CCP Estimation of Dynamic Discrete Choice Demand Models with Segment Level Data and Continuous Unobserved Heterogeneity: Rethinking EV Subsidies vs. Infrastructure

Cheng Chou Tim Derdenger*

January 2025

Abstract

When multiple groups of consumers reside in the same market, we determine that we can write each group's conditional choice probabilities (CCP) as a function of unobserved consumer heterogeneity. Moreover, we can specify choice probabilities of one group as a function of another by shifting the unobserved component. Armed with our novel CCP estimator, we develop an approach to identify and estimate a dynamic discrete demand model for durable goods with non-random attrition of consumers and continuous unobserved consumer heterogeneity but without the usual need for value function approximation or reducing the dimension of state space by ad hoc behavioral assumptions. We illustrate the empirical value of our method by estimating consumer demand for electric vehicles in the state of Washington during the period of 2016–2019. We also determine the impact of a different federal tax credit based on the electric range of a car rather than the size of the battery, which was the existing policy during the data period, and we evaluate how best to seed a nascent market that presents indirect network effects to drive faster adoption. Should the government incentivize adoption through consumer tax credits or through EV infrastructure?

Keywords: Dynamic discrete choice, dynamic selection, Electric Vehicles (EVs)

1 Introduction

Dynamic discrete choice models play a key role in modeling consumer demand due to their ability to incorporate the dynamics of the state of the market and the intertemporal preferences of consumers. The incorporation of these dynamic aspects comes at the cost of the complexity of estimation and the obscurity of identification. Specifically, defining a tractable state space while accounting for all the products in the market is often a difficult task, leading some to adopt ad hoc approximation methods. The task becomes even more challenging when the researcher wants to include multidimensional unobserved state variables while having access to only aggregate sales data. Besides the estimation difficulties, it is also

*Corresponding Author: Tim Derdenger is an Associate Professor of Marketing and Strategy at the Tepper School of Business, Carnegie Mellon University. Email:derdenge@andrew.cmu.edu. Note: the authors thank Bob Miller and Matt Shum for comments, conference participants from the Dynamic Structural Econometrics Conference (DSE) at the University of Wisconsin and the UT Dallas BASS Conference, seminar participants at the University of Texas, Austin, University of Rochester, UC-Riverside, Cambridge University, and the University of Delaware.

uncertain whether (or which of) the structural parameters are identified when there are both continuous unobserved consumer heterogeneity and continuous unobserved product characteristics. In the market of durable products where consumers leave after purchasing, we have the additional problem that the distribution of unobserved consumer heterogeneity (e.g. random price coefficients), for those consumers who remain in the market, is likely to change over time. It is necessary to understand the consequences of such non-random attrition of consumers (also known as dynamic selection), which usually causes estimation bias in panel data analysis if ignored.

Our main contribution is to develop a novel approach using market-level data to model, identify, and estimate a dynamic discrete choice demand model without replacement for durable goods with dynamic selection, continuous unobserved consumer heterogeneity, and continuous unobserved product characteristics, in addition to the commonly included individual-product idiosyncratic errors. The unobserved product characteristics are specified as serially correlated and correlated with the observed product characteristics, particularly price. The continuous unobserved consumer heterogeneity can be multidimensional, and its distribution varies over time due to the non-random attrition of consumers. We provide a new method to estimate all model primitives, including the consumer's discount factor, without the need to reduce the dimension of the state space or by other approximation techniques such as discretizing state variables. We also provide new identification results that show the model is identified while being agnostic about how consumers form their beliefs regarding the state transition distribution. Yet, we still require that their beliefs are consistent with what is observed in the data.[1] The implementation of our new estimator only involves nonlinear least squares (NLS) and 2-stage least squares (2SLS). The simplicity of the estimation allows researchers to estimate multiple model specifications at little computational cost.

We illustrate the empirical value of our method by estimating the consumer demand for electric vehicles (EV) in the state of Washington during the period of 2016–2019. Our estimates determine that demand for EVs is driven in part by price, miles per gallon equivalent (MPGe), horsepower, electic range, gas prices, and the network size associated with electric charging stations. We also determine that there are significant levels of unobserved consumer heterogeneity associated with EV price and MPGe. After estimation, we leverage our estimates to further determine the impact of a different federal tax credit based on the electric range of a car rather than the size of the battery, which was the existing policy during the data period.[2] We conclude a simple tax credit of \$30.54 per mile of electric range reduces CO_2 emissions by 11.03%, with 361 more electric vehicles on the road at no additional cost relative to the existing policy. We further address the question of how best to seed a nascent market that presents indirect network effects to spur adoption. Should the government incentivize adoption through consumer tax credits or EV infrastructure? Would such money be better directed at EV infrastructure by building a larger Level 3 charging network? We determine that by eliminating consumer tax credits and focusing on EV infrastructure by building a larger Level 3 charging network, the adoption of battery powered

[1] Recently, An, Hu and Xiao (2020) use individual-level panel data to identify the agent's preference and their subjective beliefs, which do not need to be rational expectations or myopic. Our results are based on market-level data, the discount factor in this paper will be identified without belief restriction (the discount factor is assumed to be known in their paper), and our state variables, observed and unobserved, are all continuous (the state variables, excepting for the conventional utility shocks, are discrete in their paper). By no means, we are claiming that our results are more general. We limit our research scope to the market of durable goods, where purchasing can be viewed as a terminal action, hence simplifying the task, but their paper focuses on general dynamic discrete choice models.

[2] The proposed policy is treated as a permanent change.

electric vehicles (BEVs) increases by 25.75%, with a 51.33% reduction in CO2.

The proposed approach relies on our new idea of estimating conditional choice probability (CCP) functions. In its original form (Hotz and Miller, 1993), the CCP function is a function of observed state variables.[3] Applying the original CCP estimator to the market of durable goods has two major difficulties. The first is the large dimension of product space and/or product characteristics space. The second is the continuous multidimensional unobserved state variables (unobserved consumer preference heterogeneity) whose unknown distribution could also vary over the course of time due to nonrandom attrition of consumers. We provide a new perspective by exploring market-level data on multiple demographic groups of consumers in the same market. *Instead of viewing the CCP as a function of all observed state variables as in individual-level panel data, our objective is to estimate the CCP as a function of unobserved consumer heterogeneity for each observed group and market.* Recovering directly the CCP for each group and market evaluated at each value of unobserved consumer heterogeneity is the central pillar of our estimator and essential for addressing the dynamic selection problem due to the nonrandom attrition of consumers after purchasing.

We discover that when we observe the market share of a product for multiple groups of consumers in the same market, we can easily estimate the CCP function of unobserved consumer heterogeneity by NLS. To see the intuition, note that for a demographic group g, the known market share in this group is the integrated unknown CCP of group g with respect to the unknown distribution of unobserved consumer heterogeneity in group g. This can be viewed as one moment condition. If the number of unknown CCP functions grows with the number of groups, we can never recover these unknown CCPs. The key insight is that when there are multiple groups of consumers residing in the same market, we can write each group's CCP as a function of unobserved consumer heterogeneity because these consumers face the same state of the market. Moreover, we can specify the conditional choice probabilities of one group as a function of another by shifting the unobserved component.

The presence of multiple groups of consumers in the same market creates within-market variation, which also plays a key role in simplifying our CCP estimation.[4] By exploiting the variation of group market shares within the same market, we can avoid estimation issues due to the unobserved product characteristics and the possibly high dimension of product characteristics (since they are fixed given one particular market). We explicitly show how the effect of the state of the market on demand is aggregated into the parameters of our CCP as a function of unobserved consumer heterogeneity, which is then further mapped to the parameters of consumer flow utility functions.

Having highlighted the paper's innovations, we believe it is important to discuss the data requirements for implementing our new methodology. Up to now, researchers who employ market-level sales data have been in search of a methodology that is able to accommodate unobserved state variables as well as continuous forms of unobserved consumer heterogeneity in preference parameters, but without the cost

[3]Arcidiacono and Miller (2011) made important progress so that the CCP function can depend on an agent's unobserved discrete type.

[4]The use of within-market variation is not new in the literature of demand identification and estimation. Recently, Berry and Haile (2020) developed new results of nonparametric identification of demand function using the variation in the choice probabilities of different individual consumers within the same market. Within a market, the state of the market (including demand shocks within the market) does not change, but the observed consumer heterogeneity can still shift demand quantity—observably different consumers have different choice probabilities. Hence, the within-market variation of observed consumer heterogeneity is a natural instrument for demand quantities.

of reducing the state space via approximation. With our methodology and panel data of product sales for two or more consumer groups (or repeated cross-sectional data of individual consumer purchases where researchers can construct group sales from it), researchers can now account for both needs at little cost. The researcher must also be in possession of data from the origination of the market and that the existing product's replacement value is zero, if there is one, to remove any bias associated with the well-known initial conditions problem. Because customer segmentation is a standard marketing practice, it is easy to obtain panel data at the consumer segment level.[5] For example, the data company NPD provides such consumer segment-level panel data for many industries from apparel to video games.[6]

In the remainder of the introduction, we discuss the literature. Our identification results are novel relative to the literature on identifying dynamic discrete choice (DDC) models. Our model for durable goods can be understood as a general DDC model in which a subset of unobserved state variables (unobserved product characteristics herein) are continuous, serially correlated, and correlated with other observed state variables. The existing identification results (Magnac and Thesmar, 2002; Norets, 2009; Kasahara and Shimotsu, 2009; Arcidiacono and Miller, 2011, 2018; Hu and Shum, 2012; Hu et al., 2017) in the literature of DDC models cannot be applied here.

Most of the research that focuses on individual-level data does not include persistent unobservable state variables (e.g. Bajari et al., 2016; Daljord, Nekipelov and Park, 2018).[7] The following exceptions involving persistent unobservables are worth noting. Hu and Shum (2012) study dynamic binary choice models with continuous unobserved state variables, but their identification result is limited to the conditional choice probabilities and state transition distribution functions, not to model primitives like flow utility functions and the discount factor. Norets (2009) does include a serially correlated unobservable *idiosyncratic error*, which is individual-specific rather than an aggregate product shock like in our case. Arcidiacono and Miller (2011) model persistent unobservables, but limit them to a discrete set of values.

Our estimation approach is also new relative to the literature on estimating DDC models. First, our estimation approach does not rely on the validity of specific approximations like interpolation or other value function approximations, or behavioral assumptions that consumers only consider some function of the state space and not the entire state (Melnikov, 2013; Gowrisankaran and Rysman, 2012). Second, we estimate more model primitives than the current literature since our method recovers not just the preference parameters but also the discount factor.

Our work builds on several foundational papers in the demand estimation literature. First is the result that the difference between choice-specific payoff is a function of individual choice probabilities (Hotz and Miller, 1993) in static and dynamic settings. The work of Berry (1994) and the BLP model (Berry, 1994; Berry, Levinsohn and Pakes, 1995; Berry and Haile, 2014) on demand estimation with market-level data including unobservable product characteristics have been extensively used. This is similar to our setting but focused on a static environment.

Extending the BLP models to a dynamic setting with forward-looking agents is a challenge. Some

[5]Segment level panel data are also substantially cheaper than individual-level data if researchers have to purchase data.

[6]See "Checkout Segmentation & Survey Insights," NPD.com, accessed September 7, 2021, `https://www.npd.com/products/checkout-segmentation-survey-insights/`.

[7]We note that Daljord, Nekipelov and Park (2018) presents an innovative way to identify the discount factor in DDC models with individual data. The primary difference is that our setting involves persistent unobservable state variables, whereas those are not present in the aforementioned paper.

researchers either do not model persistent unobserved shocks (Song and Chintagunta, 2003), or make them time-invariant (Goettler and Gordon, 2011). Others have focused on improving the computational speed of fixed-point estimators with a variety of approaches. Dubé, Fox and Su (2012) have employed a constrained optimization approach from (Su and Judd, 2012) to estimate static and dynamic structural models. Melnikov (2013) and Gowrisankaran and Rysman (2012) develop an approximation based on inclusive value sufficiency that allows the researcher to collapse the multi-dimensional state into one dimension, making the problem much more computationally tractable when using aggregate data. Moreover, the formal identification in the paper is not specified. Derdenger and Kumar (2019) have studied the approximation properties of this approach, and have shown that it is a biased and inconsistent estimator when consumers track the full set of state variables. Sun and Ishihara (2019) takes a different approach to include unobserved heterogeneity in a dynamic demand estimator. They transform the generalized method of moments (GMM) nested fixed point estimator by transforming it into a quasi-Bayesian (Laplace type) estimator. They develop a novel MCMC method that efficiently solves fixed-point problems.

Another paper in the field of dynamic discrete choice demand is that of Chou, Derdenger and Kumar (2019). They present a dynamic discrete choice estimator that eliminates the approximation of the expected value function by simply leveraging the Hotz and Miller (1993) inversion. Yet, the Chou, Derdenger and Kumar (2019) estimator does not include unobserved heterogeneity in consumer preferences. Moreover, the model uses only "standard" aggregate sales data. Our paper differs substantially from this work. We provide a model that includes observed consumer heterogeneity (in the form of income) AND unobserved consumer preference heterogeneity—as presented in Gowrisankaran and Rysman (2012) and Sun and Ishihara (2019). Our estimator also requires a very different data structure than that of Chou, Derdenger and Kumar (2019), Gowrisankaran and Rysman (2012), and Sun and Ishihara (2019). We believe that we are the first paper to present an estimator for a dynamic discrete choice demand model that includes observed and unobserved consumer preference heterogeneity without the use of individual-level data as in Arcidiacono and Miller (2011) and under a weaker assumption about how consumers form expectations for state variables. We do so with a new data structure that is a hybrid between pure aggregated sales data and pure individual purchase data, again a sizeable departure from the works of Chou, Derdenger and Kumar (2019), Gowrisankaran and Rysman (2012) and Sun and Ishihara (2019).

2 The Dynamic Discrete Demand Model

2.1 The Dynamic Model

The timing of our model is the following. In each period t, a forward-looking consumer i observes the state of market Ω_{it} and considers whether or not to purchase a durable product from the available goods $1, \ldots, J$.[8] Note, the state of the market for individual i does not include the value of the existing product that is to be replaced, if there is one, as this is assumed to be zero for all individuals. The associated expected *lifetime* payoffs are $v_{i1t}, \ldots, v_{iJt}$. The lifetime payoff is a "sum" of expected discounted per period or flow utilities. We first describe the flow utilities. If consumer i does not purchase in period t, she receives the flow utility ε_{i0t} in period t and stays in the market. When consumer i purchases product

[8] For the simplicity of exposition, we let the product space be fixed. The arguments do not change if we consider a time-varying choice set.

j at time t, her indirect flow utility *during the purchase period* t is

$$f_{ijt} = \delta_j + \gamma' X_{jt} + \xi_{jt} - \alpha_i P_{jt} + \varepsilon_{ijt}.$$

She then receives the identical flow utility $\delta_j + \gamma' X_{jt} + \xi_{jt}$ in each period following her purchase and exits the market. Thus, purchasing a product is a terminal action a la Hotz and Miller (1993)–we do not allow the consumer to make a replacement decision in future t periods. Here X_{jt} is a vector of observable product attributes other than price, P_{jt} is the price, and ξ_{jt} is unobserved product characteristics. Following Berry (1994), the term $\delta_j + \xi_{jt}$ can be interpreted as the mean of consumers' valuation of unobserved product characteristics, such as quality, in period t. We let ξ_{jt} have mean zero over T periods, so δ_j is the product fixed effect, and therefore ε_{ijt} denotes the individual deviation from this mean. Let $X_t \equiv (X'_{1t}, \ldots, X'_{Jt})$, and P_t and ξ_t are defined similarly. Let $\Omega_{it} \equiv (X'_t, P'_t, \xi'_t, \varepsilon'_{it})'$.

Suppose that there are $G \geq 2$ demographic groups of consumers (such as income bracket, age group, sex, etc.) in the market facing the same price and product characteristics, observed and unobserved. For a market, we only observe group market shares, prices, and observed characteristics of products over T periods indexed by $t = 1, \ldots, T$. Consumers choose from products $1, \ldots, J$ with 0 being the outside option.

Consumers are heterogeneous in their price coefficient α_i, which depends on discrete demographic groups and unobserved continuous heterogeneity U_i. For exposition simplicity, we limit our attention to the consumer heterogeneity in price coefficient α_i here and defer the extension to the multidimensional consumer heterogeneity until our empirical application in Section 6 with further details in Appendix D. We use a vector of dummy variables $D_i \equiv (D_i^{(1)}, \ldots, D_i^{(G)})'$ to indicate the membership—$D_i^{(g)} = 1$ if consumer i belongs to group g, and $D_i^{(g)} = 0$ otherwise. Assume $\sum_{g=1}^{G} D_i^{(g)} = 1$. Consider

$$\alpha_i = \alpha^{(1)} + \tau^{(2)} D_i^{(2)} + \cdots + \tau^{(G)} D_i^{(G)} + \omega U_i, \quad \alpha_i \sim \mathcal{N}(\alpha^{(1)} + \tau^{(g)} D_i^{(g)}, \omega)$$

where $\tau^{(g)}$ captures the *between group variation*, and ωU_i is idiosyncratic unobserved price preference, which captures the *within group variation* of the price coefficient. We normalize the variance of U_i to be 1, hence $\omega \geq 0$ controls the size of within-group variation, and $U_i \sim \mathcal{N}(0, 1)$.[9] It will be convenient to define $\tau^{(1)} = 0$. Hereafter, we say that a *consumer* i *is of type-*(g, U) if she is from group g and $U_i = U$, hence her price coefficient $\alpha_i = \alpha^{(1)} + \tau^{(g)} + \omega U$.

For a consumer i of type-(g, U), we write $v_{jt}^{(g)}(U)$ to denote her expected lifetime payoffs v_{ijt} from product j. Given the fact that consumers exit the market after the purchase of any product, a consumer's expected lifetime payoff can be written as the sum of the current period t utility and the stream of utilities in periods following purchase:

$$v_{jt}^{(g)}(U) = \frac{\delta_j + \gamma' X_{jt} + \xi_{jt}}{1 - \beta} - (\alpha^{(1)} + \tau^{(g)} + \omega U) P_{jt}, \qquad j = 1, \ldots, J. \tag{1}$$

The discount factor is $\beta \in [0, 1)$. Recall $\alpha^{(1)} + \tau^{(g)} + \omega U$ is the price coefficient of a consumer of type (g, U). To formalize the option value of "waiting" (choosing the outside option), we make the following assumptions, so a consumer's decision becomes a dynamic programming problem. The first

[9] To take account of heteroskedasticity (group varying within-group variation), it is straightforward to consider the more flexible specification $\alpha_i = \alpha^{(1)} + \tau^{(2)} D_i^{(2)} + \cdots + \tau^{(G)} D_i^{(G)} + (\omega^{(1)} D_i^{(1)} + \cdots + \omega^{(G)} D_i^{(G)}) U_i$, where $\omega^{(g)}$ controls the variation of α_i within group g.

three assumptions regulate how the observed data are generated. The last assumption specifies consumer belief's about the data generation process (DGP).

Assumption 1 (Type-I EVD). *Assume that utility shocks* ε_{ijt} *are serially independent, follow type 1 extreme value distribution (EVD), and are independent of* $(D_i, U_i, X_t, P_t, \xi_t)$.

Assumption 2 (Markov Process). $\Pr(\Omega_{i,t+1} \mid \Omega_{it}, \Omega_{i,t-1}, \ldots) = \Pr(\Omega_{i,t+1} \mid \Omega_{it})$.

Assumption 3 (Conditional Independence). *For all periods* t, *we have (i)* $\Omega_{i,t+1} \perp\!\!\!\perp \varepsilon_{it} \mid (X_t, P_t, \xi_t)$; *(ii)* $\varepsilon_{i,t+1} \perp\!\!\!\perp \Omega_{it} \mid (X_{t+1}, P_{t+1}, \xi_{t+1})$;

Assumption 4 (Consumer Belief). *(i) Consumers believe Assumptions 1 to 3; (ii) consumers have unconditional rational expectation—the unconditional expectation of* $g(X_{t+1}, P_{t+1}, \xi_{t+1})$, *where* $g(\cdot)$ *denotes a generic function, is the same as the unconditional expectation according to the DGP; (iii) when two consumers* i *and* j *have the same price coefficient, i.e.* $\alpha_i = \alpha_j$, *they two have the same belief about the conditional expectation of* $g(X_{t+1}, P_{t+1}, \xi_{t+1})$ *given* (X_t, P_t, ξ_t), *which may or may not be the same as the rational expectation.*

Additional comments about Assumption 4 are due here. Assumption 4 is weaker than assuming that consumers have the rational expectations about the DGP. One interesting feature of our method is that we can identify and estimate the flow utility functions and the discount factor under this weaker assumption of a consumer's belief. The ubiquitous rational expectation assumption implies Assumption 4.(i) and 4.(ii). Assumption 4.(iii) holds as well because all consumers have the same belief as the DGP. However, Assumption 4.(iii) does not assume that consumers believe the conditional distribution of $g(X_t, P_t, \xi_t)$ given (X_t, P_t, ξ_t) is the same as the conditional distribution in the DGP—researchers can be agnostic about such a conditional distribution as long as what consumers believe does not violate Assumption 4.(i) and 4.(ii). Taking EV price p_t for example, we do assume that consumers know $\mathrm{E}(p_{t+1})$, which is also the expectation as observed in data; we do not assume that $\mathrm{E}(p_{t+1} \mid p_t)$ according to a consumer's belief must be what we observed in the actual data.

Consider a consumer i of type (g, U), and let $V_t^{(g)}(\Omega_{it}, U)$ denote her value function.[10] Then the expected option value of waiting to purchase is

$$v_{0t}^{(g)}(U) = \beta\,\mathrm{E}\Big[V_{t+1}^{(g)}(\Omega_{i,t+1}, U) \;\Big|\; \Omega_{it}\Big] = \beta\,\mathrm{E}\Big[\bar{V}_{t+1}^{(g)}(X_{t+1}, P_{t+1}, \xi_{t+1}, U) \;\Big|\; X_t, P_t, \xi_t\Big],$$

where $\bar{V}_{t+1}^{(g)}(X_{t+1}, P_{t+1}, \xi_{t+1}, U) \equiv \mathrm{E}\Big(V_{t+1}^{(g)}(\Omega_{i,t+1}, U) \;\Big|\; X_{t+1}, P_{t+1}, \xi_{t+1}, U\Big)$. The second identity follows from applying Assumption 3.

[10] We can write the Bellman equation in terms of the value function $V_t^{(g)}(\Omega_{it}, U)$ as follows:

$$V_t^{(g)}(\Omega_{it}, U) = \max\left(\varepsilon_{i0t} + \beta\,\mathrm{E}\Big[V_{t+1}^{(g)}(\Omega_{i,t+1}, U) \;\Big|\; \Omega_{it}\Big],\; \max_{j \in \{1,\ldots,J\}} v_{jt}^{(g)}(U) + \varepsilon_{ijt}\right).$$

The first term within brackets is the present discounted utility associated with the decision to not purchase any product, i.e. choosing the outside option $j = 0$, in period t. The choice of not purchasing in period t provides flow utility ε_{i0t}, and a term that captures expected future utility conditional on the current state being Ω_{it}. This last term is the option value of waiting to purchase. The second term within brackets indicates the value associated with the purchase of a product.

