 3 (2011) coexist, it becomes challenging to assert that the preferences for a bundle such as the Toyota Prius (2015) and Mazda 3 (2015) are independent of the bundle comprising the Toyota Prius (2015) and Mazda 3 (2011), given the substantial similarity between the two Mazda 3 models. Such interdependence would considerably complicate the estimation process.

As for the electric vehicle market, Jakobsson et al. (2016) study whether an electric vehicle is feasible as the second car in multi-car households. Leveraging data from household travel surveys in Germany and Sweden, along with daily driving GPS data, they estimate the time required for households to adapt to electric vehicles. They find that multi-car households take less time to adapt to EVs as the second cars. Abotalebi, Scott, and Ferguson (2019) apply similar methods using Canada data, and they discuss possible policy advice, including better knowledge of EVs, marketing more to households with longer annual miles, and setting financial incentives. Karlsson (2017) further calculates the value and implication of replacing BEV as a second vehicle in a two-car household and finds that the flexibility within the car portfolio in a two-car household valuing about $6,000 at an early stage when EV came out. To augment this body of literature, I delve into the role of household preference heterogeneity induced by vehicle holding in conjunction with various socio-demographic variables. Building upon this analysis of heterogeneous demand, I further quantify the policy implications of promoting electric vehicle purchases.

The third strand of literature focuses on the subsidy policy in electric vehicle adoption, including financial and non-financial incentives. For policies focusing on financial incentives, most research discusses the cost and benefit of different consumer subsidy schemes. Existing literature had provided evidence of the positive relationship between financial incentives such as tax rebate and electric vehicle adoption in the early stage when electric vehicle was introduced (Wee, Coffman, and La Croix (2018), Zambrano-Gutiérrez et al. (2018)). Beresteanu and S. Li (2011) examine different federal support schemes that encourage hybrid vehicle adoption and find that a flat rebate scheme that provides equal subsidy to the same model could encourage hybrid sales more than an income tax subsidy. In contrast, Muehlegger and Rapson (2018) focus on an electric vehicle subsidy policy focusing on California's low- and middle-income households, using the policy as a quasi-experiment to analyze the policy effect. They find that the elasticity for EVs is -3.3, and there is still $12-$18 billion in total subsidies required to meet the goal of having 1.5 million EVs on the road by 2025. Ku and Graham (2022) compare the cost and benefit of California's electric vehicle rebate program, and they find that the cost distribution is slightly regressive. However, the benefit of the rebate is highly regressive, and overall net financial impacts are regressive. Hardman et al. (2021) synthesize research about the transition to EV and investigate policy implications to address equity issues in the EV market. Hao (2022) investigates the effect of driving road restriction policies on EV demand and the role of car ownership. I complement the research by quantifying the implications of electric vehicle subsidies from the perspective of preference heterogeneity induced by vehicle holding rather than income heterogeneity. The closest paper to mine is Hao (2022), which also studies the heterogenous effect of car ownership. The difference between this paper and Hao (2022) is that she ignores the complementarity and substitution of vehicle types concerning both vehicle holding and the decision to purchase. I examine the complementarity and substitution pattern between EV and GV for households with vehicles, which is pivotal in assessing the effects of new products on existing ones.

Besides allocating subsidies to consumers, another rising branch in electric vehicle subsidy research focuses on the network effect in charging station infrastructure. Since electric vehicles rely on charging infrastructure to provide energy, especially on long trips, the number of charging stations and EVs on the road are positively related. Zhou and S. Li (2018) study the EV subsidy policy from the perspective of indirect network effects of EVs. They find that subsidizing EVs to pass the critical mass is essential in achieving a high EV adoption rate. S. Li et al. (2017) quantifies the role of indirect network effects on the EV market and government subsidy implication. Springel (2021) finds EV purchases and charging stations respond positively to each other, and subsidizing charging stations is more cost-effective to promote EV adoption. J. Li (2019) discusses the impact of uniforming the charging mode of fast charging stations in promoting the adoption of electric vehicles. To account for the influence of charging stations on electric vehicle demand, I also incorporate the number of charging stations on the market into the consumer utility function. The outcomes align with the existing literature, reinforcing the positive relationship between an increased number of charging stations and higher demand for EVs.

Another body of literature delves into the effiveness of non-financial incentives on EV adoption, including privileges on parking or access to bus lanes (Langbroek, Franklin, and Susilo (2016)), exempting driving or purchase restriction (Ma, Xu, and Fan (2019),Wang, Tang, and Pan (2017), Hao (2022)). The studies indicate that non-financial incentives could wield a more significant impact than financial incentives in promoting EV adoption, and they also point out the importance of consumer heterogeneity (Wang, Pan, and Zheng (2017), Mersky et al. (2016)). Most of the research discussing non-financial incentives uses survey data from consumers. I complement the research by providing a framework including consumer heterogeneity in EV policy analysis by combining the survey data with market share data to analyze the welfare impact. The framework provided in this paper is readily applicable in analyzing non-financial incentives in electric vehicles.

3 Institutional Background and Data

3.1 Overview of the EV Market and Policy Background

EVs are road vehicles powered by batteries that can be recharged by plugging into the electric grid. Currently, there are two types of EVs: battery electric vehicles (BEVs), which are powered exclusively through electricity, and plug-in hybrid electric vehicles (PHEVs), which use the electric motor as the primary power source and the internal combustion engine as a backup.[4] Compared with traditional internal combustion engine vehicles (ICEV, or gasoline vehicle/GV), EVs have a higher manufacturer cost due to their large battery and new features like regenerative braking, engine stop-start, and novel transmission system (Palmer et al. (2018)). In this paper, I focus on passenger vehicles including cars and SUVs. Table 1 presents the sales and sales-weighted average prices of EV and GV during 2014-2016 in California. The average price of the EV was \$47.6K, while the GV counterpart was \$32.4K. On the other hand, EVs usually have lower running costs from cheaper annual fuel costs, taxes, and maintenance. The running cost can partly offset the EV price premium, but EVs still need subsidies to fill the gap (Palmer et al. (2018)).

Table 1: summary statistics for GV and EV in California

category	year	sales	avg price($)	mkt share (%)
EV	2014	47,147	40,674	0.356
EV	2015	46,292	46,277	0.350
EV	2016	58,811	54,167	0.444
GV	2014	1,260,672	31,897	9.528
GV	2015	1,342,444	32,284	10.146
GV	2016	1,315,012	32,989	9.939

In addition to the higher price of EVs, the other reason the government should intervene in the EV market is out of externalities. When the producers or consumers do not consider all social costs or benefits when they make decisions, the market can not produce the socially optimal amount of products. Rapson and Muehlegger (2023) summarizes two main types of externalities in the EV market: one type is the externality created by the operation of EVs (the "intensive" margin), like carbon emissions or local pollution. The other externality arises from the production or stock of EVs on the road (the "extensive" margin), such as learning-by-doing and network effect.

To promote electric vehicle adoption, governments worldwide set ambitious targets for

[4] EVs are different from fuel cell electric vehicle (FCEV) (*e.g.* Toyota Mirai) and traditional hybrid electric vehicle (HEV) (*e.g.* Toyota Prius) since FCEV and HEV can not be recharged in the electric grid.

adopting electric vehicles or phasing out GV gradually. In the US, California is at the forefront of this commitment to put 1.5 million "Zero Emissions Vehicles" (ZEVs, the vehicles that do not emit exhaust gas or other pollutants from the onboard source of power, including EVs and other human-powered vehicles) on the road by 2025 and 5 million by 2030[5], and effectively bans sales of new GVs by 2035[6].

To spread EVs, the US federal government started a tax credit program for PEV purchases in 2009. EVs made after December 31, 2009, are offered non-refundable tax credits (IRS, 2009). The federal tax credit is offered based on the battery size: starting from $2,500 for PEV of 4 kWh or less, the federal subsidy increases $417 for 1 kWh over 4 kWh, up to $7,500 in total. With bigger batteries, BEVs usually get more federal subsidies than PHEVs. Most popular BEV brands, like Tesla models and Chevrolet Bolt, get the full $7,500 federal tax credits. These granted federal tax credits on PEV vehicles will phase out after the manufacturer sells 200,000 EVs in the US. In July 2018, Tesla Inc. was the first manufacturer to pass 200,000 sales, and the entire federal tax credit was available until the end of 2018, with the phase-out beginning in January 2019. In 2023, the 200,000 cap for federal tax credit was lifted. GM and Tesla are eligible for federal tax rebates again. Within this paper's vehicle sales data period, all EV manufacturers are unaffected by the phasing-out policy.

In addition to the federal subsidy policies, some states in the US also provide their state subsidy policies to promote EV purchases. California state subsidy is controlled by the Clean Vehicle Rebate Program (CVRP). The evolution of California state subsidy on the EV market is presented in figure 1. Starting from the end of 2009, maximum standard rebate amounts for FCEV, BEV, and PHEV were $5,000, $5,000, and 0 (PHEV came to market in 2011). Maximum rebate amounts decreased to $2,500 and $1,500 in 2011-2012.

In March 2016, CVRP implemented income caps and increased rebate levels for lower-income consumers. This policy gives different rebate levels for different income consumers.

[5] https://opr.ca.gov/planning/transportation/zev.html

[6] https://www.cnbc.com/2022/01/10/california-proposes-6point1-billion-in-new-incentives-for-electric-vehicles-.html

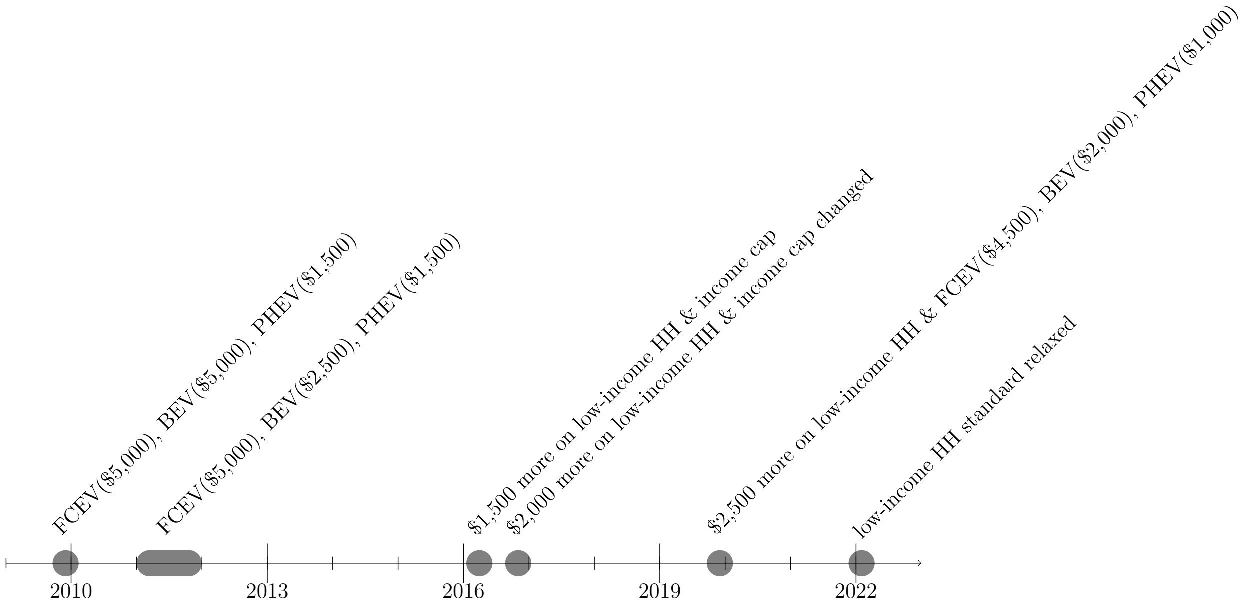


Figure 1: Evolution of CA subsidy on EV market

From March 29, 2016, to October 31, 2016, the rebate for consumers of household income $\leq$ 300% of the federal poverty level increased by $1,500, and consumers with income over $250,000 for a single filer, $340,000 for head-of-household filers, and $500,000 for joint filers are not eligible for rebate. From November 2016 to December 2019, the income cap decreased to $150,000 for single filers, $204,000 for head-of-household filers, and $300,000 for joint filers, and the increase of rebate for lower-income consumers increased from $1,500 to $2,000. After December 2019, the increased rebate for low-income consumers moved to $2,500 more than the standard rebate level, and the standard rebate level decreased to $4,500, $2,000, and $1,000 for FCEV, BEV, and PHEV. Since Feb 2022, the increased rebate expanded consumers from 300 percent of the federal poverty level to 400 percent, lowering the income cap from $150,000 to $135,000. The data in this paper is based on yearly sales of vehicles, and I cannot tell the purchasing time for the survey data, so I do not treat subsidy to vary in household income.

3.2 Data

I mainly use the following datasets: the aggregate vehicle sales data in California during 2014-2016 are registration data from IHS Markit (formerly R.L.Polk). The registrations are collected by the state's Department of Motor Vehicles and reflect the new vehicle purchases. The dataset reports the number of registrations by car model, geographic area, purchase time, and car model defined by vehicle make, model name, model year, and engine type. The registration data are collected at the zip code level, and I aggregate the data into the core-based statistical area (CBSA) to define geographic markets.[7] The panel includes 35 CBSAs over 2014-2016 and thus 105 markets in total. The car sales data are merged into model-level characteristics information from www.teoalida.com (including Manufacturer's Suggested Retail Price (MSRP), horsepower, car classification, and five vehicle size-related variables: length, width, height, wheelbase, and curb weight) and Ward's Automotive Yearbook, as well as the number of charging station information from Department of Energy's Alternative Fuels Data Center (AFDC).

