Essays in Industrial Organization

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Boyoung Seo

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy

Amil Petrin, Advisor

May, 2016

Acknowledgements

I am greatly indebted to my advisor, Amil Petrin, for his guidance and support. I am not a type of person who is smart and persistent enough to thrive without help. Without intense training and advices from my committee members, Amil Petrin, Thomas J. Holmes, Joel Waldfogel, and Naoki Aizawa, it would have been difficult to pull through.

This research has benefitted from conversations with Ellen McGrattan, Kyoo-il Kim, Keaton Miller, Thomas Quan, Kevin Williams, Kailin Clarke, Ethan Singer, and Bitmaro Kim as well as the members of the Applied Micro workshop and participants at seminars I have presented at the University of Minnesota. I give special thanks to my best friend and colleague, Matthew H. Shapiro, for all his help.

This dissertation was supported by the University of Minnesota Doctoral Dissertation Fellowship. I thank Melissa Norton at the Washington State Liquor and Cannabis Board for providing data.

Abstract

This dissertation contains three essays, each of which is pertinent to topics in empirical Industrial Organization. The second and third chapters are coauthored with Matthew H. Shapiro and Amil Petrin, respectively.

In the first chapter[1] , I quantify the consumer benefit of introducing one-stop shopping as a new shopping choice and estimates the benefit to firms in turn. Firm scope benefits consumers by allowing them to purchase multiple goods at one location like supercenters, malls, or department stores. I exploit an exogenous change in the allowable firm scope in Washington State, which recently deregulated the retail liquor industry to allow liquor sales in grocery stores. After deregulation, the number of liquor-selling stores is increased fourfold, and 75% of the liquor shopping has been done by one-stop shopping with groceries. Moreover, the liquor quantity sold has increased despite the increased after-tax price of liquor, implying that the choice set of shopping trips has improved. To disentangle the value of one-stop shopping from the value of reduced shopping distance due to more liquor-selling stores, I build a structural demand model of choices for shopping trips. I use household panel and retailer sales data from both before and after deregulation and extend the standard method to allow for endogenous prices to the setting where a store can have two separate qualities of grocery and liquor sections. The estimated consumer benefit of one-stop shopping is $2.52 per trip per household, which is 8% of the household's expenditure on liquor. Selling liquor inside of a grocery store increases its grocery sales by 4.5%, and liquor sales are increased by 30% compared to being sold outside of the grocery store.

In the second chapter, we describe a model and estimation strategy to assess the efficacy of several US federal investment projects amounting to $130 million to build out the foundation of an plug-in electric vehicle (PEV) charging network and to encourage purchase of these vehicles. Using a new micro-level data set of electric vehicle purchases in California to estimate a rich discrete choice model of automobile demand, we analyze

[1] Any empirical evidence presented in this chapter is derived based on data from the Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

whether these charging stations have had a significant role in the adoption of electric vehicles in California over the past several years and weigh them against several other policy alternatives.

In the last chapter, we suggest an alternative approach to identify demand and supply in discrete choice demand set up. A major contribution of Berry et al. (1995) (BLP) is to show how to transform the market shares in a non-linear discrete choice demand system into product-quality indices that are linear in price, product characteristics, and the demand error, so standard IV techniques can once again be used to estimate demand. They treat price as endogenous but assume that the product characteristics observed by the researcher are uncorrelated with the demand error, that is, the characteristics observed by consumers and producers but not by the researcher. As in Spence (1976) (e.g.) if firms set observed and unobserved characteristics at the same time then this assumption may not hold. We mimic BLP exactly but instead of using their mean independence assumption we learn about demand and supply by assuming firms are maximizing expected profits given their beliefs about preferences, costs, and competitors' actions when they choose observed and "unobserved" product characteristics. We allow firms' information sets at the time they choose characteristics to potentially include other firms' product characteristics, demand and cost shocks, signals on all of these, or no information at all on them in the setting of Hansen (1982). Ex-post firms may wish they had made different decisions and our identification is based on the assumption that firms are correct in their choices on average. Using the same automobile data from BLP, we find that some of the slightly puzzling parameter estimates of BLP go away as all of our parameter estimates are of the correct sign. We find significantly more precise estimates given the same exact data equivalent to approximately a sixteen-fold increase in the number of observations. We strongly reject the standard identification assumption; the conditional correlations of the demand and cost unobservables with observed characteristics equal 0.85 and 0.7 respectively.

Contents

List of Tables

List of Figures

Chapter 1

Firm Scope and the Value of One-Stop Shopping in Washington State's Deregulated Liquor Market

1.1 Introduction

Firms have increasingly expanded their scope of products over time to offer consumers economies of scope in purchasing; the convenience of saving time by one-stop shopping. For example, food retailers have evolved to include pharmacies, banks, and clothing, such as Wal-Mart, providing one-stop shopping for general merchandise and groceries. However, little empirical work has been done to quantify the value of one-stop shopping.

This paper estimates the value of one-stop shopping to consumers and to stores with four different data sources. The deregulation of Washington's liquor market creates a rare environment where one-stop shopping for groceries and liquor is introduced to the consumer choice set. As part of this paper, though not the main purpose, I also analyze the welfare effects of the deregulation of policies that limit the scope of products carried by retailers. Washington privatized liquor stores in 2012 and allowed grocery stores to sell liquor. Before deregulation, consumers could purchase liquor only from state stores which did not sell groceries, but now consumers have the option of one-stop shopping for groceries and liquor. The number of liquor-selling stores has increased fourfold and 97% of this entry is from grocery stores.

The market outcomes in Washington provide qualitative evidence that the consumer choice set of liquor shopping has improved. Post deregulation, Nielsen household panel data shows that 75% of liquor shopping has been done by one-stop shopping with groceries. Moreover, state tax revenue data reveals that the quantity sold has increased despite the increase in average after-tax liquor price, due to an extra tax imposed on the liquor retailers as part of deregulation policies. These empirical facts suggest that the choice set of shopping trips has improved. However, this improvement could stem from the introduction of one-stop shopping with groceries as a new type of shopping, or from reduced distances of general liquor shopping trips due to the introduction of a large number of new liquor selling stores.

To disentangle the impact of one-stop shopping from other market outcomes of deregulation, I construct a structural demand model of choices of shopping trips where a trip consists of store choices for groceries, liquor, or both. I include extensive control variables in the model: seasonality and store-chain fixed effects by linking the store-level data from Nielsen and from the state. The value of one-stop shopping is identified by tracking consumers' switching patterns of shopping trips for groceries and liquor in response to the change in scope of offerings at grocery stores. From the household panel data, I observe where consumers purchase liquor, whether they shop for groceries at the same time, how much they paid, and demographic information. Consumers' switching to one-stop shopping after deregulation enables me to identify the value of one-stop shopping, conditional on demographics, price, and travel distance.

Even though the demand model uses a rich set of control variables, I do not observe some characteristics related to shopping, such as the quality of the store's grocery and liquor selection, promotion, and advertisement. Therefore, a common endogeneity problem in demand estimation can exist; the price coefficient is biased upward if prices are positively correlated with the unobserved characteristics. I extend the standard approach suggested by Berry (1994) and Berry et al. (1995) (henceforth BLP) to allow for endogenous prices to the setting where there are two separate unobserved shopping characteristics per choice – unobserved store qualities of grocery sections and liquor sections. Extending the BLP's method, I control for the store-grocery specific terms and store-liquor specific terms. I locate those terms by matching model-predicted market shares of grocery sales and liquor sales of each store to data. I obtain a precise measure of the market shares by linking store-level revenue data from Nielsen and from the state. It is new in the demand estimation literature to allow for two demand errors and their correlation with prices.

I find that the value of one-stop shopping is \$2.52 per trip per household, which is 8% of the expenditure on liquor, or \$4.46 per household annually. Allowing one-stop shopping contributes approximately half of the overall consumer gain from deregulation as the overall gain is \$9.35 per household in a year, which is a 55% increase in consumer surplus. The rest half of the gain stems from reduced distances of liquor shopping trips. Even though the average consumer welfare improves after deregulation, it was not Pareto improving: consumers in rural areas lose 63 cents of their consumer surplus while those in cities gain \$10.93. The reason is that rural areas have almost no increase in the number of liquor-selling stores but they face the highest increase in prices. Moreover, the gain is not distributed uniformly across consumers: consumer gains increase as population density and income level increase.

Offering one-stop shopping opportunities is also beneficial to firms. Selling liquor inside of a grocery store increases its grocery sales by 4.5% on average, and a stand-alone grocery store loses 0.5% of its grocery sales. The aggregate grocery sales remain unchanged when allowing liquor sales in grocery stores. On the other hand, liquor sales are increased by 30% if liquor sections are located inside of grocery stores, rather than being located outside. In contrast, liquor sales at a stand-alone liquor store decrease by 0.81% on average. The aggregate liquor sales increase by 20% when liquor can be sold in grocery stores. These results suggest that there are strong incentives for stores to expand their scope of product offerings not only because it raises revenue but also because stores lose revenue if they do not expand product lines.

The convenience benefit of one-stop shopping has important implications to firms' assortment decisions or decisions on the scope of product offerings. Previous literature on firm scope has primarily focused on the supply side benefits as a motivation to expand firm scope, such as efficiency gains or cost reduction (Holmes (2001); Ellickson

(2011); and Basker et al. (2012)). As a flip side of putting a new category of products into a store, this paper addresses the demand-side economies of scope, which arise from one-stop shopping convenience. Therefore, my paper opens up the discussion about the one-stop shopping benefit as a determinant of which categories of products generate the most attraction if sold under one roof.

There are several differences between this paper and previous research on one-stop shopping. Berger et al. (1996), Yuan and Phillips (2008), and Cummins et al. (2010) examine the benefits of one-stop shopping to financial firms. Their studies rely on the assumptions of firms' optimal decisions and focus on measuring the revenue raised from the variety of financial products to firms without modeling demand. Likewise, Sen et al. (2013) estimate revenue economies of scope to retailers raised by one-stop shopping for groceries and gas, but without accounting for price. In contrast, my paper directly models the demand for one-stop shopping and identifies the benefits to consumers based on consumer switching behavior in shopping. Second, most papers (Messinger and Narasimhan (1997) etc.) approximate the extent of one-stop shopping by the store size or the variety of products in a store. However, these proxies measure the *love of variety* rather than the convenience benefit of one-stop shopping. Finally, Arentze et al. (2005) study consumers' choices of one-stop shopping, which arise from complementarities between groceries and other goods, but without including prices in demand.

Consumer benefits from one-stop shopping can be interpreted as complementarities between product categories as Betancourt and Gautschi (1990) pointed out. In turn, the demand setup resembles the setup in the existing empirical studies on complementarity, such as Gentzkow (2007) and Wakamori (2015). Gentzkow's paper differs from my paper in that the choice set in his paper is small enough to have observations on demand of each choice, and some prices are zero. Wakamori's notion of complementarities is closer to the idea of a love of variety, raised within a product category (automobiles), whereas complementarities from one-stop shopping in this paper are raised in between categories (grocery and liquor).

This paper is also different from previous studies on the welfare effects of deregulation on consumers in the retail industry. Seim and Waldfogel (2013) examined the effects of privatization in Pennsylvania as a counterfactual experiment. They consider reduced travel distance due to the entry of private stores as the source of potential gains to consumers, holding price fixed. In contrast, I account for economies of scope from the expanded offerings by retailers and changes in prices in addition to store entry in the welfare analysis with the realized data.

The rest of the paper is organized as follows. Section (1.2) documents changes in market outcomes as a result of privatizing the retail liquor industry and describes household panel data used for the documentation. The demand model is described in Section (1.3). Section (1.4) describes store-level datasets used in demand estimation in addition to the household data. Section (1.5) details the identification and estimation approach used in the paper. Then, the results for estimated demand and changes in welfare from privatization are presented in Section (1.6).

1.2 Liquor Deregulation in Washington

1.2.1 Background

Initiative 1183 was approved in November 2011 and enacted in June 2012 to privatize liquor stores and introduce license system in Washington[1] . Prior to privatization, liquor for off-premises consumption was only sold through state stores. The number of stores had been capped by 330; 167 state-run and 163 contract stores. Washington auctioned off the state-run stores and granted the rights to the operators of the former contract stores.

There were also three primary policies implemented as part of privatization: (1) introducing 17% additional tax, leading to 37.5% effective tax on revenue, (2) requiring a minimum store size of 10,000 square feet, which is the typical size of a full-service neighborhood grocery store[2] , and (3) allowing grocery stores to sell liquor.

Change in the Number of Liquor-Selling Stores

As 83% of the existing grocery stores satisfying the minimum store size requirement started selling liquor, the number of liquor retailers increased from 330 to 1,373 by

[1] Beer and wine sales have been allowed in grocery stores since 1969. The Washington State Liquor and Cannabis Board (WSLCB) uses the term spirits instead of liquor.

[2] The only exemptions are granted to the former state stores, and none of them satisfy the requirement.

six months after privatization. Figure (1.1) compares store locations before and after privatization. The red dots indicate the former state stores, which have exited after privatization. The green dots indicate those which still operate. Lastly the blue dots are the entrants as of February 2015. The number of liquor-selling stores increased in most areas, but the distribution of the increase is not uniform across urban and rural areas. I account for the distributional effects of deregulation and examine the gains and losses across areas in the results section.

Combined with the minimum store size restriction, allowing grocery stores in the liquor retail business led 97% of the new liquor-selling stores are existing grocery stores. Table (1.1) summarizes the change in the number of liquor stores by grocery and stand-alone liquor stores. Post privatization 80% of the stores are grocery stores, and they account for 75% of the total market share of liquor. On the other hand, there are only 20 to 40 stand-alone liquor stores besides the former state stores. The number of former state stores has been decreasing down to two third. Notice that *liquor store* in this paper refers not only stand-alone liquor stores but also grocery stores selling liquor.

Increase in Price and Quantity Sold

Table (1.2) summarizes the change in liquor price and quantity sold between a year before and a year after deregulation. The post-tax prices (consumer prices) increased from \$20.71 to \$23.03 on average, which is an 11% increase. Price is defined as the average post-tax price per liter weighted by liter sold, deflated in January 2006 dollars, by using the aggregate sales and quantity data from the Washington State Department of Revenue. On the other hand, the average pre-tax price (shelf price) fell from \$15.02 to \$14.61, which may have been the consequence of the increase in competition or efficiency gains. Despite the increase in price, liquor quantity sold also increased by 5.97%, which is equivalent to 1.75 liters per household who has purchased liquor from a store at least once a year. These empirical facts provide first pass evidence that the choice set for liquor shopping trips has improved. See Appendix (A) for robustness of the increases in price and quantity conditional on the week, month, national trend, and brand.

1.2.2 Consumer Shopping Behavior

Following suggestive evidence from the household panel data shows that the improvement of the choice set of shopping trips can arise from two channels – *convenience* of one-stop shopping, which saves time and effort of having to find a parking spot, waiting at the checkout line, and carrying children between stores, and reduced total distance of shopping trips.

Household Panel Data

From the Nielsen Consumer Panel dataset, I describe consumer's shopping behavior for groceries and liquor. Each year about 1,500 registered panelists in Washington logs every store visit and purchase from the store, including store, visit date, products purchased, quantity purchased, and the price paid. The panelists are given a device to scan each product purchased and manually enter the quantity and price into the device. About one-third of these panelists purchase liquor at least once each year. The dataset also provides the panelist's demographics, including home zip code, income, race, education, and the number of children under 7. Zip code is used to measure the travel distance to the store. Even though each store is given a unique ID by Nielsen, the actual store location is masked in the dataset. Therefore, I approximate the distance by assuming that the panelist visited the closest store conditional on matching the store type. Appendix (A) describes how the distance is derived, how grocery and liquor shopping trips are defined as well as summary statistics of the panelists.

Switching to One-Stop Shopping

After privatization, 75% of the liquor shopping is done by one-stop shopping with groceries. Conditional on purchasing liquor, Table (1.3) summarizes share of shopping trips by type: liquor-shopping-only (L only), grocery and liquor shopping at two stores (GL2), and grocery and liquor shopping at one store (GL1), which is one-stop shopping. The one-stop shopping received zero shares before privatization because it was not an available option. By tracking consumers' choices of shopping trips in the data, I observe 65% of the households who used to choose GL2 the most before privatization switched to GL1 after. Similarly, 60% of the households who chose L only the most before switched to GL1 after. Moreover, given that a stand-alone liquor store is less than 0.1 miles away from a grocery store which also sells liquor, 86% of the times consumers chose one-stop shopping at the grocery store. All these pieces of empirical evidence show that *convenience* of making fewer stops is the important margin which led to switches to one-stop shopping.

Reduced Shopping Trip Distance

Due to the increase in liquor-selling stores, the average distance to the closest liquor store from zip code is reduced by 0.33 miles (0.19 in median), which is a 19% decrease (17%), shown in Table (1.4). Distance between a zip code and a store is derived by averaging distance between each census block within the zip code and the store, weighted by the population of the block. It suggests that reduced distance to a liquor store is another potential channel of the consumer gains from deregulation. The table also shows that the reduction in distance is not evenly distributed across the state because the increase number of stores is not evenly distributed; the higher the population density, the larger the percentage decrease in distance to the closest liquor store.

By using the household panel data, I compare the total trip distance conditional on shopping for groceries and liquor in Table (1.5). Before privatization purchasing groceries and liquor in a given day required traveling 8.5 miles on average, which is the triangular distance between home, a store for groceries, and another store for liquor. After privatization, it takes 6.9 miles to purchase groceries and liquor, mainly because one-stop shopping is only 6.3 miles long on average. In other words, consumers save 35% of the trip distance (27% in median) by switching from shopping at two stores before privatization to one-stop shopping after. The reduced distance of one-stop shopping stems from avoiding traveling between a store for groceries and another store for liquor, which were 3.8 miles apart on average (2.11 in median).

In summary, the above evidence shows that there are two possible avenues through which consumers gain from deregulation: one-stop shopping and reduced distance. To disentangle these two effects, I build a structural demand model and apply the household panel data.

1.3 Model

A market is defined as a week. Each week consumer i considers a shopping trip for groceries, liquor, or both. Let the set of all stores be $S \subset \mathbb{N} \cup \{0\}$. A trip choice is defined as a 2 by 1 vector of store pair $j = (g, \ell)$, which is an element of $T \subset S \times S$, where g indexes a store choice for grocery purchase and ℓ indexes a store choice for liquor purchase. T only includes the subset of possible trips that consumers could make and includes $(\cdot, 0)$ for the choices of *not shopping for liquor* and $(0, \cdot)$ for the choices of *not shopping for groceries*. For example, if store 1 only sells liquor, then $(0, 1) \in T$ and $(1, 0) \notin T$. Before privatization, available choices were no shopping where $j = (0, 0)$; shopping at a grocery store g for groceries only (G) where $j = (g, 0),\ g > 0$; shopping at a liquor-selling store ℓ for liquor only (L) where $j = (0, \ell),\ \ell > 0$; or shopping at two separate stores for groceries and liquor ($GL2$) where $j = (g, \ell),\ g \neq \ell,\ g, \ell > 0$. After privatization, a new choice, shopping at one store for both groceries and liquor ($GL1$) where $j = (g, \ell)$ – $g = \ell,\ g, \ell > 0$ – is added to the choice set.

Consumer i receives the utility u_{ij} from choosing a trip $j = (g, \ell)$. u_{ij} is defined by

$$\begin{aligned} u_{ij} &= \delta_j + \Gamma_j + \mu_{ij} + \varepsilon_{ij} \\ \delta_j &= \lambda \delta_g^G + (1 - \lambda)\, \delta_\ell^L \\ \delta_0^G &= \delta_0^L = 0. \end{aligned} \tag{1.1}$$

The first two terms δ_j and Γ_j are choice specific and do not vary by consumer. δ_j is defined by the convex combination of the quality of groceries δ_g^G at the store chosen for groceries g and the quality of liquor δ_ℓ^L at store ℓ. The weight parameter is λ. The subscripts g and ℓ indicate stores, and the superscripts G and L indicate the product categories – groceries and liquor, respectively. If groceries are purchased at store g, the consumer receives δ_g^G with weight λ. Likewise, if liquor is purchased at store ℓ, then the consumer receives δ_ℓ^L with weight $1 - \lambda$. Grocery-shopping-only trip yields $\delta_j = \lambda \delta_g^G$ and liquor-shopping-only trip yields $\delta_j = (1 - \lambda)\, \delta_\ell^L$. Γ_j contains the fixed effects for the type of choices. The consumer-trip specific term μ_{ij} includes demographics and interaction of price and income. ε_{ij} is assumed to be i.i.d. across consumers and choices and follows Type 1 Extreme Value distribution.

The store-grocery and store-liquor specific utilities, δ^G and δ^L, are defined by

$$\begin{aligned} \lambda \delta_g^G &= -\alpha_0 p_g^G + X_g' \beta + \xi_g^G \\ (1-\lambda)\, \delta_\ell^L &= -\alpha_0 p_\ell^L + X_\ell' \beta + \xi_\ell^L \end{aligned} \tag{1.2}$$

$\lambda \delta_g^G$ consists of price p_g^G, other observed characteristics X_g^G, and unobserved (to econometricians) characteristics ξ_g^L and the same applies to $(1-\lambda)\, \delta_\ell^L$. p_g^G is the price index for groceries at store g and p_ℓ^L is the price index for liquor at store ℓ. When choosing $j = (g, \ell)$, the consumer faces prices $p_j = p_g^G + p_\ell^L$ if both groceries and liquor are purchased, $p_j = p_g^G$ if only groceries are purchased, and $p_j = p_\ell^L$ if only liquor is purchased. The price coefficient α_0 measures the base marginal utility of income. β is a vector of coefficients of the observed store characteristics X_g^G and X_ℓ^L. These characteristics include 13 season-varying quarter fixed effects and the fixed effects for 40 store-chain brands, controlling for heterogeneity in liquor or grocery quality across store-chain brands. Two unobserved qualities to econometricians in my setting – one for groceries ξ_g^G and the other for liquor ξ_ℓ^L – may include advertisements, assortment size, and product selection specific to grocery or liquor sections of the store. If these unobserved qualities are positively correlated with prices, then the estimate of α_0 is biased upward. I ultimately instrument for prices by modifying the standard trick suggested by Berry (1994) and Berry et al. (1995) (BLP) in Section (1.5).

The choice specific constant term Γ_j contains the fixed effects for types of trip choices. Letting G be the left-out choice, Γ_j is defined as

$$\Gamma_j \quad = \quad \gamma_{GL1} \, \{GL1\}_j + \gamma_{GL2} \, \{GL2\}_j + \gamma_L \, \{L\}_j$$

where $\{\cdot\}_j$ is an indicator variable of type of choice j. *Convenience* of one-stop shopping is given by the difference between γ_{GL1} and γ_{GL2}. Households with young children may have different preference towards one-stop shopping; one-stop shopping reduces hassle for parents to get the children in and out of the car multiple times. As an extended model, I allow the fixed effects to interact with an indicator variable of young children under 7, k_i:

$$\Gamma_j \quad = \quad \left(\gamma_{GL1} + \gamma_{GL1}^k k_i\right) \{GL1\}_j + \left(\gamma_{GL2} + \gamma_{GL2}^k k_i\right) \{GL2\}_j + \left(\gamma_L + \gamma_L^k k_i\right) \{L\}_j.$$

If a household has at least one young child, the value of one-stop shopping is $\gamma_{GL2} + \gamma_{GL2}^k - \left(\gamma_{GL1} + \gamma_{GL1}^k\right)$.

μ_{ij} varies across consumers, and it is specified by

$$\mu_{ij} = \sum_{b=2}^{4} \alpha_b p_j \, \{b\}_i - \beta_d d_{ij} + Z_i' \beta_z.$$

Consumers are assumed to have different price sensitivities, varying by their income levels. $\{b\}_i$ is an indicator variable for whether i's income level belongs to the income bin b where $b = 1$ is the base income bin. A consumer in the base income bin has the marginal utility of income α_0 whereas a consumer in a higher income bin has the marginal utility of income $\alpha_0 - \alpha_b$. d_{ij} indicates the total trip distance in miles. If a trip was made to two stores, then d_{ij} measures the distance of the triangular path between home, the grocery store, and the liquor store. Otherwise, it is a round trip distance to the store. β_d measures disutilities from traveling. β_z is the vector of coefficients for demographics Z_i which control for race and education.

This model nests a standard demand model of single store choice; by removing the option of shopping for multiple purposes either at one or more stores, i.e. all γ's equal 0, the model is reduced to a single store choice demand. Specifically, $\lambda = 1$ reduces the model to a single store choice for grocery shopping, whereas $\lambda = 0$ reduces the model to a single store choice for liquor shopping.

Let J^G and J^L be the number of grocery and liquor-selling stores, respectively. Given some λ, define $\boldsymbol{\delta}^*$ is defined as $\left(\lambda\delta_1^G,\ ...,\ \lambda\delta_{JG}^G,\ (1-\lambda)\,\delta_1^L,\ ...,\ (1-\lambda)\,\delta_{JL}^L\right)$, which is a vector of length $J^G + J^L$. θ is defined as $\left((\alpha_b)_{b=2,\ ...,\ 4}, \beta_d, \beta_z, \lambda\right)$, which is comprised of the parameters outside of δ_g^G and δ_ℓ^L. The market share of choice $j = (g, \ell)$ is derived by aggregating the individual probability of choosing j:

$$s_j\left(\boldsymbol{\delta}^*; \theta\right) = \int \frac{exp\left(\delta_j + \Gamma_j + \mu_{ij}\right)}{1 + \sum_{j' \in T \setminus \{(0,0)\}} exp\left(\delta_{j'} + \Gamma_{j'} + \mu_{ij'}\right)} di. \tag{1.3}$$

1.4 Data

In addition to the household level data described in subsection (1.2.2), two store-level datasets are used for demand estimation. From the WSLCB, I obtained quarterly store-level liquor sales revenue from both before and after the deregulation to construct the market shares of liquor, which help controlling the correlation between price and unobserved store quality. For pre-privatization, the dataset also includes retail prices, which are used to create the liquor price index[3] . The liquor price index is defined by the quarterly average quantity-weighted-prices for 1.75 liters of liquor, which is the median quantity purchased per trip per every quarter, observed from the household data. The last column in Table (1.6) shows the average and standard deviation of the liquor price index. Moreover, the state stores' wholesale prices and operational costs, including: wages, benefits, rents, etc., are given by the data, and I use those costs to calculate the retail surplus without estimating the costs as in Miravete et al. (2014). Finally, locations of all grocery and liquor stores before and after, are used to derive shopping trip distance alongside with zip code information from the household panel data.

The Nielsen Retail Scanner dataset provides the price and quantity of every product sold in four major grocery chain stores, two major discount chain stores, and two drug chain stores in Washington. All establishments whose parent company belongs to the Nielsen chain stores are included in the dataset, and I designate these stores as *Nielsen-affiliated stores.* They account for 50% of the liquor sales and at least 48% of the grocery sales in Washington (Lazich and Burton (2014)). Out of 679 Nielsen-affiliated stores, 676 all started selling liquor in June 2012. From this dataset I construct market shares of groceries and grocery price index before and after, as well as liquor price index after. The grocery price index is the sum of the average sales-weighted-prices of the 11 most commonly purchased products per quarter[4] . Table (1.6) summarizes grocery price index before and after.

For the non-Nielsen-affiliated stores, I use the household data to construct the price index and market shares of groceries before and after and price index of liquor after. All panelists are included in the analysis even if some did not purchase liquor during the sample period because they still affect the demand for groceries.

The sample period runs from 2011 to 2013. Since the sales and price data is defined by quarter, I assume that the weekly market shares and price indices remain the same within a quarter. There are about 600 grocery stores observed as visited by the panelists both before and after deregulation whereas the observed liquor stores in the household

[3] Since liquor prices before privatization were set uniformly every month by the WSLCB, technically there was no price variation across stores. However, the difference in product selection across stores generates variation in liquor price index among the state stores.