Up to now, we have formalized what the payoffs $v_{i0t}, v_{i1t}, \ldots, v_{iJt}$ are. They are $v_{ijt} = v_{jt}^{(g)}(U)$ for a consumer of type (g, U). Correspondingly, the CCP of type-(g, U) is

$$\sigma_{jt}^{(g)}(U) = \frac{\exp(v_{jt}^{(g)}(U))}{\exp(v_{0t}^{(g)}(U)) + \sum_{k=1}^{J} \exp(v_{kt}^{(g)}(U))}.$$

Lastly, the market share of product j in group g in period t is

$$S_{jt}^{(g)} = \int \sigma_{jt}^{(g)}(u) \, \mathrm{d}\, F_t^{(g)}(u). \tag{2}$$

The observed "group market share" $S_{jt}^{(g)}$ is defined from averaging unobserved individual CCP $\sigma_{jt}^{(g)}(U)$ over unobserved price sensitivity U where $F_t^{(g)}(u)$ is the cumulative distribution function (CDF) of the unobserved price sensitivity U_i of consumers within group g in period t. Let $f_t^{(g)}(u)$ denote the respective probability density function (PDF) of $F_t^{(g)}(u)$. The observed group market share is either directly observed, constructed from observed aggregate sales (to different consumer segments), or constructed from individual panel data by the researchers. Since consumers leave the market after purchasing, $f_t^{(g)}(u)$ varies from group to group and over time.

Assume in period one of the market (at the launch of the market) the unobserved consumer price sensitivity follows the standard normal distribution. We then can write $f_t^{(g)}(u) = \phi(u)\mathit{\Gamma}_t^{(g)}(u)$, where $\mathit{\Gamma}_t^{(g)}(u)$ is recursively defined by the CCP $\sigma_{01}^{(g)}(u), \ldots, \sigma_{0,t-1}^{(g)}(u)$ as shown in Proposition 1. Here $\phi(u)$ denotes the PDF of the standard normal distribution, and let $\Phi(u)$ be the respective CDF.

2.2 Dynamic Selection Problem

Dynamic selection plays in an integral role in the adoption of new innovative technology such as smartphone, digital payments, smart home devices, EVs and, more generally to durable goods. It is widely discussed within the academic literature on innovation that heterogeneous consumers make adoption decision at different periods in time. One byproduct from this literature is the Rogers diffusion curve–also known today as the technology adoption curve. This curve specifically highlights that innovators and early adopters adopt first and are followed by the mass market segments of the early majority, late majority, and laggards.

The impact of the composition of consumers over time and how it depends whether or not consumers remain in the market after purchasing is an important aspect to capture in the model—i.e. on $F_t^{(g)}(U)$. For the case of non-durable goods, like ready-to-eat oatmeal, consumers remain in the market after purchasing, hence $F_t^{(g)}(u)$ does not vary across time as in a static model. In the case of durable goods, it is reasonable to assume that consumers will exit the market after purchasing. It is expected that such non-random attrition (or dynamic selection) of consumers could significantly change the distribution of unobserved price sensitivity in our model $F_t^{(g)}(u)$, depending on the rate of attrition and the length of one time period. This is especially true for new products like the electric vehicles in our empirical study. Early adopters of EV, which are pricey, are presumably less price elastic (see Palm (2020) for such behavior in green energy products and Gowrisankaran and Rysman (2012) for general durable goods). This is a "selection problem" in dynamic discrete choice. In practice, to entirely correct for dynamic selection, researchers must be in possession of data from the launch of the market and either assume there is no replacement product or that said product's value is normalized to zero; otherwise, the initial conditions problem association with dynamic selection will still remain in part. Given the necessary data, we will show that in order to fix the dynamic selection problem, it is essential to obtain the CCP $\sigma_{jt}^{(g)}(U)$ as a function of unobserved heterogeneity U.

The attrition has the following implications. First, it changes the distribution of price sensitivity U_i over the course of time even after controlling the demographic groups. It is intuitive that attrition "pushes" the distribution of U_i to concentrate more and more on the price-sensitive area over time. Second, attrition also changes the composition of groups. Attrition pushes the distribution of groups to concentrate more on price-sensitive groups—over time, we see bigger and bigger weights on price-sensitive groups. Lastly, the rate of attrition is different for different groups. Consumers in the group with lower average price elasticity would leave the market faster.

Depending on whether or not there is attrition, Proposition 1 provides a formula of the distribution of price sensitivity for the subsequent periods in terms of the CCP function $\sigma_{jt}^{(g)}(U)$.[11]

Assumption 5 (Initial distribution of unobserved price sensitivity)**.** *For each of the G groups, assume that in $t = 1$ prior to purchase, the unobserved consumer price sensitivity U follows the standard normal distribution, that is $F_1^{(g)}(u) = \Phi(u)$.*[12]

Proposition 1 (Distribution of unobserved heterogeneity due to attrition)**.** *Suppose Assumptions 1 to 5 hold. Let $f_t^{(g)}(U)$ be the PDF of the unobserved price sensitivity U in period t and group g. We have that*

$$f_t^{(g)}(u) = \phi(u) \times \mathit{\Gamma}_t^{(g)}(u),$$

where $\mathit{\Gamma}_t^{(g)}(u)$ satisfies the following.

(i) (Case I: No attrition) If consumers remain in the marker after purchasing, $\mathit{\Gamma}_t^{(g)}(u) = 1$ for all (u, t, g);

(ii) (Case II: Attrition) If consumers left the market after purchasing,

$$\mathit{\Gamma}_1^{(g)}(u) = 1, \qquad \mathit{\Gamma}_t^{(g)}(u) = \prod_{s=1}^{t-1} \frac{\sigma_{0s}^{(g)}(u)}{S_{0s}^{(g)}}, \qquad t \geq 2.$$

Note that the definition of $\mathit{\Gamma}_t^{(g)}(u)$ implies the following recursive formula:

$$\mathit{\Gamma}_1^{(g)}(u) = 1, \qquad \mathit{\Gamma}_{t+1}^{(g)}(u) = \mathit{\Gamma}_t^{(g)}(u) \times \frac{\sigma_{0t}^{(g)}(u)}{S_{0t}^{(g)}}.$$

3 Identification of CCPs

In this section we highlight from eq. (2) that $\sigma_{jt}^{(1)}(U), \ldots, \sigma_{jt}^{(G)}(U)$ are identified and can be viewed as *one* unknown CCP function after some transformations. Specifically, We have G equations from the definition

[11] Proposition B.1 in the appendix describes the variation of group composition when there is consumer attrition.

[12] Note, this assumption does not preclude the initial distribution of consumer price heterogenity at the beginning of period 1 to differ across groups as ω can be group specific. It does not also preclude other distributions from being used.

of group market shares $S_{jt}^{(1)}, \ldots, S_{jt}^{(G)}$ for a given j product in time t,

$$\begin{aligned} S_{jt}^{(1)} &= \int \sigma_{jt}^{(1)}(u)\, \mathrm{d}\, F_t^{(1)}(u), \\ &\vdots \\ S_{jt}^{(G)} &= \int \sigma_{jt}^{(G)}(u)\, \mathrm{d}\, F_t^{(G)}(u). \end{aligned} \tag{3}$$

There are G unknowns $\sigma_{jt}^{(1)}(U), \ldots, \sigma_{jt}^{(G)}(U)$ in the above G equations, for a given product j and period t and it appears hopeless for $\sigma_{jt}^{(g)}(U)$ to be identified. Below, we argue that this is not the case.

3.1 Key Observation: Shifting CCP Across Consumer Groups

Our argument is based on the following observation: these G unknown functions $\sigma_{jt}^{(1)}(U), \ldots, \sigma_{jt}^{(G)}(U)$ indeed can be viewed as the *one* unknown CCP function after some transformations. To see this, suppose there are two income brackets, 1 (high-income) and 2 (low-income), in the sample. Note that the CCP is determined by comparing the expected payoffs of different alternatives. The expected lifetime payoff of product $j = 1, \ldots, J$ is simply

$$v_{ijt} = \frac{\delta_j + \gamma' X_{jt} + \xi_{jt}}{1 - \beta} - \overbrace{\left(\alpha^{(1)} + \tau^{(2)} D_i^{(2)} + \omega U_i\right)}^{\text{Consumer Heterogeneity}} P_{jt}.$$

Moreover, the expected payoff of the outside option depends on the expected future $v_{ij,t+1}, v_{ij,t+2}, \ldots$ for all products $j = 1, \ldots, J$.[13] The unknown consumer's type affects the choice probabilities by altering the expected payoffs, which can be only through the term $\alpha_i = \alpha^{(1)} + \tau^{(2)} D_i^{(2)} + \omega U_i$ and a consumer's belief about the state transition distribution. Now consider consumer h from a high-income group 1, and consumer ℓ from a low-income group 2. Let U_h and U_ℓ be the idiosyncratic price sensitivity relative to their respective group for the two consumers h and ℓ, respectively. Quite simply, if U_h and U_ℓ satisfy the condition that $U_h = U_\ell + \tau^{(2)}/\omega$, we conclude that these two consumers have the same price coefficient hence the same conditional expectations about $v_{ij,t+1}, v_{ij,t+2}, \ldots$ according to Assumption 4.(iii). Consequently, these two consumers have the same conditional choice probabilities $\sigma_{jt}^{(2)}(U_\ell) = \sigma_{jt}^{(1)}(U_h)$.

Intuitively, this says that the shape of the conditional choice probability function for each group, given a time period t and product j, are identical. Moreover, the two CCPs $\sigma_{jt}^{(1)}(U)$ and $\sigma_{jt}^{(2)}(U)$, viewed as a function U, are essentially identical in that we can obtain one by *shifting* the other along the axis of U by $\tau^{(2)}/\omega$, i.e.

$$\sigma_{jt}^{(2)}(U) = \sigma_{jt}^{(1)}(U + \tau^{(2)}/\omega).\text{[14]}$$

Figure 1 illustrates this observation using the CCP of two groups from our simulation studies. The essential observation is that the underlying structural model implies certain restrictions that can be used

[13] This is most transparent by considering a two-period dynamic model. In period 2 (terminal period), $v_{i,j=0,t=2} = 0$. The expected optimal payoff in period 2 is $\ln(1 + \sum_j \exp(v_{ij2}))$. Then the payoff of the outside option in period 1 is $\beta\, \mathrm{E}[\ln(1 + \sum_j \exp(v_{ij2})) \mid X_1, P_1, \xi_1]$. The statement can be generalized to infinite horizon dynamic programming problems easily.

[14] This observation does not imply any extensive overlap of consumer price sensitivities between groups to hold. Below, we discuss the data required for estimation.

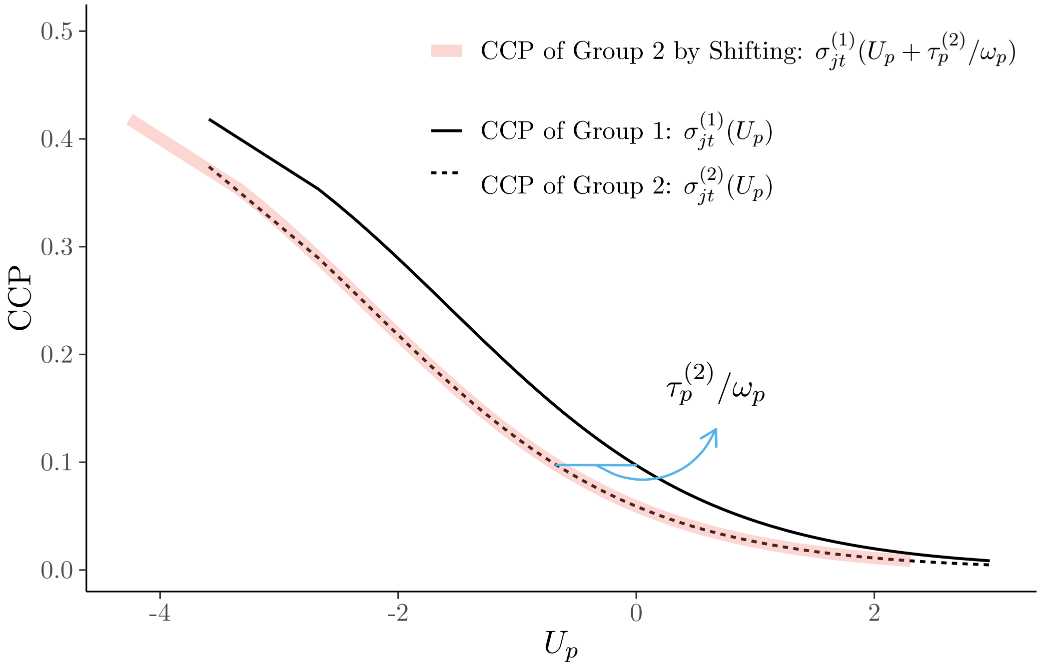

Figure 1: Shift CCP of Group 1 to Obtain CCP of Group 2

to transform the CCP functions of one focal group to get the CCP of the other groups. In addition, the underlying structural models also impose restrictions on CCP function $\sigma_{jt}^{(1)}(U)$. The application of these restrictions can be better seen after expressing $\sigma_{jt}^{(1)}(U)$ using a series multinomial logit.

We can now rewrite the G group market share equations about G unknown CCP functions at the beginning, eq. (3), as the following G equations about one unknown $\sigma_{jt}^{(1)}(U)$ by applying our shifting observation $\sigma_{jt}^{(g)}(U) = \sigma_{jt}^{(1)}(U + \tau^{(g)}/\omega)$. Recall that we defined $\tau^{(1)} = 0$ for convenience. We have

$$\begin{aligned} S_{jt}^{(1)} &= \int \sigma_{jt}^{(1)}\left(u + \frac{\tau^{(1)}}{\omega}\right) \mathrm{d}\, F_t^{(1)}(u) \\ &\vdots \qquad\qquad\qquad\qquad\qquad\qquad j = 1, \ldots, J, \quad t = 1, \ldots, T. \\ S_{jt}^{(G)} &= \int \sigma_{jt}^{(1)}\left(u + \frac{\tau^{(G)}}{\omega}\right) \mathrm{d}\, F_t^{(G)}(u) \end{aligned} \tag{4}$$

For each period, and for all J products, we now have $J \cdot G$ equations, but only $J + G$ unknowns, where J refers to the unknown CCP $\sigma_{1t}^{(1)}(U), \ldots, \sigma_{Jt}^{(1)}(U)$ for group 1, and G comes from the unknown $\tau^{(2)}, \ldots, \tau^{(G)}$ and ω, which are common for all markets and products. It is hardly a surprise that we can recover $\sigma_{jt}^{(1)}(U)$ (hence the other $\sigma_{jt}^{(2)}(U), \ldots, \sigma_{jt}^{(G)}(U)$ by shifting) from the above equations by parameterizing $\sigma_{jt}^{(1)}(U)$ (so it is known up to finite number of parameters).

4 Estimation

This section shows how to directly estimate the CCP $\sigma_{jt}^{(g)}(U)$ as a function of unobserved U. We propose a novel method of estimating $\sigma_{jt}^{(g)}(U)$ using observed group market share data and certain constraints implied by the underlying structural model. The estimation of CCPs only involve solving an NLS with linear constraints.

4.1 Series Multinomial Logit Approximation of CCP

We now discuss the details of solving the CCP $\sigma_{jt}^{(1)}(U)$ from eq. (4). The unknown $\sigma_{jt}^{(1)}(U)$ is a continuous function of scalar U, whose range is between 0 and 1, and $\sum_{j=0}^{J}\sigma_{jt}^{(1)}(U)=1$ for any U. We approximate the CCP $\sigma_{jt}^{(1)}(U)$ (as a function of U) by interpolation using a "series multinomial logit", which is a simple extension of the series logit in Hirano, Imbens and Ridder (2003). For $j=1,\ldots,J$, let

$$\sigma_{jt}^{(1)}(U;\rho_t)=L_j(R_K(U;\rho_{1t}),\ldots,R_K(U;\rho_{Jt})),$$

where L_j is a multinomial logit model,

$$L_j(c_1,\ldots,c_J)\equiv\frac{\exp(c_j)}{1+\sum_{k=1}^{J}\exp(c_k)},$$

and $R_K(U;\rho_{jt})$ is a polynomial function,

$$R_K(U;\rho_{jt})\equiv\rho_{jt1}+\rho_{jt2}U+\rho_{jt3}U^2+\cdots+\rho_{jtK}U^{K-1}.$$

Let $\rho_{jt}\equiv(\rho_{jt1},\ldots,\rho_{jtK})'$, and let ρ_t be the collection of $\rho_{1t},\ldots,\rho_{Jt}$. Lastly,

$$\sigma_{0t}^{(1)}(U;\rho_t)\equiv 1-\sum_{j=1}^{J}\sigma_{jt}^{(1)}(U;\rho_t).$$

The idea of a series (multinomial) logit is to use the power series $R_K(U;\rho_{jt})$ to approximate the log odds ratio $\ln[\sigma_{jt}^{(1)}(U)/\sigma_{0t}^{(1)}(U)]$. Let ρ be the $(K\cdot J\cdot T)\times 1$ vector from stacking ρ_t over all T periods. In practice, we found the polynomial of degree 2, i.e. $K=3$, is sufficient for approximation.

We estimate (τ,ω,ρ) using an NLS procedure below. Knowing ρ, we know group 1 CCP $\sigma_{jt}^{(1)}(U;\rho_t)$ for each alternative j and each period t. Knowing τ and ω, we know the CCP of the other groups by shifting: $\sigma_{jt}^{(g)}(U)=\sigma_{jt}^{(1)}(U+\tau^{(g)}/\omega)$.

For the rest of this section, we will explain the estimation of (τ,ω,ρ). The estimation is based on

$$\hat{S}_{jt}^{(g)}\equiv\int\sigma_{jt}^{(1)}\Big(u+\frac{\tau^{(g)}}{\omega};\rho_t\Big)f_t^{(g)}(u)\,\mathrm{d}\,u=\int\sigma_{jt}^{(1)}\Big(u+\frac{\tau^{(g)}}{\omega};\rho_t\Big)\Gamma_t^{(g)}(u)\phi(u)\,\mathrm{d}\,u,$$

where

$$\sigma_{jt}^{(1)}\Big(U+\frac{\tau^{(g)}}{\omega};\rho_t\Big)=\frac{\exp\Big[\rho_{jt1}+\rho_{jt2}\Big(U+\frac{\tau^{(g)}}{\omega}\Big)+\rho_{jt3}\Big(U+\frac{\tau^{(g)}}{\omega}\Big)^2\Big]}{1+\sum_{k=1}^{J}\exp\Big[\rho_{kt1}+\rho_{kt2}\Big(U+\frac{\tau^{(g)}}{\omega}\Big)+\rho_{kt3}\Big(U+\frac{\tau^{(g)}}{\omega}\Big)^2\Big]}.$$

The above equation can be rewritten as follows

$$\hat{S}_{jt}^{(g)}=\mathrm{E}\Big[\sigma_{jt}^{(1)}\Big(U^*+\frac{\tau^{(g)}}{\omega};\rho_t\Big)\Gamma_t^{(g)}(U^*)\Big],\qquad U^*\sim\mathcal{N}(0,1),$$

for all $j=1,\ldots,J$, $g=1,\ldots,G$, and $t=1,\ldots,T$. In practice, it is straightforward to compute the above expectation by Gauss–Hermite quadrature:

$$\mathrm{GH}_{jt}^{(g)}(\tau,\omega,\rho)\equiv\sum_{i=1}^{n}\zeta_i\cdot\Big[\sigma_{jt}^{(1)}\Big(u_i+\frac{\tau^{(g)}}{\omega};\rho_t\Big)\Gamma_t^{(g)}(u_i)\Big].$$

Here $u_1, \dots, u_n$ are n nodes, and $\zeta_1, \dots, \zeta_n$ are the respective weights. Both nodes and weights are predetermined known constants.[15] This approximation is valid only when $U^* \sim \mathcal{N}(0,1)$ as in Assumption 5. If U^* follows other distributions, the approximation will be incorrect and cause bias in CCP estimation.

We then can estimate the unknown parameters (τ, ω, ρ) by the NLS:

$$(\hat{\tau}, \hat{\omega}, \hat{\rho}) \equiv \arg\min_{\tau,\omega,\rho} \sum_{j=1,g=1,t=1}^{J,G,T} \left[S_{jt}^{(g)} - \mathrm{GH}_{jt}^{(g)}(\tau, \omega, \rho) \right]^2 \tag{5}$$

but subject to a set of linear constraints:

$$\rho_{jt2} - \rho_{1t2} = -\omega(P_{jt} - P_{1t}) \qquad \text{and} \qquad \rho_{jt3} - \rho_{1t3} = 0, \qquad j = 2, \dots, J. \tag{6}$$

In the Appendix (under the proof of Proposition 2 below), we provide the derivation of the constraints eq. (6) implied by our dynamic model.

Identification of $\tau^{(g)}$ and ω

For customers from group g, their price coefficients $\alpha_i = \alpha^{(1)} + \tau^{(g)} + \omega U_i$ follow the normal distribution $(\alpha^{(1)} + \tau^{(g)}, \omega^2)$. The estimator does not depend on how much the bell shaped PDFs of price coefficients from different groups overlap each other. The model identifies $\tau^{(g)}/\omega$ by shifting the unobserved consumer heterogeneity of group 1 until the conditional choice probability of group 1 overlays with the choice probability of group g. The distance it travels to overlay on to group g's CCP is equal to $\tau^{(g)}/\omega$ as is illustrated in Figure 1.

In order to understand the necessary requirements for identifying $\tau^{(g)}/\omega$, we can view the CCP estimation equations in eq. (5) as a GMM approach. There are $J \cdot G \cdot T$ number of moment equations with the following form

$$\mathrm{E}\left[S_{jt}^{(g)} - \sigma_{jt}^{(1)}\Big(U^* + \frac{\tau^{(g)}}{\omega}; \rho_t\Big) \mathit{\Gamma}_t^{(g)}(U^*) \right] = 0,$$

where the expectation is taken over $U^* \sim \mathcal{N}(0,1)$, and the group market shares $S_{jt}^{(g)}$ are constants. The NLS estimator in eq. (5) is the GMM estimator but with an identity weighting matrix. The shift amount $\tau^{(g)}/\omega$ as a whole is identified because only the correct shift amount will guarantee the identity of $\sigma_{jt}^{(g)}(U^*) = \sigma_{jt}^{(1)}(U^* + \tau^{(g)}/\omega)$, hence satisfy the above moment equation.