In this paper, the consumer choice sets are defined by all available vehicle models within California in a purchasing year since consumers could travel within the state to buy a new car. Some new models could be sold with zero quantities in some markets, which generates zero market shares. Zero market shares could cause the numerical challenges when applying the inversion step in S. T. Berry (1994) and S. Berry, Levinsohn, and Pakes (1995) since zeros are not applicable in logarithm. To deal with this problem, I take the method similar in J. Li (2019) by shrinking the data toward an empirical Bayesian prior formed from similar markets and bump the zero market shares to some small positive numbers. The detailed algorithm I use to address the zero market share problem is presented in section 9.1. I also drop some luxury models and models sold less than 100 in all markets in a year.[8] After these

[7]A core-based statistical area (CBSA) is a U.S. geographic area defined by the Office of Management and Budget (OMB). CBSA contains the Metropolitan Statistical Area (MSA) and the Micropolitan Statistical Area (MicroSA).

[8]Luxury models here include Bentley, Aston Martin, Mclaren, Lamborghini, Ferrari, and Maserati.

steps, there are 56,385 observations in total, and 537 models are available for a consumer in each market on average.

The demographic information comes from the California Vehicle Survey. Three primary waves of surveys include the electric vehicle: 2013, 2017, and 2019. Since the vehicle sales data are from 2014-2016, I adopt the 2017 survey only for consistency in household behavior and micro-moment matching. 3,614 households took the survey in 2017.[9] Another widely used survey is the National Household Travel Survey (NHTS), which covers a larger number of respondents. There are a few advantages of using California Vehicle Survey data. Firstly, California Vehicle Survey contains all vehicles listed in a household, the purchasing year, vehicle model and model year, vehicle characteristics, and rich participants' demographic information like zip code, household income, and whether the household installed solar panels. Secondly, the vehicle purchasing time is precise for each responding household in California Vehicle Survey, so I can order the vehicle purchasing time within a household and determine the purchasing order for each household, while NHTS does not have a clear purchasing time for all vehicles.

3.3 Household Heterogeneity

Figure 2 shows the frequency of households according to the number of vehicles owned in the sample. Among all households, 14 households did not own any vehicle and did not buy any vehicle in the survey.[10] 1,244 households own one vehicle, 1,636 own two vehicles, and over 600 own three or more vehicles. Multi-car households occupy over 60% of the households in California. These numbers show the vehicle access rate at the survey time. For micro-moment construction, I need to categorize these households further according to whether the households are buying new vehicles or not in the market.

[9] In California Vehicle Survey 2017, no household is observed to purchase a new vehicle in 2017, so I treat all the responding households purchasing the new vehicle in the 2016 market to match the micro-moments. I randomly choose the order if two vehicles are purchased in the same year.

[10] This number is small because the households without vehicle access are under-sampled in the 2017 survey. However, it will not affect the further analysis because it does not equal the number of households

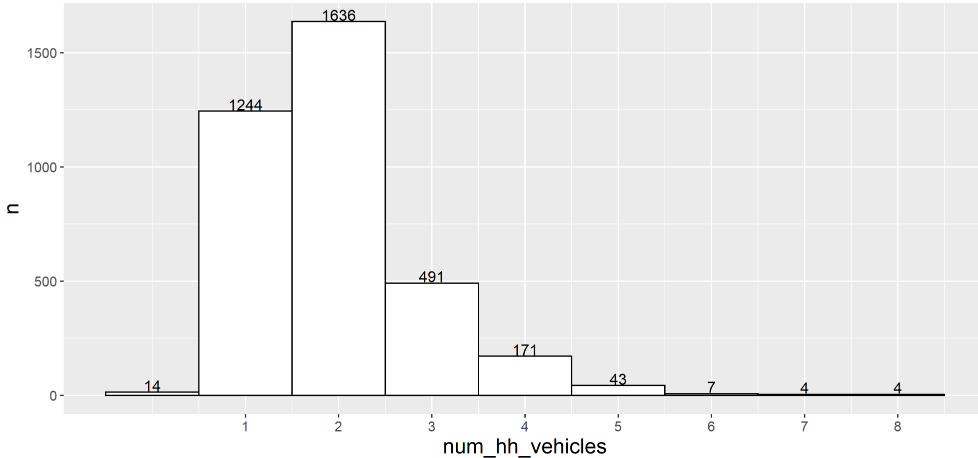


Figure 2: Household distribution according to vehicle numbers owned

Then, I treat all households from the survey in the 2016 market since there are no households making decisions to purchase vehicles in 2017. Households can choose to buy an EV or a GV in the new car market or choose the outside option (which includes buying other new vehicles, buying used vehicles, or not buying anything). I define a vehicle as a new vehicle in the survey if its purchase mileage does not exceed 200 miles and the vehicle model year is the same or above the vehicle purchase year.[11] The outside option is labeled as "not buy" for simplicity.

Table 2 presents the frequency of different actions in 2016 from the survey. Column (2) presents the total sample frequency. About 90% households chose the outside option, and 10% chose to purchase a new vehicle. These numbers are comparable with the new car registration (1.85 to 2.09 million out of 13 million households during 2014-2016, which is about 13.9% to 15.3%. Because of some data mismatching in car characteristics, the vehicle purchasing rate in the survey is lower than the sales data from IHS).[12]. Among the households buying new vehicles, the relative purchasing rate of GV and EV is 8:1.

Table 2 further presents household heterogeneity in columns (3) to (5). Here, households are different in their vehicle holdings: "none" stands for households with no other vehicle

choosing outside options.

[11] Here, households could buy new vehicles of model year 2016 or 2017 in 2016.

[12] https://www.cncda.org/wp-content/uploads/Cal-Covering-4Q-19.pdf

Table 2: Frequency tables given vehicles in hand

(a) Frequency table

	Full Sample	none	EV	GV
not buy	3259	36	242	2981
buy GV	314	73	16	225
buy EV	41	2	8	31
Total	3614	111	266	3237

(b) Conditional Probability (in percentage)

	Full Sample	none	EV	GV
not buy	90.2	32.4	91.0	92.1
buy GV	8.7	65.8	6.0	7.0
buy EV	1.1	1.8	3.0	1.0
Total	100.0	100.0	100.0	100.0

to form the household vehicle portfolio. "GV" stands for households with only gasoline vehicle to form the household vehicle portfolio, and "EV" stands for households with at least one electric vehicle to form the household vehicle portfolio. If a household is going to replace a vehicle by buying a new one in the market, then the replaced vehicle is not counted as a vehicle holding. The household vehicle holding is an exogenous state variable in this paper.[13] There is considerable heterogeneity across households with different vehicle holdings. For households with no vehicle holding, the probability of choosing the outside option is 32.4%, while for households with vehicles, the probability increases to above 90%. Moreover, preferences for EV and GV are also different among households with different vehicle holdings. For households with no vehicle holding, households prefer to buy a GV much more than an EV. For households with at least one EV in hand, the probabilities of buying a new GV and EV are 6% and 3% (the relative ratio is about 2:1). For households with GV in hand, the probabilities of buying a new GV and EV are 7% and 1% (the relative ratio is about 7:1).

Table 3 characterizes the relation between family income and the household vehicle hold-

[13]Since the household survey is cross-sectional, I cannot observe the household vehicle holdings over time, and I cannot tell whether the household will replace the current vehicle holding with a new vehicle or add the new vehicle to the current vehicle holding. Instead, I only focus on how the vehicle still in hand (or the vehicle holding) affects the purchase decision.

ing or household vehicle purchase decision. The upper panel in table 3 is the frequency between income and household vehicle holding. As household income increases, the GV vehicle holding rate decreases from 92.9% to 75.1%, and the EV vehicle holding rate increases from 3.7% to 22.3%. The lower panel in table 3 presents the frequency between income and household vehicle purchase decisions. Middle- and high-income households are more likely to purchase a vehicle than low-income households. The differences between high-income and middle-income groups in purchasing frequency are not large.

Table 3: demographic given vehicles in hand

(a) Income and vehicle holding frequency (in percentage)

	Full Sample	low income	mid income	high income
none	3.1	3.4	2.6	2.6
EV	7.4	3.7	9.7	22.3
GV	89.6	92.9	87.6	75.1
Total	100.0	100.0	100.0	100.0

(b) Income and purchase frequency (in percentage)

	Full Sample	low income	mid income	high income
not buy	90.2	92.8	86.5	85.4
buy GV	8.7	6.6	11.9	11.4
buy EV	1.1	0.6	1.6	3.1
Total	100.0	100.0	100.0	100.0

"Low income" is defined if family annual income is less than $100,000. "Mid income" is defined if family annual income is between $100,000 and $200,000. "High income" is defined if a family's annual income is higher than $200,000.

Table I2 further presents the relation between household vehicle holding and other demographic information, including family size, education, and race. Households with EVs typically have higher income, higher education, and smaller white families. In summary, household vehicle purchase decisions are heterogeneous depending on household characteristics like vehicle holdings and household income.

4 Reduced-form Evidence

In this section, I will investigate the households' vehicle choice and how the vehicle holding would affect vehicle purchase decisions by providing some reduced-form evidence. I run the following linear probability model (LPM) using the survey data and treat all households in the 2016 market as mentioned in section 3:

$$EV_{it} = \beta_1 hold_i + \beta_2 income_i + \beta X_i + \gamma_t + \epsilon_{it} \tag{1}$$

The dependent variable is an indicator variable for buying an electric vehicle. Here, I only focus on comparing EV and GV, so I exclude the households who choose the outside option. The independent variable $hold_i$ is an indicator of whether households hold any vehicle (when the indicator equals 1). Other independent variables include $income_i$ (middle-income and high-income indicators) and other demographic variables X_i (including household family size, education, and ethnicity). γ_t is the market fixed effect.

Table 4 presents the results from the linear probability model. Columns (2) - (4) present that holding a vehicle has a significant positive relation with EV purchase. Household with a vehicle has 11% higher probability of buying an EV than households with no vehicle. The positive impact is still significant after controlling other demographic information (especially the income effect) and the market fixed effect. Column (5) - (7) further decompose the vehicle holding into $hold_{GV}$ and $hold_{EV}$ to investigate the heterogeneity effect within households with vehicle holding. The positive impact is consistent; households with EVs tend to buy EVs more often than households with GV. Table I4 shows the logit model based on the same setup, and the results are consistent with LPM.

4.1 Using Charging Stations as Instruments for Vehicle Holdings

In the previous setup, household vehicle holdings are assumed to be exogenous. One concern is that the exogeneity assumption may not hold since the current vehicle holdings

Table 4: Relation between vehicle purchase and vehicle holding

	Dependent variable:					
	buyEV					
	LPM	LPM	LPM	LPM	LPM	LPM
hold	0.113***	0.122**	0.150***			
	(0.041)	(0.047)	(0.048)			
hold GV				0.094**	0.115**	0.146***
				(0.041)	(0.047)	(0.048)
hold EV				0.307***	0.328***	0.308***
				(0.073)	(0.082)	(0.084)
mid income		0.014	−0.012		0.002	−0.019
		(0.038)	(0.038)		(0.038)	(0.038)
high income		0.077	0.047		0.035	0.014
		(0.054)	(0.055)		(0.055)	(0.057)
family size	No	Yes	Yes	No	Yes	Yes
Education	No	Yes	Yes	No	Yes	Yes
Ethnicity	No	Yes	Yes	No	Yes	Yes
CBSA FE	No	No	Yes	No	No	Yes
Observations	355	355	355	355	355	355
R^2	0.021	0.099	0.226	0.048	0.123	0.239
Adjusted R^2	0.018	0.050	0.119	0.043	0.073	0.130

Note: *p<0.1; **p<0.05; ***p<0.01
Dependent variable: whether to buy electric vehicle. Endow is indicator of whether the household endowed a vehicle. Mid income indicates whether the household annual income is between 100k dollars and 200k dollars. High income indicates whether the household annual income is higher than 200k dollars.

are vehicles purchased before, and household preferences could be consistent across products purchased. For example, the households who purchased EVs could prefer to buy an EV in later purchases. The ideal approach to deal with the exogeneity problem is using random experiments to randomly assign households to hold GV or EV, which is hard to achieve. The other approach is using instrumental variables for vehicle holdings. Here I use the numbers of level two and level three public charging ports for EVs at the time when the vehicle holdings were purchased as the instrument variables for the vehicle holding. Level two and level three charging ports differ according to the charging speed.

As described in Springel (2021) and J. Li (2019), the demand for EV is positively related to the number of available public charging ports due to network effects. For example, the decision for buying GV or EV in 2016 is related to the number of available charging ports in 2016. Suppose the vehicle held was purchased in 2012, then the purchasing decision for this vehicle held is related to the number of charging ports in 2012. On the other hand, the decision in 2016 is not directly related to the number of charging ports in 2012 since the number of charging ports evolved during 2012-2016.

Table 5 presents the results with instrument variables. I only keep the households with vehicle holdings since the households with no car do not have vehicle holding observed (and then no matched charging ports observed). The dependent variable is again whether the household buys EV or GV in the 2016 market. For independent variable, now I only put $hold_{EV}$ showing households holding EV (if $hold_{EV} = 1$) or GV (if $hold_{EV} = 0$). The last column shows the positive relation between vehicle holding type and purchase type from the OLS model, similar to equation 1. The first and second stage IV results are presented in columns (2)-(3). From the first stage, the vehicle holding type to be EV is positively related to the number of charging ports (I take the logarithm of the number of charging ports plus one) of level three. In contrast, the relation between vehicle holding type and level two charging ports is insignificant. The reason for insignificance could be from the relation between level three and level two charging ports. The positive correlation between vehicle holding type and vehicle to buy still holds in the IV specification (although the magnitude is larger than the result from the OLS model). I will keep assuming vehicle holding is exogenous for now and later in the structural model since assuming the vehicle holding is exogenous does not change the sign of relation between types of vehicle holding and types of vehicles to buy.

