Title
Essays in Quantitative Marketing

Author
Sinyashin, Alexey

Publication Date
2022

Peer reviewed|Thesis/dissertation

Essays in Quantitative Marketing

by

Alexey Sinyashin

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Business Administration

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor J. Miguel Villas-Boas, Chair
Professor Przemyslaw Jeziorski
Professor Benjamin Handel
Professor James Sallee

Spring 2022

Essays in Quantitative Marketing

Abstract

Essays in Quantitative Marketing

by

Alexey Sinyashin

Doctor of Philosophy in Business Administration

University of California, Berkeley

Professor J. Miguel Villas-Boas, Chair

The dissertation has two chapters. In the first chapter, called "Optimal Policies for Differentiated Green Products: Characteristics and Usage of Electric Vehicles", I study the issue of policy design for electric vehicles. When designing policies for electric vehicles (EVs) policymakers need to decide how to allocate policy support among EVs with different characteristics, since different EVs are likely to have differences in attractiveness to consumers and usage patterns and, hence, differences in environmental impact. In this paper, I build and estimate a structural model of the U.S. auto market that is able to predict market shares and usage patterns of electric and traditional vehicles with different characteristics under various market conditions and is able to assess the effects of policies differentiating on characteristics of EVs. On the demand side, I introduce the concept of consumer inconvenience costs of charging, which depend on EV battery range, charging infrastructure development level, consumer's driving needs, and other individual-specific factors. On the supply side, I model firm choice of prices and battery ranges. The estimation results show that the inconvenience costs have a dramatic effect on consumer purchase decisions and usage patterns of EVs, and, hence, their environmental impact. Also, the results indicate that firms are more likely to adjust battery ranges when policy support depends on battery range. I use the model estimates to evaluate the effects of two major U.S. policies for EVs, the federal subsidy and California's Zero Emission Vehicle regulation, on the environment, consumer surplus, firm profits, and social welfare. I also

experiment with alternative structures of the federal subsidy that differentiate on type and battery range of EVs. I find that more efficient structures can improve the environmental effect of the subsidy by 4.6% and the welfare effect by 1.6%. Interestingly, the more efficient structures result in fewer EVs sold, but in more electric miles traveled and more gasoline miles replaced.

In the second chapter, called "Do Big Businesses Influence Media? The Case of Amazon.com and The Washington Post.", I study whether media outlets bias their coverage of the news about their owners or companies the owners have vested interests in. To shed some light on this question, I look at how the acquisition of the Washington Post, a major U.S. daily newspaper, by Jeff Bezos, the founder and CEO of Amazon.com, affected the coverage of the news about Bezos and Amazon.com. Using data on news stories in several major newspapers, I document that the acquisition resulted in an increase in the number of mentions of Bezos, Amazon, or Amazon's products in the Washington Post, relative to other newspapers and news stories about other big tech companies. From a simple sentiment analysis, however, I found no evidence of change in the sentiment of the stories. I discuss potential mechanisms that can explain the results, including a conflict of interest, a shift in preferences of the readership, improved access of the Washington Post to information about Bezos and Amazon, and a shift in the beliefs of the newspaper's editors and journalists about the importance of news about Amazon and Bezos.

Acknowledgments

I am grateful to my advisor J. Miguel Villas-Boas for the continuous support, encouragement, and responsiveness during my PhD journey. I also want to thank James Sallee, Przemyslaw Jeziorski, Benjamin Handel, and Kei Kawai for support and helpful discussions. This work also benefited from valuable comments from the marketing group.

I thank the Institute for Business Innovation at Haas for the generous support of the data purchase used in Chapter 1.

Finally, I am grateful to my wife, my son, and my parents who have always been with me and have always supported me.

Chapter 1

Optimal Policies for Differentiated Green Products: Characteristics and Usage of Electric Vehicles

1 Introduction

Many countries use a variety of policies to encourage adoption of electric vehicles (EVs) in order to reduce local air pollution and overall greenhouse gas (GHG) emissions from the transportation sector, one of the biggest GHG emission contributors.[1] Designing policies for EVs, e.g., purchase subsidies, is complicated by the fact that an electric car is a differentiated product and EVs with different characteristics may potentially have different attractiveness to consumers, driving patterns, and, hence, environmental benefits. To account for these differences, some policies do differentiate among EVs with different characteristics.[2] In particular, policy support often depends on two characteristics: type and battery (or electric) range of EVs, with the two most common types being battery-electric (or all-electric) vehicles (BEVs) and plug-in hybrid electric vehicles (PHEVs),[3] and the battery range being the maximum distance a car can travel on one full battery charge using only electric energy from the battery.

In this paper, I ask three research questions. First, how do characteristics of EVs, in particular, type and battery range, affect consumer willingness to buy and driving patterns of these vehicles? Second, how do policies differentiating on these two characteristics affect firm choice of battery ranges, a key characteristic of an EV? And, finally, how much should policies differentiate among BEVs and PHEVs with different battery ranges?

To address these questions, I build an empirical model that is able to predict market shares and usage patterns of electric and traditional vehicles, as well as firm choice of battery ranges under different market conditions and policy scenarios. Incorporating usage of EVs and traditional cars is an important element of the model: environmental damages from burning gasoline or generating electricity for charging EVs directly depend on how much con-

[1]In 2018, the transportation sector accounted for 28% of total U.S. GHG emissions with passenger cars and light-duty trucks being the largest contributors (59%) within the category (U.S. Environmental Protection Agency, 2020).

[2]Examples in the U.S. include the federal subsidy, the Zero Emission Vehicle regulation, and some state subsidies. More details are provided in Section 2.2.

[3]BEVs run purely on electric energy stored in built-in rechargeable batteries that can be charged from an electric outlet. PHEVs also have batteries that can be charged, but, in addition, they have back-up gasoline engines that turn on once the electric battery is depleted. Thus, PHEVs can be operated in two modes, electric and hybrid. Generally, when operated in electric mode, PHEVs are similar to BEVs, and when operated in hybrid mode they are similar to conventional hybrids. PHEVs should not be confused with conventional hybrid cars (e.g., Toyota Prius), which have much smaller electric batteries and cannot be charged from an external source of electricity.

sumers drive their cars. The previous literature studying the effectiveness of EV incentives largely ignored the aspect of usage, focusing instead on adoption rates of EVs. This paper fills this gap by modeling how various factors, including vehicle characteristics, availability of charging stations, heterogeneity in consumer driving needs and other consumer-specific factors, affect not only purchase rates but also usage patterns of different EV models.

What affects how consumers use electric cars and are EVs driven in similar ways to traditional cars? On the one hand, fuel and maintenance costs per mile are generally lower for EVs,[4] which should encourage higher mileage drivers to adopt electric cars and, hence, should result in higher than average mileage driven. On the other hand, long charging times,[5] limited electric ranges and limited availability of public charging stations may discourage consumer adoption of EVs, especially among higher-mileage drivers, who need to charge more often and, hence, experience more inconvenience of charging. This inconvenience may result in less than average driving, especially for BEVs. In the case of PHEVs, there is always a possibility of fueling up at a conventional gas station and running on gasoline, hence the inconvenience is likely to only limit the number of miles traveled on electricity. The existing evidence on usage of EVs (Davis 2019, Burlig et al. 2021, UC Davis PHEV Center 2020), including the data I use in this paper (California Air Resources Board, 2017), suggests that the inconvenience of charging is an important factor that negatively affects usage of BEVs and PHEVs in electric mode and therefore should be accounted for in the analysis.[6] To the best of my knowledge, this paper is the first to estimate the inconvenience costs and to study their relation to battery range, charging stations, and other factors.

The model is organized in the following way. On the demand side, consumers choose which car to buy. They are endowed with an expected number of miles they need to drive per year. Consumers make their decisions by maximizing a utility function that depends on vehicle characteristics, purchase price, expected annual fuel costs and two terms capturing inconvenience of charging BEVs and PHEVs, which are referred to as *inconvenience cost* terms. Each

[4]For example, in 2018, the U.S. average fuel cost per mile for BEVs was 4.7c, assuming residential electricity prices; for conventional hybrids it was 6.9c, and for gasoline cars it was 11c.

[5]The most common, level 2, chargers give 10-20 miles per hour of charging. DC fast chargers may give at least 60 miles in 20 minutes, but they are expensive to build and, hence, scarce. For more information, see https://www.epa.gov/greenvehicles/plug-electric-vehicle-charging. On the upside, a convenient feature of EVs is their ability to be charged at home.

[6]Section 3 presents more details on usage patterns of various EVs.

inconvenience cost term is an increasing convex function of expected annual mileage, in the case of BEVs, or expected annual mileage driven on electricity, in the case of PHEVs. Both terms also depend on battery range, development level of charging infrastructure and individual-specific factors. For PHEVs, consumers are assumed to optimally choose the fraction of miles they expect to drive on electricity, trading off the fuel cost savings and the PHEV inconvenience costs.

On the supply side, carmakers choose prices and battery ranges.[7] I model choice of battery ranges for two reasons. First, policy incentive schemes that depend on electric ranges or battery capacities are likely to affect producer decisions on what size batteries to put in the cars. Carmakers are likely to choose battery sizes carefully because battery cost is a substantial part of the marginal production cost of EVs.[8] Second, modeling battery choice allows me to more precisely estimate battery costs. This provides useful information about the evolution of the battery costs over time and about how close the production costs of EVs with different ranges are to those of traditional cars. In addition, the model estimates of the battery costs can be compared to the industry estimates, as an auxiliary check on the model.

I estimate the model in the context of the U.S. auto market from 2013 to 2018. The estimation results confirm that inconvenience costs of charging play an important role in purchase decisions of consumers with different driving needs. In the case of BEVs, electric range and charging station availability are important inconvenience costs determinants. For example, in 2018, for an average consumer, the inconvenience cost over the car lifetime was around \$7,000 for the Tesla Model 3 with a 310 mile battery range, while for the 151 mile Nissan Leaf it was around \$39,000. In California, a state with relatively well developed charging infrastructure, these numbers were much lower: \$800 for Model 3 and \$5,000 for Leaf. Consumers who actually buy these BEVs have much lower inconvenience costs than an average consumer: \$1,600 for Model 3 and \$3,000 for Leaf throughout the country, and \$600 for Model 3 and \$1,600 for Leaf in California. The difference in inconvenience costs between the actual buyers and an average consumer within the same geographical area is driven by heterogeneity in driving needs and other consumer-specific factors.

In the case of plug-in hybrids, according to the model estimates, the share of driving on electricity strongly depends on the PHEV electric range and the

[7] In fact, I will model battery choice only for BEVs produced by American manufacturers. More discussion is presented below.

[8] For example, UBS (2017) estimates the 2017 Chevy Bolt battery pack cost to be \$12,300, or 34% of its MSRP, and the 2018 Tesla Model 3 \$9,075 or 26% of its MSRP.

difference in the cost of an electric and gasoline mile, but charging station availability doesn't appear as an important factor, which is likely due to the presence of a back-up gasoline engine. Inconvenience cost levels are generally lower for PHEVs than in the case of BEVs, because PHEV consumers can control the level of inconvenience by adjusting how much they drive on electricity. Drivers of longer-range PHEVs have higher inconvenience costs than drivers of shorter-range PHEVs, on average, because the marginal benefit of charging is higher on average for longer-range PHEVs.

The supply side estimates of the battery costs in general agree with the estimates from the industry. I compute average battery costs per kWh by calendar year and compare them to the estimates from the surveys conducted by the Bloomberg New Energy Finance (BNEF).[9] According to the model predictions, the average battery cost dropped from $574 in 2013 to $214 in 2018, while the BNEF estimates are $650 in 2013 and $176 in 2018. Also, I look at expert estimates of battery pack costs that are available for some EV models and find that they are similar to the corresponding model predictions.

I use the model estimates to run two counterfactual exercises. In the first exercise, I study the effects of two existing programs, the federal subsidy for buyers of new EVs and California's Zero Emission Vehicle (ZEV) regulation. Both programs have comparable scales. One prominent difference is that, while the ZEV regulation has separate incentive schemes for BEVs and PHEVs and the amount of support it allocates depends on battery range, the federal subsidy does not distinguish between EV types and is essentially flat for BEVs. This difference plays an important role. First, I find that while the ZEV regulation has a significant impact on BEV ranges, resulting in range increases from 4% to 36%, with a larger effect for more affordable models, the federal subsidy program, with its flat structure, has almost no effect on BEV ranges. Second, relative to the federal subsidy, each EV added due to the ZEV regulation is driven on average more electric miles (14,638 vs. 11,962 miles) and generates more environmental benefits ($3,636 vs. $3,214). Besides the environmental effects, both programs improve consumer surplus and producer profits, with the welfare effect of the federal subsidy being $718M and that of the ZEV regulation being $416M per quarter.

In the second counterfactual exercise, I solve for the optimal federal subsidy structure that maximizes either environmental benefits or social welfare, holding the program budget fixed. I allow the subsidy structure to distinguish between BEVs and PHEVs and to depend piece-wise linearly on the battery

[9]see https://www.statista.com/statistics/883118/global-lithium-ion-battery-pack-costs/

range, similarly to the ZEV regulation. I find that, relative to the current subsidy, the subsidy that maximizes environmental benefits allocates more support to BEVs, with the subsidy size increasing in range, and less support to PHEVs, with PHEVs with less than 25 miles of range not getting any subsidy. The optimal subsidy results in 4% fewer EVs added by the program, but 4.6% more electric miles traveled by added EVs, and in an \$8.1M per quarter, or 4.6%, increase in the environmental benefit of the program. Relative to the subsidy that maximizes environmental benefits, the subsidy that maximizes social welfare gives less support to BEVs with small and intermediate ranges, but PHEVs with less than 25 miles of range now get a non-zero subsidy. Essentially, this scheme is focusing more on adoption of EVs with lower inconvenience costs, which is why it is resulting in a smaller environmental improvement, \$4.9M per quarter, or 2.8%. The welfare-maximizing subsidy results in \$11.2M per quarter, or 1.6%, welfare improvement. Also, similarly to the subsidy that maximizes environmental benefits, it results in fewer EVs added by the program (-1.7%) but more electric miles traveled by added EVs (+2.8%), which suggests that focusing on EV adoption rates can be misleading if the real goal is maximizing environmental benefits or social welfare.

Literature review. This paper relates to three strands of literature. First, there is a growing body of literature studying the effects of various government policies for EVs. DeShazo et al. (2017) and Muehlegger and Rapson (2021) assess California's EV rebate program. Li et al. (2017) and Springel (2021) build a two-sided market framework to evaluate the effectiveness of subsidies for EVs and charging stations in the U.S. and Norway, respectively. Li (2019) studies the effect of a charging standard compatibility mandate in the U.S. Xing et al. (2021) study what kind of vehicles EVs replace and assess alternative subsidy designs, targeting lower-income households. Remmy (2020) studies how subsidies affect battery ranges of EVs in the German market. Jenn et al. (2018) assess the effects of various EV policies in the U.S. and also provide a nice overview of papers studying incentives for conventional hybrid cars and EVs. My paper contributes to this literature by formally introducing the concept of consumer inconvenience costs of charging and modeling how this inconvenience affects consumer utilities and usage patterns of EVs with different characteristics under different policy scenarios and market conditions. In addition, I model battery range choice by firms,[10] and evaluate the effec-

[10] Remmy (2020) is a concurrent paper that also endogenizes EV battery choice, but for the German market. My approach is different in that I explicitly model the link between battery capacity and range, in that I make different assumptions about which manufacturers respond to local policy changes with battery updates, and in some other details.

tiveness of subsidy structures that differentiate on type and battery range of EVs.

Second, this paper contributes to the literature modeling vehicle demand and utilization (Mannering and Winston 1985, Goldberg 1998, West 2004, Bento et al. 2009, D'Haultfœuille et al. 2014, Grigolon et al. 2018 etc.). The paper argues that modeling usage of EVs requires taking into account the inconvenience of charging and proposes a model that is able to predict usage patterns of EVs consistent with the data. The demand side of the model builds on Berry et al. (1995) and Petrin (2002) and captures key trade-offs that consumers with different driving needs face when choosing among electric and gasoline vehicles.

Finally, this work adds to the literature on endogenous product positioning (Draganska et al. 2009, Eizenberg 2014, Wollmann 2018, Fan 2013, Remmy 2020, Crawford et al. 2019 etc.) by proposing an approach on how to endogenize choice of EV ranges by carmakers, taking into account technological details, in particular the link between battery capacity and range.

The paper proceeds as follows. Section 2 provides an overview of the U.S. market and policies for EVs, Section 3 provides a descriptive summary of average mileage driven by EVs and traditional cars. Section 4 presents the model. Section 5 describes the dataset. Section 6 discusses estimation and identification of model parameters. Section 7 presents the estimation results. Section 8 reports counterfactual analysis results and Section 9 concludes.

2 Market and Policy Overview

2.1 Market overview

The first modern mass-produced electric vehicles came to the U.S. market in the early 2010s as a result of tightening environmental regulation, government investments, and a series of improvements in the battery technology.[11] High battery production costs didn't allow carmakers to use batteries that are large enough for long-distance driving and, at the same time, keep the prices at a reasonable level. For example, an all-electric Nissan Leaf, one of the best selling EVs introduced in late 2010, had only 73 miles of battery range[12] and was priced at around $33,000 (before any subsidies). To facilitate long-distance driving, carmakers came up with the idea of a plug-in hybrid car, which lets

[11] A more detailed history of EVs can be found at https://www.energy.gov/articles/history-electric-car.

[12] All ranges are as rated by the U.S. Environmental Protection Agency (EPA).

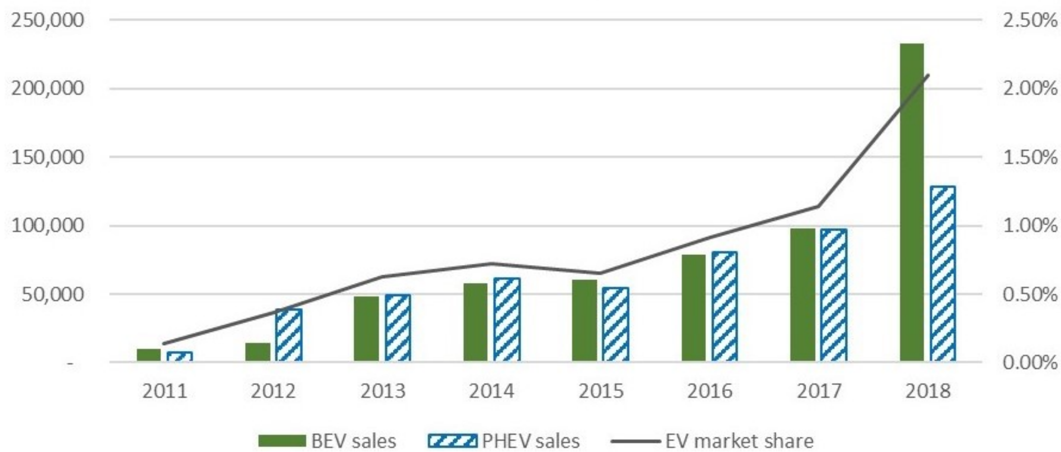


Figure 1: Sales of EVs over time in the U.S.

Notes: This figure shows BEV and PHEV sales by calendar year and combined share (BEV + PHEV) in the total sales of new passenger cars and light duty trucks. Data sources: U.S. DoE Alternative Fuels Data Center and Statista.com.

consumers do short trips using only energy from the battery and use a backup gasoline engine for longer trips. The first commercial plug-in hybrid car was the Chevrolet Volt, also released in late 2010, with 35 miles of battery range and a price tag of $40,280. In 2011, around 9,700 Leafs and 7,700 Volts were sold in the US, which comprised 98% of total EV sales and 0.14% of all new car sales in the U.S. (see Figure 1).

Over time, declining battery costs and development of charging infrastructure led to a significant expansion of EVs. Manufacturers introduced more models with various price and range options. Tesla Model S, introduced in 2012, was the first long-range mass-produced car with range options varying from 139 to 265 miles, priced at $57,400 - $77,400. More affordable long-range BEVs went on sale later: Chevrolet Bolt, 238 miles, priced at $36,620, in late 2016, and Tesla Model 3, 310 miles, priced at $46,500, in late 2017. Table 1 shows the evolution of EV models, prices and ranges for the 2013-2018 period. The average prices consumers paid for BEVs and PHEVs before any incentives[13] were quite stable over time, about $53,000 for BEVs and $37,000 for PHEVs, while the average ranges consumers were getting for these prices improved dramatically for BEVs, from 149 to 281 miles, and to some extent for

[13]For example, federal and state consumer subsidies may add up to $13,500 depending on various parameters such as state, time, EV type, range, MSRP etc.

Table 1: Evolution of the number of models, prices and ranges of EVs.

	2013	2014	2015	2016	2017	2018
BEVs						
N models	9	13	12	13	14	15
MSRP						
Mean	$50,861	$44,766	$53,388	$60,027	$53,896	$55,621
Min	$27,010	$25,560	$25,560	$25,510	$28,955	$29,120
Max	$79,900	$79,900	$85,000	$83,000	$83,000	$96,000
Range, mi						
Mean	149	129	159	183	224	281
Min	62	75	76	76	76	84
Max	265	265	270	270	310	335
PHEVs						
N models	6	10	11	18	26	30
MSRP						
Mean	$35,970	$34,642	$35,466	$38,899	$37,294	$38,606
Min	$32,000	$29,990	$29,990	$28,800	$27,120	$27,900
Max	$40,100	$76,400	$77,200	$78,700	$78,700	$99,600
Range, mi						
Mean	26.4	26.4	36.1	35.8	32.5	33.7
Min	11	11	11	11	9	9
Max	38	72	72	97	97	97

Notes: This table shows the evolution of the number of BEV and PHEV models, prices and ranges in the U.S. by calendar year. All mean values are sales-weighted.

PHEVs, from 26.4 to 33.7 miles. Overall, as Figure 1 shows, the EV sales had reached 2% of all new car sales by 2018, with BEVs and PHEVs contributing similarly until 2018. In 2018, right after its introduction, Tesla Model 3 gained 39% of all EV sales, making a huge contribution to BEV sales.

It is worth noting that for the period of the study the largest local market for EVs in the U.S. was California, with the state EV sales comprising about half of the national EV sales and the within-state market share of EVs reaching 8% in 2018. Several factors can potentially explain high adoption rates in California, including the state's strict environmental regulation, relatively high gasoline prices and "green" preferences of the population.

2.2 Policy Overview

In this paper, I focus on two types of policies for EVs in the U.S.: the federal subsidy and the Zero Emissions Vehicle (ZEV) regulation adopted in some states. These policies, and state subsidies for EVs, will be explicitly accounted for in the model.[14] I describe them in detail below.

Federal subsidy. The federal subsidy comes in the form of the federal income tax credit and is available to buyers of new EVs. The size of the subsidy depends on the battery capacity, which must be at least 4 kWh. The subsidy size formula is $2,500 plus $417 for each kWh over 4 kWh with a maximum value of $7,500. Although how the battery kWh capacity translates into the car range depends on other car characteristics, during the study period all BEVs and longest-range PHEVs received the full subsidy, while some smaller-range PHEV models got a partial subsidy.[15,16]

State subsidies. To further encourage consumer adoption of electric vehicles, some states introduced their own purchase subsidies in addition to the federal subsidy. During the study period, 16 states offered subsidies to consumers, with subsidy conditions varying across states and time. Depending on EV type, range, MSRP and household income, consumers were able to receive from $500 to $6,000 of a state subsidy in these states.

ZEV regulation. The ZEV regulation was designed and adopted by California in 1990 to achieve its long-term emission reduction goals by requiring auto manufacturers to sell zero emission vehicles. Later it was adopted by nine other states: Connecticut, Maine, Maryland, Massachusetts, New York, New Jersey, Oregon, Rhode Island and Vermont.[17]

The ZEV program requires manufacturers to earn a certain amount of credits each year by selling BEVs, PHEVs[18] and some other types of "clean" vehi-

[14] Other, not explicitly accounted for, policies for EVs include subsidies for charging equipment, access to carpool lanes, free parking etc. Some of these policies will be captured by the model implicitly through the fixed effects.

[15] For example, Chevy Volt (35-53 miles) and Honda Clarity PHEV (48 miles) received the full subsidy while Toyota Prius Prime (25 miles) was qualified for a $4,500 subsidy. The full list of credit amounts can be found at https://www.fueleconomy.gov/feg/taxevb.shtml

[16] For each manufacturer, the federal subsidy starts to phase out once the manufacturer has sold 200,000 qualified EVs. During the study period, Tesla and GM passed this mark, in July and November 2018, respectively. However, buyers of Tesla and GM EVs were able to receive the full credit during the entire study period.