When the first-order condition of the NLS problem in eq. (5) cannot guarantee a unique solution, our method does not hold like a GMM estimator. In empirical studies, it helps to check the singularity of the Hessian matrix of eq. (5), which indicates the non-uniqueness of the optimization solution. As one example, consider the case when group 2 customers are so price sensitive relative to group 1 (i.e. $\tau^{(2)}$ is very big) that $S_{0t}^{(2)} = 1$ and $S_{jt}^{(2)} = 0$ (customers do not buy anything now). Then $\tau^{(2)}$ cannot be identified because any value greater than the true value of $\tau^{(2)}$, meaning greater price coefficient which makes product j even less desirable for group 2, will also make $S_{jt}^{(2)}$ zero for any $j \neq 0$ and $S_{0t}^{(2)} = 1$. This observation is connected to the singularity of the Hessian matrix. The argument is saying that when the

[15] To be clear, let $u_1^*, \dots, u_n^*$ be the n nodes of Gauss-Hermite quadrature, and let $\zeta_1^*, \dots, \zeta_n^*$ be the respective weights. In our simulation, we used $n = 15$ nodes. The nodes and associated weights are determined by the Hermite polynomial, and they do not depend on the function to be approximated, which is $\sigma_{jt}^{(1)}(U + \tau^{(g)}/\omega; \rho_t)$ herein. For $i = 1, \dots, n$, define $u_i = \sqrt{2}u_i^*$ and $\zeta_i = \zeta_i^*/\sqrt{\pi}$.

first-order condition with respect to $\tau^{(2)}$ evaluated at its true value is flat, i.e. the second-order condition is 0 and $\tau^{(2)}$ cannot be identified. In summary, to identify $\tau^{(g)}/\omega$, one necessary condition is that for all groups and periods, the group-level market share of at least one product is greater than 0.[16]

The variance in consumer preferences (ω) is identified from the variation in product characteristics. For example, as characteristics change for one product, substitution to products with similar characteristics indicates the presence of heterogeneity; on the other hand, if consumers substitute equally to all goods, then consumers are more homogeneous in their preferences. The panel data structure also provides another source of identification through the endogenous shift in the distribution of consumer valuations over time. If the heterogeneity in α_i is substantial, then the responses of consumers over time to price changes will be different. Heterogeneity in price sensitivity can be identified if earlier purchasers respond less to price changes than later purchasers. This identification strategy is common in dynamic discrete choice models with aggregate data (Gowrisankaran and Rysman (2012)).

4.2 Post-CCP Estimation: Structural Parameters

In order to estimate the structural linear parameters in consumer preferences, including the discount factor, one simply needs to run two linear regressions, eq. (Linear-Reg-1) and eq. (Linear-Reg-2), presented below. Moreover, we provide the intuition for why we can estimate the model while being agnostic about consumers' belief about the state transition and a remark explaining the consequence of dynamic selection on the model estimation.

Model parameters except for discount factor and product fixed effect

Identification and estimation of model parameters outside of the discount factor and product fixed effects start from the following observation. Conditional on purchasing in period t, a consumer's choice about which one to buy does not depend on the unknown continuation value $v_{0t}^{(g)}(U)$. We choose product 1 as the reference product and focus on consumer group 1, which results in

$$\ln\left[\frac{\sigma_{jt}^{(1)}(U)}{\sigma_{1t}^{(1)}(U)}\right] = v_{jt}^{(1)}(U) - v_{1t}^{(1)}(U).$$

By the definition of payoff functions $v_{jt}^{(1)}(U)$ in eq. (1), and integrating the above display over U with respect to its distribution function in period t, we have the *first condition*

$$\int \ln\left[\frac{\sigma_{jt}^{(1)}(U)}{\sigma_{1t}^{(1)}(U)}\right] \, \mathrm{d}\, F_t^{(1)}(U) + \omega(P_{jt} - P_{1t}) \int U \, \mathrm{d}\, F_t^{(1)}(U) = \\ \frac{\delta_j - \delta_1}{1-\beta} + (X_{jt} - X_{1t})' \frac{\gamma}{1-\beta} - \alpha^{(1)}(P_{jt} - P_{1t}) + \frac{\xi_{jt} - \xi_{1t}}{1-\beta}. \quad (7)$$

The next proposition, whose proof is in the Appendix, provides the interpretation of ρ_{jt1} from the power function earlier using the structural parameters. Note that the left-hand side of the above is exactly Y_{jt} defined in Proposition 2.

[16] We examine this case as well as the case of having other product characteristics that consumers positively value, in the simulation section below.

Proposition 2 (Interpretation of series logit parameters ρ_{jt1} in dynamic model)**.** *Define*

$$Y_{jt} \equiv \int \ln\left[\frac{\sigma_{jt}^{(1)}(U)}{\sigma_{1t}^{(1)}(U)}\right] \mathrm{d}\, F_t^{(1)}(U) + \omega(P_{jt} - P_{1t}) \int U \;\mathrm{d}\, F_t^{(1)}(U).$$

We have

$$Y_{jt} = \rho_{jt,1} - \rho_{1t,1}.$$

Thus, $Y_{jt} = \rho_{jt,1} - \rho_{1t,1}$ is known after CCP estimation. We conclude that

$$Y_{jt} = \frac{\delta_j - \delta_1}{1-\beta} + (X_{jt} - X_{1t})' \frac{\gamma}{1-\beta} - \alpha^{(1)}(P_{jt} - P_{1t}) + \frac{\xi_{jt} - \xi_{1t}}{1-\beta}, \qquad \text{(Linear-Reg-1)}$$

for $j = 2, \ldots, J$. *We then identify* $(\delta_2 - \delta_1)/(1-\beta), \ldots, (\delta_J - \delta_1)/(1-\beta)$*,* $\gamma/(1-\beta)$*,* $\alpha^{(1)}$ *using 2SLS and the standard Berry (1994) identification strategy.*

Model parameters: discount factor and product fixed effect

Identification and estimation of β and δ_1 originates from $\ln[\sigma_{1t}^{(g)}(U)/\sigma_{0t}^{(g)}(U)] = v_{1t}^{(g)}(U) - v_{0t}^{(g)}(U)$ and describes the trade-off from buying now and waiting. By the definition of the payoffs, it becomes

$$\ln\left[\frac{\sigma_{1t}^{(g)}(U)}{\sigma_{0t}^{(g)}(U)}\right] \;=\; \frac{\delta_1}{1-\beta} + X_{1t}' \frac{\gamma}{1-\beta} - (\alpha^{(1)} + \tau^{(g)} + \omega U) P_{1t} + \frac{\xi_{1t}}{1-\beta} \;-\; \beta\, \mathrm{E}[\bar{V}_{t+1}^{(g)}(U) \mid X_t, P_t, \xi_t]. \quad (8)$$

We will identify the discount factor β and product fixed effect δ_1 using the second condition. Before that, we will first show that for any *fixed* unobservable price sensitivity $\tilde{U}$ (e.g. $\tilde{U} = 0$),

$$\mathrm{E}[W_t^{(g)}(\tilde{U})] = \delta_1 + \beta\, \mathrm{E}[W_{t+1}^{(g)}(\tilde{U}) + \ln \sigma_{0,t+1}^{(g)}(\tilde{U})], \qquad \text{(Linear-Reg-2)}$$

where $W_t^{(g)}(\tilde{U})$ is estimable and defined below. This equation will give rise to an estimable formula of $(\beta, \delta_1)'$.

We obtain eq. (Linear-Reg-2) from eq. (8) with four steps. Note that after running 2SLS of eq. (Linear-Reg-1), we already know many parameters including $\sigma_{jt}^{(g)}(U)$, τ, ω, $\gamma/(1-\beta)$, and $\alpha^{(1)}$. *Step 1* is to define $W_t^{(g)}(U)$ by combining the terms that are already known in eq. (8). Let

$$W_t^{(g)}(U) \equiv \ln\left[\frac{\sigma_{1t}^{(g)}(U)}{\sigma_{0t}^{(g)}(U)}\right] - \left[X_{1t}' \frac{\gamma}{1-\beta} - (\alpha^{(1)} + \tau^{(g)} + \omega U) P_{1t}\right].$$

Note that $W_t^{(g)}(U)$ is *known* for any U after the CCP estimation and the above 2SLS linear regression. *Step 2* is to convert the unknown integrated value function $\bar{V}_{t+1}^{(g)}(U)$ into something we already know using the well-known expectation-maximization formula for the logit model of Arcidiacono and Miller (2011): $\bar{V}_{t+1}^{(g)}(U) = v_{1,t+1}^{(g)}(U) - \ln \sigma_{1,t+1}^{(g)}(U)$. We have

$$\bar{V}_{t+1}^{(g)}(U) = -[W_{t+1}^{(g)}(U) + \ln \sigma_{0,t+1}^{(g)}(U)] + \frac{\delta_1}{1-\beta} + \frac{\xi_{1,t+1}}{1-\beta},$$

In *step 3*, we rewrite eq. (8) in terms of $W_t^{(g)}(U)$ and conclude

$$W_t^{(g)}(U) = \delta_1 + \frac{\xi_{1t}}{1-\beta} + \beta\, \mathrm{E}\left(W_{t+1}^{(g)}(U) + \ln \sigma_{0,t+1}^{(g)}(U) - \frac{\xi_{1,t+1}}{1-\beta} \,\middle|\, X_t, P_t, \xi_t \right),$$

Lastly, in *step 4*, for a fixed unobserved price sensitivity $\tilde{U}$, we take *unconditional expectation* with respect to (X_t, P_t, ξ_t), and use the condition $\mathrm{E}(\xi_{1t}) = \mathrm{E}(\xi_{1,t+1}) = 0$ and the law of iterated expectation to reach eq. (Linear-Reg-2).

We now show how to identify the discount factor β and product fixed effect δ_1 using eq. (Linear-Reg-2). The expectations $\mathrm{E}[W_t^{(g)}(\tilde{U})]$ and $\mathrm{E}[\ln \sigma_{0t}^{(g)}(\tilde{U})]$ are taken over (X_t, P_t, ξ_t) only with $\tilde{U}$ being fixed for each group g and each t. This expectation can be estimated by $T^{-1}\sum_{t=1}^{T} W_t^{(g)}(\tilde{U})$ when (X_t, P_t, ξ_t) satisfies certain stationarity conditions.[17] With at least two groups (say 1 and 2), we have

$$\begin{aligned}
\mathrm{E}[W_t^{(1)}(\tilde{U})] &= \delta_1 + \beta\,\mathrm{E}[W_{t+1}^{(1)}(\tilde{U}) + \ln \sigma_{0,t+1}^{(1)}(\tilde{U})] \\
\mathrm{E}[W_t^{(2)}(\tilde{U})] &= \delta_1 + \beta\,\mathrm{E}[W_{t+1}^{(2)}(\tilde{U}) + \ln \sigma_{0,t+1}^{(2)}(\tilde{U})].
\end{aligned}$$

We can solve the discount factor β from the above linear system of equations, and obtain

$$\beta = \frac{\mathrm{E}[W_t^{(1)}(\tilde{U})] - \mathrm{E}[W_t^{(2)}(\tilde{U})]}{\mathrm{E}[W_{t+1}^{(1)}(\tilde{U})] - \mathrm{E}[W_{t+1}^{(2)}(\tilde{U})] + \mathrm{E}[\ln \sigma_{0,t+1}^{(1)}(\tilde{U})] - \mathrm{E}[\ln \sigma_{0,t+1}^{(2)}(\tilde{U})]}.$$

The discount factor can be estimated by the sample analog of the above formula. Knowing the discount factor β, we recover the fixed effect δ_1. The other product fixed effects $\delta_2, \ldots, \delta_J$ are automatically determined since we know $(\delta_j - \delta_1)/(1-\beta)$.

Identification of β

The identification proceeds in two steps. For product $j = 1, \ldots, J$, $v_{ijt} = \frac{\delta_j + \gamma' X_{jt} + \xi_{jt}}{1-\beta} - \alpha_i P_{jt}$. Step 1 recovers the revealed preference from consumer choices among the non-outside option products $1, \ldots, J$. This step does not identify the discount factor β as β and product preference parameters (e.g. γ) are not separated. However, this step can identify a significant portion of how consumers value a product, which lays the foundation for step 2. For example, comparing two cars with different electric range, we can identify a consumer's lifetime valuation of range, though we cannot separate how much the consumer discounts the future value of long range. In step 2, we leverage consumer choices between buying one product and not buying anything now. In this step, we can identify the discount factor. The value of buying one product today has been largely determined in step 1. The value of "not buying anything now" is the **discounted** option value of buying the best value product in the next period, and such "best value" is determined by step 1, using the product attributes in the next period.

Our model differs from those of Magnac and Thesmar (2002) and Abbring and Daljord (2020) in that we do not require an exclusion restriction to identify β. For our estimator, the identification of β comes from observing how customers trade off buying product 1 now against waiting,

$$\ln\left[\frac{\sigma_{1t}^{(g)}(U)}{\sigma_{0t}^{(g)}(U)}\right] = v_{1t}^{(g)}(U) - \beta\,\mathrm{E}[\bar{V}_{t+1}^{(g)}(U) \mid X_t, P_t, \xi_t]. \tag{9}$$

Let's assume that there are no unobserved product characteristics ξ_{jt} and $\delta_j = 0$. In this simpler case, the product value is $v_{jt}^{(g)}(U) = X_{jt}'\gamma/(1-\beta) - (\alpha^{(1)} + \tau^{(g)} + \omega U)P_{jt}$. After the CCP estimation and the 2SLS of regression eq. (Linear-Reg-1), we know $v_{jt}^{(g)}(U)$ for all periods, because $\gamma/(1-\beta)$, price

[17] We need the time series (X_t, P_t, ξ_t) is ergodic, and for the chosen fixed $\tilde{U}$, $W_t^{(g)}(\tilde{U})$, as a function of (X_t, P_t, ξ_t), satisfies certain continuity conditions.

coefficients for all groups $\alpha^{(1)} + \tau^{(g)}$, and ω are all known. We also know the integrated value function $\bar{V}^{(g)}_{t+1}(U)$ because

$$\bar{V}^{(g)}_{t+1}(U) = v^{(g)}_{1,t+1}(U) - \ln \sigma^{(g)}_{1,t+1}(U).$$

Identifying the discount factor without knowing a consumer's belief regarding the law of state transition is less transparent. The easiest way to avoid specifying the consumer belief embedded in $\mathrm{E}[\bar{V}^{(g)}_{t+1}(U)|X_t, P_t, \xi_t]$ is to eliminate the conditional expectation entirely by using the law of iterated expectation, which gives us

$$\begin{aligned}\mathrm{E}\left(\ln\left[\frac{\sigma^{(g)}_{1t}(U)}{\sigma^{(g)}_{0t}(U)}\right]\right) &= \mathrm{E}\Big(v^{(g)}_{1t}(U)\Big) - \beta\,\mathrm{E}\Big(\bar{V}^{(g)}_{t+1}(U)\Big), \quad g = 1, \ldots, G.\\ &= \mathrm{E}\Big(v^{(g)}_{1t}(U)\Big) - \beta\,\mathrm{E}\Big(v^{(g)}_{1,t+1}(U) - \ln \sigma^{(g)}_{1,t+1}(U)\Big)\end{aligned}$$

These expectations are estimable from data because we made the rational unconditional expectation assumption (Assumption 4.(ii)).[18] Note, even if we assumed rational expectation (that is $\mathrm{E}[\bar{V}^{(g)}_{t+1}(U)|X_t, P_t, \xi_t]$ according to consumer belief is just the same as its data counterpart) we can still estimate this conditional expectation and identify β.

To compare our method to those using exclusion conditions, specifically Abbring and Daljord (2020), we illustrate identification of β with the generic notation. For the simplicity of exposition, let X_t be the only observed state variable. The alternative specific value function of each choice $j \in \{0, 1, \ldots, J\}$ is $v_{jt}(X_t) = u_j(X_t) + \beta\,\mathrm{E}(\bar{V}_{t+1}(X_{t+1}) \mid X_t, j)$, where $u_j(X_t)$ is the unknown expected flow utility function. The expectation $\mathrm{E}(\cdot \mid X_t, j)$ is conditional on current state X_t and the current choice being j. It is common to normalize $u_0(X_t) = 0$ and to assume that $\bar{V}_{t+1}(X_{t+1})$ is known for the identification of the discount factor.[19] Like eq. (9), we have

$$\ln\left[\frac{\sigma_{jt}(X_t)}{\sigma_{0t}(X_t)}\right] = v_{jt} - v_{0t} = u_j(X_t) + \beta\,\mathrm{E}(\bar{V}_{t+1}(X_{t+1}) \mid X_t, j) - \beta\,\mathrm{E}(\bar{V}_{t+1}(X_{t+1}) \mid X_t, 0),$$

where $u_0(X_t)$ disappear because of the normalization. One difference from our approach is that $u_j(X_t)$ here is *unknown*, but its counterpart $v^{(g)}_{1t}(U)$ in eq. (9) is *known*. To proceed in this general model, one can consider the exclusion condition.

The exclusion restriction states that expected flow utility of choice j in state x is equivalent to the utility of choice l in state x', i.e. $u_j(X_t = x) = u_l(X_t = x')$. Consider the difference

$$\begin{aligned}\ln\left[\frac{\sigma_{jt}(X_t)}{\sigma_{0t}(X_t)}\right] - \ln\left[\frac{\sigma_{lt}(X_t)}{\sigma_{0t}(X_t)}\right] &= v_{jt}(x) - v_{lt}(x')\\ &= u_j(x) - u_l(x') + \beta\Big[\mathrm{E}(\bar{V}_{t+1}(X_{t+1}) \mid X_t = x, j) - \mathrm{E}(\bar{V}_{t+1}(X_{t+1}) \mid X_t = x', l)\Big]\\ &= \beta\Big[\mathrm{E}(\bar{V}_{t+1}(X_{t+1}) \mid X_t = x, j) - \mathrm{E}(\bar{V}_{t+1}(X_{t+1}) \mid X_t = x', l)\Big].\end{aligned}$$

Remark 1 (How does the attrition affect the estimation?)**.** *Non-random attrition affects the estimation in two ways. The first is apparent. Ignoring the attrition is to let $F^{(1)}_t(u)$ be the distribution function*

[18] The expectation is over t if the data is stationary or over m markets if multiple markets are used.

[19] For example, when the dynamic programming discrete choice model is stationary, $\bar{V}_{t+1}(X_{t+1})$ is time invariant and is the unique solution of the integral equation $\bar{V}_t(X_t) - \beta\,\mathrm{E}(\bar{V}_{t+1}(X_{t+1}) \mid X_t, 0) = u_0(X_t) - \ln \sigma_{0t}(X_t)$ (e.g. see Srisuma and Linton (2012)). Note that $\bar{V}_t(X_t)$ and $\bar{V}_{t+1}(X_{t+1})$ are the same function; $u_0(X_t)$ is normalized to 0, and $\sigma_{0t}(X_t)$ and state transition distribution are both known from data.

of U_i in the first period, i.e. $F_1^{(1)}(u) = \Phi(u)$. Reading the definition of the dependent variable Y_{jt} of eq. (Linear-Reg-1)*, it is apparent that ignoring the attrition will misspecify the $F_t^{(1)}(U)$, causing bias in estimating preference parameters. The second is more subtle. Attrition will create a nonstationarity problem, which can be clearly seen from the estimation of the discount factor.*

Reading the equation of identifying the discount factor, eq. (Linear-Reg-2)*, one may wonder why not integrate out the unobserved price sensitivity U since we also know its distribution function $F_t^{(g)}(U)$? It turns out this will lead us to a biased estimator of the discount factor in the presence of attrition. To understand why, consider $\bar{W}_t^{(g)} \equiv \int W_t^{(g)}(U)\,\mathrm{d}\,F_t^{(g)}(U)$. For a given period t, we have*

$$\bar{W}_t^{(g)} = \int \ln\left[\frac{\sigma_{1t}^{(g)}(U)}{\sigma_{0t}^{(g)}(U)}\right]\,\mathrm{d}\,F_t^{(g)}(U) - X_{1t}'\frac{\gamma}{1-\beta} + (\alpha^{(1)} + \tau^{(g)})P_{1t} + \omega P_{1t}\int U\,\mathrm{d}\,F_t^{(g)}(U),$$

by the definition of $W_t^{(g)}(U)$. The integrated term $\bar{W}_t^{(g)}$ is still estimable using our approach for each period t. It is also easy to verify that eq. (Linear-Reg-2) *becomes*

$$\mathrm{E}(\bar{W}_t^{(g)}) = \delta_1 + \beta\,\mathrm{E}\left[\bar{W}_{t+1}^{(g)} + \int \ln\sigma_{0,t+1}^{(g)}(U)\,\mathrm{d}\,F_t^{(g)}(U)\right].$$

We then have an alternative formula of the discount factor,

$$\beta = \frac{\mathrm{E}(\bar{W}_t^{(1)}) - \mathrm{E}(\bar{W}_t^{(2)})}{\mathrm{E}(\bar{W}_{t+1}^{(1)}) - \mathrm{E}(\bar{W}_{t+1}^{(2)}) + \mathrm{E}[\int \ln\sigma_{0,t+1}^{(1)}(U)\,\mathrm{d}\,F_t^{(1)}(U)] - \mathrm{E}[\int \ln\sigma_{0,t+1}^{(2)}(U)\,\mathrm{d}\,F_t^{(2)}(U)]}.$$

with at least two groups 1 *and* 2*.*

The problem is how to estimate $\mathrm{E}(\bar{W}_t^{(g)})$ and $\mathrm{E}[\int \ln\sigma_{0,t+1}^{(g)}(U)\,\mathrm{d}\,F_t^{(g)}(U)]$? Taking $\mathrm{E}(\bar{W}_t^{(g)})$ for example, it is tempting to use $T^{-1}\sum_{t=1}^{T}\bar{W}_t^{(g)}$ as the estimator, however, it is an inconsistent estimator when there is non-random attrition of consumers. The underlying reason is that even though (X_t, P_t, ξ_t) satisfies certain stationarity conditions, $\bar{W}_t^{(g)}$ is still non-stationary due to the attrition of consumers. In particular, both

$$\int \ln\left[\frac{\sigma_{1t}^{(g)}(U)}{\sigma_{0t}^{(g)}(U)}\right]\,\mathrm{d}\,F_t^{(g)}(U) \qquad \text{and} \qquad \int U\,\mathrm{d}\,F_t^{(g)}(U)$$

in the definition of $\bar{W}_t^{(g)}$ are nonstationary. Intuitively, consumers who are less price sensitive purchase and leave the market earlier making the average $\int U\,\mathrm{d}\,F_t^{(g)}(U)$ drift upward over time. Due to this non-stationary property (caused by attrition), the temporal average will not converge in probability to $\mathrm{E}(\bar{W}_t^{(g)})$, which is indeed only well defined for a fixed period. In order to estimate $\mathrm{E}(\bar{W}_t^{(g)})$, one needs access to a large number of cross-sectional markets for each period. Such data access is usually unavailable in empirical studies. The same comments apply to the integral term of CCP functions.