Table 5: Relation between vehicle purchase and vehicle holding (IV)

	Dependent variable:		
	hold EV	buy EV	
	1st-stage	IV	OLS
hold EV		0.610^{*}	0.146^{*}
		(0.364)	(0.080)
log(station lv3)	0.035^{*}		
	(0.021)		
log(station lv2)	0.011		
	(0.015)		
mid income	0.052	−0.053	−0.025
	(0.037)	(0.057)	(0.047)
high income	0.201^{***}	−0.096	0.004
	(0.051)	(0.109)	(0.067)
HH char	Yes	Yes	Yes
CBSA FE	Yes	Yes	Yes
Observations	280	280	280
R^2	0.279	0.127	0.243
Note:	*p<0.1; **p<0.05; ***p<0.01 Dependent variable: whether to buy electric vehicle. Endow is indicator of whether the household endowed a vehicle. Mid income indicates whether the household annual income is between 100k dollars and 200k dollars. High income indicates whether the household annual income is higher than 200k dollars.		

4.2 Multinomial Logit Results

Table 6 uses a multinomial logit model on the choice of EV and GV purchase to investigate the impact of vehicle holding. The model takes the outside option as the reference choice. Column (2)-(3) takes only the $hold_{GV}$ and $hold_{EV}$ as explanation variables, and other columns contain income and other household characteristics as controlled variables. The dependent variable "GV" means household choosing GV, and "EV" means household choosing EV in the same model. The significant number for vehicle holding means that households with no car have higher propensity to buy a vehicle than households with vehicles.

Table 7 presents the impact of income on vehicle holding using the multinomial logit model. Columns (2) and (3) present middle- and high-income households are more likely to be with a vehicle. Columns (4)-(5) add the market fixed effect (CBSA FE), and the results are robust.

In summary, the empirical findings presented in this section highlight the connection between household vehicle decisions and the composition of their vehicle holdings. While vehicle holdings are correlated with family income, they continue to exert an independent impact on vehicle choices even when controlling for income. Nevertheless, it's crucial to emphasize that a structural model remains essential for the precise assessment of welfare implications and the execution of counterfactual simulations to evaluate the impact of various policy scenarios.

5 Model

In this section, I will build a structural model to quantify the household demand for electric vehicles. The framework follows S. Berry, Levinsohn, and Pakes (1995) and Petrin (2002).

Table 6: Multinomial logit model between vehicle purchase and vehicle holding

	Dependent variable:							
	GV	EV	GV	EV	GV	EV	GV	EV
hold EV	−3.424***	−0.523	−4.028***	−1.211	−4.156***	−1.096	−4.331***	−2.194*
	(0.329)	(0.810)	(0.351)	(0.847)	(0.360)	(0.906)	(0.420)	(1.165)
hold GV	−3.291***	−1.679**	−3.616***	−1.869**	−3.747***	−1.569*	−4.402***	−0.768
	(0.215)	(0.748)	(0.233)	(0.767)	(0.241)	(0.835)	(0.306)	(1.119)
mid income			0.791***	0.931**	0.813***	0.887**	0.853***	0.733*
			(0.146)	(0.392)	(0.148)	(0.399)	(0.156)	(0.436)
high income			0.769***	1.323***	0.876***	1.224**	0.852***	0.875
			(0.219)	(0.467)	(0.222)	(0.480)	(0.234)	(0.552)
family size	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Education	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity	No	No	Yes	Yes	Yes	Yes	Yes	Yes
CBSA FE	No	No	No	No	Yes	Yes	Yes	Yes
Brand FE	No	No	No	No	No	No	Yes	Yes
Akaike Inf. Crit.	2,341.339	2,341.339	2,305.920	2,305.920	2,357.388	2,357.388	2,275.665	2,275.665

Note: *p<0.1; **p<0.05; ***p<0.01

Dependent variable: 'EV' indicates that households buy electric vehicle, and 'GV' indicates that household buy GV. Mid income indicates whether the household annual income is between 100k dollars and 200k dollars. High income indicates whether the household annual income is higher than 200k dollars.

Table 7: Multinomial logit model result between endowed vehicle and demographic information

	Dependent variable:			
	EV	GV	EV	GV
mid income	0.876^{***}	0.020	0.869^{***}	0.039
	(0.279)	(0.237)	(0.285)	(0.241)
high income	1.558^{***}	-0.198	1.516^{***}	-0.163
	(0.411)	(0.380)	(0.419)	(0.386)
family size	Yes	Yes	Yes	Yes
Education	Yes	Yes	Yes	Yes
Ethnicity	Yes	Yes	Yes	Yes
CBSA FE	No	No	Yes	Yes
Akaike Inf. Crit.	2,716.738	2,716.738	2,762.577	2,762.577

Note: *p<0.1; **p<0.05; ***p<0.01
Dependent variable: 'EV' indicates that household is endowed with at least one electric vehicle, and 'GV' indicates that houshold is endowed with only gasoline vehicle. Mid income indicates whether the household annual income is between 100k dollars and 200k dollars. High income indicates whether the household annual income is higher than 200k dollars.

5.1 Demand

Consider a household i with vehicle holding $h_i \in \{\emptyset, GV, EV\}$ is deciding whether to buy a new vehicle $j \in J$, where J is the set of all available vehicles in the market, or choose the outside option $\emptyset$. The household utility from choosing the vehicle j is

$$u_{ij} = \alpha_i log(y_i - p_j) + \gamma_{iG} 1(fuel_j = G) + \gamma_{iE} 1(fuel_j = E) + x_j \beta_i + \xi_j + \epsilon_{ij} \tag{2}$$

where y_i is the household income, p_j is the price that the household pays out of pocket (that is, $p_j = MSRP_j - subsidy_j$, where $MSRP_j$ is the listed price and $subsidy_j$ is the subsidy on product j). α_i is the household marginal utility of income, and households get utility $\alpha_i log(y_i - p_j)$ from the commodity goods. To capture the household preference for GV and EV, $1(fuel_j = G)$ and $1(fuel_j = E)$ are indicators of whether vehicle j is GV and EV. Then γ_{iG} and γ_{iE} capture the household preference to GV and EV. Other vehicle characteristics x_j enter the utility function linearly (including a constant term). Term ξ_j is the unobserved (by the econometrician) product characteristic of vehicle j, and ϵ_{ij} is the unobserved taste shock following type one extreme value distribution.

For household preferences on GV and EV, the parameters vary as functions of individual characteristics. The individual characteristics consist of two parts: observed demographics D_i and unobserved characteristics ν_i. I denote the preference on GV and EV as $\gamma_i = (\gamma_{iE}, \gamma_{iG})'$, and it can be modeled as

$$\gamma_i = \gamma + \Pi D_i + \Sigma \nu_i \tag{3}$$

For simplicity, D_i and ν_i are assumed to be independent. Π is the (K $\times$ 2) matrix of coefficients that measure how the taste to GV and EV change with K demographic variables, and Σ is a 2×2 matrix of parameters. I allow that household preference to EV/GV depends on the vehicle holding h_i. The preference to GV γ_{iG} and preference to EV γ_{iE} consist of three parts: preference fixed effects, observed heterogeneity parts based on vehicle holding, and unobserved heterogeneity parts:

$$\begin{aligned}
\gamma_{iG} &= \gamma_G + \Gamma_{same}1(h_i = G) + \Gamma_{hybrid}1(h_i = E) + \sigma_G \nu_{iG} \\
\gamma_{iE} &= \gamma_E + \Gamma_{hybrid}1(h_i = G) + \Gamma_{same}1(h_i = E) + \sigma_E \nu_{iE}
\end{aligned} \tag{4}$$

The preference fixed effects parts γ_G and γ_E are constant for all households. For households with no car ($\emptyset$) (the reference group), the preference for EV is characterized by $\gamma_E + \sigma_E \nu_{iE}$. For households with GV, the preference for EV is characterized by $\gamma_E + \Gamma_{hybrid} + \sigma_E \nu_{iE}$. For households with EV, the preference for EV is $\gamma_E + \Gamma_{same} + \sigma_E \nu_{iE}$. Thus, Γ_{hybrid} terms measure the extent to which the added utility of consuming GV increases if the household has an EV, and Γ_{same} terms measure the extent to which the added utility of consuming EV increases if the household has an EV. Intuitively, the terms Γs capture the utility of owning the combination of the vehicles. For simplicity, I only allow the heterogeneity terms to differ on whether the household owns the same or different types of vehicles after purchase. For a household with a GV(EV) to buy an EV(GV), this household will own diversified types of vehicles. For a household with a GV(EV) to buy a GV(EV), this household will own a uniform combination type of vehicle. I assume that ν_{iG} and ν_{iE} follow the standard normal distribution. The preferences for the GV and EV are then

$$\begin{aligned}
\gamma_{iE} &\sim N(\gamma_E + \Gamma_{hybrid}1(h_i = G) + \Gamma_{same}1(h_i = E), \sigma_E^2) \\
\gamma_{iG} &\sim N(\gamma_G + \Gamma_{same}1(h_i = G) + \Gamma_{hybrid}1(h_i = E), \sigma_G^2)
\end{aligned} \tag{5}$$

Households with different vehicles have different preferences for the next vehicle, and the preference heterogeneity induced by vehicle holding is reflected in the mean part.

Table 8 summarizes the model specification for the observed preference heterogeneity for fuel types depending on vehicle holding. Panel (a) is the simplest version of the model mentioned above. Panel (b) describes the case allowing households with GV and EV to differ in owning the same types of vehicle combination. That is, households consuming the combination of "GG" and "EE" can be different. Panel (c) is the full model, allowing all heterogeneity terms to differ.

Table 8: Model specification

(a) Simplest Model

	none	GV	EV
not buy	0	0	0
GV	γ_G	$\gamma_G + \Gamma_{same}$	$\gamma_G + \Gamma_{hybrid}$
EV	γ_E	$\gamma_E + \Gamma_{hybrid}$	$\gamma_E + \Gamma_{same}$

(b) With Heterogeneity on GV & GV and EV & EV

	none	GV	EV
not buy	0	0	0
GV	γ_G	$\gamma_G + \Gamma_{GG}$	$\gamma_G + \Gamma_{hybrid}$
EV	γ_E	$\gamma_E + \Gamma_{hybrid}$	$\gamma_E + \Gamma_{EE}$

(c) With Full Heterogeneity

	none	GV	EV
not buy	0	0	0
GV	γ_G	$\gamma_G + \Gamma_{GG}$	$\gamma_G + \Gamma_{EG}$
EV	γ_E	$\gamma_E + \Gamma_{GE}$	$\gamma_E + \Gamma_{EE}$

For income effect in estimation, instead of including the term $\log(y_i - p_j)$, which gives rise to a host of numerical problems, I follow S. Berry, Levinsohn, and Pakes (1999) and use its first-order linear approximation, $-p_j/y_i$ in estimation part. I keep using the logarithm form $\log(y_i - p_j)$ for model setup.

Given the random shock following the extreme value distribution, the equilibrium CCP to choose alternative j takes the logit form (here normalizing utility to outside option as zero and assuming that α_i and β_i are the same for all households):

$$s_{ij} = \frac{exp(\alpha log(y_i - p_j) + x_j\beta + \gamma_{iG}1(fuel_j = G) + \gamma_{iE}1(fuel_j = E) + \xi_j)}{1 + \sum\limits_{j \in J} exp(\alpha log(y_i - p_l) + x_l\beta + \gamma_{iG}1(fuel_l = G) + \gamma_{iE}1(fuel_l = E) + \xi_l)} \tag{6}$$

Integrating across the households, the market share for product j is

$$s_j = \int_i \frac{exp(\alpha log(y_i - p_j) + x_j\beta + \gamma_{iG}1(fuel_j = G) + \gamma_{iE}1(fuel_j = E) + \xi_j)}{1 + \sum_{j\in J} exp(\alpha log(y_i - p_l) + x_l\beta + \gamma_{iG}1(fuel_l = G) + \gamma_{iE}1(fuel_l = E) + \xi_l)} dF_i \quad (7)$$

where F_i is the CDF of the household demographic variable.

5.2 Supply

The supply side is the standard Bertrand competition model. Consider the profit of firm f, which controls several products J_f and sets price p_j (The MSRP of product j). The first-order conditions of the profit function are as follows:

$$\max_{p_j:j\in J_f} \sum_{j\in J_f} s_j(\mathbf{p})(p_j - c_j),$$

$$s_j(\mathbf{p}) + \sum_{j\in J_f} \frac{\partial s_j(\mathbf{p})}{\partial p_j}(p_j - c_j) = 0 \quad (8)$$

Writing the above first-order condition in matrix form:

$$\mathbf{p} - \mathbf{c} = \eta = \Delta^{-1}(\mathbf{p})\mathbf{s}(\mathbf{p}) \quad (9)$$

where $\Delta^{-1} = -\mathcal{H} \odot \frac{\partial \mathbf{s}(\mathbf{p})}{\partial \mathbf{p}}$ is the element-by-element product. $\mathcal{H}$ is the ownership matrix where (j, k) element equal to one means that the same firm produces product j and k, and $\frac{\partial \mathbf{s}(\mathbf{p})}{\partial \mathbf{p}}$ is the derivatives and (j, k) element is equal to $\frac{\partial s_j(\mathbf{p})}{\partial p_k}$. Denote θ to be the full parameters to be estimated. Then I can recover the marginal cost $c_j = p_j - \eta_j(\theta)$, which allows me to construct the supply side moments.

5.3 Identification and micromoments

The main challenging identification issue is in the heterogeneity part of γ_i, especially the preference heterogeneity terms induced by vehicle holding $\Gamma = \{\Gamma_{GG}, \Gamma_{GE}, \Gamma_{EG}, \Gamma_{EE}\}$.