[17] For more information on the ZEV regulation see https://ww2.arb.ca.gov/our-work/programs/zero-emission-vehicle-program

[18] PHEVs are referred as transitional zero emissions vehicles (TZEVs) in the ZEV regulation.

cles[19] in the states that adopted the ZEV regulation. For each manufacturer, the required number of credits per year is proportional to the total number of cars sold by this manufacturer in California on average each year, thus requiring bigger manufacturers to sell more EVs. The requirement is tightening over time. For example, the required number of credits was 4.5% of the total sales in 2018 and 22% in 2025. Each EV can earn more or less than one credit depending on its type and battery range. Prior to 2018, depending on the range and technology used, each PHEV was earning up to 2.5 credits and each BEV was earning up to 4 credits.[20] Starting from 2018, the number of credits has been proportional to the range and capped at 1.4 for PHEVs and at 4 per BEVs. In case a carmaker does not satisfy the regulation's requirements, it has to pay a $5,000 fine for each missing credit.

In order to incorporate the ZEV regulation into the model, I need to estimate the "market prices" of credits for BEVs, called ZEV credits, and credits for PHEVs, called transitional ZEV (TZEV) credits. These two type of credits have different values, with TZEV credits being less valuable because the regulation can be fully satisfied with ZEV credits and only partially satisfied with TZEV credits. Carmakers are allowed to trade ZEV and TZEV credits, hence they have some monetary values, which are, however, not publicly known. While we can observe how many credits were transferred between carmakers in a given year and state, information on the corresponding monetary transfers is not available. The only relevant publicly available information can be found in Tesla's quarterly financial reports prior to 2019, in which Tesla disclosed its revenues from selling ZEV credits.[21] I combine Tesla's revenues from ZEV credit sales and the number of ZEV credits transferred from Tesla to other manufacturers each year to estimate ZEV credit prices. I found that the average ZEV credit price was around $2,200 for the 2015-2018 period.[22] This number is very similar to the estimates by McConnell et al. (2019), who used the same methodology. I will assume that this price does not change

[19]For example, fuel cell electric vehicles (FCEVs). However, for the period of the study, the sales of other vehicle types that earn the same credits as BEVs and PHEVs were extremely small.

[20]More precisely, carmakers could earn up to 9 credits in the case of BEVs if they were able to demonstrate battery swapping ability. Only Tesla was able to do it and was earning from 5 to 7 credits per car in 2012-2014. However, after the regulation changed in mid-2014, no manufacturers were earning more than 4 credits.

[21]Since Tesla does not sell any gasoline cars, it's not subject to the ZEV regulation requirements. However, it still earns ZEV credits and sells them to other carmakers. According to its quarterly reports, Tesla earned around $165 million on average each year in 2013-2018 from selling ZEV credits.

[22]For earlier years, information on credit transfers is not available for some ZEV states.

over time, which is reasonable because carmakers can bank their credits, i.e., credits earned in a given year do not expire and can be used in the future to satisfy the ZEV program requirements.

Recovering TZEV credit prices is more complicated because Tesla earns and sells only ZEV credits and no other data on ZEV or TZEV credit revenues is available. However, I overcome this problem by analyzing credit transfers between carmakers where credits of different types were exchanged. For example, sometimes carmakers traded ZEV credits for TZEV credits. I found several such occasions. Assuming that they didn't involve any monetary transfers, I estimate the TZEV credit price to be about 30% lower than the ZEV credit price, or about $1,540. More details on the calculation of the ZEV and TZEV credit prices are provided in Appendix A.

To summarize, credits that carmakers earn by selling EVs in the ZEV states are estimated to have dollar values of up to $8,800 per car in the case of BEVs, up to $3,850 per car in the case of PHEVs before 2018, and up to $2,156 per car in the case of PHEVs starting in 2018. The exact values depend on car ranges. Given that about 30% of all new cars and 60% of all EVs in the U.S. are sold in the ten ZEV states, the ZEV regulation is likely to be an important driver of car manufacturer decision-making at the state and national level.

3 Miles driven by EVs

In order to understand the environmental benefits of EVs and to be able to make more informed policy decisions, it is important to look at how much BEVs with different characteristics are driven relative to traditional cars and what fraction of mileage PHEV drivers drive in the electric mode. Electric cars are typically cheaper to drive per mile than gasoline cars due to lower fuel and maintenance costs,[23] which should encourage more adoption among drivers who need to drive more. However, inconvenience of charging, related to long charging times, limited availability of charging stations, and limited battery ranges, may discourage adoption among the higher-mileage drivers, or, in the case of PHEVs, discourage driving in the electric mode. These two factors affect how EV adopters use their vehicles, as well as how much consumers are willing to substitute EVs for traditional cars.

The existing evidence on how much EVs are driven is scarce. Davis (2019) analyzes the data from the 2017 National Household Travel Survey (NHTS),

[23]See https://www.energy.gov/articles/egallon-how-much-cheaper-it-drive-electricity for price comparison of a gasoline gallon and an "e-gallon" in different states.

which includes a relatively small sample of adopters of early EV models, and finds that on average BEVs are driven 38% and PHEVs 24% fewer miles per year than an average gasoline car, with no data available on how much PHEVs are driven on electricity versus gasoline. Burlig et al. (2021) analyze changes in electricity consumption of households who recently purchased electric vehicles in California during the period between 2014 and 2017 and find that these changes indicate substantially lower usage of EVs (usage in the electric mode, in the case of PHEVs) relative to gasoline cars. The UC Davis PHEV Center (2020) conducted a representative survey of EV owners in California during the 2015-2018 period.[24] They document higher average usage of BEVs and PHEVs than Davis (2019). Using logger[25] data on 275 EVs they found that longer-range BEVs are driven more miles on average than shorter-range BEVs and that longer-range PHEVs are driven more electric miles than shorter-range PHEVs.

In this paper, I use another data source for usage of EVs, which has several models of EVs, includes vehicles in and outside California, and has a large sample size.[26] The average mileage numbers for Californian EVs from this dataset are similar to those from the UC Davis PHEV Center (2020). The data were collected by the California Air Resources Board (California Air Resources Board, 2017) directly from car manufacturers for electric vehicles nationwide. Seven car manufacturers submitted data for eleven EV models during the 2014-2016 period to the California Air Resources Board (CARB). The data summary is presented in Table 2. Although the individual car level data are not available, the report has data on average mileage that is, for some models, disaggregated by model-year and state (California vs. national), which allows me to construct a rich set of moment conditions for the model estimation.

To compare how much EVs are used relative to traditional cars, I add the vehicle usage data from the 2017 National Household Travel Survey (the 2017 NHTS, Federal Highway Administration 2017). The survey includes data on odometer readings and model-years of household vehicles, from which I compute annual mileage for each car. I focus only on passenger cars and pick-

[24]Their analysis is also complemented by a smaller nationwide survey, but no loggers were installed outside California, hence there is no data on PHEV usage on electricity versus gasoline for non-Californian vehicles.

[25]A logger is a device that is plugged into a vehicle to record data on various usage parameters.

[26]One concern about these data is that its representativeness cannot be confirmed. However, for some vehicle models the number of observations is close to the overall sales for a given model-year, which should alleviate the concern.

Table 2: EV usage summary statistics

Model	Type	Range, miles	N vehicles	Model years	Mean VMT, miles	Mean eVMT, miles	% eVMT
Tesla Model S	BEV	265	37,635	2012-2015	13,494	13,494	100%
Nissan Leaf	BEV	84	12,215	2011-2014	10,294	10,294	100%
Honda Fit	BEV	82	645	2012-2013	9,789	9,789	100%
Ford Focus	BEV	76	4,218	2012-2015	9,741	9,741	100%
BMW i3	BEV	81	4,193	2014-2015	7,916	7,916	100%
BMW i3 REX	PHEV	72	8,309	2014-2015	9,063	8,387	93%
Chevrolet Volt	PHEV	38	2,154	2011-2013	12,403	8,924	72%
Ford Fusion Energi	PHEV	20	12,842	2013-2016	15,076	4,776	32%
Ford C-MAX Energi	PHEV	20	10,253	2013-2015	13,920	4,574	33%
Honda Accord	PHEV	13	189	2012	15,221	3,246	21%
Toyota Prius Plug-in	PHEV	11	1,523	2013	15,283	2,304	15%

Notes: The table shows the number of vehicles in the sample, model-years of these vehicles, the average annual vehicle-miles traveled (VMT) and the average annual electric vehicle-miles traveled (eVMT). All miles of BEVs are electric, hence VMT and eVMT are equal for BEVs.
Data source: California Air Resources Board (2017). The data were collected in 2013-2016.

up trucks under five years old.[27,28] The CARB's report compares the mileage of EVs to the average mileage for gasoline cars computed using the California smog check data for cars with ages 0 to 4, which is about 14,600 miles. As a check, I compute the same statistic using the data I constructed from the 2017 NHTS and obtain a very similar number, 14,400 miles.

Figure 2 shows average mileage for EV models from the CARB data and average mileage for gasoline and conventional hybrid cars from the 2017 NHTS. Conventional hybrid cars are driven more miles on average than gasoline cars (15,800 vs. 16,600 miles), which is consistent with the hypothesis that higher-mileage drivers prefer more fuel-efficient vehicles. However, all the BEV models, despite their lower cost per mile, are driven fewer miles than gasoline and conventional hybrid cars. This is true even for a relatively long range (265 miles) Tesla Model S, with average mileage of 13,494. Other, smaller range

[27] The 2017 NHTS contains odometer reading data for a number of electric vehicles as well, but I decided not to use these data because the number of observations for each EV model is small and the exact purchase date is unknown, which makes estimation of average mileage imprecise. Moreover, the survey has no information on how much PHEVs are used in electric and gasoline modes. Davis (2019) used the 2017 NHTS data to analyze how much EVs are driven distinguishing only between BEVs and PHEVs.

[28] More details on how the vehicle usage data were constructed from the 2017 NHTS are provided in Appendix C.1.

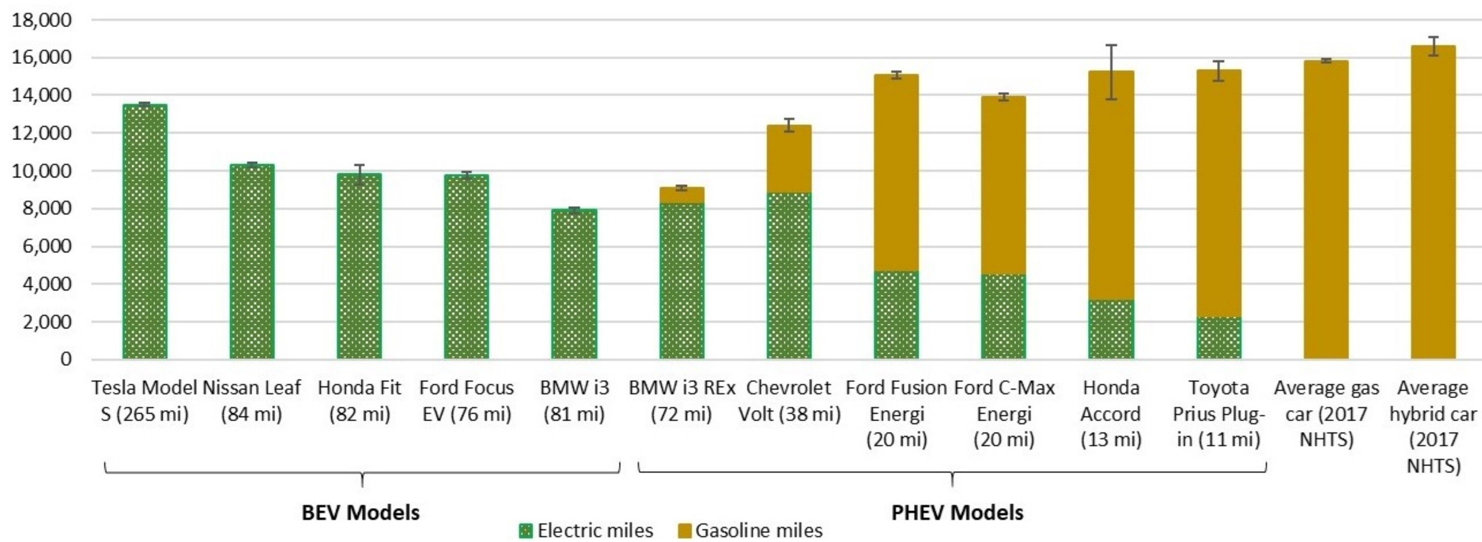
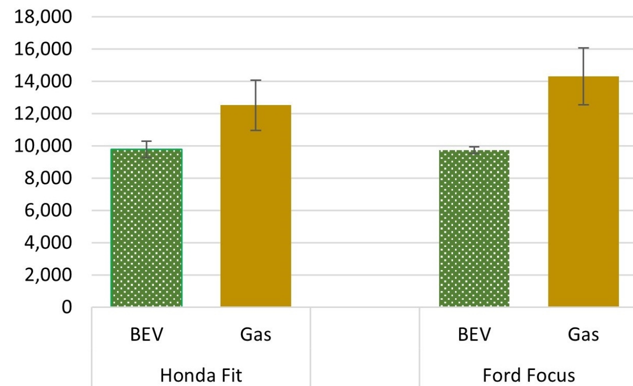


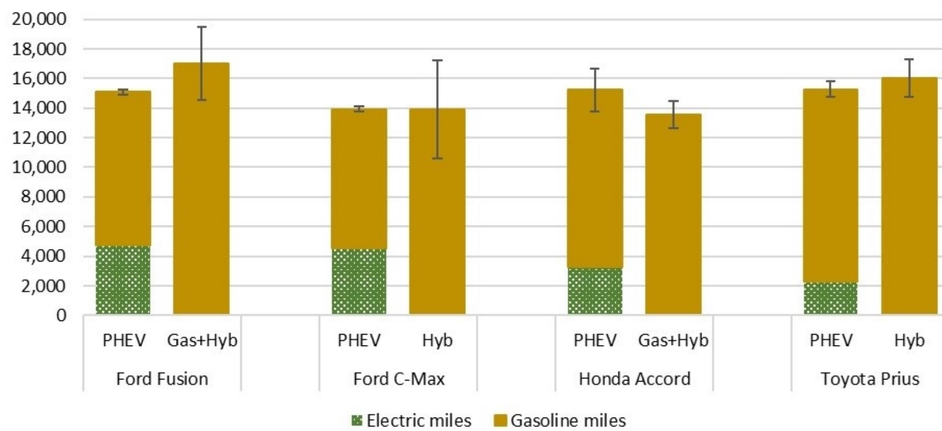
Figure 2: Average annual mileage of electric and gasoline/hybrid cars.

Notes: The figure shows average electric and total mileage of various BEV and PHEV models with battery ranges in parentheses. Average mileage of gasoline and hybrid cars are computed for vehicles with ages less than 5 years. 95% confidence intervals are presented. Confidence intervals for EVs are approximated using the number of observations and the 2017 NHTS data.

Data sources: California Air Resources Board (2017), 2017 NHTS.



(a) BEV vs. gas version of the same model



(b) PHEV vs. gas/hyb version of the same model

Figure 3: Average annual mileage of electric versus gasoline/hybrid versions.

Notes: Subfigure (a) compares mileage of BEV and gasoline versions of same model. Subfigure (b) compares mileage of PHEV and gasoline/hybrid versions of same model. 95% confidence intervals are presented. Confidence intervals for EVs are approximated using the number of observations and the 2017 NHTS data.
Data sources: California Air Resources Board (2017), 2017 NHTS.

BEVs (76-84 miles) are driven around 10,000 or even fewer miles per year on average.

Besides range and per-mile costs, other car characteristics can be important determinants of average mileage. Tesla Model S is a large sedan; other BEVs in the data are smaller, with BMW i3 being the smallest car.[29] For a closer comparison, Figure 3a looks at the electric and gasoline versions of the same model. There are two cars for which this comparison is possible: Honda Fit and Ford Focus. In both cases electric versions are driven significantly fewer miles, around 9,800 for both BEVs versus 12,500 for gasoline Honda Fit and 14,300 for gasoline Ford Focus.

In the case of plug-in hybrids, how much consumers drive in electric mode is strongly related to a car's battery range, as Figure 2 shows. While drivers of Chevrolet Volt with 38 miles of battery range drive on average almost 9,000 electric miles per year (72% of total annual mileage), 20-mile range Ford Fusion Energi drivers drive around 4,800 electric miles (32% of total mileage) and 11-mile range Toyota Prius Plug-in drivers drive only 2,300 electric miles per year (15% of total mileage).

From Figures 2 and 3b we see that the average total (gas plus electric) mileage of the smaller range PHEVs is comparable to the average mileage of their gasoline and conventional hybrid counterparts. However, the longer range PHEVs, Chevrolet Volt and BMW i3 REX,[30] seem to be driven less on average than gasoline cars. While, to some extent, this can be explained by the relatively small sizes of the cars, especially in the case of BMW i3 REX, another potential explanation is that consumers who purchase long range PHEVs intend to drive them mostly on electricity, which is harder for consumers who need to drive more. For a high-mileage driver, a conventional hybrid car can be a better choice because it is generally less expensive than a PHEV and may be more fuel efficient than long-range PHEVs in gasoline mode.[31]

To summarize, the evidence presented suggests that the inconvenience of charging is an important force that affects usage of EVs, with the battery range being a key factor affecting the inconvenience.[32] In the next section, I will introduce a structural model that will formally address how various factors affect the inconvenience of charging, and how consumers with different driving

[29] From the 2017 NHTS, I found that smaller gasoline cars are driven less on average than larger gasoline cars.

[30] With REX meaning "range extender", BMW i3 REX is a modification of BMW i3 that has a small gasoline engine and a small gasoline tank.

[31] Heavy batteries of long-range PHEVs make them less fuel efficient because they have to carry more weight.

[32] One can also think about other mechanisms that affect average usage of EVs. For example, if environmentally concerned consumers are more likely to drive less, e.g., because they are more likely to substitute driving with public transport, cycling or walking, and have higher preferences for EVs, this may result in lower average EV mileage. However, this doesn't explain the relationship between mileage patterns and battery ranges. Nevertheless, the structural model will address this issue.

needs and other characteristics choose which vehicle to buy.

4 Model

The model consists of two parts: vehicle demand and vehicle supply. On the demand side, consumers make static choices on which car to purchase. Besides purchase price and other vehicle characteristics, consumers account for the future fuel expenses and, in the case of EVs, the cost of inconvenience of charging. The two latter components depend on how many miles consumers need to drive. I assume that each consumer is endowed with a certain number of miles they expect to drive every year and this number is the same across the cars in the choice set. One concern about this assumption is that it ignores the possibility that consumers may drive more if they pay less per each mile traveled, a form of what is called the rebound effect.[33] Although there are papers that allow for non-zero elasticities of individual demand for driving with respect to fuel prices, car fuel economy, and other car characteristics,[34] I do not adopt that approach for several reasons. First, the data I use do not allow me to reliably identify these elasticities and, at the same time, fully account for heterogeneity in driving needs.[35] Second, a number of papers report small elasticities of demand for miles traveled with respect to fuel prices for gasoline cars, which are also declining over time (e.g., Mannering and Winston 1985, Goldberg 1998, Small and Dender 2007, Hughes et al. 2008, West et al. 2017). Third, given that, in the presence of the inconvenience costs, EVs are on average driven less than gasoline cars, the rebound effect does not seem to be a first-order issue in this case.[36]

[33]For a review of the literature on the rebound effect see Gillingham et al. (2015) and Linn (2016).

[34]These papers typically use the discrete-continuous framework of Dubin and McFadden (1984) and focus on gasoline/diesel cars. For example, see Mannering and Winston (1985), Goldberg (1998), West (2004), Bento et al. (2009), D'Haultfœuille et al. (2014).

[35]For example, some studies estimate the fuel price elasticity of miles traveled from observing monthly household gasoline consumption together with gasoline prices. I do not have this kind of data.

[36]Another aspect that can potentially be important and is ignored by the model is the possibility of usage substitution among vehicles of households with multiple vehicles. For example, households with a BEV and a gasoline car can potentially choose which vehicle to use for each trip, trying to maximize usage of the BEV to minimize fuel expenses. Hence, the problem becomes similar to that of a PHEV driver, who chooses the optimal share of driving on electricity, as will be described below. The model, in principle, allows households to have multiple vehicles, but it assumes that households are endowed with a fixed number of driving needs, which are also fixed, which means that households do not optimize over

On the supply side, carmakers maximize static profits by choosing prices and EV battery ranges. When deciding on battery ranges, carmakers are assumed to face the trade-off between consumer valuation of ranges and marginal production costs of batteries, holding other car characteristics constant. While this assumption seems to be realistic for BEVs, it is likely to be unrealistic for PHEVs because of the more complicated nature of the battery size choice problem. In addition to higher production costs of larger batteries, PHEV manufacturers face serious space constraints because they have to fit together electric equipment, including battery and motor, and gasoline equipment, including internal combustion engine and gas tank. Therefore, adding an extra kWh of battery energy can be quite challenging and is likely to affect other car characteristics. For this reason, I will model battery choice only for BEV models. To be even more realistic, I will focus only on BEVs produced by American manufacturers because, unlike foreign manufacturers, they are more likely to adjust battery ranges in response to changes in U.S. policies.

Both demand and supply sides of the model assume static behavior of agents, although dynamic aspects can potentially be important, given ongoing improvements in the EV battery technology. On the demand side, one can argue that consumers may have incentives to strategically postpone vehicle purchases, waiting for the arrival of more affordable and longer-range EV models, as well as for more charging stations to be built. However, the value of postponing is likely to be limited for several reasons. First, consumers may place a much larger weight on their current, rather than future, driving needs, when deciding when to buy a car. Second, consumers are likely to face a lot of uncertainty about when better EVs will arrive on the market and what their prices and characteristics will be, as well as whether more charging stations will be built in places where they need to drive, all of which should discourage forward-looking behavior. Third, there is a rich set of traditional vehicles to choose from at any point of time, which makes waiting less attractive relative to buying now. Even if, nevertheless, consumers are still strategically waiting for better EVs in the future, this should be captured by BEV and PHEV fixed effects in the model. As long as counterfactual policies do not have a noticeable effect on consumer beliefs about future EVs, the model should give valid

how much to drive each vehicle depending on which vehicle bundle they decide to own. The main reason for this approach is data limitations. The only source available to me that has information on vehicle bundles and mileage driven by each vehicle is the 2017 NHTS, which has few observations of households owning EVs, with these EVs being mainly early models. When estimating the demand, I tried to account for the effect of multiple vehicles by including the number of vehicles in the inconvenience cost terms (presented below), but the coefficients were not economically significant.

predictions.[37]

On the supply side, while modeling pricing decisions in a static manner seems plausible, battery range choices may potentially entail dynamic considerations in reality, if battery updates are costly. However, this may have a limited impact if battery ranges are updated mainly during scheduled vehicle redesigns. Also, as Wollmann (2018) points out, vehicle manufacturers tend to approximate the solutions to complicated budgeting and discounting problems by using solutions to simpler problems. The approach he uses suggests that firms make updates to their vehicle lineups if the ratio of the expected profit gain in the next period over the investment exceeds a certain threshold, i.e., firms, essentially, solve a static problem. Consistent with this approach, I model firm battery range choice as a static optimization over the next period, solving a range optimization problem only when I observe a vehicle entry or a battery update.[38] However, I do not attempt to recover firms' profit change thresholds for battery updates because of a limited number of update occasions in the data and concerns that update timing may coincide with scheduled vehicle redesigns. Hence, for the counterfactual analysis I will assume that firms always respond with updating battery ranges.

4.1 Demand

The demand-side model adopts a discrete choice random coefficient utility framework. Consumers arrive at the market and choose from one of the $j = 1, ..., J$ inside options or the outside option ($j = 0$). Consumer i expects to drive d_i miles per year, with d_i being independent of j. Her utility from car j in geographic market m is:[39]

$$U_{ijm} = u_{ijm} + \epsilon_{ijm} = x_j \beta_i^x - \alpha_i^p p_{jm}^c - \alpha_i^f fuelcost_{ijm} - BEV_j \cdot c_{bev}(d_i, r_j, chst_{jm}, v_i^b) \\ - PHEV_j \cdot c_{phev}(\psi_{ijm}^* d_i, r_j, chst_{jm}, v_i^p) + FEs + \xi_{jm} + \epsilon_{ijm}, \quad (1)$$

where x_j is a vector of K observed car characteristics, p_{jm}^c is consumer price, which is equal to MSRP minus federal and state subsidies, $fuelcost_{ijm}$ is ex-

[37]Another argument for validity of the static approach can be found in Springel (2021), who estimates a dynamic version of the demand model in addition to her main static specification. She finds no evidence of consumers strategically postponing EV purchases in Norway.