Our approach works here because we can recover the CCP function at such a precise level that the CCP for any given unobserved price sensitivity, i.e. $\sigma_{jt}^{(g)}(\tilde{U})$ herein, can be obtained. We then can avoid the problem of attrition by focusing on one type of consumer (in terms of fixing U). After fixing U, all variables, like $W_t^{(g)}(U)$, involve only stationary process (X_t, P_t, ξ_t). We then can use the temporal average to estimate them and the discount factor. This again highlights the importance of recovering the CCP $\sigma_{jt}^{(g)}(U)$ in order to fix the dynamic selection problem.

4.3 Estimation Summary

Below we provide a summary of the estimation routine.

1. Estimate $\tau^{(g)}$, ω and ρ via NLS

$$(\hat{\tau}, \hat{\omega}, \hat{\rho}) \equiv \arg\min_{\tau,\omega,\rho} \sum_{j=1,g=1,t=1}^{J,G,T} \left[S_{jt}^{(g)} - \mathrm{GH}_{jt}^{(g)}(\tau, \omega, \rho)\right]^2$$

subject to

$$\rho_{jt2} - \rho_{1t2} = -\omega(P_{jt} - P_{1t}) \qquad \text{and} \qquad \rho_{jt3} - \rho_{1t3} = 0, \qquad j = 2, \ldots, J,$$

where

$$\mathrm{GH}_{jt}^{(g)}(\tau, \omega, \rho) \equiv \sum_{i=1}^{n} \zeta_i \cdot \left[\sigma_{jt}^{(1)}\left(u_i + \frac{\tau^{(g)}}{\omega}; \rho_t\right) \mathit{\Gamma}_t^{(g)}(u_i)\right],$$

and

$$\sigma_{jt}^{(1)}\left(U + \frac{\tau^{(g)}}{\omega}; \rho_t\right) = \frac{\exp\left[\rho_{jt1} + \rho_{jt2}\left(U + \frac{\tau^{(g)}}{\omega}\right) + \rho_{jt3}\left(U + \frac{\tau^{(g)}}{\omega}\right)^2\right]}{1 + \sum_{k=1}^{J} \exp\left[\rho_{kt1} + \rho_{kt2}\left(U + \frac{\tau^{(g)}}{\omega}\right) + \rho_{kt3}\left(U + \frac{\tau^{(g)}}{\omega}\right)^2\right]}.$$

2. Form $Y_{jt} = \rho_{jt,1} - \rho_{1t,1}$.

3. Estimate $(\delta_2 - \delta_1)/(1-\beta), \ldots, (\delta_J - \delta_1)/(1-\beta)$, $\gamma/(1-\beta)$, $\alpha^{(1)}$ using 2SLS

$$Y_{jt} = \frac{\delta_j - \delta_1}{1-\beta} + (X_{jt} - X_{1t})' \frac{\gamma}{1-\beta} - \alpha^{(1)}(P_{jt} - P_{1t}) + \frac{\xi_{jt} - \xi_{1t}}{1-\beta}.$$

4. Estimate β: for $\tilde{U} = 0$ calculate (if β is specified skip to step 5)

$$\beta = \frac{\mathrm{E}[W_t^{(1)}(\tilde{U})] - \mathrm{E}[W_t^{(2)}(\tilde{U})]}{\mathrm{E}[W_{t+1}^{(1)}(\tilde{U})] - \mathrm{E}[W_{t+1}^{(2)}(\tilde{U})] + \mathrm{E}[\ln \sigma_{0,t+1}^{(1)}(\tilde{U})] - \mathrm{E}[\ln \sigma_{0,t+1}^{(2)}(\tilde{U})]}.$$

5. Estimate δ_1: for $\tilde{U} = 0$ calculate

$$\delta_1 = \mathrm{E}[W_t^{(1)}(\tilde{U})] - \beta\, \mathrm{E}[W_{t+1}^{(1)}(\tilde{U}) + \ln \sigma_{0,t+1}^{(1)}(\tilde{U})].$$

5 Simulation

In order to determine how well our estimator performs in small samples, we run several simulations that vary the degree of within-group variation, across group heterogeneity, and allow for the degree of within-group variation to vary across groups. We designed our numerical experiments to illustrate the applicability of our estimator and to understand the following empirically relevant questions:

(i) How does little to no overlap in cross-group consumer price heterogeneity distribution affect the estimation?

(ii) How does the outside good's market share equal 1 affect the estimation?

(iii) How does heteroscedasticity in ω affect the estimation?

We address each question in the results section 5.2 below. We make section 5.2 self-contained so that readers, who are not interested in the data generating process (DGP) details, can skip the DGP section and jump to the results. Lastly, in the online appendix, we report the simulation study when consumers have multi-dimensional heterogeneity.

5.1 Data Generating Process

In our DGP, the flow utility function follows the specification in Section 2.1. When consumer i of group g purchases product j in period t in market m, she receives the following utility

$$u_{ijtm} = \frac{f(X_{jtm}, \xi_{jtm})}{1-\beta} - \alpha_i P_{jtm} + \varepsilon_{ijtm},$$

and receives $f(X_{jtm}, \xi_{jtm})$ as flow utility in each period post-purchase in period t where

$$\alpha_i = \alpha^{(1)} + \tau^{(2)} D_i^{(2)} + \cdots + \tau^{(G)} D_i^{(G)} + \omega U_i.$$

In all the simulations below (except where noted) we let

$$f(X_{jtm}, \xi_{jtm}) = \delta_j + X_{jtm}\gamma + \xi_{jtm} = -0.1 + X_{jtm} \cdot 0.03 + \xi_{jtm},$$

for any product j. Thus, $\gamma = 0.03$ and $\delta_j = -0.1$ for any product j. For price coefficient α_i, let $\alpha^{(1)} = 0.1$, $\tau^{(2)} = 0.12$, $\tau^{(3)} = 0.24$, the within-group variation ω will take one value from $(0, 0.01, 0.03)$, and let U_i be a random variable drawn from the standard normal distribution. Products are differentiated by the observed price, P_{jtm}, observed product characteristic X_{jtm}, and unobserved characteristics, ξ_{jtm}. The discount factor β is set to 0.90.

We next describe the data generation process of price, X_{jtm}, and the unobserved product characteristics. We specifically account for the correlation between ξ_{jtm} and P_{jtm}. Such a formulation is motivated by the price endogeneity problem researchers face when employing aggregate data, where firms can observe ξ_{jtm} and then set prices optimally. In practice, we allow for multiple markets where $M = 2$. We use a reduced-form price model to characterize this dependence. Specifically,

$$\begin{aligned} X_{jtm} &= r_m + \phi_m^x X_{j,t-1,m} + \nu_{jtm}^x, \\ \xi_{jtm} &= \phi^\xi \xi_{j,t-1,m} + \nu_{jtm}^\xi, \\ P_{jtm} &= c + MC_{jtm} + \nu_{jtm}^p, \\ MC_{jtm} &= d_j + \phi_j^{MC} MC_{j,t-1,m} + \nu_{jtm}^{MC}, \end{aligned}$$

where $(\nu_{jtm}^x, \nu_{jtm}^\xi, \nu_{jtm}^p, \nu_{jtm}^{MC})'$ is independent and identically distributed across products, time periods and markets, and follows a multivariate normal distribution,

$$\begin{pmatrix} \nu_{jtm}^x \\ \nu_{jtm}^\xi \\ \nu_{jtm}^p \\ \nu_{jtm}^{MC} \end{pmatrix} \sim \mathcal{N}\left(0, \begin{pmatrix} \sigma_x^2 & 0 & 0 & 0 \\ 0 & \sigma_\xi^2 & \rho\sigma_\xi\sigma_p & 0 \\ 0 & \rho\sigma_p\sigma_\xi & \sigma_p^2 & 0 \\ 0 & 0 & 0 & \sigma_{MC}^2 \end{pmatrix}\right).$$

Here MC_{jtm} denotes the marginal cost of product j at time t in market m. We will use MC_{jtm} as the instrumental variable in estimation.

In our simulations, the number of products is 8, and we assign the following parameter values. We let $c = 3$, $(d_1, \ldots, d_8) = (0.21, 0.28, 0.35, 0.42, 0.49, 0.56, 0.63, 0.7)$, $(r_{m=1}, r_{m=2}) = (0.35, 0.55)$, $\phi^{\xi} = 0$, $(\phi_1^{MC}, \ldots, \phi_8^{MC}) = (0.965, 0.94, 0.925, 0.91, 0.895, 0.88, 0.865, 0.85)$ and $(\phi^x_{m=1}, \phi^x_{m=2}) = (0.35, 0.55)$. For the initial state of MC_{j0m}, we let $(MC_{1,0,m}, \ldots, MC_{8,0,m}) = (9.5, 9.25, 9.00, 8.75, 8.50, 8.25, 8.00, 7.75)$. Such specification ensures that product marginal cost, MC_{jtm}, has a declining trajectory, which is consistent with durable goods models. As for the X variable, the initial starting values do not differ across j, but do so across markets with $(X_{0,m=1}, X_{0,m=2}) = (0.525, 0.825)$. Finally, we let $\sigma_x = 0.15$, $\sigma_\xi = 0.05$, $\sigma_p = 0.25$, $\sigma_{MC} = 0.1$ and $\rho = 1$.

It is important to note the specified DGP produces static own-price elasticities (when all 8 goods are available) in the range of -1 for type 1 consumers to -4.25 for type 3 in period 1.[20] In addition, each set of simulation results is based on 50 replications.

5.2 Results

Effects of little to no overlap in price heterogeneity

In Table 1, we present the results of a Monte Carlo simulation where we specifically vary ω from 0 to 0.03. When $\omega = 0$ there is no within group heterogeneity. Moreover, as ω increases within group heterogeneity increases but given the measures of $\tau^{(g)}$ there is little overlap across groups. For example, when $\omega = 0.01$, consumers who are above the fifth standard deviation of the mean price sensitivity will overlap with users from another group (when $\omega = 0.03$ the cutoff lowers to fourth standard deviation).

We see the estimator is able to recover model primitives across all values of ω. In the extreme case when there is no within group price heterogeneity, the model preforms well and pins down all $\tau^{(g)}$ and ω. Thus, as discussed above in the identification of τ and ω, identification of $\tau^{(g)}$ are due to the shifting of CCPs and not the overlap of consumer price heterogeneity across groups.

Biased estimation when group 3's outside good market share=1

The next set of simulations expands on Table 1 to illustrate when the estimator fails. As we discussed above, as long as there are consumers in the market with positive market shares for J goods, the model can find a consumer in group 2 or 3 who has the same CCP as the consumer in group 1 by changing the consumer price sensitivity in group 1 by τ^g/ω for group 2 or 3. To showcase such, we fix within group consumer heterogeneity ($\omega = 0.03$) and vary τ^3 from $0.50 - 1.50$, resulting in group three's outside good's market share to vary from 0.9826 to 1 in Table 2. The simulations illustrate that the estimator breaks down when a specific group does not participate in the market, leading to the group's conditional choice probability for the outside option being equal to 1. This is due to the estimator not being able to uniquely shift group 1's CCP to overlay with group 3 given $\sigma^3_{jt}(U) = 0$ for all values of U.

[20] This is approximately close to the magnitudes estimated in our empirical application which varies from -4 to -10 across consumer types, markets and vehicles in period 1

Table 1: Simulation Results: Comparison Across Within Group Heterogeneity with Little to No Overlap
DGP: $M = 2$, $T = 36$ and $J = 8$

	$\omega = 0$	$\omega = 0.01$	$\omega = 0.03$
$\delta = -0.1$	-0.1024 (0.0030)	-0.1024 (0.0030)	-0.1018 (0.0030)
$\gamma = 0.03$	0.0306 (0.0020)	0.0306 (0.0020)	0.0305 (0.0020)
$\alpha^{(1)} = 0.10$	0.0999 (0.0027)	0.0999 (0.0027)	0.0996 (0.0027)
$\tau^{(2)} = 0.12$	0.1200 (5.51e-6)	0.1200 (1.94e-6)	0.1200 (7.39e-6)
$\tau^{(3)} = 0.24$	0.2401 (1.26e-5)	0.2401 (3.86e-6)	0.2403 (1.26e-5)
ω	0.0029 (2.84e-4)	0.0105 (2.57e-5)	0.0310 (3.21e-5)
$\beta = 0.90$	0.8989 (3.97e-4)	0.8989 (3.97e-4)	0.8992 (3.95e-4)

Note: Mean and standard deviation (in parenthesis) for 50 simulations.
Starting values follow the procedure in the appendix and vary with each simulation run.
The starting value for $\omega = 0.25$ for all simulation runs.

Table 2: Simulation Results: Comparison of Across Group Heterogeneity with No Overlap as $\sigma_0^3 \to 1$
DGP: $M = 2$, $T = 36$, $\omega = 0.03$, and $J = 8$

	$\tau^{(3)} = 0.50$	$\tau^{(3)} = 0.75$	$\tau^{(3)} = 1$	$\tau^{(3)} = 1.25$	$\tau^{(3)} = 1.50$
$\delta = -0.1$	-0.0994 (0.0030)	-0.0998 (0.0030)	-0.1303 (0.0051)	-0.1995 (0.0197)	-0.4929 (9.2528)
$\gamma = 0.03$	0.0303 (0.0020)	0.0303 (0.0020)	0.0320 (0.0022)	0.0359 (0.0029)	0.0345 (0.8640)
$\alpha^{(1)} = 0.10$	0.0985 (0.0027)	0.0990 (0.0027)	0.1003 (0.0027)	0.1009 (0.0027)	0.0968 (0.0033)
$\tau^{(2)} = 0.12$	0.1205 (3.11e-5)	0.1204 (4.17e-5)	0.1199 (3.27e-5)	0.1197 (6.7e-5)	0.1211 (4.90e-4)
$\tau^{(3)}$	0.5031 (1.68e-4)	0.7535 (2.49e-4)	1.0054 (0.0011)	1.2589 (0.0117)	1.3094 (0.0323)
$\omega = 0.03$	0.0333 (1.65e-4)	0.0321 (2.27e-4)	0.0292 (2.06e-4)	0.0276 (5.08e-4)	0.0367 (0.0029)
$\beta = 0.90$	0.9001 (3.29e-4)	0.9001 (3.08e-4)	0.8942 (8.86e-4)	0.8813 (0.0043)	0.8596 (2.9072)
$\bar{\sigma}_0^{(3)}$	0.9826	0.9972	0.9996	0.9999	1

Note: Mean and standard deviation (in parenthesis) for 50 simulations. Starting values follow the procedure in the appendix and vary with each simulation run.The starting value for $\omega = 0.25$ for all simulation runs.

Table 3: Simulation Results: Heteroscedastic ω

	$\omega_1 = 0.03$ $\omega_2 = \omega_3 = 0.015$	$\omega_1 = \omega_3 = 0.015$ $\omega_2 = 0.03$
$\delta = -0.1$	-0.1022 (0.0030)	-0.1021 (0.0030)
$\gamma = 0.03$	0.0306 (0.0020)	0.0306 (0.0020)
$\alpha^{(1)} = 0.10$	0.0996 (0.0027)	0.0996 (0.0027)
$\tau^{(2)} = 0.12$	0.1201 (7.55e-5)	0.1201 (7.05e-5)
$\tau^{(3)} = 0.24$	0.2403 (1.93e-4)	0.2403 (1.81e-4)
ω_1	0.0311 (4.67e-4)	0.0164 (7.88e-4)
ω_2	0.0167 (9.35e-4)	0.0307 (5.07e-4)
ω_3	0.0173 (0.0011)	0.0168 (0.0011)
$\beta = 0.90$	0.8989 (3.94e-4)	0.8990 (3.94e-4)

Note: Mean and standard deviation (in parenthesis) for 50 simulations.

Starting values follow the procedure in the appendix and vary with each simulation run.

The starting value for $\omega_g = 0.25$ for all simulation runs.

Heteroscedasticity in ω

With this set of Monte Carlo simultions we illustrate the robustness of our estimator when we allow for differing levels of within group heterogeneity. We proceed with the above parameterization and allow ω to vary across groups. Let ω_g denote the standard deviation of price coefficient within the customers in group g. In the first simulation, we allow ω_1 to differ from $\omega_2 = \omega_3$. The second set of simulations allows for $\omega_1 = \omega_3$ and have ω_2 to differ. The results are presented in Table 3 and illustrate the model is able to estimate model parameters without bias with heterscedasticity in ω.

6 Empirical Application: Demand of Electric Vehicles

We illustrate the empirical value of our method by estimating consumer demand for electric vehicles in the state of Washington during the period of 2016–2019. We further ascertain the impact of a different federal tax credit based on a car's electric range rather than battery size, which was the existing policy during the data period. We also evaluate an infrastructure focused policy where the existing consumer tax credits when to government or third party installers of Level 3 DC charging stations.

6.1 Data

The main data originate from new EV registration records from the Washington State Department of Licensing. The electric vehicles include battery electric vehicles (BEV) and plug-in hybrid electric vehicles (PHEV). For each vehicle, we observe the first ten digits of the vehicle identification number (VIN), from which we collect vehicle characteristics by using a VIN decoder. In the registration records, we also observe the ZIP code of an owner's residence, with which we obtain the Internal Revenue Service (IRS) ZIP code household income of the owner.

In the data, a product corresponds to a model car such as the Nissan Leaf. We excluded the models with extremely small market share such as the Smart Fortwo EV. The models we included in the sample account for more than 99% of the total sales of EVs in Washington. If two trims of a model are very similar in terms of fuel efficiency and battery range (like BMW 530E and BMW 530E AWD), we collapse them as one model and use their average characteristics as the product characteristics of the collapsed model. In total, we have 29 distinct models from 17 makers.

Our dynamic discrete demand model below describes how a few groups of consumers, who reside in different markets, make discrete choices about purchasing EVs over a number of periods. In the data, a geographical market is a county, one time period is a quarter of a year, and the group of consumers is defined based on IRS ZIP code household income. There are three geographical markets in the data: King County, Snohomish County, and Pierce County. Together, these counties account for 77% of the EV market in Washington state. We observed these markets from quarter 1 (Q1) of 2016 to quarter 4 (Q4) of 2019. There are three income groups in the data. These income groups are based on IRS ZIP code household income in 2016. Groups 1, 2, and 3 consist of households with ZIP code incomes above the 90-th percentile ($100,966), between the 75-th ($74,353) and 90-th percentile, and below the 75-th percentile, respectively.

Our data contain only the sales of electric vehicles. We do not observe the sales of conventional internal combustion vehicles. Consequently, we define the market as consumers who only consider EVs for purchase. One may think that this would create a potential problem in our model specification where the outside option in a period is to delay the purchase rather than to purchase a gasoline car. This potential issue is mitigated thanks to the 2015-2017 California Vehicle Report.[21] According to the report, 38% of the respondents intended to only purchase an EV as their next vehicle for their household, while 62% only consider an internal combustion engine vehicle.[22] Additionally, Figure 2 illustrates the similarities in adoption rates across the states of California and Washington by county. These surveys, we believe, provide strong support for our market definition. Therefore, we let the initial market size be half of the number of households in that income group and county in 2016.

We recognize that the definition of the size of the potential market is an important aspect in estimating consumer preferences. In our method, the model primitives are estimated via two separate channels: $\ln[\sigma_{jmt}^{(g)}(U)/\sigma_{1mt}^{(g)}(U)]$ (related to eq. (7))and $\ln[\sigma_{1mt}^{(g)}(U)/\sigma_{0mt}^{(g)}(U)]$ (related to eq. (8)).The first ratio $\sigma_{jmt}^{(g)}(U)/\sigma_{1mt}^{(g)}(U) = \frac{Sales_{jmt}^{(g)}(U)}{Sales_{1mt}^{(g)}(U)}$ concerns only the sales of the two car models j and 1, and it does not depend on the market size. So the model primitives including the "life-time" preference of car characteristics ($\gamma/(1-\beta)$) and price coefficient ($\alpha^{(1)}$) are not impacted by the specification of market size. An incorrectly specified market size, however, will impact the second ratio $\sigma_{1mt}^{(g)}(U)/\sigma_{0mt}^{(g)}(U)$, because it relates to the sales of car model 1 and the number of customers in the market who do not purchase in period t. Consequently, product fixed effect term for product 1, the discount factor etc. will be affected. In our case, we believe that if the market size is incorrect, we overestimate it, leading to a larger market share for product 0 in all periods. This will likely overestimate the discount factor.

Table 4 shows the initial market size of each group. In the estimation strategy, we frequently use the term "group market share" of a product. Here we use Nissan Leaf as an example to explain how such

[21] No such detailed survey data was found for Washington State and so we use California data as a proxy for Washington.
[22] https://www.energy.ca.gov/sites/default/files/2023-02/CEC-200-2018-006.pdf

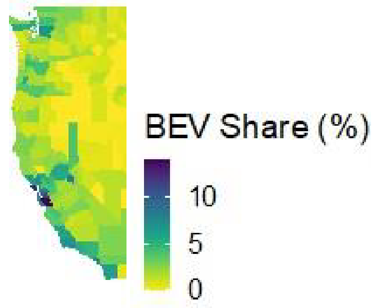

Figure 2: County-level market share of battery electric-vehicle sales in 2020 from Gillingham et al. (2023)

Table 4: Initial (Q1 of 2016) Market Size by Zip Income Group and County

	Group 1 >90-th Percentile	Group 2 75-th ~ 90-th Percentile	Group 3 < 75-th Percentile
King County	238,750	132,175	165,085
Snohomish County	24,275	50,000	110,480
Pierce County	11,715	29,935	155,440

a "group market share" was constructed. In Q1 of 2016, the market size of the lowest income group in King County was 165,085, and the number of sales of Nissan Leaf in this income group and King County during Q1 of 2016 was 17. Then the market share of Nissan Leaf in the lowest income group in King County in Q1 of 2016 is $17/165,085$.

We obtain additional data from various sources. The gasoline prices by county and period are from the Cost of Living Index. The number of electric charging stations in each county and period is from the Alternative Fuels Data Center in the US Department of Energy. The number of public facilities in each county and period, which will be used as an instrumental variable of the number of charging stations, is from County Business Patterns from the US Census Bureau. The price is mostly from the transaction price in the registration data. In the registration data, 25.3% of the vehicles do not have valid transaction prices. This is due to the random incomplete submission of the registration data to the state which we believe is not systematic. When the transaction price is unavailable, we use the list price after the deduction of all federal and state-level tax incentives. The available rebates for each model car and the amount of total CO2 emission per model car in tons per year including the production of the electricity are from the US Department of Energy.[23]

Table 5 reports the means of key vehicle characteristics and the sales in different income groups. We note that because electric cars are expensive even after a variety of government financial incentives,[24]

[23]Pre-owned vehicles purchased before 2023 don't qualify for a used ($4,000) credit. https://www.fueleconomy.gov/feg/taxused.shtml

[24]The average new car prices in 2016 and 2019 were $34,077 and $36,718, according to Edmunds.