[4] See Appendix (A) for details.

data is increased from 104 before and 229 after. The average number of trip choices observed in the data per quarter is 797 before and 849 after deregulation. I treat the pair of stores which were never chosen by the panelists as part of outside option to ensure the manageable size of the choice set. The first column of the Table (1.6) shows that grocery-shopping-only type of choices are chosen most frequently. The two-store choice previously had the second largest trip shares but it has shifted to the one-stop shopping.

1.5 Identification and Estimation

1.5.1 Allowing for Endogenous Price

If prices are positively correlated with either of the unobserved store qualities ξ_g^G or ξ_ℓ^L, ignoring this correlation will lead to upwardly biased estimates of the price parameter, making consumers look less price sensitive than they are. For example, qualities of grocery and liquor selections which econometricians do not observe can be positively correlated with prices. In addition, stores' promotions and advertisements can also have positive correlation with prices if stores increase prices due to advertisement cost. On the other hand, it is possible that the promotional activities are negatively correlated with prices if promotions include price discounts, biasing the price coefficient downward.

To address endogenous prices in nonlinear demand, I extend the standard approach from Berry (1994) and Berry et al. (1995) (BLP) to the setting, which allows up to two unobserved qualities – store characteristics of grocery sections and liquor sections – for each store. Extending the BLP's idea to this setting, I control for the store-grocery and store-liquor specific terms – δ_g^G and δ_ℓ^L – so that the correlation between prices and unobserved qualities is buried inside of those terms. I modify the BLP's method to locate those two terms by matching the model's predicted shares of grocery sales and liquor sales to the shares observed in the data through a contraction mapping operator which is tweaked based on the operator used in the BLP papers. The precise measure of stores' market shares for grocery sales and for liquor sales is obtained from the WSLCB and Nielsen dataset.

Let $\boldsymbol{\delta} = \left(\delta_1^G, \ ..., \ \delta_{J^G}^G, \ \delta_1^L, \ ..., \ \delta_{J^L}^L\right)$ be a vector of length $J^G + J^L$, which is the stacked up store qualities of groceries and liquor without weight λ unlike $\boldsymbol{\delta}^*$. Given any

$\lambda \in [0,1]$ and θ, I prove that the following operator f, the modified operator of BLP, has a unique fixed point $\boldsymbol{\delta}(\theta)$ which matches the model-predicted market shares of grocery sales and liquor sales for each store, $\boldsymbol{s}(\boldsymbol{\delta}^*;\theta)$, to the market shares in the data:

$$\begin{aligned} f : R^{J^G+J^L} &\rightarrow R^{J^G+J^L} \\ f(\boldsymbol{\delta}) = \boldsymbol{\delta}^* + log&\left(\boldsymbol{s}^{\text{data}}\right) - log\left(\boldsymbol{s}\left(\boldsymbol{\delta}^*;\theta\right)\right) \end{aligned} \tag{1.4}$$

where s_g^{data} is the observed market shares of groceries across J^G grocery stores and shares of liquor across J^L liquor-selling stores. Appendix (A) provides a proof that f is a contraction mapping with modulus less than 1. The operator f in (1.4) is different from the BLP's operator in two ways. First, while the BLP's operator identifies choice-specific constant, the operator f identifies the store-specific constant for groceries and for liquor. Second, the formulation of f in (1.4) includes the weighting parameter $\lambda \in [0,1]$ to guarantee the existence of the unique fixed point. One part of the BLP's proof requires that the inside share should be less than 1 so that there exists a unique $\boldsymbol{\delta}$ at which the predicted shares match with data. However, in my setup, there are two inside shares, one for groceries and the other for liquor, and the sum of those inside shares is not necessarily less than 1. By formulating $\delta_j = \lambda\delta_g^G + (1-\lambda)\,\delta_\ell^L$, the weighting parameter λ guarantees that f is a contraction and has a unique fixed point that matches predicted market shares with data.

Once parameters θ are estimated while holding fixed store-specific constant for groceries and liquor, the resulting $\boldsymbol{\delta}^*(\theta)$, which is linear in price and unobserved store qualities, is then regressed on its arguments with instrumented price to control for the potential endogeneity.

1.5.2 Estimation

Let the market be indexed by t. The first step parameters θ are estimated by maximizing the log likelihood function,

$$\max_{\theta}\ log\ L(\theta) = \sum_t \sum_j \sum_i\ w_{it}\ Y_{ijt}\ log\ s_{ijt}\left(\boldsymbol{\delta}_t^*(\theta), \theta\right) \tag{1.5}$$

where w_{it} is the weight on each household i within a market t and $Y_{ijt} = 1$ if i chose j in week t and 0 otherwise. At each iteration of nonlinear parameter search, $\boldsymbol{\delta}_t^*(\theta)$ is identified by using the contraction mapping in (1.4).

Given the estimated $\hat{\theta}$ and $\boldsymbol{\delta}^*\left(\hat{\theta}\right)$ from the first step, marginal utility of income, α_0, is estimated in the second step by using the system of equations in equation (1.2):

$$\begin{aligned} \boldsymbol{\delta}^{*G} &= -\boldsymbol{p}^G\alpha_0 + \boldsymbol{X}^G\beta + \boldsymbol{\xi}^G \\ \boldsymbol{\delta}^{*L} &= -\boldsymbol{p}^L\alpha_0 + \boldsymbol{X}^L\beta + \boldsymbol{\xi}^L \end{aligned} \tag{1.6}$$

where the dependent variable $\boldsymbol{\delta}^{*G}$ is the first J^G elements of $\boldsymbol{\delta}^*$, which is a stacked vector of $\lambda\delta_g^G$. $\boldsymbol{\delta}^{*L}$ is the rest of the elements of $\boldsymbol{\delta}^*$, which is a stacked vector of $(1-\lambda)\,\delta_\ell^L$. $\boldsymbol{p}^G$, $\boldsymbol{X}^G$, and $\boldsymbol{\xi}^G$, respectively, are stacked vectors of p_g^G, X_g^G, and ξ_g^G. The same is applied to $\boldsymbol{p}^L$, $\boldsymbol{X}^L$, and $\boldsymbol{\xi}^L$. The effects of store-chain brands and seasons with respect to grocery products are allowed to be different from those with respect to liquor products. I use the average price of other stores outside of a 5-mile radius of a store as an instrument, similar to Hausman (1996)'s. The price coefficient is identified by assuming that the statewide supply side cost, such as wholesale price, distribution cost, or wage, is correlated with price but it is uncorrelated with the local store quality or promotions, conditional on store-chain brand and seasons in $\boldsymbol{X}^G$ and $\boldsymbol{X}^L$.

Standard errors are derived by adjusting the overall number of observations. Weights between the first step moment condition (the score of MLE) and the second step moment from (1.6) are determined according to Arellano and Meghir (1992) since the observations for the likelihood function are from the distribution of households while those for the linear equation are from the distribution of grocery and liquor stores.

1.6 Results

1.6.1 Demand Results

Table (1.7) shows the nonlinear demand estimates of the first step. The estimates represent the marginal effects on utility, not on purchasing probability. The first column is the model specification where the fixed effects for type of choices are not interacted with the indicator variable for whether there is at least one young child whereas second column allows the interaction. The fixed effects for type of choices in each specification reveal that the most preferred shopping is grocery-shopping-only (G), followed by one-stop shopping ($GL1$) and grocery and liquor shopping at two stores ($GL2$). This order of preference lines up with the order of trip shares in Table (1.6). It is reasonable that

G type is most preferred since two third of the consumers do not buy liquor at all during the sample period. Conditional on purchasing liquor, one-stop shopping is most preferred. That is, consumers value the *convenience* of one-stop shopping. The average difference in utility from two-stop and one-stop shopping for households with children (0.8) is twice larger than that for other households with no young children (0.4). It implies that switching from two-stop to one-stop shopping increases utility more when shopping with young children than without children.

Consumers also value their time and dislike traveling longer distances. If a trip's total distance increases 1%, the probability of choosing that trip decreases 1.7%. The estimated weight between quality of groceries and liquor, λ, still lies between 0 and 1 without imposing the restriction. Consumers with income of $35,000 or more are slightly less elastic to price than those with less than $35,000 of income. This result is consistent with findings of previous works. For example, Goolsbee and Petrin (2004) showed that price elasticities decreases as household income increases. Race and education are also important factors in determining shopping probability.

Table (1.8) contains the results of the second stop estimation on α_0. The first two columns are OLS results and the last two columns are 2SLS results. The price coefficients are almost identical whether trip type fixed effects are interacted with young children indicator. The estimated price coefficient without instrumenting price is overestimated (biased upward) by six times compared to the instrumented price coefficient. This implies that, after controlling for the store chain and seasonality, there is still correlation between price and unobserved store qualities of groceries and liquor. Based on the instrumented price coefficient with the base specification, -0.23, the mean elasticity with respect to the liquor basket price at own store is about -8.34. This is consistent with the literature on store choice model. For example, Smith (2004) found that store level own price elasticities for supermarkets lie between -7 and -9. Furthermore, by translating the nonlinear parameters into dollar values from the marginal utility of income α_0, the estimated travel cost per mile is about $1.76. This is the consumer's willingness to pay to reduce a given trip's distance by one mile[5] . The estimated travel cost is also consistent with the literature to what others have found. For example, Seim and Waldfogel (2013) estimated travel costs of $1.01 per mile.

[5] See Chapter 3 in Train (2009).

Switching from $GL2$ type of shopping before privatization to one-stop shopping after can benefit consumers through two channels. First, *convenience* of $GL1$ compared to $GL2$ translates to \$1.65: a consumer must be compensated by \$1.65 to be indifferent between $GL1$ and $GL2$ types of trips with exactly the same price, distance, store quality. This is equivalent to 5% of the average expenditure on liquor per shopping trip. Second, the gains from reduced distance (1.57 miles) realized by switching from $GL2$ to $GL1$ is \$2.77, holding all other trip characteristics constant. Combining the gains from these two channels, consumers gain \$4.42 when switching from a $GL2$ trip before privatization to $GL1$ after, holding all other characteristics constant, and *convenience* accounts for 37% of such gains. This exercise assumes the choice is fixed.

1.6.2 The Benefit of One-Stop Shopping to Consumers

To measure the value added by one-stop shopping to consumers who shop for liquor, I select households which purchased liquor at least once in the given year and normalize utility components from grocery shopping to be 0. The value of one-stop shopping, *convenience*, is derived by simulating consumer surplus with and without the one-stop shopping option in the shopping choice set, holding all other variables constant: distance, price, a total number of liquor stores, store brand, and store quality. In other words, I disentangle the gains of one-stop shopping from other potential gains from increased number of liquor-selling stores and from potential loss from increased price. This exercise is different from subsection (1.6.1) because consumers can switch to other choices. I use the estimates of the base model without children interaction but the results are almost identical if the extend model estimates are used.

First of all, the per-trip value of *convenience* by one-stop shopping is \$2.52, which is about 8% of the expenditure on liquor per trip. Column "One-Stop Shopping" in Table (1.9) summarizes the compensating variation for banning one-stop shopping; the monetary compensation given to consumers such that they are indifferent between the realized economic environment after privatization and the counterfactual environment where grocery stores are not allowed to sell liquor. The counterfactual experiment assumes that the liquor sections inside of grocery stores are relocated right next to those stores, holding fixed prices and qualities. On average, a household is willing to pay \$2.52 per liquor shopping trip to have an option of one-stop shopping or \$4.46 annually. On average, households with young children value the one-stop shopping convenience by \$5.23 while those without young children value the convenience by \$4.58. By allowing one-stop shopping, consumers gain aggregate 4.5 million dollars in a year. This gain is a 31% increase of consumer surplus.

Second, deregulation expands the choice set of shopping trips through two margins: the introduction of one-stop shopping as a new type of trip and reduced travel distance due to the increased number of liquor selling outlets. The following exercise derives the benefits from increased number of stores. I first simulate the consumer surplus where the total number of stores is capped at the same as before privatization to estimate the value generated by the influx of stores, possibly through a reduction in trip distance. The counterfactual assumes that there are only 330 stores, which are the top 330 stores by their market shares of liquor sales in 2013, and their store characteristics remain fixed. Column "Reduced Distance" in Table (1.9) shows that a household's valuation of the increased number of stores is \$3.36 per liquor shopping trip or \$5.85 per year on average, which is equivalent to 45% increase of consumer surplus. Thus, the value of *convenience* of one-stop shopping is about two-thirds of the value of the reduced distance, which is consistent with the simple exercise in subsection (1.6.1).

Third, I evaluate the gross gains of expanded choice set for shopping trips holding price constant. Holding price constant, the gross gains from deregulation is derived by the compensating variation between the economic environment after privatization and the counterfactual environment where one-stop shopping is not allowed and the total number of stores is limited to 330, holding all other variables fixed. This compensating variation is reported in column "Combined Gains" and reveals how much consumers gain from the expansion of the choice set. Holding price constant, deregulation benefit a household \$5.32 per liquor shopping trip or \$9.35 annually on average. The value of *convenience* accounts for approximately 37 to 47% of this benefit from deregulation.

Fourth, I estimate the costs from increased price. The compensating variation between the benchmark economy setup before privatization and the world where prices are increased as 11% without any benefit of improved choice set is \$1.56 per household per liquor shopping trip or \$2.86 per year. The increase in price reduces the consumer surplus by 21%.

Finally, the overall gain from deregulation is \$3.86 per household per liquor shopping trip or \$6.70 per year. This gain is equivalent to a 56% increase of the consumer surplus compared to 2011. This is a net gain from the the combined effects – the increased price and expansion of the choice set, which is led by one-stop shopping and the increased number of stores. The value added by one-stop shopping is equivalent to two-thirds of the net gain from deregulation.

1.6.3 The Benefit to Stores by Expanding the Scope

The consumer benefit of one-stop shopping has implications to stores' sales revenue. If a store sells both groceries and liquor, it attracts more consumers and more frequent visits to the store, and in turn, raises the store revenue. In this subsection, I compare grocery sales in 2013 to those in the counterfactual world where grocery stores are not allowed to sell liquor, holding all other factors constant. It is the same experiment as shutting down all liquor sections inside of grocery stores and placing those liquor sections adjacent to the grocery stores without sharing the same roof. Selling liquor inside of grocery stores increases grocery sales by 4.5% on average. On the other hand, stand-alone grocery stores lose 0.5% of its grocery sales on average. Standard deviations are reported in Table (1.10). Overall grocery sales remain almost unchanged, implying that the sales increase in grocery stores selling both product categories is mainly due to business stealing from those not selling liquor. This result is sensible because most grocery sales arise from "groceries only" type of shopping, which has no connection with liquor shopping, and therefore, allowing liquor sales in grocery stores or not has little impact on aggregate grocery sales.

Table (1.11) describes the change in liquor sales when stores are allowed to accommodate both groceries and liquor. Liquor sales are increased by 30% on average if liquor sections are located inside of a grocery store, rather than being located outside. Approximately one-third of the increase is due to the business stealing effect from other stand-alone liquor stores, which lose 0.81% of their liquor sales. The aggregate liquor sales are increased by 20%. This increase in liquor quantity sold in stores does not imply that total liquor consumption is increased; consumers could have substituted liquor purchases at bars or restaurants for the on-premises consumption, which is considered as an outside option in this paper's setup, with purchases at stores for the off-premises consumption. See Appendix (A) for more detailed evidence.

This result suggests that there is a strong incentive for a store to expand their scope of product offerings because it raises sales revenue and it also prevents stores from losing sales revenue. If the incentive exists for other combinations of product categories, such as general merchandise and groceries, some retailers with resources could potentially grow larger and the others, especially specialty shops, may exit the market, leading to a concentrated market structure similar to the current retail industry.

1.6.4 Distribution of Consumer Gains from Deregulation

Even though a household benefits \$3.86 per trip or \$6.70 per year from deregulation on average, it is not Pareto improving; there are some consumers who lose after deregulation. Moreover, among the consumers who are better off, the gains are not distributed uniformly across demographics and density of the area. Table (1.12) summaries distribution of the annual consumer gains. The zip code areas where per square mile density is less than 83 people lose the most: the consumer surplus decreases by 63 cents, which is a 34% decrease on average. This is because reduced travel distance in rural areas is not big enough to compensate for the increase in price in those areas. In addition, consumers in areas where the median income is less than \$35,000 lose 1.14% of the consumer surplus. The rest of the demographical groups benefit from deregulation but at a different rate. Gains increase with population density, income, and non-Caucasian population. The distribution of gains is also consistent with data observations on the distribution of the number of liquor stores, price increase, and decrease in distance in the Appendix (A). Moreover, the percentage gain for consumers with young children is twice as large as that for consumers without children. This result is consistent with the fact that $\gamma_{GL2}-\gamma_{GL1}$ for consumers with children is twice higher than that for consumers without children. On the other hand, the level of gain for consumers with children is lower than the others because households with children generally dislike liquor shopping more than those without children.

One-stop shopping and the increased number of liquor stores also have different impacts across consumers. Table (1.13) shows the percentage change in consumer surplus by demographics under different counterfactual experiments. Column "One-Stop Shopping" shows that the gains from one-stop shopping are relatively uniformly distribution across area density, income, and race. In contrast, the effect of more stores in column

"Reduced Distance" is disproportionate to different demographic groups.

1.6.5 Total Welfare Change From Deregulation

After privatization the total welfare in liquor industry increases by 34 million dollars or $34 per household (including non-liquor purchasers) per year, a 15% increase. The welfare change between 2011 and 2013 is presented in Table (1.14). Among all economic players, the government makes the most surplus both before and after privatization, and its gains are the largest. The state makes 58% more revenues due to increases in the tax rate and quantity sold. The annual consumer surplus increases by $6.70, which is a 55% increase. The retailer surplus in 2013 is derived by change in profits, assuming that the marginal cost remains unchanged since 2011[6] . Negative retailer surplus in 2013 suggests that the marginal cost or markup must have decreased, since retailers have the option to exit the market rather than make negative profits. Therefore, the estimated total welfare can be interpreted as the lower bound. If the retailer surplus were zero, then gains from deregulation almost double to a 28% increase in total welfare.

1.7 Conclusion

While the impact of expanding the firm scope has been studied mainly on the cost side, little is known on the consumer side. Deregulation in the liquor market in Washington provides an environment where grocery stores are allowed to expand the scope of product categories to liquor. This paper exploits the change in retailer scope to evaluate the value of one-stop shopping as a consumer side mechanism of economies of scope. By using the sets of household and store level data before and after deregulation, I trace consumer's switching behavior on shopping trips before and after to separately identify the value of one-stop shopping from the other effects of deregulation, such as reduced distance, a number of stores, and price change. The value of one-stop shopping per trip per household is estimated to be $2.52, which is 8% of the liquor expenditure. I find that the complementarities between grocery and liquor products are significant: the benefits from one-stop shopping account for 47% of the gross gains from deregulation.

[6] The retail surplus does not include the fixed cost of entering into the retail liquor industry. Since most new liquor-selling stores are existing grocery stores, the fixed cost of adding a few aisles is expected to be small.

Moreover, if liquor is sold inside of a grocery store instead of outside, it raises grocery sales by 5% and the liquor sales are increased by 30%.

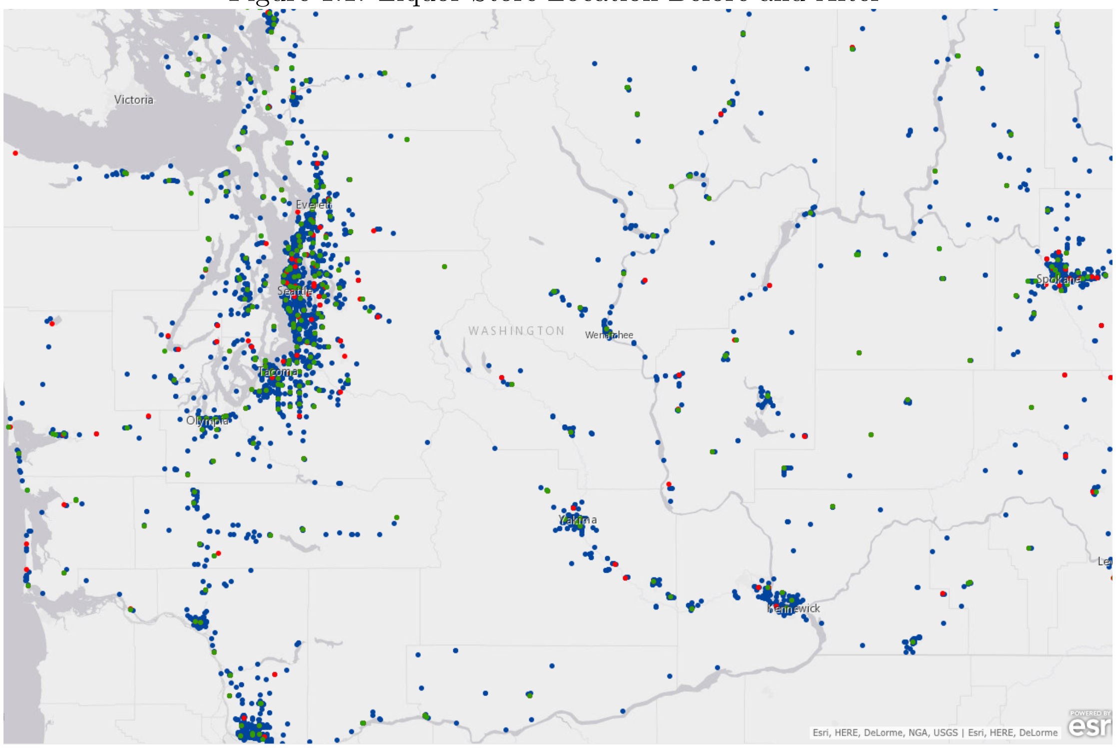
The results suggest that retailers have incentives to add more categories of products into their stores if there are enough complementarities between categories. This finding opens up the possibility that the recent trend in big box stores or supercenters may have been motivated to raise revenue by offering one-stop shopping convenience. Moreover, the implication that some retailers accommodate a wide range of assortment while the specialty shops exit the market is consistent with the concentrated market structure of both on-line and off-line retail industries.

The framework in this paper can be applied to study the consumer side motivation for expanding firm scope. It identifies complementarities between combinations of products and sheds light on firm scope decision affected by consumer side economies of scope. Moreover, the framework provides an identification strategy to allow for correlation between price and multiple unobserved store qualities. This framework is useful even when the store level data is limited to the marginal market share of each product category rather than the joint share of each combination of product categories. The applicable examples of this framework include measuring the consumer benefits from shopping at Wal-Mart, department stores, or a mall. Moreover, this framework can be used to find the optimal combination of product categories which generates the largest economies of scope when sold by the same firm.

The application is not limited to the retail industry: it can be applied to examine the economies of scope for occupational licensing or supply chain. Some regulations limit what an occupation can do. For example, in some states physicians are not allowed to dispense drugs. The supply chain of alcoholic beverage in the U.S. are regulated such that a firm can only operate in one of the three tiers; supplier, distributor, or retailer. Studying the economies of scope which arise from complementarities between occupations or firms can shed light on complementarities between different tasks and the costs of the regulations which limit the firm scope.

1.8 Tables and Figures

Figure 1.1: Liquor Store Location Before and After



Notes: The longitude and latitude was derived by using ArcGIS API from the business address given by the WSLCB dataset.

- Red: former state stores, which no longer exist, as of January 2012.
- Green: former state stores, which still exist, as of February 2015.
- Blue: new stores after privatization.

Table 1.1: Number of Liquor Stores

Year	Total	Grocery[a]	Stand-alone[b]	
			Former State	New
Jan.-May 2012	330	0	-	-
June-Dec. 2012	1373	1075	276	22
2014	1398	1125	235	38

[a] Grocery stores selling liquor

[b] Stand-alone liquor stores

Data source: the Washington State Liquor and Cannabis Board

Table 1.2: Change in Price and Quantity Sold

	Before[a]	After	% Change
Post-tax price ($)	20.71	23.03	11.20
Pre-tax price ($)	15.02	14.61	-2.73
Quantity sold (million liter)	29.65	31.42	5.97

Note: Price is defined as the average price per liter weighted by liter sold, deflated in January 2006 dollars.

[a] "Before" refers to a year prior to privatization and "After" refers to a year after.

Data source: Spirits Tax Collections and Sales data from the Washington State Department of Revenue

Table 1.3: Share of Trips by Type Conditional on Purchasing Liquor

Type	Trip Shares		
	Before	After	% Change
GL1[a]	N/A	0.7457	
GL2[b]	0.5005	0.1009	-79.85
L only[c]	0.4995	0.1534	-69.29
N[d]	2751	5294	92.44
Panelists[e]	423	625	-0.60

Note: Deregulation occurred in June 2012. “Before” refers to from January 2011 to May 2012 and “After” is from June 2012 to December 2013.

[a] GL1: shopping for groceries and liquor at one store

[b] GL2: Shopping for groceries and liquor at two different store

[c] L only: Shopping for liquor only

[d] The number of observed shopping trips with liquor purchase

[e] The number of panelists who have purchased liquor at least once during the period.

Data source: the Nielsen Consumer Panel

Table 1.4: Distance the Closest Liquor Store

	Upper Bound	Before	After	% Change
Mean		1.13	0.94	-16.99
Median		1.71	1.38	-19.69
Zip density[a]	13	3.23	2.98	-7.59
	116	2.18	1.84	-15.49
	1185	1.71	1.32	-22.76
	Above	1.14	0.84	-26.45

Notes: Distance between a zip code and a store is derived by averaging distance between each census block within the zip code and the store, weighted by the population of the block.

[a] The zip code density is population divided by square miles. The results do not change much if population is only 21+.

Table 1.5: Trip Distance Conditional on Shopping for Grocery and Liquor

	Before	After	% Change
Mean	8.51	6.94	-18.45
Median	5.06	3.49	-31.03
Mean GL1[a]		6.34	
Median GL1		3.49	

Notes: Distance of total shopping trip in miles. Deregulation occurred in June 2012. "Before" refers to from January 2011 to May 2012 and "After" is from June 2012 to December 2013.

[a] Round trip distance of one-stop shopping for groceries and liquor

Table 1.6: Trip Shares and Price per Trip Type

Type	Trip Shares		Grocery Price[a]		Liquor Price[b]	
	Before	After	Before	After	Before	After
GL1		0.0545		62.30		25.10
		(0.0036)		(30.03)		(13.79)
GL2	0.0236	0.0090	70.53	65.78	32.21	34.41
	(0.0041)	(0.0016)	(30.18)	(26.61)	(18.76)	(23.75)
L only	0.0179	0.0081			31.86	32.80
	(0.0038)	(0.0012)			(15.52)	(22.06)
G only	0.9584	0.9283	61.91	60.43		
	(0.0073)	(0.0061)	(29.29)	(28.87)		

Notes: Mean values and standard deviation in the parenthesis. Deregulation occurred in June 2012. “Before” refers to from January 2011 to May 2012 and “After” is from June 2012 to December 2013.

[a] Price index for groceries, i.e. sum of sales weighted prices of the 11 most commonly purchased products. See Appendix (A).

[b] Price index for liquor, i.e. liter-sold weighted price per 1.75 liters.