The identification for other terms follows S. Berry, Levinsohn, and Pakes (1995). Following Petrin (2002), I use the California Vehicle Survey data on the households to construct micro-moments. The moments match the average GV and EV purchase probability, conditional on vehicle holding. The moments match the average model predictions of EV/GV purchase probability to the observed averages from the survey respondents. These moments are given by

$$\mathbb{E}[\{\text{i purchases GV}\}|\{h_i = GV\}],$$

$$\mathbb{E}[\{\text{i purchases GV}\}|\{h_i = EV\}],$$

$$\mathbb{E}[\{\text{i purchases GV}\}|\{h_i = \emptyset\}],$$

$$\mathbb{E}[\{\text{i purchases EV}\}|\{h_i = EV\}],$$

$$\mathbb{E}[\{\text{i purchases EV}\}|\{h_i = GV\}],$$

$$\mathbb{E}[\{\text{i purchases EV}\}|\{h_i = \emptyset\}],$$

where $\{i \text{ purchases GV}\}$ is the event that household i purchases a GV, and $\{h_i = GV\}$,$\{h_i = EV\}$,$\{h_i = \emptyset\}$ are, the events that household i holds GV, EV, or nothing in hand, respectively. Denoting the probability for households with vehicle fuel type $v_0 \in \{\emptyset, GV, EV\}$ choosing option j of fuel type $v_1 \in \{\emptyset, GV, EV\}$ as $s_i(v_1|h_i = v_0)$, I can express the first micromoment by $s_i(G|h_i = G) = \mathbb{E}[\{\text{i purchases GV}\}|\{h_i = GV\}]$ (other moments can be expressed similarly). Using the sample analog, I can get the vehicle-holding specific choice probability in the survey data.

The preference heterogeneity terms induced by vehicle holdings can be identified through the vehicle-holding specific choice probabilities:

$$\begin{aligned}
\Gamma_{GE} &= log\frac{s_i(E|h_i=G)}{s_i(\emptyset|h_i=G)} - log(\frac{s_i(E|h_i=\emptyset)}{s_i(\emptyset|h_i=\emptyset)}), \\
\Gamma_{EE} &= log\frac{s_i(E|h_i=E)}{s_i(\emptyset|h_i=E)} - log(\frac{s_i(E|h_i=\emptyset)}{s_i(\emptyset|h_i=\emptyset)}), \\
\Gamma_{GG} &= log\frac{s_i(G|h_i=G)}{s_i(\emptyset|h_i=G)} - log(\frac{s_i(G|h_i=\emptyset)}{s_i(\emptyset|h_i=\emptyset)}), \\
\Gamma_{EG} &= log\frac{s_i(G|h_i=E)}{s_i(\emptyset|h_i=E)} - log(\frac{s_i(G|h_i=\emptyset)}{s_i(\emptyset|h_i=\emptyset)}), \qquad (10)
\end{aligned}$$

The derivation of equation 10 is in the Appendix.

5.4 Estimation

In this section, I describe the estimation of parameters in demand. My estimation strategy resembles the generalized method of moments (GMM) taken by S. Berry, Levinsohn, and Pakes (1995) and Petrin (2002). I supplement the macro moments with micro-moments the household survey data provides.

5.4.1 The macro moments

The first set of moments matches the market-level disturbances ($\xi_j(\theta)$). The unobserved demand disturbances are assumed to be uncorrelated with observed demand-side variables of all vehicles in that year. Then demand side moment is

$$\mathbb{E}[\xi_j(\theta)Z_j^D] = 0, \qquad (11)$$

where Z^D is the demand side instruments. The instruments used for demand estimations are:

- steel price $\times$ car weight (1 IV).
- BLP instruments (rival case) for HPwt and wheelbase (2 IVs).

- Differentiation instruments (rival case) for HPwt and wheelbase (2 IVs) following Gandhi and Houde (2019).

- Differentiation instruments (rival case) for HPwt and wheelbase $\times$ EV indicator (2 IVs).

The second set of moments matches the model's share predictions $s_j(\delta(\theta), \theta)$, to the shares in the data, S_j:

$$s_j(\delta(\theta), \theta) = S_j, \tag{12}$$

where $\delta_j(\theta) = x_j\beta + \gamma_E 1(fuel_j = E) + \gamma_G 1(fuel_j = G) + \xi_j$ is the mean utility. The other nonlinear part of utility is denoted as $\mu_{ij} = \alpha log(y_i - p_j) + \Pi D_i + \Sigma\nu_i$. This set of moments enters the estimation in finding a fixed point of mean utility.

5.4.2 The micro moments

The set of micro-moments, as mentioned above, matches the model predictions of EV/GV purchase probability from the macro sales data and those from the households survey:

$$\mathbb{E}[\{\text{i purchases GV}\}|\{h_i = GV\}],$$

$$\mathbb{E}[\{\text{i purchases GV}\}|\{h_i = EV\}],$$

$$\mathbb{E}[\{\text{i purchases GV}\}|\{h_i = \emptyset\}],$$

$$\mathbb{E}[\{\text{i purchases EV}\}|\{h_i = EV\}],$$

$$\mathbb{E}[\{\text{i purchases EV}\}|\{h_i = GV\}],$$

$$\mathbb{E}[\{\text{i purchases EV}\}|\{h_i = \emptyset\}],$$

5.4.3 The objective function

The program with macro moment $g_1(\theta)$ and micro-moment $g_2(\theta)$ can be summarized as following:

$$\begin{aligned}
\min_{\theta} q(\theta) &\equiv g(\theta)'\tilde{W}g(\theta) \\
g(\theta) &= \begin{bmatrix} g_1(\theta) \\ g_2(\theta) \end{bmatrix}; \\
\xi_j &= \delta_j - x_j\beta - \gamma_E 1(fuel_j = E) - \gamma_G 1(fuel_j = G), \\
\eta_j &= \Delta^{-1}(\theta)\mathbf{s}, \\
S_j &= s_j(\delta_j;\theta) \qquad (13)
\end{aligned}$$

where the weighting matrix is

$$\tilde{W} = \begin{bmatrix} W_1 & 0 \\ 0 & W_2 \end{bmatrix} = \begin{bmatrix} (Z_D Z_D')^{-1} & 0 \\ 0 & W_2 \end{bmatrix} \qquad (14)$$

where W_2 at the lower right part matches the number of respondents with different vehicle holdings in the survey data.

The estimation procedures are similar to Petrin (2002), and I summarise them as follows:

1. Guess a value of θ.

2. Given θ, solve the nonlinear equation 12 to get the mean utility $\hat{\delta}_j(\theta)$.

3. Solve the linear IV problem for β, γ through the GMM objective function.

4. Update θ, and iterate until convergence.

I deal with variance of linear part of parameters $\theta_1 = [\beta', \gamma']'$ and variance of non-linear parameters $\theta_2 = [\alpha, \Gamma, \sigma]'$ separately. The variance of θ_2 comes from the GMM objective function. Since $\theta_1 = C\delta(\theta_2)$ (C here comes from the expression of mean utility), the variance for the linear parameters can be derived using the Delta method:

$$var(\theta_1) = \nabla(C\delta(\theta_2))'\Sigma_{\theta_2}\nabla(C\delta(\theta_2)) \tag{15}$$

and $se(\theta_1) = var(\theta_1)^{1/2}$.

6 Results

Table 9 reports the results for the demand-side models: Column (2) presents demand estimates without preference heterogeneity induced by vehicle holding ("w/o EH"). The fixed effect for GV is much higher than that for EV, although both are negative. For other characteristics, I include the vehicle size (carsize), Horsepower/weight (HPwt), wheelbase, and the number of charging stations for EVs, which interacted with the indicator for EVs. The estimates for these characteristics are similar across all specifications.

Column (3) presents the demand model with preference heterogeneity induced by vehicle holding ("w/ EH").[14] I report the estimated results under the simplest model mentioned in table 8: Γ_{same} represents the preference for the fuel type when the vehicle holding fuel type is the same as the fuel type to buy, and Γ_{hybrid} represents the preference for the fuel type when the vehicle holding fuel type is different from the fuel type to buy. The fixed effect terms for GV and EV are much higher than those in column (2), while the preference heterogeneity terms induced by vehicle holding (Γ_{same} and Γ_{hybrid}) are both negative. The results of fixed effect terms are consistent with the summary statistics that households with no car tend to buy a vehicle more than households with cars. Moreover, Γ_{same} is significantly larger than Γ_{hybrid} (with t-stat equals to 3.302 in the lower panel of table 9). Households prefer to have the same types of vehicles rather than diversify the fuel types. Keeping the same car characteristics, for a representative median income household with a GV in hand, EV should

[14] The income for each household in the survey data is drawn from the empirical distribution from the survey. The median income for households with no car, with GV, and with EV are 80,659, 82,805, 137,924 USD, respectively.

be $1.592/11.937 \times 82,805 \approx 11,043$ USD cheaper to make the household with GV feel the same of holding "GG" combination and "GE" combination. This gap is much larger than the at-home charging port average cost ($1,000). There could still be the cost of holding a hybrid combination of vehicles other than the installation costs of charging stations.

Column (4) presents the demand model with vehicle holding and income heterogeneity ("w/ income & EH"). I separate all households into two groups: *lowinc* and *highinc*, depending on whether household annual income is lower or higher than $150K. The results are similar to column (3), but only $\Gamma_{same}^{lowinc} - \Gamma_{hybrid}^{lowinc}$ is significant for the preference heterogeneity induced by vehicle holding. Low-income households have significant preference differences between GV and EV, depending on their vehicle holding.

To better understand the results from demand estimates, I present the utility of getting a vehicle with representative vehicle characteristics for a representative household in each group of vehicle holding. The representative household in each group takes the group median income. The representative characteristics for each vehicle include: carsize = 0.7995 (inch^3/1,000,000), HPwt = 0.06172 (HP/lb), wheelbase = 1.093 (inch/100), price = 36,417 USD. The unobserved characteristics ξ_j and individual specific term ν_i are assumed to be zero. Utility from the outside option is normalized to zero for each group. Taking the coefficient from column (3) in table 9 as an example, I present the utility for getting each option for a representative household in table 10. First, a typical household's total utility from buying a vehicle is smaller than buying nothing, consistent with the small market share for each vehicle model (compared to households choosing the outside option in the market). Second, the total utility from buying a GV is larger than that from buying an EV for each group. The total utility difference (ΔU) for each group can be decomposed into two parts: the first part is contrast in utility for GV and EV *i.e.* $\gamma_i = (\gamma_{iE}, \gamma_{iG})'$ (which includes two components: the preference fixed effect $\gamma_G - \gamma_E = 6.028$, and preference heterogeneity induced by vehicle holdings $\Gamma_{same} - \Gamma_{hybrid} = 1.592$), and second components are utility from other characteristics, which concentrates on preference difference from charging stations when buying EV (0.749).

Table 9: Demand System Estimates

	w/o EH	w/ EH	w/ income & EH
log(y-p)	13.362	11.937	10.648
	(1.639)	(1.207)	(0.443)
γ_G	-8.497	-1.724	-0.875
	(1.336)	(0.794)	(0.504)
γ_E	-17.571	-7.752	-8.894
	(2.675)	(2.288)	(2.322)
carsize	3.308	3.279	3.381
	(0.098)	(0.059)	(0.056)
HPwt	24.705	23.961	25.746
	(1.156)	(1.057)	(1.036)
wheelbase	1.028	0.978	1.037
	(0.206)	(0.091)	(0.090)
log(station) $\times$ I(EV)	0.224	0.216	0.190
	(0.019)	(0.010)	(0.029)
σ_G	3.318	2.997	-0.561
	(0.907)	(0.852)	(1.968)
σ_E	5.553	4.460	4.565
	(0.933)	(0.969)	(0.982)
Γ_{same}		-7.045	
		(1.128)	
Γ_{hybrid}		-8.637	
		(1.181)	
$\Gamma_{same}^{highinc}$			-7.412
			(1.092)
$\Gamma_{hybrid}^{highinc}$			-7.846
			(1.087)
Γ_{same}^{lowinc}			-5.382
			(1.142)
Γ_{hybrid}^{lowinc}			-6.575
			(1.237)
Γ_{same} - Γ_{hybrid}		1.592	
t_{diff}		(3.302)	
$\Gamma_{same}^{highinc}$ - $\Gamma_{hybrid}^{highinc}$			0.434
t_{diff}			(1.302)
Γ_{same}^{lowinc} - Γ_{hybrid}^{lowinc}			1.193
t_{diff}			(2.752)

Note: numbers in parenthesis in the upper panel are standard errors. Numbers in parenthesis in the lower panel are t-statistics for the differences between Γs.

Taking households with GV as an example, total utility differences for getting GV and EV $\Delta U = (\gamma_{iG} - \gamma_{iE}) - 0.749 = (\gamma_G - \gamma_E) + (\Gamma_{same} - \Gamma_{hybrid}) - 0.749 = 6.028 + 1.592 - 0.749 = 6.871$. While the common preference fixed effect for GV and EV dominates, preference heterogeneity induced by vehicle holdings still plays an essential role in explaining utility contrast among the three types of households.

Table 10: Utility for a Representative Household and Vehicle

Utility for EV and GV			
	none	GV	EV
not buy	0	0	0
buy GV	-1.724	-8.769	-10.361
buy EV	-7.752	-16.389	-14.797
Utility from other characteristics			
	none	GV	EV
not buy	0	0	0
buy GV	-0.146	0.000	2.109
buy EV	0.602	0.749	2.857
Total utility			
	none	GV	EV
not buy	0	0	0
buy GV	-1.870	-8.769	-8.252
buy EV	-7.150	-15.640	-11.940
ΔU= U_{GV} - U_{EV}	5.279	6.871	3.687

Note: The table shows the utility decomposition for a representative household and vehicle with representative characteristics. *Utility for EV and GV* is the utility from buying EV or GV (γ_i) (shown in table 8). *other characteristics* is the utility from other characteristics of the vehicle, including carsize (inch^3/1,000,000), HPwt (HP/lb), wheelbase (inch/100), price effects, and charging station for household buying EV. *total utility* is the total utility for the household. ΔU= U_{GV} - U_{EV} is the difference between the total utility from GV and EV.

