Counterfactual Analyses with New Goods in Differentiated Products Demand Systems

Ali Eren Çamur*

Tilburg University (CentER)
Research Master Thesis

Main Supervisor: Dr. Christoph Walsh †
Second Reader: Prof. Dr. Tobias Klein ‡

Defense Date : August 31^{th}, 2021

Abstract

Researchers in the field of the empirical industrial organization perform copious amounts of counterfactual analyses, either in the context of the introduction of new good(s) or various merger policies. Yet, they make some simplistic assumptions on the nature of the unobserved product characteristics. The choice on the nature of these unobserved product characteristics, in return, can result in a broad difference in terms of the prices and welfare for policy experiments. This master's thesis intends to analyze this phenomenon: By synthetically producing a dataset, we assume that a conjectural researcher estimates the general model and performs the counterfactual simulations in the context of the introduction of the existing product, based on the parameter estimates. The resulting profits and welfare of the prementioned simulation exercises demonstrate the importance concerning the inference on the nature of the unobserved product attributes. Therefore, our findings indicate that the applied researchers should choose the nature of these unobserved product characteristics in their policy experiments carefully.

*Research Master Student in Economics, Tilburg University
†Department of Econometrics and Operations Research, Tilburg University
‡Department of Econometrics and Operations Research, Tilburg University

Contents

1 INTRODUCTION

In the last 50 years, the empirical industrial organization literature primarily evolved around the demand systems in characteristic space, especially pioneered by the works of McFadden around the 1970s.[1] As Ackerberg et al. (2007)[1] pointed out in their famous and comprehensive work, the empirical industrial organization literature was mainly focused on the demand systems in product space with the common representative agent framework, before these lofty seminal works of McFadden. The drawbacks of this latter approach, however, are significant and cannot be easily ignored.[2] The demand systems in characteristic space can deal with these drawbacks that product space systems face, mainly thanks to the fact that the products are fundamentally bundles of several characteristics, and the preferences of consumers are defined on those characteristics (Ackerberg et al. (2007)). However, note that the demand systems in characteristic space might not fully circumnavigate the aforementioned issues which the demand systems in product space suffer from.[3] Nevertheless, they are currently the most appealing toolbox the empirical researchers use.

Arguably, the most notorious advance in the context of discrete-choice differentiated products systems is due to the well-known ”BLP” framework (1995)[5]. Mainly using the insights of Berry (1994)[4], the authors devised a framework in which the usage of market-level data suffices to exploit far more reasonable substitution patterns and parameter estimates, compared to the plain logit specification.[4] By contriving a contraction mapping in the estimation procedure, they developed a niche algorithm to adapt Berry's findings (1994) to the mixed logit framework. They claimed their contribution to the literature at the last section of their paper as follows:

”Our model is defined in terms of four primitives and a Nash equilibrium assumption in prices. The primitives are the utility surface that assigns values to different possible combinations of product characteristics as a function of consumer characteristics, a cost function which determines the production cost associated with different combinations of product characteristics,

[1] In this regard, he made several important contributions, but presumably his most sonorous papers are [12] and [13]. In the former paper, he introduced the conditional logit analysis with a specific emphasis on the well-known I.I.A. property; and he extended his analysis to urban travel demand with the introduction of the multinomial logit specification in his latter work.

[2] Such as the curse of dimensionality issues (e.g. with having J products with this approach, one needs to estimate J^2 parameters); the inability to incorporate the introduction of the new good to the general procedure; and the lack on the possible differences in ”tastes” concerning heterogeneity among consumers. For a comprehensive review on these issues, one can consult Ackerberg et al. (2007).

[3] For example, as mentioned in Ackerberg et al. (2007), if the newly introduced good is ”too” new, then it might not be reasonable to deduce the possible preferences for this good since the characteristics of this new good will be presumably a non-convex combination of the characteristics of already existing products.

[4] Not only he introduced the specification of the unobserved attributes to the empirical industrial organization literature, Berry (1994) provided an elegant way for solving the problem of having to use the nonlinear instrumental variables in the context of discrete-choice differentiated products framework: He proved that the observed market shares and model market shares must be equal to each other at the true value of the mean utility components, and he provided an inversion technique to ensure linearity between these factors.

a distribution of consumer characteristics, and a distribution of product characteristics. Conditional on these primitives the model can solve for the distribution of prices, quantities, variable profits, and consumer welfare."

The advantages of the BLP framework, as it has been shown above from the direct citation, are numerous: Not only this model can solve for the main determinants of a given industry, but it can do that with relatively low information at hand. In addition, their model is a very flexible one in the sense that general adjustments can be made in most places of their approach to adapt their setting to different setups and situations. This is one of the main reasons which renders this framework a highly popular one within the field of the empirical industrial organization.