[38]When possible, I also check what happens if firms actually optimize over several future periods. I didn't find that the results would be much different.

[39]I drop time subscripts for ease of notation. A geographic market here is a state, with several exceptions.

pected annual fuel expenses and β_i^x, α_i^p and α_i^f are individual-specific coefficients on these terms; $c_{bev}(d_i, r_j, chst_{jm}, v_i^b)$ and $c_{phev}(\psi_{ijm}^* d_i, r_j, chst_{jm}, v_i^p)$ are BEV and PHEV inconvenience costs, which depend on annual mileage d_i, car electric range r_j, market m's charging infrastructure level of development $chst_{jm}$, unobserved individual-specific factors v_i^b and v_i^p and, for PHEVs, expected share of driving on electricity ψ_{ijm}^* (also called *the utility factor*); FEs is a set of brand and geographic market fixed effects; ξ_{jm} is unobserved vehicle-market taste, and ϵ_{ijm} is an i.i.d. type I extreme value error term. The mean utility of the outside option is normalized to 0.

I assume the following functional form of the taste parameters for characteristics, prices and fuel expenses:

$$\begin{pmatrix} \beta_i^x \\ \log \alpha_i^p \\ \log \alpha_i^f \end{pmatrix} = \begin{pmatrix} \beta^x \\ \alpha^p \\ \alpha^f \end{pmatrix} + BZ_i + \Sigma v_i, \quad v_i \sim N(0, I_{K+2}),$$

where Z_i is a vector of demographic characteristics (e.g., income, family size), B is a matrix of coefficients and Σ is a diagonal matrix of standard deviations of the elements of v_i.

The fuel cost term $fuelcost_{ijm}$ includes only one year of fuel expenses, hence the fuel cost sensitivity coefficient α_i^f incorporates discounting over the expected number of years of car ownership together with possible consumer myopia with respect to the future fuel expenses. For gasoline, conventional hybrid cars and BEVs expected annual fuel costs depend on local fuel costs per mile w_{jm} and the consumer's annual mileage d_i. For PHEVs, fuel expenditures also depend on the fraction of miles a consumer expects to drive on electricity, ψ_{ijm}^*, as well as costs per mile in electric and gasoline modes, $w_{e,jm}$ and $w_{g,jm}$. Hence,

$$fuelcost_{ijm} = \begin{cases} w_{jm} d_i & \text{if } j \text{ is gasoline, hybrid or BEV,} \\ \left[\psi_{ijm}^* w_{e,jm} + (1 - \psi_{ijm}^*) w_{g,jm}\right] d_i & \text{if } j \text{ is PHEV.} \end{cases}$$

The goal of including BEV and PHEV inconvenience cost terms is to explain the driving patterns of EVs presented in section 3. I assume that these two inconvenience cost functions are convex and increasing functions of mileage d_i, in the case of BEVs, and electric mileage $\psi_{ijm}^* d_i$, in the case of PHEVs.[40] I assume that other factors that can potentially affect consumer inconvenience

[40] Appendix D also presents a simplified demand model to give more intuition on how the inconvenience cost terms help rationalize the usage patterns of EVs.

include the car's electric range r_j, development level of the charging infrastructure $chst_{jm}$ and unobserved individual factors v_i^b and v_i^p, assumed to be independent standard normal. These individual factors may capture, for example, (in)convenience of charging station locations for a particular driver, access to charging infrastructure at work and at home, specific driving needs and patterns (e.g., long versus short trips) and, also, consumer "green" preferences, with more environmentally conscious consumers willing to incur higher inconvenience costs in order to make a lower environmental impact. For BEVs, I assume the following functional form of the inconvenience cost:

$$c_{bev}(d_i, r_j, chst_{jm}, v_i^b) = d_i^2 \cdot \exp(\theta_{b1} + \theta_{b2} \log r_j + \theta_{b3} chst_{jm} + \theta_{b4} v_i^b),$$

where $\theta_b = (\theta_{b1}, \theta_{b2}, \theta_{b3}, \theta_{b4})$ is a vector of parameters to be estimated.[41]

Unlike BEV drivers, PHEV drivers can optimally choose their inconvenience cost level by deciding on what fraction of their mileage to drive on electricity. In order to save more money on fuel costs or to minimize their environmental impact, PHEV drivers may want to exert more effort to drive more on electricity by adjusting their trip schedules and driving routes or spending more time at the charging stations. I assume that potential PHEV buyers evaluate all the relevant factors and form consistent beliefs about their optimal share of driving on electricity. In particular, they are assumed to solve the following problem:

$$\min_{\psi} \quad [\psi w_{e,jm} + (1-\psi) w_{g,jm}] d_i + \frac{c_{phev}(\psi d_i, r_j, chst_{jm}, v_i^p)}{\theta_{ph5}}.$$

The first term represents annual fuel expenses, which are decreasing in ψ, assuming $w_{g,jm} > w_{e,jm}$, i.e., driving on electricity is cheaper. The second term is annualized inconvenience cost (PHEV lifetime inconvenience cost divided by multiplier θ_{ph5}), which is increasing in ψ. Assuming that

$$c_{phev}(\psi d_i, r_j, chst_{jm}, v_i^p) = \theta_{ph5} (\psi d_i)^2 \exp(\theta_{ph1} + \theta_{ph2} \log r_j + \theta_{ph3} chst_{jm} + \theta_{ph4} v_i^p),$$

and denoting $\Theta_{ijm}^{ph} = \theta_{ph1} + \theta_{ph2} \log r_j + \theta_{ph3} chst_{jm} + \theta_{ph4} v_i^p$, the solution is given by:

$$\psi_{ijm}^* = \min \left[1, \frac{w_{g,jm} - w_{e,jm}}{2 d_i \exp(\Theta_{ijm}^{ph})} \right]. \tag{2}$$

[41] Likely, consumers anticipate that more charging stations will be built in the future. Hence, coefficient θ_{b3} captures consumer expectations about the number of charging stations in the future, given the current state of charging infrastructure development, $chst_{jm}$.

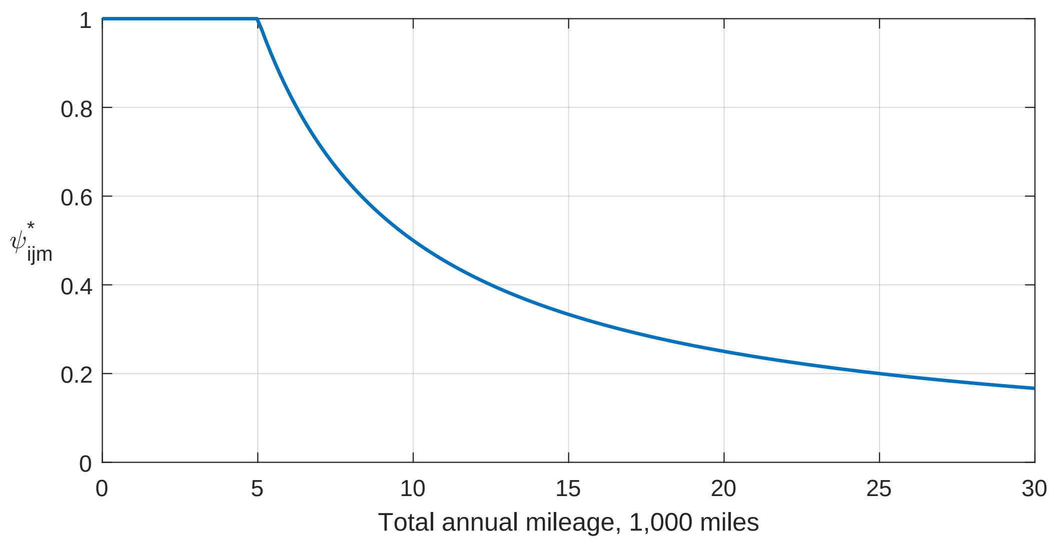


Figure 4: An example of PHEV optimal share of driving on electricity (the utility factor) as a function of the total annual mileage

Notes: This figure shows an example of a PHEV driver's optimal share of mileage driven in electric mode as a function of her total mileage, given by equation (2), for some set of parameters, holding everything else fixed.

Figure 4 depicts ψ^*_{ijm} as a function of annual mileage d_i. The optimal share of driving on electricity is decreasing in total mileage, reflecting that it is more difficult for higher-mileage drivers to drive a higher fraction of their miles in the electric mode. On the contrary, if d_i is small enough, all the driving can be done using only electricity, i.e., $\psi^*_{ijm} = 1$. Changes in fuel prices, car's electric range, charging infrastructure quality or unobserved consumer-specific factors can shift the curve, resulting in more or less driving on electricity for the same total mileage. The relationship between ψ^*_{ijm} and d_i predicted by (2) is consistent with Figure B1 from Appendix B, which shows a scatterplot of the individual utility factors for the PHEV models from the CARB data. From this scatterplot, it can be noticed that higher-mileage drivers tend to have a lower percentage of electric miles, for all the PHEV models.

Optimal shares of driving on electricity for PHEVs, ψ^*_{ijm}, are then plugged into PHEV inconvenience and fuel cost expressions, which are part of the utility specification (1). Product market shares are then given by familiar multinomial logit formulas:

$$s_{jm} = \int \frac{\exp(u_{ijm})}{1 + \sum_{l=1}^{J} \exp(u_{ilm})} dF(v^i) dF(z_i, d_i), \tag{3}$$

where $F(v^i)$ is the joint (multivariate normal) CDF of $(v_i,\, v_i^b,\, v_i^p)$ and $F(z_i, d_i)$ is the joint CDF of consumer demographic variables z_i and annual mileage d_i.

4.2 Supply

On the supply side, I model manufacturer choice of prices and BEV ranges, assuming other car characteristics are exogenously given. I assume that, while carmakers set prices every period, battery range decisions are made less frequently (e.g., during scheduled redesigns). The model first order conditions for battery ranges will be applied only when range updates are actually observed or when new BEV models are introduced. Also, I will limit my analysis of battery range decisions only to American manufacturers (Tesla, General Motors, Ford, Fiat-Chrysler), since they are likely to base their range decisions mainly on the U.S. market policies and conditions, and to those models of foreign manufacturers that were developed specifically for the U.S. market (e.g., Honda Fit EV and Toyota RAV4 EV). I assume that the timing of range updates and new model introductions is exogenous.

Each firm f, offering an exogenous set of products $\mathcal{F}_f$, is assumed to maximize the following static profit function:

$$\Pi_f = \sum_{j \in \mathcal{F}_f} \sum_m (p_j - mc_j + ZEV_{jm}) s_{jm} M_m,$$

where s_{jm} are market shares given by (3), M_m is the size of geographic market m, p_j are car prices (MSRPs), mc_j are car marginal costs and ZEV_{jm} are dollar values of ZEV credits that carmakers earn in the ZEV states for selling electric cars ($ZEV_{jm} = 0$ if m is not a ZEV state or j is not an EV). Assuming ZEV credit dollar values are the same across the ZEV states, let $ZEV_j = ZEV_{jm}$ if m is a ZEV state. The first order conditions with respect to price and battery range are then given by:

$$\frac{\partial \Pi_f}{\partial p_j} = S_j + \sum_{k \in \mathcal{F}_f} (p_k - mc_k) \frac{\partial S_k}{\partial p_j} + \sum_{k \in \mathcal{F}_f} ZEV_k \frac{\partial S_{zev,k}}{\partial p_j} = 0, \tag{4}$$

$$\frac{\partial \Pi_f}{\partial r_j} = -\frac{\partial mc_j}{\partial r_j} S_j + \sum_{k \in \mathcal{F}_f} (p_k - mc_k) \frac{\partial S_k}{\partial r_j} + \frac{\partial ZEV_j}{\partial r_j} S_{zev,j} + \sum_{k \in \mathcal{F}_f} ZEV_k \frac{\partial S_{zev,k}}{\partial r_j} = 0. \tag{5}$$

Here, S_j are sales of model j in all states and $S_{zev,j}$ are sales only in the ZEV states. The first two terms of the price equation (4) are standard and the third one reflects the effect of price changes on ZEV credit revenues through the impact on the market shares. According to the range equation (5), the battery range affects the profits in four ways: through the direct impact on marginal costs (first term) and consumer utility, hence market shares (second term), and through the impact on ZEV credit revenues, by affecting the credit amount per car (third term) and overall sales in ZEV states (fourth term).

Equations (4) and (5) allow us to recover marginal costs, mc_j, and marginal cost derivatives with respect to range, $\partial mc_j/\partial r_j$. For the counterfactual analysis, I need to specify a functional relationship between marginal costs and battery range. The main characteristic of a battery pack is its energy capacity, usually expressed in kWh, which does not depend on other car characteristics. In contrast, a battery range is a function of the battery energy capacity *and* the car's characteristics, such as weight, size and shape. For example, a battery pack with the same energy capacity would typically translate into more miles of range for smaller cars than for bigger cars because the latter consume more energy. For this reason, I expect less variation in battery costs per kWh than per mile across manufacturers and car models, and therefore I model marginal costs as a function of battery energy capacity rather than range. More specifically, I assume the following relationship between range and capacity:

$$r_j = \mu_{1j} \log(1 + \mu_2 b_j). \tag{6}$$

Here, range r_j is an increasing concave function of battery energy capacity b_j. This functional form allows me to capture decreasing returns to adding more kWh's of battery energy: an extra kWh increases the car's weight, thus increasing the car's energy consumption per mile traveled. The variables μ_{1j} and μ_2 are positive constants with μ_{1j} being model-specific, capturing differences in car sizes, shapes, body types, horsepower and other characteristics that affect cars' energy consumption. For the empirical application, I calibrate μ_2 using data on battery capacities and ranges of some BEV models that offer several range options.[42] I assume that $\mu_2 = 0.007$. After that, I compute μ_{1j} for each BEV model using their battery capacity and range specifications.[43]

The relationship (6) is used to convert the marginal cost derivatives with

[42]These are mainly Tesla models: S, X and 3.

[43]Equation (6) assumes that manufacturers know exactly how battery capacity translates into range, while in reality the range is learned only after regulatory agencies have conducted their range tests. Hence, I assume that carmakers are able to correctly predict the test results.

respect to range, $\partial mc_j/\partial r_j$, into derivatives with respect to battery capacity, $\partial mc_j/\partial b_j$:

$$\frac{\partial mc_j}{\partial b_j} = \frac{\mu_{1j}\mu_2}{1+\mu_2 b_j}\frac{\partial mc_j}{\partial r_j}.$$

Then, I assume the following simple marginal cost function:

$$mc_j = \lambda x_j + batcost_j + \varepsilon_j, \tag{7}$$

where x_j is a vector of car characteristics, $batcost_j$ is the cost of all EV battery components that vary with the battery energy capacity, which I will refer to as the battery pack cost,[44] and ε_j are i.i.d. errors.

I assume that the battery pack cost term, $batcost_j$, is a quadratic function of the battery capacity and time t:

$$batcost_j = (\gamma_0 + \gamma_1 t + \gamma_2 b_j + \gamma_3 t \cdot b_j + \eta_j)b_j,$$

where η_j is a model-specific error term. This specification allows me to capture two things. First, including interactions with the time trend t allows me to capture the substantial decline in battery costs per kWh over the period of the study. Second, including quadratic terms of battery capacity allows me to capture the nonlinear relationship between battery pack cost and battery capacity.[45] Marginal cost derivatives with respect to battery capacity are then given by

$$\frac{\partial mc_j}{\partial b_j} = \gamma_0 + \gamma_1 t + 2\gamma_2 b_j + 2\gamma_3 t \cdot b_j + \eta_j. \tag{8}$$

5 Data

I combine several data sources to estimate the model. First, the U.S. vehicle sales data for the calendar years from 2013 to 2018 come from IHS Markit/R.L. Polk and Co. It reports the quarterly number of new vehicle registrations by state, and by designated market area (DMA) for California. The data are broken down by brand, model name, model year, fuel type and, in some cases, model trim. For each car model included in the analysis, I distinguish between different fuel types (gasoline, hybrid, plug-in hybrid or all-electric)[46]

[44]Conventional hybrids have electric batteries, too, but they are usually much smaller than EV batteries. I assume that their costs are captured by conventional hybrid dummies.

[45]There is some evidence that larger battery packs have lower costs per kWh; see, e.g., The International Council on Clean Transportation (2019).

[46]The sales of diesel passenger cars are quite small, so I will not distinguish them from gasoline cars.

and assume the most popular model trim and model year for each fuel type at a given time point. To avoid dealing with zero market shares, I combine states with relatively small sales of EVs into several regional markets based on these states' geographic proximity and drop several states with low population and extremely low market shares of EVs. On the other hand, I split California, the state with the largest number of EVs in the U.S., into four markets with the centers in San Francisco, Sacramento, Los Angeles and San Diego. I end up with 34 geographic markets and 816 market-quarters.

Consumer choice sets are constructed by including all EVs available in a given quarter and state[47] and around 100 gasoline and conventional hybrid cars. To select these cars, I analyzed sales data in San Francisco DMA, a DMA with the highest share of EV sales in California, and picked top-selling passenger cars in each calendar year, assuming they are closest substitutes for EVs. The market size is defined as the total number of passenger vehicles and light duty trucks sold in a given market-quarter. This is a convenient assumption that allows me, first, to work with the empirical distribution of consumer driving needs without making parametric assumptions, which would be needed to recover the driving needs of those who opt out of a new vehicle purchase. Second, I do not need to make any assumptions about emissions generated by consumers who choose not to buy a new car.

The sales data are complemented with vehicle prices (MSRPs) and characteristics collected from two online sources: msn autos (`www.msn.com/en-us/autos`) and `www.fueleconomy.gov`. Information on federal and state subsidies for EVs is collected from online legislative records. Average gasoline and electricity prices are collected from the U.S. Energy Information Administration (EIA). I use residential electricity prices to compute EV charging costs since charging at home is common among EV drivers[48] and since I do not have more precise data on charging locations and costs of EV drivers. All prices and subsidies are converted into 2018 U.S. dollars using the BLS CPI.

The data on charging stations with exact locations and opening dates come from the U.S. Department of Energy Alternative Fuels Data Center. I construct a measure of the regional charging infrastructure development level by dividing the number of level 2 and level 3 charging outlets in a state or a region by the number of gasoline stations in the same state or region.

Information on vehicle mileage and consumer demographics completes the

[47] Sometimes carmakers do not offer their EVs nationwide but rather focus on some of the ZEV states or other selected regions.

[48] For example, according to `https://www.energy.gov/eere/electricvehicles/charging-home`, more than 80% of EV charging happens at home.

dataset. Data on average mileage for 11 electric vehicle models comes from California Air Resources Board (2017) and is described in Section 3. The 2017 National Household Travel Survey (Federal Highway Administration, 2017) contains data on demographics and travel behavior of about 130,000 U.S. households in all 50 states. I use this survey to construct a joint distribution of driving needs and consumer demographics for each geographic market and to compute aggregate statistics to build moment conditions for the estimation. I focus only on relatively new vehicles, with model years from 2013 to 2017, i.e., not older than five years at the time of the survey. I end up with a sample of 49,902 vehicles, including information on vehicle make, model, model year, fuel type, odometer reading and owner's location and demographics.

6 Estimation and Identification

Demand. The demand estimation procedure is similar to Petrin (2002). I estimate the parameters by minimizing a generalized method of moments (GMM) objective function using two sets of moment conditions. The first set are orthogonality conditions between the vector of instruments Z_{jm} and the structural econometric error term ξ_{jm}, i.e., $\mathbb{E}[Z_{jm}\xi_{jm}] = 0$. I assume that vehicle characteristics x_j are exogenously given, hence they are valid instruments for themselves, but prices p^c_{jm} are potentially correlated with unobserved vehicle characteristics and demand shocks included in ξ_{jm}.[49] To instrument for price, I use state subsidies for EVs, which vary across states, time and EV models.[50,51] One concern about using this instrument is that the subsidies may be set endogenously, in response to unobserved local preferences for electric cars. To address this issue, I include market fixed effects and their interactions with the EV dummy variable in order to capture local preferences and policies that do not vary over the period of the study, and that may be correlated with the subsidy amounts chosen by the states. The remaining concern is

[49]Notice that x_j does not include battery range and charging station variables, which enter only through the inconvenience cost terms and, in the case of PHEVs, indirectly, through the fuel cost term. Endogeneity of charging station entry was recognized by previous literature (Li et al. 2017, Li 2019, Springel 2021, and Shriver 2015, with the latter focusing on ethanol fueling stations). I address this issue by identifying the parameters related to charging stations and battery ranges using the second set of the moment conditions, described below.

[50]A rich set of demographic moment conditions described further below allows me to not use BLP instruments, commonly criticized for potentially being endogenous.

[51]Li (2019) also uses federal subsidies as a price instrument. I found that in my setting they perform poorly, likely because they do not vary across states and time.

that the exact timing of subsidy introduction or discontinuation, or changes in subsidy amounts, may be caused by unobserved changes in local preferences. However, anecdotal evidence suggests that this is unlikely because of the long-term nature of policy planning, with legislation procedures being unable to react promptly to slight changes in consumer preferences.

The second set of moment conditions matches consumer demographic and mileage statistics to their model predictions. First, using the 2017 NHTS data, I match average household income and household size of buyers of expensive (MSRP is over \$43,000), moderately expensive (MSRP is between \$26,000 and \$33,000) and least expensive (MSRP is below \$21,000) cars to identify the parameters on the interaction terms between price and these demographics. I also add moment conditions for average income and household size of hybrid, plug-in hybrid and all-electric car buyers, which helps with the identification of the interaction between fuel costs and these demographics.

Second, I construct the moment conditions for average mileage of various vehicle models. Using the CARB data, I construct 22 moment conditions for EVs, including average mileage for five BEVs, and average total and electric mileage for six PHEVs. For four of these models, I am able to construct the moment conditions separately for Californian and non-Californian drivers. Next, using the 2017 NHTS data, I form average mileage moment conditions for 43 gasoline and conventional hybrid models. Together with variation in fuel and purchase prices, quality of the charging infrastructure, and vehicle characteristics, including fuel consumption rate and electric range, these moment conditions help identify the parameters inside the utility terms that depend on consumer mileage, including the fuel expenses term and the BEV and PHEV inconvenience cost terms. Intuitively, consumers with different driving needs are trading off purchase prices, future fuel expenses and, in the case of EVs, inconvenience costs of charging. Changes in fuel prices, charging infrastructure or vehicle ranges affect both market shares and average mileage. Hence, knowing market shares and average mileage of vehicles with different prices, costs per mile and electric ranges, which are purchased in regions or time periods with different levels of charging infrastructure development, allows me to identify the corresponding parameters.

Supply. The supply side model is estimated separately from the demand side. First, demand side estimates are used to recover marginal costs and marginal cost derivatives with respect to range using firms' first order conditions (4) and (5). Next, supply side parameters (λ, γ) are estimated using equations (7) and (8), assuming $\mathbb{E}[\varepsilon_j \mid x_j, b_j, \eta_j] = \mathbb{E}[\eta_j \mid b_j] = 0$. Make dummies are included in equation (7) to capture correlation between observed characteristics and unobserved ones, such as quality.[52]

7 Results

7.1 Demand

Reduced form results. Before presenting the estimates from the full model, I first discuss performance of the state subsidies as a price instrument using simple OLS and IV Logit regressions. The second and third columns of Table 3 show that the absolute value of the price coefficient increases by more than 4 times, from 0.034 to 0.14, when the instrument is used. This translates to an absolute mean own price elasticity increase from 1.06 to 4.32. These results suggest that the price endogeneity problem is serious and the instrument plays an important role in correcting the issue. The last column reports the first stage results of the 2SLS IV procedure. The F-statistic of the instrument relevance test is very large, suggesting that the instrument is strong.