I report the micro-moment matching result in table I5. I match the summarized market shares for buying a GV and EV. The model is based on column (2) for table 9. Table 11 presents a sample of own and cross-price elasticities implied by the demand estimates. I chose the MSA 31080 (Los Angeles Area) market in 2016. The cross-price elasticities are larger among similar products, and EVs have larger own-price elasticities than GVs. Overall, the elasticities estimated are in line with those from the literature in automobile demand estimation.

Table 11: Elasticity for Selected Vehicle Models (in the market of MSA 31080 in 2016)

		NISSAN LEAF	TESLA MODEL S	BMW I3	HONDA CIVIC	TOYOTA COROLLA	HONDA CR-V	HONDA ACCORD	KIA SOUL	TOYOTA PRIUS	CHEVROLET VOLT
NISSAN,LEAF	E	-3.1390	0.4175	0.0337	0.0314	0.0026	0.0113	0.0173	0.0052	0.0103	0.6559
TESLA,MODEL S	E	0.0271	-5.8012	0.0098	0.0179	0.0013	0.0096	0.0145	0.0022	0.0082	0.1598
BMW,I3	E	0.0976	0.4375	-4.2698	0.0229	0.0018	0.0097	0.0148	0.0034	0.0086	0.5576
HONDA,CIVIC	G	0.0008	0.0069	0.0002	-2.3168	0.0102	0.0502	0.0767	0.0194	0.0452	0.0043
TOYOTA,COROLLA	G	0.0009	0.0067	0.0002	0.1331	-2.3599	0.0519	0.0793	0.0208	0.0469	0.0046
HONDA,CR-V	G	0.0006	0.0075	0.0002	0.1014	0.0080	-2.6261	0.0672	0.0148	0.0390	0.0032
HONDA,ACCORD	G	0.0006	0.0074	0.0002	0.1021	0.0081	0.0443	-2.5958	0.0149	0.0392	0.0032
KIA,SOUL	G	0.0010	0.0063	0.0002	0.1439	0.0118	0.0543	0.0830	-2.2300	0.0493	0.0051
TOYOTA,PRIUS	HEV	0.0006	0.0074	0.0002	0.1057	0.0084	0.0452	0.0689	0.0156	-2.5917	0.0034
CHEVROLET,VOLT	PHEV	0.1124	0.4222	0.0330	0.0291	0.0024	0.0108	0.0165	0.0047	0.0098	-2.8251

Table 12 presents the demand estimation results under other model specifications in table 8. The second column shows the demand estimates under three preference heterogeneity terms induced by vehicle holding (that is, I set Γ_{GG} and Γ_{EE} differently), the estimates for Γ_{GG} and Γ_{EE} are pretty similar. Column (4) presents the results under full preference heterogeneity terms induced by vehicle holding. The estimation for Γ_{GE} and Γ_{EE} are not significant from zero, and they are not significantly different from Γ_{GG} and Γ_{EG} either. The insignificance in the full heterogeneity model could be because the sample size for households purchasing EVs is low in the survey data. The estimation will be more accurate with a larger sample size of households buying EVs.

Figure 3 presents the marginal cost distribution implied by the demand estimation. The median marginal cost for GVs is about 22,800 USD, and the median marginal cost for EVs is about 25,219 USD. Marginal costs for EVs are higher than those of GV.

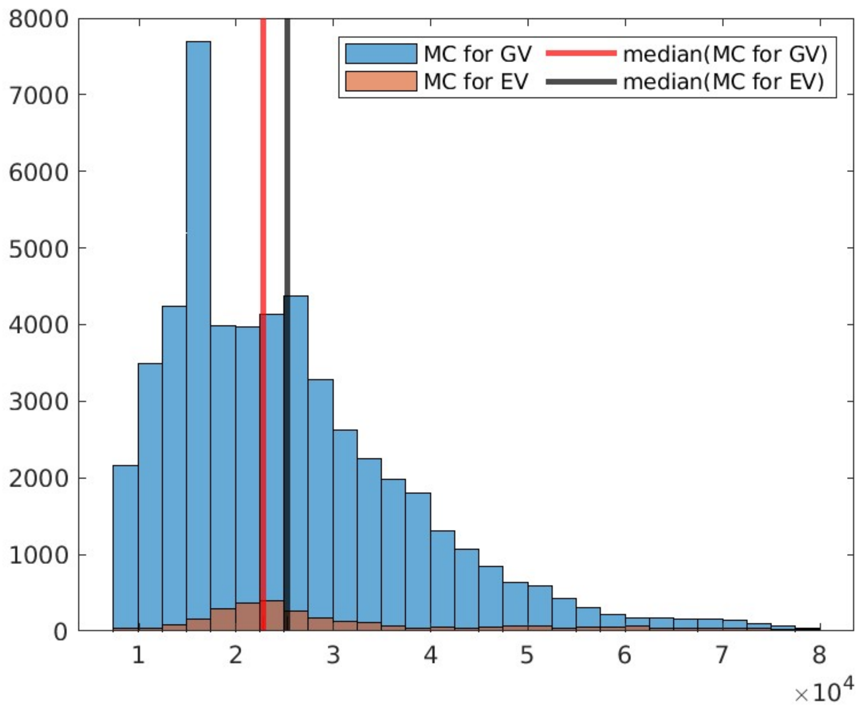


Figure 3: Distribution of Marginal Cost for GV and EV

Overall, household vehicle preferences exhibit a strong correlation with their existing

Table 12: Demand Estimation for other specifications

Intermediate case		Full heterogeneity	
γ_G	-1.478	γ_G	-1.419
	(0.784)		(0.758)
γ_E	-7.559	γ_E	-10.207
	(3.628)		(4.077)
carsize	3.274	carsize	3.280
	(0.060)		(0.062)
HPwt	23.879	HPwt	23.378
	(1.095)		(1.078)
wheelbase	0.983	wheelbase	0.823
	(0.096)		(0.097)
log(station) $\times$ PEV	0.217	log(station) $\times$ PEV	0.196
	(0.017)		(0.015)
σ_G	2.894	σ_G	2.885
	(0.861)		(0.895)
σ_E	4.442	σ_E	3.645
	(1.487)		(0.806)
log(y-p)	11.946	log(y-p)	11.866
	(1.234)		(1.210)
Γ_{GG}	-7.190	Γ_{GG}	-7.255
	(1.138)		(1.157)
Γ_{hybrid}	-8.844	Γ_{GE}	-2.871
	(1.184)		(3.411)
Γ_{EE}	-7.188	Γ_{EG}	-9.124
	(1.710)		(1.240)
		Γ_{EE}	-3.722
			(3.566)
Γ_{GG} - Γ_{hybrid}	1.654	Γ_{GG} - Γ_{GE}	-4.384
t_{diff}	3.353	t_{diff}	-1.249
Γ_{hybrid} - Γ_{EE}	-1.656	Γ_{EG} - Γ_{EE}	-5.402
t_{diff}	-1.089	t_{diff}	-1.446

numbers in parenthesis in the upper panel are standard errors. Numbers in parenthesis in the lower panel are t-statistics for the differences between Γs.

vehicle holdings. It is evident that households without any vehicles tend to express a significantly stronger inclination toward acquiring a new vehicle in comparison to those already owning vehicles, and households with EVs prefer an EV more than households with GVs. That is, households prefer to have the same fuel types rather than diversify vehicle fuel types when choosing between GV and EV. This preference partly explains why EV adoption is slow. With these demand and supply estimates in mind, I can now explore their implications for the design of subsidy policies.

7 Counterfactual Analysis

In this section, I will evaluate market outcomes and welfare implications in the counterfactual policy regime of a tax rebate for electric vehicles. I look at two types of experiments: the first one is repealing the subsidies for EVs. I mainly compare the results from the models with and without preference heterogeneity induced by vehicle holdings. The second one is changing the subsidy scheme from uniform to heterogeneous subsidies based on household vehicle holdings. In both experiments, I leverage the parameter estimates from the previous section to recompute the equilibrium price under the aforementioned model, and subsequently, I compute the pertinent market outcomes. It is essential to underscore that these computations are grounded on the assumption that the vehicle models, marginal costs, and demographic characteristics remain the same as the observed data. In terms of subsidies, I focus on the federal tax credits, which span a range from 2,500 to 7,500 USD, contingent upon the battery size of the EV.

The common approach to solve the equilibrium price is to treat equation 9 as a fixed point and iterate on the equation. The problem is that the equation is not necessarily a contraction mapping. Instead, I follow the faster and more reliable method by Morrow and Skerlos (2011) and Conlon and Gortmaker (2020) who reformulate equation 9 by breaking up the demand derivatives into two parts: a diagonal $J \times J$ matrix Λ, and a $J \times J$ dense matrix Γ:

$$\begin{aligned} \frac{\partial \boldsymbol{s}}{\partial \boldsymbol{p}}(\boldsymbol{p}) &= \Lambda(\boldsymbol{p}) - \Gamma(\boldsymbol{p}), \\ \Lambda_{jj} &= \int \frac{\partial u_{ij}}{\partial p_j} s_{ij} dF_i, \\ \Gamma_{jk} &= \int \frac{\partial u_{ij}}{\partial p_j} s_{ij} s_{ik} dF_i \end{aligned} \tag{16}$$

The problem then can be reformulated as a different fixed point that is specific to mixed logit demands:

$$\begin{aligned} \boldsymbol{p} &= \boldsymbol{c} + \zeta(\boldsymbol{p}), \\ \text{where } \zeta(\boldsymbol{p}) &= \Lambda(\boldsymbol{p})^{-1}[\mathcal{H} \odot \Gamma(\boldsymbol{p})](\boldsymbol{p} - \boldsymbol{c}) - \Gamma(\boldsymbol{p})^{-1}\boldsymbol{s}(\boldsymbol{p}) \end{aligned} \tag{17}$$

Morrow and Skerlos (2011) argue that numerical methods based on equation 17 can have entirely different properties from equation 9 because they are different functions, and these two functions coincide only at stationary prices.

7.1 Simulation One: Subsidy Repealing Designs

Table 13 presents the market outcomes before and after the subsidies on EVs are repealed. Columns 2-5 report the results under the model without preference heterogeneity induced by vehicle holding. Column 2 presents the market outcomes under the current federal tax credits. I report the median price for GVs and EVs, and the vehicle sales for households with no car, GVs, and EVs. Column 3 presents the model predicted market outcomes if all federal tax credits are repealed. Column 4 shows the difference between the cases with and without federal tax credits. Furthermore, Column 5 shows the respective percentage change. After the federal tax credits are repealed, the EV prices increase slightly, while GV prices stay almost the same. Total vehicle sales decrease, and EV sales decrease by 30%, especially in households with no vehicle (decrease by 57%). GV sales decrease slightly.

Columns 5-9 show the counterpart results under the model with preference heterogeneity induced by vehicle holding. Like the model without such heterogeneity, EV prices increase, and GV prices remain similar. Overall sales are also similar. The differences focus on the change in EV sales. Under the model with preference heterogeneity induced by vehicle holding, EV sales decrease by 32.51%. Households with vehicles and without vehicles perform similarly. The reason is that households with different vehicle holdings dislike EVs in the same way under the model without preference heterogeneity induced by vehicle holding. In the model with such heterogeneity, the households with no vehicle enjoy EVs more, which is consistent with the summary statistics that households without vehicles have a higher propensity to buy a vehicle.

Table 14 presents the welfare changes after federal tax credits are repealed under two models. I report the change in total welfare, consumer surplus (CS), producer surplus (PS), and total subsidy expenditure. Calculating consumer surplus and producer surplus follows the literature of demand estimation in differentiated products like J. Li (2019). I also report the change of CS for households with different vehicle holdings, the change of PS for GVs and EVs, and the subsidy expenditure the different households get. The last row is the average subsidy per EV. All numbers share a unit of 10K USD. Similar to the market outcome in table 13, both models predict similar overall welfare changes. The total welfare changes are negative after the federal tax rebate is repealed. Consumer surpluses decrease, producer surpluses decrease, and subsidy expenditure decreases to zero.

However, households with different vehicle holdings get different welfare levels. Under the model without preference heterogeneity induced by vehicle holding, the CS for no-vehicle households is small due to the predicted sales. Under the model with preference heterogeneity induced by vehicle holding, households with no vehicle are predicted to purchase more EVs and lose more consumer surplus if the subsidies are repealed. A similar pattern comes to the subsidy amount changes for these households with no vehicle.

Table 13: CF1: Market outcome after the subsidy repealed

	w/o EH				w/ EH			
	w/ subsidy	w/o subsidy	difference	difference (%)	w/ subsidy	w/o subsidy	difference	difference (%)
median price (EV)	41,450	43,425	909	2.04%	41,450	44,278	1,124	2.36%
median price (GV)	36,417	36,412	-6	-0.02%	36,417	36,409	-13	-0.03%
total sales	4,065,804	4,017,397	-48,407	-1.19%	4,064,702	4,020,308	-44,394	-1.09%
EV sales	147,874	103,202	-44,672	-30.21%	146,347	98,775	-47,572	-32.51%
EV sales (no car HH)	3,391	1,456	-1,936	-57.08%	41,503	26,964	-14,540	-35.03%
EV sales (with GVs)	129,464	91,391	-38,074	-29.41%	80,709	54,496	-26,213	-32.48%
EV sales (with EVs)	15,018	10,356	-4,662	-31.05%	24,135	17,316	-6,819	-28.25%
GV sales	3,917,930	3,914,195	-3,735	-0.10%	3,918,355	3,921,532	3,177	0.08%
GV sales (no car HH)	120,784	120,550	-234	-0.19%	717,763	724,347	6,584	0.92%
GV sales (with GVs)	3,312,350	3,309,168	-3,182	-0.10%	2,961,448	2,958,037	-3,411	-0.12%
GV sales (with EVs)	484,796	484,477	-319	-0.07%	239,143	239,148	5	0.00%

Note: this table presents the counterfactual market outcomes after the federal tax rebates are taken out. “w/ subsidy” column shows the result under the current federal tax rebates from 2014-2016, and “w/o subsidy” column shows the result after all federal tax rebates taken out. Model “w/o EH” is the model without preference heterogeneity induced by vehicle holding, and the model “w/ EH” is the simplest preference heterogeneity induced by vehicle holding in table 9.