Table 1.7: Nonlinear Demand Estimates

	No Int.	Kids
GL1	-4.3239	-4.3634
	(0.6827)	(0.6606)
GL2	-4.7067	-4.7655
	(0.3264)	(0.5917)
L	-6.3008	-6.2727
	(2.5297)	(2.2314)
Kids*GL1[a]		-0.8706
		(0.0622)
Kids*GL2		-1.2685
		(0.0902)
Kids*L		-1.2204
		(0.0871)
Distance	-0.4094	-0.4097
	(0.0315)	(0.0265)
λ[b]	0.3596	0.3601
	(0.0244)	(0.0247)
p*Income1[c]	0.0005	0.0005
	(0.0000)	(0.0000)
p*Income2	0.0003	0.0003
	(0.0000)	(0.0000)
p*Income3	0.0003	0.0004
	(0.0000)	(0.0000)
Education[d]	-0.4035	-0.3935
	(0.0268)	(0.0261)
Race[e]	-0.4774	-0.4877
	(0.0346)	(0.0343)
Dist. Elas.	-1.874	-1.2037

[a] Kids: indicator variable if there is at least one child under 7

[b] Weight on unobserved characteristics for grocery (λ) is derived without restriction of $0 \leq \lambda \leq 1$.

[c] Baseline income: less than \$35,000. Income1: between \$35,000 and \$59,999. Income2: between \$60,000 and \$99,999. Income3: above \$100,000.

[d] 1 if holding at least college degree, 0 otherwise

[e] The proportion of Caucasians compare to other races

Table 1.8: Linear Demand Estimates

	OLS		2SLS[a]	
	No Int.[b]	Kids	No Int.	Kids
Intercept	0.8212	0.8653	9.567	9.7796
	(0.0718)	(0.0725)	(0.2696)	(0.274)
Price	-0.0452	-0.0463	-0.2321	-0.2371
	(0.0014)	(0.0014)	(0.0059)	(0.006)
Chain FX	Y	Y	Y	Y
Season FX	Y	Y	Y	Y
Mean Elas.[c]	-1.9216	-1.9803	-8.3366	-8.5699
Entire Elas.[d]	-0.4796	0.1734	-6.5529	-6.1869
Travel Cost[e]	7.62	7.44	1.76	1.73
Convenience of OSS[f]	7.13	7.31	1.65	1.70
Reduced distance[g]	11.96	11.69	2.77	2.71
Total $GL2 \rightarrow GL1$[h]	19.09	19.00	4.42	4.41

Dependent variable: $\boldsymbol{\delta}^*$

[a] Instrument: average price at other stores outside of the 5-mile radius of the store

[b] Fixed effects for type of choices are not interacted with kids.

[c] Elasticities with respect to the entire liquor price of own store

[d] Elasticities with respect to the liquor price in the entire state

[e] Dollar value of traveling one mile

[f] Dollar value of switching from $GL2$ before to $GL1$ after, conditional on distance, price, and other characteristics

[g] Dollar value of reduced distance (1.57 miles) by switching from $GL2$ before to $GL1$ after

[h] Total effects of switching from $GL2$ before to $GL1$ after, conditional on price and other characteristics

Table 1.9: Gains from One-Stop Shopping and Reduced Distance

2013	One-Stop Shopping[a]		Reduced Distance[b]		Combined Gains[c]	
	CV	% Change	CV	% Change	CV	% Change
Total	4,503	30.89	5,904	44.80	9,438	97.85
Per hhld[d]	4.46	30.89	5.85	44.80	9.35	97.85
Per trip	2.52	30.43	3.36	45.56	5.32	98.03

Notes: All values are in \$1,000 except "Per" and "% Change" which are in \$. The choice set for shopping trips is improved by (1) having the one-stop shopping option and (2) reduced distance. The benchmark economy is 2013 when one-stop shopping is allowed and entry is not limited.

[a] Grocery stores are not allowed to sell liquor under the same roof, holding prices, locations of stores, and the total number of liquor-selling stores fixed. The compensating variation is the consumer's valuation of *convenience* from one-stop shopping.

[b] The total number of liquor-selling stores is limited to 330 – the same as before privatization – so that the trip distance of liquor shopping is approximately the same as before, holding all other variables constant. One-stop shopping is allowed. The compensating variation is the value of reduced shopping distance.

[c] One-stop shopping is banned and the number of liquor-selling stores is limited to 330. The compensating variation is the total gains from the improved choice set by deregulation, holding price constant.

[d] Average annual value per household who has purchased liquor at least once in 2013. There are 1,009 thousand households who purchased liquor at least once in 2013.

Table 1.10: Average Percentage Change in Grocery Sales

	Stand-Alone Grocery Store	Store Selling Both
Per store	-0.50	4.51
	(0.01)	(0.28)
Aggregate Change	0.74	

Notes: Sales weighted average percentage change in grocery sales across grocery stores. Numbers in parentheses are standard deviations.

Grocery sales in 2013 are compared to those in the counterfactual world where grocery stores are not allow to sell liquor, holding all other factors constant. It is the same experiment as shutting down all liquor sections inside of grocery stores and placing those liquor sections adjacent to the grocery stores without sharing the same roof.

Table 1.11: Average Percentage Change in Liquor Sales

	Stand-Alone Liquor Store	Store Selling Both
Per Store	-0.81	29.27
	(0.03)	(0.23)
Aggregate Change	20.10	

Notes: Sales weighted average percentage change in liquor sales across liquor-selling stores. Numbers in parentheses are standard deviations.
Liquor sales in 2013 are compared to those in the counterfactual world where grocery stores are not allow to sell liquor, holding all other factors constant. It is the same experiment as shutting down all liquor sections inside of grocery stores and placing those liquor sections adjacent to the grocery stores without sharing the same roof.

Table 1.12: Distribution of Consumer Gains from Deregulation

	Upper Bound	CV	% Change
Zip density[a]	83	-0.63	-34.44
	508	1.21	16.29
	3,222	6.39	67.36
	Above	10.93	60.41
Income[b]	34,999	-0.16	-1.14
	59,999	5.37	44.33
	99,999	7.07	71.35
	Above	6.20	58.16
Race[c]	0.74	9.67	68.22
	0.86	6.38	55.92
	0.91	3.47	34.03
	Above	0.23	4.65
Kids[d]	Yes	6.02	108.36
	No	6.37	51.32

Notes: Annual consumer gains from deregulation per household (with 11% price increase). The overall consumer gain is $6.70 on average.

[a] The zip code density is population divided by square miles. The results do not change much if population is only 21+.

[b] Median income in zip code tabulated area.

[c] Race is the proportion of Caucasians compared to other races.

[d] Whether the household has any children under age 7. The extended model which allow interaction between the type fixed effects and children dummy variable. Using these estimates, the overall annual consumer gain from deregulation is $5.98.

Table 1.13: Counterfactual Consumer Gains Across Demographics

	Upper Bound	One-Stop Shopping[a] CV	One-Stop Shopping[a] % Change	Reduced Distance[b] CV	Reduced Distance[b] % Change	Combined Gains[c] CV	Combined Gains[c] % Change
Zip density	83	0.33	37.49	0.51	73.36	0.71	146.19
	508	2.12	32.51	2.82	48.19	4.32	99.59
	3,222	4.02	33.91	3.83	31.79	7.17	82.50
	Above	6.66	29.79	9.55	49.02	14.90	105.37
Income	34,999	2.99	27.92	5.02	57.95	7.06	106.70
	59,999	4.16	31.20	5.13	41.44	8.45	93.40
	99,999	4.13	32.14	5.07	42.59	8.41	98.39
	Above	3.37	25.01	9.11	117.68	11.25	200.71
Race	0.74	5.36	29.04	8.01	50.58	12.38	108.03
	0.86	4.61	35.00	4.73	36.21	8.38	89.20
	0.91	3.32	32.04	3.97	40.99	6.54	91.93
	Above	1.21	29.92	1.39	36.08	2.30	78.05
Kids	Yes	5.23	82.33	3.42	41.96	4.28	58.66
	No	4.58	32.28	5.81	44.86	2.54	15.63

Notes: Annual net gains per household.

[a] Grocery stores are not allowed to sell liquor under the same roof, holding prices, locations of stores, and the total number of liquor-selling stores fixed. The compensating variation is the consumer's valuation of *convenience* from one-stop shopping.

[b] The total number of liquor-selling stores is limited to 330 – the same as before privatization – so that the trip distance of liquor shopping is approximately the same as before, holding all other variables constant. One-stop shopping is allowed. The compensating variation is the value of reduced shopping distance.

[c] One-stop shopping is banned and the number of liquor-selling stores is limited to 330. The compensating variation is the total gains from the improved choice set by deregulation, holding price constant.

Table 1.14: Total Welfare Change From Privatization

	2011	2013	Change	% Change
Consumer Surplus[a]	12,323	19,083	6,760	54.86
Retailers Surplus[b]	43,253	-30,255	-73,508	-169.95
Government Revenue	175,025	276,201	101,175	57.81
Total	230,601	265,029	34,427	14.93
CS per hhld[c]	12.21	18.91	6.70	54.86
RS per store	131	-22	-152	-116.56
GR per hhld[d]	58.04	91.59	33.55	57.81
Per hhld	76.48	87.90	11.42	14.93

Notes: All values are per year in $1,000 except "Per hhld" and "% Change" which are in $.

[a] Estimated by 1,009 thousand households who had purchased liquor at least once in 2013. The rest of the "Per hhld" measures are based on 3,016 thousand population households.

[b] Assuming that the marginal cost of operating a liquor store remains unchanged over time.

[c] Per household who has purchased liquor at least once a year: 1,009 thousand households.

[d] Per any household: 3,015 thousand households.

Table 1.15: Welfare Change Under Counterfactual Experiments

	One-Stop Shopping[a]		Reduced Distance[b]		Combined Gains[c]	
	Change	% Change	Change	% Change	Change	% Change
Consumer Surplus	4,503	30.89	5,904	44.80	9,438	97.85
Retailer Surplus	-8,267	-26.23	-3,689	-11.70	-9,749	-30.93
Government Revenue	57,988	19.85	116,985	40.05	162,671	55.68
Total	54,225	76.97	119,201	96.55	162,360	184.46
CS per hhld	4.46	30.89	5.85	44.80	9.35	97.85
PS per store	-6	-26.23	-3	-11.70	-7	-30.93
GR per hhld	19.23	19.85	38.79	40.05	53.94	55.68
Per hhld	17.98	19.39	39.53	42.62	53.84	58.05

Notes: All values are in $1,000 except “Per hhld” and “% Change” which are in $.

[a] Grocery stores are not allowed to sell liquor under the same roof, holding prices, locations of stores, and the total number of liquor-selling stores fixed. The compensating variation is the consumer’s valuation of *convenience* from one-stop shopping.

[b] The total number of liquor-selling stores is limited to 330 – the same as before privatization – so that the trip distance of liquor shopping is approximately the same as before, holding all other variables constant. One-stop shopping is allowed. The compensating variation is the value of reduced shopping distance.

[c] One-stop shopping is banned and the number of liquor-selling stores is limited to 330. The compensating variation is the total gains from the improved choice set by deregulation, holding price constant.

Chapter 2

Cash or Charge: Assessing Charging Station Build Out and Incentive Programs' Roles in Electric Vehicles Adoption

2.1 Introduction

In the wake of mounting attention toward investment in energy efficiency and renewable resources, the US allocated approximately $400 million of the 2009 American Recovery and Reinvestment Act toward investment and research into electric vehicles. Of this allocation $115 million was specifically designated for the deployment and study of a first-stage charging station network for plug-in electric vehicles (PEVs) across several states. The EV Project and ChargePoint America, the beneficiary projects of this grant, deployed nearly 5000 charging stations in California homes and public locations from May 2010 to the end of December 2013 starting from the approximate 480 in place at the beginning of the period.[1] The federal and California state governments concurrently rolled out other incentive programs to encourage the adoption of electric cars, including hefty rebates, tax credits, and electricity discounts.

[1] See Table B.2 in Appendix B.

We study these developments with a focus on understanding the role these programs actually played in spurring plug-in electric vehicle (PEV) adoption and whether alternative policy choices could have resulted in more favorable adoption outcomes. Our work fits with (1) a growing body of policy research around encouraging PEV purchases and developing a suitable infrastructure for their use to reduce carbon emissions and with economic literature on indirect network effects.

There are numerous policy studies on the rebate programs and deployment of charging network. However, they do not specified how much of the PEV demand is led by the government support. This paper quantifies the portion of PEV market growth attributable to the government support. Furthermore, we suggest the reduced amount of carbon emissions, resulted from substituting gas vehicles with PEVs under the current policy. We also simulate the reduction in carbon emissions under different policies and evaluate the effectiveness of each policy. The results shed light on optimal government support for expansion of PEV sales and reduction in carbon emissions in turn.

Consumers' price sensitivity determines whether the monetary compensation policies are effective in encouraging PEV purchase. We estimate demand for automobile and evaluate the increase in PEV shares due to the rebates. If consumers are highly sensitive to price, then the rebate programs do provide significant incentives to the consumers for purchasing PEVs.

The effectiveness of deploying public charging stations, such as EV Project and ChargePoint America did, depends on the indirect network effect between PEV and charging station markets. Indirect network effects exist when provision of secondary market (charging station market) has spillovers to demand in a primary market (PEV market). In other words, prevalence of charging stations increases the value of PEVs, leading to high demand in PEVs. The paper looks for the presence of and measure the magnitude of PEV charging network's impact on demand for PEVs. If the indirect effect is significant, then the government support on charging station network is effective in spurring PEV market growth

Using a new data set featuring household-level electric vehicle purchase data with detailed geographic and timing information and supplemented with behavioral data from recent studies on electric car adopters, we estimate a rich discrete choice model of demand for the automobile market in California. The data set permits identification of consumer heterogeneity heretofore impossible in the existing literature. The current version of our paper describes the model and estimation strategy using this dataset.

The paper is organized as follows. In Section 2.2 we describe specific obstacles for widespread adoption of PEVs and features of consumers and incentives in California that we deemed important to model. Section 2.3 places our work into the context of the larger indirect network effect literature. Section 2.4 presents the model of consumer demand and Section 2.5 the data used to estimate the model. Section 2.6 presents the moments we use in a GMM estimation procedure to derive the parameters of the demand model. We outline how features in our data sets allow the estimation procedure to identify the parameters of demand in Section 2.7.

2.2 The Market for Electric Vehicles in CA

Many peculiarities of both PEVs and California differentiate its market for electric and gasoline-fueled vehicles and make the region ripe for a focus of this study. Here we discuss various obstacles inhibiting the mainstream adoption of electric vehicles, argue that *a priori* the charging station network may have an appreciable impact on demand, and describe particulars of the California market and PEV purchasers.

PEVs can be classified into two categories: one category is full electric, such as Nissan LEAF and Tesla Model S, and the other is plug-in hybrid electric vehicle (PHEV), such as Chevy Volt and Toyota Prius Plug-In Hybrid. PHEV is different from full electric vehicles because it has an option to fuel the car with gasoline and its range with electric battery only is shorter than full electric vehicles.[2] We use PEV to include both fully electric and plug-in hybrid electric vehicles.

2.2.1 Charging Electric Vehicles

Before delving into obstacles facing the electric vehicles, it is critical to understand how PEVs "fuel" differently than traditional gas vehicles. Fully electric vehicles and plug-in hybrids recharge their electric engine by, surprisingly, plugging into specially designated charging docks. Most of the charging outlets are compatible to any PEV models. The main difference is the charging speed. Technically any standard electric plug (Level 1

[2] These details will be elaborated upon later. See Table 2.2

charger) can be used to charge a PEV, but current models, even with relatively small batteries, take anywhere from 20 hours to 10 hours to fully charge.[3] These charging times are clearly unreasonable long for a consumer that plans to use their vehicle in daily usage.

The current PEV-specific charging standard is Level 2 chargers, which can fully charge the Nissan LEAF in approximately 8 hours and the Chevrolet Volt in 4 hours. The cost for faster charging time is the necessity of installing a special Level 2 charging port. But, these times still amount to an unreasonable wait for a traveler looking to drive beyond the single-charge range of the vehicle. The newest standard attempts to address this concern. The DC Fast charger can charge vehicles in less than 30 minutes.[4]

Unlike Level 1 and 2 chargers, DC fast is is only available at non-residential charging stations. See Table 2.1 for a comparison of the speed standards. Because compared to Level 2 chargers Level 1 chargers are impractical charging solutions for modern electric vehicles, we do not consider them as part of the charging station network in our analysis.

We further classify charging stations into three groups: residential, public, and private. Residential charging outlets are only accessible to the home owner, and we exclude it from charging network market. The public and private charging stations are installed at public parking lots, malls, restaurants, and so on. We refer them to "charging stations."

The lengthy time to charge these vehicles not only poses a convenience problem but also a congestion problem for public charging stations. According to a charging station usage statistics of PEV owners by the EV Project (2014), owners leave their car charging for 6.2 to 7.4 hours on private away-from-home Level 2 chargers and 3.5 to 4.9 hours on public away-from-home Level 2 chargers. If a PEV driver is looking for a public charging station as a solution for a drained battery, the low turnover at these stations decrease the likelihood of finding an open charger. The current version of the paper does not separately identify inconvenience from these two sources.

[3] The Nissan LEAF, a fully electric vehicle, takes about 20 hours to charge and can get 78 miles per charge. The Chevy Volt, a plug-in hybrid, takes 10 hours to charge and travel 38 miles on its electric engine.

[4] Tesla also features special charging standards for its own vehicles. Tesla Superchargers are stationed around the country with long-range travelers in mind and boast charging speeds of less than 30 minutes. The company is also testing a "battery swap" program for this stations, which swaps the drained battery in a Tesla with a fully charged battery in less than 90 seconds.

2.2.2 Challenges in PEV Adoption: High Price and Range Anxiety

Government support for PEVs has focused on the burden of their high price and the inconvenience from their limited driving range because they are the most common concerns in driving a PEV. Here, we provide descriptive evidence that the consumers indeed face relatively high price and range anxiety. The disutility magnitude of these inconveniences are not identified by the descriptive statistics, and therefore, we model the PEV demand.

High Price

Electric vehicles tend to have higher upfront prices than gas vehicles. In the bottom of Table 2.2, we compare the average price of PEVs and gas vehicles. The average price of PEVs is more than 33% higher than gas vehicles in 2013. More specifically, the most popular fully electric vehicle, the Nissan LEAF had a $28,800 baseline price tag in 2013. In contrast the Honda Civic, the most popular vehicle, cost only $16,555. The most expensive electric vehicle, the Tesla Model S, which cost at least $69,900, while the most expensive gas vehicle among the top 80 CA car models in 2013 cost $52,800.

Another upfront cost in electric vehicle comes from installation of residential charger. According to the Center for Sustainable Energy (2013a), an organization responsible for administrating rebates for electric vehicle purchases in CA, 90% of the PEV purchasers installed a residential charger. For the time period of analysis, the permitting and installation cost for these chargers added about $1,700 in the upfront cost of PEVs.

While PEVs have higher upfront costs, they tend to have lower maintenance and fuel costs than gas vehicles. Without discounting future costs or accounting for changes in the price of oil, the maintenance and fuel costs are about $33,728 for gas vehicle, $20,460 for plug-in hybrid electric vehicle, and $19,344 for full electric vehicles purchased in 2013[5] . Under the assumptions that generate those estimates, the PEV is still overall more expensive than gas vehicles.

[5] The estimates assume 12,400 annual mileage and a 16-year lifespan. Fuel costs for the gas vehicle are calculated based on 26.7 miles per gallon and a gas price of $3.5 per gallon. The fuel cost for PEV is calculated by assuming $0.13 per kWh, taken from the average CA price as of January 2012, and a mileage of 4.42 miles per kWh. The maintenance and repair costs include battery replacement, engine oil change, tire rotation, etc. The assumptions are taken from Alternative Fuel Life-Cycle Environmental and Economic Transportation (AFLEET) 2013. Numerous data sources include AFDC Price Report, AEO Report, Argonne National Laboratory, and so on. See User Guide for AFLEET Tool 2013.

The cost burden of PEVs is of often a binding constraint when consumers are making their vehicle purchase decision. Two surveys, Center for Sustainable Energy (2013a) and Center for Sustainable Energy (2014), reveal that 90% of the respondents consider the rebate was an important decision factor in purchasing their PEV. Additionally, 50% claimed that subsidy for residential charger was also important to their decision.

Range Anxiety

Range anxiety when driving a PEV stems from both their short range and the lack of charging stations. As shown in Table 2.2, the full electric range of PEVs is on average only 15% of the range of gas vehicles. The Nissan LEAF carries a 72-mile range whereas Honda Civic has more than 400-mile range. The Tesla Model S is the only fully electric vehicle with comparable range to traditional gasoline vehicles. According to the Center for Sustainable Energy (2013a), nearly 40% of survey respondents were not satisfied with the electric range of their purchased vehicle (see Table B.3in Appendix B).

In addition to the range issue, public electric vehicle charging stations are still lagging far behind the ubiquity of gas stations. In 2013 there were approximately 9,000 gas stations in California compared to 3,000 charging stations. Table B.4 in Appendix B describes the number of charging stations over time. While the volume of gas cars is significantly higher, these vehicles do not suffer from either the range problems or extensive filling times that electric vehicles do. According to the Center for Sustainable Energy (2013a), 71% of the PEV purchasers were reportedly unsatisfied with the public charging infrastructure (see Figure B.1 in Appendix B). Potential PEV buyers then need to remain highly cognizant of charging stations in areas frequently traveled or avoid long trips out of comfortable range entirely.

Because every standard electricity outlet is effectively a charging station, albeit slow, for an electric vehicle, the importance of a charging station network may not be immediate despite some evidence on the pervasiveness of range anxiety. Further evidence that suggests charging stations should have little impact on demand comes from CA driver habits. The 2010-2012 California Household Travel Survey shows the average Californian drives approximately 38 miles, well within the single-charge range of the Nissan LEAF or Tesla Model S. Table B.6 in Appendix B indicates that the drive range for LEAF owners is even shorter (24.8 miles) per day. These figures suggest that charging away from home may be unnecessary for most use cases.

Other evidence, however, reveals these shorter driving ranges may be a symptom of range anxiety. Table B.6 shows that Volt, which has the range of a standard gasoline vehicle, owners tend to use their vehicles lengthier trips. In addition, despite the short trips PEV owners tend to take every opportunity to charge their vehicles. 27% of charging events were away from home for the LEAF while only 16% were for the Volt in Q4 2013 according to EV Project (2014). Additionally, the survey Center for Sustainable Energy (2013a) noted the importance of having work-place chargers in their decision to purchase their electric vehicle (see Figure B.2).

Even more direct evidence of range anxiety comes from the different charging behavior between full electric vehicles and plug-in hybrid electric vehicles. Figure 2.1 reports how much of the battery remains when a charging event takes place, taken from a sample of participants in an EV Project field study (EV Project (2014)). For the range-impaired Nissan LEAF, charging events take place the most when the battery still is 50% full or more. LEAF drivers charge even though the battery is not close to empty. In contrast, the Chevy Volt is charged primarily when the battery is fully exhausted. It reveals that LEAF drivers make a concerted effort to keep the battery at least half full all the time while Volt drivers let them run out. If range were not a concern for LEAF drivers, we should expect to see similar charging event distribution of Volt.

Range anxiety is frequently cited as a key problem facing the widespread adoption of electric vehicles. The immediate solution is the recent trend among car makers to push plug-in hybrid electric vehicles (PHEV) that use an electric drive engine until the power is exhausted and then switches to a traditional gas engine. While still more expensive than traditional gas vehicles, PHEVs provide the comparable range that fully electric vehicles still lack. Table 2.2 compares the market share, price, and range of the most popular PEVs in California as of Q4 2013.

2.2.3 Government Incentives Encouraging Adoption

To overcome the challenges in adopting electric vehicles, the federal government, California, and local California governments have all offered varying incentive packages.

Easing High Price

Both California, via the Clean Vehicle Rebate Program, and the federal government offer tax credits for the purchase of PEVs. The federal government offers up to a $7500 for fully electric vehicles, while California offers up to $2500. Some districts in California offer rebates on top of these two programs.[6] Table 2.3 provides further details on rebate amounts for specific models.

Other monetary incentives include lower electricity rates for charging electric vehicles at home during off-peak hours. Table B.5 in Appendix B details potential fuel cost savings from these deals. Additionally, work charging, parking lot charging, and public charging stations typically are offered at nominal to no charge.

Easing Range Anxiety

The effort to expand the charging network infrastructure was made by the subsides from the government. In addition to the nearly 2000 charging stations installed in California by ChargePoint America and the EV Project since 2010, over 2500 chargers have been added to the network. They account for 43% of the total charging stations in California as of December 2013. While their origin is not available, representatives of the Department of Energy have suggested many businesses and workplaces have begun installing stations as a convenience to customers or employees. Table B.2 in Appendix B has details about the projects.

Other Incentives

The benefits from the programs describe above are included in our analysis. There are some other incentives we do not explicitly model in the paper. For example, limited numbers of PHEVs and fully electric vehicles can gain access to HOV lanes on California highways. To those that have endure Los Angeles traffic, this program could easily be more valuable than any monetary benefit.[7]

[6] Certain counties in the San Joaquin Valley feature rebates up to $3000.

[7] In fact according the Center for Sustainable Energy (2013a) survey of PEV users, 59% claimed HOV access was an “important” consideration in their purchase decision.

2.2.4 Other Relevant Purchase Factors

A survey of vehicle owners who received vehicle rebates from the CVRP Survey reveals that 38%, 18%, and 16% of LEAF, Volt, and Prius Plug-In Hybrid consumers, respectively, were motivated by environmental concerns in the purchase of their vehicles. LEAF, the only zero emission vehicle in that list, purchasers in particular responded that the environment was their biggest motivator. Other than environmental concerns, high income is also a distinctive feature of PEV purchasers. The first panel of Table B.7 compares the income distribution of households conditional on vehicle purchase versus PEV purchase. The distribution is skewed more to the right for PEV consumers. More than 50% of the PEV consumers have more than \$150,000 annual income.

Because of how important these two demographic features are in distinguishing PEV purchasers from typical vehicle purchasers, we take care to include both in our model of consumers, to be detailed in section 2.4. To proxy for environmental concern in our estimation and in descriptive analysis of the data, we consider political preferences. Consumers are split into three groups — Republicans, independents, and Democrats — with the latter most concerned with environmental issues. While based on party lines, political preference need not be wholly irrelevant in understanding actual consumption preferences. Costa and Kahn (2013) found liberal communities are more likely to participate in "voluntary restraint", that is consume less electricity than conservative but otherwise identical households.[8]

Table B.8 and Figure B.5 both provide descriptive evidence regarding how political preferences and income might affect PEV purchase. The second and third columns in Table B.8 show that more PEVs are sold in high income counties compared to relatively low income counties. However, income is not the only factor explaining the PEV purchase behavior. The fourth column reveals that San Francisco, Marin, and Sacramento counties prefer LEAF over Volt while Orange and Riverside prefer Volt over LEAF. A critical difference in the make up of these counties are political preferences. The former counties tend to have more Democrats than Republicans whereas the latter counties have more Republicans than Democrats. This is shown in Figure B.5. The left panel of contains two wealthy counties (San Francisco and Marin and Orange) but with different

[8] See also Kahn (2007).

average political tendencies. The right panel has two less wealthy counties (Sacramento and Riverside) also with different political preferences. The statistics suggest that income affects the external margin of the PEV purchase decision while political preference may affect which PEV to purchase.

2.3 The Indirect Network Effect Literature

Most recent indirect network literature has emphasized competing platform markets benefiting from the growth of a platform-specific complementary market. Nair et al. (2004), Clements and Ohashi (2005), Dubé et al. (2010), Goolsbee and Klenow (2002), Gandal et al. (2000) analyzed the impact of an indirect network effect from software markets in varying tech industries. This line of research differs from our current project in several ways. First, except for Goolsbee and Klenow (2002), thiese papers have featured non-compatible platforms competing for market concentration. In this market, with a few exceptions, charging stations have been standardized to work across all PEVs.[9] Our ultimate question of interest is in the growth of the relatively new electric vehicle market space, rather than competition among firms in a mature market. In this respect our setting is most similar to Goolsbee and Klenow (2002) which analyzes the diffusion of home computers in the 1990s, where the network size and specific tools, like e-mail, drive adoption more than platform-specific software. Additionally, we may consider the role of a similar peer effect in our setting using the methodology proposed by Bollinger and Gillingham (2012) in which the adoption of solar panels is encouraged by neighbors' adoption.