Many papers have been written regarding the BLP framework thereafter. Some of these papers tried to facilitate the interface of the BLP framework, whereas the other papers provided further insights related to the micro details of its general procedures. They also tried to implement some practices regarding the methods that have been employed.[5]In addition, there exists some exhaustive meta-studies that have been conducted in respect of the BLP setup.[6]

As it has been made clear, the BLP framework is hugely popular among the empirical researchers, and the main advantage of this framework is to enable the researchers to conduct several policy experiments, such as mergers and the introduction of new product(s) for the markets of interest. One important, yet generally disregarded fact is that when the applied researchers perform these counterfactual exercises, they need to decide how they should infer the values of the unobserved product characteristics, generally in

[5]To give some examples, Nevo (2000x)[14] provided further clarification on the nature of the BLP specification, and he also demonstrated how product-specific dummy variables can be added to the common procedure; Petrin (2002)[17] exemplified the introduction of a new product (in his case, minivans) and how additional moments related to average demographics can ameliorate the welfare results; Berry et al. (2007)[6] introduced the "pure characteristics" model in which they omitted the idiosyncratic shock from the demand side and checked its general properties. Lastly, Dubé et al. (2012)[9] argued the importance of both inner and outer loop tolerance levels therein the process of BLP estimation. They also adjusted the estimation routine in the sense that they constrained a new GMM minimization problem with additional equilibrium constraints (in terms of shares and moment restrictions), and they checked the difference -mainly in terms of efficiency and timing- between the conventional approach and their so-called "MPEC" approach. To learn more about these mentioned papers, please consult them.

[6]The works of Knittel et al. (2014)[11] and Conlon et al. (2020)[8] can be considered as comprehensive meta-studies related to the BLP framework: In the former one, the authors engaged attention to the general drawbacks of the BLP model: By using the original BLP dataset and a pseudo data for Nevo's setting[15], they showed that in most instances, the optimization routine fails to fulfill the proper first and second-order optimality conditions, and there is a huge variety both in terms of welfare and elasticity results when they convey the estimation routine with different optimization algorithms. In the latter one, the authors tried to present several practices in the context of BLP framework: They investigated various different algorithmic advances regarding the general BLP structure, such as the incorporation of fixed effects, integration of the heterogeneity, and solving for shares, all with diverse approaches. This paper has also a very clear exposition on how the optimal instruments can be constructed, by mainly following the prior footsteps of Amemiya (1977)[3] and Chamberlain (1987)[7], as they mentioned. For further details, one can consult these mentioned papers.

ex-ante scenarios. Yet, they generally prefer to make some simplifying assumptions on the nature of these unobservables.[7]The choice of these unobserved characteristics, however, might gravely affect the counterfactual prices, profits, and consumer welfare. The motivation of this thesis is to present to the readers how various different picks of those components could affect both counterfactual scenario profits and total consumer surplus. In other words, this thesis intends to present a sensitivity check on the counterfactual prices and consumer surplus calculations with respect to the choice of unobserved product characteristics of the brands, in the context of the introduction of the existing product.

To convey these simulation exercises, we first construct our model specification and synthetically produce a dataset.[8]After handling the data simulation part, we assume that a conjectural researcher -having access to the premade data- estimates the model and performs the counterfactual analyses where an existing good enters into a new market. When the researcher performs these counterfactual exercises and infers the value of the unobserved product characteristics of the ”missing” brand, s/he has several options to deduce its unobserved product attributes.

Indeed, our results show us that the counterfactual prices, profits, and total consumer welfare depend on the pick of the unobserved product attributes: By inferring the unobserved product characteristics of the ”missing” product from several different means, we arrive at different results both for the counterfactual profits and total consumer welfare. These counterfactual results indicate that the choice of the unobserved product characteristics in policy simulations is indeed momentous. Therefore, the applied researchers in the field of the empirical industrial organization should be more careful while inferring the values of the unobserved product characteristics in their policy experiments.

The rest of this thesis is constructed as follows: In section **2**, we briefly present the baseline model specification that we use both in the plain logit and mixed logit frameworks. We present the methodology of the thesis and describe the general data generating process, as well as the estimation strategy and the details of the counterfactual analyses in section **3**. Section **4** presents the results of the estimation procedures, together with the results of the policy experiments. Eventually, section **5** concludes, as well as pointing out some future research.

[7]For example, Igami (2017)[10] investigated the creative destruction process in the HDD industry, and he performed several related counterfactual exercises with different patent systems and different license fees regarding the innovation in this industry. In his work, he inferred the counterfactual values of the unobserved product characteristics from some time-series (specifically AR(1)) specification. In another paper, Olesiński (2020)[16] tried to evaluate the ex-ante effects of EU-related cigarette prohibitions on the Polish market by synthetically removing the menthol cigarettes from the Polish market and by introducing some adjustments on the tax schedule regarding the tobacco industry in the same market. In these counterfactual experiments, he inferred the values of the unobserved components directly from the error term coming from the joint BLP estimation as single data points.

[8]In fact, the ideal approach consists of simulating many datasets and performing both the estimation and counterfactual analyses for all of these ”synthetic” datasets, and lastly, getting the average. We talk about this issue in subsection **3.2**.

2 MODEL SPECIFICATION

Herein, we present the general sketch of the model that we will use for the aforementioned counterfactual exercises by discussing both the demand and supply specifications. This part of the thesis aims to introduce the readers to the chosen model specification in a brief sense.[9]

2.1 Demand Side

In this subsection, we present to the readers the demand side specification for both the plain logit and mixed logit models.