Table 5: Means of Key EV Characteristics

Variable	PHEV	BEV	2016	2019
Price (thousand of $)	42.7	48.0	42.9	47.0
	(17.3)	(28.7)	(21.9)	(23.2)
Federal Tax Credit ($)	4831	6789	6015	5312
	(1354)	(1633)	(1579)	(1817)
Battery Size (kWh)	21.1	155.2	69.6	107.5
	(8.75)	(73.59)	(67.25)	(97.92)
Electric Range (miles)	21.1	155.2	69.6	107.5
	(8.75)	(73.59)	(67.25)	(97.92)
Miles per gallon equivalent (MPGe)	48.7	100.9	72.4	75.1
	(18.4)	(14.8)	(31.0)	(31.0)
Horsepower	213	236	212	245
	(88.1)	(138.3)	(89.5)	(136.4)
Total Tailpipe CO2 emission per car (ton/year)	3.744	1.506	2.750	2.580
	(1.278)	(0.279)	(1.520)	(1.448)
Sales: all groups	5318	24921	4422	9120
% sales: income below 75th percentile	28.1	17.1	16.7	20.1
% sales: income between 75th and 90th percentile	21.4	21.4	21.1	21.7
% sales: income greater 90th percentile	47.6	61.5	62.2	58.2

Note: CO2 emission of PHEV is based on 15,000 miles driving distance a year, and 37% electric driving in real life (Plötz et al., 2020). The standard deviation is in the parenthesis.

buying an EV is more popular among affluent Americans. The top 10% of consumers in terms of income own 47.6% of PHEVs and 61.5% of BEV, which is more expensive than PHEVs on average, and they own 62.2% and 58.2% of electric cars in 2016 and 2019, respectively.

Though it is possible to study the demand for EVs using the familiar static BLP model (Berry, Levinsohn and Pakes, 1995), we found that the EV market features a variety of market dynamics that are relevant to a forward-looking consumer's car buying decisions and better understood using a dynamic discrete demand framework. Using our data of EV sales we plot the average electric range of EVs in miles, the number of charging stations in the state, the average federal tax credit, and sales across time in Figure 3. We observe a substantial improvement in electric range and rapid deployment of EV charging stations over time. On the other hand, we see a declining federal tax credit in Figure 3. This is because the tax credit is phased out over time after a manufacturer reaches a total sales (200,000) milestone since 2010.[25] We also see a decline in EV sales in Q1 of 2019 that is caused by the state of Washington removing its sales and use tax exemption for EV vehicles in Q3 of 2018. Later in Q3 of 2019, the state reinstalled that exemption. Intuitively, the shrinking tax credit incentivizes consumers to buy EVs earlier than later. Additionally, we should highlight that no tax incentives were provided for used EVs during

[25]Taking Tesla Model 3 for example, its credit was $7,500 before 12/31/2018, then $3,750 from 1/1/2019 to 6/30/2019, and lastly $1,875 from 7/1/2019 to 12/31/2019.

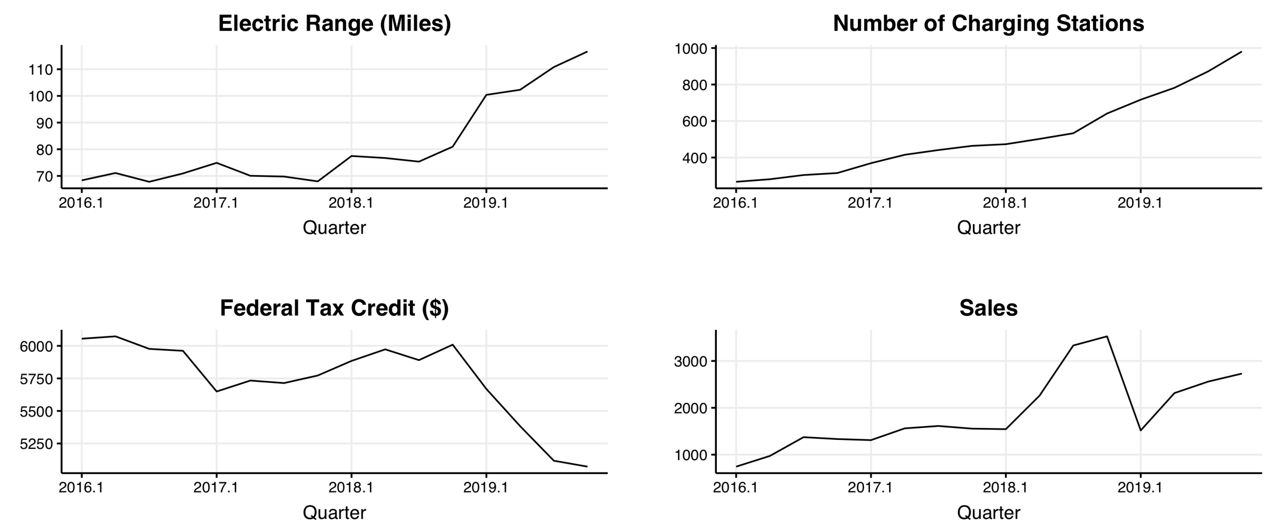

Figure 3: Time Series of Key Variables in the EV Market

our data period. It was not until the Inflation Reduction Act of 2023 that used EVs were eligible for EV tax credits. On a different note, a growing electric range and a network of charging stations motivate delaying an EV purchase. It is important to incorporate these possible dynamic considerations when modeling consumer choice about purchasing an EV.

6.2 Consumer Dynamic Demand Model for EV

Our empirical model closely follows the theoretical framework with a few changes to address specific empirical issues. Most notable is the inclusion of multidimensional unobserved heterogeneity and model characteristics that evolve post-purchase and impact future flow utilities. It is also important to highlight upfront that we do not model EV purchasing as a replacement problem akin to an iPhone upgrading each year due to the lack of initial vehicle data. That said, we further support the use of the above model with additional evidence. While we acknowledge that EV technology is quickly evolving during the studied time frame, the duration a consumer holds/owns a car is longer than the four years of data we possess–the average is 11.6 years.[26] Moreover, for the few EVs sold before our data period, the length of ownership would likely take it beyond our observed data period. Lastly, we further assume that the replacement value associated with an individual's current automobile at the time of EV consideration is normalized to zero. We believe this is a reasonable assumption since the average resale value of a car declines "approximately 20 percent of its value in the first year and 15 percent each year after that until, after 10 years, it's worth around 10 percent of what it originally cost".[27] To further support this assumption, note the average price of a new car sold in 2006 was $\$22,500$ leaving a depreciated value of $\$2,250$ in 2016.[28] Because we do not have individual level data, the heterogeneous value of one's current vehicle can only be treated as a part of ε_{i0t}.

[26] https://www.bts.gov/content/average-age-automobiles-and-trucks-operation-united-states

[27] https://smallbusiness.chron.com/average-depreciation-rate-vehicle-64934.html

[28] https://www.aei.org/carpe-diem/new-real-car-prices-fell-by-2500-from-1998-2006/

Consumers are heterogeneous in price and MPGe coefficients. Similar to MPG for gasoline automobiles which tracks fuel efficiency, we allow preferences for "fuel" efficient vehicles to vary across consumers for electric vehicles through MPGe.

Individual heterogeneity is determined by income group membership $D_i^{(g)}$ and continuous heterogeneity $U_i = (U_{i,mpge}, U_{i,p})'$. We let

$$\begin{pmatrix} \eta_i \\ \alpha_i \end{pmatrix} = \begin{pmatrix} \eta^{(1)} \\ \alpha^{(1)} \end{pmatrix} + D_i^{(2)} \begin{pmatrix} \tau_{mpge}^{(2)} \\ \tau_p^{(2)} \end{pmatrix} + \cdots + D_i^{(G)} \begin{pmatrix} \tau_{mpge}^{(G)} \\ \tau_p^{(G)} \end{pmatrix} + \Omega \begin{pmatrix} U_{i,mpge} \\ U_{i,p} \end{pmatrix},$$

where Ω is a diagonal matrix with the diagonal elements ω_{mpge} and ω_p. Assume that U_i follows the standard bivariate normal distribution. A consumer i is said to be of type (g, U) with $U = (U_{mpge}, U_p)'$ when $\eta_i = \eta^{(1)} + \tau_{mpg}^{(g)} + U_{mpge}$ and $\alpha_i = \alpha^{(1)} + \tau_p^{(g)} + U_p$. Let $\tau = (\tau_{mpge}^{(2)}, \ldots, \tau_{mpge}^{(G)}, \tau_p^{(2)}, \ldots, \tau_p^{(G)})'$ and $\omega = (\omega_{mpge}, \omega_p)'$.

The expected lifetime payoff is a "sum" of expected discounted per period or flow utilities. We first state the flow utilities for a consumer. If consumer i does not purchase in period t, she receives the flow utility ε_{i0t} in period t and stays in the market. When consumer i purchases EV j at time t, her indirect flow utility *during the purchase period t* is

$$f_{ijt} = \sum_{k=1}^{K} \delta_{make,k} Make_{jk} + \gamma_1' X_{jt}^{ev} + \gamma_2' X_t^{cnty} + \xi_{jt} + \eta_i MPGe_{jt} - \alpha_i P_{jt} + \varepsilon_{ijt}$$

where f_{ijt} is the flow utility in purchase period t. Here $Make_{jk}$ is a dummy variable for make $k = 1, \ldots, K$. If EV j is manufactured by make k, $Make_{jk} = 1$ and $= 0$ otherwise. So $\delta_{make,k}$ is the make fixed effect. There are $K = 17$ makes in the sample. We let the Nissan Leaf be the base product 1, because it appears in all markets, time periods, and groups. Also, let Nissan be make 1. X_{jt}^{ev} consists of various vehicle characteristics including horsepower, natural log of electric range, a dummy variable for BEV, a dummy variable of sport utility vehicle (SUV), the interaction between BEV and SUV indicator, the interaction term between the dummy variable of BEV and the gasoline price in the market, and the interaction term between the dummy variable of BEV and the number of charging stations in the market. We assume the value of the number of EV charging stations in a county is normalized to zero for PHEVs as an electric charging network does not provide any utility given its gas consumption and small battery pack. The vector X_t^{cnty} consists of the county market characteristic, gasoline price. P_{jt} is defined similarly to the discussion above but includes all tax incentives as immediate price discounts.[29]

As for ξ_{jt}, we refer to ξ_{jt} as the unobserved characteristics of product j at time t. A possible interpretation of these unobservable product characteristics is product quality. "If the firm has quality control in the production process, then there is likely some degree of randomness or stochasticity in the manufacturing process. This would vary by product period and fit the assumptions about ξ_{jt}" Chou, Derdenger and Kumar (2019). In our application ξ_{jt} can be thought of as the quality of EVs, battery life, durability, etc. The other possible interpretation is that of advertising or the combination of quality and advertising as ξ_{jt} is the term that captures all unobserved aspects of the product. We do not have product-period advertising data in our application. It is important to highlight that the model estimation

[29]This is a slight deviation from practice as state incentives are immediately accounted for due to the state of Washington not having an income tax whereas the federal tax credit in practice is recovered by filing a federal tax return. We simplify the model and assume both credits are immediately accounted for in the price of the EV at the time of purchase.

and the counterfactual analysis do not need to assume that ξ_{jt} are independent across different products j. For example, the unobserved characteristics of the two car models from the same manufacturer (e.g. Tesla Model 3 and Model Y) could be correlated even after controlling for make fixed effect.

Following her purchase, she expects to receive the following flow utility from period $\tau > t$

$$f_{ij\tau} = \sum_{k=1}^{K} \delta_{make,k} Make_{jk} + \gamma_1' \,\mathrm{E}(X_{j\tau}^{ev} \mid X_{jt}^{ev}) + \gamma_2' \,\mathrm{E}(X_{\tau}^{cnty} \mid X_{t}^{cnty}) + \eta_i MPGe_{jt} + \xi_{jt}$$

The conditional expectations are to capture consumers' expectations about the value of future state variables, such as the expected number of Level 3 DC charging stations nearby in the next quarter of a year given the information consumers have at present. In practice, for the variables of X_{jt}^{ev} and X_t^{cnty} that evolve, we assume that consumers believe they increase at a fixed growth rate. More specifically, $\mathrm{E}(X_{j\tau,k}^{ev} \mid X_{jt}^{ev}) = q_{ev,k}^{\tau-t} X_{jt,k}^{ev}$ for $k = 1, \dots, d_{ev} = \dim(X_{jt}^{ev})$, and similarly $\mathrm{E}(X_{\tau,k}^{cnty} \mid X_{t}^{cnty}) = q_{cnty,k}^{\tau-t} X_{t,k}^{cnty}$ for $k = 1, \dots, d_{cnty} = \dim(X_t^{cnty})$. If a product characteristic or market specific variable does not vary over time, the associated growth rate is just 1.

For a consumer i of type-(g, U), we write $v_{jt}^{(g)}(U)$ to denote her expected lifetime payoffs v_{ijt} from EV j. Given the fact that consumers exit the market after the purchase of any EV, a consumer's expected lifetime payoff can be written as the sum of the current period t utility and the stream of discounted utilities in periods following purchase:

$$v_{jt}^{(g)}(U) = \frac{\sum_{k=1}^{K} \delta_j Make_{jk} + \eta_i MPGe_{jt} + \xi_{jt}}{1-\beta} + \sum_{k=1}^{d_{ev}} \frac{\gamma_{1k}}{1-\beta q_{ev,k}} X_{jt,k}^{ev} + \sum_{k=1}^{d_{cnty}} \frac{\gamma_{2k}}{1-\beta q_{cnty,k}} X_{t,k}^{cnty} - \alpha_i P_{jt} \quad (10)$$

The discount factor is $\beta \in [0, 1)$. The expected lifetime value for all variables take the form $\frac{\gamma_{1k}}{1-\beta q_{ev,k}}' X_{jt,k}^{ev}$ or $\frac{\gamma_{2k}}{1-\beta q_{cnty,k}}' X_{t,k}^{cnty}$ from the property of the infinite summation of a geometric series where the geometric component is $\beta q_{ev,k}$ or $\beta q_{cnty,k}$, which is less than 1. It is important to note that all the listed variables except for gasoline price and the number of charging stations, do not change after purchase. For the variables that do (the number of level 3 charging stations, gasoline price $\times$ EV, and gasoline price), the growth rates are greater than 1.

From above, the structural estimation relies on[30]

$$Y_{jmt} = \sum_{k=1}^{K} \frac{\delta_{make,k} - \delta_{make,1}}{1-\beta} Make_{jk} + \sum_{k=1}^{d_{ev}} (X_{jmt,k}^{ev} - X_{1mt,k}^{ev}) \frac{\gamma_{1k}}{1-\beta q_{ev,k}} - \alpha^{(1)} (P_{jmt} - P_{1mt}) + \frac{\eta^{(1)}}{1-\beta} [MPGe_{jmt} - MPGe_{1mt}] + \frac{\xi_{jmt} - \xi_{1mt}}{1-\beta}. \quad (11)$$

Note that γ_2 does not appear in and cannot be identified from the above equation. This is because the county market variables do not vary across the products in the county. From this first stage, we obtain the estimates of the make effect $(\delta_{make,2} - \delta_{make,1})/(1-\beta), \dots, (\delta_{make,K} - \delta_{make,1})/(1-\beta)$, $\gamma_{1k}/(1-\beta q_{ev,k})$ $(k = 1, \dots, d_{ev})$, $\eta^{(1)}/(1-\beta)$ and price coefficient $\alpha^{(1)}$. In this regression, car price P_{jmt} and the interaction term between the BEV dummy variable and the number of charging stations are endogenous. The instrumental variables (IV) for the price of the product j in our estimates is the cost of the battery pack multiplied by the natural log of the electric range, which can be viewed as a proxy of the battery size

[30] Appendix D collects the details of estimating CCP function with multidimensional unobserved heterogeneity.

of an EV. The average lithium battery pack price in 2019 was about half of its price in 2016 according to Bloomberg.[31] For this reason we specify battery pack price in its log form. We also include its log battery pack price without the interaction with the natural log of the electric range. Note, the IV employed accounts for the difference between product 1 and product j as seen in eq. (7)—battery pack cost in period t times the difference in electric range between products 1 and j in period t. These sets of instruments vary over time due to the declining nature of battery pack prices and cross-sectional variation from the difference between product one and the product j's electric range. This estimator is similar to the set of instruments where a time-varying instrument interacts with product dummy variables to allow for the time-varying variable to impact individual products differently (Villas-Boas, 2007). The F-statistic for the first stage excluded instruments is 11.6105 indicating non-weak instruments for price. For the number of charging stations times BEV, we follow a variant of Zhou and Li (2018) and use the log number of public facilities (such as universities, hospitals, supermarkets, etc.) which is time-varying and declining over our observed data period and its interaction with BEV. This comes from the observation that many EV charging stations are located in the parking lots of these public facilities, which were built before there was an EV market–thus uncorrelated with ξ. Charging stations are typically under the ownership of public facilities and are strategically installed to draw in customers and enhance their environmental image, among other motivations. The parking lots of these establishments present ideal locations for charging stations, offering EV drivers the potential convenience of charging their vehicles while they shop, work, or learn–hence is correlated with the number of charging stations in market m and period t. The weak instrument test of the excluded instruments for the number of stations is $6.7425e + 03$.[32]

6.3 Empirical Results

Given our methodology focuses on estimating unobserved heterogeneity, the most important parameter estimates are those linking to MPGe and price, which recover respective sensitivities ($\eta^{(1)}, \tau_{mpg}^{(g)}, \alpha^{(1)}, \tau_p^{(g)}$) for each consumer type as well as ω_{mpg} and ω_p, which again measures the variation of unobserved consumer heterogeneity within each group.

Our estimates in Table 6 indicate the price sensitivity for those whose income resides in the 90-th percentile or higher (High Income) is 2.5576, followed by group 2 (income between 75-th and 90-th percentile or Medium Income in the table) whose price sensitivity is $\alpha^{(1)} + \tau_p^{(g=2)} = 3.2127$. On average, Group 3 consumers whose income is below the 75th percentile (Low Income) are the most price-sensitive consumers with $(3.2131 = 2.5576 + 0.6555)$. Additionally, there is a statistically significant estimate of within-group unobserved heterogeneity ($\omega_p = 0.4971$) in price sensitivity. To provide economic meaning to these price estimates, we determine that a permanent 1% price increase for **all** EVs would lead to a fall in sales by 3.9284%. With respect to $MPGe_{jt}$, we find that sensitivities decrease. Consumer segment 1 whose income is above the 90th percentile is statistically not different from zero. Consumers in groups 2 and 3 negatively value MPGe compared to group 1, with sensitivity estimates of (-0.0084) and (-0.2383), respectively. There is also statistically significant unobserved heterogeneity with respect

[31] See "Battery Pack Prices Fall to an Average of $132/kWh, But Rising Commodity Prices Start to Bite," BloombergNEF, accessed May 3, 2022, `https://about.bnef.com/blog/battery-pack-prices-fall-to-an-average-of-132-kwh-but-rising-commodity-prices-start-to-bite/`

[32] Per the request of the review team, discussion of the estimation of $\delta_{make,1}$, β and $\gamma_{2k}/(1 - \beta q_{cnty,k})$ are in Appendix E.

to MPGe ($\omega_{mpg} = 0.2926$). The goodness-of-fit statistic of our first stage nonlinear least-squares model is 0.99.

We include several other observable product characteristics in our estimation, in addition to price and MPGe. We report and discuss these results in terms of their lifetime effect ($\gamma_1/(1-\beta)$). Most notably, we include make fixed effects in addition to variables for ln[Electric Range], SUV, BEV, and the interaction for BEV× SUV. The variable for ln[Electric Range] is found to be positive and significant, indicating consumers value longer-range electric vehicles. Additionally, the indicator variable for BEVs is negative and significant. This suggests that after controlling the price and electric range difference, between BEVs and PHEVs, Americans prefer PHEVs over BEVs. Furthermore, consumers have a preference for larger vehicles (SUV=1.2123) and for BEV SUVs (3.4435). Consumers of EVs also illustrate utility to horsepower as the parameter estimate is statistically significant at (2.0532) which may be driven by the out-sized sales of Telsa and its powerful engines.

In the following, we discuss the results associated with state variables whose measures change after purchase and impact future flow utilities. We report parameter estimates in terms of their lifetime effect (i.e. $\gamma_{1k}/(1 - \beta q_{ev,k})$). Our results determine that BEV demand is driven by the network of charging stations in the residing county (0.0042). This effect indicates that consumers value the size and existence of the network of electric charging stations for BEVs. Our estimates also indicate that a strong driver of the adoption of BEVs is the current price of gas, with a greater effect for BEVs (0.7102). The general term of gasoline for all electric vehicles is statistically significant with an estimate of 13.790.[33]

Finally, we discuss our discount factor estimate. The first note of interest is that we are able to precisely estimate it. Additionally, the magnitude (0.8625) is in line with the standard assumption of a monthly discount factor of $\beta = 0.965$. In order to put this estimate into perspective our estimate of $\beta = 0.8625$ equates to consumers valuing utility 13 years in the future—by the 52^{th} quarter or 13^{th} year after the purchase date, the associated flow utility is valued at near zero by the consumer in period $t = 1$ ($\beta^{4\times13} = 0.0004$). How reasonable is this estimate? It is quite sensible according to the leading industry magazine, *Consumer Reports* reports EV battery packs last 13-17 years, which is based upon a total of $200,000$ miles.[34]

6.4 Short-Term Temporary Policy Simulation

Since 2010, the federal government has incentivized the purchase of new electric vehicles with federal tax credits. To date, many BEV and PHEV buyers have benefited from this program. With consumer preference estimates in hand, we are able to compare the existing federal tax credit based on battery size to a new policy that is based on electric range. To make a fair comparison of these two policies, we simulate the impact on sales and CO2 emissions for only period 9 in our data (the middle of the data period), as the new policy must be treated as a temporary unexpected change to the existing policy as we have not estimated any state transition distributions. Moreover, we hold the cost of the new policy equal to the existing cost of the policy in period 9 across all three counties.

Before we discuss the details of our new policy and its results, it is important to understand how the

[33] We determine the growth rate for the number of charging stations is 1.0301%

[34] Ceyhan Cagatay, "How Long Should An Electric Car's Battery Last?", MyEV, accessed May 16 2022, `https://www.myev.com/research/ev-101/how-long-should-an-electric-cars-battery-last`.