Table 14: CF1: Welfare outcome after subsidy repealed

	w/o EH			w/ EH		
	w/ subsidy	w/o subsidy	difference	w/ subsidy	w/o subsidy	difference
total welfare (10K USD)	20,836,134	20,575,863	-54,378	22,865,375	22,606,853	-54,721
CS	15,732,272	15,588,578	-143,694	17,603,145	17,457,447	-145,698
CS (no car HH)	789,481	787,286	-2,196	5,355,604	5,327,046	-28,558
CS (with GVs)	12,207,555	12,078,198	-129,357	11,015,049	10,916,571	-98,478
CS (with EVs)	2,735,235	2,723,094	-12,141	1,232,493	1,213,831	-18,662
PS	5,000,916	4,987,285	-13,631	5,160,329	5,149,406	-10,924
PS (for GV)	4,730,812	4,730,442	-370	4,894,583	4,899,622	5,039
PS (for EV)	270,103	256,843	-13,260	265,746	249,783	-15,963
total subsidy	102,947	0	-102,947	101,900	0	-101,900
subsidy (no car HH)	2,428	0	-2,428	29,013	0	-29,013
subsidy (with GVs)	89,929	0	-89,929	55,976	0	-55,976
subsidy (with EVs)	10,590	0	-10,590	16,911	0	-16,911
avg subsidy	0.6962	0	-0.6962	0.6963	0	-0.6963

Note: this table shows the counterfactual welfare comparison after the federal tax rebates are taken out. "w/ subsidy" column shows the result under the current federal tax rebates from 2014-2016, and "w/o subsidy" column shows the result after all federal tax rebates taken out. Model "w/o EH" is the model without preference heterogeneity induced by vehicle holding, and the model "w/ EH" is the simplest preference heterogeneity induced by vehicle holding in table 9.

7.2 Simulation Two: Heterogenous Subsidy Policy Design

In this simulation, I will redesign the policy on EV subsidy and the new subsidy scheme heterogenous on household vehicle holding. I first set grids for different federal tax rebates: For each group (none, GV, EV), the new subsidy varies from 50% to 150% of the current rebate level (the grid width is 5%). Then, I compute the new market outcome, welfare, and total subsidy expenditure for each grid point. Finally, I pick the best scenarios under different policy goals. I show two federal tax rebate allocations for different policy goals. Case one targets to raise consumer surplus given similar subsidy budgets. Case two targets to raise the EV sales given similar subsidy budgets.

Case One: Raising CS and EV Sales

In this case, I target to raise consumer surplus and EV sales simultaneously. I raise the federal tax rebates by 10% for households with GVs and reduce the federal tax rebates by 45% for households with EVs, relative to the original federal tax rebates (households with no car still get the original federal tax rebates). Table 15 shows the price and sales change under this new subsidy policy. The second column presents results under the current uniform federal tax rebates for all households. The third column presents results under the heterogeneous federal tax rebates design. Column 4 shows the difference between the two scenarios, and column 5 shows the percentage change. Under this new subsidy policy, prices for EVs and GVs do not change much overall, and total sales and EV sales increase by a small amount. EV sales for households with EV decreases by 11.4% while EV sales for households with GV increase and compensate for it. Households with EVs are less elastic to price, so cutting the rebates in this group will not cut the sales much. Table 16 reports the welfare change under case one—total welfare increases by 5.59 million dollars, and total consumer welfare increases. Households with EVs benefit less from the new subsidy scheme, and households with GVs benefit more to compensate for the loss. Total subsidy expenditure decreases by 690K USD, and total consumer surplus increases by 15.8 million USD. Thus, total welfare and EV sales can be improved with even smaller subsidy expenditures while the magnitude is limited.

Table 15: Simulation Two: Case One: market outcome under heterogenous subsidy

	current subsidy	new subsidy	difference	difference (%)
median price (EV)	41,450	41,432	-12	-0.03%
median price (GV)	36,417	36,418	1	0.00%
total sales	4,064,702	4,066,095	1,394	0.03%
EV sales	146,347	147,404	1,057	0.72%
EV sales (no car HH)	41,503	41,901	398	0.96%
EV sales (with GVs)	80,709	84,113	3,404	4.22%
EV sales (with EVs)	24,135	21,390	-2,745	-11.37%
GV sales	3,918,355	3,918,691	337	0.01%
GV sales (no car HH)	717,763	717,664	-99	-0.01%
GV sales (with GVs)	2,961,448	2,961,957	509	0.02%
GV sales (with Evs)	239,143	239,071	-73	-0.03%

Note: this table presents the counterfactual market outcomes after the federal rebates are heterogenous for different households: federal tax rebates increase by 10% for households with GVs, and the federal tax rebates decrease by 45% for households with EVs, relative to the original federal tax rebates (households with no car still get the original federal tax rebates).

Table 16: Simulation Two: Case one: welfare analysis under heterogenous subsidy

	current subsidy	new subsidy	difference
total welfare (10K USD)	22,865,375	22,865,796	559
CS	17,603,145	17,604,724	1,578
CS (no car HH)	5,355,604	5,355,957	354
CS (with GVs)	11,015,049	11,023,396	8,348
CS (with EVs)	1,232,493	1,225,370	-7,123
PS	5,160,329	5,159,241	-1,089
PS (for GV)	4,894,583	4,894,812	229
PS (for EV)	265,746	264,429	-1,318
total subsidy	101,900	101,831	-69
subsidy (no car HH)	29,013	29,288	276
subsidy (with GVs)	55,976	64,357	8,380
subsidy (with EVs)	16,911	8,186	-8,725
avg subsidy	0.6963	0.6908	-0.0055

this table shows the counterfactual welfare changes after the federal rebates are heterogenous for different households: federal tax rebates increase by 10% for households with GVs, and the federal tax rebates decrease by 45% for households with EVs, relative to the original federal tax rebates (households with no car still get the original federal tax rebates).

Case two: Just Raising EV Sales

Although household welfare (consumer surplus) and EV sales exhibit simultaneous increases in case one, the extent of the surge in EV sales remains rather modest. In case two, I target to raise EV sales by redistributing the subsidies on EVs. Specifically, the approach involves a 25% increase in federal tax rebates for households with no vehicles, coupled with a 40% reduction in federal tax rebates for households owning GVs, relative to the original federal tax rebate scheme (households with EVs continue to receive the original federal tax rebates). Table 17 presents the market outcome changes under case two. While the total vehicle sales and median vehicle prices do not change much under the revised subsidy scheme, EV sales surge by a substantial 8%, equating to an additional 11,714 EV units sold, a stark contrast to the outcomes observed in case one. If the heterogenous subsidies on EV are not allowed, the government would need to allocate an additional $81.6 million to attain the same increase in EV sales through the existing subsidy scheme.[15] For context, California had

[15] This is calculated from the EV sales increment 11,714 times the average federal EV subsidy of $6,963

expended approximately 900 million USD in the pursuit of promoting EV sales by mid-2020. That is, the expenditure saving for promoting EV sales is similar to one-year state-level EV subsidy budget in California during 2010s.

Table 18 reports the welfare change under case two. While the total subsidy expenditure decreases by 7.67 million USD, with an average subsidy decrease of 565 dollars, the overall welfare experiences a decline of 206 million USD. Households with GVs derive considerably less benefit under the revised subsidy framework, with the other two household groups unable to fully compensate for this reduction. It is essential to acknowledge that while total welfare and consumer surplus exhibit declines in this scenario, these decreases are relatively modest in relation to the overall scale of the welfare measures (roughly decrease by 0.1%).

In summary, the reallocation of subsidies across different households based on their vehicle holdings yields a marginal enhancement in both total welfare and EV sales, even when the total subsidy expenditure is reduced. However, if the objective is to bolster EV sales while maintaining a comparable subsidy budget, certain households may experience a reduction in welfare, although the extent of this loss remains relatively minor. The intuition behind this increase in EV sales under the new subsidy scheme is straightforward: when consumers cannot be distinguished, a higher subsidy expenditure is necessitated to encourage a broader base of consumers to opt for EVs. Conversely, improvements in EV purchases can be achieved without the need for additional subsidies when consumers are segmented (in this paper through vehicle holdings), consumers within different segments exhibit varying elasticities. It's important to note that this process does entail a reduction in consumer surplus, as exchange among the different groups are restricted, leading to inefficiencies. The concept bears resemblance to third-degree price discrimination, albeit with the government now orchestrating this discrimination through subsidies.

Table 17: Simulation Two: Case two: market outcome under heterogenous subsidy

	current subsidy	new subsidy	difference	difference (%)
median price (EV)	41,450	40,255	-55	-0.11%
median price (GV)	36,417	36,419	2	0.00%
total sales	4,064,702	4,067,633	2,931	0.07%
EV sales	146,347	158,061	11,714	8.00%
EV sales (no car HH)	41,503	57,789	16,286	39.24%
EV sales (with GVs)	80,709	72,531	-8,178	-10.13%
EV sales (with EVs)	24,135	27,741	3,606	14.94%
GV sales	3,918,355	3,909,572	-8,783	-0.22%
GV sales (no car HH)	717,763	712,085	-5,678	-0.79%
GV sales (with GVs)	2,961,448	2,958,417	-3,031	-0.10%
GV sales (with Evs)	239,143	239,070	-74	-0.03%

Note: this table shows the counterfactual market outcomes after the federal rebates are heterogenous for different households: I raise the federal tax rebates by 25% for households with no car, and reduce the federal tax rebates by 40% for households with GVs, relative to the original federal tax rebates (households with EVs still get the original federal tax rebates).

Table 18: Simulation Two: Case two: welfare analysis under heterogenous subsidy

	current subsidy	new subsidy	difference
total welfare (10K USD)	22,865,375	22,843,225	-20,615
CS	17,603,145	17,568,566	-34,579
CS (no car HH)	5,355,604	5,371,440	15,837
CS (with GVs)	11,015,049	10,962,412	-52,636
CS (with EVs)	1,232,493	1,234,713	2,220
PS	5,160,329	5,173,526	13,197
PS (for GV)	4,894,583	4,890,320	-4,263
PS (for EV)	265,746	283,206	17,460
total subsidy	101,900	101,133	-767
subsidy (no car HH)	29,013	51,623	22,610
subsidy (with GVs)	55,976	29,880	-26,096
subsidy (with EVs)	16,911	19,630	2,719
avg subsidy	0.6963	0.6398	-0.0565

Note: this table shows the counterfactual welfare changes after the federal rebates are heterogenous for different households: I raise the federal tax rebates by 25% for households with no car, and reduce the federal tax rebates by 40% for households with GVs, relative to the original federal tax rebates (households with EVs still get the original federal tax rebates).

8 Conclusion

The issue of electric vehicle (EV) adoption has been a subject of substantial debate in both economic research and policy discussions. Most research has concentrated on vehicle adoption without accounting for the intricacies of households' vehicle holding. Notably, most vehicles in the US are owned by multi-car households. This paper develops a structural model to estimate the preference heterogeneity induced by vehicle holdings in demand for new vehicles, employing data from household vehicle surveys and market sales. In addition to income heterogeneity documented in the literature, whether the household owns a car and, if so, its fuel type, also plays a substantial role in household automobile choices between GV and EV, given other car characteristics the same.

In the counterfactual experiments, I compare the market outcomes and welfare implications under the models with and without preference heterogeneity induced by vehicle holding, particularly in scenarios where EV market subsidies are removed. The results reveal that both models yield analogous outcomes in terms of overall EV sales and welfare implications. However, it is essential to note that predictions diverge across different households within these two models. I further experiment with different heterogenous subsidy schemes and propose a better subsidy scheme aimed at promoting EV adoption. This analysis suggests that while total welfare and EV sales can be enhanced to some extent without necessitating a larger subsidy budget, a more efficient approach involves redistributing subsidies from households with GVs to those without any vehicles, which could boost EV sales by 8%, albeit with a minor decrease in overall consumer surplus. To achieve the same level of EV sales under the existing subsidy scheme, the government would need to allocate an additional $81.6 million to the subsidy budget, a sum analogous to a one-year state-level subsidy budget for EVs in California during the 2010s.

Admittedly, it is essential to acknowledge the limitations of this papar. First, the data used in this paper are solely from California, a state at the forefront of early EV adoption in the United States. A natural direction for future work lies in extending this investigation to incorporate data from other states or regions. Holland et al. (2016) show great spatial heterogeneity exists regarding the environmental benefits associated with EV promotion,[16] and

[16] Xing, Leard, and S. Li (2021) provide the environmental benefit analysis in nation wide.

the policies adopted in regions with higher reliance on fossil fuel-based electricity generation may diverge substantially, potentially including taxation rather than subsidies. However, the framework developed in my paper, which integrates the household preference heterogeneity induced by vehicle holdings, remains highly adaptable and can be readily extended to accommodate other geographic regions through adjustments to subsidy amounts among different groups.