2.1.1 Demand Side in Plain Logit Specification

Regarding the specification of the demand side for the plain logit case, we mainly follow the prior footsteps of Nevo $(2000x)$. In this setting, the consumers have the following utility schemes:[10]

$$U_{ijmt} = x_{jmt} \cdot \beta - \alpha \cdot p_{jmt} + \xi_{jmt} + \epsilon_{ijmt} \tag{1}$$

First of all, we assume that there exists M markets, T time periods, I_{mt} consumers and J_{mt} brands across those markets and time periods. In above specification, the term x_{jmt} corresponds to the observed attributes of brand j in market m at time period t. Analogously, p_{jmt} and ξ_{jmt} refer to the observed market-level prices and unobserved product characteristics respectively. Besides, ϵ_{ijmt} represents the idiosyncratic taste shock of the individual i for brand j in market m at time period t. Moreover, the β coefficients correspond to the linear demand side parameters, and α represents the price sensitivity parameter, in this setup.[11]

With independent and identically distributed ϵ_{ijmt} random variables belonging to the Type 1 extreme value distribution, we get the closed-form market-level shares expression as follows (Nevo $(2000x)$):[12]

$$s_{jmt}(\delta_{.mt}; \beta, \alpha) = \frac{exp(x_{jmt} \cdot \beta - \alpha \cdot p_{jmt} + \xi_{jmt})}{1 + \sum_{\mathfrak{z} \in J_{mt}} exp(x_{\mathfrak{z}mt} \cdot \beta - \alpha \cdot p_{\mathfrak{z}mt} + \xi_{\mathfrak{z}mt})} = \frac{exp(\delta_{jmt})}{1 + \sum_{\mathfrak{z} \in J_{mt}} exp(\delta_{\mathfrak{z}mt})}\,^{13} \tag{2}$$

[9] Additionally, one can find an itemized version of the model's primitives in subsection 6.1.

[10] U_{ijmt} stands for the overall utility individual i gets by choosing brand j in market m at time period t.

[11] The general characterization in this paragraph also applies to the mixed logit framework, albeit with some changes in the nature of the price sensitivity parameter.

[12] Type 1 extreme value distribution is also known as Gumbel distribution.

[13] Throughout this thesis, the expression $.mt$ indicates the dependency of all goods in the specific market m and time period t combination. We infer this approach from Conlon et al. (2020).

Here, the term $\delta_{jmt}(x_{jmt}, p_{jmt}, \xi_{jmt}; \beta, \alpha)$ corresponds to the *mean utility* consumers get from choosing product j in market m at time period t. Arguably, the biggest advantage of the plain logit framework is that the analytical formula governing market-level shares is closed-form and it is computationally very cheap. However, the main downside of this approach is that the consumer heterogeneity enters the functional form only via the idiosyncratic error term, which results in some counter-intuitive aspects (Ackerberg et al. (2007)). For a comprehensive investigation on those aspects and further explanations, one can consult the first chapter of Ackerberg et al. (2007).

2.1.2 Demand Side in Mixed Logit Specification

Closely following the model specifications both on Nevo (2000x) and on Conlon et al. (2020), yet only incorporating one random coefficient (which is on price variable), the consumers have the following utility schemes in our mixed logit framework:

$$U_{ijmt} = x_{jmt} \cdot \beta - \alpha_i \cdot p_{jmt} + \xi_{jmt} + \epsilon_{ijmt} \tag{3}$$

The interpretation of the variables x_{jmt}, p_{jmt}, ξ_{jmt}, and ϵ_{ijmt}, as well as the linear demand side parameters β is identical compared to the given explanations in the plain logit specification in subsection **2.1.1**. Nonetheless, the general structure of the price sensivity parameter changes: Now, we specify the *individual-specific taste coefficients* take the form $\alpha_i = \alpha + \sigma_\alpha \cdot v_{i\alpha}$, where σ_α stands for the deviation parameter for price, and the term v_i represents the unobserved individual characteristics, interacted with the deviation parameter.[14]

With this novel characterization on the utility schemes, we can write equation (3) as:

$$U_{ijmt} = x_{jmt} \cdot \beta - \Big[\alpha + \sigma_\alpha \cdot v_{i\alpha}\Big] \cdot p_{jmt} + \xi_{jmt} + \epsilon_{ijmt} \tag{4}$$

So that the general structure of utility schemes concludes as:

$$U_{ijmt} = \underbrace{x_{jmt} \cdot \beta - \alpha \cdot p_{jmt} + \xi_{jmt}}_{\delta_{jmt}(x_{jmt}, p_{jmt}, \xi_{jmt}; \beta, \alpha)} - \underbrace{p_{jmt} \cdot \sigma_\alpha \cdot v_{i\alpha}}_{\mu_{ijmt}(p_{jmt}, v_i; \sigma_\alpha)} + \ \epsilon_{ijmt}{}^{15} \tag{5}$$

In equation (5), the term $\delta_{jmt}(x_{jmt}, p_{jmt}, \xi_{jmt}; \beta, \alpha)$ corresponds to the *mean utility* individuals obtain from choosing product j in market m at time period t; whereas the terms

[14] Where $v_i \sim P_v(v)$, following Nevo (2000x). Throughout this work, we assume that $P_v(v) \equiv N(0, 1)$.
[15] Note that both in this specification and in plain logit framework, we assume $U_{i0mt} = \epsilon_{i0mt}$