Table 3: OLS/IV Logit demand results

	OLS	IV	First Stage
Price	-0.034 (0.001)	-0.140 (0.016)	
State subsidy (Instrument)			-1.015 (0.063)
Characteristics	Yes	Yes	Yes
Make FEs	Yes	Yes	Yes
Market, Market×EV FEs	Yes	Yes	Yes
Time trend	Yes	Yes	Yes
N obs.	82,830	82,830	82,830
Mean own price elasticity	-1.06	-4.32	
First stage F-statistic			259.85

Notes: Heteroskedasticity robust standard errors are in parentheses. Prices are in $1,000. The second column reports simple OLS logit demand estimates. The third column reports 2SLS logit estimates using state subsidies as an instrument, with the first stage results shown in the last column.

[52] I also tried including make fixed effects in equation (8); this didn't affect the estimates significantly, and due to a relatively small number of observations in (8), I decided to drop them.

Vehicle characteristics. Table 4 reports the estimates from the full model, which, in addition to the reported parameters, also includes make, market, market×EV, and quarter dummies. The base coefficients on vehicle characteristics are statistically significant and have the expected signs. The EV indicator is interacted with the time trend, consumer mileage, and unobserved consumer heterogeneity term ($v_{1,i}$). The interaction with the time trend is included to capture the ongoing changes in preferences for EVs among consumers shopping for new vehicles. These changes may include trends in consumer environmental preferences, confidence in the new technology, awareness about EVs etc., i.e., trends that are common within the general population. Also, this term may capture changes in the composition of the new vehicle buyers. In particular, the negative sign of the coefficient may suggest that consumers with higher preferences for EVs were more likely to come to the market earlier in time. Next, the interaction with consumer mileage is included to capture the possible correlation between environmental preferences of consumers and their attitudes towards driving. For example, one may expect that environmentally concerned consumers have higher preferences for EVs (as well as other "green" products) and drive less on average, because, e.g., they prefer more environmentally friendly ways of transportation, like public transport, cycling or walking. If this is true, then we should expect the coefficient on the interaction between EV and mileage to be negative. The estimate of this coefficient is, indeed, negative, but statistically insignificant and very close to zero, indicating that this effect, if present, has a very limited impact. Finally, the interaction with the unobserved heterogeneity term is intended to capture consumer heterogeneity in preferences for EVs that is unrelated to how much consumers need to drive.

There are two more demographic interaction terms for vehicle characteristics I include in the final specification.[53] The first one is the footprint of the car interacted with log mileage. This term is included to capture the idea that consumers who need to drive more miles prefer bigger cars. The second term is the interaction between household size and minivan, which captures the preferences of large households for minivans.

Price coefficient. The parameter estimates of the price coefficient suggest that price sensitivity is decreasing in income, which is intuitive. The estimated mean own-price elasticity is equal to -6.26, which is somewhat higher in absolute terms than elasticities computed by Berry et al. (1995), which range from -6.5 to -3. However, given that the period of my study begins more

[53] I started with a larger number of interaction terms and eventually kept those that were consistently impactful.

Table 4: Demand estimates

	Parameter estimate	Standard error
Vehicle characteristics		
Constant	-9.706	0.360
AWD	0.185	0.094
Horsepower/Weight	34.655	3.521
Footprint (L*W), ft^2	0.0369	0.0043
Interior Volume, ft^3	0.0328	0.0030
SUV	0.443	0.084
Minivan	-30.942	0.294
Hybrid	-2.121	0.079
PHEV	-1.170	0.122
BEV	0.592	0.134
Time	-0.0507	0.0028
EV × Time	-0.0607	0.0053
EV × Mileage	-0.0085	0.0091
EV × $v_{1,i}$	0.202	1.005
Footprint × Log Mileage	0.0276	0.0018
Minivan × Log Household Size	18.696	1.033
Log coefficient on Price		
Const.	-0.864	0.077
Income	-0.120	0.0062
$v_{2,i}$	0.017	0.336
Log coefficient on Fuel Expenses		
Const.	-1.010	0.069
Income	0.232	0.0047
BEV Inconvenience Cost term		
Const.	9.354	0.416
Log Battery range	-2.530	0.138
Charging stations	-1.676	0.319
v_i^b	1.250	0.101
PHEV Inconvenience Cost term		
Const.	-2.560	0.667
Log Battery range	-0.998	0.218
Charging stations	-0.075	0.065
v_i^p	0.641	0.398
PHEV Cost multiplier	6.942	3.840
N obs.		82,830

Notes: The table shows GMM parameter estimates of the utility function (1). The specification also includes make, market, market×EV, and quarter dummies. Asymptotic standard errors are reported. $v_{1,i}$, $v_{2,i}$, v_i^b, and v_i^p are i.i.d. standard normal. Mileage, income and household size are drawn from a joint empirical distribution, built from the 2017 NHTS.

than 20 years after the last year of the data used in Berry et al. (1995), this difference can be due to some significant changes that happened in the U.S. auto industry, such as an increased number of products and more intensive competition.

Fuel cost coefficient. Unlike price sensitivity, the estimated fuel cost sensitivity is increasing in income, which echoes a result from a seminal work by Hausman (1979), who studied consumer tradeoff between purchase price and operating costs for room air conditioners and found that lower-income households had considerably larger implied discount rates for future utilization costs relative to higher income ones. Hausman (1979) argues that this result can be due to different marginal tax rates and availability of credit, as well as lack of savings and uncertainty of income streams of lower-income consumers.

In the case of passenger cars, understanding how much consumers discount future fuel expenses is important for designing policies aimed at reducing transportation emissions. For example, if consumers discount future fuel costs by a lot, then subsidizing purchase prices of more fuel efficient vehicles can be more effective than increasing gasoline tax rates (see, e.g., Grigolon et al., 2018). I follow Hausman's work and compute fuel cost discount rates implied by the estimates of my model for consumers from different income groups. I assume that consumers expect future gasoline and electricity prices to be equal to the prices at the moment of purchase (Anderson et al., 2013) and that consumers expect to drive the same number of miles every year in the future. Then the expected discounted fuel expenses can be written as:

$$G_{ijm} = \sum_{s=1}^{S} \frac{1}{(1+\delta_i)^s} fuelcost_{ijm},$$

where δ_i is the individual discount rate, $fuelcost_{ijm}$ is the expected annual fuel cost (see utility specification (1)) and S is the expected car lifespan, which is assumed to be 15 years. G_{ijm} has to be multiplied by the price coefficient α_i^p when it enters the utility function, hence the relationship between α_i^p and the fuel cost coefficient α_i^f is:

$$\alpha_i^f = \alpha_i^p \sum_{s=1}^{S} \frac{1}{(1+\delta_i)^s}$$

and δ_i can be recovered from the ratio of fuel cost and purchase price coefficients. I compute the averages of these ratios by income group to recover discount rates for the population of new vehicle buyers. The results, presented in Table 5, show that the discount rate varies greatly with income,

Table 5: Estimated discount rates by income group

Income group	Share of new car buyers	Implied discount rate
<$35,000	11.7%	42.9%
$35,000-$50,000	9.6%	23.3%
$50,000-$75,000	18%	15.3%
$75,000-$100,000	16%	8.9%
$100,000-$150,000	24.2%	1.7%
>$150,000	20.5%	-6.7%
All incomes		3.2%

Notes: The table uses average ratios of fuel cost and purchase price coefficients by income group to compute annual discount rates of future fuel expenses.

from 42.9% for households with income less than $35,000 to -6.7% for households with income over $150,000. In addition to the reasons mentioned by Hausman (1979), this result can also be explained by potential correlation of income with consumer myopia toward future expenses and by "green" preferences, with "green" preferences being related to utility gains from driving more fuel efficient vehicles because they are more environmentally friendly, on top of utility gains from lower future fuel expenditures.[54]

The last row of Table 5 reports the discount rate computed for all income groups. It is equal to 3.2%. The recent literature on consumer myopia about future fuel costs, including Grigolon et al. (2018), Sallee et al. (2016), Allcott and Wozny (2014) and Busse et al. (2013), assume a discount rate of 5-6%, estimate a myopia parameter,[55] and find that consumer undervaluation of fuel costs is moderate to absent on average. By comparing their assumed discount rates with the discount rate for all income groups computed here, I can conclude that my results imply no undervaluation, consistent with these papers, and even slight overvaluation of fuel costs on average. However, as was discussed earlier, consumer heterogeneity plays an important role. In the context of EVs, the results indicate that policies aimed at reducing purchase

[54]Notice that "green" preferences captured here are not specific to EVs. In the model, "green" preferences specific to EVs are captured by the inconvenience cost terms, BEV and PHEV dummies and market fixed effects interacted with the EV dummy.

[55]Estimating consumer myopia by income group requires making assumptions about discount rates, which are likely to be different for different income groups. This is out of the scope of this paper.

prices, e.g., purchase subsidies, can potentially be a more effective tool to encourage adoption among lower-income consumers relative to policies taxing usage of gasoline cars or subsidizing charging costs for electric cars.

BEV inconvenience costs. Coefficient estimates of the BEV inconvenience cost term from Table 4 indicate that both battery range and charging infrastructure are important factors in consumer inconvenience related to charging and driving a BEV. Also, the coefficient on the random term v_i^b suggests that there is considerable heterogeneity in the unobserved factors. To make economic sense of these estimates, I compute dollar values of the inconvenience costs for several BEV models with different battery ranges over time.

Figures 5a and 5b show evolution of average inconvenience costs nationally and in California for four BEV models: relatively inexpensive and short-range Nissan Leaf, expensive and long-range Tesla Model S, and two relatively affordable and long-range models introduced in 2017: Chevrolet Bolt and Tesla Model 3. Solid lines represent inconvenience costs conditional on purchase, i.e., those of actual buyers of these vehicles, and dashed lines are unconditional inconvenience costs, i.e., inconvenience costs of the general population of new vehicle buyers. First, these graphs indicate a large difference in inconvenience costs between BEV adopters and an average consumer. For example, from Figure 5a, the average inconvenience costs of Nissan Leaf buyers are estimated to fall from around \$7,000 in 2013 to \$3,000 in 2018, while average costs among all consumers, which is outside the plot, decreased from \$250,000 to \$39,000 at the same time.[56] The longer-range models show lower average inconvenience costs for both buyers and, especially, the general population. The difference between the two groups is much smaller than in the case of Leaf, indicating that, for longer-range BEVs, other factors start to play less important roles, but it is still considerable.

Limited availability of charging stations is an important factor in BEV inconvenience. This can be noted from negative time trends of the inconvenience costs graphs, especially those for the general population, since more charging stations were built over time. I note that some of this effect can be due to vehicle range updates. To get a clearer picture, Figure 5b shows inconvenience cost estimates for the same models in California, a state with

[56]One might think that the numbers for an average consumer are unrealistically large, but recall that the model assumes that consumers have to drive a given number of miles every year regardless of what car they purchase. For example, if a consumer needs to drive long distances regularly in places without adequate charging infrastructure, then buying a 75-mile Nissan Leaf would make it impossible to satisfy those driving needs. This will be captured in the model by a huge cost of inconvenience.

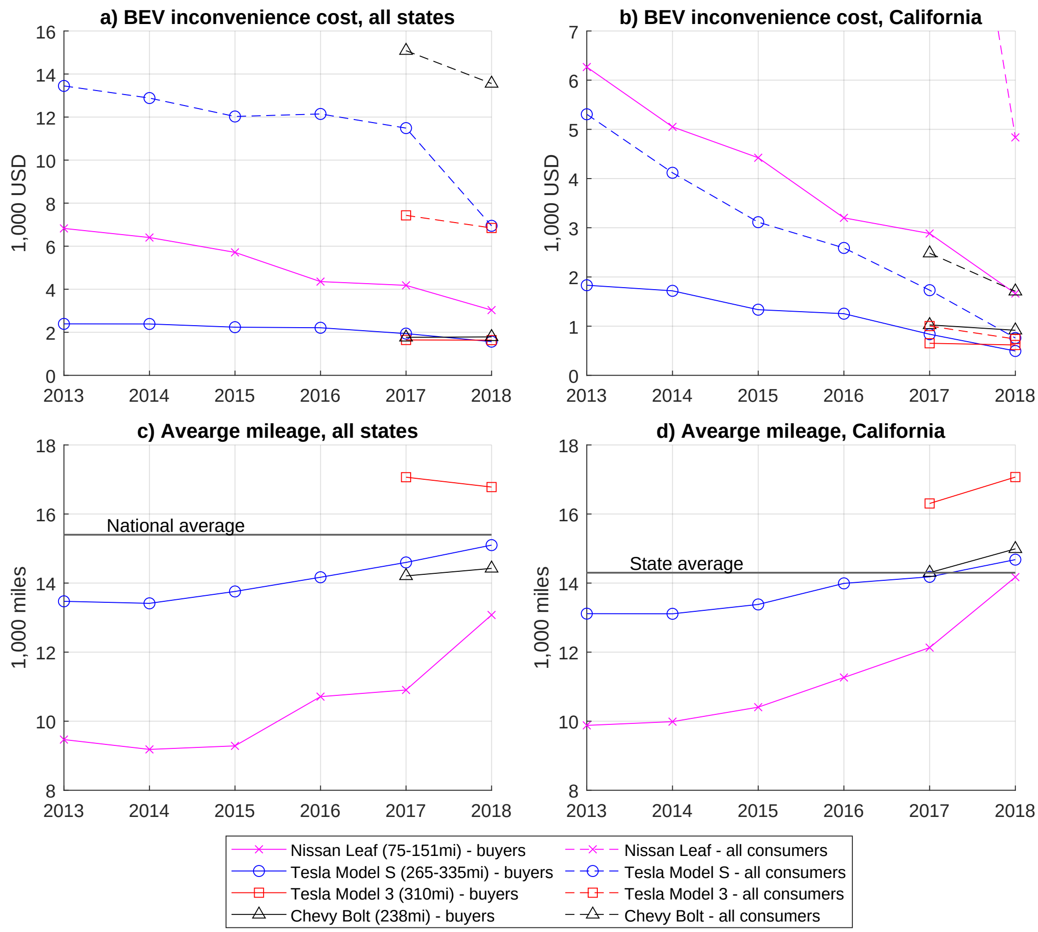


Figure 5: BEV inconvenience costs and average mileage predicted by the model
Notes: Figures a) and b) show the average estimated inconvenience costs for four BEV models in all states and California, respectively. Solid lines represent estimates conditional on purchase; dashed lines are average population estimates. Average population costs for Nissan Leaf are outside the plots; they decrease from \$250,000 in 2013 to \$39,000 in 2018 for all states, and from \$106,000 to \$4,800 for California. Figures c) and d) show average model predicted mileage for the same cars together with national and state averages for all passenger vehicles, computed from the 2017 NHTS. BEV ranges are in parentheses. Nissan Leaf range was updated in 2014 from 75mi to 84mi, in 2016 to 107mi and in 2018 to 151mi. Tesla Model S range is assumed to be updated in 2018 from 265mi to 335mi.

a relatively large number of charging stations.[57] The inconvenience costs in California are much lower for an average consumer, i.e., unconditional on purchase. This is especially prominent for shorter-range BEVs. For example, in 2018, the average inconvenience costs for the 151-mile Nissan Leaf are $39,000 nationally and $4,800 in California, while for the 335-mile Tesla Model S these numbers are around $7,000 and $700. The difference in inconvenience costs of actual buyers is not that large between California and all states. However, due to more charging stations, there are relatively more consumers with lower inconvenience costs in California. Noticeably, the inconvenience costs of buyers and an average driver in California have essentially converged by 2018 for BEVs with more than 300 miles of range (Tesla Model S and 3), but are still above zero, at the level of about $500-$1,000.

Figures 5c and 5d show model predicted average mileages for the same BEV models. The estimation objective function includes moment conditions for Model S and Leaf in 2013-2014, and the model predictions match these moment conditions well. Over time, as the inconvenience costs are falling, the average mileages of Model S and Leaf are increasing and get close to the state average in California by 2018. Nationally, though, Leaf mileage is still considerably below the national average. Interestingly, despite having a range similar to the Tesla Model S (310mi and 335mi), the Tesla Model 3 is driven about 2,000 miles more on average. Since Model S is much more expensive than Model 3 (over $80,000 vs. $46,500), Model S buyers have higher incomes on average and, hence, are less sensitive to purchase price and more sensitive to future fuel expenses. Thus, they do not need to drive many miles for a large enough difference in fuel costs relative to gasoline cars to justify the purchase of Model S. On the contrary, Model 3 buyers are relatively more price sensitive, hence, they need to drive relatively more miles to justify their purchase of Model 3. A similar argument explains why Chevrolet Bolt, which has a smaller range (238mi) and much lower price ($36,600) than Model S, is driven a number of miles that is similar to Model S.

PHEV inconvenience costs. Finally, the last portion of Table 4 reports parameter estimates of the PHEV inconvenience cost term. Similarly to BEVs, battery range is an important factor. However, charging stations do not seem to play an important role; the corresponding coefficient is neither statistically nor economically significant.[58] A possible explanation is that PHEV drivers

[57]According to the measure used in this paper, the level of charging infrastructure development in California is on average 5-6 times higher than in other states.

[58]This can be checked by plugging in values of the charging stations variable, whose mean is equal to 0.34 and the maximum value is equal to 1.67.

charge their cars mostly at home and, when needed, rely on the gas engine rather than public charging stations.

Understanding how much PHEVs are driven on electricity versus gasoline is crucial for understanding the environmental impact of plug-in hybrids and for policy design. Policymakers typically rely on assumed PHEV utility factors, i.e., proportions of electric mileage in total mileage driven by a typical driver. In my model, individual utility factors are given by ψ^*_{ijm} and the model is able to predict how these utility factors are changing under various circumstances, e.g., changes in battery ranges, fuel prices, market conditions etc., both for buyers and non-buyers of PHEVs. To illustrate, Figure 6a shows predicted average utility factors for three PHEV models with different battery ranges: Chevrolet Volt with 38 and then 53 miles of range, Ford Fusion Energy with 20 and then 26 miles of range and the plug-in version of Toyota Prius with 11 and then 25 miles of range. These predicted utility factors are markedly different for PHEVs with different battery ranges, e.g., in 2013, the average utility factor for a 38-mile Volt is predicted to be about 95%, for a 20-mile Fusion about 55%, and for an 11-mile Prius Plug-in about 20%. The model also predicts strong response to fuel prices: when gasoline prices dropped drastically in 2015, so did the utility factors. The model predicts that the difference in utility factors between buyers and non-buyers is relatively small, and sometimes non-buyers even have larger utility factors,[59] although, as shown in Figure 6b, buyers generally have lower inconvenience costs than non-buyers.

According to Figure 6b, overall levels of PHEV inconvenience costs are lower than those in the case of BEVs. This is also true for the difference between PHEV buyers and the general population. This is not surprising because consumers can control their level of inconvenience by adjusting utility factors. The optimal inconvenience costs are higher for longer-range PHEVs, because, on the margin, each unit of the inconvenience results in more electric miles driven and, hence, larger fuel cost savings for longer range PHEVs. Also, not surprisingly, when gasoline prices fall, driving on gasoline becomes relatively cheap and consumers choose smaller utility factors with lower levels of inconvenience.

[59]Overall, buyer and non-buyer utility factors are determined by multiple factors, including purchase and fuel prices, and fuel consumption of competitive cars and PHEVs in electric and gasoline modes. For example, lower mileage drivers may have larger utility factors, but, since they do not drive a lot of miles, their potential fuel cost savings do not justify paying a higher purchase price for a PHEV, which may result in the utility factor being larger on average for non-buyers.

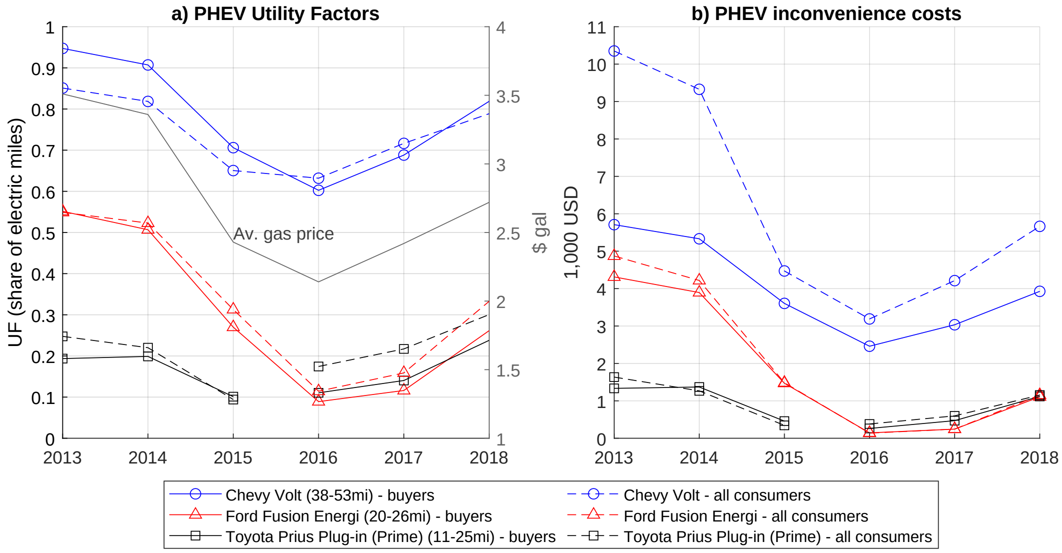


Figure 6: PHEV utility factors and inconvenience costs predicted by the model
Notes: This figure shows the average consumer utility factor (share of electric miles in total mileage) and inconvenience cost estimates for three PHEV models. Solid lines represent estimates conditional on purchase; dashed lines are average population estimates. Electric ranges are in parentheses. Chevy Volt range was updated in 2016 from 38mi to 53mi. Ford Fusion Energi range was updated at the end of 2018 from 20mi to 26mi. Toyota Prius Plug-in (11mi) was discontinued in 2015 and was reintroduced as Prius Prime (25mi) at the end of 2016.

7.2 Supply

Markups implied by the model generally match public financial disclosures.[60] For gasoline and conventional hybrids the average profit margin is predicted to be 19%, while for EVs it is 13%, if ZEV credits are included in the revenues, and 6%, if not. Lower markup estimates for EVs can be explained by their relatively high production costs and limited demand due to charging inconvenience. Interestingly, the model predicts that EVs that are not available outside the ZEV states are not profitable on average if ZEV credits are excluded from the revenues, with corresponding average margins equal to -0.2%, or -2.2% if weighted by sales. This finding suggests that manufacturers of these vehicles find it more profitable to sell these cars only in the ZEV states

[60]The estimates of the marginal costs are reasonable, too. Chevrolet Spark with an MSRP of about $13,000 has the lowest estimated marginal cost of about $8,700. Tesla Model X with an MSRP of $96,000 and 100kWh battery pack has the highest marginal cost estimate of about $86,000.

Table 6: Marginal cost parameter estimates

	Parameter estimate	Standard error
Specs except battery (λ)		
AWD	0.857	0.308
Footprint (L*W)	0.262	0.013
SUV	0.584	0.279
Minivan	-3.132	0.393
Gas engine HP	0.0751	0.0028
Electric motor HP	0.0493	0.0113
BEV	9.823	1.486
PHEV	5.940	1.279
Hybrid	3.126	0.161
PHEV battery cost (PHEV γ)		
kWh	0.915	0.265
t*kWh	-0.0509	0.0068
kWh^2	-0.0095	0.0096
t*kWh^2	0.0013	0.00034
BEV battery cost (BEV γ)		
kWh	0.842	0.0422
t*kWh	-0.0235	0.0019
kWh^2	-0.0051	0.00045
t*kWh^2	1.61e-4	2.07e-5
N obs.		2623
R^2		0.948

Notes: Make dummies are included in equation (7). Heteroskedasticity robust standard errors are reported.

at lower prices and make profits through earning ZEV credits, rather than to sell them nationally at higher prices and thereby earning fewer ZEV credits.