Table 6: Estimation of EV Demand

Variables	Estimate	SE
Price	2.5576**	(0.7747)
Price × I[Medium Income]	0.6555**	(0.0130)
Price × I[Low Income]	0.6551**	(0.0198)
MPGe	3.0477	(2.1431)
MPGe × I[Medium Income]	−0.0084	(0.0172)
MPGe × I[Low Income]	−0.2383**	(0.0413)
ln[Electric Range] × BEV	0.7881*	(0.4616)
BEV	−2.7099**	(0.9975)
SUV	1.2123**	(0.4795)
SUV × BEV	3.4435**	(1.2727)
Horsepower	2.0532**	(0.7274)
Gas Price	13.7900**	(4.9848)
Gas Price × BEV	0.7102**	(0.3278)
EV Stations × BEV	0.0042**	(0.0012)
ω_{mpge}—MPGe	0.2926**	(0.0341)
ω_p—Price	0.4971**	(0.0510)
Discount Factor (β)	0.8625**	(0.2888)

Note: Make fixed effects are not reported to save space. Price is scaled by 1/10,000. MPGe and HP are scaled by 1/100. Estimates for observed charac. report values $\gamma/(1-\beta)$ or $\gamma/(1-\beta q)$. NLS stage has a reported $R^2 = 0.99$. ** 95 percent significance; * 90 percent significance.

tax credit works, and which cars are eligible. First, the vehicle must be new and it must be purchased rather than leased. The IRS tax credit for 2016 ranged from \$2,500 to \$7,500 per new electric vehicle purchased. The exact credit amount is based upon the *size* of the EV battery. For instance, there is a base payment of \$2,500 for any battery size. For EVs with larger batteries, the policy provides an additional credit of \$417 per kilowatt hour for batteries that are in excess of 5 kilowatt hours and is capped at \$5,000. The total federal tax credit is therefore limited to \$7,500.[35] [36]

Our new policy leverages vehicle electric range to incentivize EV adoption, rather than battery capacity.[37] For our policy simulation we propose a new federal tax incentive that is a linear function of electric range: Tax Credit (\$) = Electric Range×\$35.48. This latter measure is found by aggregating the tax credits from all EVs sold during Q1 of 2018 in the 3 counties of Washington state.

The new proposed policy rewards not only cars with larger batteries but also more efficient cars compared with the existing policy. To put this into perspective, a Tesla Model S in 2018 had an electric range of 250 miles and a 95kWh battery and received a federal tax credit of \$7,500. Under our proposed policy that tax credit would increase to \$8,870. Contrast that with a BMW X5 PHEV with an electric range of 14 miles and a 9.1 kWh battery in 2018. Under the current policy, the consumer received

[35] https://www.efile.com/electric-vehicle-car-tax-credits/

[36] There are other details of the policy that are omitted as they are not relevant to the current counterfactual simulation (e.g. phasing out of the tax credit over time after a manufacturer reaches a total of 200,000 BEV or PHEV vehicles sold nationally since 2010).

[37] Electric range captures the efficacy of the car and size of the electric battery.

a $\$4,668$ tax credit whereas under our policy a consumer would only receive \$496.72—a roughly 89% reduction in the federal tax credit.

In order to simulate the results of our new temporary policy, we assume the change in the tax rebate was unexpected to the consumers and is believed to be temporary (one period). We leverage the structural parameter estimates along with the series multinomial logit estimates of the CCP (ρ_{jmt1}, ρ_{jmt2}, and ρ_{jmt3}) to simulate new choice probabilities under the new proposed tax policy. We show below that only ρ_{jmt1} and ρ_{jmt2} change from their initial estimates. Note, that policies that are not temporary cannot leverage the series multinomial logit CCP parameters to simulate new CCPs as such policies would require a change in the beliefs and/or evolution of the state space.

Proposition 3 (Transformation of the series multinomial logit parameters for a temporary and unexpected policy change)**.** *Define $\check{\rho}_{jmt1}$, $\check{\rho}_{jmt2}$ and $\check{\rho}_{jmt3}$ as CCP series multinomial logit parameters under the new policy. We have*

$$\check{\rho}_{jmt1} = \rho_{jmt1} - \alpha^{(1)}\Delta P_{jmt}, \qquad \check{\rho}_{jmt2} = \begin{pmatrix} \check{\rho}_{jmt2,1} \\ \check{\rho}_{jmt2,2} \end{pmatrix} = \begin{pmatrix} \rho_{jmt2,1} \\ \rho_{jmt2,2} - \Delta P_{jmt}\omega_2 \end{pmatrix}, \qquad \check{\rho}_{jmt3} = \rho_{jmt3},$$

where ΔP_{jmt} is the (after-tax-credit) price difference between the new and old policy for product j in period t and county m.

The results of our proposed policy are reported in Table 7, and highlight a sizable reduction in total CO2 emissions (11.78%) with a (-0.61%) reduction in the total number of EVs sold (10 car). This indicates under the new policy consumers substituted away from PHEV vehicles for more efficient BEV vehicles (209 more BEVs were sold) but at the cost of 10 gasoline-powered cars remaining on the road. In Q1 of 2018, the original policy sold 1,619 units whereas the new policy sold 10 fewer cars at 1,609 cars. Given our policy emphasizes electric range (a combination of battery efficiency and size rather than simply capacity), we analyze the policy's effect on CO2 emissions. Leveraging data from the US Department of Energy we can determine the change in CO2 due to the change in tax policy. It is important to note that the results that pertain to CO2 emissions *do* include emissions from the generation of electricity to charge the EVs. To account for the 10 fewer cars sold in our policy experiment, we add the differential in total carbon emissions between PHEV and gasoline automobiles times the number of fewer EVs sold to the total CO2 policy calculation. Using the same data from the Department of Energy we determine in the state of Washington that gasoline-powered cars emit 2.3 times more total CO2 than PHEVs and equates to 55.92 tons of CO2 in our simulation. Even with 10 fewer EVs sold, our proposed policy reduces total CO2 emissions by 11.78% or 354 tons of CO2 from $3,006$ tons to $2,652$ tons. In summary, we show a cost-neutral temporary policy that emphasizes electric range over battery storage can reduce CO2 emissions by incentivizing consumers to switch from PHEVs to BEVs, and in particular lighter more efficient BEVs.

6.5 Long-Term Policy Simulations

We present two additional permanent long-term policies. The first of these two is identical to the one above but such a policy is now permanent and runs over the entire data period. We again determine the credit per mile in electric range that equates to the total cost between the existing storage-focused policy

Table 7: Policy Simulation Results: Temporary Electric Range Tax Credit

	Existing Policy (kWh)	Proposed Policy (Elec. Range)	% change
Sales (units)	1,619	1,609	-0.61%
BEV Sales (units)	1,172	1,381	17.83%
PHEV Sales (units)	447	228	-49.01%
Total CO2 (tons)	3,006	2,652	-11.78%
Total Cost ($)	10,920,283.60	10,920,314.98	-2.8734e-04%

and our new efficiency policy for the observed data period. The second policy simulation addresses the trade-off between incentivizing EV purchases through tax credits versus an investment in infrastructure in order to provide a deeper understanding of how to seed a market with indirect network effects.

It is important to note that to determine the impact of each policy we must specify the state space and the assumption on how consumers form expectations about the evolution of those state variables. This is required as our methodology for estimation is agnostic to such. Moreover, we determine that to identify the time, model, and market-specific structural errors (ξ_{jtm}) we also must make such an assumption. In practice, many papers have used the inclusive value sufficiency (IVS) assumption of Gowrisankaran and Rysman (2012) with aggregate data to estimate dynamic demand models. Yet, Derdenger and Kumar (2019) have highlighted the IVS estimation procedure is a biased and inconsistent estimator unless beliefs follow such an assumption. At this given time there is not a method to determine whether the IVS assumption is a valid assumption for a given data set. That said, consumer beliefs are required to proceed with our counterfactual simulations given our estimator is belief agnostic. Any specified consumer beliefs (including IVS) may well be wrong. Note again, the benefit of our estimator is that (potentially) incorrect beliefs are not assumed until after estimation and are only used in counterfactual simulation rather than for estimation AND counterfactual simulations such as other estimators.

In practice, we select the IVS assumption for its computational ease. Since all utility preferences are estimated and known at this stage, including parameters that capture observed and unobserved heterogeneity in price and MPGe, we leverage the IVS assumption to identify the model $\xi's$. This involves matching the simulated market shares given the estimated utility parameters and the assumption of IVS to the observed data. The IVS assumption reduces the state space to one dimensional inclusive value. For consumer i of type (g, U), the inclusive value ι_{it} is defined as follows,

$$\iota_{it} = \mathrm{E}_{\varepsilon}\left[\max_{j=1,\ldots,J} v_{ijt} + \varepsilon_{ijt}\right] = \ln\left(\sum_{j=1,\ldots,J} \exp\left(v_{jt}^{(g)}(U)\right)\right)$$

where $v_{jt}^{(g)}(U)$ is from eq. (10).

The Bellman equation can consequently be expressed in terms of the inclusive values of the car models on the market,

$$\bar{V}_{it}\left(\iota_{it}\right) = \ln\left(\underbrace{\exp\left(\iota_{it}\right)}_{\text{Purchase}} + \underbrace{\exp\left(\beta\,\mathrm{E}\left[\bar{V}_{it}(\iota_{i,t+1}) \;\middle|\; \iota_{it}\right]\right)}_{\text{No Purchase}}\right)$$

Table 8: Long-Term Policy Simulation Results: Electric Range Tax Credit

	Existing Policy (kWh)	Proposed Policy (Electric Range)	% change
Sales (units)	30,638	30,999	1.18%
BEV Sales (units)	25,287	27,399	8.35%
PHEV Sales (units)	5,352	3,600	-32.73%
Total CO2 (tons)	51,872	46,150	-11.03%
Total Cost ($)	180,813,613	180,813,613	2.29e-04%

Note: Policy Rebate=30.54 per mile of electric range

with the evolution of the inclusive value being specified as evolving according to an AR(1) process with Gaussian random shocks.[38] After the model $\xi's$ have been identified, we proceed with counterfactual simulations where we adjust the relevant model characteristics, as well as assume that the AR(1) process of the inclusive value adjusts given that consumers form new beliefs about the new inclusive value term.

In Table 8 we present the result of a permanent long-term policy focused on battery efficiency. Like the first policy simulation above, we determine the cost-neutral price (between the old and new policy) is $30.54 per mile of electric range. The policy increases the total number of BEVs sold by $2,112$ but the number of PHEVs sold decreases by $1,752$ units for a total increase of 1.18%. Yet, CO2 emissions decrease by 6.78% or $3,519$. This number under-reports the reduction in CO2 emissions as it does not adjust for the 361 additional EV vehicles that were sold (hence 361 fewer internal combustion engine vehicles on the road). Consequently, we reduce the total calculated CO2 emissions by $2,209.32 = 361 \times 6.12$ to provide a total CO2 emission of $46,150$ tons. The policy reduces CO2 emissions by 11.03%. With that, it should be highlighted that the total C02 emissions under the new policy are lower than the observed data $(51,872)$ amount even without the adjustment $(48,353)$.

To glean more insight into the result, we present in Table 9 the analysis of the impact by income group. What is evident is that under the new policy, many of the more price-sensitive low-income consumers (those who fall under the $90th$ percentile) exit the overall market (BEV plus PHEV) by not purchasing any EVs. Whereas those who are relatively less price sensitive substitute PHEVs for more efficient BEVs as a whole and within the class of BEVs. Thus, the impact of such a policy has a disparate impact on differing income groups as BEVs are on average $6k more expensive than PHEVs.

It is important to note that our counterfactual results do not account for how EV characteristics may change because of the new policy. That said, we believe this new policy based on the electric range of the EV could incentivize EV manufacturers to change and design more tax-efficient vehicles. The implication is that under such a setting, the tax rate per electric mile would decrease due to the increase in sales caused by the positive preference for a larger range for BEVs holding the policy cost neutral to the existing one. The net effect on price may be minimal as long as the price per mile is reduced by the same percentage as the increase in the electric range. If this is the case, the impact on CO2 emissions is a lower bound.

Our last long-term permanent policy simulation looks to address how best to seed a market with indirect network effects. The question of whether individual tax credits versus investments in infrastructure

[38] We refer the reader to Gowrisankaran and Rysman (2012) for details.

Table 9: Long-Term Policy Simulation Results: Electric Range Tax Credit

	Income ≤ 75th Percentile	Income between 90th and 75th Percentile	Income ≥ 90th Percentile
BEV Sales (units)	4,274	5,372	15,640
BEV Sales CF (units)	4,630	5,675	17,093
PHEV Sales (units)	1,493	1,292	2,566
PHEV Sales CF (units)	835	951	1,812

Note: Policy Rebate=30.54 per mile of electric range. "CF" stands for counterfactual policy.

through the subsidizing of a larger charging network is an important question to address. To implement this policy, we eliminate tax credits to consumers and replace them with investment in level 3 charging stations in the three markets. The US Department of Energy reports the max cost of purchasing and installing a level 3 DC station is $\$90,000$ in November of 2015.[39] The DOE also reports that the state of Washington offered a \$15k tax incentive for the purchase and installation of each level 3 charger leading to a price of $\$75,000$ per level 3 charger. The policy we run looks to determine the increase in the number of new charging stations in each period that would make such a policy cost-neutral to the existing tax credit policy based on battery size over the observed data period.[40] For each of the three counties, we determine that an increase in the mean growth rate from 6.89% to 12.34% makes this policy cost-neutral. Note that any new station built is present in each of the subsequent periods.

Table 10 presents the results of the policy. Specifically, we see a 25.75% increase in BEVs sold ($6,511$) and a reduction of 32.13% in PHEV sales ($2,288$). Thus, a policy that emphasizes infrastructure and incentivizes consumer adoption through a larger and more dense network of level 3 charging stations provides a greater effect on BEV adoption than consumer tax credits. To address the impact on CO2, as we do in the above policies, we must assume the type of EVs sold for the extensive margin. Here we assume that the additional $4,792$ EVs are BEVs and replace gasoline-powered automobiles, leading to a reduction in carbon emissions from gasoline-powered cars of 7.65 tons per total CO2 to 1.53 tons. Consequently, we reduce the total calculated CO2 emissions by $29,327 = 4,792 \times 6.12$ to provide a total CO2 emission of $25,245$ tons (without this adjustment the total CO2 is 54,572).

We acknowledge that there is the potential for changes to product characteristics due to policy changes (e.g. PHEVs with larger batteries). Moreover, we assume that the price before any tax incentives remains constant with what is observed in the data. As is always the case with demand-side counterfactuals, when the modeling of other supply-side decisions does not occur, such counterfactual results will only present a partial equilibrium viewpoint. This is a limitation of our counterfactual analysis—potentially price before tax incentives may increase due to the incentive to increase battery size and its subsequent cost.

[39] https://afdc.energy.gov/files/u/publication/evse˙cost˙report˙2015.pdf

[40] We accomplish this by estimating an AR1 process for the existing number of charging stations and adjust the AR1 coefficient to find the cost neutral allocation. We then determine the growth rate for the existing and new simulated set of charging stations to adjust for the new coefficient $\gamma/(1-\beta q')$

Table 10: Long-Term Policy Simulation: Level 3 Charging Infrastructure Network

	Existing Policy	New Infrastructure Policy	% change
Sales (units)	30,638	35,430	15.64%
BEV Sales (units)	25,287	31,798	25.75%
PHEV Sales (units)	5,352	3,064	-32.13%
Total CO2 (tons)	51,872	25,245	-51.33%
Total Cost ($)	180,813,613	180,827,340	0.01%

Note: All prior federal tax credits are eliminated

7 Conclusion

We study the effects of a few competing government incentives on EV adoption by building a structural dynamic discrete demand model. We develop a new approach to estimating the model with aggregate consumer *group* sales data. We find that building an EV charging network is better than awarding a tax rebate for purchasing an EV in terms of nurturing the EV market. Even for the design of a tax rebate program, we find the rebate based on electric range is more effective than the current rebate program that is based on battery size for bringing more EVs on the road and reducing CO2 emission.

The estimation method we develop in this paper has a broader application than the study of EV adoption. In estimating dynamic discrete choice demand models for durable goods, it is essential to account for unobserved consumer heterogeneity and product characteristics to obtain unbiased estimates of important parameters like price elasticities. However, in the implementation of such models, it is inevitable to address the curse of dimensionality caused by the large number of products on the market and the high dimension of product characteristics. Our new approach using group market share data includes continuous unobserved consumer heterogeneity and unobserved product characteristics but avoids the curse of dimensionality. As a result, our methodology can be used in the markets with many differentiated products.

The implementation of our method is simple and requires only NLS and 2SLS. It allows researchers to consider various model specifications at little computational and programming cost. In addition to these practical benefits, our method also has a few theoretically appealing properties. We find that the identification of the dynamic demand model requires the same conditions as the identification of the static demand model, which requires instrumental variables to address the endogeneity of variables like price.

Our data requirement is group market shares (or sales). This is weaker than requiring consumer-level panel data, from which we can construct the group market shares. To collect customer-level panel data, companies need to track customers over time. For durable goods, this could be very costly and impractical due to their longer life span (last at least 3 years by the definition of the US Census Bureau[41]). Additionally, in the digital age where privacy is a concern, customers are becoming more unwilling to be tracked over time, and companies are unwilling to share their customer-level data with researchers.[42] As

[41]See for example: "Manufacturers Place Orders When Economy Improves," United States Census Bureau, accessed December 24, 2020,https://www.census.gov/library/stories/2018/07/manufacturers-durable-goods.html

[42]Dropbox once shared its customer data with researchers and incurred ethical concerns. See the report by

a result, we believe the need for such a method will grow. For instance, Google recently announced that its Chrome internet browser will stop supporting third-party cookies (a user-tracking technology) by late 2023 making it very difficult for digital advertising companies to individually target consumers.[43] As an alternative, Google has been testing a new tool called Floc which allows advertisers to follow *cohorts of users* rather than individuals. We therefore expect our method to be applicable to the data generated by Floc or other similar privacy amicable technologies.

8 Funding and Competing Interests

None

References

Abbring, Jaap H., and Øystein Daljord. 2020. "Identifying the discount factor in dynamic discrete choice models." *Quantitative Economics*, 11(2): 471–501.

An, Yonghong, Yingyao Hu, and Ruli Xiao. 2020. "Dynamic Decisions under Subjective Expectations: A Structural Analysis." *Journal of Econometrics*, Forthcoming.

Arcidiacono, P., and R.A. Miller. 2011. "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity." *Econometrica*, 79(6): 1823–1867.

Arcidiacono, Peter, and Robert Miller. 2018. "Identifying Dynamic Discrete Choice Models off Short Panels." Duke University Memo.

Bajari, Patrick, Chenghuan Sean Chu, Denis Nekipelov, and Minjung Park. 2016. "Identification and Semiparametric Estimation of a Finite Horizon Dynamic Discrete Choice Model with a Terminating Action." *Quantitative Marketing and Economics*, 14(4): 271–323.

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

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

Berry, Steven T, and Philip A Haile. 2014. "Identification in Differentiated Products Markets Using Market Level Data." *Econometrica*, 82(5): 1749–1797.

Berry, Steven T, and Philip A Haile. 2020. "Nonparametric Identification of Differentiated Products Demand Using Micro Data." Yale University Memo.

Chou, Cheng, Tim Derdenger, and Vineet Kumar. 2019. "Linear Estimation of Aggregate Dynamic Discrete Demand for Durable Goods: Overcoming the Curse of Dimensionality." *Marketing Science*, 38(5): 888–909.

Emily Dreyfuss, "Was It Ethical for Dropbox to Share Customer Data with Scientists?", Wired, July 24, 2018, https://www.wired.com/story/dropbox-sharing-data-study-ethics/.

[43] Patience Haggin, Sam Schechner and Suzanne Vranica, "Google Delays Cookie Removal to Late 2023", The Wall Street Journal, June 24, 2021, https://www.wsj.com/articles/google-delays-cookie-removal-to-late-2023-11624542064?st=sund6fzua5hdf76&reflink=desktopwebshare_permalink.

Daljord, Øystein, Denis Nekipelov, and Minjung Park. 2018. "A Simple and Robust Estimator for Discount Factors in Optimal Stopping Dynamic Discrete Choice Models." University of Chicago Memo.

Derdenger, Timothy, and Vineet Kumar. 2019. "Estimating Dynamic Discrete Choice Models with Aggregate Data: Small Sample Properties of the Inclusive Value Approximation." *Quantitative Marking and Economics*, 17: 359–384.

Dubé, Jean-Pierre, Jeremy T Fox, and Che-Lin Su. 2012. "Improving the Numerical Performance of Static and Dynamic Aggregate Discrete Choice Random Coefficients Demand Estimation." *Econometrica*, 80(5): 2231–2267.

Gillingham, Kenneth T., Arthur A. van Benthem, Stephanie Weber, Mohamed Ali Saafi, and Xin He. 2023. "Has Consumer Acceptance of Electric Vehicles Been Increasing? Evidence from Microdata on Every New Vehicle Sale in the United States." *AEA Papers and Proceedings*, 113: 329–35.

Goettler, Ronald L, and Brett R Gordon. 2011. "Does AMD Spur Intel to Innovate More?" *Journal of Political Economy*, 119(6): 1141–1200.

Gowrisankaran, Gautam, and Marc Rysman. 2012. "Dynamics of Consumer Demand for New Durable Goods." *Journal of Political Economy*, 120(6): 1173–1219.

Hirano, Keisuke, Guido W. Imbens, and Geert Ridder. 2003. "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score." *Econometrica*, 71(4): 1161–1189.

Hotz, V. Joseph, and Robert A. Miller. 1993. "Conditional Choice Probabilities and the Estimation of Dynamic Models." *The Review of Economic Studies*, 60(3): 497–529.

Hu, Yingyao, and Matthew Shum. 2012. "Nonparametric Identification of Dynamic Models with Unobserved State Variables." *Journal of Econometrics*, 171(1): 32–44.

Hu, Yingyao, Matthew Shum, Wei Tan, and Ruli Xiao. 2017. "A Simple Estimator for Dynamic Models with Serially Correlated Unobservables." *Journal of Econometric Methods*, 6(1).

Kasahara, Hiroyuki, and Katsumi Shimotsu. 2009. "Nonparametric Identification of Finite Mixture Models of Dynamic Discrete Choices." *Econometrica*, 77(1): 135–175.

Magnac, Thierry, and David Thesmar. 2002. "Identifying Dynamic Discrete Decision Processes." *Econometrica*, 70(2): 801–816.

Melnikov, Oleg. 2013. "Demand for Differentiated Durable Products: The Case of the US Computer Printer Market." *Economic Inquiry*, 51(2): 1277–1298.

Norets, Andriy. 2009. "Inference in Dynamic Discrete Choice Models with Serially Correlated Unobserved State Variables." *Econometrica*, 77(5): 1665–1682.

Palm, A. 2020. "Early adopters and their motives: Differences between earlier and later adopters of residential solar photovoltaics." *Renewable and Sustainable Energy Reviews*, 133: 110142.