$\mu_{ijmt}(p_{jmt}, v_i; \sigma_\alpha)$ and ϵ_{ijmt} stand for the *individual deviations* from the specified mean utility (Berry et al. (1995)).[16]

The consumers make their purchasing decisions based on the overall utility they get from all existing products in a given market and time period. Assuming they can purchase at most 1 unit of the existing products (as in Berry et al. (1995)), the choice indicator of a product for each market and its corresponding time period is given by (following Conlon et al. (2020)):

$$Choice_{ijmt} = \begin{cases} 1 & if \quad U_{ijmt} > U_{i\mathfrak{z}mt} \quad \forall \mathfrak{z} \neq j, \quad \forall \{\mathfrak{z}, j\} \in J_{mt} \\ 0 & otherwise \end{cases} \tag{6}$$

We can infer from the expression (6) to get aggregate market shares, as pointed out by Berry et al. (1995) and Conlon et al. (2020), across markets and corresponding time periods as follows:

$$s_{jmt} = \int Choice_{ijmt}(\delta_{.mt}, \mu_{i.mt}) \cdot dP_v(v) \cdot d\epsilon_{i.mt} \tag{7}$$

As in the plain logit framework, with independent and identically distributed ϵ_{ijmt} random variables belonging to the Type 1 extreme value distribution, we get (by following Nevo (2000*x*)):

$$s_{jmt}(\delta_{.mt}; \beta, \alpha, \sigma_\alpha) = \int \frac{exp(x_{jmt} \cdot \beta - \alpha \cdot p_{jmt} + \xi_{jmt} - p_{jmt} \cdot \sigma_\alpha \cdot v_{i\alpha})}{1 + \sum_{\mathfrak{z} \in J_{mt}} exp(x_{\mathfrak{z}mt} \cdot \beta - \alpha \cdot p_{\mathfrak{z}mt} + \xi_{\mathfrak{z}mt} - p_{\mathfrak{z}mt} \cdot \sigma_\alpha \cdot v_{i\alpha})} \cdot dP_v(v) \tag{8}$$

The equation (8) is equivalent to the following expression:

$$s_{jmt}(\delta_{.mt}; \beta, \alpha, \sigma_\alpha) = \int \frac{exp(\delta_{jmt} + \mu_{ijmt})}{1 + \sum_{\mathfrak{z} \in J_{mt}} exp(\delta_{\mathfrak{z}mt} + \mu_{i\mathfrak{z}mt})} \cdot dP_v(v) \tag{9}$$

As opposed to the plain logit specification, the market-level shares equation in mixed logit framework doesn't have a closed-form expression. This issue mainly stems from the fact that the individual choice probabilities are now evaluated for consumers with different characteristics.[17][18]To be able to arrive to the numerical values of these aforementioned market-level shares in mixed logit specification, several different techniques

[16]From this point, we will use the simplified expressions for mean utilities and individual-specific deviations from these mean utilities, i.e. $\delta_{jmt}(x_{jmt}, p_{jmt}, \xi_{jmt}; \beta, \alpha) = \delta_{jmt}$ and $\mu_{ijmt}(p_{jmt}, v_i; \sigma_\alpha) = \mu_{ijmt}$.

[17]As pointed out in Nevo (2000*x*), $s_{ijmt} = \frac{exp(\delta_{jmt} + \mu_{ijmt})}{1 + \sum_{\mathfrak{z}} exp(\delta_{\mathfrak{z}mt} + \mu_{i\mathfrak{z}mt})}$ corresponds to the choice probability of individual i for product j in market m at time period t.

[18]The dissimilarity among consumers is mainly embedded into the model with the term v_i.

might be employed. Conceivably, the most popular technique is to employ some pseudo Monte Carlo draws and evaluate the expression (9) for the different values of v_i.[19] This procedure will be discussed in section **3**.

Before proceeding to the supply specification, there are mainly three things worth keeping in mind: First, we don't actually execute the aforementioned counterfactual exercises for the plain logit case. Nevertheless, we also incorporate this framework. The main reason behind this choice is to progress the general structure in a gradual way and to ease the general follow-up. Furthermore, since we assume that the conjectural researcher also estimates the demand model for both *OLS Logit* and *2SLS Logit* cases, we decided to add the demand side plain logit framework into the general model specification section.

Second, note that we don't have any demographics in our mixed logit specification: Apart from the idiosyncratic demand-side shock, the only component that is heterogeneous across consumers is unobserved consumer characteristics, i.e. v_is.

Last but not least, both model specifications are static in nature: As one can see, we incorporate both markets and time periods, but neither the consumers nor the firms actually calculate the present discounted value of their utilities and profits, respectively. In that sense, the nature of the proposed setting is completely one-shot.

2.2 Supply Side for Mixed Logit Specification

Since we will employ the aforementioned counterfactual analyses just with the mixed logit model, we only specify the supply side of that framework in this subsection, albeit the general structure can also be easily adjusted for the plain logit specification.