Equations (7) and (8) are used to estimate marginal cost parameters γ and λ. Table 6 reports the results. The coefficient estimates can be interpreted as partial effects measured in \$1,000. For example, adding one square foot to a car's footprint is estimated to cost on average \$262 and transforming a two-wheel drive vehicle into an all-wheel drive one costs \$857. Battery cost function parameters are estimated more precisely for BEVs than for PHEVs, because for BEVs the estimation is aided by the marginal cost derivative in

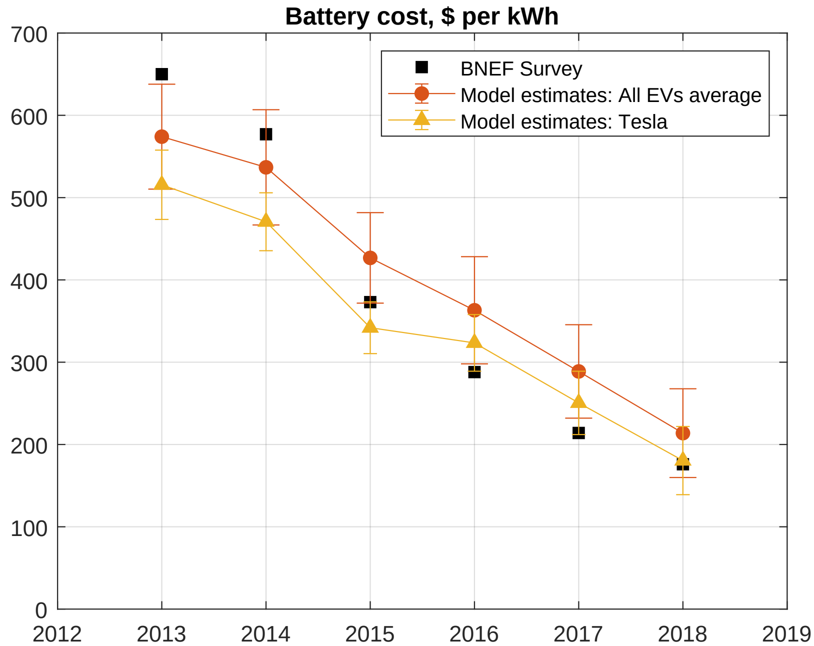


Figure 7: Model vs. real world battery cost estimates

Notes: Vertical bars represent 95% confidence intervals. BloombergNEF survey results can be found at https://about.bnef.com/.

Equation (8). Equation (8) also allows me to recover vehicle-specific shocks to battery costs, hence different BEV models are allowed to have different battery cost structures, which I will use for the counterfactual analysis. The results indicate that the battery cost function is concave in its capacity and decreasing in time, which is consistent with industry technical reports.

To check whether the battery cost estimates agree with other sources, I compute average battery costs per kWh over time and compare them to the BloombergNEF (BNEF) annual battery price surveys, which consider EV and stationary storage batteries. Figure 7 shows the results. The BNEF numbers and the model estimates are very similar. The model predicts that the average battery costs per kWh declined from \$574 to \$214 during 2013-2018, while corresponding BNEF numbers are \$650 and \$176. The figure also shows battery cost estimates for Tesla, an anecdotally recognized leader during this period. Consistently with the anecdotal evidence, the model predicts that Tesla paid on average \$53 less per kWh than other carmakers, with the difference declining to \$33 by 2018. It is worth noting that Tesla's estimated battery cost advantage comes from two sources. The first is from the concave structure of the battery cost function and relatively large batteries that Tesla puts in the cars. The second is from model specific battery cost shocks, which are on average negative for Tesla models. Finally, I check publicly available battery cost estimates of some EV models against the model predictions. For example, UBS (2017) estimates the battery pack cost of the 2017 Chevrolet Bolt to be \$11,500-\$12,500, while the model estimate is \$13,000. The same report's estimate for the 2018 Tesla Model 3 battery pack is \$12,000-\$14,200, while the model prediction is \$14,900.

8 Counterfactual Analysis

In this section, I use the model estimates to perform two counterfactual exercises. In the first exercise, I analyze the effects of two major American policy programs for EVs: federal purchase subsidies for consumers and the ZEV regulation. In the second exercise, I solve for the federal subsidy structure that maximizes either environmental benefits or total welfare holding the program budget fixed, allowing the subsidy to distinguish between BEVs and PHEVs and to depend linearly on battery range.

For each exercise, I assume that carmaker product offerings are given exogenously, but carmakers respond by adjusting prices of their vehicles. Also, American manufacturers are assumed to respond with battery range adjustments of their BEVs while holding other vehicle characteristics fixed. The analysis is performed for the fourth quarter of 2018. In this quarter, there are four BEVs for which the battery range is assumed to be endogenous. These vehicles are sold nationally and include three Tesla models - model S with 335 miles of range and an MSRP of \$94,000, model X (295mi, \$96,000) and model 3 (310 mi, \$46,500) - and Chevrolet Bolt (238mi, \$36,620). These four vehicles together make up 88% of the total BEV sales in this quarter.

I compute policy environmental effects and the effects on consumer surplus and producer profits. The effects on consumer surplus are computed using the compensating variation formula (Small and Rosen, 1981):

$$\Delta CS = \int \frac{1}{\alpha_i^p} \left[\ln\left(1 + \sum_{j=1}^{J} \exp(u_{ijm}^1)\right) - \ln\left(1 + \sum_{j=1}^{J} \exp(u_{ijm}^0)\right) \right] dF(v^i) dF(z_i, d_i),$$

where 0 indicates a baseline scenario and 1 indicates a counterfactual scenario. The environmental effects are approximated by the effects on total CO2 emissions.[61] I start by computing annual emissions for each vehicle. For gasoline cars, I use gasoline consumption ratings (MPG) and model predicted annual mileages to compute annual gasoline consumption, which is then converted into CO2 emissions. For BEVs and PHEVs in electric mode there are no tailpipe emissions, but emissions are produced when electricity that is used for charging is generated. To account for this, I use data on CO2 emissions per kWh of electric energy produced in a given state from the U.S. Energy Information Administration's state electricity profiles.[62] These data are then combined with vehicle electricity consumption per mile (kWh per mile) and model predicted average mileages to compute annual CO2 emissions. In the next step, dollar values of CO2 emission damages are calculated assuming a $50 social cost of CO2 per metric ton (U.S. Environmental Protection Agency, 2016) and 15 years of average car lifespan.

8.1 Federal Subsidy and ZEV Regulation Effects

To evaluate effects of the federal subsidies and the ZEV regulation, I will proceed by removing one program at a time. Thus, the baseline is a market where both programs are in place.[63] The ZEV regulation will be "removed" by setting ZEV credit prices to zero. In reality this would mean relaxing the ZEV regulation requirements to a degree that they do not affect seller decision-making.

A detailed description of the ZEV regulation (as well as the federal subsidies) is provided in Section 2.2. For the purpose of this counterfactual, one can think about the ZEV regulation as a system of subsidies to sellers, where the subsidy size depends on the EV's type and range. After having converted ZEV credits into their dollar values, Figure 8 compares the structures of the two programs. The federal subsidy does not distinguish between BEVs and PHEVs and, essentially, gives a flat $7,500 subsidy to all BEVs (the lowest BEV range in the given quarter is 84mi), with some differences for ranges relevant to plug-in hybrids.[64] The structure of the ZEV regulation is different in a

[61] I consider only CO2 for simplicity. Some studies, however, also take into account other pollutants; see, e.g., Holland et al. (2016).

[62] I use average emission rates, although, ideally, I need marginal emissions rates instead. However, the latter is not readily available to me.

[63] An alternative approach would be to remove both programs and then add one at a time to see their effects. However, since each program is estimated to have a large impact on the EV segment of the market, it seems implausible to believe that the product offerings would be the same if both programs were removed.

[64] The federal subsidy is actually a function of the battery capacity, not range, hence the

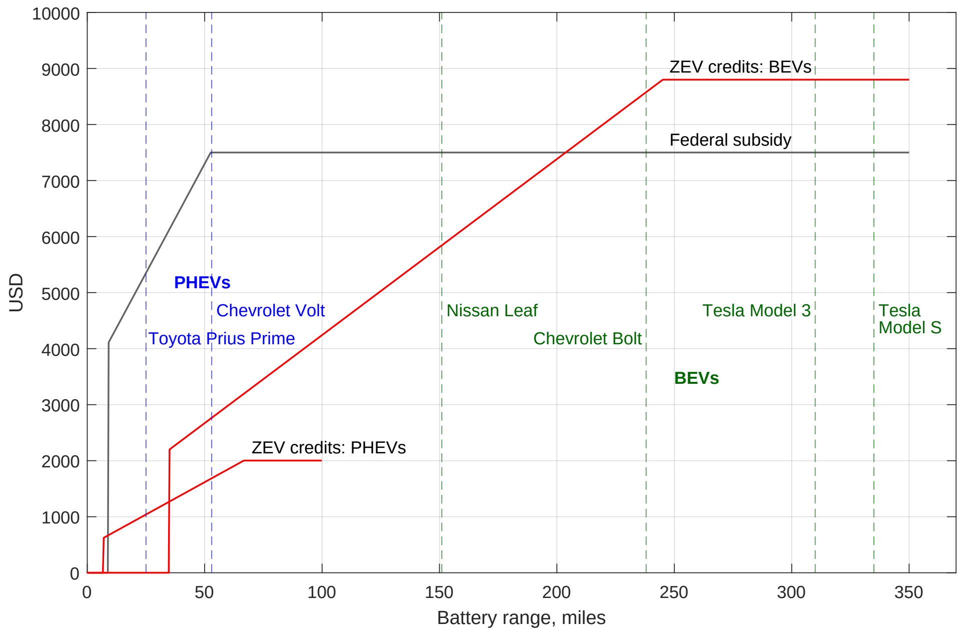


Figure 8: Federal subsidy and ZEV credit structures

Notes: The figure compares the structures of the federal subsidy for EV buyers and the ZEV regulation. The non-flat part of the federal subsidy is approximate, since the actual subsidy is a function of battery capacity rather than range. ZEV credit dollar values are my estimates. The vertical dashed lines represent the battery ranges of some BEV and PHEV models available in 2018.

couple of ways. First, it distinguishes between BEVs and PHEVs and it gives much greater support to BEVs – up to \$8,800 – while for PHEVs it is up to around \$2,000. Second, both BEV and PHEV schemes have non-flat regions relevant to corresponding EV type ranges: up to 240 miles for BEVs and up to 70 miles for PHEVs. Which structure is better for the environment or social welfare is an empirical question because of the presence of multiple factors, including the trade-off between charging inconvenience and production costs of EV batteries, discussed earlier.

I start my analysis by exploring the effects of the two programs on battery ranges. As Table 7 shows, while the federal subsidy, which is flat for BEVs, has no noticeable effect on BEV ranges, the effect of the ZEV regulation, which gives more support to longer range BEVs, is significant, especially for more affordable models: the range of Chevrolet Bolt decreases from 238 miles to

non-flat part on the picture is approximate.

Table 7: Federal subsidy and ZEV regulation effects on BEV ranges

	Range, mi		
Model	Actual range	no Fed subsidy	no ZEV regulation
Chevrolet Bolt	238	240	153
Tesla Model 3	310	312	265
Tesla Model S	335	338	321
Tesla Model X	295	297	282

Notes: The last two columns report the range predictions if the federal subsidy is removed (column 3) or the ZEV credit prices are set to zero (column 4).

153 miles and of Tesla Model 3 from 310 to 265 miles if the ZEV regulation is removed.

Table 8 shows the estimated program effects on other outcomes of interest. First, to make sure the programs have comparable scales, I compute the total amount of the federal subsidies received by consumers and the total dollar value of the ZEV credits earned by sellers. The federal subsidy has a larger scale of \$577M, while the ZEV regulation's scale is \$301M; however, the numbers are quite comparable. Both programs have a large impact on sales of EVs and total annual electric miles traveled; without the federal subsidy, EV sales would drop by 65% and electric mileage by 67%, and, without the ZEV regulation, EV sales would drop by 38% and electric mileage by 48%.

The ZEV regulation adds relatively more electric miles and, hence, replaces more gasoline miles: an EV added due to the ZEV regulation is driven on average 14,638 electric miles while an EV added due to the federal subsidy is driven 11,962 electric miles on average. This translates into larger environmental benefits of an added electric vehicle due to the ZEV regulation: \$3,636, versus \$3,214 in the case of the federal subsidy. This can be attributed to the differences in the program structures: the ZEV regulation gives relatively more support to BEVs and even more support to longer-range BEVs.

Both programs are estimated to improve consumer surplus, which is not surprising, because they make EVs more affordable while not much affecting the prices of traditional vehicles. An interesting question is to what extent consumers capture the program benefits. For the federal subsidy, the estimated pass-through rate to the consumers is 101%, i.e., consumers fully capture the subsidy. This finding is in line with the results by Sallee (2011), who studied tax credits for Toyota Prius in the 2000s, and Muehlegger and Rapson (2021),

Table 8: Federal subsidy and ZEV regulation effects

	no Federal Subsidy	no ZEV Regulation
Program scale	\$577M	\$301M
ΔEV sales	-54,756 (-65%)	-32,177 (-38%)
ΔE-miles	-655M (-67%)	-471M (-48%)
Average e-mileage of added EVs	11,962mi	14,638mi
CO2 emissions benefit	-\$176M	-\$117M
CO2 benefit per added EV	\$3,214	\$3,636
ΔConsumer surplus	-\$373M	-\$229M
ΔProfit	-\$169M	-\$44M
ΔWelfare	-\$718M	-\$390M
ΔWelfare (plus tax dollars)	-\$141M	-\$390M
Pass-through rate to consumers	101%	61%

Notes: The effects are relative to the market with both programs present. Program scale is the total value of subsidies received by consumers or ZEV credits earned by carmakers. CO2 emission benefits are computed assuming \$50 social cost of CO2 per metric ton and 15 years of average car lifespan. Welfare is the sum of CO2 benefit, consumer surplus and firm profits. All numbers are computed for the market as of the fourth quarter of 2018.

who studied subsidies for EVs in California. Both papers found that consumers fully capture the subsidy. In the case of the ZEV regulation, the pass-through rate of ZEV credits to consumers is lower and equal to 61%. This lower value is because sellers earn ZEV credits only in the ZEV states, but set the prices nationally. Hence, they have incentives to reduce the prices of their EVs, but they do not reduce them too much in order to maintain profit margins for sales outside the ZEV states.

To compute the effect on carmaker profits, revenues from ZEV credits are counted as a part of the profit function and the regulation compliance costs are subtracted when necessary. Removing the federal subsidy hurts seller profits, because the subsidy encourages consumers to buy EVs, which helps the sellers earn ZEV credits. When the ZEV regulation is removed, the carmakers don't earn revenues from ZEV credits and, also, don't pay costs of compliance with the ZEV regulation. This is beneficial for traditional carmakers, whose profits are estimated to increase by \$100M. However, for Tesla, which is the only carmaker that doesn't sell any gasoline cars and, hence, doesn't incur any cost of compliance, removing the ZEV regulation means only losing profits, with the profit loss estimated to be $144M. Thus, the estimated overall effect of removing the ZEV regulation on the industry profits is negative and equal to -$44M. This negative effect, however, may potentially be reversed in the future, because the requirements of the ZEV regulation are tightening over time, meaning increasing compliance costs. Hence, in the long run, the ZEV regulation may be hurting the overall industry profits.

Finally, both programs are estimated to improve social welfare: the federal subsidy effect is $718M and the ZEV regulation effect is $390M. However, the federal subsidy is funded directly by taxpayers, hence, the program scale can be seen as the social cost of the subsidy. After subtracting it from the welfare effect above, the net welfare effect is equal to $141M. In the case of the ZEV regulation, the aforementioned number is already the net effect, although, for a more complete picture, it's important to take into account the long run effect of the ZEV regulation on the firms' profits.

8.2 Optimal Federal Subsidy

Next, I investigate whether it is possible to improve performance of the federal subsidy by employing a structure similar to that of the ZEV regulation, i.e., distinguishing between BEVs and PHEVs and depending piece-wise linearly on battery range. I will solve for the optimal subsidy scheme holding the program scale fixed and assuming either of two objectives: maximizing social welfare or maximizing the environmental effect of the program. Focusing on the environmental effect rather than social welfare may be useful for two reasons. First, in this case, assumed dollar costs of emission damages do not affect the solution,[65] while the subsidy that optimizes social welfare depends on assumed dollar values of CO2 emissions, which some readers may find too ad hoc. Second, it may better reflect long-term regulation goals because an emission-minimizing subsidy may be more efficient at encouraging future development and entry of vehicles with more environmentally favorable characteristics.

Figure 9 presents the optimal structures of the federal subsidy depending on which objective function is optimized. I will start by comparing the scheme maximizing CO2 benefit with the status quo subsidy. First, the CO2 optimal subsidy allocates more support to BEVs with relatively long ranges (more than 125 miles) and less support to PHEVs. The maximum subsidy is $8,260 for BEVs and $5,335 for PHEVs, while the current subsidy scheme gives at most

[65]However, I will still present the estimated effects in dollars for ease of interpretation.

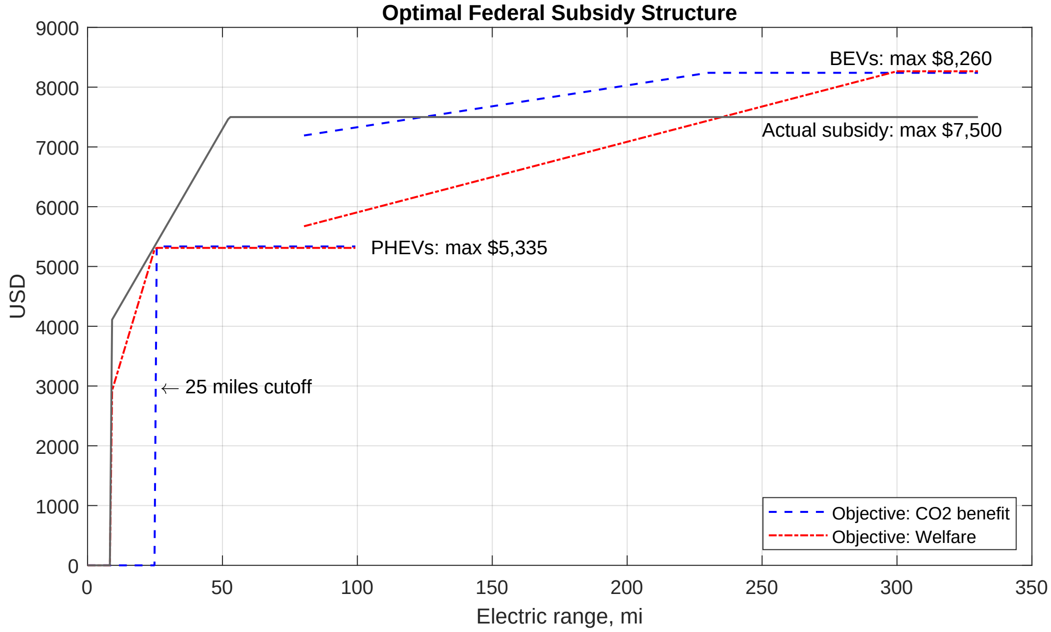


Figure 9: Optimal federal subsidy

Notes: The figure shows an approximate structure of the actual federal subsidy and the optimal structures maximizing CO2 benefit or social welfare, assuming fixed subsidy budget.

$7,500 for both BEVs and PHEVs. The optimal subsidy gives progressively more support to longer-range BEVs. For example, a 100-mile BEV receives around $7,300 while 240-mile and longer range BEVs receive the maximum, $8,260. For the ZEV regulation, the corresponding numbers are $4,200 and $8,800, i.e., the optimal subsidy is much flatter because the ZEV regulation is already steep enough.

The CO2 optimal subsidy does not give any support to PHEVs with ranges less than 25 miles. At the time of the counterfactual analysis, most plug-in hybrids with ranges below 25 miles were plug-in versions of vehicles of luxury brands such as BMW, Mercedes or Audi, which are relatively fuel-inefficient and not expected to be driven a lot on electricity. Hence, it's not surprising that reallocating subsidies from these vehicles to BEVs or longer-range and more affordable PHEVs is beneficial for the environment.

Table 9 reports the effects of the optimal subsidy schemes relative to the status quo, with percentage changes being relative to the effects of the actual subsidy (reported in Table 8). The CO2 optimal subsidy results in lower

Table 9: Optimal federal subsidy effects

Objective	ΔEV sales	ΔE-miles	ΔCO2 effect	ΔCons. Surpl.	ΔProfit	ΔWelfare
CO2 benefit	-4.0%	+4.6%	+\$8.1M (+4.6%)	-\$9.3M (-2.5%)	+\$4.5M (+2.7%)	+\$3.2M (+0.4%)
Welfare	-1.7%	+2.9%	+\$4.9M (+2.8%)	-\$0.0M (-0.0%)	+\$6.3M (+3.7%)	+\$11.2M (+1.6%)

Notes: Percentage effects are relative to the federal subsidy effects reported in Table 8.

sales of EVs (-4.0%), but more electric miles traveled (+4.6%) and higher environmental impact of the program (+\$8.1M, or +4.6%), which means that a subsidy scheme maximizing CO2 benefits does not necessarily maximize sales of EVs. This result illustrates that focusing on EV sales can be misleading if the true goal is minimizing emissions (or maximizing social welfare, as will be discussed later).

The CO2 optimal subsidy improves seller profits since it encourages adoption of EVs that earn more ZEV credits, but hurts consumer surplus, which is intuitive, because it gives lower or no subsidy to vehicles with the lowest expected inconvenience costs, which are plug-in hybrids, especially plug-in hybrids with smaller battery ranges.

Relative to the CO2 optimal subsidy, the welfare optimizing scheme gives some support to PHEVs with ranges below 25 miles by giving a lower subsidy to BEVs with small and intermediate battery ranges. By this, it essentially reallocates support from vehicles with the highest inconvenience costs to vehicles with the lowest inconvenience costs. As a result, this scheme does not hurt consumer surplus but results in a lower improvement in the environmental benefit, which drops from \$8.1M to \$4.9M. The total welfare improvement is estimated to be \$11.2M, which is 1.6% of the effect reported in Table 8, or 7.9% if compared to the net effect, after subtracting the program budget. Similarly to the CO2 optimal subsidy, the welfare optimizing scheme results in lower EV sales added by the program (-1.7%) but more electric miles traveled (+2.9%).

9 Conclusion

Many countries seek to reduce greenhouse gas emissions from the transportation sector by promoting adoption of electric vehicles. Designing policies for electric cars, in particular purchase subsidies, requires understanding of the environmental benefits and production costs of different types of electric vehicles with different characteristics. Some current policies allocate differential support to all-electric and plug-in hybrid electric vehicles with different battery ranges, recognizing the differences in production costs, attractiveness to consumers, usage patterns, and environmental benefits. This paper develops an empirical framework that allows evaluation of environmental and welfare effects of various policies for electric vehicles, taking into account vehicle usage and firm pricing and battery range choice decisions. The framework also allows evaluation of the inconvenience costs of charging electric cars depending on such factors as consumer driving needs, available charging infrastructure, and battery range.

I use the framework to evaluate the effects of two major U.S. policies, the federal subsidy and the Zero Emission Vehicle regulation. Also, I solve for the optimal federal subsidy structure that distinguishes between BEVs and PHEVs and depends on battery range. The optimal structure improves the environmental effect of the subsidy by 4.6% and the social welfare effect by 1.6%, depending on whether the environmental effect or the social welfare is maximized, holding the subsidy budget fixed.

Chapter 2

Do Big Businesses Influence Media? The Case of Amazon.com and the Washington Post

1 Introduction

On October 1, 2013, Jeff Bezos, the founder and CEO of Amazon.com, officially acquired the Washington Post from the Graham family. The acquisition raised concerns about a potential conflict of interest: the Washington Post coverage of Amazon.com may have become biased to benefit the company and the new owner.

A situation, where media owners have vested interests in other companies, is not uncommon. For example, Rupert Murdoch through News Corp. is the owner of a number of news outlets, including the Wall Street Journal, the New York Post, Fox News, and he also controls the 20th Century Fox movie studio. The Walt Disney Company owns ABC television network, as well as a film studio, Disney parks etc. Warner Media controls HBO, CNN and a film studio. Should one be concerned about these media outlets being biased to favor affiliated businesses?

To shed some light on this issue, in this paper I study how the Washington Post coverage of Amazon.com changed after the acquisition of the newspaper by Jeff Bezos. For the analysis, I use data on news stories from four large American newspapers – the Washington Post, the New York Times, the New York Post, and the Los Angeles Times. The data spans over several years before and after the acquisition, from June 2009 to June 2018.

I start with investigating how the acquisition affected the amount of coverage of Amazon.com by the Washington Post. First, I do a simple difference-in-difference analysis on the number of mentions of Amazon.com and its products (e.g. Amazon Alexa, Amazon Prime etc.), comparing the number of mentions in the Washington Post to the number of mentions in the other three newspapers before and after the acquisition. I find that the acquisition resulted in a statistically significant increase in the number of mentions, with the effect being persistent over time.