Another inherent limitation of this study is its concentration on the static welfare implications of subsidy policies determined by exogenous vehicle holdings. In reality, household automobile choices are often characterized by sequential and dynamic decision-making processes, with current vehicle holdings reflecting past choices. Incorporating the temporal dimension into the analysis could provide valuable insights into the transitions from gasoline vehicles to electric vehicles. This endeavor necessitates access to more extensive and representative sales and consumer survey panel data, which are often scarce, particularly in the early stages of EV adoption. As a result, the modeling of household dynamic transition patterns is a prospective avenue for future research. However, it's important to underscore that the static analysis presented here can be regarded as evidence of the short-term responses in the market. Despite these caveats, my work opens up possibilities for research exploring intra-household transitions in automobile choices and other new technology adoption problems.

References

Abotalebi, Elnaz, Darren M Scott, and Mark R Ferguson (2019). “Can Canadian households benefit economically from purchasing battery electric vehicles?” In: *Transportation Research Part D: Transport and Environment* 77, pp. 292–302.

Archsmith, James et al. (2020). “Attribute substitution in household vehicle portfolios”. In: *The RAND Journal of Economics* 51.4, pp. 1162–1196.

Beresteanu, Arie and Shanjun Li (2011). “GASOLINE PRICES, GOVERNMENT SUPPORT, AND THE DEMAND FOR HYBRID VEHICLES IN THE UNITED STATES*”. In: *International Economic Review* 52.1, pp. 161–182. DOI: https://doi.org/10.1111/j.1468-2354.2010.00623.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-2354.2010.00623.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-2354.2010.00623.x.

Berry, Steven T. (1994). “Estimating Discrete-Choice Models of Product Differentiation”. In: *The RAND Journal of Economics* 25.2, pp. 242–262. ISSN: 07416261. URL: http://www.jstor.org/stable/2555829 (visited on 08/17/2023).

Berry, Steven, James Levinsohn, and Ariel Pakes (1995). “Automobile Prices in Market Equilibrium”. In: *Econometrica* 63.4, pp. 841–890. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/2171802 (visited on 07/21/2023).

— (1999). “Voluntary export restraints on automobiles: Evaluating a trade policy”. In: *American Economic Review* 89.3, pp. 400–431.

Conlon, Christopher and Jeff Gortmaker (2020). “Best practices for differentiated products demand estimation with pyblp”. In: *The RAND Journal of Economics* 51.4, pp. 1108–1161.

Davis, Lucas (2021). “Electric Vehicles in Multi-Vehicle Households”. In: *Work. Pap* 322.

Gandhi, Amit and Jean-François Houde (2019). “Measuring substitution patterns in differentiated-products industries”. In: *NBER Working paper* w26375.

Gentzkow, Matthew (2007). "Valuing new goods in a model with complementarity: Online newspapers". In: *American Economic Review* 97.3, pp. 713–744.

Hao, Yiran (2022). "On Electric Vehicle Demand, Price Competition and Environmental Policies". PhD thesis. University of Toronto (Canada).

Hardman, Scott et al. (2021). "A perspective on equity in the transition to electric vehicles". In: *MIT Science Policy Review* 2, pp. 46–54.

Holland, Stephen P et al. (2016). "Are there environmental benefits from driving electric vehicles? The importance of local factors". In: *American Economic Review* 106.12, pp. 3700–3729.

Jakobsson, Niklas et al. (2016). "Are multi-car households better suited for battery electric vehicles?–Driving patterns and economics in Sweden and Germany". In: *Transportation Research Part C: Emerging Technologies* 65, pp. 1–15.

Karlsson, Sten (2017). "What are the value and implications of two-car households for the electric car?" In: *Transportation Research Part C: Emerging Technologies* 81, pp. 1–17.

Ku, Arthur L and John D Graham (2022). "Is California's Electric Vehicle Rebate Regressive? A Distributional Analysis". In: *Journal of Benefit-Cost Analysis* 13.1, pp. 1–19.

Langbroek, Joram HM, Joel P Franklin, and Yusak O Susilo (2016). "The effect of policy incentives on electric vehicle adoption". In: *Energy Policy* 94, pp. 94–103.

Li, Jing (2019). "Compatibility and Investment in the US Electric Vehicle Market". In: *Unpublished manuscript, MIT.*

Li, Shanjun et al. (2017). "The market for electric vehicles: indirect network effects and policy design". In: *Journal of the Association of Environmental and Resource Economists* 4.1, pp. 89–133.

Ma, Shao-Chao, Jin-Hua Xu, and Ying Fan (2019). "Willingness to pay and preferences for alternative incentives to EV purchase subsidies: An empirical study in China". In: *Energy Economics* 81, pp. 197–215.

McFadden, Daniel (1974). *Conditional logit analysis of qualitative choice behavior*. New York [u.a.]

Mersky, Avi Chaim et al. (2016). "Effectiveness of incentives on electric vehicle adoption in Norway". In: *Transportation Research Part D: Transport and Environment* 46, pp. 56–68.

Morrow, W. Ross and Steven J. Skerlos (2011). "Fixed-Point Approaches to Computing Bertrand-Nash Equilibrium Prices Under Mixed-Logit Demand". In: *Operations Research* 59.2, pp. 328–345. ISSN: 0030364X, 15265463. URL: http://www.jstor.org/stable/23013173 (visited on 08/18/2023).

Muehlegger, Erich and David S Rapson (2018). *Subsidizing mass adoption of electric vehicles: Quasi-experimental evidence from California*. Tech. rep. National Bureau of Economic Research.

Palmer, Kate et al. (2018). "Total cost of ownership and market share for hybrid and electric vehicles in the UK, US and Japan". In: *Applied energy* 209, pp. 108–119.

Petrin, Amil (2002). "Quantifying the benefits of new products: The case of the minivan". In: *Journal of political Economy* 110.4, pp. 705–729.

Rapson, David S and Erich Muehlegger (2023). "The economics of electric vehicles". In: *Review of Environmental Economics and Policy* 17.2, pp. 000–000.

Springel, Katalin (2021). "Network externality and subsidy structure in two-sided markets: Evidence from electric vehicle incentives". In: *American Economic Journal: Economic Policy* 13.4, pp. 393–432.

Wakamori, Naoki (2015). *Portfolio considerations in differentiated product purchases: An application to the Japanese automobile market*. Tech. rep. SFB/TR 15 Discussion Paper.

Wang, Ning, Huizhong Pan, and Wenhui Zheng (2017). "Assessment of the incentives on electric vehicle promotion in China". In: *Transportation Research Part A: Policy and Practice* 101, pp. 177–189.

Wang, Ning, Linhao Tang, and Huizhong Pan (2017). "Effectiveness of policy incentives on electric vehicle acceptance in China: A discrete choice analysis". In: *Transportation Research Part A: Policy and Practice* 105, pp. 210–218.

Wee, Sherilyn, Makena Coffman, and Sumner La Croix (2018). "Do electric vehicle incentives matter? Evidence from the 50 US states". In: *Research Policy* 47.9, pp. 1601–1610.

Xing, Jianwei, Benjamin Leard, and Shanjun Li (2021). "What does an electric vehicle replace?" In: *Journal of Environmental Economics and Management* 107, p. 102432. ISSN: 0095-0696. DOI: https://doi.org/10.1016/j.jeem.2021.102432. URL: https://www.sciencedirect.com/science/article/pii/S0095069621000152.

Zambrano-Gutiérrez, Julio C et al. (2018). "The role of public policy in technology diffusion: the case of plug-in electric vehicles". In: *Environmental science & technology* 52.19, pp. 10914–10922.

Zhou, Yiyi and Shanjun Li (2018). "Technology adoption and critical mass: The case of the US electric vehicle market". In: *The Journal of Industrial Economics* 66.2, pp. 423–480.

9 Appendices

9.1 Zero Market Shares

For zero market share products, it could cause problems when applying the inversion step in S. T. Berry (1994) and S. Berry, Levinsohn, and Pakes (1995) since zeros are not applicable in logarithm. The dataset covers all new vehicle registrations for each market, so the observed zeros are not due to sampling error like from survey to local level. As McFadden (1974) and S. Berry, Levinsohn, and Pakes (1995), if consumer's choice is an independent draw form a multinomial distribution with a set of purchasing probabilities, then the market shares are the aggregation over the consumers' multinomial draws. Even when the consumer sample is the full population, the finite market size and small purchasing probabilities can cause zero observed market shares. Table I1 indicates taht 24.5% of market share data for GVs and 39.3% of market share data in EVs are 0. Zero market shares in GVs are similar over 2014-2016, while zero market shares in EVs are slightly increasing, ranging from 36.7% to 42.6%. This is caused by new EV models entering the market and generates more zeros in the data.

To deal with the zero market share issue, I use a parametric empirical Bayes or shrinkage estimator to smooth the market shares following J. Li (2019). The Bayes estimator will generate the positive posterior market shares for each alternative. Each market's empirical Bayes prior is generated using similar markets defined by the closest CBSAs in market size (*i.e.* total household number in each market). The concrete method is as follows:

The quantities purchased of each vehicle j in each market m, K_{jm}, are modeled as a draw from a binomial distribution with N_m trials and purchase probability s^0_{jm}. Here N_m is the total market size. The purchase probability s^0_{jm} are different for each vehicle and market and are drawn from a Beta prior distribution with hyperparameters λ_{1jm} and λ_{2jm}. That is,

Table I1: Unit Sales, Market Shares, and Empirical Bayes Posterior Market Shares

Variable	Mean	Std.Dev	Min	10%	Median	90%	% Zeros	N
GV sales	73.6	513.3	0.0	0.0	4.0	99.0	24.5	53,270
2014	69.1	454.9	0.0	0.0	4.0	96.0	24.4	18,235
2015	74.8	537.5	0.0	0.0	4.0	98.0	24.7	17,955
2016	77.0	545.0	0.0	0.0	4.0	103.0	24.4	17,080
EV sales	48.9	240.0	0.0	0.0	1.0	67.0	39.3	3,115
2014	53.9	253.1	0.0	0.0	2.0	72.6	36.7	875
2015	47.2	223.6	0.0	0.0	1.0	59.0	37.2	980
2016	46.7	243.2	0.0	0.0	1.0	68.0	42.6	1,260
GV observed mkt share	0.000135	0.000361	0	0	3.13E-05	0.000322	24.5	53,270
EV observed mkt share	6.85E-05	0.000177	0	0	1.17E-05	0.000169	39.3	3,115
GV posterior mkt share	0.000139	0.000357	3.06E-11	6.38E-06	3.49E-05	0.000321	0.0	53,270
EV posterior mkt share	6.96E-05	0.000173	3.07E-11	1.48E-06	1.59E-05	0.000161	0.0	3,115

Note: This table shows summary statistics of vehicle sales, observed market shares (observed market share), and estimates of empirical Bayesian posterior mean market shares (posterior market share). Each observation corresponds to outcomes for an available vehicle model, market (CBSA), and year based on data from IHS Markit from 2014 to 2016.
The top panel shows unit sales of GV by year, followed by unit sales of EV by year.
The bottom panel depicts observed market shares (observed mkt shares) for GV and EV and estimates of empirical Bayes posterior mean market shares (posterior mkt shares).

$$K_{jm} \sim Binomial(N_m, s^0_{jm}), \quad s^0_{jm} \sim Beta(\lambda_{1jm}, \lambda_{2jm}) \tag{18}$$

The posterior distribution of the purchase probability is also a Beta distribution,

$$s_{jm} \sim Beta(\lambda_{1jm} + K_{jm}, \lambda_{2jm} + N_m - K_{jm}) \tag{19}$$

with mean given by

$$\hat{s}_{jm} = \frac{\lambda_{1jm} + K_{jm}}{N_m + \lambda_{1jm} + \lambda_{2jm}} \tag{20}$$

The observed shares are simple fractions between observed sales and market size,

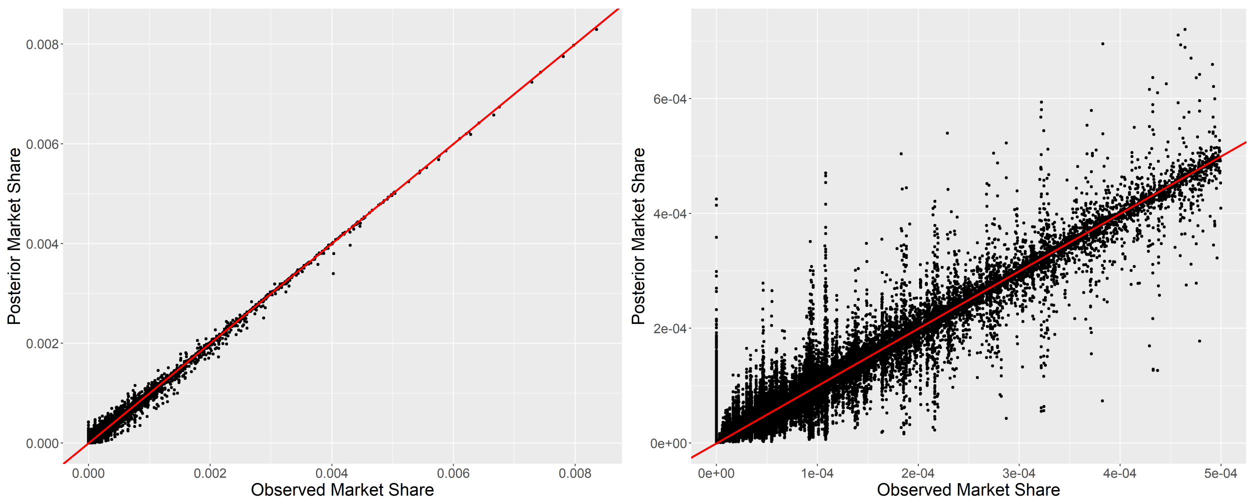
$$\hat{s}^{obs}_{jm} = \frac{K_{jm}}{N_m} \tag{21}$$

Since the hyperparameters λ_{1jm} and λ_{2jm} are strictly positive, the posterior mean is strictly positive. When the samples are large, the posterior would be much closer to the observed shares since the data provide more information than the prior distribution.