An important feature of market growth in industries with indirect network effects is the feedback mechanism between the main and secondary markets. In this setting if charging stations increase demand for electric vehicles, greater electric vehicle purchases may in turn encourage the growth of the charging station network or eventually turn it to a profitable venture for the private sector.

The current contributions of this paper to this literature rest in several differentiating factors about our setting. First, a subsample of our observations in the charging

[9] A notable exception might be Tesla's expanding supercharger network. These chargers are meant for long-distance trips, however, rather than every day use.

station network are free from the typical endogeneity problems borne by the feedback effect. Government-sponsored stations were allocated and predetermined before the electric vehicle market resopnded to their presence. Second, the value of the charging station network features significant heterogeneity depending on the geographic location of consumers; the heterogeneity will prove a valuable source for identification. Third, the setting permits demand that can accommodate heterogeneity in the utility towards the indirect network effect, the notable difference from the setting in Gowrisankaran et al. (2014). Hence we do not need to impose that there is a strictly positive or negative relationship between charging stations and demand and can accommodate nuance in tastes for the network currently absent from our model.[10]

2.4 Demand Specification

In the static model consumers make choices in a specific market m defined by a time q and location g. Our estimation will ultimately consider each county-quarter pair in California a specific market. Let $q(m)$ denote the quarter of market m, and $g(m)$ denote the county of market m. Each consumer has four relevant characteristics: income y, political affiliation d, charging stations within a useful range c, and idiosyncratic tastes $\{\nu\}$ toward specific product characteristics. Let $D_i = (y_i, d_i, c_i, \{\nu_i\})$. Specific details on how we model which charging stations are relevant to consumers is in Appendix B. Consumer i in market m purchases one product from the choice set $J_{q(m)}$ or an outside good to maximize her utility. Note we assume the choice set is fixed across geographic markets in a given time period.

Consumer utility is modeled by a variation on the standard random coefficients discrete choice model. Specifically, consumer i's utility from purchasing a vehicle model j has the following form:

$$\begin{aligned} u_{ij} = \alpha \log(y_i - p_j) + X_j\beta + \sum_{k=1}^{K} X_{jk}\sigma_k\nu_{ik} + PEV_j\left[\beta_e + d_i\beta_d + \sigma_e\nu_{ie}\right] \\ + PEV_j c_i\left[\beta_c + \sigma_c\nu_{ic}\right] + \xi_j + \varepsilon_{ij} \end{aligned}$$

[10] For example, we still have not arrived at a convincing proxy for congestion at charging stations. More charging stations are not necessarily better if the usage by other PEV owners is also significantly higher.

Each vehicle model has $K + 2$ observable product characteristics denoted by $W_j = (p_j, X_j, PEV_j)$, where X_j is a vector of K specifications for vehicle j and PEV_j is a indicator variable for plug-in electric or fully electric vehicles. ξ_j is an unobserved (to the econometrician, not consumer) product-specific demand shock, which can be correlated with price p_j, and ε_{ij} is an idiosyncratic preference shock.

This utility model features 2K + 6 parameters of interest $\theta = (\beta', \beta_e, \beta_d, \beta_c, \sigma', \sigma_e, \sigma_c, \alpha)'$. βs are the mean utility derived from a unit of a specific product characteristic, and σs scale the heterogeneity in tastes for each characteristic. We can decompose the utility into components specific to each product, to each consumers, and to PEVs.

$$u_{ij} = \delta_j(\theta) + V_{ij}(\theta) + EV_{ij}(\theta) + \varepsilon_{ij} \tag{2.1}$$

$$\text{where} \qquad \delta_j(\theta) = X_j\beta + \xi_j$$

$$V_{ij}(\theta) = \alpha \log(y_i - p_j) + \sum_{k=1}^{K} X_{jk}\sigma_k\nu_{ik}$$

$$EV_{ij}(\theta) = PEV_j\left[\beta_e + d_i\beta_d + \sigma_e\nu_{ie}\right] + PEV_j c_i\left[\beta_c + \sigma_c\nu_{ic}\right]$$

The terms specific to PEVs are highlighted to emphasize additions to the standard random coefficients model. $[\beta_e + d_i\beta_d + \sigma_e\nu_{ie}]$ captures the utility of purchasing a PEV and may depend on the consumer's political preference. $[\beta_c + \sigma_c\nu_{ic}]$ is the impact of the charging station network on the utility of purchasing a PEV. The magnitude of this term is our measure for the indirect network effect on PEV demand through the charging station network.

Under the assumption that ε_{ij} shocks are iid type-1 extreme value over products and consumers, the probability consumer i in market m purchases vehicle j takes on the familiar logit form.

$$P_\theta(j|W, D_i) = \frac{\exp(\delta_j(\theta) + V_{ij}(\theta) + EV_{ij}(\theta))}{1 + \sum_{j' \in J_{q(m)}} \exp(\delta_{j'}(\theta) + V_{ij'}(\theta) + EV_{ij'}(\theta))}$$

Integrating over all of the individuals in a market yields the aggregate market share. We allow the distribution of individual characteristics to differ by market according to $f_m(y, d, c, \nu)$ but assume the distribution of tastes is $f(\nu)$ is common for all markets and independent from other characteristics. Hence $f_m(y, d, c, \nu) = f_m(y, d, c)f(\nu)$. Finally, we assume that tastes follow a multivariate standard normal distribution with a diagonal variance-covariance matrix. Under these assumptions the market share for vehicle j in market m is

$$S_j^m(\theta|W) = \int P_\theta(j|W, D_i) f_m(y, d, c) f(\nu) d(y, d, c, \nu)$$

The geographic markets we have considered partition California. Within a given time period shares can be further aggregated to the state level by weighting each county with its share of consumers. Let $\pi_q(m)$ denote the fraction of consumers in market m in time period $q(m)$. Then

$$S_j^q(\theta|W) = \sum_m S_j^m(\theta|W)\pi_q(m) \tag{2.2}$$

gives the vehicle j's market share for the state of California in quarter q.

2.5 Data

The California New Car Dealers Association (CNCDA) provides new car registration data for the top 80 car models (accounting for approximately 70% of total sales) every quarter in their Auto Outlook newsletter. This aggregate CA car share data is available for most quarters between 2007 to 2013. In total we have 1674 vehicle-quarter share observations, and an average of 76 vehicles per quarter.[11]

Household-level purchase data is available for PEVs via the California Air Resources Board Clean Vehicle Rebate Project (CVRP). The CVRP dataset includes all PEVs submitted for a CA clean vehicle rebate between March 2010 to the present; according to the Center for Sustainable Energy (2013b) (CSE), which manages the project, from March 2010 to 2013 76% of eligible PEVs were submitted for rebates. Additional eligibility criteria require the vehicle to be a new purchase / lease and for the lease to last at least 3 years. Of these transactions 85% participated in the rebate program.[12]

Each record in the dataset includes the zip code of the purchaser, the vehicle make,

[11] Newsletters in Q3 of 2010 and 2011 and Q2 of 2012 do not report shares. The CNCDA actually reports new car registrations to date in that year. Therefore, quarterly new registrations are backed out by differencing the sales for a particular model between two reports. When one quarter is missing data, the data of the subsequent quarter are also unusable for our estimation.

[12] Other concerns might include varying participation rates by geography or vehicle type. While participation varies from 55% to nearly 100%, it does not appear correlated with the region's income or general political affiliation. A more robust statistic is, unfortunately, not possible with the report's data. Additionally, PHEV rebate participation is slightly lower but generally similar to fully electric vehicle participation.

which typically yields the vehicle model, and the date of the rebate application. The latter is used to approximate the quarter of purchase. A representative from the CSE reported between 50 and 60% of participants apply within five days of purchasing the vehicle while nearly 90% apply within 50 days. The mean delay between purchase and rebate application is approximately 25 days.

For each vehicle we collect price (MSRP of the base model) and various standard characteristics from AutoTrader.com and Edmunds. The characteristics included in the estimation are horsepower-weight, length-width ratio, drive type dummy (2 or 4 wheel base), miles per dollar, driving range, and fuel type (electric vehicle dummy). Several are included as standards in the literature, others to emphasize characteristics that typically differentiate electric vehicles from gasoline models.

The numerous incentive programs for electric vehicles described in section 2.2 require adjustments to the baseline specifications for these vehicles. In particular, we consider electric vehicles characteristics taking into account 1) the federal and state rebate, 2) the cost of installing a charging station for their home, and 3) miles per dollar using home charging prices at the cheapest price under their energy provider.[13] We get these rates from the dominant energy provider in the county.

Charging station data is collected from the Alternative Fuels Data Center (AFDC), which in turn collects from major charging station operators Blink, SemaCharge, and Chargepoint, and Open Charge Map (OCM). These data sets provide longitude and latitude coordinates thus allowing great flexibility in how we spatially aggregate the data for use in the model. Data on charging station opening dates are provided in older snapshots (before 2014) of the datasets. More recent snapshots do not include opening dates and deduced by assuming stations that appear between snapshots opened in that period.[14]

[13] It might be unreasonable to assume the consumer always charges at the cheapest rate at home, but this method also accounts for the possibility that a lot of charging is actually free from work chargers. We assume consumers install Level 2 chargers at home based on the Center for Sustainable Energy (2013a) survey of CVRP benefactors that found approximately 90% of respondents own a Level 2 home charger for their PEV.

[14] Operators provide the AFDC with opening dates of new stations, but OCM also relies on "community submissions" which might be less reliable than the information provided by operators. One may also worry OCM stations are consistently reported "late" after there is enough of an electric vehicle community in the area to report the station.

We use the Federal Election Commission (FEC) and American Community Survey (ACS) of the US Census Bureau to estimate parameters on political orientation and aggregated income distributions. The ACS is the Census Bureau's recent replacement of the decennial long-form survey, now featuring a continuous survey of up to 250,000 households every month covering a wide range of subjects. Among the available statistics is an annual income distribution estimates across 10 income brackets for most counties in California. Smaller counties are covered by estimates generated from 3 or 5 years of estimates. FEC data provide information on all political contributions over \$250. Records are also sufficiently detailed to match most contributions with a recipient political party. We discuss the procedure for generating the estimate of the joint distribution $f_m(y, d, c)$ from the ACS and FEC in Appendix B.1.

We also utilize demographic data specific to the subset of households that purchase automobiles and electric vehicles with the 2010 to 2012 California Department of Transportation CA Household Travel Survey (CHTS) and demographic survey data collected by the CVRP.[15] The former is updated every 10 years and collects demographic information. In total the 2012 to 2013 survey included 42,431 households.[16] 5,717 of the participant households were additionally given GPS devices to track detailed movement information. We use both sets of data to construct highly aggregated statistics on the income of new vehicle purchasers as well as to understand travel behavior to determine relevant charging stations for individuals in certain areas. The latter point is discussed further in Appendix B.2. CVRP survey participants also report their incomes and thus is used to aggregate statistics on the income of electric vehicle purchasers.

2.6 Estimation

2.6.1 Reducing the Parameter Space

To simplify the estimation we reduce our parameter space by a standard technique in the literature. Recall from Equation 2.1, $\delta_j(\theta) = X_j\beta + \xi_j$, the product-specific term common to all consumers. Because the mean parameters β are linear in δ, we can back out estimates for them after the nonlinear search over δ. Therefore, we refine

[15] See Table B.10 in Appendix B for survey size details.
[16] Despite the name of the survey, households were surveyed from 2012 to early 2013.

$\theta = (\beta_e, \beta_d, \beta_c, \sigma', \sigma_e, \sigma_c, \alpha)'$, now with only $K + 6$ parameters. Searching over the J components of δ is expensive, however, so we utilize the share inversion technique introduced by Berry (1994) to "concentrate out" these parameters. This technique requires the restriction that CA-level vehicle shares should match the predicted shares generated by the model at the true parameter values. That is

$$S_j^{DATA} - S_j^{CA}(\delta, \theta_0) = 0 \quad \forall j \tag{2.3}$$

Given our distributional assumption on consumer tastes, Berry (1994) demonstrates that for each θ, there is a unique $\delta(\theta)$ such that Equation 2.3 holds. This technique proves useful not only as a mechanism to reduce the parameter space but also to mitigate the endogeneity problem with price by conditioning on the component of the error with which it is correlated, i.e. ξ_j. Given this restriction three more sets of moments identify the parameters θ.

2.6.2 BLP Moments

The first set of moments are the standard demand moments from Berry et al. (1995) using the data set $V = \{V_j\}_{j=1}^{J}$, the collection of car characteristics W_j and corresponding CA state shares S_j^{DATA} for the J vehicle shares observed. For the benefit of imposing any restriction on the joint distribution of price and ξ_j, these moments require the relatively strong assumption that the demand shock $\xi_j(\theta)$ is mean independent of the observed product characteristics X_j for j in the same markets $(m(j))$ at the true parameter values, i.e.

$$E[\xi_j(\theta_0)|(X_j)_{j\in m(j)}] = 0 \tag{2.4}$$

Under condition 2.4 the equation for δ is a standard linear regression model, and $\xi_j(\theta)$ can be backed out as the error from $\delta_j(\theta) - X_j\hat{\beta}_{OLS}$. We are also still exploring an alternative set of moments proposed by Petrin and Seo (2014) allowing for X to be endogenous.

We follow BLP in generating three types of instruments for our first set of moment conditions. For each product the corresponding instruments are

$$Z_j = (X_j, \sum_{j'\in J_f\setminus j} X_{j'}, \sum_{j'\in J\setminus J_f} X_{j'})'$$

where J_f are the set of products offered by firm (make) in the particular market. Hence the instruments are the observable characteristics of the vehicle, the sum of characteristics across the firm except product j, and the sum of characteristics across competitors' product. These instruments proxy for supply-side variables that factor into firm and industry-wide costs correlated with prices but uncorrelated with demand shocks. Theoretically, Z_j is a vector of length $3K$ by 1.[17] These first set of moments can be combined into a $3K$ by 1 vector

$$E[\psi_1(\theta_0, V_j)] \equiv E[\xi_j(\theta_0) Z_j] = 0 \tag{2.5}$$

with corresponding sample moment at arbitrary θ, $\frac{1}{J}\sum_{j=1}^{J}\psi_1(\theta, V_j)$.

2.6.3 Maximum Likelihood Moments

The second set of moments are the score of a maximum simulated likelihood estimator of θ using the CVRP micro purchase data $M = \{M_i\}_{i=1}^{N}$. Unlike the aggregated data, estimators using this sample can take advantage of much greater variation in charging stations and purchasing observations across zip codes and quarters.

Because we observe the time of purchase, model of purchase, and location of individuals in this data set, it is straightforward to write down a likelihood function. For consumer i in market m let I_i be a vector of length $J_{q(m)}$ by 1 of purchase indicators (1 for the vehicle the consumer purchased and specially denoted $I_{ij(i)}$). Conditional on all relevant characteristics of person i, i.e. $D_i = (y_i, d_i, c_i, \{\nu_i\})$, the log probability of observing I_i under parameter θ is given by

$$\frac{1}{N}\sum_{i=1}^{N} L(\theta; M_i) = \frac{1}{N}\sum_{i=1}^{N} I_{ij(i)} \log(P_\theta(j(i)|W, D_i))$$

Following the argument of Goolsbee and Petrin (2004) we claim that conditioning purchase probability calculations on the value of $\delta(\theta)$ described previously eliminates the specification error that might arise because of the endogeneity of price.

However, the available data M_i for the individual does not include income, political affiliation, or idiosyncratic tastes so the probability of purchase must be simulated for

[17] In practice we test for strong collinearity among the instruments and drop those that fail this test.

each individual by integrating over the distribution $P_m(y,d)P(v)$.[18]

After replacing the purchase probability for a consumer with characteristics M_i with the simulated probability $\hat{P}_\theta(\cdot|W,M_i)$, the sample log likelihood function is

$$\frac{1}{N}\sum_{i=1}^{N}\hat{L}(\theta;M_i) = \frac{1}{N}\sum_{i=1}^{N} I_{ij(i)} \log\left(\hat{P}_\theta(j(i)|W,M_i)\right)$$

The score of the log likelihood function generates $K+6$ moment conditions

$$E[\psi_2(\theta_0,M_i)] \equiv E\left[\frac{\partial L(\theta_0;M_i)}{\partial \theta_0}\right] = 0 \tag{2.6}$$

with corresponding sample moment at arbitrary θ, $\frac{1}{N}\sum_{i=1}^{N}\psi_2(\theta,M_i)$.

2.6.4 Matching Income Distribution Moments

The last set of moments takes advantage of aggregated demographic statistics derived from survey participants in the CHTS and a survey of CVRP participants, data sets denoted by $A = \{A_r\}_{r=1}^{R}$ and $B = \{B_l\}_{l=1}^{L}$, respectively. Following the insight of Imbens and Lancaster (1994) these aggregated statistics are simply aggregations of micro data; hence aggregate model predictions should match these statistics. Similar to matching moments considered by Petrin (2002) we consider three conditions matching observed income distributions of various categories of car purchasers against model predictions.

The first condition matches 13 income category densities conditional on purchasing a particular make of electric vehicle. Using the data from the CVRP survey and invoking Bayes' rule, i.e. $P(A|B) = P(A,B)/P(B)$, the sample statistics derived are

$$\hat{P}(income_i \in [q_k, q_{k+1}]|i \text{ purchases EV from firm } f) = \frac{\sum_{l=1}^{L}\mathbb{1}(l \text{ purchases EV from firm } f) * \mathbb{1}(income_i \in [q_k,q_{k+1}])}{\sum_{l=1}^{L}\mathbb{1}(l \text{ purchases EV from firm } f)}$$

for $k \in \{1,\ldots,13\}$ and for f, the six firms that produce electric vehicles in the sample time period, for a total of 78 statistics. q_k for $k \geq 2$ are income cutoffs specified in the CVRP survey data. We assume that the samples are unbiased so the sampling

[18] Note c is not simulated. Since we observe the time of purchase and the zip code of the purchaser, we can assign the effective number of charging stations to the individual.

error is 0 in expectation.[19] We will denote the population version of this probability distribution by $\mu = E[\hat{P}(income_i \in [q_k, q_{k+1}]|i \text{ purchases EV from firm } f)]$, where the expectation is taken over sampling error.

The corresponding statistics generated by aggregated model predictions are derived by the following equation.

$$\hat{P}_\theta(income_i \in [q_k, q_{k+1}]|i \text{ purchases EV from firm } f) = \frac{\sum_{m=1}^{M}\sum_{i=1}^{ns}\left[\mathbb{1}(income_i \in [q_k, q_{k+1}]) * \sum_{j\in\{J_q(m)\cap J_f\cap EV\}} \hat{P}_\theta(j|W, D_i)\right]}{\sum_{m=1}^{M}\sum_{i=1}^{ns}\sum_{j\in\{J_q(m)\cap J_f\cap EV\}} \hat{P}_\theta(j|W, D_i)}$$

where $\hat{P}_\theta(j|W, D_i)$ is simulated by integrating over the distribution $P_m(y, d, c)P(v)$.[20] The summation is only over markets covered by the CVRP survey, and EV is used here to denote the set of electric vehicles.

The moment restriction imposes that at the true parameter the model's aggregated statistic and the population statistic should be equal.

$$\begin{aligned} E[\psi_3(\theta_0, B_l)] \equiv \\ &\mu(income_i \in [q_k, q_{k+1}]|i \text{ purchases EV from firm } f) \\ &\quad - P_{\theta_0}(income_i \in [q_k, q_{k+1}]|i \text{ purchases EV from firm } f) = 0 \end{aligned}$$

The sample equivalent evaluated at arbitrary θ is $\frac{1}{L}\sum_l \psi_3(\theta, B_l)$ and reflects that all sampling error is driven by the sample distribution used in place of the population distribution μ. Similarly, we construct 20 more moments by matching 10 income category densities conditional on purchasing any vehicle and on purchasing any electric vehicle based on CHTS survey data. The CHTS-based moments are denoted by $\psi_4(\theta, A_r)$.

[19] This assumption only requires 1) the CVRP participants taking the survey are not a biased sample of the pool of all CVRP participants and 2) the CHTS sample of households is also not a biased sample of auto purchasers in California.

[20] Note unlike the MLE simulation, here we must also simulate the location of an individual within a market and hence the effective number of charging stations.

2.6.5 GMM Estimator

These four sets of moments can be stacked into a single vector. Formally, we assume that θ_0 uniquely satisfies the population moment conditions described above.

$$E\left[\psi(\theta_0, U)\right] = E\begin{bmatrix}\psi_1(\theta_0, V)\\ \psi_2(\theta_0, M)\\ \psi_3(\theta_0, B)\\ \psi_4(\theta_0, A)\end{bmatrix} = 0 \tag{2.7}$$

with sample analog $\hat{\psi}(\theta, U)$, where $U = (V, M, A, B)$. The GMM estimator $\hat{\theta}$ solves the following criterion function

$$\hat{\theta} = \arg\min_{\theta} \hat{\psi}(\theta)' W \hat{\psi}(\theta)$$

where W is a weighting matrix. In practice we follow the Hansen (1982) two-step method, by first estimating parameters where W is the identity matrix and then with W as the inverse of the asymptotic variance matrix of the moments derived from the first step.

2.7 Identification

2.7.1 Identifying Taste Heterogeneity

Our utility model requires identifying $K + 6$ non-linear taste parameters from variations across choice sets, charging stations, and demographics in our four data sets. The standard argument for identifying heterogeneity in tastes using aggregate data follows from Berry et al. (1995). Changes in market shares as choice sets vary across different markets serve as proxies for substitution patterns that the parameters on consumer tastes must explain. One reason we conduct our analysis on all vehicle types rather than just PEVs is to increase variation in choice sets across markets. Moments matching aggregated statistics conditioned on consumer demographics complement potentially weak identification from the aggregate share data. These moments penalize parameter values that fail to reproduce aggregated statistics.

Identifying parameters on terms related to PEVs rely on our data specific to PEV purchases, in particular our rich micro data set from the CVRP. Because electric vehicles, particularly plug-in hybrid electric vehicles, are similar in performance and range to traditional gas-powered vehicles, differences in shares of these electric vehicles from similar gasoline vehicles must be driven largely by PEV-specific tastes.

We also observe significant variation in purchases of electric vehicles as a category and within that category across counties and time. Across geographic markets consumers differ on three dimensions that can explain the differences in purchasing behavior: charging stations, income, and political affiliation. The latter two vary most across geographic markets rather than time, while charging stations vary significantly even over short periods of time as well as across counties. Fixing the number of charging stations at a snapshot in time can thus help pin down income and political orientation parameters, i.e. (α, β_d).[21]

Significant variation in the number of charging stations across time within markets can pin down parameters relating to charging stations. Because markets are defined by time periods as short as a quarter, we do not worry about general improvements in the perception of electric vehicle usage absorbing most of the change in purchase behavior of PEVs from quarter to quarter. Additionally, given a market, the only county-specific factors that significantly change over this short time period are the number of charging stations.[22]

2.7.2 Endogeneity

Two sources of endogeneity could potentially affect the estimation procedure. The first is the typical assumption that unobserved product characteristics ξ_j are correlated with price p_j. *A priori* these two variables should be positively correlated and generate a positive bias in price elasticity (price elasticity is less negative). Because electric vehicles still tend to carry a higher price than similar gasoline equivalents, the model with bias price would predict consumers are less hesitant to buy high-price electric vehicles than at

[21] See again, for example, Table B.8.

[22] We are, however, concerned about other trends that might be more sensitive to short periods of time. For example the growing peer effect over time might impact purchase of electric vehicles more than new charging stations. In future iterations of this draft, we intend to use our detailed purchase data to proxy for this more direct network effect. A second concern is that charging stations are always increasing over time; we do not observe many instances of charging stations being shut down. For the future research, we also consider a more natural definition for charging stations, which considers congestion. Effective charging stations by this definition do not necessarily increase over time hence breaking the potential conflation of time and charging stations.

the true parameter. Ultimately, the bias can force down the magnitude of PEV-specific utility terms to explain the low shares of electric vehicles. As detailed in section 2.6 we address this issue with the standard tactic in this literature by directly specifying the component of the error with which price is correlated (ξ_j) in our calculation of demand.

A more serious issue for analysis is the potential that charging stations are also endogenous. While we expect demand to increase the number of charging stations, it is also reasonable to suspect the number of charging stations in an area are driven by local demand. The assumption that the unobserved characteristics ξ_j is uniform across counties can amplify the endogeneity. If a county has strong preference towards PEVs, charging stations may be deployed more in the county. However, uniform unobserved demand shock ξ_j may not capture the county-specific preference, leading to an omitted variable problem. The correlation between omitted county specific preference and charging station may result in upward bias of the coefficient.

We deal with this problem in three ways. First, government programs determined the location of 43% of charging stations opened in our sample time period. In conversations with representatives involved in these projects, we learned that the government targeted large areas, such as Los Angeles or San Diego, for receipt of charging stations but specific locations were determined independent of demand. The projects placed charging stations wherever willing partners could be found. Second, we find that charging stations at the household's home zip code are a worse predictor of demand than charging stations in associated "work" zip codes (as determined by travel information in the CHTS).[23]

Hence PEV demand is very sensitive to the specific location of these stations. The potential endogeneity problem can then be mitigated by 1) blunt placement of charging stations and 2) considering consumers at the high level of granularity we do.

Finally, charging stations might trend with market-level shocks favoring electric vehicles. Since ξ_j for PEV vehicles is precisely this shock, specifying this component in the calculation of demand mitigates the endogeneity problem.[24]

[23] See Appendix B.2 for details.

[24] Of course, if the number of charging stations is actually highly correlated with shocks specific to a tight array, specifying ξ_j will not completely eliminate the endogeneity problem. In this case, we would need a more sophisticated model of how charging stations are placed. A future iteration on this paper may revisit that question.

2.8 Conclusion

The primary goal of this paper is to accurately estimate the magnitude of a secondary market's impact on the growth of a first and to apply that technique to understanding the role incentives and the charging station network have on the growth of plug-in electric vehicle adoption. This draft takes a preliminary step toward that goal by estimating a rich discrete choice model of demand with heterogeneous tastes toward the "network" good in a static setting. Our initial results suggest that the elasticity of demand for PEVs with respect to charging stations is not negligible. In our sample period PEV market share could have increased by as much as 1.6 to 7% (depending on specific PEV model) with a 1% increase in charging stations across California. We also found the role of other incentives was significant, however. We estimate PEV penetration would have been less by .78% without California or federal tax incentives.

We warn, however, that ignoring the feedback mechanism between the primary and secondary markets is a potentially egregious error that may underestimate the importance of early movers in either of these markets to the long-term growth of the market. For a classically durable good like automobiles, shutting down the ability for consumers to forward look with respect to potential growth in the secondary market is an another significant issue that may understate the importance of the network to consumers. We hope to address both of these concerns with the next draft of the paper.