Next, I investigate whether this result may reflect the change in the Washington Post's coverage policy, which could be trying to focus more on the Internet users, who, in their turn, could be more interested in tech companies. To address this issue, besides Amazon, I include in the analysis other big tech companies, namely, Facebook, Google and Netflix, to control for the changes in the coverage of these companies. The triple-difference analysis produces similar results, showing an increase in the number of mentions, but with a slightly lower magnitude.

Next, I investigate what types of news stories drive the result. I classify all the stories that mention Amazon.com into 3 types: stories with a high concentration of Amazon.com mentions, i.e., those where Amazon.com is likely to be the main theme, stories with a moderate concentration of Amazon.com mentions, and stories with a low concentration of Amazon.com mentions, i.e., those that mention Amazon.com occasionally. The results indicate that the increase in the coverage is driven by the first two types, i.e. by the stories where Amazon.com is the main or one the main topics.

I next turn to discussing potential mechanisms. First, seven months before the acquisition there was a change of the executive editor of the Washington Post. Hence, the change in the Amazon.com coverage may simply reflect the preferences of the new editor. Although, the date of the new editor appointment does not exactly coincide with the acquisition date, the period between the two events largely overlaps the negotiation period, hence the Amazon.com coverage in this period could potentially be affected both by the executive editor change and the anticipation of the future acquisition. Although, I cannot completely rule out the effect of the executive editor change, I can investigate how the coverage of Amazon.com changed at the Boston Globe, where the Washington Post's new editor had an appointment as the executive editor before. After including data from the Boston Globe in the analysis I didn't find any evidence that the coverage of Amazon.com in the Boston Globe had changed after the executive editor was changed there, suggesting that the change in the coverage of Amazon.com in the Washington Post, likely, was not due to the executive editor change.

There are several potential explanations of an increased coverage of Amazon.com in the Washington Post that the analysis in the paper cannot rule out. The first one is a potential conflict of interest, where the newspaper provides more coverage about Amazon.com and related products and services to benefit the owner, for example, by advertising Amazon and its products this way. Second, readers that are interested in Bezos and Amazon.com may be more likely to subscribe to the Washington Post after the Bezos' acquisition, and the newspaper could be simply addressing their demands by providing more stories about Amazon.com and Bezos. Third, as a result of the acquisition, the Washington Post may have got cheaper or faster access to the news about Amazon.com, which resulted in more stories about the company. Finally, since the acquisition was a major event for the Washington Post's journalists and editors, they may have updated their beliefs about the importance of the news about Amazon.com and, as a result, they started to publish more stories.

In the second part of the paper, I explore whether the acquisition affected the tone of the coverage of Amazon.com in the Washington Post. To do this, I use an approach similar to Tetlock (2007) and Tetlock et al. (2008). I construct a simple sentiment measure which is equal to the fraction of negative words in the total number of words in a news story, assuming that it reflects the overall tone of the article. I use two lists of negative words from two dictionaries, the commonly used Harvard IV-4 psychosociological dictionary and the Loughran/McDonald dictionary, constructed by Loughran and McDonald (2011) for financial applications. The latter one can be useful since some of the news stories discuss Amazon's financial indicators, such as stock prices, profits, sales etc.

One can think about several ways of how the owner-related news stories in a given outlet can be biased. First, media outlets may give more coverage to favorable stories and underplay stories that are critical of the owner. Second, the news stories can be given more positive outlook in general, even the critical ones. The sentiment measure used here can capture both mechanisms, but cannot distinguish between them.[66]

The results of the sentiment analysis are inconclusive. While the measure based on the Loughran/McDonald dictionary provides some weak evidence that the coverage of Amazon in the Washington Post became more positive after the acquisition, the measure based on the Harvard IV-4 dictionary doesn't detect any statistically significant change of the coverage tone.

Literature review. The literature on media bias has been focusing mainly on the political side of the issue (Groseclose and Milyo (2005), Gentzkow and Shapiro (2010), Larcinese et al. (2011), Durante and Knight (2012)). A few papers study media bias due to pressure of advertisers (Ellman and Germano (2009), Beattie et al. (2017)). In particular, Beattie et al. (2017) in their empirical study find that newspapers provide less coverage of car safety recalls by auto manufacturers if these auto manufacturers buy ads in the newspapers. The closest study to my paper is Dellavigna and Hermle (2017), where they explore media bias due to cross-ownership. In particular, they consider movie reviews in media outlets owned by News Corp and Time Warner and test whether these outlets provide biased ratings for the movies produced by the affiliated film studios 20th Century Fox and the Warner Bros, respectively, and find no evidence of bias. The paper gives three possible explanations of this finding: high reputational costs, high distance between the outlets and the movie studios, and a low return to bias.

The paper proceeds as follows. Section 2 describes the data. Section 3 studies the effect on the coverage volume. Section 4 studies the effect on the story sentiment and Section 5 concludes.

[66] Also, a general issue of this method is its inability in some cases to detect the true sentiment because negative words do not always identify negative context (e.g sarcasm, negation, disagreement with negative opinions, etc.).

2 Data

I use digital archives of the news stories from four major US newspapers: the Washington Post, the New York Times, the New York Post and the Los Angeles Times. The archives of the Washington Post and the New York Post come from LexisNexis, and the archives of the New York Times and the Los Angeles Times come from ProQuest. The data spans from June 2009 to June 2018, i.e., approximately four years before and five years after the acquisition of the Washington Post by Jeff Bezos.

The Washington Post and the New York Times' archives contain articles both from the printed versions of the newspapers and from the online blogs on the newspapers' websites. I use only articles from the printed versions because: 1) the archives of other newspapers don't have stories from their online blogs, and 2) the stories from the online blogs for the Washington Post and the New York Times are missing for some time periods.

The data cleaning process consists of two steps. First, for each newspaper, I search for the news stories that contain keywords "Amazon" or "Jeff Bezos". This gives 14,450 stories in total from the all four newspapers. However, some of these stories are irrelevant for the analysis since, for example, they are related to the Amazon river or rainforest or they are book reviews with a link to Amazon website. Hence, in the second step, I analyze a subset of the news stories by identifying some keywords that are likely to indicate that a given story is about the Amazon river or rainforest or a book review. For instance, most stories that mention keywords "river", "forest", or "rainforest" and don't mention Amazon.com or Jeff Bezos in the story's metadata are not about Amazon.com. I drop those stories from the analysis. Also, I drop stories that mention Amazon only within a web link, since those links were almost always related to book reviews.

3 Effect on Coverage Volume

In this section, I study how the acquisition of the Washington Post by Jeff Bezos affected the volume of the coverage of Amazon in the Washington Post. As a main measure of the coverage volume, I will focus on how many times Amazon, its products, and Jeff Bezos are mentioned every month in each newspaper from my dataset. Then, I will also split all the news stories into three groups according to the concentration of the Amazon-related keywords – from the stories with the highest concentration, where Amazon, its products, or Jeff Bezos are likely to be the main theme of the story, to the stories with the lowest concentration, i.e., where Amazon, its products, or Jeff Bezos are only mentioned occasionally. I will study the effect of the acquisition on the number of stories from each group. That will allow us to learn more about the types of the news stories where the effect on the coverage volume is coming from.

3.1 Model-free Evidence

I start with presenting some model-free evidence on how the acquisition affected the number of Amazon-related mentions. For illustration, for each newspaper I normalize the monthly number of mentions by its pre-acquisition mean. Figure 1 shows the results. The figure has four subplots: three subplots compare the number of mentions in the Washington Post to each newspaper from the "control" group – the New York Times, the New York Post, and the Los Angeles Times, and the fourth one plots the number of Amazon-related mentions for the New York Post versus the New York Times.

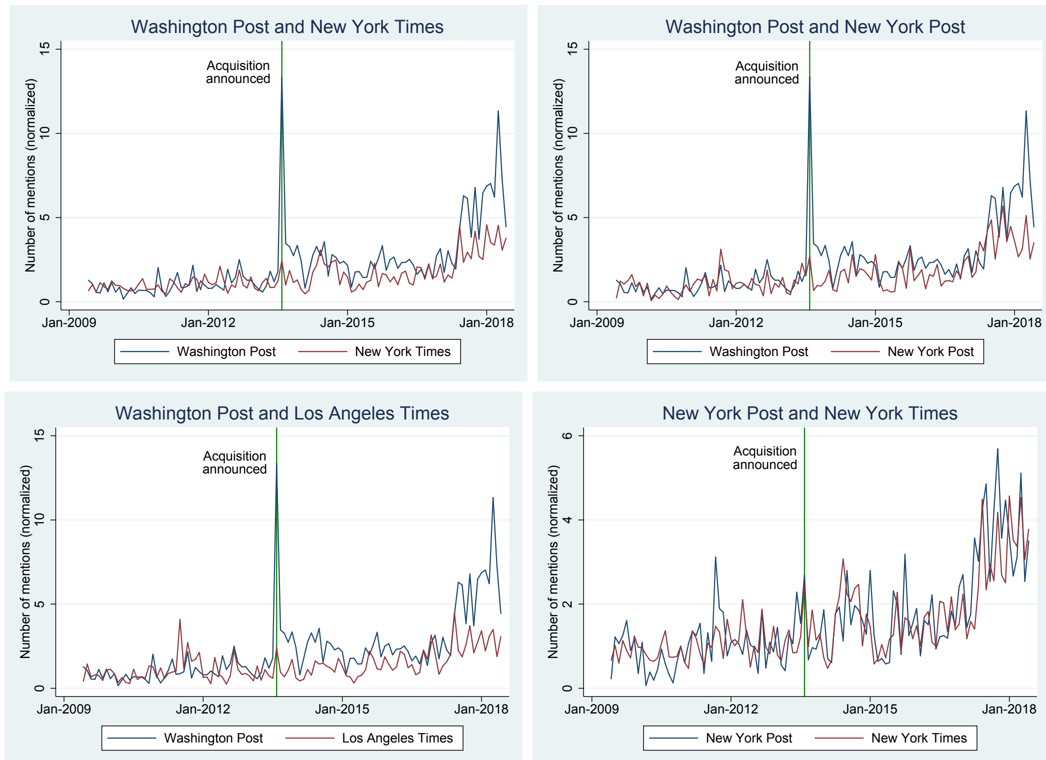
One can see a huge spike in the number of mentions for the Washington Post at the time of the acquisition. This spike is coming mainly from the news stories related to the acquisition itself and to the vision of the newspaper's future. The effect vanishes quickly. At the same time, the other newspapers did not see a noticeable increase in the number of Amazon-related mentions when the acquisition of the Washington Post by Jeff Bezos was announced.

From the three subplots that compare the number of mentions in the Washington Post versus other newspapers one can see a noticeable gap during the period after the acquisition, which indicates a relative increase in the number of mentions of Amazon and related keywords for the Washington post as a result of the acquisition. Noticeably, this gap is persistent over time. There is no noticeable gap when comparing the number of mentions between the New York Post and the New York Times (shown on Figure 1), as well as other pairs of the "control" group newspapers (not shown on Figure 1).

3.2 Difference-in-Difference Analysis

Next, I do formal statistical analysis to establish the effect of the acquisition on the Amazon coverage volume in the Washington Post. I start with difference-in-difference analysis, where I study how the number of Amazon-related mentions in the Washington Post changed after the acquisition relative to "control" group newspapers – the New York Times, the Los Angeles Times, and the New York Post. I estimate the following regression equation:

Figure 1: Model-free evidence on the effect on the coverage volume



Notes: The graphs show monthly number of Amazon-related mentions normalized by pre-acquisition averages for each newspaper.

$$LogMentions_{it} = \alpha_i + \sum_{k=-3}^{5} \delta_k D_{kt} + \sum_{k=-3}^{5} \beta_k D_{kt} \cdot WPost_i + \theta size_{it} + \varepsilon_{it}, \quad (9)$$

The dependent variable is the natural logarithm of the number of Amazon-realted mentions in newspaper i in month t. On the right-hand side, α_i is the newspaper i's fixed effect, D_{kt} is a dummy variable for year k, where k runs from -3 to 5. $WPost_i$ is an indicator for the Washington Post newspaper. Coefficients β_k represent the acquisition effect in year k, with β_{-3} to β_0 representing the years before the acquisition, and β_1 to β_5 representing the years after the acquisition. I normalize coefficient β_{-1} to zero, i.e., the estimated acquisition effects for each year should be interpreted as being relative to the year between two and one years before the acquisition. The year right before the acquisition, represented by coefficient β_0, may have potentially been affected by negotiations between Bezos and Graham Holdings, then-owner of the newspaper, while the earlier periods, from -3 to -1, are unlikely to be affected by the future acquisition. I exclude the period between the acquisition announcement and the actual transaction, i.e., the period from August to October 2013, from the analysis. During this period the Washington Post released a large number of stories about the acquisition. Outside this period there are a very few stories about the acquisition. Finally, I control for newspaper size, $size_{it}$, which is defined as the total number of stories in a given month.

Table 1 presents the estimation results. Panel A shows the estimates of the acquisition effect by year. Columns 1 to 3 show the estimation results when one of the "control" group newspapers is used for comparison, and the regression from column 4 uses all the three newspapers together.

Most of the coefficients for the post-acquisition years, 1 to 5, are positive and statistically significant, indicating that the acquisition resulted in increased coverage of Amazon, its products, and Jeff Bezos in the Washington Post. The coefficients β_{-3} and β_{-2} have a smaller magnitude and are not statistically significant, supporting the "parallel trend" assumption used by the difference-in-difference strategy. The coefficient β_0 is positive and has a larger magnitudes relative to β_{-3} and β_{-2} for each regression, which may indicate some effect on coverage during the negotiations, however, it's still not statistically significant.

Panel B of Table 1 shows the "average" acquisition affect, which is an outcome of a regression, where all the pre-acquisition years are combined into one time period and all the post-acquisition years are combined into another time period. The acquisition effect is positive and statistically significant. Panel C shows the average effect from a similar regression, where the year just before the acquisition is dropped from the analysis so that the results are not affected by the negotiations before the acquisition. This results in even larger estimate of the acquisition effect on Amazon-related coverage in the Washington Post.

3.3 Controlling for the News Coverage of Other Big Tech Companies

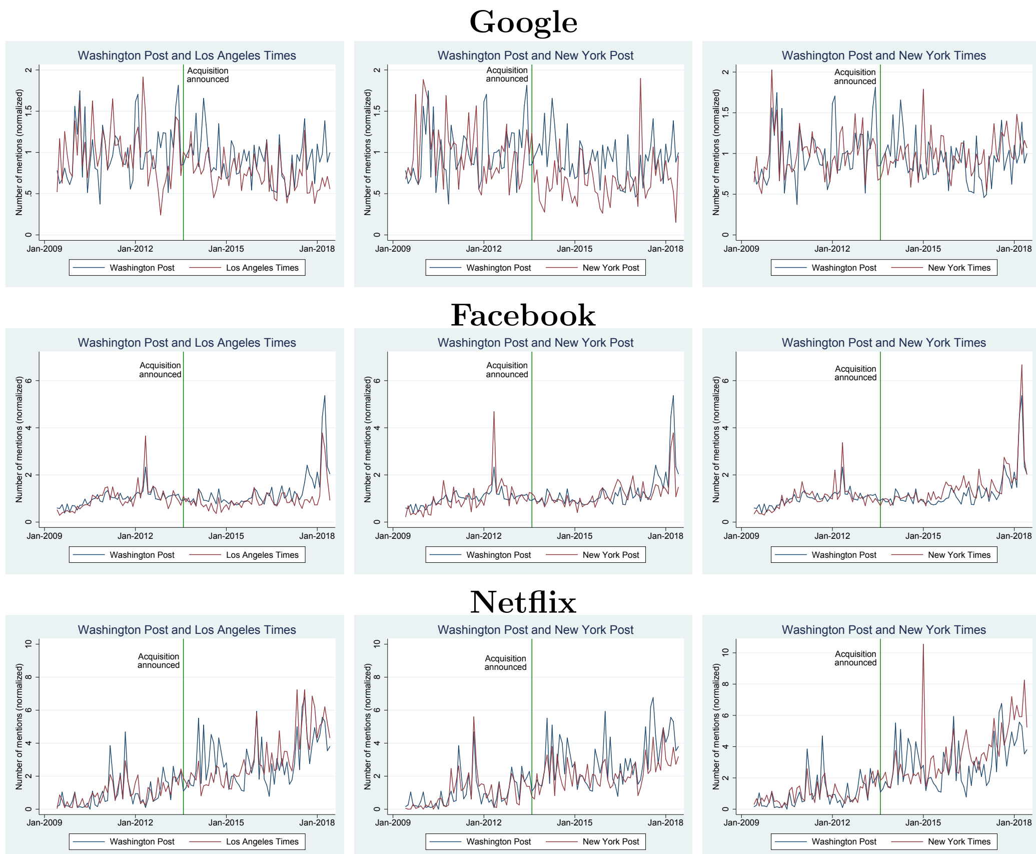
Potentially, an increase in the Amazon-related coverage in the Washington Post can be an indicator of a change in the overall editorial policy that could be affecting the coverage of the entire tech sector. Hence, besides Amazon, we may see an increase in the coverage of other tech companies in the Washington

Table 1: The effect on coverage volume: difference-in-difference analysis.

Dependent variable: Log Number of Amazon Mentions				
	(1)	(2)	(3)	(4)
Control group:	NY Times	NY Post	LA Times	All newsp
Panel A: Yearly effects				
β_{-3}	0.0119	0.286	0.0539	0.123
	(0.249)	(0.254)	(0.307)	(0.208)
β_{-2}	0.166	0.223	-0.189	0.098
	(0.252)	(0.257)	(0.382)	(0.225)
β_{-1}	0	0	0	0
	(-)	(-)	(-)	(-)
β_0	0.267	0.430*	0.272	0.307
	(0.221)	(0.226)	(0.279)	(0.186)
β_1	0.573**	0.752***	0.788***	0.678***
	(0.230)	(0.201)	(0.277)	(0.186)
β_2	0.577***	0.617***	0.570**	0.587***
	(0.211)	(0.209)	(0.273)	(0.182)
β_3	0.472**	0.559***	0.445*	0.511***
	(0.214)	(0.166)	(0.253)	(0.165)
β_4	0.378*	0.259	0.254	0.319
	(0.225)	(0.183)	(0.283)	(0.186)
β_5	0.659***	0.703***	0.723***	0.728***
	(0.209)	(0.199)	(0.264)	(0.178)
Observations	214	214	214	428
R-squared	0.783	0.677	0.603	0.742
Panel B: Average effect, year before acquisition included				
Average effect	0.417***	0.344***	0.464***	0.420***
	(0.115)	(0.117)	(0.127)	(0.088)
Observations	214	214	214	428
R-squared	0.779	0.668	0.588	0.739
Panel C: Average effect, year before acquisition excluded				
Average effect	0.498***	0.404***	0.603***	0.481***
	(0.147)	(0.131)	(0.172)	(0.106)
Observations	190	190	190	380
R-squared	0.785	0.685	0.604	0.743

Notes: Time periods -3 to 0 correspond to pre-acquisition years and time periods 1 to 5 correspond to post-acquisition years. All the regressions include newspaper and year FEs and control for the total monthly number of articles. Standard errors are clustered by month.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure 2: Coverage of other big tech companies. Model-free evidence.



Notes: The graphs show monthly number of mentions normalized by pre-acquisition averages for each newspaper and company.

Post. To account for that, I include in the analysis the news stories about some big tech companies, namely, Google, Facebook, and Netflix.

I start with presenting some model-free evidence on how the coverage volume of Google, Facebook, and Netflix in the four newspapers evolved before and after the acquisition of the Washington Post by Jeff Bezos. Figure 2 compares the number of mentions of Google, Facebook, and Netflix in the Washington Post against each of the newspaper from the control group. As before, the number of mentions is normalized by the pre-acquisition mean for each company and newspaper. The evidence is inconclusive. In some cases there is a noticeable gap in the post-acquisition coverage, e.g., in case of Google, between the Washington Post and the New York Post. However, in general, it's hard to conclude whether the acquisition affected the coverage of Google, Facebook, or Netflix in the Washington Post. Hence, I am turning to the formal analysis.

Formally, I include the stories about Google, Facebook, and Netflix from all the four newspapers in another control group and run a triple-difference analysis. In particular, I estimate the following regression:

$$LogMentions_{ijt} = \alpha_i+\gamma_j+\sum_{k=-3}^{5}\delta_k D_{kt}+\sum_{k=-3}^{5}\eta_k D_{kt}\cdot WPost_i+\sum_{k=-3}^{5}\sigma_k D_{kt}\cdot Amazon_j$$
$$+\,\psi\cdot WPost_i\cdot Amazon_j+\sum_{k=-3}^{5}\beta_k D_{kt}\cdot WPost_i\cdot Amazon_j+\theta\cdot size_{it}+\epsilon_{it}, \tag{10}$$

Here, α_i and γ_j are newspaper and company fixed effects, respectively; D_{kt} is an indicator for year k; $WPost_i$ is a dummy for the Washington Post; $Amazon_j$ is a dummy for Amazon. We are interested in the coefficients β_k, which represent the acquisition effect on Amazon-related coverage in the Washington Post in year k.

Table 2 shows the estimation results. In this table, columns 1-3 represent the regressions where all the three control group newspapers but only one of the control group companies, Google, Facebook, or Netflix, are included in the data. Columns 4-6 show the estimation results, where all the control group companies but only one of the control group newspapers are included in the data. Finally, the last column regression includes all the companies and newspapers. The results are robust to the composition of the control group, so I will focus on the estimates in the last column of Table 2. The coefficient estimates for yearly effects are similar to those previously reported in Table 1. The average effects reported in Panels B and C of Table 2 are also similar to the corresponding estimates from Table 1. Hence, including the news coverage data for other big tech companies did not affect the results substantially.

3.4 What kind of stories drive the effect?

Next, I investigate what type of Amazon-related coverage drives the previously established effect. Since I measure the coverage volume using the number of Amazon-related mentions, the results don't tell us much about whether the increase in the number of mentions is coming from an increase in the number of stories where Amazon, its products, or Jeff Bezos are the main theme, or

Table 2: The effect on coverage volume relative to other newspapers and other big tech companies.

Dependent variable: Log Number of Mentions							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Control group:	All newspapers			All companies			All comp.
	Google	Facebook	Netflix	NY Times	NY Post	LA Times	All newsp.
Panel A: Yearly effects							
β_{-3}	0.186	-0.268	-0.152	-0.0990	0.0496	-0.185	-0.0780
	(0.201)	(0.231)	(0.292)	(0.219)	(0.267)	(0.266)	(0.192)
β_{-2}	0.123	-0.0486	-0.0881	-0.00278	0.228	-0.239	-0.00446
	(0.199)	(0.256)	(0.278)	(0.232)	(0.282)	(0.316)	(0.215)
β_{-1}	0	0	0	0	0	0	0
	(-)	(-)	(-)	(-)	(-)	(-)	(-)
β_0	0.189	0.298	0.00695	0.134	0.271	0.0892	0.165
	(0.231)	(0.212)	(0.193)	(0.205)	(0.249)	(0.272)	(0.185)
β_1	0.447**	0.594***	0.264	0.450**	0.408*	0.447	0.435**
	(0.200)	(0.214)	(0.228)	(0.212)	(0.214)	(0.279)	(0.184)
β_2	0.521**	0.542**	0.448**	0.696***	0.446**	0.369	0.504***
	(0.200)	(0.217)	(0.190)	(0.207)	(0.219)	(0.264)	(0.173)
β_3	0.389**	0.620***	0.430*	0.710***	0.388*	0.342	0.480***
	(0.175)	(0.189)	(0.220)	(0.202)	(0.198)	(0.238)	(0.158)
β_4	0.336*	0.266	0.198	0.575***	0.0676	0.156	0.266
	(0.200)	(0.199)	(0.172)	(0.189)	(0.196)	(0.256)	(0.164)
β_5	0.496***	0.440**	0.624***	0.845***	0.323*	0.391	0.520***
	(0.184)	(0.211)	(0.172)	(0.197)	(0.175)	(0.263)	(0.162)
Observations	856	856	856	856	856	856	1,712
R-squared	0.800	0.819	0.755	0.808	0.685	0.635	0.727
Panel B: Average effect, year before acquisition included							
Average effect	0.307***	0.510***	0.443***	0.641***	0.193	0.426***	0.420***
	(0.0866)	(0.0989)	(0.125)	(0.0951)	(0.123)	(0.112)	(0.0878)
Observations	856	856	856	856	856	856	1,712
R-squared	0.767	0.806	0.728	0.800	0.674	0.622	0.719
Panel C: Average effect, year before acquisition excluded							
Average effect	0.327***	0.609***	0.465***	0.683***	0.236*	0.482***	0.467***
	(0.0919)	(0.112)	(0.146)	(0.107)	(0.138)	(0.128)	(0.0983)
Observations	760	760	760	760	760	760	1,520
R-squared	0.765	0.804	0.745	0.796	0.664	0.612	0.713

Notes: Regressions 1-3 include all the newspapers and regressions 4-6 include all the companies. Time periods -3 to 0 correspond to pre-acquisition years and time periods 1 to 5 correspond to post-acquisition years. All the regressions include newspaper, company, and year fixed effects and control for the total monthly number of articles. Standard errors are clustered by month.