For each vehicle j in market m, the Beta prior is formed using the 10 markets closest in market size, $l \in \mathcal{B}_m$, where l is a market from the set of similar markets $\mathcal{B}_m$. Hyperparameters λ_{1jm} and λ_{2jm} are estimated from maximizing the log of the likelihood over the outcomes in the markets forming the priors,

$$f(K_{jl}, l \in \mathcal{B}_m | \lambda_{1jm}, \lambda_{2jm}) = \prod_{l \in \mathcal{B}_m} C(N_l, K_{jl}) \frac{\Gamma(\lambda_{1jm} + \lambda_{2jm})\Gamma(\lambda_{1jm} + K_{rl})\Gamma(N_l - K_{jl} + \lambda_{2jm})}{\Gamma(\lambda_{1jm})\Gamma(\lambda_{2jm})\Gamma(N_l + \lambda_{1jm} + \lambda_{2jm})} \tag{22}$$

The bottom panel of Table I1 summarizes the observed and empirical Bayes posterior market shares. All posterior market shares are positive, and the mean of the observed and empirical Bayes posterior market shares are quite similar, 0.0001312 and 0.0001349, respectively. Observed zero market shares have posterior mean estimates ranging from 3.063e-11 to 4.253e-04. Figure 4 presents the observed market shares against posterior market shares. Most points are near the 45-degree line.



(a) All market shares

(b) Zooming in on observed share ranging from 0 to 0.0005

Figure 4: Empirical Bayes Posterior Mean vs. Observed Market Shares

Note: This figure plots the empirical Bayes posterior means and the observed market shares. Posterior mean estimates may be larger or smaller than the original observed market shares, represented through the scatter plots as being above or below the 45-degree line (the red line). Subfigure (a) shows all data points. Subfigure (b) zooms into the smallest market shares.

9.2 Idenfication derivation

The probability for households with vehicle fuel type $v_0 \in \{\emptyset, GV, EV\}$ choosing option j of fuel type $v_1 \in \{\emptyset, GV, EV\}$ $s_i(v_1|h_i = v_0)$ is the summation of all vehicle j with fuel type v_1:

$$
\begin{aligned}
s_i(E|h_i=\emptyset) &= \sum_{j\in\{j\in J:fuel_j=E\}} \frac{exp(-\alpha p_j/y_i + x_j\beta + \beta_E + \sigma_E\nu_{i,fuel} + \xi_j)}{\sum\limits_{l\in J\cup\emptyset} exp(-\alpha p_l/y_i + x_l\beta + \sum\limits_{v\in G,E}(\beta_v + \sigma_v\nu_{i,fuel})1(fuel_l = v) + \xi_l)}, \\
s_i(\emptyset|h_i=\emptyset) &= \frac{1}{\sum\limits_{l\in J\cup\emptyset} exp(-\alpha p_l/y_i + x_l\beta + \sum\limits_{v\in G,E}(\beta_v + \sigma_v\nu_{i,fuel})1(fuel_l = v) + \xi_l)}, \\
s_i(G|h_i=\emptyset) &= \sum_{j\in\{j\in J:fuel_j=G\}} \frac{exp(-\alpha p_j/y_i + x_j\beta + \xi_j)}{\sum\limits_{l\in J\cup\emptyset} exp(-\alpha p_l/y_i + x_l\beta + \sum\limits_{v\in G,E}(\beta_v + \sigma_v\nu_{i,fuel})1(fuel_l = v) + \xi_l)}, \\
s_i(E|h_i=G) &= \sum_{j\in\{j\in J:fuel_j=E\}} \frac{exp(-\alpha p_j/y_i + x_j\beta + \beta_E + \Gamma_{GE} + \sigma_E\nu_{i,fuel} + \xi_j)}{\sum\limits_{l\in J\cup\emptyset} exp(-\alpha p_l/y_i + x_l\beta + \sum\limits_{v\in G,E}(\beta_v + \Gamma_{Gv} + \sigma_v\nu_{i,fuel})1(fuel_l = v) + \xi_l)}, \\
s_i(\emptyset|h_i=G) &= \frac{1}{\sum\limits_{l\in J\cup\emptyset} exp(-\alpha p_l/y_i + x_l\beta + (\beta_v + \Gamma_{Gv} + \sigma_v\nu_{i,fuel})1(fuel_l = v) + \xi_l)}, \\
s_i(G|h_i=G) &= \sum_{j\in\{j\in J:fuel_j=G\}} \frac{exp(-\alpha p_j/y_i + x_j\beta + \beta_G + \Gamma_{GG} + \sigma_G\nu_{i,fuel} + \xi_j)}{\sum\limits_{l\in J\cup\emptyset} exp(-\alpha p_l/y_i + x_l\beta + (\beta_v + \Gamma_{Gv} + \sigma_v\nu_{i,fuel})1(fuel_l = v) + \xi_l)}, \\
s_i(E|h_i=E) &= \sum_{j\in\{j\in J:fuel_j=E\}} \frac{exp(-\alpha p_j/y_i + x_j\beta + \beta_E + \Gamma_{EE} + \sigma_E\nu_{i,fuel} + \xi_j)}{\sum\limits_{l\in J\cup\emptyset} exp(-\alpha p_l/y_i + x_l\beta + (\beta_v + \Gamma_{Ev} + \sigma_v\nu_{i,fuel})1(fuel_l = v) + \xi_l)}, \\
s_i(\emptyset|h_i=E) &= \frac{1}{\sum\limits_{l\in J\cup\emptyset} exp(-\alpha p_l/y_i + x_l\beta + (\beta_v + \Gamma_{Ev} + \sigma_v\nu_{i,fuel})1(fuel_l = v) + \xi_l)}, \\
s_i(G|h_i=E) &= \sum_{j\in\{j\in J:fuel_j=G\}} \frac{exp(-\alpha p_j/y_i + x_j\beta + \beta_v + \Gamma_{EG} + \sigma_v\nu_{i,fuel} + \xi_j)}{\sum\limits_{l\in J\cup\emptyset} exp(-\alpha p_l/y_i + x_l\beta + (\beta_v + \Gamma_{Ev} + \sigma_v\nu_{i,fuel})1(fuel_l = v) + \xi_l)},
\end{aligned}
$$

Taking EV as example,

$$
\begin{aligned}
s_i(E|h_i=\emptyset)/s_i(\emptyset|h_i=\emptyset) &= \sum_{j\in\{j\in J:fuel_j=E\}} exp(-\alpha p_j/y_i + x_j\beta + \beta_E + \sigma_E\nu_{i.fuel} + \xi_j), \\
s_i(E|h_i=G)/s_i(\emptyset|h_i=G) &= exp(\Gamma_{GE}) \sum_{j\in\{j\in J:fuel_j=E\}} exp(-\alpha p_j/y_i + x_j\beta + \beta_E + \sigma_E\nu_{i.fuel} + \xi_j), \\
s_i(E|h_i=E)/s_i(\emptyset|h_i=E) &= exp(\Gamma_{EE}) \sum_{j\in\{j\in J:fuel_j=E\}} exp(-\alpha p_j/y_i + x_j\beta + \beta_E + \sigma_E\nu_{i.fuel} + \xi_j),
\end{aligned}
$$

By comparing the equations, the preference heterogeneity terms induced by vehicle holding can be expressed as

$$\begin{aligned}\Gamma_{GE} &= log\frac{s_i(E|h_i = G)}{s_i(\emptyset|h_i = G)} - log(\frac{s_i(E|h_i = \emptyset)}{s_i(\emptyset|h_i = \emptyset)}),\\ \Gamma_{EE} &= log\frac{s_i(E|h_i = E)}{s_i(\emptyset|h_i = E)} - log(\frac{s_i(E|h_i = \emptyset)}{s_i(\emptyset|h_i = \emptyset)})\end{aligned} \tag{23}$$

The terms for GV can be expressed in similar way.

9.3 More summary statistics

Table I3 summarizes vehicle in hand, income, family size, solar panel adoption rate, driving frequency, eduation, race, and employment for households of different new vehicles. The first column shows the characteristics of households who don't buy new vehicles. The second column shows the characteristics of households who buy a new GV, and the thrid column is for households who buy a new EV. The last column shows the characteristics of households who buy a new vehicle in general. Households of new veh purchases has higher rate of EV in hand, higher income, higher solar panel adoption rate, drive more frequently, higher education, and higher full time job employment rate than household not buying new vehicles. In particular, for EV purchasers, their high income rate is highest (35%) and the rate of EV in hand is much higher than GV purchasers. Comparing with households not buying new vehicle, households who buy a new vehicle (either GV or EV) have lower rate of a vehicle in hand. This is consisitent with decreasing marginal utility (need to control the family size here). The difference between keeping rate of GV and EV reflects the preference for different portfolios.

9.4 More tables in estimation results

Table I4 presents counterpart results under logit model rather than linear probability model. The results are similar as those in LPM in the sign and significance level. The

Table I2: demographic given vehicles in hand

(a) Mean family size

	Full Sample	none	EV	GV
not buy	2.446	1.944	2.570	2.442
buy GV	2.532	1.671	3.250	2.760
buy EV	2.585	1.000	2.750	2.645

(b) Education distribution

	Full Sample	none	EV	GV
College graduate (4-year degree)	0.314	0.270	0.323	0.314
Community college graduate (Associate degree, 2-year degree)	0.091	0.090	0.079	0.092
High school graduate/GED	0.082	0.036	0.045	0.086
Less than high school	0.018	0.018	0.011	0.019
Post-graduate degree	0.218	0.261	0.331	0.207
Post-graduate work	0.071	0.090	0.109	0.067
Some college	0.169	0.198	0.083	0.175
Technical school/professional business school	0.038	0.036	0.019	0.039
Total	1.000	1.000	1.000	1.000

(c) Ethnicity distribution

	Full Sample	none	EV	GV
American Indian or Alaska Native	0.012	0.018	0.015	0.011
Asian	0.137	0.180	0.117	0.137
Black or African American	0.033	0.000	0.011	0.036
Hispanic or Latino	0.108	0.054	0.056	0.114
Native Hawaiian or Other Pacific Islander	0.005	0.000	0.008	0.005
Other, please specify	0.017	0.009	0.030	0.016
Prefer not to answer	0.038	0.027	0.060	0.037
White	0.650	0.712	0.703	0.643
Total	1.000	1.000	1.000	1.000

Table I3: Average household characteristics in select groups

	not buy	GV	EV	new veh
Keep_G	0.92	0.74	0.77	0.74
Keep_E	0.07	0.06	0.14	0.07
Keep	1.00	0.79	0.91	0.81
Midinc	0.28	0.42	0.47	0.43
Highinc	0.09	0.13	0.25	0.15
Family size	2.42	2.68	2.56	2.67
Solar	0.14	0.22	0.30	0.22
Freq_driver	0.83	0.88	0.84	0.87
College	0.86	0.88	0.93	0.88
White	0.65	0.64	0.67	0.64
Fulltime	0.48	0.58	0.54	0.58

Source: California Vehicle Survey.
Note: Keep_G is a binary variable for the households with a GV in hand, and Keep_E is a binary variable for the households with an EV in hand. Midinc is the group with household income falls between \$100,000 and \$200,000. Highinc is the group with household income falls higher than \$200,000. Family size is the number of household members. Solar is a binary variable for the households with an solar panel installed on permanent residence. Freq_driver is a binary variable for the households is driving frequently (everyday). College is a binary variable for the households with college or more education. White is a binary variable for the white households. Fulltime is a binary variable for the households with full time job.

average marginal effect (AME) for variable *endow* is 13.3%, and AMEs for $hold_{GV}$ and $hold_{EV}$ are 19.3% and 27.1% in logit models.[17] The impact of vehicle holding on EV choice is slightly larger in logit model than in linear probability model.

[17] The marginal effects are calculated based on column (4) and column (7) in table I4.

Table I4: Logit regression of vehicle purchase and vehicle holding

	Dependent variable:					
	buyEV					
	Logit	Logit	Logit	Logit	Logit	Logit
hold	1.776**	1.970**	3.159***			
	(0.737)	(0.808)	(1.172)			
hold GV				1.615**	1.933**	3.222***
				(0.742)	(0.821)	(1.201)
hold EV				2.904***	3.339***	4.504***
				(0.837)	(0.989)	(1.397)
mid income		0.174	−0.090		0.023	−0.230
		(0.426)	(0.476)		(0.437)	(0.490)
high income		0.614	0.211		0.251	−0.095
		(0.504)	(0.578)		(0.542)	(0.607)
family size	No	Yes	Yes	No	Yes	Yes
Education	No	Yes	Yes	No	Yes	Yes
Ethnicity	No	Yes	Yes	No	Yes	Yes
CBSA FE	No	No	Yes	No	No	Yes
Observations	355	355	355	355	355	355
Log Likelihood	−122.248	−108.615	−86.556	−118.987	−105.689	−84.717
Akaike Inf. Crit.	248.496	255.230	261.112	243.975	251.378	259.435

Note: *p<0.1; **p<0.05; ***p<0.01
Dependent variable: whether to buy electric vehicle. Endow is indicator of whether the household endowed a vehicle. Mid income indicates whether the household annual income is between 100k dollars and 200k dollars. High income indicates whether the household annual income is higher than 200k dollars.

Table I5: Micro moment matching result

	prediction	survey
s_G^{keep0}	56.163%	65.766%
$s_G^{keep_G}$	7.823%	6.951%
$s_G^{keep_E}$	7.351%	6.015%
s_E^{keep0}	3.351%	1.802%
$s_E^{keep_G}$	0.375%	0.958%
$s_E^{keep_E}$	3.471%	3.008%