2.9 Tables and Figures

Table 2.1: Charging Speed by Charger Type

Charger Type	Vehicle	Level 1	Level 2	DC Fast
Full Charge Time	LEAF	20 hr	8 hr	30 mins[a]
	Volt	10 hr	4 hr	15 mins[a]
Distance with 1 hr of Charge		2-5 mi	10-20 mi	60 mi
Home Installation Fee		Free	~$1700	N/A

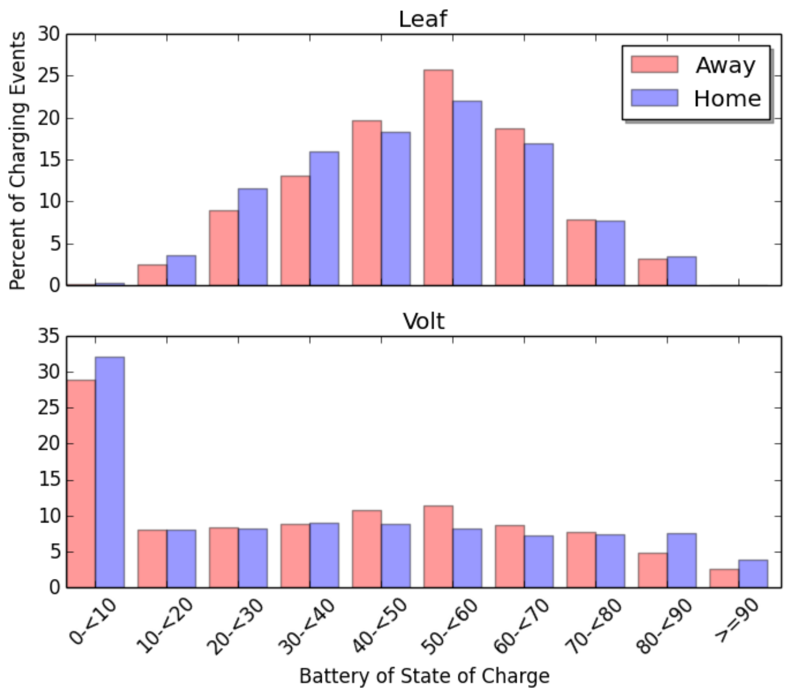
[a]Time for battery to be 80% charged.

Table 2.2: List and Characteristics of PEVs Available in Q4 2013 in California

Make	Model	PHEV	Share[a]	Price	Range[b]	Miles / $[c]
Chevrolet	Spark EV		0.0007	$26685	81.94	24.80
Nissan	LEAF		0.0078	$28880	72.98	20.70
FIAT	500e		0.0038	$31800	87.01	24.02
Smart	ForTwo Electric Drive		0.0009	$25000	76.05	22.62
Chevrolet	Volt	YES	0.0081	$39145	380 (38)	20.22
Honda	Fit EV		0.0002	$36625	81.98	24.88
Ford	Focus Electric		0.0005	$39200	75.95	20.17
Toyota	Prius Plug-In Hybrid	YES	0.0062	$29990	540 (11)	10.03
Ford	CMAX Energi	YES	0.0046	$35340	550 (20)	18.42
Toyota	RAV4 EV		0.0007	$49800	102.87	15.82
Honda	Accord Plug-In Hybrid	YES	0.0002	$32000	570 (13)	23.97
Tesla	Model S		0.0070	$69900	209.12	19.13
Average	PEV		0.0034	$36947	235.66 (72.49)	20.40
Average	Gas (Top 70%)[d]		0.0155	$27656	439.27	6.68

[a] Clean Vehicle Rebate Project rebate data set.
[b] For PHEVs electric drive range is in parentheses.
[c] Miles per dollar (MP$) is calculated assuming average Time of Use rate offered by utility companies in California. Miles per gallon for PEVs is substituted by MPGe, which uses the equivalency 33.7kWh = 1 gallon.
[d] We only took the list of gas vehicles within the top 70% of the market share.

Figure 2.1: Distribution of Battery Charge at the Start of Charging Events



Source: EV Project Electric Vehicle Charging Infrastructure Summary Report, Q4 2013

Table 2.3: Government Monetary Incentives for PEV

	Clean Vehicle Rebate Project	Federal Tax Credits[a]
Project Period	2010 - current	2009 - current
Area Covered in CA	All	All
Funding Institute	California Air Resources Board	IRS
Funding Amount	$58.5 million dollar[b]	Until manuf. sells 200K
Rebate Amount	$1,500 to $2,500	$2,500 to $7,500
Leaf Rebate	$2,500	$7,500
Volt Rebate	$1,500	$7,500
Prius Rebate	$1,500	$2,500
Eligibility	Zero emission, plug-in hybird	Zero emission, plug-in hybird
	Lease $\geq$ 36 months	Battery capacity $\geq$ 5KWH
	Battery, Hydrogen Fuel EV	Only battery EV
Total Rebates Issued[c]	35,769	

[a] Plug-in Electric Drive Vehicle Credit (IRC 30D) by IRS

[b] Distributed thus far. FY 2014 - 2015 alone has a new allocation of over $100 million.

[c] As of December 31, 2013.

Chapter 3

Identification and Estimation of Discrete Choice Demand Models when Observed and Unobserved Characteristics are Correlated

3.1 Introduction

The identification of discrete choice demand models since Berry (1994) and Berry et al. (1995) (BLP) has relied on the assumption that observed product characteristics (excluding price) are uncorrelated with the unobserved product characteristics. Questions have been raised as to whether the assumption is reasonable and what the intuition is for identification of the heterogeneity in taste parameters. Noisy demand estimates are often the outcome with market level data.

Since Spence (1976) economists have understood that firms' decisions about prices and characteristics are driven by their beliefs about the distribution of consumer preferences as both marginal and infra-marginal consumers enter the first-order condition for profit maximization. In this paper we show how to use these first-order conditions to estimate demand and supply parameters in a setting with product-specific unobserved demand and cost shocks. We endogenize all product characteristics by letting multi-product firms choose prices and observed and even unobserved characteristics. We allow firms information sets at the time they choose characteristics to potentially include other firms' product characteristics, demand, and cost shocks, signals on all of these, or no information at all on them. Ex-post firms may wish they had made different decisions and our identification is based on the assumption that firms are correct in their choices on average (Hansen (1982)).

As with the GMM approach of BLP, we also write our estimator in terms of the moments implied by the first order conditions. The estimation procedure is exactly the same as in BLP, except we use replace their moment conditions with firms' first oder condtions. Thus, just as with BLP one can directly add supplemental moments that may further help with identification, as in Petrin (2002) or Berry et al. (2004).

Applying our approach on the same automobile data BLP used, point estimates are found reasonable for mean and variance in tastes as well as cost side parameters. Some puzzling point estimates from BLP estimation, such as negative mean marginal utility on miles per dollar and some cost shifters – miles per gallon and size, go away with our approach. They are *very precisely estimated* with same market-level data compared to BLP estimates. We find significantly more precise estimates given the same exact data equivalent to approximately a sixteen-fold increase in the number of observations. In turn, we show that only the half of the samples are enough to identify parameters significantly.

Given the estimates, we test if conventional assumption in identifying heterogeneity in taste is valid. We strongly reject the standard identification assumption; the conditional correlations of the demand and cost unobservables with observed characteristics equal 0.85 and 0.7 respectively. In addition, we find that BLP instrumented prices are still highly correlated with the demand error.

Our approach provides an economically intuitive interpretation on data. We find that the implied elasticities are higher and markups are lower than BLP estimates. Moreover, the mean marginal utility of purchasing a new automobile is positive while BLP predicted it negative. Our approach fits the demand model to the data – 10% market share of automobile industry – by interpreting that consumers like a new car but they are too price sensitive to purchase one, leading to only 10% of market share.

On the other hand, BLP's fits the model to the data by predicting that consumers are not quite price sensitive but they generally do not like a car, leading to 10% market share.

The identification in the paper is carried out by exploiting optimal decisions in identification. The idea is similar to the approaches taken by Hansen (1982), Pakes et al. (2015) (henceforth PPHI), Spence (1976), and Veiga and Weyl (2014) which used optimization for identification. Some previous work, including Mazzeo (2002), Sweeting (2007), Crawford and Shum (2007), Lustig (2008), Gramlich (2009), Eizenberg (2014), Fan (2013), Blonigen et al. (2013)[1] Multiple studies (Mazzeo (2002), Sweeting (2007), Crawford and Shum (2007), etc.) argue that the product characteristics are endogenously chosen by the firm.allowed for endogenous product characteristics in estimation. However, all of these papers maintain the assumption that the demand and supply shocks are orthogonal to product characteristics while we do not impose restrictions on joint distribution of observed and unobserved characteristics. In addition, this paper is unique since we also endogenize unobserved characteristics.

In Section (3.2), we specify demand and supply system and describe strength and weakness of the identification strategy that previous studies used in the discrete choice demand estimation since BLP. In Section (3.3), we present our approach to identify heterogeneity in tastes and supply parameters by using market-level data. Estimation and choice of instruments are suggested in Section (3.4). We apply our approach to the same automobile data BLP used in (3.5). The demand and supply estimation results using the BLP data are shown in Section (3.6).

3.2 BLP Demand and Supply

The random coefficients discrete choice demand system from Berry et al. (1995) (BLP) considers a market with a set of products J, each of which is defined as a vector of K observed characteristics and price $(X_j, p_j) \in \mathbb{R}^{K+1}$ and unobserved (to the econometrician) characteristic ξ_j. Product $j = 0$ typically refers to outside option with $X_0 = p_0 = \xi_0 = 0$, anchoring the level of the indirect utility. Given the common taste parameters θ consumer i is defined by random coefficients $(v_{ik})_{k=1}^{K}$ and "taste" shock ε_{ij}

[1] See review in Crawford (2012) for details.

which is assumed to be independent and identically distributed across consumers and products. Utility that consumer i derives from good j – $u_{ij}(\theta)$ – is a function of model parameters $\theta = (\alpha, \beta, \sigma) \in \mathbb{R}^{1+K+K}$, observed and unobserved product characteristics j, prices, and consumer i's characteristics:

$$\begin{aligned} u_{ij}(\theta) &= \delta_j(\theta) + \sum_{k=1}^{K} \sigma_k v_{ik} X_{jk} + \epsilon_{ij} \\ \text{where} \qquad \delta_j(\theta) &= -\alpha p_j + X_j' \beta + \xi_j .^2 \end{aligned}$$

δ_j represents the utility component common to all consumers and includes unobserved characteristic ξ_j. α refers to marginal utility of income parameter, β is the mean taste for observed characteristics, and σ_k indicates the extent of heterogeneity in tastes among consumers for characteristic k.

Consumer chooses one and only one product j which yields the highest utility:

$$u_{ij}(\theta) \geq u_{ij'}(\theta), \quad \forall j'.$$

The individual probability of purchasing a good j, s_{ij}, is derived as following if idiosyncratic demand shock ε_{ij} follows i.i.d. Type 1 Extreme Value distribution:

$$s_{ij}(p, X, \xi;\ \theta) = \frac{exp\left(-\alpha p_j + X_j' \beta + \xi_j + \sum_k \sigma_k v_{ik} X_{jk}\right)}{\sum_{j' \in J} exp\left(-\alpha p_{j'} + X_{j'}' \beta + \xi_{j'} + \sum_k \sigma_k v_{ik} X_{j'k}\right)}.$$

Integrating over over consumers who prefers good j the most yields the market share of good j given a density $F(\cdot)$ of heterogenous tastes:

$$s_j(p, X, \xi;\ \theta) = \int_{\{u_{ij}(\theta) \geq u_{ij'}(\theta),\ \forall j'\}} s_{ij}(p, X, \xi;\ \theta)\ dF(v, \varepsilon).$$

The share is a nonlinear function of characteristics and price. One can estimate the parameters with MLE with either individual level data on choices or market-level data on product shares.

Ignoring the unobserved characteristic ξ in demand estimation often results in low price elasticities, sometimes even positive as shown in Trajtenberg (1989), due to endogenous price. If price is positively correlated with unobserved product quality, then the price coefficient is upward biased, leading to low elasticities. Since the market share is nonlinear in price and ξ, one cannot apply instrumental variable approach to correct the endogeneity problem.

One of contributions of the BLP method is to control correlation of prices and unobserved characteristics by inverting out δ. BLP proved that if the goods are weak substitutes, at any given nonlinear parameter σ, there exists a unique vector of $\delta(\sigma) = (\delta_j(\sigma))_{j \in J}$ which exactly matches observed shares in data to predicted shares:

$$s_j\left(p, X, \xi\left(\delta(\sigma), \theta\right); \theta\right) = s_j^{\text{data}}, \quad \forall j. \tag{3.1}$$

Note that once one has $\delta_j(\sigma)$ then conditional on α and β one knows $\xi_j(\theta) \equiv \xi_j\left(\delta(\sigma), \theta\right), \forall j$. The approach controls for the unobserved characteristic $\xi_j = \delta_j + \alpha p_j - X_j'\beta$ by holding δ_j constant (i.e. holding ξ_j constant) during estimation of σ.[3] Since price and the unobserved characteristics are linear to δ, price can be instrumented to derive price coefficient α. We make use of this inversion to "back out" the unobserved characteristic for any value of θ allowing us to treat it as another observed characteristic over which the firm makes choices.

BLP's solution to allow for both correlation between price and unobserved characteristics and heterogeneity in tastes has been to assume that the observed and unobserved product characteristics are mean independent:

$$E\left[\ \xi_j \mid X\ \right] = 0 \quad \forall j. \tag{3.2}$$

Any function of $X = (X_j)_{j \in J}$ can serve as instruments for ξ_j under the assumption (3.2). Instruments used by BLP are own product characteristics X_j, other product characteristics within own firm, and competitor's product characteristics to instrument price and taste heterogeneity parameters. However, if firms set their observed and unobserved qualities at the same time, there may be reason to believe that this assumption (3.2) is violated. In fact, we exploit this optimization for identification without relying on the BLP assumption.

As estimating marginal cost parameters BLP assume that prices conditional on observed an unobserved demand and cost factors are set according to Bertrand-Nash

[3] In practice researchers have found that the set of observed product characteristics they observe can explain only part of demand conditional on price, leading much of the demand model to load on the ξ_j's, as in (e.g.) Petrin (2002) or Goolsbee and Petrin (2004).

competition. Solving the profit maximization problem by each firm indexed by f yields J first order conditions with respect to prices:

$$s_j + \sum_{j' \in J_f} \left(p_{j'} - mc_{j'}\right) \frac{\partial s_{j'}}{\partial p_j} \quad = \quad 0, \quad \forall j \in J. \tag{3.3}$$

where J_f is the set of goods produced by firm f. Marginal costs mc_j $\forall j$ are derived by inverting the J first order conditions (FOCs) in Equation (3.3). Supply side parameters are then estimated by the following equation:

$$ln\left(mc_j\right) = W_j'\gamma + \omega_j$$

where W_j are cost shifters, typically X_j itself or log of it.

The BLP identifies cost parameters γ by assuming that the cost shocks are uncorrelated with cost shifters – analogous to the demand identification condition (3.2):

$$E\left[\ \omega_j\left(\theta_0, \gamma_0\right)\ \mid W\ \right] = 0 \quad \forall j. \tag{3.4}$$

Since the observed cost shifters generally include the product characteristics observed by the researcher the realized cost shock will generally contain the unobserved product characteristics. Supply side identification thus rests heavily on the same maintained assumption on the demand side but for the marginal cost equation. Our approach avoids this assumption (3.4) by coupling the price FOCs from BLP in (3.3) with FOCs with respect to the optimal choice of characteristics.

3.3 Identification: Competition in Prices and Characteristics

Instead of relying on identification conditions (3.2) and (3.4) used in the literature since BLP, we use the equilibrium conditions where multi-product firms choose prices and observed and unobserved characteristics in Bayes Nash manner. We exploit the firms' optimal choices on characteristics and price to infer firms' beliefs about consumer demand as both marginal and infra-marginal consumers determine firms' decisions. That is, firms' first order conditions (FOCs) with respect to characteristics contain information on marginal and infra-marginal consumers, and we identify firms' beliefs on the distribution of consumer tastes from those FOCs. Same idea applies to identifying cost parameters as they enter the FOCs. Since do not assume that the observed and unobserved characteristics are uncorrelated, we can test mean independence hypotheses on both demand and supply side.

Firms compete with each other in two steps. In the first step, they compete in characteristics (both observed and unobserved) conditional on some information set, knowing that they will ultimately compete in prices in Bertrand Nash fashion in the second step given the realized portfolio of products in the market. The information set of a firm f, I_f, may include other firms' product characteristics, cost shifters, demand and cost shocks, some signals on the shocks, or no information at all on them. For example, in our application to automobile case, we consider one possible information set containing competitors' observed and unobserved characteristics as well as prices from the contemporaneous year. One may include some forecasts on them based on previous years. Once the characteristics are realized, they compete in price in the second step. Equilibrium is Bayes Nash in the choices of product characteristics and prices conditional on characteristics, given each firm's information set.

Formally, firms' equilibrium choices are described as following. Define $Z_j = (X_j, W_j, \omega_j)$, and $Z = (Z_j)_{j \in J}$. Redefine K be the number of characteristics, including ξ, and let $\theta = (\alpha, \beta, \sigma, \gamma)$. In the first step, firm f chooses vectors $X_f = (X_j)_{j \in J_f}$ and $\xi_f = (\xi_j)_{j \in J_f}$ given some information set I_f by maximizing the expected profit:

$$\begin{aligned} \max_{X_f} \quad & E\left[\Pi_f \mid I_f\right] \\ \Pi_f \quad = & \sum_{j' \in J_f} \left(p_{j'}\left(Z, \xi; \theta\right) - mc_{j'}\left(Z_{j'}, \xi_{j'}; \theta\right)\right) \ s_{j'}\left(p(Z, \xi; \theta), Z, \xi; \theta\right). \end{aligned}$$

with prices determined after characteristics are set Bertrand-Nash. Conditional on the information set the FOC for characteristic k of product j is given by

$$\begin{aligned} FOC_{jk}\left(\theta\right) &= \frac{d\,\Pi_f}{d\,X_{jk}} \\ &= \sum_{j' \in J_f} \Bigg[\left(p_{j'}\left(Z, \xi; \theta\right) - mc_{j'}\left(Z_{j'}, \xi_{j'}; \theta\right)\right) \frac{d\ s_{j'}\left(p(Z, \xi; \theta), Z, \xi; \theta\right)}{d\ X_{jk}} \qquad (3.5) \\ & \quad + s_{j'}\left(p(Z, \xi; \theta), Z, \xi; \theta\right)\ \frac{d\ \left(p_{j'}\left(Z, \xi; \theta\right) - mc_{j'}\left(Z_{j'}, \xi_{j'}; \theta\right)\right)}{d\ X_{jk}}\Bigg]. \qquad (3.6) \end{aligned}$$

The FOC illustrates that multi-product firms internalize the externality of changing X_j on its other products $j' \in J_f$. The term in (3.5) represents the change in profits attributable to *marginal* consumers while the second term in (3.6) captures those attributable to *infra marginal* consumers. Note that all variables in FOC_{jk} are "observed" in the data and in turn can be evaluated by the econometrician: even though ξ is initially unobserved to econometricians, one can "observe" unique ξ by matching model-predicted shares with data as described in (3.1) at each (θ) following BLP.

Firms also anticipate the change in equilibrium prices that will occur in the second step if they change their product characteristics as these price changes affect demand. This shows up in the FOC in the derivative of shares with respect to characteristics X_j:

$$\frac{d\ s_{j'}\left(p(Z,\xi;\theta),Z,\xi;\theta\right)}{d\ X_{jk}} \quad = \quad \frac{\partial s_{j'}\left(p,Z,\xi;\theta\right)}{\partial X_{jk}} + \sum_{j''\in J} \frac{\partial s_{j'}\left(p,Z,\xi;\theta\right)}{\partial p_{j''}} \frac{\partial p_{j''}(Z,\xi;\theta)}{\partial X_{jk}}$$

The effect of changing X_j on prices of other products of the market is in $\frac{\partial p_{j''}(Z,\xi;\theta)}{\partial X_{jk}}$. This derivative is derived by applying Implicit Function Theorem to (3.3) (See Fan (2013)).

The optimal choice of X_f solves the system of equations in expectation as

$$E\left[FOC_{jk}\left(\theta\right)\middle|I_f\right] \quad = \quad 0, \quad \forall j,k. \tag{3.7}$$

Sometimes firms may over or undershoot the choice of X and p depending on the realization of random variables which are not included in the information set. However, the mistakes will be zero on average at the optimal choice by construction. These JK FOCs with respect to X in (3.7) alongside with J FOCs with respect to p in (3.3) serve as our identification condition.

3.4 Estimation

We define our source of error ν_{jk}, uncorrelated with I_f, in the optimal decision in (3.7) by

$$\nu_{jk}\left(\theta\right) \quad \equiv \quad FOC_{jk}\left(\theta\right)$$

ν_{jk} may include expectational error. For example, expectational errors exists due to firms' incomplete information on own cost shock ω_f and asymmetric information on other firms' characteristic and price choices (X_{-f}, p_{-f}) if the information set is null. Also, ν_{jk} may have approximation (econometrician's expectational error) and measurement errors[4] . Nevertheless, ν_{ij} is zero on average because firms' optimal decisions are made such that

$$E\left[\nu_{jk}\left(\theta\right)\right]=0, \quad \forall j,k \tag{3.8}$$

at Bayes Nash equilibrium. We assume the firm's profit Π_f be the approximation of the true profit function[5] . Potential specification error between the firm's true optimal decision and our approximation may exist, such as adjustment cost in characteristics and other variables that firms take into account when choosing characteristics while they are unobserved by econometricians.

Instruments

Suppose we only knew K FOC conditions (3.8) for each j but we have more than K unknown parameters. To deal with the shortage of moment conditions we use the conditional moment restriction to generate instruments. Conditional moment restriction is given as

$$E\left[\nu_{jk}\left(\theta_0\right)\ \mid I_f\right]=0 \quad \forall k, j\in J_f,\ \forall f \tag{3.9}$$

Although any functions of I_f are possible instruments, Chamberlain (1987) showed that the efficient set of instruments are given by the derivatives of the conditional moment restriction with respect to the parameter evaluated at the truth θ_0:

$$H_{jkl}=E\left[\frac{\partial\nu_{jk}\left(\theta_0\right)}{\partial\theta_l}\ \middle|\ I_f\right]=0 \quad \forall k,l,j\in J_f,\ \forall f. \tag{3.10}$$

Now $L=K\cdot|\theta|$ number of instruments are available in estimation. As Berry et al. (1999) noted, the instruments are intuitive by placing larger weights on the observations which respond sensitively to θ.

There are several challenges in deriving the instruments. I_f has to be specified by a researcher who may not know on which information firms choose characteristics, even

[4] ν is equivalent to ν_1^{π} and ν_1^{r} in Pakes et al. (2015).

[5] i.e. $\nu_2=0$ in PPHI

though the conditional moment restriction itself does not restrict the choice of I_f. The second challenge is that H_{jkl} requires integration over the variables which do not belong to the information set. For example, if I_f only contains the firm's own information, then H_{jkl} requires calculating own best responses in (p, X) at each possible set of competitors' portfolios – prices, characteristics, and cost shocks – and some other underlying shock upon which ν is built, taking derivatives at those best responses, and then integrating the derivatives over the distribution of competitors' values and underlying shocks. This process is computationally demanding and requires us to specify the distribution of the underlying shocks. The last challenge is that the optimal instrument needs to be evaluated at the true θ_0, which is unknown in the beginning of the estimation.

Here, we suggest a way to approximate the optimal instrument H_{jkl} by evaluating the derivative $\frac{\partial \nu_{jk}(\theta_0)}{\partial \theta_l}$ for various I_f. Let $X_t = (X_{jt})_{j \in J_t}$ be a vector of characteristics of all products available in year t, $X_{f,t} = (X_{jt})_{j \in J_{ft}}$ be the set of firm f's products, and $X_{-f,t} = (X_{jt})_{j \notin J_{ft}}$ be the set of firm f's competitors' products. One information set could contain the contemporaneous market condition:

$$I^1_{f,t} \quad = \quad \{X_{-f,\ t},\ \xi_{-f,\ t},\ p_{-f,\ t},\ W_{-f,\ t},\ \omega_t\}.$$

This information set implies that firm f knows competitors' choices of characteristics as well as prices at the time of decision. Therefore, the derivative H_{jkl} is evaluated at realized values of X, ξ, W, p, and ω.

Estimation Procedure

Using the instruments suggested above, the moment condition is given by

$$\begin{aligned} G_{kl}(\theta) &\equiv E\left[H_{jkl}\ \nu_{jk}(\theta)\right] \quad \forall k, l \\ &= 0 \end{aligned} \tag{3.11}$$

We can apply the two step GMM as in BLP. Let $G(\theta)$ be a stacked vector of $G_{kl}(\theta)\ \ k, l$. The GMM estimator is the solution to

$$\min_{\theta} Q(\theta) \quad = \quad G(\theta)'\ \Omega\ G(\theta)$$

where Ω is the optimal weighting matrix. Using the GMM approach allows one to augment with other moment conditions such as matching demographics conditional on purchase. For example, micro moment conditions used in Petrin (2002) and Berry et al. (2004) can be coupled with the moment (3.11).

The computation routine to calculate the GMM objective function $Q(\cdot)$ follows the exact BLP procedure except it uses different moment conditions. Given some initial value at θ_1,

1. Invert out $\delta(\theta_1)$ (and then $\xi(\theta_1)$) by matching market shares between model prediction and data according to equation (3.1).

2. Using equation (3.3), invert out $mc(Z, \xi;\ \theta_1)$.

3. Evaluate FOCs in equation (3.5) at (p, Z, ξ), $\delta(\theta_1)$, and $mc(Z, \xi;\ \theta_1)$.

4. Interact the instruments and FOCs to derive Q.

5. Move on to θ_2 and repeat the Step 1-4 at θ_2.

The only difference between our approach and BLP is in Step 3 where we use different moment conditions.

3.5 Application to BLP Data

We apply our estimation method to the same data used in BLP. There are 2,217 market-level observations on prices, quantities, and characteristics of new U.S. automobiles. Market is defined yearly from 1971 to 1990. Since the instruments for the first year is not available, $J = 2,125$. There are five observed characteristics; constant term, the ratio of horsepower to weight, interior space (length time width), whether air conditioning is standard (a proxy for luxury), and miles per dollar. There are also five cost shifters; unobserved characteristics, log of ratio of horsepower to weight, log of interior space, air conditioning, and log of miles per gallon. We assume that firms set unobserved characteristic (not separate from the constant term), horsepower weight, size, air conditioning, and miles per dollar. Therefore, $K = 5$ characteristics are chosen yielding K sets of J FOCs (3.7). In BLP specification the β parameter on constant is not separately identified from ξ_j. Therefore, $\beta \in \mathbb{R}^{K-1}$ and $\theta \in \mathbb{R}^{3K}$. Parameter θ consists of $\beta \in \mathbb{R}^{K-1}$, $\gamma \in \mathbb{R}^K$, $\sigma \in \mathbb{R}^K$, and $\alpha \in \mathbb{R}$.

To be consistent with the exact utility specification used in BLP, we include income effect:

$$u_{ij}(\theta) \;=\; \alpha\ \ln(y_i - p_j) + \delta_j + \sum_{k=1}^{K} \sigma_k v_{ik} X_{jk} + \varepsilon_{ij}$$

where the product specific utility component δ_j is defined as

$$\delta_j = X_j'\beta + \xi_j.$$

As in BLP we use importance samples to minimize simulation errors. We draw importance samples at an initial estimate θ_1, and then evaluate instruments H and optimal weighting matrix Ω at θ_1 for GMM estimation. Once the first step estimates are converged at θ_2, we re-draw importance samples, re-derive instruments, and re-evaluate optimal weighting matrix at θ_2. Then, we repeat the search over θ.

3.6 Results

Table [3.1] shows the demand and supply estimates, comparing our approach to BLP. The first column has BLP results, using conditional moments (3.2) and (3.4) with BLP instruments, and columns labeled with FOC are derived with conditional moments (3.9) with optimal instruments (3.10). Full IV refers to using the entire $3K^2$ instruments while Part IV refers to using $(K-1)^2 + (K-2)$ subset of instruments [6] .

Even if BLP and our approach (FOC) used the same data set, the results differ with each other in several dimensions. First of all, all estimates following FOC moments are reasonable. In contrast, some BLP estimates have negative signs on mean marginal utility towards miles per dollar (MP$) and cost parameters for size and miles per gallon (MPG); consumers get negative marginal utility from increase in miles per dollar and firms marginal costs decreases as size and miles per gallon improve. However, by using FOC identification, consumers appear to like high miles per dollar because they get positive marginal utility (1.959) from miles per dollar.