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

Table 3: The composition of news stories that mention Amazon by story type.

Overall	Total N articles	Type 1	Type 2	Type 3
The Washington Post	3613	258 (7%)	943 (26%)	2412 (67%)
The New York Times	7398	503 (7%)	1100 (15%)	5795 (78%)
The New York Post	1922	234 (12%)	356 (19%)	1332 (69%)
The Los Angeles Times	3636	380 (10%)	530 (15%)	2726 (75%)
Before the acquisition	Total N articles	Type 1	Type 2	Type 3
The Washington Post	1023	39 (4%)	197 (19%)	787 (77%)
The New York Times	2179	170 (8%)	379 (17%)	1630 (75%)
The New York Post	535	64 (12%)	111 (21%)	360 (67%)
The Los Angeles Times	1154	159 (14%)	206 (18%)	789 (68%)
After the acquisition	Total N articles	Type 1	Type 2	Type 3
The Washington Post	2481	198 (8%)	718 (29%)	1565 (63%)
The New York Times	5121	323 (6%)	703 (14%)	4095 (80%)
The New York Post	1361	167 (12%)	237 (18%)	957 (70%)
The Los Angeles Times	2427	211 (9%)	313 (13%)	1903 (78%)

Notes: The table shows the number and the fraction of stories of different types that mention Amazon by newspaper. Type 1 stories has the highest concentration of Amazon-related mentions, type 2 stories has intermediate concentration, and Type 3 stories has the lowest concentration.

whether it is coming from an increase in the number of occasional mentions in news stories about various things, or it is a combination of those two.

To study this aspect of coverage, I split all the news stories that mention Amazon into three groups – those with the highest concentration of Amazon-related keywords, i.e., where Amazon is likely to be the main topic (type 1 stories), those with intermediate concentration, i.e., where Amazon is a substantial part of the story, but is potentially not the main theme (type 2 stories), and, finally, stories with the lowest concentration, where Amazon is, likely, being mentioned just occasionally (type 3 stories). To define the thresholds of Amazon-related keywords concentration for the three types, I read and analyzed a subset of stories that mention Amazon from each newspaper and manually set the threshold values.

Table 3 shows the number of stories of each type, as well as the fraction of stories of each type in the total number of stories for each newspaper. A vast majority of the stories (67%-78%) where Amazon is mentioned mention it occasionally. The stories where Amazon is the main theme constitute only about 7% to 12%. The table shows a substantial increase in the fraction of type 1 and type 2 stories for the Washington Post in the post-acquisition period relative to the period before the acquisition.

Tables 4-6 present the estimation results of regressions similar to (9) and (10) for each type of stories. Table 4 shows the results for type 1 stories, i.e., stories with the highest concentration of Amazon-related keywords. Columns 1 and 2 show the estimates of difference-in-difference regressions, similar to (9), and columns 3 and 4 show the estimates of triple-difference regressions, similar to (10). For the triple-difference strategy, I include type 1 stories about Google, Facebook, and Netflix, where I use the same keyword concentration thresholds as for Amazon to classify stories that mention these companies into three types. As a dependent variable I use the natural logarithm of either the number of mentions or the number of stories in a newspaper-month. The estimates from Table 4 show a statistically significant effect of the acquisition on the number of type 1 stories, as well as on the overall number of Amazon-related mentions in type 1 stories. Hence, the acquisition resulted in a higher number of stories mainly devoted to Amazon or Jeff Bezos.

Table 5 shows the estimation results for type 2 stories, where Amazon is likely to be a substantial part of the story, but not necessarily the main one. The estimates show a statistically significant increase in the coverage related to type 2 stories as well.

Finally, Table 6 shows the results for type 3 stories, i.e. stories that occasionally mention Amazon. For this type of stories we do not see any significant positive effect, some of the coefficients are even negative and statistically significant. Hence, the effect, if any, is likely to be negative for the stories that mention Amazon occasionally.

To summarize, I found an empirical evidence that the previously established positive effect of the acquisition on Amazon-related coverage is coming from the stories where Amazon is at least a substantial part of the story, and it's not coming from occasionally mentioning Amazon in stories where Amazon is not a substantial part of the story.

Table 4: The effect of the acquisition on Type 1 stories.

	(1)	(2)	(3)	(4)
	Diff-in-diff		Diff-in-diff-in-diff	
Dep. var:	logMentions	logNstories	logMentions	logNstories
Panel A: Yearly effects				
β_{-3}	0.971*	0.356*	0.271	0.0733
	(0.580)	(0.203)	(0.675)	(0.260)
β_{-2}	0.485	0.216	0.325	0.154
	(0.633)	(0.284)	(0.686)	(0.298)
β_{-1}	0	0	0	0
	(-)	(-)	(-)	(-)
β_0	0.491	0.193	0.715	0.256
	(0.549)	(0.190)	(0.701)	(0.252)
β_1	2.278***	0.775***	1.959***	0.628***
	(0.413)	(0.172)	(0.610)	(0.237)
β_2	1.856***	0.570***	1.761***	0.595**
	(0.516)	(0.188)	(0.643)	(0.237)
β_3	1.776***	0.551***	1.549**	0.475*
	(0.561)	(0.197)	(0.693)	(0.260)
β_4	1.547***	0.401**	1.055*	0.229
	(0.480)	(0.200)	(0.633)	(0.249)
β_5	2.487***	1.020***	1.511***	0.560**
	(0.398)	(0.178)	(0.553)	(0.233)
Observations	428	428	1,712	1,712
R-squared	0.442	0.419	0.464	0.474
Panel B: Average effect, year before acquisition included				
Average effect	1.481***	0.456***	1.244***	0.376***
	(0.268)	(0.106)	(0.303)	(0.120)
Observations	428	428	1,712	1,712
R-squared	0.432	0.406	0.448	0.458
Panel C: Average effect, year before acquisition excluded				
Average effect	1.476***	0.445***	1.367***	0.419***
	(0.311)	(0.128)	(0.329)	(0.135)
Observations	380	380	1,520	1,520
R-squared	0.429	0.399	0.438	0.446

Notes: Type 1 stories are stories with the highest concentration of Amazon-related keywords. All regressions include newspaper and year fixed effects and control for the total monthly number of articles. Standard errors are clustered by month.
*** p<0.01, ** p<0.05, * p<0.1

Table 5: The effect of the acquisition on Type 2 stories

	(1)	(2)	(3)	(4)
	Diff-in-diff		Diff-in-diff-in-diff	
Dep. var:	logMentions	logNstories	logMentions	logNstories
Panel A: Yearly effects				
β_{-3}	-0.290	-0.296	-0.659	-0.526**
	(0.438)	(0.234)	(0.407)	(0.212)
β_{-2}	0.455	0.101	0.0284	-0.0867
	(0.332)	(0.193)	(0.353)	(0.206)
β_{-1}	0	0	0	0
	(-)	(-)	(-)	(-)
β_0	0.180	0.0892	-0.0775	0.0149
	(0.328)	(0.216)	(0.294)	(0.212)
β_1	0.475	0.463**	0.636**	0.592***
	(0.332)	(0.223)	(0.289)	(0.203)
β_2	0.695**	0.525**	0.611*	0.517**
	(0.332)	(0.206)	(0.335)	(0.229)
β_3	0.678**	0.529***	0.891***	0.814***
	(0.281)	(0.198)	(0.280)	(0.201)
β_4	0.512*	0.562***	0.815***	0.910***
	(0.283)	(0.201)	(0.269)	(0.180)
β_5	0.863***	0.866***	0.791***	0.913***
	(0.295)	(0.201)	(0.254)	(0.180)
Observations	428	428	1,712	1,712
R-squared	0.600	0.628	0.591	0.630
Panel B: Average effect, year before acquisition included				
Average effect	0.603***	0.636***	0.943***	0.905***
	(0.157)	(0.0978)	(0.170)	(0.109)
Observations	428	428	1,712	1,712
R-squared	0.562	0.582	0.584	0.622
Panel C: Average effect, year before acquisition excluded				
Average effect	0.600***	0.650***	0.980***	0.962***
	(0.193)	(0.112)	(0.199)	(0.118)
Observations	380	380	1,520	1,520
R-squared	0.597	0.633	0.591	0.626

Notes: Type 2 stories are stories with intermediate concentration of Amazon-related keywords. All regressions include newspaper and year fixed effects and control for the total monthly number of articles. Standard errors are clustered by month.
*** p<0.01, ** p<0.05, * p<0.1

Table 6: The effect of the acquisition on Type 3 stories.

	(1)	(2)	(3)	(4)
	Diff-in-diff		Diff-in-diff-in-diff	
Dep. var:	logMentions	logNstories	logMentions	logNstories
Panel A: Yearly effects				
β_{-3}	-0.0609	-0.0623	0.0113	0.0394
	(0.157)	(0.147)	(0.159)	(0.148)
β_{-2}	-0.0937	-0.116	0.130	0.0800
	(0.122)	(0.103)	(0.135)	(0.106)
β_{-1}	0	0	0	0
	(-)	(-)	(-)	(-)
β_0	0.358***	0.215**	0.331***	0.219**
	(0.105)	(0.101)	(0.117)	(0.0992)
β_1	0.396***	0.242**	0.302**	0.177
	(0.129)	(0.116)	(0.149)	(0.133)
β_2	0.102	0.0488	0.271**	0.266**
	(0.119)	(0.109)	(0.124)	(0.107)
β_3	-0.126	-0.260**	0.139	0.0177
	(0.136)	(0.120)	(0.138)	(0.122)
β_4	-0.310**	-0.629***	-0.0169	-0.377***
	(0.122)	(0.110)	(0.124)	(0.104)
β_5	-0.260**	-0.436***	0.0339	-0.237**
	(0.120)	(0.118)	(0.130)	(0.112)
Observations	428	428	1,712	1,712
R-squared	0.882	0.872	0.836	0.851
Panel B: Average effect, year before acquisition included				
Average effect	-0.0766	-0.198***	0.0382	-0.102
	(0.0716)	(0.0719)	(0.0653)	(0.0673)
Observations	428	428	1,712	1,712
R-squared	0.853	0.830	0.829	0.844
Panel C: Average effect, year before acquisition excluded				
Average effect	0.00563	-0.154*	0.107	-0.0595
	(0.0791)	(0.0798)	(0.0737)	(0.0747)
Observations	380	380	1,520	1,520
R-squared	0.869	0.852	0.826	0.841

Notes: Type 3 stories are stories with the lowest concentration of Amazon-related keywords. All regressions include newspaper and year fixed effects and control for the total monthly number of articles. Standard errors are clustered by month.
*** p<0.01, ** p<0.05, * p<0.1

3.5 Alternative Explanations: Change of the Executive Editor

On January 2, 2013, seven months before the acquisition, the Washington Post appointed a new chief executive editor, Martin Baron, who previously had held a similar position at the Boston Globe. Hence, the increase in the Amazon-related coverage at the Washington Post that we have found can potentially be attributed to the change of the chief executive editor, especially given that we see some weak evidence of the coverage increase during the negotiation stage before the acquisition.

In this subsection I investigate whether the change in the Amazon-related coverage at the Washington Post can, indeed, be attributed to the preferences of the new chief executive editor. To do that, I will study how the Amazon-related coverage changed at the Boston Globe after Martin Baron left his position there. I need to acknowledge the limitations of this strategy. For example, even if don't see any change in news coverage at the Boston Globe, that could be explained by inertia or similar tastes of the new chief editor at the Boston Globe.

Nevertheless, in addition to the data I have I collected the data on the coverage of Amazon.com, as well as Google, Facebook, and Netflix, from the Boston Globe archives and ran regressions similar to (9) and (10), replacing the number of Amazon-related mentions in the Washington Post with the number of Amazon-related mentions in the Boston Globe. Table 7 presents the results. Some of the yearly coefficients are positive and statistically significant, indicating an increase in Amazon.com coverage in these periods. However, this can be explained by the coverage of local news related to Amazon.com. For example, in period 0 Amazon.com purchased a robot-maker company Kiva, located in Massachusetts. Also, there was a debate around the sales tax deal between Amazon.com and the state in this period. Increased coverage in period 5 can be explained by the news about Amazon.com choosing a location for its second headquarters, where Boston was one of the candidates. A small and statistically insignificant estimate of the average effect suggests that the Amazon-related coverage in the Boston Globe was not affected by the executive editor change. Keeping in mind the limitations I discussed above, these results suggest that the change in the Amazon-related coverage at the Washington Post at the time of the acquisition is, likely, not due to the chief editor change happened seven months before the acquisition.

Table 7: The effect of the chief editor change on Amazon coverage in the Boston Globe.

Dependent variable: Log Number of Amazon mentions	(1) Diff-in-diff	(2) Diff-in-diff-in-diff
Panel A: Yearly effects		
β_{-2}	0.621***	0.102
	(0.172)	(0.179)
β_{-1}	0	0
	(-)	(-)
β_0	0.706***	0.600***
	(0.189)	(0.179)
β_1	0.303	0.256
	(0.198)	(0.196)
β_2	0.0211	-0.194
	(0.192)	(0.212)
β_3	0.624***	0.275
	(0.209)	(0.205)
β_4	0.547*	0.221
	(0.316)	(0.206)
β_5	0.629**	0.686***
	(0.242)	(0.180)
Observations	436	1,744
R-squared	0.731	0.735
Panel B: Average effect		
Average effect	-0.0811	0.0759
	(0.114)	(0.107)
Observations	436	1,744
R-squared	0.718	0.720

Notes: All regressions include newspaper and year fixed effects and control for the total monthly number of articles. Standard errors are clustered by month.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

3.6 Discussion of Potential Mechanisms

So far, we have established that the acquisition of the Washington Post by Amazon's founder Jeff Bezos resulted in increased coverage of topics related to Amazon, its products, and Bezos. Here, I will discuss some potential mechanisms that can explain the effect.

One potential explanation is a conflict of interest, where either the new owner demands more coverage of Amazon or the newspaper voluntarily gives more coverage of Amazon for the benefit of the owner and/or for the benefit of the newspaper. For instance, more coverage may raise more awareness and serve as an advertising tool for Amazon and its products.

Another explanation is a possible change in the readership preferences and the newspaper's response to this change. For instance, after the acquisition, consumers who were interested in Amazon or Bezos became more likely to subscribe to the Washington Post, and the newspaper responded to an increased demand by providing more stories about Amazon.

Next, it's possible that the Washington Post got an easier access to information about Amazon relative to other newspapers and responded by providing more coverage.

Finally, there could be a behavioral explanation, where journalists and editors of the Washington Post have updated their beliefs about the importance of news about Amazon relative to other news topics, and this resulted in more coverage about what they believe became more important.

The empirical strategy I use in this paper cannot distinguish among these possible mechanisms and, hence, it cannot say which of them are more or less likely to explain the results.

4 Effect on Coverage Tone

In this section I study another aspect of the possible acquisition effect – the effect on coverage sentiment, i.e., whether the acquisition resulted in more positive (or negative) coverage of Amazon and Jeff Bezos in the Washington Post.

4.1 Sentiment Measure

I construct two sentiment measures using two dictionaries: Harvard IV-4 psychosocial dictionary and Loughran/McDonald dictionary (LM dictionary onwards, Loughran and McDonald (2011)). The latter one was developed for sentiment analysis in financial texts, which can be relevant in this analysis since some stories about Amazon concern Amazon's financial indicators.

The Harvard dictionary has 77 predetermined word categories and the LM dictionary has six categories. However, according to some previous research,[67]

[67] see, e.g., Tetlock (2007), Tetlock et al. (2008), Loughran and McDonald (2011)

the most useful category for sentiment analysis is negative words[68]. Sometimes researchers also use positive words, but they are more often negated than negative words, which makes the inference more limited.

Thus, I construct a sentiment score for each news story where Amazon or Jeff Bezos are the main theme[69] using a list of negative words from each of the two dictionaries. The score is equal to the fraction of negative words in the total number of words in a story. The average sentiment score of a story for the Harvard dictionary is 0.088 with the standard deviation of 0.031. For the LM dictionary the average is 0.040 and the standard deviation is 0.024. The correlation between the two scores is 0.54, which indicates that there is some disagreement between the two measures on how positive or negative a story is.

4.2 Empirical Strategy and Results

To study the effect of the acquisition on Amazon-related coverage sentiment I use the difference-in-difference strategy, where the control group includes the same three newspapers as before. I estimate the following model:

$$neg_{it} = \alpha_i + \sum_{k=-1}^{4} \delta_k D_{kt} + \sum_{k=-2}^{4} \beta_k D_{kt} \cdot WPost_i + \varepsilon_{it}, \tag{11}$$

where neg_{it} is a sentiment score based either on the Harvard or LM dictionary; α_i - newspaper i fixed effect; $WPost_i$ - indicator for the Washington Post; and D_{kt} - a dummy variable for year k. As before, I normalize year -1, the year between two and one years before the acquisition, to zero.

Table 8 presents the estimation results. The estimates based on the Harvard dictionary don't show any evidence that the acquisition affected the sentiment of Amazon-related stories in the Washington Post. There are positive and negative coefficients, none of them are statistically significant. For the LM dictionary, the coefficients are negative, meaning that the stories became more positive, but only one coefficient is statistically significant at 95% level.

Overall, there is no convincing empirical evidence that the sentiment of the coverage of Amazon-related news changed as a result of the acquisition. This can be because there was no such a change in reality, e.g., because of potential reputational damages or because of a high degree of independence of

[68] The full list of categories and words can be found on http://www.wjh.harvard.edu/~inquirer/ for the Harvard dictionary and https://www3.nd.edu/~mcdonald/Word_Lists.html for LM dictionary

[69] I drop stories shorter than 250 words because they are too short to extract sentiment.

Table 8: The effect of the acquisition on the coverage sentiment

	(1)	(2)
	neg_{LM}	neg_{Harv}
β_0	-0.0160	-0.0125
	(0.0118)	(0.0123)
β_1	-0.0193*	0.00497
	(0.0107)	(0.0113)
β_2	-0.0235**	-0.0105
	(0.0114)	(0.0111)
β_3	-0.0149	0.00347
	(0.0108)	(0.0118)
β_4	-0.0130	0.0121
	(0.0105)	(0.0112)
Newspaper FE	Yes	Yes
Time FE	Yes	Yes
Clusters	103	103
Observations	768	768
R-squared	0.101	0.081

Notes: The coefficient for the year -1 is normalized to zero. Standard errors are clustered by newspaper-quarter
*** p<0.01, ** p<0.05, * p<0.1

the editorial board. Alternatively, this can be because the sentiment measure or the empirical strategy is not accurate enough to detect the effect and a different approach is required.

5 Conclusion

This work studies a conflict of interest that may potentially arise in a situation where media outlets are owned by someone who has vested interest in other businesses. These media outlets may benefit their owners by providing more favorable coverage of the affiliated companies. I study this issue in the context of the Washington Post, a major U.S. daily newspaper, that was acquired in 2013 by Jeff Bezos, the founder of Amazon.com.

I study the effect of the newspaper acquisition on coverage volume and coverage sentiment of Amazon-related stories. I found that the acquisition resulted in increased coverage of Amazon, but I didn't find convincing evidence on the effect on the coverage tone. The paper also discusses some potential mechanisms that can explain the results.

References

Allcott, H. and Wozny, N. (2014). Gasoline Prices, Fuel Economy, and the Energy Paradox. *The Review of Economics and Statistics*, 96(5):779–795.

Anderson, S. T., Kellogg, R., and Sallee, J. M. (2013). What do consumers believe about future gasoline prices? *Journal of Environmental Economics and Management*, 66(3):383–403.

Beattie, G., Durante, R., Knight, B., and Sen, A. (2017). Advertising Spending and Media Bias: Evidence from News Coverage of Car Safety Recalls. Working Paper 23940, National Bureau of Economic Research.

Bento, A. M., Goulder, L. H., Jacobsen, M. R., and von Haefen, R. H. (2009). Distributional and efficiency impacts of increased U.S. gasoline taxes. *American Economic Review*, 99(3):667–99.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890.

Burlig, F., Bushnell, J., Rapson, D., and Wolfram, C. (2021). Low energy: Estimating electric vehicle electricity use. *AEA Papers and Proceedings*, 111:430–35.

Busse, M. R., Knittel, C. R., and Zettelmeyer, F. (2013). Are consumers myopic? Evidence from new and used car purchases. *American Economic Review*, 103(1):220–56.

California Air Resources Board (2017). California's Advanced Clean Cars midterm review, https://ww2.arb.ca.gov/resources/documents/2017-midterm-review-report.

Crawford, G. S., Shcherbakov, O., and Shum, M. (2019). Quality overprovision in cable television markets. *American Economic Review*, 109(3):956–95.

Davis, L. W. (2019). How much are electric vehicles driven? *Applied Economics Letters*, 26(18):1497–1502.

Dellavigna, S. and Hermle, J. (2017). Does Conflict of Interest Lead to Biased Coverage? Evidence from Movie Reviews. *The Review of Economic Studies*, 84(4):1510–1550.

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

D'Haultfœuille, X., Givord, P., and Boutin, X. (2014). The Environmental Effect of Green Taxation: The Case of the French Bonus/Malus. *The Economic Journal*, 124(578):F444–F480.

Draganska, M., Mazzeo, M., and Seim, K. (2009). Beyond plain vanilla: Modeling joint product assortment and pricing decisions. *Quant Mark Econ*, (7):105–146.

Dubin, J. A. and McFadden, D. L. (1984). An econometric analysis of residential electric appliance holdings and consumption. *Econometrica*, 52(2):345–362.

Durante, R. and Knight, B. (2012). Partisan control, media bias and viewer responces: Evidence from Berlusconi's Italy. *Journal of the European Economic Association*, 10(3):451–481.

Eizenberg, A. (2014). Upstream innovation and product variety in the U.S. home PC market. *The Review of Economic Studies*, 81(3 (288)):1003–1045.

Ellman, M. and Germano, F. (2009). What do the Papers Sell? A Model of Advertising and Media Bias. *Economic Journal*, 119(537):680–704.

Fan, Y. (2013). Ownership consolidation and product characteristics: A study of the US daily newspaper market. *American Economic Review*, 103(5):1598–1628.

Federal Highway Administration (2017). 2017 National Household Travel Survey, U.S. Department of Transportation, Washington, DC. Available online: https://nhts.ornl.gov.

Gentzkow, M. and Shapiro, J. M. (2010). What Drives Media Slant? Evidence From U.S. Daily Newspapers. *Econometrica*, 78(1):35–71.

Gillingham, K., Rapson, D., and Wagner, G. (2015). The Rebound Effect and Energy Efficiency Policy. *Review of Environmental Economics and Policy*, 10(1):68–88.

Goldberg, P. K. (1998). The effects of the Corporate Average Fuel Efficiency standards in the US. *The Journal of Industrial Economics*, 46(1):1–33.

Grigolon, L., Reynaert, M., and Verboven, F. (2018). Consumer valuation of fuel costs and tax policy: Evidence from the European car market. *American Economic Journal: Economic Policy*, 10(3):193–225.

Groseclose, T. and Milyo, J. (2005). A Measure of Media Bias. *The Quarterly Journal of Economics*, 120(4):1191–1237.

Hausman, J. A. (1979). Individual discount rates and the purchase and utilization of energy-using durables. *The Bell Journal of Economics*, 10(1):33–54.

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

Hughes, J. E., Knittel, C. R., and Sperling, D. (2008). Evidence of a shift in the short-run price elasticity of gasoline demand. *The Energy Journal*, 29(1):113–134.

Jenn, A., Springel, K., and Gopal, A. R. (2018). Effectiveness of electric vehicle incentives in the United States. *Energy Policy*, 119:349 – 356.

Larcinese, V., Puglisi, R., and Snyder, J. M. (2011). Partisan bias in economic news: Evidence on the agenda-setting behavior of U.S. newspapers. *Journal of Public Economics*, 95(9):1178 – 1189.

Li, J. (2019). Compatibility and investment in the U.S. electric vehicle market. *Working Paper*.