Plötz, Patrick, Cornelius Moll, Georg Bieker, Peter Mock, and Yaoming Li. 2020. "Real-World Usage of Plug-In Hybrid Electric Vehilces Fuel Consumption, Electric Drivign, and CO2 Emissions." The International Council of Clean Transportation White Paper.

Song, Inseong, and Pradeep K Chintagunta. 2003. "A Micromodel of New Product Adoption with Heterogeneous and Forward-Looking Consumers: Application to the Digital Camera Category." *Quantitative Marketing and Economics*, 1(4): 371–407.

Srisuma, Sorawoot, and Oliver Linton. 2012. "Semiparametric estimation of Markov decision processes with continuous state space." *Journal of Econometrics*, 166(2): 320–341.

Su, Che-Lin, and Kenneth L Judd. 2012. "Constrained Optimization Approaches to Estimation of Structural Models." *Econometrica*, 80(5): 2213–2230.

Sun, Yutec, and Masakazu Ishihara. 2019. "A computationally efficient fixed point approach to dynamic structural demand estimation." *Journal of Econometrics*, 208(2): 563–584.

Villas-Boas, Sofia Berto. 2007. "Vertical Relationships between Manufacturers and Retailers: Inference with Limited Data." *The Review of Economic Studies*, 74(2): 625–652.

Zhou, Yiyi, and Shanjun Li. 2018. "Technology Adoption and Critical Mass: The Case of the US Electric Vehicle Market." *The Journal of Industrial Economics*, 66(2): 423–480.

(For Online Publication)

Online Appendix

A Asymptotics

The asymptotics is built on $T \to \infty$ with the number of products J and the number of groups G being fixed. Let $\theta \equiv (\theta_1, \theta_2)$, where $\theta_1 \equiv (\tau, \omega, \rho, (\delta_2 - \delta_1)/(1-\beta), \ldots, (\delta_J - \delta_1)/(1-\beta), \gamma/(1-\beta), \alpha^{(1)})$, and $\theta_2 \equiv (\beta, \delta_1)$. We decompose θ into these two parts, because the estimation of θ_2 relies on the estimation of θ_1. Let $\hat{\theta}_1$ and $\hat{\theta}_2$ denote the estimator of θ_1 and θ_2, respectively.

It is easier to derive the asymptotic distribution backwardly from $\theta_2 \equiv (\beta, \delta_1)$. We use eq. (Linear-Reg-2) to estimate θ_2. In the estimation of θ_2, we fix U at certain number. In particular, we let $U = 0$ here to simplify the discussion. For exposition simplicity, we ignore the dependence on U below. Let

$$\hat{W}^{(g)}(\hat{\theta}_1) \equiv T^{-1} \sum_{t=1}^{T} W_t^{(g)}(U=0, \hat{\theta}_1), \qquad \text{and} \qquad \hat{\ell}^{(g)}(\hat{\theta}_1) \equiv T^{-1} \sum_{t=1}^{T} \ln \sigma_{0t}^{(g)}(U=0; \hat{\theta}_1).$$

Then

$$\hat{\theta}_2 = [A(\hat{\theta}_1)' A(\hat{\theta}_1)]^{-1} A(\hat{\theta}_1)' Y_2(\hat{\theta}_1).$$

where $A(\hat{\theta}_1)$ is $G \times 2$ matrix, and $Y_2(\hat{\theta}_1)$ is $G \times 1$ vector defined below:

$$A(\hat{\theta}_1) \equiv \begin{bmatrix} 1 & \hat{W}^{(1)}(\hat{\theta}_1) + \hat{\ell}^{(1)}(\hat{\theta}_1) \\ \vdots & \vdots \\ 1 & \hat{W}^{(G)}(\hat{\theta}_1) + \hat{\ell}^{(G)}(\hat{\theta}_1) \end{bmatrix} \qquad \text{and} \qquad Y_2(\hat{\theta}_1) \equiv \begin{bmatrix} \hat{W}^{(1)}(\hat{\theta}_1) \\ \vdots \\ \hat{W}^{(G)}(\hat{\theta}_1) \end{bmatrix}.$$

They are defined in this way so that Y_2 is the "dependent variable" and A is the "design matrix" of eq. (Linear-Reg-2). Because $\hat{\theta}_2$ is analytical function of random variables $\hat{W}^{(g)}(\hat{\theta}_1)$ and $\hat{\ell}^{(g)}(\hat{\theta}_1)$, its variance can be easily obtained by simulation provided that we know the asymptotic distribution of $\hat{W}^{(g)}(\hat{\theta}_1)$ and $\hat{\ell}^{(g)}(\hat{\theta}_1)$.

Now, we derive the distribution of $\hat{W}^{(g)}(\hat{\theta}_1)$ and $\hat{\ell}^{(g)}(\hat{\theta}_1)$, whose definition requires $W_t^{(g)}(0; \hat{\theta}_1)$ and $\sigma_{0t}^{(g)}(0; \hat{\theta}_1)$. We have

$$\begin{aligned} W_t^{(g)}(0; \hat{\theta}_1) &= \ln \left[\frac{\hat{\sigma}_{1t}^{(g)}(0)}{\hat{\sigma}_{0t}^{(g)}(0)} \right] - X_{1t}' \widehat{\left(\frac{\gamma}{1-\beta} \right)} + (\hat{\alpha}^{(1)} + \hat{\tau}^{(g)}) P_{1t} \\ &= \ln \left[\frac{\hat{\sigma}_{1t}^{(1)}(\hat{\tau}^{(g)}/\hat{\omega})}{\hat{\sigma}_{0t}^{(1)}(\hat{\tau}^{(g)}/\hat{\omega})} \right] - X_{1t}' \widehat{\left(\frac{\gamma}{1-\beta} \right)} + (\hat{\alpha}^{(1)} + \hat{\tau}^{(g)}) P_{1t} \\ &= \left[\hat{\rho}_{jt1} + \hat{\rho}_{jt2} \frac{\hat{\tau}^{(g)}}{\hat{\omega}} + \hat{\rho}_{jt3} \left(\frac{\hat{\tau}^{(g)}}{\hat{\omega}} \right)^2 \right] - X_{1t}' \widehat{\left(\frac{\gamma}{1-\beta} \right)} + (\hat{\alpha}^{(1)} + \hat{\tau}^{(g)}) P_{1t}, \end{aligned}$$

and $\sigma_{0t}^{(g)}(0; \hat{\theta}_1) = \sigma_{0t}^{(1)}(\hat{\tau}^{(g)}/\hat{\omega}; \hat{\theta}_1)$ has the series logit expression. Both $W_t^{(g)}(0; \hat{\theta}_1)$ and $\sigma_{0t}^{(g)}(0; \hat{\theta}_1)$ are analytical functions of $\hat{\theta}_1$. We then can determine the variance of $\hat{W}^{(g)}(\hat{\theta}_1)$ and $\hat{\ell}^{(g)}(\hat{\theta}_1)$ by randomly drawing samples from the asymptotic distribution of $\hat{\theta}_1$.

Lastly, we derive the distribution of $\hat{\theta}_1$. We estimate θ_1 by

$$\hat{\theta}_1 \equiv \arg\min_{\theta_1 \in \Theta_1} T^{-1} \sum_{t=1}^{T} h_t(\theta)' h_t(\theta)$$

subject to constraints eq. (B.4) below,

where

$$h_t(\theta_1)' \equiv (h_{1t}(\theta_1)', h_{2t}(\theta_1)', \ldots, h_{Jt}(\theta_1)'),$$

and

$$h_{jt}(\theta_1) \equiv \begin{bmatrix} S_{jt}^{(1)} - \mathrm{GH}_{jt}^{(G)}(\tau, \omega, \rho) \\ \vdots \\ S_{jt}^{(G)} - \mathrm{GH}_{jt}^{(G)}(\tau, \omega, \rho) \\ X_{jt}^{IV}\left[Y_{jt} - \frac{\delta_j - \delta_1}{1-\beta} - (X_{jt} - X_{1t})' \frac{\gamma}{1-\beta} + \alpha^{(1)}(P_{jt} - P_{1t})\right] \end{bmatrix}.$$

Here X_{jt}^{IV} is a vector of IV in eq. (Linear-Reg-1) satisfying $\mathrm{E}[X_{jt}^{IV}(\xi_{jt} - \xi_{1t})] = 0$. This is a standard M-estimation problem, so under the regularity conditions, $\sqrt{T}(\hat{\theta}_1 - \theta_1) \to_d \mathcal{N}(0, \Sigma_1)$. The asymptotic variance Σ_1 is readily reported by most statistical softwares, or computed using numerical score and Hessian matrices.

B Proofs

Proof of Proposition 1. First, without attrition, the pool of consumers does not change with time. So that $f_t^{(g)}(u) = f_1^{(g)}(u) = \phi(u)$ for any period t. In the rest, we focus on the case of attrition.

When $t = 1$, $f_1^{(g)}(u) = \phi(u)$ by Assumption 5. We will prove

$$f_t^{(g)}(u) = \phi(u) \times \prod_{s=1}^{t-1} \frac{\sigma_{0t}^{(g)}(u)}{S_{0t}^{(g)}}, \qquad t \geq 2. \tag{B.1}$$

by the induction. Define a few notations for exposition. Let A_{it} denote the discrete purchasing choice made by consumer i in period t. Particularly, $A_{it} = 0$ means not purchase in period t. Let $Z_t \equiv (X_t', P_t', \xi_t')'$ denote the vector of product characteristics in period t. Also recall that $D_i^{(g)} = 1$ denotes that i is from group g.

When we randomly draw a consumer i with unobserved price sensitivity U_i from group g, $f_t^{(g)}(u)$ is the probability that $U_i = u$ provided that consumer i still exists in period t. Because consumers leave the market after purchasing, a consumer would exist in period t if and only if she had chosen not to purchase in all the previous periods given the past product characteristics. In other words, $f_t^{(g)}(u)$ is the probability that $U_i = u$ conditional on that $D_i^{(g)} = 1$ (so i is from group g) and consumer i did not purchase from period 1 to $t - 1$ with the past product characteristics $Z_1, \ldots, Z_{t-1}$. That is

$$f_t^{(g)}(u) = \Pr(U_i = u \mid D_i^{(g)} = 1, A_{i1} = 0, \ldots, A_{i,t-1} = 0, Z_1, \ldots, Z_{t-1}).$$

The above conditioning variables just restrict the population to be the remaining consumers after $t - 1$ periods. Because all the remaining consumers in period t face the same product state variables Z_t, we also have the following conditional independence,

$$
\begin{aligned}
f_t^{(g)}(u) &= \Pr(U_i = u \mid D_i^{(g)} = 1, A_{i1} = 0, \dots, A_{i,t-1} = 0, Z_1, \dots, Z_{t-1}) \\
&= \Pr(U_i = u \mid D_i^{(g)} = 1, A_{i1} = 0, \dots, A_{i,t-1} = 0, Z_1, \dots, Z_{t-1}, Z_t).
\end{aligned} \tag{B.2}
$$

We now prove eq. (B.1) by the induction. Starting with period 2, we have

$$
\begin{aligned}
f_2^{(g)}(u) &= \Pr(U_i = u \mid D_i^{(g)} = 1, A_{i1} = 0, Z_1) \\
&= \frac{f(U_i = u \mid D_i^{(g)} = 1, Z_1) f(A_{i1} = 0 \mid U_i = u, D_i^{(g)} = 1, Z_1)}{f(A_{i1} = 0 \mid D_i^{(g)} = 1, Z_1)} \\
&= \phi(u) \times \frac{\sigma_{01}^{(g)}(u)}{S_{01}^{(g)}}.
\end{aligned}
$$

The third line used $f(U_i = u \mid D_i^{(g)} = 1, Z_1) = f(U_i = u \mid D_i^{(g)} = 1)$ because all consumers in group g face the same product characteristics Z_1 in the first period.

Suppose eq. (B.1) holds for period t. We will prove that this equation also holds for period $t+1$. We have

$$
\begin{aligned}
f_{t+1}^{(g)}(u) &= \Pr(U_i = u \mid D_i^{(g)} = 1, A_{i1} = 0, \dots, A_{i,t-1} = 0, A_{it} = 0, Z_1, \dots, Z_t) \\
&= \Pr(U_i = u \mid D_i^{(g)} = 1, A_{i1} = 0, \dots, A_{it-1} = 0, Z_1, \dots, Z_t) \\
&\quad \times \frac{f(A_{it} = 0 \mid U_i = u, D_i^{(g)} = 1, A_{i1} = 0, \dots, A_{it-1} = 0, Z_1, \dots, Z_t)}{f(A_{it} = 0 \mid D_i^{(g)} = 1, A_{i1} = 0, \dots, A_{it-1} = 0, Z_1, \dots, Z_t)} \\
&= f_t^{(g)}(u) \times \frac{f(A_{it} = 0 \mid U_i = u, D_i^{(g)} = 1, A_{i1} = 0, \dots, A_{it-1} = 0, Z_1, \dots, Z_t)}{f(A_{it} = 0 \mid D_i^{(g)} = 1, A_{i1} = 0, \dots, A_{it-1} = 0, Z_1, \dots, Z_t)} \qquad \text{by eq. (B.2)} \\
&= f_t^{(g)}(u) \frac{\sigma_{0t}^{(g)}(u)}{S_{0t}^{(g)}} = \phi(u) \times \prod_{s=1}^{t} \frac{\sigma_{0s}^{(g)}(u)}{S_{0s}^{(g)}}.
\end{aligned}
$$

Note that the purchase choice A_{it} in period t depends only on the payoffs v_{ijt}, which are functions of $(U_i, D_i^{(g)}, Z_t)$ only. So that $A_{it} \perp\!\!\!\perp (A_{i1}, \dots, A_{it-1}, Z_1, \dots, Z_{t-1}) \mid (U_i, D_i^{(g)}, Z_t)$, and the last line follows. This completes the proof. ■

Proposition B.1 (Group composition due to attrition)**.** *Suppose consumers leave the market after purchasing. Let $\pi_t^{(g)}$ denote the proportion of group g consumers in period t, and let S_{0t} denote the share of consumers who choose the outside option (not purchase) in period t. We have*

$$
\pi_t^{(g)} = \pi_1^{(g)} \times \left(\prod_{s=1}^{t-1} \frac{S_{0s}^{(g)}}{S_{0s}} \right). \tag{B.3}
$$

Proof. The proof is similar to the proof of Proposition 1 and we keep using the notation defined in that proof. We prove by induction. Starting from period 2, we have the following by definition,

$$
\begin{aligned}
\pi_2^{(g)} &= \Pr(D_i^{(g)} = 1 \mid A_{i1} = 0, Z_1) \\
&= \frac{\Pr(D_i^{(g)} = 1 \mid Z_1) \Pr(A_{i1} = 0 \mid D_i^{(g)} = 1, Z_1)}{\Pr(A_{i1} = 0, Z_1)} \\
&= \pi_1^{(g)} \times \frac{S_{01}^{(g)}}{S_{01}}.
\end{aligned}
$$

Suppose eq. (B.3) holds for period t. We want to show the statement holds for period $t+1$, we have

$$\begin{aligned}\pi_{t+1}^{(g)} &= \Pr(D_i^{(g)} = 1 \mid A_{i1} = 0, \ldots, A_{i,t-1} = 0, A_{it} = 0, Z_1, \ldots, Z_t) \\ &= \Pr(D_i^{(g)} = 1 \mid A_{i1} = 0, \ldots, A_{i,t-1} = 0, Z_1, \ldots, Z_{t-1}, Z_t) \\ &\quad \times \frac{\Pr(A_{it} = 0 \mid D_i^{(g)} = 1, A_{i1} = 0, \ldots, A_{i,t-1} = 0, Z_1, \ldots, Z_t)}{\Pr(A_{it} = 0 \mid A_{i1} = 0, \ldots, A_{i,t-1} = 0, Z_1, \ldots, Z_t)} \\ &= \pi_t^{(g)} \frac{S_{0t}^{(g)}}{S_{0t}}.\end{aligned}$$

The last line follows because the vector product characteristics Z_t is the same for different groups of remaining consumers after $t-1$ periods, so that $\Pr(D_i^{(g)} = 1 \mid A_{i1} = 0, \ldots, A_{i,t-1} = 0, Z_1, \ldots, Z_{t-1}, Z_t) = \Pr(D_i^{(g)} = 1 \mid A_{i1} = 0, \ldots, A_{i,t-1} = 0, Z_1, \ldots, Z_{t-1}) = \pi_t^{(g)}$. This completes the proof. ■

Proof of Proposition 2. We claim that the dynamic model implies the following constraints on the parameters ρ_t in the CCP function:

$$\rho_{jt2} - \rho_{1t2} = -\omega(P_{jt} - P_{1t}) \qquad \text{and} \qquad \rho_{jt3} - \rho_{1t3} = 0, \qquad j = 2, \ldots, J. \tag{B.4}$$

The above constraints eliminate many parameters, leaving the following to estimate in the NLS problem, eq. (5):

$$(\rho_{1t1}, \ldots, \rho_{Jt1})', \quad \rho_{1t2}, \quad \rho_{1t3}, \quad \omega, \quad \tau, \qquad \text{for } t = 1, \ldots, T.$$

The degree of freedom of the NLS problem is $JGT - JT - 2T - G$, where JGT is the number of observations, JT results from ρ_{jt1} for each $j = 1, \ldots, J$ and $t = 1, \ldots, T$, $2T$ comes from (ρ_{1t2}, ρ_{1t3}) for each $t = 1, \ldots, T$, and G refers to one ω plus $(G-1) \times 1$ vector τ. The most stringent case is when $G = 2$, in which we need at least three products ($J \geq 3$) and $(J-2)T > G$. In practice, such NLS with the above constraints takes very little time and is robust to the choice of initial guess.

To see how we obtain the above constraints, note that in dynamic model, we have $\ln[\sigma_{jt}^{(g)}(U)/\sigma_{1t}^{(g)}(U)] = v_{jt}^{(g)}(U) - v_{1t}^{(g)}(U)$. Using the definition of the payoffs, we can compute the derivative:

$$\frac{\mathrm{d} \ln\left[\sigma_{jt}^{(g)}(U)/\sigma_{1t}^{(g)}(U)\right]}{\mathrm{d}U} = -\omega(P_{jt} - P_{1t})$$

By the series logit approximation, we have

$$\ln\left[\frac{\sigma_{jt}^{(g)}(U)}{\sigma_{1t}^{(g)}(U)}\right] = (\rho_{jt1} - \rho_{1t1}) + \left[\left(\frac{\rho_{jt2}}{\omega}\right) - \left(\frac{\rho_{1t2}}{\omega}\right)\right](\omega U + \tau^{(g)}) + \left[\left(\frac{\rho_{jt3}}{\omega^2}\right) - \left(\frac{\rho_{1t3}}{\omega^2}\right)\right](\omega U + \tau^{(g)})^2, \tag{B.5}$$

which implies

$$\frac{\mathrm{d} \ln\left[\sigma_{jt}^{(g)}(U)/\sigma_{1t}^{(g)}(U)\right]}{\mathrm{d}U} = \omega\left[\left(\frac{\rho_{jt2}}{\omega}\right) - \left(\frac{\rho_{1t2}}{\omega}\right)\right] + 2\left[\left(\frac{\rho_{jt3}}{\omega^2}\right) - \left(\frac{\rho_{1t3}}{\omega^2}\right)\right](\omega U + \tau^{(g)})\omega.$$

Equalizing the two formulas of the same derivative gives rise to the conclusion in eq. (B.4).

A useful conclusion is the following. Applying the constraints about ρ to eq. (B.5) for the first group, $g = 1$, we have

$$\ln\left[\frac{\sigma_{jt}^{(1)}(U)}{\sigma_{1t}^{(1)}(U)}\right] = (\rho_{jt,1} - \rho_{1t,1}) - \omega U(P_{jt} - P_{1t}).$$

The dependent variable Y_{jt} of eq. (Linear-Reg-1), whose definition is copied below, has a simple expression,

$$Y_{jt} \equiv \int \ln\left[\frac{\sigma_{jt}^{(1)}(U)}{\sigma_{1t}^{(1)}(U)}\right] \mathrm{d}\, F_t^{(1)}(U) + \omega(P_{jt} - P_{1t}) \int U \;\mathrm{d}\, F_t^{(1)}(U) = \rho_{jt,1} - \rho_{1t,1}.$$

This is useful, because the NLS step will estimate $\rho_{jt,1} - \rho_{1t,1}$. After which, one can estimate $(\delta_j - \delta_1)/(1-\beta)$, $\gamma/(1-\beta)$ and $\alpha^{(1)}$ by running 2SLS of $(\rho_{jt,1} - \rho_{1t,1})$ on $(X_{jt} - X_{1t})$ and $(P_{jt} - P_{1t})$ according to eq. (Linear-Reg-1). This also proves Proposition 2. ■

Proof of Proposition 3. For expositional simplicity, we omit the market index m in this proof. Also note that $U = (U_1, U_2)'$ involves two-dimensional unobserved heterogeneity, and U_1 and U_2 are associated with a variable MPGe and price, respectively. Correspondingly, $\tau^{(g)} = (\tau_1^{(g)}, \tau_2^{(g)})'$. Thus, what is below is a generalized proof of for two dimensions of heterogenity and not just one. We begin by first specifying the log market share ratio under each policy where the current policy follows from eq. (E.1), and the new policy takes the form of

$$\ln\left[\frac{\check{\sigma}_{jt}^{(g)}(U)}{\check{\sigma}_{0t}^{(g)}(U)}\right] = \frac{\delta_j}{1-\beta} + X_{jt}'\frac{\gamma_i}{1-\beta} - (\alpha^{(1)} + \tau_2^{(g)} + \omega_2 U_2)\tilde{P}_{jt} + \frac{\xi_{jt}}{1-\beta} - \beta\,\mathrm{E}[\bar{V}_{t+1}^{(g)}(U) \mid X_t, P_t, \xi_t], \quad \text{(B.6)}$$

where $\tilde{P}_{jt}$ is the new price of the EV vehicle with the proposed tax credit. Its series representation is that

$$\ln\left[\frac{\check{\sigma}_{jt}^{(g)}(U)}{\check{\sigma}_{0t}^{(g)}(U)}\right] = \tilde{\rho}_{jt1} + \tilde{\rho}_{jt2}'\Omega^{-1}\Big(\Omega U + \tau^{(g)}\Big) + \Big(\Omega U + \tau^{(g)}\Big)'\Omega^{-1}\tilde{\rho}_{jt3}\Omega^{-1}\Big(\Omega U + \tau^{(g)}\Big),$$

with unknown coefficients $\tilde{\rho}_{jt1}, \tilde{\rho}_{jt2}$, and $\tilde{\rho}_{jt3}$. Ω is a diagonal matrix whose diagonal elements are ω_1 and ω_2. The objective is to solve for $\tilde{\rho}_{jt1}, \tilde{\rho}_{jt2}$, and $\tilde{\rho}_{jt3}$. Define $\tilde{P}_{jt} = P_{jt} + \Delta P_{jt}$. First note that by eq. (E.1), we have

$$\begin{aligned}
\ln\left[\frac{\check{\sigma}_{jt}^{(g)}(U)}{\check{\sigma}_{0t}^{(g)}(U)}\right] &= \ln\left[\frac{\sigma_{jmt}^{(g)}(U)}{\sigma_{0mt}^{(g)}(U)}\right] - (\alpha^{(1)} + \tau_2^{(g)} + \omega_2 U_2)\Delta P_{jt} \\
&= \rho_{jt1} + \rho_{jt2}'\Omega^{-1}\Big(\Omega U + \tau^{(g)}\Big) + \Big(\Omega U + \tau^{(g)}\Big)'\Omega^{-1}\rho_{jt3}\Omega^{-1}\Big(\Omega U + \tau^{(g)}\Big) \\
&\quad - (\alpha^{(1)} + \tau_2^{(g)} + \omega_2 U_2)\Delta P_{jt} \\
&= \rho_{jt1} - (\alpha^{(1)} + \tau_2^{(g)})\Delta P_{jt} + \rho_{jt2}'\Omega^{-1}\Big(\Omega U + \tau^{(g)}\Big) + \Big(\Omega U + \tau^{(g)}\Big)'\Omega^{-1}\rho_{jt3}\Omega^{-1}\Big(\Omega U + \tau^{(g)}\Big) \\
&\quad - \omega_2 U_2 \Delta P_{jt}.
\end{aligned}$$