Second, our estimates almost always have lower standard errors BLP's, yielding all estimates significant. The t-statistics are four times higher with our approach on average

[6] For X_k except ξ, we selected four instruments. For each k, derivatives in the instrument with respect to $(\alpha, \beta_k, \gamma_k, \sigma_k)$ are selected. For ξ, we selected three, $(\alpha, \gamma_k, \sigma_k)$.

– equivalent effect as if having twice more observations. It is also reasonable to find that the standard errors when using the full set of instruments are lower than those with part of instruments. Improvement in standard errors even allows us to identify the demand and supply only using half of the observations. We report parameter estimates for 1970s and 1980s separately, each of which use 834 and 1,175 observations, respectively, in Table (3.2). Although demand and supply system do not change much over the two decades, the standard errors with half samples are low enough to make all estimates significant. The standard errors with half samples are almost always higher than those with entire samples, and the magnitude of difference is close to $\frac{1}{\sqrt{2}}$.

Third, our estimates yield a different interpretation on how the model fits the data, and in turn, we find different elasticities and markups. On average only 10% of the population buys a new automobile in a given year. BLP estimates replicate this inside share by predicting that consumers are not too price sensitive ($\alpha = 43.501$) but receive negative mean utility of buying a car ($\bar{\xi} = -7.10$) on average compared to not buying one ($\xi_0 = 0$). In contrast, our estimates imply that consumers do receive positive mean utility when buying a car ($\bar{\xi} = 5.43$) but they are very price sensitive ($\alpha = 138.585$), leading to 10% inside share. The high disutility from price cancels out with the positive mean utility so that the predicted level of u_{ij} remains almost the same as in BLP. Therefore, the mean income conditional on purchasing a car in a given year is estimated quite closely – \$216,218 with our estimates whereas BLP predicts \$182,617. We find that demand is more elastic and markups are lower than BLP estimates. Implied price elasticities are reported in Table [3.3]. It is consistent with the fact that our α is higher than BLP's.

Lastly, the marginal cost parameter on unobserved characteristic ξ is positive and significant (0.133). It is intuitive result since products with high quality or favorable attributes tend to be more costly. In addition, cost estimates with FOC sums up to 1.038, implying constant returns to scale in production, even without imposing the constant return assumption.

We test if BLP's conditional moment restrictions (3.2) and (3.4) are valid at our estimates. If the unobserved characteristics were mean independent of observed characteristics, interaction between BLP instruments and unobserved characteristics should be mean zero. However, we find that they are highly correlated. Table [3.4] and [3.5] show regression of ξ and ω on BLP demand and supply instruments[7] , respectively. R-squared suggest that unobserved and observed demand characteristics have 0.79 to 0.86 correlation. Correlation between unobserved and observed cost shifters range from 0.55 to 0.74. All coefficients on BLP instruments are away from zero and significant except one.

One role of instruments is to alleviate correlation between price and unobserved shocks. According to BLP, the instrumented price is supposed to be orthogonal to unobserved characteristic ξ. We also test how well BLP instruments reduce that correlation. We first regressed price on BLP instruments. Then, we regressed the predicted price $\hat{p}(IV_X)$ (i.e. BLP instrumented price) on unobserved demand shock. The results are shown in Table [3.6]. Instrumented price is still highly correlated with unobserved demand shock, with correlation ranging from 0.52 to 0.62[8] . If ξ were linear in price, the instruments does not sufficiently remove correlation between price and unobserved demand shock.

3.7 Conclusions

We exploit optimal choices of product characteristics and prices by firms to infer the distribution of consumer tastes in demand and supply estimation. We assume multi-product firms to choose prices and product characteristics, including both observed and unobserved, in Bayes Nash Equilibrium set up. The identification is achieved without assuming that observed and unobserved characteristics are uncorrelated, which has typically been used in the literature. Using our approach, we find much more precise and reasonable estimates of taste parameters given the same market-level data. We also find strong evidence that characteristics are highly correlated with each other, rejecting the traditional identification assumption.

---

[7] BLP demand instruments are own product characteristic X_{jk}, $\forall k$, sum of characteristic across own-firm products $\sum_{j' \neq j, j' \in J_f} X_{j'}$, and sum of all characteristics across competing firms, $\sum_{j' \notin J_f} X_{j'}$. BLP supply instruments are achieved in the same way by replacing X with W.

[8] Since ξ is not linear in price, instrumented price may still be correlated with ξ. Therefore, this test does not necessarily imply that BLP instruments are invalid

3.8 Tables

Table 3.1: Estimated Parameters of the Demand and Supply

	Variable	BLP	FOC	
			Full IV	Part IV
Means (β's)	Constant	-7.061		
		(0.941)		
	HP/weight	2.883	1.436	2.237
		(2.019)	(0.140)	(0.369)
	Size	3.46	0.379	0.427
		(0.61)	(0.039)	(0.112)
	Air	1.521	1.452	0.093
		(0.891)	(0.196)	(0.181)
	MP$	-0.122	1.959	1.933
		(0.32)	(0.202)	(0.412)
Std. Deviations (σ_k's)	Constant	3.612	2.213	2.185
		(1.485)	(0.515)	(2.946)
	HP/weight	4.628	2.369	2.620
		(1.885)	(0.241)	(1.099)
	Size	2.056	0.613	0.218
		(0.585)	(0.084)	(0.534)
	Air	1.818	1.622	1.641
		(1.695)	(0.158)	(0.290)
	MP$	1.050	0.666	0.857
		(0.272)	(0.148)	(0.368)
Term on Price α	$ln(y-p)$	43.501	138.585	157.151
		(6.427)	(17.144)	(40.113)
Cost Side Parameters	Constant or	0.952	0.133	0.117
	Mean Charac (ξ)	(0.194)	(0.016)	(0.031)
	ln(HP/Weight)	0.477	0.074	0.098
		(0.056)	(0.010)	(0.027)
	ln(Size)	-0.046	0.081	0.064
		(0.081)	(0.010)	(0.018)
	Air	0.619	0.198	0.020
		(0.038)	(0.027)	(0.019)
	ln(MPG)	-0.415	0.552	0.465
		(0.055)	(0.058)	(0.098)
Mean Unobserved	$\bar{\xi}$	-7.10	5.43	7.44

Table 3.2: Estimated Parameters, Only Using Half Samples

	Variable	FOC	
		70s	80s
Means (β's')	HP/weight	1.378	1.214
		(0.186)	(0.137)
	Size	0.14	0.192
		(0.018)	(0.022)
	Air	1.559	1.6
		(0.218)	(0.273)
	MP$	2.959	2.122
		(0.338)	(0.240)
Std. Deviations (σ_k's')	Constant	2.21	3.194
		(0.432)	(0.503)
	HP/weight	4.259	2.651
		(0.386)	(0.227)
	Size	0.466	0.598
		(0.044)	(0.053)
	Air	1.68	1.521
		(0.269)	(0.130)
	MP$	0.762	0.709
		(0.154)	(0.143)
Term on Price α	$ln(y-p)$	145.816	160.527
		(17.250)	(20.263)
Cost Side Parameters	Constant	0.121	0.119
		(0.014)	(0.014)
	ln(HP/weight)	0.07	0.061
		(0.011)	(0.008)
	ln(Size)	0.03	0.035
		(0.004)	(0.005)
	Air	0.201	0.21
		(0.034)	(0.029)
	ln(MPG)	0.54	0.586
		(0.067)	(0.066)
Mean Unobserved	$\bar{\xi}$	2.06	7.33

Table 3.3: Implied Elasticities and Markups

	Elasticities		Markups ($)	
	BLP	FOC	BLP	FOC
Lexus LS400	-3.027	-5.106	9214.54	5453.83
Lincoln Towncar	-3.030	-5.699	8310.82	4596.23
Nissan Maxima	-4.124	-7.925	3385.84	1765.61
Ford Taurus	-3.952	-8.255	2679.14	1344.04
Chevy Cavalier	-5.899	-9.414	1327.75	787.13
Nissan Sentra	-6.304	-9.644	909.79	594.19
Mean	-4.087	-7.590	4051.87	2313.05
Median	-3.975	-8.125	2751.77	1404.51
Std. Deviation	1.120	1.810	3905.32	2491.31

Table 3.4: $E\left[\xi_j \mid X\right] \neq 0$

ξ	Full IV		Part IV	
Constant	-0.933	-8.276	0.194	-8.622
	(0.535)	(1.071)	(0.600)	(1.188)
HP/weight	8.540	6.193	8.956	6.185
	(0.534)	(0.585)	(0.599)	(0.649)
Size	5.397	4.025	5.787	4.295
	(0.265)	(0.320)	(0.297)	(0.355)
Air	1.618	1.207	3.401	2.915
	(0.122)	(0.130)	(0.137)	(0.144)
MP$	-2.142	-3.335	-2.248	-3.647
	(0.090)	(0.136)	(0.101)	(0.151)
Other BLP instruments	No	Yes	No	Yes
R-squared	0.623	0.705	0.651	0.734

Table 3.5: $E\left[\omega_j \mid W\right] \neq 0$

ω	Full IV		Part IV	
Constant	-0.231	1.671	-0.200	1.797
	(0.126)	(0.157)	(0.122)	(0.153)
ln(HP/Weight)	0.224	0.053	0.230	0.070
	(0.032)	(0.031)	(0.031)	(0.030)
ln(Size)	-0.706	-0.425	-0.610	-0.369
	(0.059)	(0.058)	(0.057)	(0.056)
Air	0.331	0.135	0.326	0.134
	(0.016)	(0.015)	(0.015)	(0.015)
ln(MPG)	0.090	-0.208	0.141	-0.161
	(0.040)	(0.042)	(0.039)	(0.041)
Other BLP instruments	No	Yes	No	Yes
R-squared	0.307	0.546	0.306	0.544

Table 3.6: Correlation Between Instrumented price ($\hat{p}(IV_X)$) and ξ

$\hat{p}(IV_X)$	Full IV	Part IV
Constant	6.551	4.550
	(0.219)	(0.227)
ξ	0.959	0.969
	(0.033)	(0.026)
R-squared	0.273	0.380

$\hat{p}(IV_X)$ is predicted price on BLP demand instruments.

References

a Pakes, J Porter, Kate Ho, and Joy Ishii, "Moment Inequalities and Their Application," *Econometrica*, 2015, *83* (1), 315–334.

Ansolabehere, Stephen and Stephen Pettigrew, "Cumulative CCES Common Content (2006-2012)," 2014.

Arellano, Manuel and Costas Meghir, "Female Labour Supply and On-the-Job Search: An Empirical Model Estimated Using Complementary Data Sets," *The Review of Economic Studies*, 1992, *59* (3), 537–559.

Arentze, Theo, Harmen Oppewal, and Harry J.P. Timmermans, "A Multipurpose Shopping Trip Model to Assess Retail Agglomeration Effects," *Journal of Marketing Research*, 2005, *42* (1), 109–115.

Basker, Emek, Shawn Klimek, and Pham Hoang Van, "Supersize It: The Growth of Retail Chains and the Rise of the "Big-Box" Store," *Journal of Economics and Management Strategy*, 2012, *21* (3), 541–582.

Berger, Allen N., David B. Humphrey, and Lawrence B. Pulley, "Do consumers pay for one-stop banking? Evidence from an alternative revenue function," *Journal of Banking and Finance*, 1996, *20* (9), 1601–1621.

Berry, Steven, "Estimating Discrete Choice Models of Product Differentiation," *The RAND Journal of Economics*, 1994, *25* (2), 242–262.

_ , James Levinsohn, and Ariel Pakes, "Automobile Prices in Market Equilibrium," *Econometrica*, 1995, *63* (4), 841–890.

— , — , and — , “Voluntary export restraints on automobiles: Evaluating a trade policy,” *American Economic Review*, 1999, *89* (3), 400–430.

— , — , and — , “Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market,” *Journal of Political Economy*, 2004, *112* (1), 68–105.

Betancourt, Roger and David Gautschi, “Demand Complementarities, Household Production, and Retail Assortments,” *Marketing Science*, 1990, *9* (2), 146–161.

Blonigen, Bruce A, Christopher R Knittel, and Anson Soderbery, “Keeping it Fresh: Strategic Product Redesigns and Welfare,” 2013.

Bollinger, Bryan and Kenneth Gillingham, “Peer Effects in the Diffusion of Solar Photovoltaic Panels,” 2012.

Center for Sustainable Energy, “California Plug-in Electric Vehicle Driver Survey Results May 2013,” Technical Report May 2013.

— , “Clean Vehicle Rebate Project Fiscal Year 2012-2013 Final Report,” Technical Report 2013.

— , “California Air Resources Board Clean Vehicle Rebate Project, Rebate Statistics,” 2014.

Chamberlain, Gary, “Asymptotic efficiency in estimation with conditional moment restrictions,” *Journal of Econometrics*, 1987, *34* (3), 305–334.

Clements, MT and Hiroshi Ohashi, “Indirect Network Effects and the Product Cycle: Video Games in the US, 1994 - 2002,” *The Journal of Industrial Economics*, 2005, *53* (4), 515–542.

Costa, Dora L. and Matthew E. Kahn, “Do Liberal Home Owners Consume Less Electricity? A Test of the Voluntary Restraint Hypothesis,” *Economics Letters*, May 2013, *119* (2), 210–212.

Crawford, Gregory S., “Endogenous product choice: A progress report,” *International Journal of Industrial Organization*, 2012, *30* (3), 315–320.

Crawford, GregoryS. and Matthew Shum, “Monopoly Quality Degradation and Regulation in Cable Television,” *The Journal of Law and Economics*, 2007, *50* (1), 181–219.

Cummins, J. David, Mary a. Weiss, Xiaoying Xie, and Hongmin Zi, “Economies of scope in financial services: A DEA efficiency analysis of the US insurance industry,” *Journal of Banking & Finance*, 2010, *34* (7), 1525–1539.

Dubé, Jean-Pierre H., Günter J. Hitsch, and Pradeep K. Chintagunta, “Tipping and Concentration in Markets with Indirect Network Effects,” *Marketing Science*, March 2010, *29* (2), 216–249.

Eizenberg, Alon, “Upstream Innovation and Product Variety in the United States Home PC Market,” *The Review of Economic Studies*, 2014, *81*, 1003–1045.

Ellickson, Paul B., “The Evolution of the Supermarket Industry: From A&P to Wal-Mart,” *SSRN Electronic Journal*, 2011, (April), 1–20.

EV Project, “EV Project Electric Vehicle Charging Infrastructure Summary Report,” Technical Report 2014.

Fan, Ying, “Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market,” *American Economic Review*, 2013, *103* (5), 1598–1628.

Gandal, Neil, Michael Kende, and Rafael Rob, “The Dynamics of Technological Adoption in Hardware / Software Systems: The Case of Compact Disc Players,” *The RAND Journal of Economics*, 2000, *31* (1), 43–61.

Gentzkow, Matthew, “Valuing new goods in a model with complementarity: Online newspapers,” *American Economic Review*, 2007, *97* (3), 713–744.

Goolsbee, Aust and PJ Klenow, “Evidence on Learning and Network Externalities in the Diffusion of Home Computers,” *Journal of Law and Economics*, 2002, *XLV* (October 2002).

Goolsbee, Austan and Amil Petrin, “The Consumer Gains From Direct Broadcast Satellites and The Competition with Cable TV,” *Econometrica*, 2004, *72* (2), 351–381.

Gowrisankaran, Gautam, Minsoo Park, and Marc Rysman, "Measuring Network Effects in a Dynamic," 2014.

Gramlich, Jacob, "Gas Prices, Fuel Efficiency, and Endogenous Product Choice in the U.S. Automobile Industry." PhD dissertation, Georgetown University 2009.

Hansen, Lars Peter, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 1982, *50* (4), 1029–1054.

Hausman, Jerry, "Valuation of new goods under perfect and imperfect competition," in "The economics of new goods," University of Chicago Press, 1996, pp. 207–248.

Holder, H D and a C Wagenaar, "Effects of the elimination of a state monopoly on distilled spirits' retail sales: a time-series analysis of Iowa.," *British journal of addiction*, 1990, *85* (12), 1615–1625.

Holmes, Thomas J, "Bar codes lead to frequent deliveries and superstores.," *RAND Journal of Economics (RAND Journal of Economics)*, 2001, *32* (4), 708–725.

Imbens, Guido and Tony Lancaster, "Combining Micro and Macro Data in Microeconometric Models," *The Review of Economic Studies*, 1994, *61* (4), 655–680.

Kahn, Matthew E., "Do Greens Drive Hummers or Hybrids? Environmental Ideology as a Determinant of Consumer Choice," *Journal of Environmental Economics and Management*, September 2007, *54* (2), 129–145.

LaVallee, Robin, Trinh Kim, and Hsiao ye Yi, "Apparent Per Capita Alcohol Consumption: National, State, and Regional Trends, 19772012," Technical Report, U.S. Department of Health and Human Services 2014.

Lazich, Ed. Robert S. and Virgil L. III. Burton, *Market Share Reporter*, 2 ed. 2014.

Lustig, Joshua, "The Welfare Effects of Adverse Selection in Privatized Medicare." PhD dissertation, UC Berkeley 2008.

Mazzeo, Michael J, "Product choice and oligopoly market structure.," *RAND Journal of Economics*, 2002, *33* (2), 221–242.

Messinger, Paul R and Chakravarthi Narasimhan, "Model of Retail Formats Based Shopping on Consumers ' Economizing on Time," 1997, *16* (1), 1–23.

Miravete, Eugenio J, Katja Seim, and Jeff Thurk, "Complexity , E ciency , and Fairness of Multi-Product Monopoly Pricing," 2014.

Nair, Harikesh, Pradeep K. Chintagunta, and Jean pierre Dube, "Empirical Analysis of Indirect Network Effects in the Market for Personal Digital Assistants," *Quantitative Marketing and Economics*, 2004, *2*, 23–58.

Petrin, Amil, "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy*, 2002, (110), 705–729.

_ **and Boyoung Seo**, "Identification and Estimation of Discrete Choice Demand Models When Observed and Unobserved Product Characteristics Are Correlated," 2014.

Rees, Rebecca, "Privatization of Liquor Stores : No Threat to Public Safety," 1997, *17110* (August).

Sacks, Jeffrey J., Jim Roeber, Ellen E. Bouchery, Katherine Gonzales, Frank J. Chaloupka, and Robert D. Brewer, "State costs of excessive alcohol consumption, 2006," *American Journal of Preventive Medicine*, 2013, *45* (4), 474–485.

Seim, Katja and Joel Waldfogel, "Public monopoly and economic efficiency: Evidence from the pennsylvania liquor control board's entry decisions," *American Economic Review*, 2013, *103* (2), 831–862.

Selway, Janet S., "Alcohol Screening and Brief Intervention," *The Journal for Nurse Practitioners*, 2006, *2* (2), 90–96.

Sen, Boudhayan, Jiwoong Shin, and K Sudhir, "Demand Externalities from Co-Location: Evidence from a Natural Experiment," 2013.

Siegel, Michael, William Dejong, Alison B. Albers, Timothy S. Naimi, and David H. Jernigan, "Differences in liquor prices between control state-operated and license-state retail outlets in the United States," *Addiction*, 2013, *108* (2), 339–347.

Smith, Howard, “Supermarket choice and supermarket competition in market equilibrium,” *Review of Economic Studies*, 2004, *71* (1), 235–263.

Spence, Michael, “Product Selection, Fixed Costs, and Monopolistic Competition,” *The Review of Economic Studies*, 1976, *43* (2), 217–235.

Sweeting, Andrew, “Dynamic Product Repositioning in Differentiated Product Markets: The Case of Format Switching in the Commercial Radio Industry,” 2007.

Train, Kenneth E., *Discrete Choice Methods with Simulation*, Vol. 2 2009.

Trajtenberg, Manuel, “The Welfare Analysis of Product Innovations, with an Application to Computed Tomography Scanners,” *Journal of Political Economy*, 1989, *94*, 444–479.

Veiga, André and E. Glen Weyl, “Product Design in Selection Markets,” 2014.

Wakamori, Naoki, “Portfolio Considerations in Automobile Purchases : An Application to the Japanese Market,” *Dissertation*, 2015, *49* (499).

Yuan, Y and Rd Phillips, “Financial integration and scope efficiency in US financial services post Gramm-Leach-Bliley,” *Journal of Banking and Finance*, 2008.

Zullo, Roland, Xi Belinda Bi, and Yu Sean Xiaohan, “The Fiscal and Social Effects of State Alcohol Control Systems May 2013,” 2013.

Appendix A

Appendix to Chapter 1

A.1 Proof of Contraction Mapping

Claim: f defined in (1.4) is a contraction mapping with modulus less than 1.

Proof: This proof uses Theorem in Appendix I from Berry et al. (1995) and follows their proof. They proved that if the following assumptions on f are satisfied, then f is a contraction mapping with modulus less than 1. For simplicity, market t is suppressed.
(0) $f : R^{J^G+J^L} \to R^{J^G+J^L}$, $d(\boldsymbol{\delta}, \boldsymbol{\delta}') = \left\|\boldsymbol{\delta} - \boldsymbol{\delta}'\right\|$ where $\|\cdot\|$ is the sup-norm in Euclidean space.
(1) $\forall \boldsymbol{\delta} \in R^{J^G+J^L}$, $f(\boldsymbol{\delta})$ is continuously differentiable, and $\forall n, m$,

$$\begin{aligned} \frac{\partial f(\boldsymbol{\delta})_n}{\partial \delta_m} &\geq 0 \\ \sum_{m=1}^{J^G+J^L} \frac{\partial f(\boldsymbol{\delta})_n}{\partial \delta_m} &< 1. \end{aligned}$$

(2) $\min_n \inf_{\boldsymbol{\delta}} f(\boldsymbol{\delta}) > -\infty$
(3) There is a value, $\bar{\boldsymbol{\delta}}$ such that if for any $\boldsymbol{\delta}$ where $\delta_n > \bar{\delta}_n$, then there exists some m such that $f(\boldsymbol{\delta})_m < \delta_m$.

The rest of the proof is to show that f defined in (1.4) satisfies the assumptions above with respect to $\boldsymbol{\delta}$. If $1 \leq n \leq J^G$, then n represents the grocery sections at store and $\delta_n = \delta^G_{g(n)}$ where $g(n)$ indexes the nth grocery store. $g(0)$ refers to no grocery shopping. If $J^G < n \leq J^G + J^L$, $\delta_n = \delta^L_{\ell(n)}$ where $\ell(n)$ indexes the $n - J^G$th liquor store. $\ell(0)$ indexes no liquor shopping. First of all, $s(\boldsymbol{\delta}^*; \theta)$, which is defined in (1.3), is continuously differentiable with respect to $\boldsymbol{\delta}$. Therefore f is, too.

Second, the marginal share of groceries at the nth grocery store is derived by summing the share of trips with grocery purchases at store $g(n)$ across all liquor store:

$$s_n^G = \sum_{m=0}^{J^L} s_{(g(n),\ell(m))}$$

If $1 \leq n \leq J^G$, the own derivative of f:

$$\begin{aligned}
\frac{\partial f(\boldsymbol{\delta})_n}{\partial \delta_n} &= \lambda - \frac{1}{s_{g(n)}^G} \frac{\partial s_{g(n)}^G}{\partial \delta_n} \\
\frac{\partial s_{g(n)}^G}{\partial \delta_n} &= \sum_{m=0}^{J^L} \frac{\partial s_{(g(n),\ell(m))}}{\partial \delta_n} \\
&= \sum_{m=0}^{J^L} \left(\lambda s_{(g(n),\ell(m))} - \lambda s_{(g(n),\ell(m))} \sum_{w=0}^{J^L} s_{(g(n),\ell(w))} \right) \\
&= \lambda s_{g(n)}^G \left(1 - s_{g(n)}^G\right) \\
\therefore \frac{\partial f(\boldsymbol{\delta})_n}{\partial \delta_n} &= \lambda s_{g(n)}^G \\
&\geq 0.
\end{aligned}$$

If $n > J^G$, own derivative changes to

$$\begin{aligned}
\frac{\partial f(\boldsymbol{\delta})_n}{\partial \delta_n} &= (1-\lambda) s_{\ell(n)}^L \\
&\geq 0.
\end{aligned}$$

Cross derivative of f given $1 \leq n \leq J^G$ and $1 \leq m \leq J^G$ is

$$\begin{aligned}
\frac{\partial f(\boldsymbol{\delta})_n}{\partial \delta_m} &= -\frac{1}{s_{g(n)}^G} \frac{\partial s_{g(n)}^G}{\delta_{g(m)}} \\
\frac{\partial s_{g(n)}^G}{\delta_{\ell(m)}} &= \lambda s_{g(n)}^G s_{g(m)}^G \\
\therefore \frac{\partial f(\boldsymbol{\delta})_n}{\partial \delta_m} &= \lambda s_{g(m)}^G \\
&\geq 0.
\end{aligned}$$

Likewise, for $m > J^G$,

$$\begin{aligned}\frac{\partial f(\boldsymbol{\delta})_n}{\partial \delta_m} &= (1-\lambda)\, s^L_{\ell(m)} \\ &\geq 0.\end{aligned}$$

Sum of the derivatives with respect to any $n = 1, ..., J^G + J^L$ is

$$\begin{aligned}\sum_{m=1}^{J^G+J^L} \frac{\partial f(\boldsymbol{\delta})_n}{\partial \delta_m} &= \sum_{m=1}^{J^G} \frac{\partial f(\boldsymbol{\delta})_n}{\partial \delta_m} + \sum_{m=J^G+1}^{J^L} \frac{\partial f(\boldsymbol{\delta})_n}{\partial \delta_m} \\ &= \sum_{m=1}^{J^G} \lambda s^G_{g(m)} + \sum_{m=J^G+1}^{J^L} (1-\lambda)\, s^L_{\ell(m)} \\ &= \lambda\left(1 - s^G_{g(0)}\right) + (1-\lambda)\left(1 - s^L_{\ell(0)}\right) \\ &< 1.\end{aligned}$$

This guarantees that the contraction mapping has modulus less than 1. Formulating δ_j in (1.1) as a sum of $\lambda\delta^G_g$ and $(1-\lambda)\,\delta^L_\ell$ ensures the sum of inside share of each good is less than 1.

Next step is to show that f is bounded from below. f_n for $1 \leq n \leq J^G$ can be rewritten by

$$\begin{aligned}f(\boldsymbol{\delta})_n &= log\left(\text{share}^{\text{data},G}_n\right) - log\left(D_n(\boldsymbol{\delta})\right) \\ D_n(\boldsymbol{\delta}) &= \sum_{m=0}^{J^L} \int_i \frac{exp\left(\mu_{i(g(n),\ell(m))} + \Gamma_{(n),\ell(m))} + (1-\lambda)\,\delta^L_{\ell(m)}\right)}{1 + \sum_{(g,l)\in J^G\times J^L\setminus\{(0,0)\}} exp\left(\mu_{i(g,l)} + \Gamma_{(g,l)} + \lambda\delta^G_g + (1-\lambda)\,\delta^L_l\right)} di \\ &\underset{\delta\downarrow-\infty}{\rightarrow} 0\end{aligned}$$

Therefore, the lower bound of f is $\underline{\delta}_n = log\left(\text{share}^{\text{data},G}_n\right)$. If $n > J^G$, then the D_n is defined the same as above except $(1-\lambda)$ in the numerator is substituted by λ.