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

Linn, J. (2016). The rebound effect for passenger vehicles. *The Energy Journal*, 37.

Loughran, T. and McDonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance*, 66(1):35–65.

Mannering, F. and Winston, C. (1985). A dynamic empirical analysis of household vehicle ownership and utilization. *The RAND Journal of Economics*, 16(2):215–236.

McConnell, V., Leard, B., and Kardos, F. (2019). California's evolving zero emission vehicle program: Pulling new technology into the market. *RFF Working Paper (19-22)*.

Muehlegger, E. and Rapson, D. S. (2021). Subsidizing low- and middle-income adoption of electric vehicles: Quasi-experimental evidence from California. *NBER Working Paper*, (25359).

Petrin, A. (2001). Quantifying the benefits of new products: The case of the minivan. *NBER Working Paper*, (8227).

Petrin, A. (2002). Quantifying the benefits of new products: The case of the minivan. *Journal of Political Economy*, 110(4):705–729.

Remmy, K. (2020). Subsidy design when firms can adjust product attributes: The case of electric vehicles. *Working Paper*.

Sallee, J. M. (2011). The surprising incidence of tax credits for the toyota prius. *American Economic Journal: Economic Policy*, 3(2):189–219.

Sallee, J. M., West, S. E., and Fan, W. (2016). Do consumers recognize the value of fuel economy? Evidence from used car prices and gasoline price fluctuations. *Journal of Public Economics*, 135:61–73.

Shriver, S. K. (2015). Network effects in alternative fuel adoption: Empirical analysis of the market for ethanol. *Marketing Science*, 34(1):78–97.

Small, K. A. and Dender, K. V. (2007). Fuel efficiency and motor vehicle travel: The declining rebound effect. *The Energy Journal*, 28(1):25–51.

Small, K. A. and Rosen, H. S. (1981). Applied welfare economics with discrete choice models. *Econometrica*, 49(1):105–130.

Springel, K. (2021). Network externality and subsidy structure in two-sided markets: Evidence from electric vehicle incentives. *American Economic*

Journal: Economic Policy (Forthcoming).

Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance*, 62(3):1139–1168.

Tetlock, P. C., Saar-Tsechansky, M., and Macskassy, S. (2008). More Than Words: Quantifying Language to Measure Firms' Fundamentals. *The Journal of Finance*, 63(3):1437–1467.

The International Council on Clean Transportation (2019). Update on electric vehicle costs in the united states through 2030. *Technical report*.

UBS (2017). Ubs evidence lab electric car teardown—disruption ahead? *Technical report*.

UC Davis PHEV Center (2020). Advanced plug-in electric vehicle travel and charging behavior final report. *Technical report*.

U.S. Environmental Protection Agency (2016). Technical update of the social cost of carbon for regulatory impact analysis - under executive order 12866. *Technical Support Document*.

U.S. Environmental Protection Agency (2020). Fast facts: U.S. transportation sector greenhouse gas emissions. *Technical report*.

Varadhan, R. and Roland, C. (2008). Simple and globally convergent methods for accelerating the convergence of any EM algorithm. *Scandinavian Journal of Statistics*, 35(2).

West, J., Hoekstra, M., Meer, J., and Puller, S. L. (2017). Vehicle miles (not) traveled: Fuel economy requirements, vehicle characteristics, and household driving. *Journal of Public Economics*, 145:65 – 81.

West, S. E. (2004). Distributional effects of alternative vehicle pollution control policies. *Journal of Public Economics*, 88(3):735 – 757.

Wollmann, T. G. (2018). Trucks without bailouts: Equilibrium product characteristics for commercial vehicles. *American Economic Review*, 108(6):1364–1406.

Xing, J., Leard, B., and Li, S. (2021). What does an electric vehicle replace? *Journal of Environmental Economics and Management*, 107.

Appendices

Appendix to Chapter 1

Appendix A Calculation of ZEV credit prices

Two sources are used to estimate the prices of ZEV credits, i.e., credits that manufacturers earn by selling BEVs in the ZEV states. The first one is Tesla's quarterly shareholder letters, where the company reports its quarterly revenues from selling ZEV credits. The second source is the ZEV states' annual reports on ZEV credit balances and transfers between manufacturers. The reports were not available for some of the states before 2015, so I only use the reports starting from 2015. Table A1 shows the number of ZEV credits transferred by Tesla in each ZEV state for each year during the 2015-2018 period, and Tesla's revenues from selling ZEV credits for the same period. ZEV credit prices are computed by dividing Tesla's revenues by the total number of credits it transferred in a given period. Yearly estimates range from \$1,506 to \$2,404 per credit. After adjusting for inflation, the average price for the entire period is estimated to be \$2,200 per credit (2018 dollars).

To estimate prices of TZEV credits, i.e., credits earned by PHEVs, I use the states' ZEV credit reports and analyze cases when manufacturers exchange credits of different types. The reports have no information on monetary transfers between carmakers, so I have to assume that no money were transferred in the exchange cases that I focus on. The idea is to infer the price of TZEV credits relative to the price of ZEV credits. For example, in 2018 in Maryland, Fiat-Chrysler transferred 58 ZEV credits to Honda and received back 650 PZEV credits. Hence, I conclude that one ZEV credit is worth 9.15 PZEV credits. Next, in 2017 in California, General Motors transferred 6,000 TZEV credits to Honda and received back 2,500 ZEV credits and 12,700 PZEV credits, which is 3,888 ZEV credits in total, after converting PZEV credits into ZEV credits. Hence, I conclude that 6,000 TZEV credits are worth 3,888 ZEV credits, i.e one TZEV credit is worth 0.65 ZEV credits. A similar analysis of several more exchange cases resulted in similar estimates, so I ended up assuming the relative price of a TZEV credit being 0.7 of the price of a ZEV credit, or \$1,540.

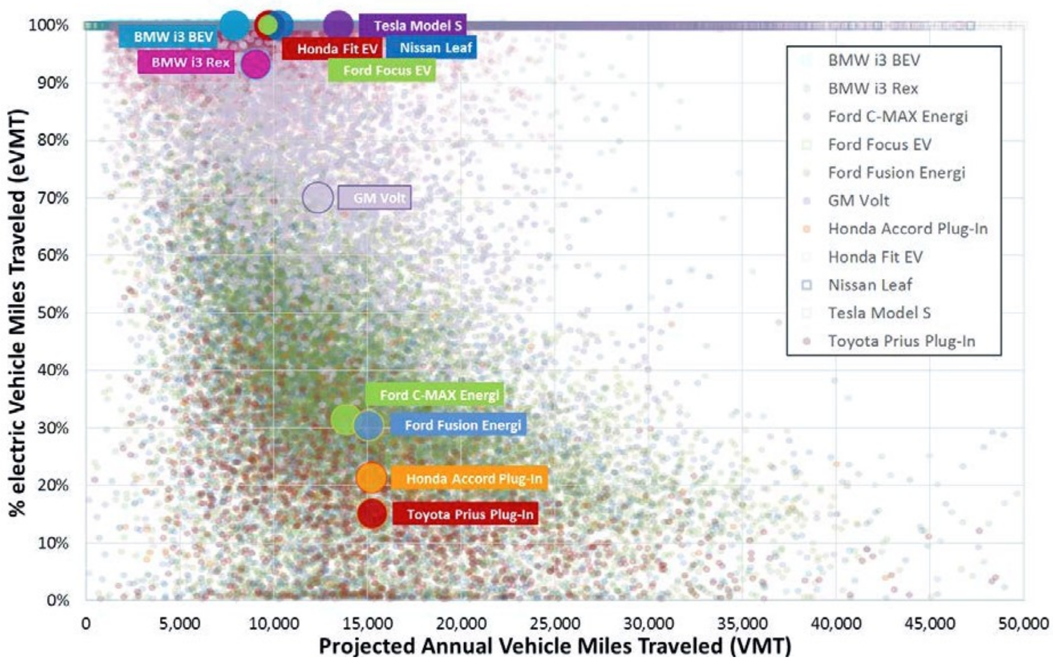
Table A1: ZEV credit prices calculated from Tesla's ZEV credit transfers and revenues

State	2015	2016	2017	2018
California	44,421	80,227	51,776	88,214
Connecticut	536	264	2,236	3,220
Maine	0	111	54	206
Maryland	756	360	2,862	6,604
Massachusetts	944	375	4,684	7,498
New Jersey	2,697	2,550	8,036	8,512
New York	37,598	850	10,827	8,854
Oregon	255	215	2,434	6,239
Rhode Island	88	0	85	0
Vermont	0	58	130	224
Total credits	86,952	84,737	79,935	123,108
Tesla revenues	\$170M	\$203.7M	\$120.4M	\$281.7
Revenue per credit (credit price)	\$1,955	\$2,404	\$1,506	\$2,288

Notes: Each cell in the upper part of the table is the number of ZEV credits transferred by Tesla to other carmakers in a given state and year. ZEV credit prices are computed by dividing Tesla's revenues from selling ZEV credits by the total number of transferred credits.

Appendix B Individual PHEV utility factors

Figure B1: Individual utility factors, CARB data



Source: The California's advanced clean cars midterm review. Appendix G. (California Air Resources Board, 2017)

Appendix C Data and computational details

C.1 Mileage and demographic data construction

I use the 2017 NHTS data to construct a joint distribution of household demographics and driving needs. The survey contains data on household demographics, including household size, income, location, ages and education levels of household members etc. It also includes data on household vehicles, including the number of vehicles, their makes, models, model years and odometer readings at the day of survey. I compute average annual mileage for each household vehicle and assume that these average mileages represent driving needs that households consider when they shop for a new car, and I assume that these driving needs are fixed and don't depend on vehicle choice. I focus only on passenger cars, SUVs, minivans and pickup trucks, i.e., excluding, for example, motorcycles, RVs etc. Also, since I estimate the model only for new vehicle purchases, I restrict my attention to relatively new vehicles from the survey, focusing on the model-years from 2013 to 2017.

Estimating average mileages is complicated by the fact that the exact purchase dates of vehicles are unknown, and making careful assumptions about this is important because of relatively short ownership periods. In general, I assume that a vehicle was purchased in the middle of the calendar year corresponding to the vehicle's model year, and define the ownership period length as the difference between the month when the survey was taken and the assumed purchase date. Whenever possible, I use other information to improve precision of the ownership period length computation. For example, the survey has data on the number of months of ownership if a vehicle was purchased less than a year ago, in which case I use this as the ownership period length. Also, I incorporate release dates of some popular vehicles if their sales start at a time point other than around the beginning of a new calendar year.

The joint distribution of mileage and demographics is then used to construct moment conditions for average mileage of gasoline and conventional hybrid vehicles, and for average demographics of buyers of various vehicle categories, as described in Section 6 of the paper. For the demand estimation, I construct separate distributions for each geographic market, so that for each market consumers are drawn at random from the corresponding distributions.

The moment conditions for average mileages of EVs are constructed using the CARB data, from which I observe only the averages and the number of observations used to compute those averages. To construct appropriate weights for these moment conditions for the demand estimation, I approximate the variances of the moments using the number of observations from the CARB data and standard deviations of mileage distributions of various vehicle models from the 2017 NHTS.[70] In particular, I assume that the standard deviation of a vehicle's mileage distribution is a linear function of the mean of this distribution. I estimate the parameters of this linear function using averages and standard deviations of mileage distributions of various vehicles from the 2017 NHTS. Then, I use the estimated relationship to infer the standard deviations of the mileage distributions for the EVs from the CARB data, and, finally, compute standard errors and variances of the average mileage estimates from the CARB data.

[70] I also use these variance approximations to construct confidence intervals for the descriptive evidence in Section 3 of the paper.

C.2 Demand estimation

The demand estimation procedure follows Berry et al. (1995) and Petrin (2002), except that it doesn't use supply side moment conditions, i.e., those derived from the first order conditions of firms' profits. The model parameters are estimated by minimizing a GMM objective function, which is a weighted average of two sets of moment conditions. The first set is formed by multiplying the vector of errors ξ_{jm}'s by a set of instruments Z_{jm}, that include state subsidies and vehicle characteristics, included in the utility function, except price and battery range. In the first estimation step, these moment conditions are weighted with a consistent estimate of matrix $E[ZZ']^{-1}$, where Z is a (column) vector of instruments. In the second step, I use the first step estimates of ξ_{jm} to construct an estimate of the optimal weighting matrix $E[Z\xi\xi'Z']^{-1}$.

The second set of moment conditions matches demographic and mileage micromoments from the 2017 NHTS and the CARB data to the corresponding model predictions. The moment conditions built using the 2017 NHTS are weighted by the inverse of an estimate of the variance-covariance matrix of these moments, which is constructed using the delta method, as described in Appendix B.1 of Petrin (2001). For the CARB data moments, the absence of individual-level data didn't allow me to estimate a variance-covariance matrix in a similar way, hence I weighted these moments by inversed approximated variances (see the previous subsection).

To solve for δ values that equate observed and predicted by the model market shares, instead of using the contraction mapping suggested in Berry et al. (1995), I used the SQUAREM method proposed by Varadhan and Roland (2008). It gave identical δ values and significantly improved the speed of the objective function computation.

C.3 Solving for counterfactual policy prices and ranges

For all the counterfactual exercises I assume that firms update prices and, in case of American carmakers, ranges of BEV. After solving for the new equilibrium prices and ranges, I compute the other outcomes of interest. The new prices and ranges are computed by the following procedure:

1. Start by guessing initial prices, p^0, and ranges, r^0.
2. Holding the range values fixed, use the price FOC to solve for a vector of new prices, p^1.
3. For the vector of new prices p^1, use the range FOC to solve for a vector of new ranges, r^1.

4. Iterate 2 and 3 until convergence.

There is no guarantee of the equilibrium uniqueness. However, it makes sense to expect that the counterfactual outcomes should not be "too far" from the status quo. Hence, I start the procedure with the status quo prices and ranges for all the exercises. For robustness, I also tried some other starting points, and arrived at the same results.

Appendix D Usage of EVs: Theory

In this section I present a simplified version of the demand model, which illustrates how the inconvenience costs of charging affect the average mileage driven by BEV and PHEV drivers relative to drivers of traditional vehicles. I assume that there is a consumer, who needs to drive d miles per year. The consumer's choice set includes three vehicles, a BEV, a PHEV, and a traditional (gasoline) vehicle, which can be a pure gasoline or a conventional hybrid car. Let the purchase prices be p_{bev}, p_{ph} and p_g for the BEV, the PHEV and the traditional car, respectively. The fuel economy of the vehicles will be characterized by prices per mile: w_e for the BEV, w_g for the traditional car, and $w_{ph,e}$ and $w_{ph,g}$ for the PHEV, when driven on electricity and gas, respectively. I assume that driving on electricity is cheaper, i.e., $w_g > w_e$, $w_{ph,g} > w_{ph,e}$ and $w_g > w_{ph,e}$. The electric ranges of the BEV and the PHEV are given by r_{bev} and r_{phev}, respectively. Other characteristics of the three vehicles are assumed to be identical across the vehicles.

I assume that the consumer utility specification includes price, fuel costs, and BEV and PHEV inconvenience costs, ignoring other vehicle characteristics, since they do not vary across the cars.

Next, I will describe consumer choice and how it depends on consumer mileage d. First, I will consider choice between the traditional car and the BEV, then, choice between the traditional car and the PHEV, and, finally, choice between the BEV and the PHEV.

BEV and traditional vehicle. Consumer utility for the traditional car is given by

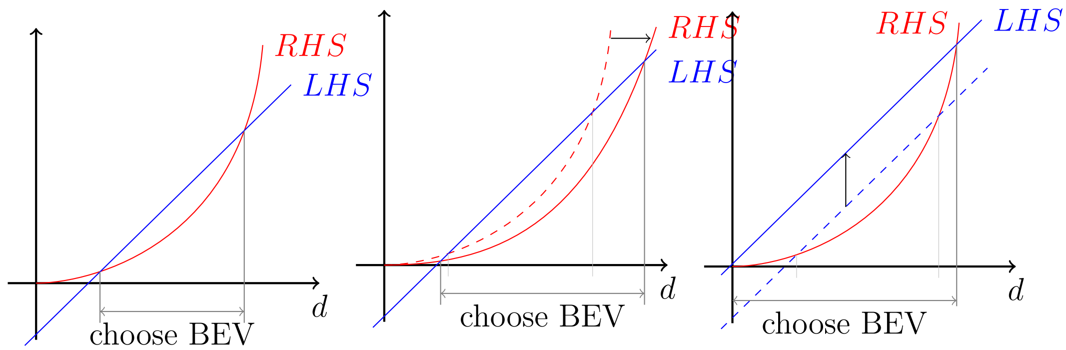
$$u_g = -\alpha^p p_g - \alpha^f w_g d,$$

and the utility for the BEV is given by

$$u_{bev} = -\alpha^p p_{bev} - \alpha^f w_e d - c_{bev}(d; r_{bev}, chst),$$

where $c_{bev}(d; r_{bev}, chst)$ is the BEV inconvenience cost, which is assumed to be continuous, strictly convex and strictly increasing in d, with $c_{bev}(0; \cdot, \cdot) = 0$.

Figure D1: Choice between BEV and traditional car



(a) Benchamrk case

(b) Increase in range or charging stations

(c) Increase in subsidy for BEV

Notes: LHS is the left-hand side and RHS is the right-hand side of inequality (12).

I also assume that $c_{bev}(d; r_{bev}, chst)$ is decreasing in range, r_{bev}, and charging station availability, $chst$. A consumer chooses the BEV when

$$\alpha^p(p_g - p_{bev}) + \alpha^f(w_g - w_e)d > c_{bev}(d; r_{bev}, chst). \tag{12}$$

Figure D1a shows the left-hand side and the right-hand side of this inequality, assuming $p_{bev} > p_g$. While the left-hand side is increasing linearly in d, the right-hand side is increasing and convex in d. Hence, as d increases, at some point, the inconvenience costs start to outweigh the fuel cost savings, and the consumer prefers to buy the traditional car despite the lower fuel costs of the BEV. An increase in battery range or the number of charging stations (Figure D1c) results in attracting drivers from a wider range of d, especially those with larger d, which is consistent with the evidence that longer range BEVs are driven more miles on average than shorter range BEVs. Figure *D1b* shows the effect of a policy that equalizes the purchase prices of the BEV and the traditional car, e.g., a subsidy for the BEV. Unlike the previous case, the policy does not affect the inconvenience costs. It attracts consumers from a wider range of d, too, but affects lower-mileage consumers to a larger extent.

PHEV and traditional vehicle. Now consider a choice between the traditional vehicle and the PHEV. The utility of the traditional vehicle is the same as before. The utility for the PHEV is given by

$$u_{ph} = -\alpha^p p_{ph} - \alpha^f \left[\psi^* w_{ph,e} + (1-\psi^*) w_{ph,g}\right] d - c_{phev}(\psi^* d; r_{phev}, chst),$$

where $c_{phev}(\psi^* d; r_{phev}, chst)$ is the PHEV inconvenience cost function, which is continuous, strictly convex and strictly increasing in $\psi^* d$, decreasing in range,

r_{bev}, and charging station availability, $chst$, with $c_{phev}(0;\cdot,\cdot) = 0$. ψ^* is the optimal share of d driven on electricity. I assume that c_{phev} has a functional form that is similar to the full model:

$$c_{phev}(\psi^* d; r_{phev}, chst) = \theta(\psi^* d)^2 f(r_{phev}, chst),$$

where $f(r_{phev}, chst)$ is decreasing in r_{phev} and $chst$. As in the full model, ψ^* is given by the solution of a problem where a consumer is minimizing the sum of fuel expenses and inconvenience costs:

$$\psi^* = \min\left[1, \frac{w_{ph,g} - w_{ph,e}}{2df(r_{phev}, chst)}\right].$$

Since in reality nearly all PHEV drivers drive less than 100% of mileage on electricity, I will focus on the case when $\psi^* < 1$. The utility for the PHEV then can be rewritten as

$$u_{ph} = -\alpha^p p_{ph} - \alpha^f w_{ph,g} d + \frac{(w_{ph,g} - w_{ph,e})^2}{2f(r_{phev}, chst)}\left(\alpha^f - \frac{\theta}{2}\right),$$

and a consumer chooses the PHEV when

$$\alpha^p(p_g - p_{ph}) + \frac{(w_{ph,g} - w_{ph,e})^2}{2f(r_{phev}, chst)}\left(\alpha^f - \frac{\theta}{2}\right) + \alpha^f(w_g - w_{ph,g})d > 0. \tag{13}$$

The last term of the left-hand side of this inequality plays a key role in defining whether drivers with higher values of d are more or less likely to adopt the PHEV. If the closest substitute for the PHEV is a vehicle with $w_g > w_{ph,g}$, i.e., a vehicle that is less fuel efficient than the PHEV in the hybrid (gasoline) mode, then higher-mileage drivers would prefer the PHEV. If, instead, the closest substitute is a car that is more fuel efficient than the PHEV in the hybrid mode ($w_g > w_{ph,g}$), then higher-mileage drivers would be less likely to adopt the PHEV. An example of the latter would be the choice between a PHEV and a conventional hybrid car. A PHEV, especially a PHEV with a larger battery, can be less fuel efficient when driven on gasoline than a similar conventional hybrid car, because it has to carry more weight from a larger battery and other related components. In this case, since higher-mileage drivers tend to drive more miles on gasoline, then the difference $w_g - w_{ph,g}$ for them is increasingly important when d is increasing, which makes them more likely to choose a conventional hybrid car. This, in particular, can explain a relatively low average mileage of Chevrolet Volt when compared to shorter-range PHEVs and conventional hybrids (Section 3 of the paper).

BEV and PHEV. Finally, I consider a choice between the BEV and the PHEV. A consumer chooses the BEV when

$$\alpha^p(p_{ph}-p_{bev})-\frac{(w_{ph,g}-w_{ph,e})^2}{2f(r_{phev},chst)}\left(\alpha^f-\frac{\theta}{2}\right)+\alpha^f(w_{ph,g}-w_e)d > c_{bev}(d;r_{bev},chst). \tag{14}$$

This inequality is similar to the one from the case of the BEV and the traditional vehicle, except for the second term in the left-hand side, which, depending on its sign, can make choice of the PHEV more or less attractive. The left-hand side of inequality (14) is increasing linearly in d, while the right-hand side is increasing and convex in d. Hence, higher-mileage drivers would prefer the PHEV to the BEV.

Appendix to Chapter 2

More on Sentiment Measure

To understand what the sentiment score measures we can take a look at stories' headlines with the highest and lowest fraction of negative words. For example, these are the top-five stories from the New York Times with the highest fraction of negative words, according to the Harvard dictionary:

1. Amazon v. the States
2. Amazon's Tax Dodge: Online retailer takes advantage of California's dysfunctional politics
3. Trump, Amazon and 'Internet Taxes': What Did He Mean?
4. Amazon and California in Deal on Tax
5. In Tax Fight, Amazon Hands Baton To eBay

And these are the top-five negative stories according to the LM dictionary:

1. Amazon v. the States
2. Amazon's E-Book Business Being Investigated in Europe
3. German Publishers Seek Amazon Inquiry
4. The Kindle Fire And a Debate On Tablets
5. Amazon's Prophet And Losses

We see that only one story is in both lists. The top negative stories according to the Harvard dictionary are all related to the sales tax dispute[71], while the LM dictionary list has stories on various topics. Most likely, the reason is that the word "tax" is classified as negative by the Harvard dictionary, unlike the LM dictionary. In general, it can be noticed that the top negative stories are generally related to some disputes or conflicts around Amazon, such as sales tax collection, conflicts with publishers, lawsuits and other.

[71] The dispute is about whether Amazon should collect sales taxes from customers in states where it doesn't have physical presence

The top five stories from the New York Times with the smallest fraction of negative words according to the Harvard dictionary include:

1. Amazon Publishing Push Grows to Children's Books
2. At Amazon Art Site, Everyone's a Critic
3. The Amazon That Readers Can Walk Into
4. Amazon to Open Manhattan Retail Store
5. Amazon to Pursue Education Technology With a Marketplace for Teachers

and according to the LM dictionary:

1. Daily Report: The Next Voice You Hear Will Be Amazon's
2. Walmart Sticks With Arkansas Headquarters as Amazon Plays the Field
3. Amazon to Open Manhattan Retail Store
4. Amazon's Boom In the Cloud
5. Voters Speak, And Amazon Orders Shows

Stories with the smallest fraction of negative words typically don't discuss any conflicts, they rather tell about Amazon's products and services.