We assume that in each market for every time period, there are $F = J_{mt}$ oligopolistic firms where each firm produces a single product j. These firms get the following profits:

$$\pi_{jmt} = (p_{jmt} - mc_{jmt}) \cdot s_{jmt}(\delta_{.mt}; \beta, \alpha, \sigma_\alpha) \tag{10}$$

In equation (10), mc_{jmt} stands for the marginal cost of producing brand j in market m at time period t. Particularly, we assume that the general structure of the marginal costs is given by:

$$mc_{jmt} = w_{jmt} \cdot \gamma + \omega_{jmt} \tag{11}$$

Regarding the established marginal cost specification given by expression (11), the term w_{jmt} corresponds to the observed cost shifters of brand j in market m at time period

[19] As it has been mentioned in Conlon et al. (2020).

t, whereas ω_{jmt} stands for the additional unobserved cost component of brand j in market m at time period t. Besides, the γ coefficients are the linear cost parameters. As one can notice in expression (11), we specify that there is a linear relationship between the observed and unobserved cost determinants and the marginal costs.

We assume that for every market and in each time period, a *unique* and *pure* Nash equilibrium exists under Bertrand price competition, so that the aforementioned single-product oligopolistic firms set their prices according to Bertrand-Nash pricing equilibrium:[20]

$$\max_{p_{jmt}} \quad (p_{jmt} - mc_{jmt}) \cdot s_{jmt}(\delta_{.mt}; \beta, \alpha, \sigma_\alpha) \tag{12}$$

Simplifying $s_{jmt}(\delta_{.mt}; \beta, \alpha, \sigma_\alpha) = s_{jmt}$, we get:

$$\frac{\partial \pi_{jmt}}{\partial p_{jmt}} = 0 \Rightarrow (p_{jmt} - mc_{jmt}) \frac{\partial s_{jmt}}{\partial p_{jmt}} + s_{jmt} = 0 \tag{13}$$

Since we have single-product firms, then with reference to the equation (13), we attain:

$$p_{jmt} - mc_{jmt} = \frac{s_{jmt}}{-\frac{\partial s_{jmt}}{\partial p_{jmt}}} \tag{14}$$

To be able to proceed with equation (14), we need the expression of the market-level shares derivative term with respect to prices, i.e. $\frac{\partial s_{jmt}}{\partial p_{jmt}}$. From the findings of Nevo $(2000x)$, we know that the general structure regarding the price elasticities of market-level shares in mixed logit framework is given by:

$$\eta_{j\mathfrak{z}mt} = \frac{\partial s_{jmt} \cdot p_{\mathfrak{z}mt}}{\partial p_{\mathfrak{z}mt} \cdot s_{jmt}} = \begin{cases} -\frac{p_{jmt}}{s_{jmt}} \int \alpha_i \cdot s_{ijmt} \cdot (1 - s_{ijmt}) \cdot dP_v(v) & if \quad j = \mathfrak{z} \\ \frac{p_{\mathfrak{z}mt}}{s_{jmt}} \int \alpha_i \cdot s_{ijmt} \cdot s_{i\mathfrak{z}mt} \cdot dP_v(v) & if \quad j \neq \mathfrak{z} \end{cases} \tag{15}$$

From expression (15), we can get the above-mentioned derivative term as follows:

$$\frac{\partial s_{jmt}}{\partial p_{jmt}} = -\int \alpha_i \cdot s_{ijmt} \cdot (1 - s_{ijmt}) \cdot dP_v(v) \tag{16}$$

[20]Unfortunately, we cannot guarantee whether such equilibrium exists in our setting: Since some of the market-level shares are greater than 0.5, our procedure fails to fulfill both the existence and uniqueness sufficiency conditions assuring such equilibrium, as indicated in the work of Aksoy-Pierson et al. (2013)[2]. Because of this issue, we can only assume that this type of equilibrium exists in our setup. Note that we attain the numerical values of those shares by following the procedures given in subsection **3.4**.

Inserting expression (16) into equation (14), we acquire:

$$p_{jmt} - mc_{jmt} = \frac{s_{jmt}}{\int \alpha_i \cdot s_{ijmt} \cdot (1 - s_{ijmt}) \cdot dP_v(v)} \tag{17}$$

Note that equation (17) characterizes a system of equations giving equilibrium prices, concerning each market and time period. Specifically, we have J_{mt} equations and J_{mt} prices for each market and time period combination, which allow us to solve for the equilibrium prices in our setting. To be able to numerically calculate these systems of equations, we again need to draw v_is and evaluate the equation (17) for those draws. The procedure is analogous to the one we mentioned in subsection **2.1.2**. This method will also be discussed in section **3**.

3 METHODOLOGY & DATA GENERATION PROCESS

In this part of the thesis, we discuss the fixation of the parameter values and other factors we described in section **2**: We briefly talk about how we simulate the data, and how we manipulate the simulated data to arrive at our desired setting before starting to analyze the aforementioned counterfactual exercises within the context of the introduction of the existing product. We will also talk about the estimation strategy of the conjectural researcher and the methods s/he employs to choose the unobserved characteristics components for the counterfactual analyses.

3.1 Choice of Parameter Values and Other Primitives

In subsections **2.1.1** and **2.1.2**, we intrinsically defined the model primitives for both plain logit and mixed logit specifications. Since we generate the data (instead of externally collecting it from an industry), we should fix the parameter values and we should also decide some other factors, such as the number of markets, number of time periods, and number of firms (i.e. the number of brands, since we have single-product firms in our setting) to be able to generate the market-level equilibrium prices and resulting market-level "observed" shares.