The last step is to find the upper bound of f. Following Berry (1994), set $\delta_m = -\infty, \forall m \neq n, 0$ for $1 \leq n \leq J^G$ and let $\bar{\delta}_n$ be such that

$$\begin{aligned}\text{share}^{\text{data},G}_0 &= s^G_0(\boldsymbol{\delta}^*; \theta) \text{ where } \delta^*_{g(n)} = \lambda\bar{\delta}_n \\ &= 1 - \int_i \frac{exp\left(\mu_{i(g(n),0)} + \Gamma_{i(g(n),0)} + \lambda\bar{\delta}_n\right)}{1 + exp\left(\mu_{i(g(n),0)} + \Gamma_{i(g(n),0)} + \lambda\bar{\delta}_n\right)} di\end{aligned}$$

$\bar{\delta}_n$ is the implied linear utility when the choice set has one element, $(n, 0)$. Likewise, if $n > J^G$, $\bar{\delta}_n$ is defined the same way as above except replacing λ with $1-\lambda$ in front of $\bar{\delta}_n$. Let the maximum as the upper bound of f be $\bar{\boldsymbol{\delta}} = \max_n \bar{\delta}_n$. Then, w.l.o.g. $1 \leq n \leq J^G$, if a $\boldsymbol{\delta}$ with $\delta_n > \bar{\delta}_n$, then $\text{share}_0^{\text{data},G} > s_0^G(\boldsymbol{\delta}^*; \theta)$. Then, the inside share implied by $\boldsymbol{\delta}$ should be larger than the inside share observed in data, which implies that there exists at least one m, such that $\text{share}_{g(m)}^{\text{data},G} < s_{g(m)}^G(\boldsymbol{\delta}^*; \theta)$ and, therefore, $f(\boldsymbol{\delta})_m < \delta_m$. □

A.2 Estimating Travel Distance

The distances between stores and panelists in the Nielsen Consumer Panel dataset are derived by panelists' home zip code and stores' estimated location based on the grocery or liquor store address from the Washington State Liquor and Cannabis Board (WALCB). The Nielsen Consumer Panel does not reveal the exact stores that panelists visited. Instead, the channel type (grocery store, warehouse club, discount store, drug store, etc.) is provided. If the store is affiliated with the Nielsen Retail Scanner, the first three digits of the zip code, county are also recorded. Based on the given store information, I narrow down the possible set of grocery or liquor stores in the WALCB licensee list by limiting the approximate location and the store types. Among the candidate stores in the WALCB licensee list, I assign the store closest to the panelist's home zip code as the store where the trip is occurred.

The straight line distances between stores and panelists' home zip codes are derived by following. I derive the longitude and latitude of the estimated store based on the address provided by the WALCB data, by using ArcGIS API. Since zip code includes areas in which no one reside, geographical centroid of a zip code does not represent the area in which most households reside. To enhance accuracy, I first measure distance between a store and each centroid of census block which belongs to a zip code. Then, I aggregate those distances by weighting population of census block to derive distance between a store and a zip code. Weighting by housing units, population, 21 or above, 18 or above, or 15 or above barely changes the distance measure.

The 95% quantile of the estimated distance of grocery trip is 8.7 miles and 7.97 miles for liquor trips, and 8.71 miles for grocery and liquor trips. Some shopping trips were made in different regions of the state, implying that the panelist was traveling.

Those cases are eliminated from analysis by cutting off trips with greater than the 99% quantile of trip distances.

A.3 Supplementary Description on Data

Household Panel Data

The panelists participate for more than 3 consecutive years on average. Table (A.1) shows that about two-thirds of the panelists do not purchase liquor in a year. This does not imply that those panelists do not drink liquor; they can consume liquor at bars or restaurants. See Appendix (A.6) for more details. The number of panelists who purchase liquor from stores for off-premises consumption has increased since 2011. The table also shows that the panelists purchase groceries three times a week. The frequency of liquor purchases is increased from 5.2 to 6.2 in 2013, implying that the fixed costs of going liquor shopping is reduced.

The panelists are selected so that they represent the population in Washington in terms of demographic composition and total sales of consumption goods. Each panelist is given a sampling weight with which the projected purchases on consumption goods match the population sales. Even though the panelists are not selected based on their liquor consumption, their reported liquor purchases consistently represent the total liquor sales for off-premises consumption in Washington. Figure (A.1) depicts that the projected liquor purchases by the panelists covary with the total liquor quantity sold accordingly each month, both before and after privatization. Moreover, the stores visited for liquor shopping by the panelists take up 40% of the liquor stores, which consist of 60% of the market share of liquor. The dataset is also geographically representative of Washington. Out of 39 counties in Washington, the panelists reside in 35 counties. The uncovered counties are rural areas where the population is less than 21,000.

Price Indices for Groceries and Liquor

Based on the Nielsen Consumer Panel data, I find that consumers purchase 11 grocery products per grocery shopping trip in median. Table (A.2) describes the average price and unit size of each of the 11 product groups. On the other hand, consumers purchase one liquor product, or equivalently 1.75 liters of liquor, per liquor shopping trip in median.

A.4 Definition of Shopping Trips and Grocery and Liquor Products

A *store visit* is defined as a physical travel to a grocery, discount, warehouse club, hypermarket, close out, military, liquor, beverage, drug, convenience, tobacco, gas mini mart, or dollar store, and etc. and exclude home delivery, online or mail order, vending machine, TV home shopping, and free samples or gifts. Moreover, the purchases have to include at least one grocery product or one liquor product.

A *shopping trip* is one or two store visits occurred on the same day or in the same week. For simplicity, I assume that there are four types of shopping trips, grocery-shopping-only, liquor -shopping-only, grocery and liquor shopping at one store, and groceries and liquor at two separate stores.

Grocery product can be either food or non-food. Food grocery products include beer, wine, dairy, deli, packaged meat, dry groceries (canned food and drinks, baking supplies, seasoning, grains, snacks, bread, etc.), fresh produce, and frozen food. Non-food grocery products consists of health and beauty care products (toothpaste, shampoo, razor, pain killer, etc.) and household supplies (laundry detergent, toilet paper, cigarettes, etc.).

Liquor product is defined as beverage which contains alcohol obtained by distillation (RCW 66.04.010 (4)) by the Washington State Liquor and Cannabis Board (WALCB). The WALCB uses the term spirits instead of liquor, though. It excludes flavored malt beverages but may include wines exceeding twenty four percent of alcohol by volume. Most fortified wines do not qualify as liquor. For example, Port, Sherry, Madeira, etc. have less than twenty one percent of alcohol by volume. The Nielsen's classification of liquor is different from the WALCB's definition of liquor. I merged brand information from the Mintel Global New Products Database and alcohol contents information from the Nielsen data to select liquor products, which satisfies the definition set by the WALCB, out of all products in the Nielsen's liquor category.

A *grocery-shopping-only trip* is a *store visit* which includes purchases of groceries but not liquor products. For example, a trip to Walmart (discount store) purchasing a diaper (non food grocery), a trip to Safeway (grocery store) purchasing milk (food grocery), and a trip to Total Wine and More (liquor store) purchasing wine (food grocery) are all considered as store visits for groceries.

A *liquor-shopping-only trip* is a *store visit* which includes purchases of liquor but not grocery products. However, if the trip includes non-liquor purchases and all of them are closely related to liquor, then the trip is treated as *liquor-shopping-only trip*. Liquor-related products include tobacco, beer, wine, carbonated beverages, glassware, ice, and so on. For example, if a trip consists of one bottle of liquor, a can of soda, and a bag of ice, it is classified as a *liquor-shopping-only trip*. If the trip further includes a pack of cigarette and a shot glass, it is still considered as a liquor only trip. However, if the same trip includes milk, lettuce, or pet food, then it is not considered as a liquor trip, rather it is a *grocery and liquor shopping trip*.

A *grocery and liquor shopping trip* (either at one store or two stores) is one or ore store visits with purchases which consist of both grocery and liquor products. For example, a trip to a convenience store to purchase milk, eggs, cigarettes, and liquor belongs to this type of shopping trip.

The rest of the store visits are dropped from the analysis because of irrelevance. These store visits either occurred in one of the non-grocery stores, such as apparel, computer, office supplies, toy, and home furnishing stores, etc., or their purchases only consist of non-grocery products, such as hardware, electronics, insecticides, automotive, stationery, gardening, and housewares, etc. For instance, purchasing a vacuum cleaner (non-grocery product) at Costco (warehouse club) or buying chocolates and chips at a toy store does not belong to either groceries only, liquor only, or groceries and liquor trip. The restriction eliminates about 25% of the total store visits reported in the Nielsen Consumer Panel dataset, resulting in approximately 200,000 trip observation each year.

A.5 Supplementary Description on Market Outcomes

Background on Liquor Regulation

Since the end of Prohibition each state has determined its own set of liquor regulations. Some states, called license states, decided to adopt a licensing system, allowing private business entities to sell liquor. In contrast, sixteen states, called control states, enacted rules to fully control liquor sales by owning liquor stores[1] . These state-run monopolies were appealing since the state could collect taxes effectively and reduce alcohol consumption. Indeed the policies effectively met their goals; it is well documented that liquor consumption in control states have been lower than in license states as their prices are maintained higher on average (Zullo et al. (2013), LaVallee et al. (2014), and Siegel et al. (2013)). However, these incentives have weakened in modern times because the focus of the control states has shifted to law enforcement. Moreover, policy makers acknowledged that private retailers can provide the same service as efficiently, if not more efficiently than the state[2] . Since the repeal, three states – Iowa in 1987 (Holder and Wagenaar (1990)), West Virginia in 1990 (Rees (1997)), and Washington in 2012 – discontinued their state monopolies. Iowa and Washington adopted licensing system whereas West Virginia introduced auction system for limited number of liquor store permits.

Change in the Number of Liquor Stores Across Areas

Table (A.3) tabulates the average number of stores within a 5 mile radius of a zip code by different characteristics of zip code, such as population density, median income level, and race. For each zip code, I count the stores if the average distance between the store and census blocks weighted by census block population within a zip code is less than 5 miles. Each demographic bins are cut off by 25th, 50th, and 75th quantile. First, there is no clear area where the number of stores have decreased. There are only few areas where the former state stores exited and no private stores started selling liquor. Second, the increase of the number of stores is not uniformly distributed across the state. In the areas with high population density, high income, or low ratio of Caucasian population have the more new stores compare to the areas the opposite. This uneven distribution of the number of stores leads to the fact that the uneven distribution of distance of liquor shopping trips across demographics and areas as mentioned in section (2) of the paper.

[1] "Beverage Alcohol Control Agency Info Sheet", National Alcohol Beverage Control Association.
[2] "Modernizing Washington's Liquor Control System", Washington Policy Center, February 2003.

Increase in Price

Figure (A.2) describes changes in the average post-tax price weighted by quantity sold in liter, based on revenue and liter sold data from the Washington State Department of Revenue. The red line represents price a year prior to privatization while the blue solid line indicates price a year after, and the blue dotted line indicates two years after. Controlling for seasonality the average consumer price after privatization has stayed above the pre-privatization price, and it did not fall back to the pre-privatization price.

Price increase in Washington cannot be explained by national shock nor influx of high quality brands. Table (A.4) shows the results of regression of price in Washington on national price, indicator for post-privatization time, brand fixed effects[3] , and year and week fixed effects. Observations are in week-UPC pair level from 2011 to 2013. "Privatization" is a dummy variable indicating the time periods from June 2012. Its coefficient represents the average price difference before and after privatization conditional on time trend, brands, and national price. The estimated change in price is still positive and significant. In fact, the estimated difference in price \$1.96 is only slightly less than the realized increase in price, which is \$2.32. This regression suggests that national supply or demand shocks, time trend, and brand quality cannot justify the increase in post-tax price, rather it is likely because of the extra tax imposed to the retailers as part of deregulation.

The price increases also varied across demographics and population density. Table (A.5) specifies percentage increase in price across zip codes, tabulated by demographics. Rural areas experience the largest price increase, and the increment becomes smaller in more urban environments. The change in price by the median income of zip codes is not linear. The ratio of Caucasian relative to other races is positively correlated with price increase. The uneven distribution of the price increase across areas and demographics also affects the uneven distribution of the gains from deregulation.

[3] Nielsen defines a brand as a product line. For example, Bacardi, Bacardi Dragon Berry, Bacardi Select are all defined as different brands. I created brand variable (e.g. Bacardi for all three) which is more general than Nielsen's definition by combining product names from Nielsen and company names from the Mintel Global New Products Database.

Increase in Quantity Sold

Figure (A.3) shows liquor quantity sold in million liters from June 2011 to May 2014, controlling seasonality. The overall increase in liquor quantity sold is 6%. Except in June 2012, post-privatization liquor sales (blue solid line) have consistently been higher month on month than pre-privatization (red solid line). The quantity remains higher than before even after two years of privatization (blue dotted line). The increase in liquor sales in Washington, which is 6%, is too large to be explained by national demand shocks. The national ethanol consumption from liquor per capita increased by 2.63% and 3.85% in West region in 2012 (LaVallee et al. (2014)). In contrast, per capita ethanol consumption in Washington increased by 4.89%.

Improvement of the Choice Set for Shopping Trips

Unless demand is upward sloping, increase in quantity despite of the increase in consumer price provides evidence that the choice set for shopping trips has improved. There is a list of suggestive evidence that proximity to stores and the new choice of one stop shopping are the two important changes which led to increase in quantity sold. First of all, increase in quantity is originated from more frequent store visits, rather than the amount purchased per visit. Table (A.1) shows that consumers purchased liquor about 5.2 times in 2012 conditional on purchasing liquor in the year. In contrast, consumers visited liquor store about 6.2 times in 2013. Meanwhile, the quantity purchased per visit remained almost the same. The fact that consumers chose to visit a liquor store one more time rather than buying more liquor per visit implies that the fixed cost of visiting a liquor store has decreased, either by reduced distance or by compensated by convenience of one stop shopping.

Second, there is correlation between the change in number of liquor stores nearby a zip code and the change in quantity in the area. Figure (A.4) plots change in liquor purchases of each household from 2011 to 2013 conditional on the household purchases liquor at least once a year, against the change in number of liquor stores within a 5 mile radius of the zip code which the household resides in. If the number of liquor retailers near a household's home zip code is increased, then the household is more likely to purchase more liquor than a year before. Likewise, if there are fewer outlets selling liquor after privatization, then quantity sold in the zip code tends to decrease. It motivates the idea that proximity to store is correlated with demand.

Switching to One Stop Shopping

Switching to one-stop shopping originated from all types of trips. Table (A.6) shows a transition matrix of household's switching trip types before and after. Abusing the notation, each household is classified as either G, GL1, GL2, or L. G type of households have never purchased liquor, and the rest of the households are assigned to a type which they chose most frequently conditional on purchasing liquor. More than half of the consumers who used to purchase – either by liquor-shopping only or grocery and liquor shopping at two stores – liquor have switched to one-stop shopping. Moreover, about a quarter of the consumers who have not purchased liquor before also switched to one-stop shopping. It is important to understand that this does not imply that non-drinkers all of sudden started consuming liquor, rather they could have switched from on-premises liquor consumption at a bar or restaurant to off-premises consumption. See Appendix (A.6) for more details.

Government Revenue and Social Costs

Government revenue has increased by \$44 million, which is 17.15% increase, over a year after privatization. Table (A.7) compares government revenue before and after. The revenue from the state-run stores' profits (\$62 million) are foregone but the state collects more through license issuance fee (\$84 million). Tax revenue has increased by 11% due to the increase in liquor sales. In addition, the state earned a one-time 31 million dollars by auctioning off the state-run stores.

The increase in liquor sales might lead to higher social costs in the form of excessive drinking, and the state may spend part of the revenue on correcting the social costs from consuming liquor. According to Sacks et al. (2013) and Selway (2006), one liter of excess drinking costs \$47.48 in Washington[4] . The cost includes health cost, productivity loss, damages from alcohol related accidents and crimes, etc., and 44.6% of it is bore by the

[4] Selway (2006) defines one liter of liquor with 80 proof is equivalent to 22.29 drinks. Sacks et al. (2013) finds that the total social cost is \$2.13 per drink in Washington.

state according to Sacks et al. (2013). Assuming that 23.3%[5] of the total alcohol consumption is due to excessive drinking, the negative externality per liter of liquor is \$10.06 and the state spends \$4.93 to mitigate the social cost. In 2013, 316 million dollars of social cost occurred and the state is estimated to spend 141 million dollars in to amend the social costs, which is 46% of the total government revenue[6] . Compared to 2011, the negative externalities are increased by 8.75%.

A.6 Extensive and Intensive Marginal Consumers of Liquor for Off-Premises Consumption

The source of increase in quantity sold originates from both extensive (on households) and intensive margins (purchase per household). I used the Nielsen Consumer Panel to extrapolate per household liquor purchases, conditional on the panelists who participated throughout 2011 to 2013. The total number of households who purchased liquor at least once from January 2011 to May 2012 was 311, and those from June 2012 to December 2013 was 425. That is, the extensive margin has increased by 37%. Table (A.8) further decomposes consumers into new and existing groups. New consumers are defined as those who did not purchase liquor from a store between January 2011 to May 2012 but started purchasing at least once between June 2012 to December 2013. 160 out of 425 consumers who purchase liquor at least once since privatization were new consumers. On the other hand, 46 out of 265 existing consumers stopped purchasing liquor from stores since privatization. The existing consumers consist of two third of total consumers from June 2012 to December 2013.

Demand growth along the intensive margin has also contributed to the overall increase in liquor sales. Table (A.9) describes liquor purchase in liter by existing and new consumers. Per household liquor purchase has increased from 10.47 liters before

[5] 23.3% of the population in Washington either binge drink or drink heavily according to BRFSS Prevalence & Trends Data, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Division of Population Health.

[6] If consumers have substituted from on-premises consumption to off-premises consumption, the total negative externalities from liquor consumption is lower than 316 million dollars. See Appendix (A.6). Moreover, \$10.06 as per liter social cost is a conservative measure since not all of the 23.3% of the liquor consumption is excessive drinking. Therefore, the social cost calculated here can be considered as an upper bound.

privatization to 11.93 liters after. Unlike the increase in extensive margins came from new consumers by definition, increase in intensive margins was driven by the existing consumers. Per household liquor purchase for existing consumers is 15.44 liters a year, which is almost twice as large as the new consumers' purchase, 8.59 liters. 46% of the total increment was contributed by the existing consumers.

It is important to note that not purchasing liquor from stores does not imply that the consumer does not consume liquor. Consumers could switch from liquor consumption at a bar or restaurant (on-premises consumption) to off-premises consumption, which would lead them to be counted as "new" liquor consumers in my analysis. In fact, the liquor quantity sold for on-premises consumption in million liters is decreased from 9.892 in 2011 to 8.424 in 2013, which is 15 % decrease. In contrast, the liquor quantity sold for off-premises consumption in million liters is increased from 29.653 to 31.424, which is 6% increase. The decrement in sales for on-premises is equivalent to 83% of the increment of sales for off-premises sales. Therefore, it is plausible that the new consumers defined in my analysis used to consume at a bar or restaurant before privatization, and then switched to off-premises consumption after privatization. In summary, liquor purchase by new consumers from stores should not be interpreted as they all of sudden starting consuming liquor since privatization.

One might worry about the possibility that the increase in quantity sold is from underreporting before privatization. Since it might be less costly to record the purchase if liquor was purchased at a Nielsen-affiliated stores after privatization. However, they have to manually input quantity regardless of the store's affiliation. Additionally, the reporting method has not changed over time. Therefore, if there is any, underreporting before privatization compared to after should be consistent. Moreover, the aggregate quantity sold data from the Washington State Department of Revenue supports that the liquor sales for off-premises consumption is increased.

A.7 Tables and Figures

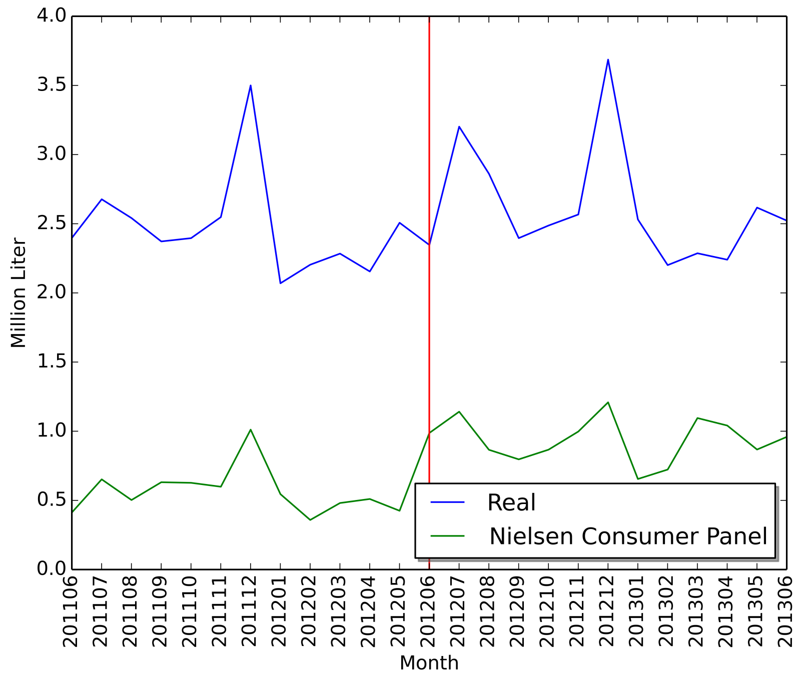
Table A.1: Summary of the Nielsen Consumer Panel Dataset

	Panelists		Grocery Purchases	Liquor Purchases	
Year	All	Liquor[a]	Weekly Frequency	Yearly Frequency	Liter[b]
2011	1592	433	3.19	5.12	2.06
2012	1568	546	3.25	5.21	1.84
2013	1545	588	3.15	6.24	1.80

[a] The number of the panelists who purchased liquor at least once in the given year

[b] Average amount of liquor in liter purchased per liquor shopping

Figure A.1: Liquor Sales by Nielsen Consumer Panel Dataset



Notes: Blue line indicates monthly liquor sold in liter for off-premises consumption in Washington, and the datasource is the Washington State Department of Revenue. The dotted red line represents projected liter sold by the panelists in the Nielsen Consumer Panel Dataset. The panelists' liquor consumption represent about 20% of the total liquor consumption consistently.

Table A.2: Grocery Basket for Price Index

Product Group	Price	Size	Unit
Bread and Baked Goods	3.56	22.5	OZ
Candy	4.07	10.0	OZ
Cheese	5.29	16.0	OZ
Deli Foods	4.41	16.0	OZ
Fresh Produce	1.65	1.0	Count
Milk	4.41	128.0	OZ
Packaged Meat	4.64	16.0	OZ
Pet Food	4.48	21.0	OZ
Frozen Food	4.59	22.0	OZ
Snacks	3.97	11.5	OZ
Yogurt	2.05	16.0	OZ
Basket Price	43.12		

Notes: Median number of product groups purchased per grocer shopping is 11. For each product group, I derived sales weighted average price per store with the median unit size.

Table A.3: Change in Number of Stores by Zip Code Characteristics

	Upper Bound	2011	2013	% Change
Median		3.00	15.00	100.00
Mean		4.31	22.98	281.29
Zip density[a]	13	0.76	1.34	76.51
	116	1.21	4.08	237.61
	1185	2.48	12.69	411.37
	Above	7.74	45.46	487.57
Income[b]	34,999	5.13	23.36	355.51
	59,999	3.69	20.47	454.18
	99,999	5.00	26.39	427.89
	Above	5.10	27.79	445.33
Race[c]	0.78	6.68	38.31	473.84
	0.89	4.90	27.50	461.75
	0.94	2.35	9.20	291.20
	Above	0.96	2.17	126.16

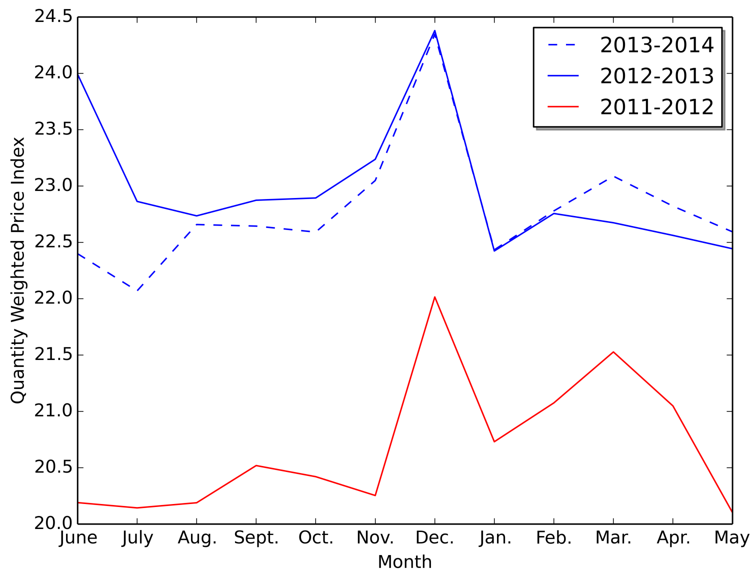
Note: Number of stores within a 5 mile radius of zip code.

[a] The zip code density is population divided by square miles. The results do not change much if population is only 21+.

[b] Median income in zip code tabulated area.

[c] Race is the proportion of Caucasians compared to other races.

Figure A.2: Average Post-Tax Price in Washington



Notes: The X-axis is month starting from June of the year. The Y-axis indicates average post-tax price per liter weighted by liter sold, adjusted by seasonal spirits CPI in January 2006 dollars. The blue lines are prices before privatization and the red line price before.

Source of data: The Washington State Department of Revenue

Table A.4: Regression of Average Per Liter Price

	No National Price		National Price	
	(1)	(2)	(3)	(4)
Intercept	11.8197	38.3830	2.7394	7.0839
	(15.2168)	(0.6669)	(0.6614)	(1.1988)
Privatization Dummy[a]	-23.9931	1.0561	1.9260	1.9558
	(45.0926)	(0.4010)	(0.3134)	(0.2927)
$P_{national}$[b]			1.0015	0.7975
			(0.0232)	(0.0505)
Brand Dummy[c]	N	Y	N	Y
Week Dummy	Y			
Year Dummy	Y			
N	9789	9777	5867	5866
R-squared	0.0082	0.6889	0.7768	0.9582

Dependent variable: weekly average per liter price in Washington. Price includes taxes.
Observation: by product (UPC) and week, 2011 - 2013
Note: Price is derived based upon the purchase data in Nielsen Consumer Panel, adjusted by seasonal spirits CPI in January 2006 dollars.

[a] 1 after May 31, 2012, 0 otherwise.

[b] Excluding Washington itself, Hawaii, and Alaska. Depending on the states, prices may or may not include taxes but the rule is consistent over time.

[c] Brand information was collected the Nielsen datasets and Mintel Global New Products Database.

Table A.5: Disproportionate Price Increase by Demographics

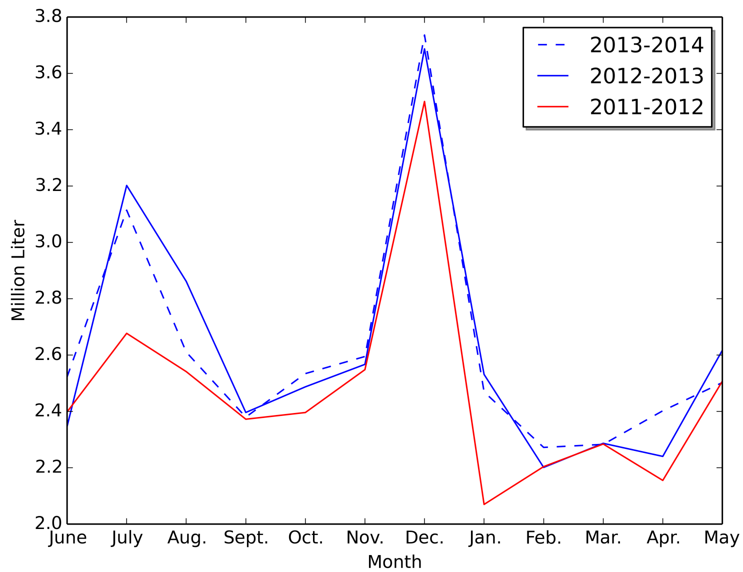
	Upper Bound	2011	2013	% Change
Median		21.36	21.81	6.97
Mean		21.71	22.10	5.80
Zip density[a]	13	19.04	21.24	11.52
	116	19.72	20.96	6.31
	1185	20.51	21.53	4.96
	Above	22.24	22.26	0.13
Income[b]	34,999	21.03	21.20	0.77
	59,999	21.23	21.82	2.77
	99,999	22.17	22.37	0.90
	Above	23.37	23.18	-0.79
Race[c]	0.78	22.11	22.27	0.72
	0.89	21.84	22.11	1.22
	0.94	20.44	21.37	4.57
	Above	19.41	21.20	9.21

[a] The zip code density is derived by population divided by square miles. The results do not change if population is only 21+.