Note that ρ_{jt2} is a two-dimensional vector:

$$\rho_{jt2} = \begin{pmatrix} \rho_{jt2,1} \\ \rho_{jt2,2} \end{pmatrix} \qquad \text{and} \qquad \rho_{jt2}'\Omega^{-1} = (\rho_{jt2,1}\omega_1^{-1}, \rho_{jt2,2}\omega_2^{-1}),$$

where $\rho_{jt2,1}$ is associated with MPGe, and $\rho_{jt2,2}$ is with price. Next, we write $\omega_2 U_2 \Delta P_{jt} = (\omega_2 U_2 + \tau_2^{(g)})\Delta P_{jt} - \tau_2^{(g)}\Delta P_{jt}$ so that

$$\ln\left[\frac{\check{\sigma}_{jt}^{(g)}(U)}{\check{\sigma}_{0t}^{(g)}(U)}\right] = \rho_{jt1} - \alpha^{(1)}\Delta P_{jt} + \begin{pmatrix} \rho_{jt2,1} \\ \rho_{jt2,2} - \Delta P_{jt}\omega_2 \end{pmatrix} \Omega^{-1}\Big(\Omega U + \tau^{(g)}\Big) + \frac{\rho_{jt3}}{\omega^2}\Big(\Omega U + \tau^{(g)}\Big)^2.$$

Thus,

$$\tilde{\rho}_{jt1} = \rho_{jt1} - \alpha^{(1)}\Delta P_{jt}, \qquad \tilde{\rho}_{jt2} \equiv \begin{pmatrix} \tilde{\rho}_{jt2,1} \\ \tilde{\rho}_{jt2,2} \end{pmatrix} = \begin{pmatrix} \rho_{jt2,1} \\ \rho_{jt2,2} - \Delta P_{jt}\omega_2 \end{pmatrix}, \qquad \tilde{\rho}_{jt3} = \rho_{jt3}.$$

■

C Initial values

Good initial values help solve the NLS in eq. (5). We need initial values of $\tau_{\text{init}} = (\tau_{\text{init}}^{(2)}, \ldots, \tau_{\text{init}}^{(G)})'$, ρ_{init}, and ω_{init}. We follow two steps to obtain the initial values, and these steps are based on the first order Taylor expansion of CCP function $\sigma_{jt}^{(g)}(U)$.

In the *first step*, we find τ_{init} by running 2SLS for the following linear regression,

$$\ln\left(\frac{S_{jt}^{(g)}}{S_{1t}^{(g)}}\right) = \frac{\delta_j - \delta_1}{1-\beta} + (X_{jt} - X_{1t})'\frac{\gamma}{1-\beta} - (\alpha^{(1)} + \tau^{(g)})(P_{jt} - P_{1t}) + \frac{\xi_{jt} - \xi_{1t}}{1-\beta}.$$

To see the rationale, recall the identity

$$S_{jt}^{(g)} = \mathrm{E}[\sigma_{jt}^{(g)}(U^*)\Gamma_t^{(g)}(U^*)], \qquad U^* \sim \mathcal{N}(0,1),$$

and consider the first order Taylor expansion of CCP function $\sigma_{jt}^{(g)}(U^*)\Gamma_t^{(g)}(U^*)$ at 0, which is the mean of $U^* \sim \mathcal{N}(0,1)$, for each group g. We have

$$\begin{aligned} S_{jt}^{(g)} &= \mathrm{E}[\sigma_{jt}^{(g)}(U^*)\Gamma_t^{(g)}(U^*)] \\ &\approx \sigma_{jt}^{(g)}(0)\Gamma_t^{(g)}(0) + \left(\frac{\mathrm{d}\sigma_{jt}^{(g)}(U^*)\Gamma_t^{(g)}(U^*)}{\mathrm{d}U^*}\right)_{U^*=0} \mathrm{E}(U^* - 0) \\ &= \sigma_{jt}^{(g)}(0)\Gamma_t^{(g)}(0). \end{aligned}$$

The first order Taylor expansion leads to $S_{jt}^{(g)} \approx \sigma_{jt}^{(g)}(0)\Gamma_t^{(g)}(0)$. Thus,

$$\frac{S_{jt}^{(g)}}{S_{1t}^{(g)}} \approx \frac{\sigma_{jt}^{(g)}(0)}{\sigma_{1t}^{(g)}(0)}.$$

The application of this conclusion to eq. (Linear-Reg-1) when $U = 0$ gives rise to the stated regression.

In the *second step*, we find $\rho_{jt1,\text{init}}$ for all $j = 1, \ldots, J$, and $(\rho_{1t2}/\omega)_{\text{init}}, (\rho_{1t3}/\omega^2)_{\text{init}}$ by running OLS for the following linear regression,

$$\ln\left(\frac{S_{jt}^{(g)}}{S_{0t}^{(g)}}\right) + (P_{jt} - P_{1t})\tau_{\text{init}}^{(g)} = \rho_{jt1} + \left(\frac{\rho_{1t2}}{\omega}\right)\tau_{\text{init}}^{(g)} + \left(\frac{\rho_{1t3}}{\omega^2}\right)(\tau_{\text{init}}^{(g)})^2.$$

We now explain how we got the above regression. It follows from series logit that

$$\begin{aligned} \ln\left[\frac{\sigma_{jt}^{(g)}(0)}{\sigma_{0t}^{(g)}(0)}\right] = \ln\left[\frac{\sigma_{jt}^{(1)}(\tau^{(g)}/\omega;\rho_t)}{\sigma_{0t}^{(1)}(\tau^{(g)}/\omega;\rho_t)}\right] &= \rho_{jt1} + \left(\frac{\rho_{jt2}}{\omega}\right)\tau^{(g)} + \left(\frac{\rho_{jt3}}{\omega^2}\right)(\tau^{(g)})^2 \\ &= \rho_{jt1} + \left(\frac{\rho_{1t2}}{\omega}\right)\tau^{(g)} - (P_{jt} - P_{1t})\tau^{(g)} + \left(\frac{\rho_{1t3}}{\omega^2}\right)(\tau^{(g)})^2. \end{aligned}$$

The second line follows from imposing the constraints eq. (B.4). By the approximation,

$$\ln\left(\frac{S_{jt}^{(g)}}{S_{0t}^{(g)}}\right) \approx \ln\left[\frac{\sigma_{jt}^{(g)}(0)}{\sigma_{0t}^{(g)}(0)}\right],$$

we have the stated regression.[1]

D Multidimensional Unobserved Heterogeneity

In this appendix, we extend our main results to include multidimensional unobserved heterogeneity. We have two observations. First, our estimation method works for multidimensional unobserved heterogeneity after some modification in the stage of CCP estimation. Second, a higher dimension of unobserved heterogeneity does not cause the curse of dimensionality for our CCP estimation that involves a series polynomial approximation of CCP as a function of multidimensional unobserved heterogeneity. This is because the structural model imposes certain restrictions that can eliminate a large number of parameters in CCP function.

D.1 Model

We now have the new dimension of unobserved heterogeneity γ_i associated with product characteristics X_{jt}. Particularly, when X_{jt} includes the product dummy variable, the above specification says that consumers could have heterogeneous valuation about the unobserved product characteristics (e.g. advertising).[2]

Using our group specification, we write

$$\begin{pmatrix}\gamma_i \\ \alpha_i\end{pmatrix} = \begin{pmatrix}\gamma^{(1)} \\ \alpha^{(1)}\end{pmatrix} + D_i^{(2)}\begin{pmatrix}\tau_1^{(2)} \\ \tau_2^{(2)}\end{pmatrix} + \cdots + D_i^{(G)}\begin{pmatrix}\tau_1^{(G)} \\ \tau_2^{(G)}\end{pmatrix} + \varOmega\begin{pmatrix}U_{i1} \\ U_{i2}\end{pmatrix}, \tag{D.1}$$

where $\varOmega$ is a diagonal matrix,

$$\varOmega \equiv \begin{pmatrix}\varOmega_1 & \\ & \omega_2\end{pmatrix}.$$

The diagonal elements $\varOmega_1$, which is also a diagonal matrix, and ω_2 determine the within group variation of γ_i and α_i, respectively. Below, let $\tau^{(g)} \equiv (\tau_1^{(g)'}, \tau_2^{(g)})'$ and let $U_i \equiv (U_{i1}', U_{i2})'$. Use q to denote the dimension of $(X_{jt}', P_{jt})'$. Again, let $\tau^{(1)} \equiv \mathbf{0}$. Let $\boldsymbol{\phi}(U)$ denote the PDF of the multivariate normal distribution $\mathcal{N}(\mathbf{0}, \mathbf{I})$.

Under the specification of (γ_i', α_i) in Equation (D.1), the expected lifetime payoff of purchasing product j at time t becomes

$$v_{jt}^{(g)}(U_i) = \frac{\delta_j + \gamma_i' X_{jt} + \xi_{jt}}{1-\beta} - \alpha_i P_{jt}, \qquad j = 1, \ldots, J.$$

[1] We do not have a clever initial value of ω_{init}. In our simulation, we varied the initial value ω_{init} substantially, and our optimization routine seems very robust.

[2] For example, suppose X_{jt} is just the product dummy variable that equals 1 for product j and 0 otherwise. The expected payoff of product j in period t then reads $v_{ijt} = \delta_j + \gamma_i + \xi_{jt} - \alpha_i P_{jt}$, and γ_i serves as random effect that explains consumer heterogeneity in the valuation of unobserved product characteristics.

Correspondingly, the CCP of type-(g, U) is

$$\sigma_{jt}^{(g)}(U) = \frac{\exp(v_{jt}^{(g)}(U))}{\exp(v_{0t}^{(g)}(U)) + \sum_{k=1}^{J} \exp(v_{kt}^{(g)}(U))}.$$

Because $v_{jt}^{(g)}(U) = v_{jt}^{(1)}(U + \Omega^{-1}\tau^{(g)})$, we still have shifting formula $\sigma_{jt}^{(g)}(U) = \sigma_{jt}^{(1)}(U + \Omega^{-1}\tau^{(g)})$. So the CCP estimation can be based on the same NLS problem excepting for the constraints about the series approximation coefficients. Once the CCP functions are known, it will be easy to estimate the model structural parameters using our procedures in the post-CCP estimation section.

The constraints about ρ_t result from the comparison between the derivatives of the log CCP ratio with respect to U in the structural demand model and the same derivatives in the series approximation. Particularly, we have

$$\begin{aligned}
\ln\left[\frac{\sigma_{jt}^{(g)}(U)}{\sigma_{1t}^{(g)}(U)}\right] &= v_{jt}^{(g)}(U) - v_{1t}^{(g)}(U) \\
&= \frac{\delta_j - \delta_1}{1-\beta} + \frac{(\gamma^{(1)} + \tau_1^{(g)})'}{1-\beta}(X_{jt} - X_{1t}) - (\alpha^{(1)} + \tau_2^{(g)})(P_{jt} - P_{1t}) + \frac{\xi_{jt} - \xi_{1t}}{1-\beta} + \\
&\quad ((1-\beta)^{-1}\Omega_1 U_1)'(X_{jt} - X_{1t}) - \omega_2 U_2 (P_{jt} - P_{1t}).
\end{aligned}$$

This gives rise to

$$\frac{\mathrm{d}\ln[\sigma_{jt}^{(g)}(U)/\sigma_{1t}^{(g)}(U)]}{\mathrm{d}U} = \begin{pmatrix} (1-\beta)^{-1}\Omega_1(X_{jt} - X_{1t}) \\ -\omega_2(P_{jt} - P_{1t}) \end{pmatrix},$$

and any higher order derivatives are zeros. Comparing the above derivatives with the resulted derivatives from series logit form, we conclude

$$\rho_{jt2} - \rho_{1t2} = \begin{pmatrix} (1-\beta)^{-1}\Omega_1(X_{jt} - X_{1t}) \\ -\omega_2(P_{jt} - P_{1t} \end{pmatrix}, \quad (\rho_{jt3} - \rho_{1t3}) = (\rho_{jt4} - \rho_{1t4}) = \cdots = 0.$$

In the series approximation, we only consider the second order approximation. So for the dynamic model with multidimensional heterogeneity, the CCP estimation stage involves the following unknowns,

$$(\rho_{1t1}, \ldots, \rho_{Jt1})', \quad \rho_{1t2}, \quad \rho_{1t3}, \quad (1-\beta)^{-1}\Omega_1, \quad \omega_2, \quad \tau, \qquad \text{for all } t.$$

Note that in this step of CCP estimation, we are only able to estimate $(1-\beta)^{-1}\Omega_1$ as a whole and cannot separately estimate β from Ω_1. Recall q is the dimension of $(X_{jt}', P_{jt})'$, and ρ_{1t3} is a $q \times q$ triangular matrix. The degree of freedom of the NLS problem is $GJT - (JT + (3q + q^2)T/2 + 2q)$. In order to make this degree of freedom be positive, it is necessary to satisfy $(G-1)J - (3q + q^2)/2 > 0$.[3] Depending on the number of dimension of unobserved heterogeneity, we may or may not need a large number of groups or products. It is interesting to point out that even when the number of groups is small, we can ensure a positive degree of freedom by including a large number of products. This manifests one advantage of our method—the number of products, rather than causing the curse of dimensionality, helps solve the curse of dimensionality if we are willing to assume that purchasing is a terminal action in the dynamic model, which is reasonable for the market of durable goods.

[3] We write $GJT - (JT + (3q+q^2)T/2 + 2q) = [(G-1)J - (3q+q^2)/2]T - 2q$. When T is relatively large, $(G-1)J - (3q+q^2)/2 > 0$ will also be sufficient to have positive degree of freedom.

Table D.1: Simulation Results: Multi Dimensional Heterogeneity
DGP: $M = 2$, $T = 48$ and $J = 8$

$\delta = -0.1$	-0.0923 (0.0097)
$\gamma = 0.03$	0.0289 (0.0017)
$\alpha^{(1)} = 0.10$	0.0963 (0.0175)
$\tau_X^{(2)} = 0.05$	0.0484 (0.0032)
$\tau_X^{(3)} = 0.10$	0.0968 (0.0063)
$\tau_X^{(4)} = 0.15$	0.1453 (0.0098)
$\tau_X^{(5)} = 0.20$	0.1942 (0.0137)
$\tau_X^{(6)} = 0.25$	0.2434 (0.0182)
$\tau_p^{(2)} = 0.05$	0.0507 (0.0016)
$\tau_p^{(3)} = 0.10$	0.1014 (0.0034)
$\tau_p^{(4)} = 0.15$	0.1519 (0.0053)
$\tau_p^{(5)} = 0.20$	0.2024 (0.0073)
$\tau_p^{(6)} = 0.25$	0.2528 (0.0092)
$\omega_X = 0.1$	0.0980 (0.0065)
$\omega_p = 0.075$	0.0754 (9.72e-5)
$\beta = 0.90$	0.9088 (0.0078)

Note: Mean and standard deviation (in parenthesis) for 50 simulations. Starting values follow the procedure in the appendix and vary with each simulation run. The starting value for $\omega = 0.25$ for all simulation runs.

D.2 Multi Dimensional Monte Carlo Simulation

In Table D.1, we report the results of Monte Carlo simulations where within group unobserved heterogeneity is present in both price and the X variable. The data generating process is identical to the uni-dimensional simulations with the exception of the Gaussian Hermite quadrature approximation of the normal distribution for the price and X coefficients uses 6 nodes rather than 12. This is done for computation time as with 12 we would have 144 individuals to simulation in the DGP process.

E Empirical Application: Estimation of $\delta_{make,1}$, β and $\gamma_{2k}/(1 - \beta q_{cnty,k})$

The next step is to use

$$\ln\left[\frac{\sigma_{1t}^{(g)}(U)}{\sigma_{0t}^{(g)}(U)}\right] = \frac{\delta_{make,1}}{1-\beta} + \sum_{k=1}^{d_{ev}} X_{1mt,k}^{ev} \frac{\gamma_{1k}}{1-\beta q_{ev,k}} + \sum_{k=1}^{d_{cnty}} X_{mt,k}^{cnty} \frac{\gamma_{2k}}{1-\beta q_{cnty,k}}$$
$$+ \left(\frac{\eta^{(1)}}{1-\beta} + \frac{\tau_{mpge}^{(g)}}{1-\beta} + \omega_{mpge} U_{mpge}\right) MPGe_{1mt}$$
$$- (\alpha^{(1)} + \tau_p^{(g)} + \omega_p U_p) P_{1mt} + \frac{\xi_{1mt}}{1-\beta} + \beta\, \mathrm{E}[\bar{V}_{t+1}^{(g)}(U) \mid X_{mt}, P_{mt}, \xi_{mt}]. \quad \text{(E.1)}$$

to recover $\delta_{make,1}$, $\gamma_{2k}/(1 - \beta q_{cnty,k})$ $(k = 1, \ldots, d_{cnty})$ and the discount factor β.

Letting

$$W^{(g)}_{mt}(U) \equiv \ln\left[\frac{\sigma^{(g)}_{1t}(U)}{\sigma^{(g)}_{0t}(U)}\right] -$$

$$\left[\sum_{k=1}^{d_{ev}} X^{ev}_{1mt,k}\frac{\gamma_{1k}}{1-\beta q_{ev,k}} + \left(\frac{\eta^{(1)}}{1-\beta} + \frac{\tau^{(g)}_{mpge}}{1-\beta} + \omega_{mpge}U_{mpge}\right)MPGe_{1mt} - (\alpha^{(1)} + \tau^{(g)}_{p} + \omega_p U_p)P_{1mt}\right], \quad \text{(E.2)}$$

we have the conclusion following the similar arguments in section 4.2:

$$W^{(g)}_{mt}(\tilde{U}) = \delta_{make,1} + \sum_{k=1}^{d_{cnty}} X^{cnty}_{mt,k}\frac{\gamma_{2k}}{1-\beta q_{cnty,k}} + \frac{\xi_{1mt}}{1-\beta} +$$

$$\beta\,\mathrm{E}\left(W^{(g)}_{m,t+1}(\tilde{U}) - \sum_{k=1}^{d_{cnty}} X^{cnty}_{m,t+1,k}\frac{\gamma_{2k}}{1-\beta q_{cnty,k}} + \ln\sigma^{(g)}_{0,m,t+1}(\tilde{U}) - \frac{\xi_{1,m,t+1}}{1-\beta}\,\middle|\, X_{mt}, P_{mt}, \xi_{mt}\right), \quad \text{(E.3)}$$

for a fixed $\tilde{U}$.

To estimate the discount factor β and γ_2, one possible approach is to consider the unconditional expectation again and to have the following conclusion:

$$\mathrm{E}[W^{(g)}_{mt}(\tilde{U})] = \delta_{make,1} + \beta\,\mathrm{E}[W^{(g)}_{m,t+1}(\tilde{U}) + \ln\sigma^{(g)}_{0,m,t+1}(\tilde{U})] - \sum_{k=1}^{d_{cnty}} \mathrm{E}[(X^{cnty}_{mt,k} - \beta X^{cnty}_{m,t+1,k})]\frac{\gamma_{2k}}{1-\beta q_{cnty,k}} = 0,$$

for a fixed $\tilde{U}$. Then we can solve β and $\gamma_{2k}/(1-\beta q_{cnty,k})$ from the above equation by the variation of groups. For this particular application, this approach is infeasible because some of our state variables are nonstationary as illustrated by the time series plot in Figure 3. For example, the number of charging stations is clearly non-stationary.

As an alternative, we note that in eq. (E.3), for a fixed $\tilde{U}$, the random variable $W^{(g)}_{mt}(\tilde{U})$ is a function of the vector $(X_{mt}, P_{mt}, \xi_{mt})$. So we have a conditional moment condition:

$$\mathrm{E}\left(W^{(g)}_{mt}(\tilde{U}) - \beta W^{(g)}_{m,t+1}(\tilde{U}) - \delta_{make,1} - \beta\ln\sigma^{(g)}_{0,m,t+1}(\tilde{U}) - \sum_{k=1}^{d_{cnty}}(X^{cnty}_{mt,k} - \beta X^{cnty}_{m,t+1,k})\frac{\gamma_{2k}}{1-\beta q_{cnty,k}} - \frac{\xi_{1mt} - \beta\xi_{1,m,t+1}}{1-\beta}\,\middle|\, X_{mt}, P_{mt}, \xi_{mt}\right) = 0.$$

Suppose there is a vector of $X_{mt,IV}$, which are elements or functions of the conditioning variables $(X_{mt}, P_{mt}, \xi_{mt})$ and satisfy

$$\mathrm{E}(X_{mt,IV}\xi_{1mt}) = \mathrm{E}(X_{mt,IV}\xi_{1m,t+1}) = 0.$$

We have the moment conditions of

$$\mathrm{E}\left[X_{mt,IV}\left(W^{(g)}_{mt}(\tilde{U}) - \beta W^{(g)}_{m,t+1}(\tilde{U}) - \delta_{make,1} - \beta\ln\sigma^{(g)}_{0,m,t+1}(\tilde{U}) - \sum_{k=1}^{d_{cnty}}(X^{cnty}_{mt,k} - \beta X^{cnty}_{m,t+1,k})\frac{\gamma_{2k}}{1-\beta q_{cnty,k}}\right)\right] = 0.$$

We then can estimate $\delta_{make,1}$, β and $\gamma_{2k}/(1-\beta q_{cnty,k})$ from the above equation using the moment estimator. In our empirical estimates, we let $\tilde{U} = 0$ and define $X_{mt,IV}$ as a vector of ones, gas^{cnty}_{mt}, the log battery price in period t, and the log number of EV makes in a given period t.