For our simulation exercises, we choose:

- $M = 20$ markets,
- $T = 10$ time periods,
- $F = 5$ single-product firms, therefore $J_{mt} = 5 \quad \forall \{m, t\}$,
- $I_{mt} = 100$ consumers $\forall \{m, t\}$,[21]

[21] Note that these consumers will be used for the estimation. For calculating the equilibrium prices, we draw 50 different individuals (whose notation is *ns*), which is independent of the consumer draws for the estimation part.

- $\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \\ 0.5 \\ 0.5 \end{pmatrix}$,
- $\alpha = 2$,
- $\sigma_\alpha = 0.5$,
- $\gamma = \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 0.5 \\ 0.5 \\ 0.5 \end{pmatrix}$

3.2 Generating the Data

Now, we describe how we generate the observed product attributes (x_{jmt}), observed cost shifters (w_{jmt}), unobserved costs (ω_{jmt}), and unobserved product characteristics (ξ_{jmt}), to be able to solve for equilibrium prices and to convey estimation. However, there is an important thing that we must emphasize here, before generating these factors:

Ideally, one should simulate many datasets and get the average results over these samples.[22][23]Unfortunately, due to time constraints, we only generate one dataset and gather the corresponding estimates and the results of the counterfactual exercises based only on this dataset.

We draw the observed product attributes (x_{jmt}), observed cost shifters (w_{jmt}), and unobserved costs (ω_{jmt}) from standard normal distribution. In that sense, we gather:

- $\{x_{1,1,1,1}, x_{1,2,1,1}, \ldots, x_{1,j,m,t}, \ldots, x_{1,5,20,10}\} \quad \sim N(0,1)$,
- $\{x_{2,1,1,1}, x_{2,2,1,1}, \ldots, x_{2,j,m,t}, \ldots, x_{2,5,20,10}\} \quad \sim N(0,1)$,
- $\{x_{3,1,1,1}, x_{3,2,1,1}, \ldots, x_{3,j,m,t}, \ldots, x_{3,5,20,10}\} \quad \sim N(0,1)$,
- $\{w_{1,1,1,1}, w_{1,2,1,1}, \ldots, w_{1,j,m,t}, \ldots, w_{1,5,20,10}\} \sim N(0,1)$,
- $\{w_{2,1,1,1}, w_{2,2,1,1}, \ldots, w_{2,j,m,t}, \ldots, w_{2,5,20,10}\} \sim N(0,1)$,
- $\{w_{3,1,1,1}, w_{3,2,1,1}, \ldots, w_{3,j,m,t}, \ldots, w_{3,5,20,10}\} \sim N(0,1)$,
- $\{\omega_{1,1,1}, \omega_{2,1,1}, \ldots, \omega_{j,m,t}, \ldots, \omega_{5,20,10}\} \qquad \sim N(0,1)$.

[22]One needs to get the average results for the equilibrium prices, market shares, estimates, and counterfactual exercises.

[23]This is particularly important to evaluate the general performance of the estimation strategy and the chosen counterfactual approaches.

Recall that we fit 3 observed product attributes and 3 observed cost shifters into our model. So that the variables x_{kjmt} and w_{kjmt} correspond to the coefficients β_k and γ_k respectively, for $k \in \{1,2,3\}$. This means that the factors matching with the coefficients β_0 and γ_0 are the constant terms.

In subsection 3.1, we chose the number of markets, time periods, and brands that are present in each market-time period combination. Since we fit 20 markets and 10 time periods into our model, and since we indicate that we also fit 5 brands into each market-time period combination, the **baseline** norm of these aforementioned factors equals to $J \times M \times T = 5 \times 20 \times 10 = 1000$. That is to say, the vectorized dimensions of these factors are (1000×1). However, we also assume that the researcher actually doesn't observe the first product in first market at first time period. The lack of this brand will actually form the base of our counterfactual exercises. In subsection 3.3, we will describe how we manipulate this generated data to ensure this situation.

3.2.1 Generating the Unobserved Product Characteristics

In this subsection, we explain how we generate the unobserved product characteristics. The method we employ is that, we draw J_{mt}, M, and T numbers from some distinct normal distributions -which differ in terms of their variances- consecutively for goods, markets, and time periods. To be more specific, let:[24][25]

- $\{x_1, \ldots, x_j, \ldots, x_5\} \sim N(0, \sigma_x^2) \Rightarrow$ *brand-specific components*
- $\{m_1, m_2, \ldots, m_m, \ldots, m_{20}\} \sim N(0, \sigma_m^2) \Rightarrow$ *market-fixed effects*
- $\{t_1, t_2, \ldots t_t, \ldots, t_{10}\} \sim N(0, \sigma_t^2) \Rightarrow$ *time-fixed effects*
- $\{\varepsilon_{1,1,1}, \varepsilon_{2,1,1}, \ldots, \varepsilon_{jmt}, \ldots, \varepsilon_{5,20,10}\} \sim N(0, \sigma_\varepsilon^2) \Rightarrow$ *random noise*

Now, we construct the unobserved product characteristics as the sum of the above-mentioned random draws in accordance with their indexes:

$$\xi_{jmt} = x_j + m_m + t_t + \varepsilon_{jmt} \tag{18}$$

$\Rightarrow$ e.g. $\xi_{3,15,4} = x_3 + m_{15} + t_4 + \varepsilon_{3,15,4}$; similarly for other ξs for each combination of j & m & t.