[b] Median income in zip code tabulated area.

[c] Race is the proportion of Caucasians compared to other races.

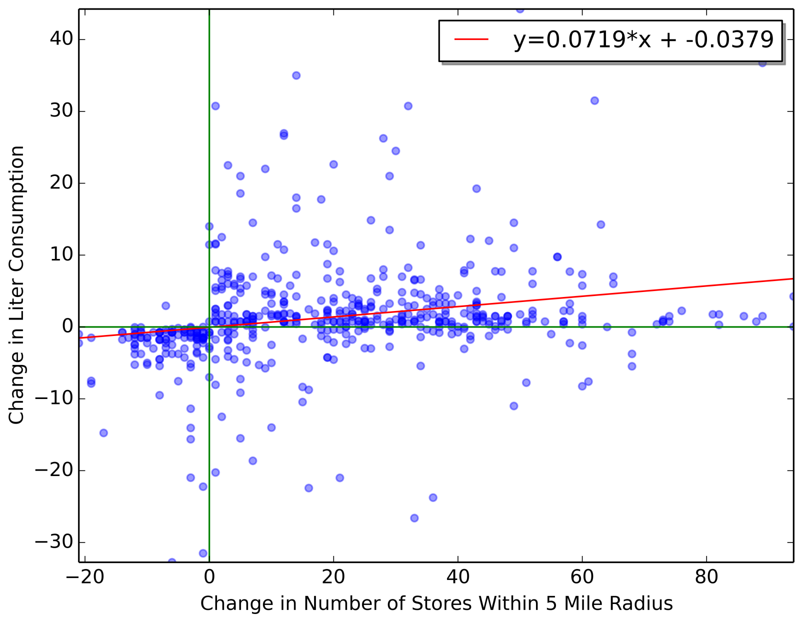
Figure A.3: Liquor Sold in Million Liter



Notes: The X-axis is month starting June of the year. The blue lines indicates liquor quantity sold in million liters for off-premises consumption in 2012-2013 and 2013-2014 while the red line represents before privatization.

Source of data: The Washington State Department of Revenue

Figure A.4: Disproportionate Increase in Liquor Consumption



Notes: The X-axis is the change in the number of stores within a 5 mile radius of each store from 2011 to 2013. The Y-axis is the change in the total liquor purchases in liter from 2011 to 2013. Each dot represents a household who purchased liquor at least once during the time period and participated the Nielsen Consumer Panel for those two years.

Table A.6: Transition of Shopping Trips by Households

Before → After	G	GL2	GL1	L	Households
G	0.7207	0.0097	0.2441	0.0256	1135
GL2	0.1943	0.0474	0.6493	0.1090	211
L	0.1587	0.0582	0.5979	0.1852	189
Households	889	32	527	87	1535

Note: Before time period is from January 2011 to May 2012 and after is from June 2012 to December 2013. Only panelists who participated throughout the time period are used. G type of households have never purchased liquor, and the rest of the households are assigned to a type which they chose most frequently conditional on purchasing liquor.

Table A.7: Government Revenue

	2011	2013	% Change
Profits ($1,000)	61,606		
Tax	197,694	219,679	11.12
License Fee		84,082	
Total	259,299	303,761	17.15

Negative Externalities[a]	-290,170	-315,572	8.75
Total - N.E.	-30,871	-11,811	61.74
State Expenditure[b]	-129,415	-140,745	8.75
Total - S.E.	129,884	163,016	25.51

[a] Externalities are estimated as $10.06 per liter of liquor according to Centers for Disease Control and Prevention.

[b] Assuming that the state spends 44.6% of the negative externalities to mitigate it.

Table A.8: Extensive Margin on Households

	Total hhlds		Existing				New[b]	Increment	
	2012-3	% Exist[a]	2011-2	2012-3	Δ	% Δ	2012-3	Total	% Int.[c]
Overall	425	0.62	311	265	-46	-0.15	160	114	-0.40

Note: Conditional on the 425 households who participated in the Nielsen Consumer Panel Dataset throughout 2011 to 2013 and have purchased at least once since privatization.

[a] Existing households are defined as the panelists who purchased liquor at least once during January 2011 to May 2012.

[b] New households did not purchase any liquor from a store from January 2011 to May 2012 but purchase liquor at least once since privatization.

[c] Percentage of intensive margin

Table A.9: Intensive Margin on Liter Sold

	Total hhlds		Exist				New[b]	Increment	
	2012-3	% Exist[a]	2011-2	2012-3	Δ	% Δ	2012-3	Total	% New
Overall	5070	0.81	3256	4091	836	0.26	979	1815	0.46
Per hhld	12		10	15	4	0.47	9	16	

Note: Conditional on the 425 households who participated in the Nielsen Consumer Panel Dataset throughout 2011 to 2013 and have purchased at least once during the three years.
[a] Existing households are defined as the panelists who purchased liquor at least once during January 2011 to May 2012.
[b] New households did not purchase any liquor from a store from January 2011 to May 2012 but purchase liquor at least once since privatization.

Appendix B

Appendix to Chapter 2

B.1 Joint Distribution of Income and Political Affiliation

Here we describe the procedure for generating a coarse joint distribution between political preferences and income. For most counties in California the ACS annually reports the percent of surveyed households in designated income brackets. Using this information we fit a log-normal income distribution to each county-year used for simulating income when integrating over $f_m(y)$.

From the FEC we can observe individuals who have donated over \$250 to a political action committee (PAC) or to a candidate for a federal election. Using information on the political affiliation of these candidates (or to the political affiliations of the candidates to which the PACs have donated), we associate each individual with a political party based on the largest recipient of their donations. Under the assumption that wealthier individuals donate more to their political causes, we then use this constructed data set to derive a multinomial distribution of three political affiliations (Republican, Democrat, and independent) for several income brackets defined by donation sizes.

We found, however, that it was unreasonable to expect individuals in the lowest income brackets to be donating over \$250 to political campaigns. Therefore, we designated a cutoff for wealth above which the FEC data was a good proxy for political affiliation and below which we use average political party affiliations to assign parties. Matching the joint distribution of income and political preference from a survey of California households in Ansolabehere and Pettigrew (2014), we found that a cut off at the median income in the county-year worked best.[1]

B.2 Determination of Charging Station Variable

Understanding which charging stations are relevant to which consumers is a critical part of our analysis and taking advantage of our frequency, high granularity purchase and charging station data. For example, relative to nearby charging stations charging stations in San Francisco are likely insignificant to a consumer in southern Los Angeles. Ultimately we found that work-area charging stations are a superior measure for a consumer's relevant charging stations than charging stations near home. While home zip code charging stations may seem like the most natural measure, home chargers for PEVs appear to mitigate the need for other proximal charging stations. According to the Center for Sustainable Energy (2013a) survey, approximately 90% of PEV owners own Level 2 home chargers.

Using the California Department of Transportation CA Household Travel Survey (CHTS), which tracked detailed travel information of participants via GPS, we determined destination zip codes for each participant's home zip code (see Table B.10 in Appendix B for more survey information). Survey participants document specific trips into four categories: to home, to work, to school, to other. The second panel of Table B.11 provides details on trips to school and to work. Work trips are typically longer (over 10 minutes with a stay of over 60 minutes) than school trips; only 1/3 of school trips qualify as "long trips" under this definition, while 2/3 of work trips do. Under the assumption that charging events are more desirable after lengthy battery-draining drives and feasible only during lengthy stays at the trip destination, we focused on work zip codes as the relevant destination zip codes. For each home zip code, there are 11.89 associated work place zip codes on average. Most home-work zip code links feature few unique households because of the sample size and level of granularity. On average 1.78 households make the trip between the home-work zip code, though some feature as many as 100.

To test whether work zip code charging stations are more relevant to consumers

[1] We do not use the joint distribution that could be obtained from this survey because of small sample issues. We would be unable to have separate political party distributions by county using the survey.

than home zip code charging stations, we use a simple heuristic comparing PEV uptake with charging stations at home versus charging stations at work at the time of purchase. Work zip code charging stations for a particular home zip code are the average of charging stations across all work zip codes weighted by the number of households in that home-work zip code pair. Table B.1 show the results of the regression. It reveals that charging stations in work zip code is more closely related to PEV purchase decision than those in home zip code. It is reasonable result because most PEV purchasers install home charging outlet, which reduce the use of public charging stations around the neighborhood. Figures B.3 and B.4 in Appendix B illustrate the same closer relationship between demand and work zip code charging stations, rather than home zip code charging stations.

While a course metric for determining whether charging stations matter in influencing purchasing behavior — that is the job of the more complete estimation model — the results do not contradict the theoretical claim that public home charging stations are less significant to PEV purchasers than work charging stations. Actual PEV owner charging behavior, featuring away-from-home charging events, as discussed in section 2.2 coupled with the common presence of in-home charging units also support that public charging stations are more relevant in work zip codes. Based on this evidence, in our estimation individuals in a certain home zip code are "assigned" the work place charging stations constructed above.

B.3 Tables and Figures

Table B.1: Charging Stations at Work Matters More for PEV Purchase

	PEV[a]	PEV
Chg Stn at Work	0.1534	0.1459
	0.002	0.002
Chg Stn at Home		0.0162
		0.001
Year	0.2445	0.2449
	0.008	0.008
Constant	-491.601	-492.511
	16.936	16.919
N	57456	57456
R^2	0.204	0.206

Observation level is (home zip code, month) pair.
Sample Period: March 2010 - February 2014
[a] PEV: number of PEV purchase at home zip code

Table B.2: Sponsored Charging Infrastructure Projects

	EV Project	ChargePoint America
Project Period	January 2011 - December 2013	May 2011 - June 2013
Area Covered in CA	Los Angeles, San Diego San Francisco	Los Angeles Sacramento, San Francisco
Funding Institute	Department of Energy (DOE) American Recovery and Reinvestment Act (ARRA)	DOE ARRA
Partnered Charging Network	Blink	ChargePoint
Funding Amount	\$130 million dollar[a]	\$18.4 million dollar
Total Charging Units Installed[b]	3182	1916
Public Charging Units Installed	933	857

Sources: Project Electric Vehicle Charging Infrastructure Summary Report (Q4 2013), ChargePoint America Vehicle Charging Infrastructure Summary Report

[a] Total budget was \$230 million and half of it was funded by the DOE. \$130 was allocated to install public or privet charging stations. The rest is operational cost and subsidy for residential chargers.

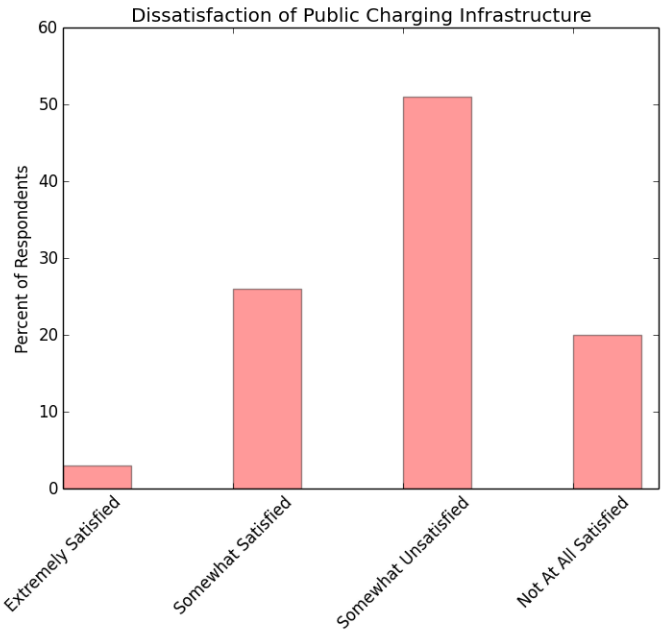
[b] Charging units are counted only in California. Q4 2013

Table B.3: Desired Range Vs. Actual Range

	Leaf	Volt	Prius
Desired Electric Range	200	100	50
Actual Electric Range	78	38	12
Dissatisfaction with Public Charging Infrastructure		71%	

Source: California Plug-in Electric Vehicle Driver Survey Results, February 2014

Figure B.1: Dissatisfaction with Charging Infrastructure



Source: California Plug-in Electric Vehicle Driver Survey Results, February 2014

Table B.4: Cumulative Charging Stations in California

	Public			Private		
	LV1	LV2	DC Fast	LV1	LV2	DC Fast
Q1 2010	0	0	0	0	0	0
Q1 2011	241	81	22	13	109	5
Q1 2012	332	1001	25	25	314	6
Q2 2012	427	1605	31	45	366	6
Q3 2012	453	1761	33	45	379	6
Q4 2012	470	1909	44	45	395	6
Q1 2013	543	2355	63	46	398	6
Q2 2013	584	2694	132	46	410	6
Q3 2013	606	2917	148	46	419	6
Q4 2013	621	3078	180	46	431	9
Total		4665			608	

Source: StationGeo

There are also 153 charging stations of which the ownership is unknown.

Table B.5: Utility Discount for PEV Charging at Home

Avg Monthly Cost by PEV	Pacific Gas & Electric	Southern California Edison	San Diego Gas & Electric	Sacramento Municipal Utility District
With Discount rate[a]	$31.91	$29.7	$52.8	$26
Without Discount rate[b]	$56.19	$58.09	$62.41	$60.59
Equivalent Avg Monthly Gas Cost[c]			$128.60	

[a] Assuming a LEAF charged during super off-peak time (12:00 AM to 6:00 AM) or off-peak times (11:00 PM to 7:00 AM) under a plan exempt from a tiered rate and is driven 38 miles a day (= 11.02 kwh a day, 330 kwh a month)

[b] Assuming that the current rate plan enjoys no Time of Use rate and charging PEV is in cost tier 101-130%.

[c] At average 2013 gas prices in California

Table B.6: Charging and Driving Behavior

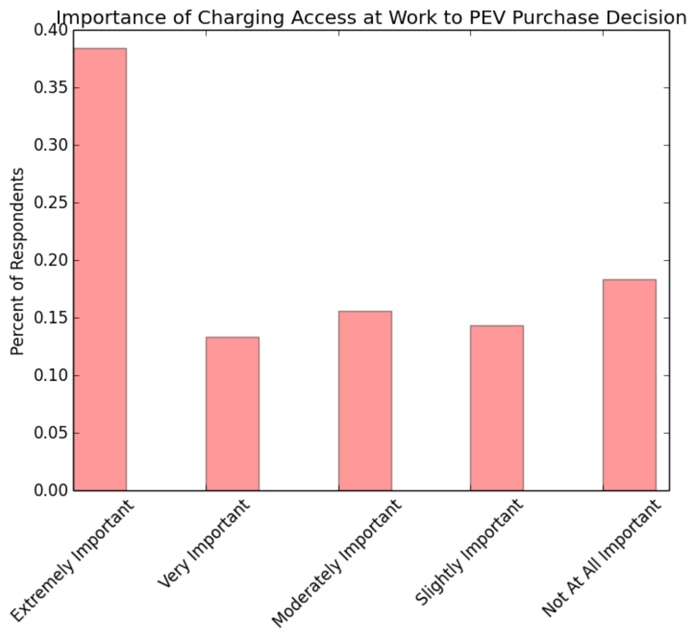
Average	LA		SD		SF		CA[b]
	Leaf	Volt	Leaf	Volt	Leaf	Volt	Gas Only
Trip Distance[a]	6.4	7.7	6.6	8.2	7.3	9.39	8.4
Distance per Day	24.8	38	26.7	39.6	27.4	41.4	38.48
Number of Trips b/w Charging Events	3.8	3.8	3.7	3.7	3.5	3	
Distance b/w Charging Events	24.2	28.9	24.6	30.1	25.7	28.1	
Charging Events per Day	1	1.3	1.1	1.3	1.1	1.5	
% Charging Events Away from Home	30	23	23	21	28	21	

Source: EV Project Nissan Leaf Summary Report, Q4 2013, EV Project Chevrolet Volt Summary Report, Q4 2013

[a] Distance is measured in miles.

[b] 2010-2012 California Household Travel Survey

Figure B.2: Charging Access at Work is Important



Source: California Plug-in Electric Vehicle Driver Survey

Figure B.3: Cumulative PEV Purchase and Charging Stations, Zip 95014

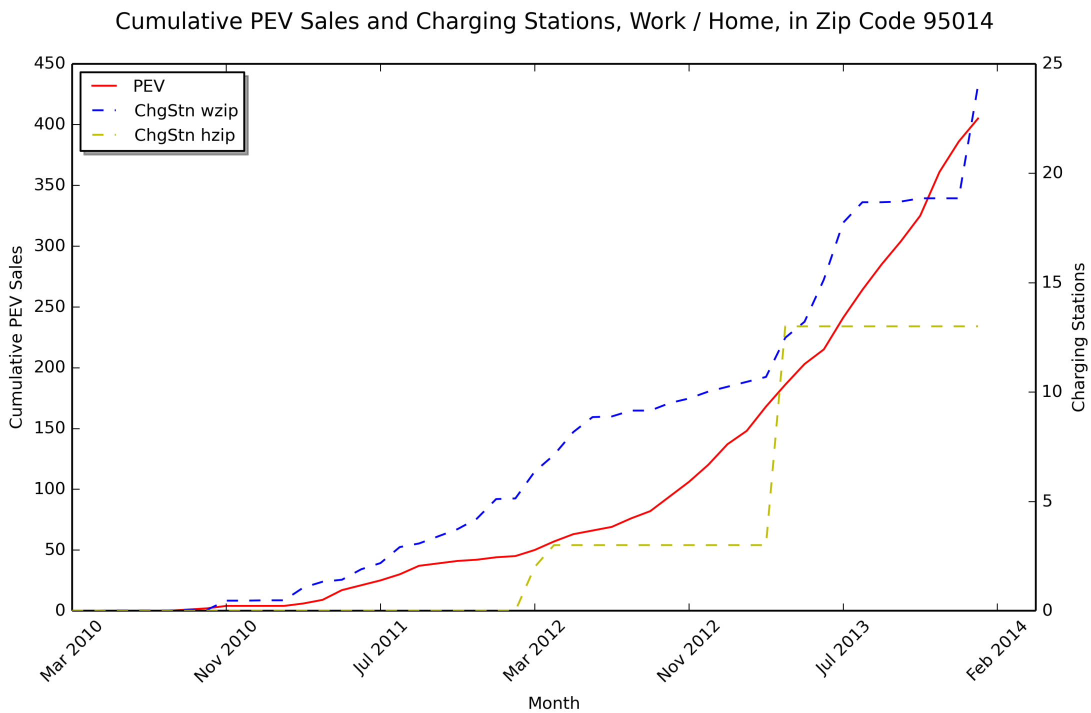


Figure B.4: Monthly PEV Purchase and Charging Stations, Zip 95014

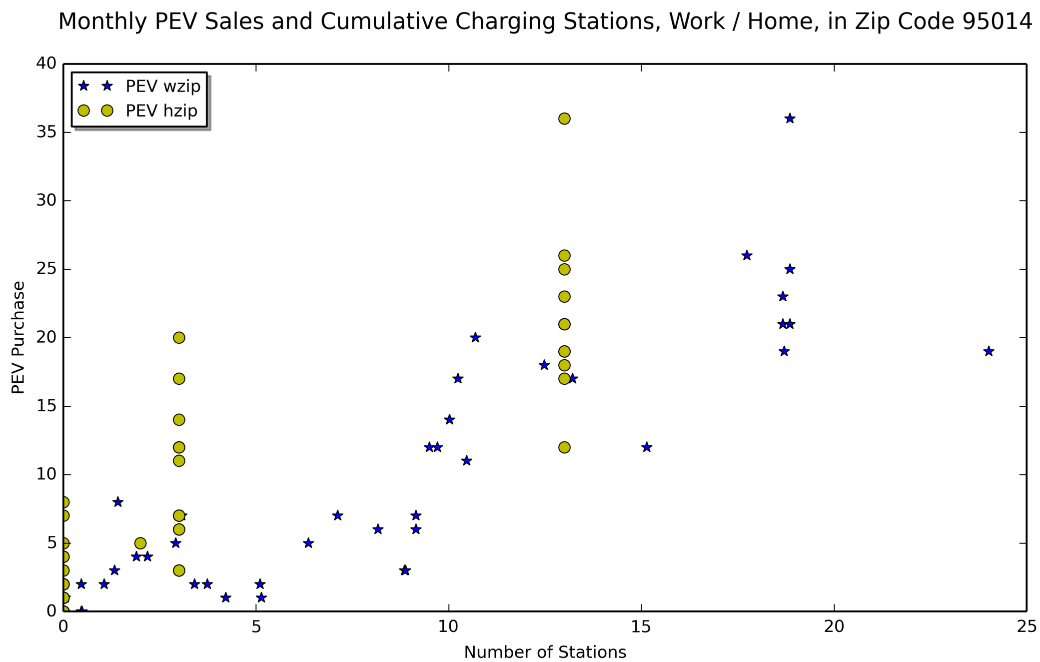


Table B.7: Income Distribution Conditional on a Vehicle or PEV Purchase

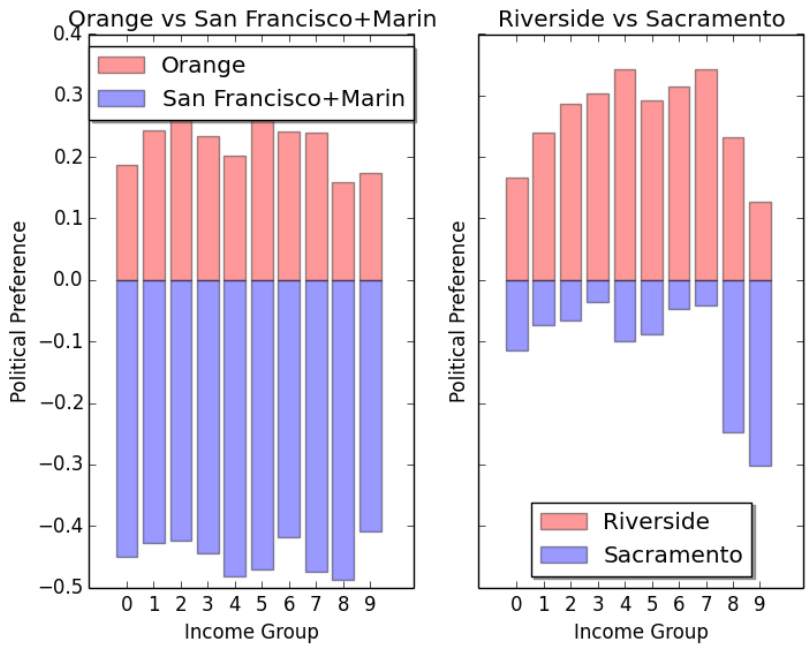
Income 2010-2012	Vehicle Purchase	PEV Purchase
Less than $74,999	0.19	0.07
$75,000 to $99,999	0.16	0.10
$100,000 to $149,999	0.27	0.24
$150,000 to $199,999	0.20	0.23
$200,000 to $249,999	0.09	0.15
More than $250,000	0.10	0.22
Number of Obs	278	92

California Department of Transportation CA Household Travel Survey

Income 2012-2014	LEAF Purchase	Tesla Purchase
Less than $49999	0.03	0.01
$50,000 to $74,999	0.07	0.02
$75,000 to $99,999	0.14	0.03
$100,000 to $124,999	0.15	0.05
$125,000 to $149,999	0.14	0.06
$150,000 to $174,999	0.12	0.07
$175,000 to $199,999	0.10	0.06
$200,000 to $249,999	0.12	0.13
$250,000 to $299,999	0.05	0.10
$300,000 to $349,999	0.03	0.08
$350,000 to $399,999	0.02	0.05
$400,000 to $449,999	0.01	0.05
More than $450,000	0.02	0.31
Number of Obs	1411	1126

Clean Vehicle Rebate Project EV Owner Demographics and Diffusion Survey

Figure B.5: Heterogeneous Income and Political Distribution by County



Source: Federal Elections Commission

Political preference: -1 if democrats, 1 if republicans, 0 if independent

Income group: percentile of FEC donation given county

Table B.8: Heterogeneous Income and Political Distribution by County

County	Income[a]	PEV[b]	LEAF/Volt[c]
Orange	$96,036	8.44	0.72
San Francisco and Marin	$108,690	5.68	1.54
Riverside	$69,835	2.61	0.62
Sacramento	$68,532	2.73	2.15

[a] Average household income in 2013, ACS

[b] Total PEV sold per 1,000 capita until August 2014, Clean Vehicle Rebate Project rebate data set

[c] Leaf/Volt: ratio between Leaf and Volt demand

Table B.9: CHTS Trips Summary

	GPS Sample[a]
Sample Period	2010-2012
Number of Households	5,717
Total Number of Trips	285,340
Number of Households	5,717
Number of Home Zip	1,519
Number of Work Zip	859
Total Trips	285,340

Trips	to Work	to School
Total Trips	29,567	30,106
Long Trips[b]	20,491	12,406
Avg Trip Duration[c]	23	16
Median Trip Duration	15	10
Min Trip Duration	1	1
Max Trip Duration	679	1080
Sd Trip Duration	23	24
Avg Stay Duration	225	223
Median Stay Duration	376	132
Min Stay Duration	1	1
Max Stay Duration	1,409	1,349
Sd Stay Duration	225	231

2010-2012 California Household Travel Survey

[a] GPS samples are subset of the total samples who recorded their trips as well.

[b] Trips which required more than 10 minutes of drive and more than 60 minutes of stay

[c] Duration is measured in minutes

Table B.10: CHTS and CVRP Survey Summary

	CHTS[a]	CVRP[b]
Survey Period	2010 - 2012	Oct 2013 - May 2014
Vehicle Purchase Period	1994 - 2012	Sept 2012 - Apr 2014
Number of Respondents	42,431	8,415
Sample Used	1588	6,602
Car Purchase	278	6,602
PEV Purchase	92	6,602

[a] California Department of Transportation CA Household Travel Survey. CVRP surveyed only PEV purchasers who applied for CVRP rebate.

[b] Clean Vehicle Rebate Project EV Owner Demographics and Diffusion Survey.

In the micro income moments we only used the samples which has the income and purchase year are known.

Table B.11: Statistics of Work and Home Zip Code Pairs

	Work Zip	Home Zip
Avg Number of Corresponding Zip Codes[a]	20.74	12.24
Med Number of Corresponding Zip Codes	10	11
Sd Number of Corresponding Zip Codes	34.30	9.10
Avg Num of Hhlds Commuting/Living	36.97	21.83
Med % of Hhlds Commuting/Living	14	17
Sd Number of Hhlds	91.83	20.46
Number Work-Home Pair	17791	
Avg Number of Hhlds for Work-Home Pair	1.78	
Med Number of Hhlds for Work-Home Pair	1	
Sd Number of Hhlds for Work-Home Pair	2.52	

California Department of Transportation CA Household Travel Survey.

[a] Corresponding zip codes of work zip codes refer to the number of associated home zip codes. If at least one household described that they commute between the two zip codes, the zip code pair is considered as ”associated”.