Note that this approach introduces correlation across markets and time periods in terms of these unobserved product characteristics. The above specification for unobserved product characteristics is particularly important for our counterfactual analyses

[24] Recall that in subsection 3.1, we fixed $J_{mt} = 5 \ \ \forall\{m,t\}$, $M = 20$, and $T = 10$.

[25] For the specification regarding the components of the unobserved product characteristics, we pick $\sigma_x = 2.4$, $\sigma_m = 1.6$, $\sigma_t = 1.2$ and $\sigma_\varepsilon = 0.5$ in our setting.

because we will need to infer the value of the unobserved product characteristics of the ”missing” good from other markets and time periods.[26] With the help of this correlation structure, the unobserved product characteristics will be dependent on which brand it actually belongs to, and in which market and time period this brand exists.

3.3 Manipulating the Data

Up to this point, we generated the observed product attributes (x_{jmt}), unobserved product characteristics (ξ_{jmt}), observed cost shifters (w_{jmt}), and unobserved costs (ω_{jmt}) for the baseline specification. However, as we pointed out briefly at the end of subsection **3.2**, we drop the first good in first market at first time period, and we assume that the conjectural researcher actually doesn't observe this specified product in that specific market-time combination.[27] This manipulation method brings two advantages:

1. With dropping the first brand in the first market at the first time period, we essentially form a solid ground for our intended counterfactual analyses in the context of the introduction of the existing product: This ”dropped” product exists across other markets and time periods. Intuitively, this means that the researcher will have access to information regarding the attributes of this product to some extent. Most importantly, he will have several means of inferring the unobserved product characteristics of this product from the unobserved product characteristics of the same product existing across other markets and time periods,[28]

2. The researcher won't be able to gather any direct information regarding this ”missing” brand. Nevertheless, we have access to these values. This status allows us to implement the true values of the unobserved product characteristics in the counterfactual analyses and to compare the general performance of the other ξ specifications relating to the ”real” value of this ”missing” unobserved product characteristics.

Since we dropped the aforementioned product and already have all the required data, we are now ready to solve for equilibrium prices and get the resulting market-level shares of all other existing products. Note that compared to section **3.2**, the vectorized dimensions of the product attributes are now (999×1).

3.4 Solving for the Equilibrium Prices

In this subsection, we generate the equilibrium prices and corresponding market-

[26] We briefly talked about this ”missing” good in the last paragraph of section **3.2**, but the actual reasoning resides in subsection **3.3**.

[27] Consequently, that means that we drop x_{111}, ξ_{111}, w_{111}, and ω_{111} from the baseline data generation. Additionally, it also implies that $J_{11} = 4$ while $J_{mt} = 5 \;\; \forall \{m, t\}$ except $m = 1$ and $t = 1$.

[28] Which the researcher will get via the joint BLP estimation.

level shares of all existing brands across markets and time periods for the mixed logit framework, following the previously discussed and implemented data manipulation.

Recall that in subsection **2.2**, we got the following expression:

$$p_{jmt} - mc_{jmt} = \frac{s_{jmt}}{\int \alpha_i \cdot s_{ijmt} \cdot (1 - s_{ijmt}) \cdot dP_v(v)} \tag{19}$$

With some tedious calculations whose details are given in appendices part (subsection **6.2**), we attain the final version of the above equilibrium pricing expression, which is given by:

$$p_{jmt}-w_{jmt}\cdot\gamma-\omega_{jmt}=\frac{\frac{1}{I}\cdot\sum_i^I \frac{exp(x_{jmt}\cdot\beta-\alpha\cdot p_{jmt}+\xi_{jmt}-p_{jmt}\cdot\sigma_\alpha\cdot v_{i\alpha})}{1+\sum_{\mathfrak{z}\in J_{mt}} exp(x_{\mathfrak{z}mt}\cdot\beta-\alpha\cdot p_{\mathfrak{z}mt}+\xi_{\mathfrak{z}mt}-p_{\mathfrak{z}mt}\cdot\sigma_\alpha\cdot v_{i\alpha})}}{\frac{1}{I}\cdot\sum_i^I(\alpha+\sigma_\alpha\cdot v_{i\alpha})\cdot\frac{exp(x_{jmt}\cdot\beta-\alpha\cdot p_{jmt}+\xi_{jmt}-p_{jmt}\cdot\sigma_\alpha\cdot v_{i\alpha})}{1+\sum_{\mathfrak{z}\in J_{mt}} exp(x_{\mathfrak{z}mt}\cdot\beta-\alpha\cdot p_{\mathfrak{z}mt}+\xi_{\mathfrak{z}mt}-p_{\mathfrak{z}mt}\cdot\sigma_\alpha\cdot v_{i\alpha})}\cdot(1-\frac{exp(x_{jmt}\cdot\beta-\alpha\cdot p_{jmt}+\xi_{jmt}-p_{jmt}\cdot\sigma_\alpha\cdot v_{i\alpha})}{1+\sum_{\mathfrak{z}\in J_{mt}} exp(x_{\mathfrak{z}mt}\cdot\beta-\alpha\cdot p_{\mathfrak{z}mt}+\xi_{\mathfrak{z}mt}-p_{\mathfrak{z}mt}\cdot\sigma_\alpha\cdot v_{i\alpha})})} \tag{20}$$

As we indicated in subsection **2.2**, concerning every market and time period, the equation (20) characterizes a system of nonlinear equations with respect to the price variable, and it is just the thoroughly processed version of the equation (17) in our setting: The nominator on the right-hand side is the approximation for market-level shares (s_{jmt}) in mixed logit specification, whereas the denominator represents the approximation for the term $\int \alpha_i \cdot s_{ijmt} \cdot (1 - s_{ijmt}) \cdot dP_v(v)$, again in mixed logit framework. Analogous to our statements in subsection **2.2**, we have J_{mt} equations and J_{mt} prices for each market-time period combination, which allow us to solve for the equilibrium prices.

Furthermore, recall that we already fixed the parameters β, α, σ_α, and γ; and we already generated x_{jmt}, w_{jmt}, ξ_{jmt}, and ω_{jmt}. Thus, the only unknown regarding the aforementioned system of equations is the price variable, i.e. p_{jmt}, $\forall\{j, m, t\}$.

For solving the above system of equations giving us the equilibrium prices for each market and time period, we use the built-in function *fsolve* in MATLAB, which serves to solve the systems of nonlinear equations. The function uses an algorithm called *Powell's dog leg method*, which minimizes the nonlinear least-squares of the given specification.[29][30]To solve for the equilibrium prices, we use $ns = 50$ individuals, which are completely different from the unobserved consumer characteristics draws, i.e. v_is. The latter will be used by the researcher for estimation and counterfactual experiments. In that sense, one might think of replacing the term I with ns in above specification, but we decided to use the notation I for general notational consistency throughout this thesis.

[29]For further notice on this method, please check Powell (1970)[18]

[30]By using different starting values for the prices in the mentioned algorithm, we arrive at the same numerical values for the equilibrium prices. To be more specific, we used the (999×1) tuples of 1s, 10s, 50s, 100s, 500s, and 1000s as the initial values for the prices. In this respect, we might claim that our procedure provides unique equilibrium prices, given other required factors to generate them.

When we get these equilibrium prices, we can simply plug them in our market-level shares approximation expression to get market-level "observed" shares in mixed logit framework, which are given by:[31]

$$s_{jmt} = \int \frac{exp(\delta_{jmt} + \mu_{ijmt})}{1 + \sum_{\mathfrak{z} \in J_{mt}} exp(\delta_{\mathfrak{z}mt} + \mu_{i\mathfrak{z}mt})} dP_v(v) \approx \frac{1}{I} \sum_{i}^{I} \frac{exp(\delta_{jmt} + \mu_{ijmt})}{1 + \sum_{\mathfrak{z} \in J_{mt}} exp(\delta_{\mathfrak{z}mt} + \mu_{i\mathfrak{z}mt})} = S_{jmt} \tag{21}$$

3.5 Estimation Strategy

Hitherto, we generated x_{jmt}, w_{jmt}, ξ_{jmt}, ω_{jmt}, p_{jmt}, and S_{jmt}, for all existing brands, markets, and time periods. Now we are ready to approach the situation from the point of view of the aforementioned researcher, who will proceed with the estimation and counterfactual analyses.

However, note that the researcher doesn't observe the unobserved product characteristics ξ_{jmt} and unobserved costs ω_{jmt}.[32]We assume that s/he observes:

- Market shares (S_{jmt}),
- Prices (p_{jmt})
- Observed product attributes (x_{jmt}),
- Observed cost shifters (w_{jmt})

Also recall that we assume this hypothetical researcher observes these values for the manipulated data, i.e. he observes $m = 1$ and $t = 1$ without the presence of brand 1.

The researcher will estimate the model with four different approaches. These approaches are:

- OLS Logit: The researcher estimates the model for the plain logit case, without using any instruments for prices,
- 2SLS Logit: The researcher estimates the model for the plain logit case with instruments: In first-stage, s/he regresses prices on instruments $Z^{Logit} = [constant \;\; x_{jmt} \;\; w_{jmt}]$, get the first-stage estimated prices; and s/he uses these estimates in the second-stage,

[31]Note that we again use the same $ns = 50$ individuals to gather the "observed" market-level shares.
[32]Nevertheless, we have access to the "true" values of these factors, as mentioned before.

- BLP Demand: The researcher estimates the model according to the findings of Berry et al. (1995). In this approach, s/he only estimates the demand side (i.e. s/he estimates the parameters β, α, and σ_α),
- BLP Joint: Again, the researcher estimates the model in accordance with the findings of Berry et al. (1995), but this time s/he also incorporates the estimation of the supply side. That is to say, the researcher estimates the joint BLP specification.

The technical details for all of these estimation approaches are given in section 6.3. One thing to note that, within the estimation of *BLP Demand* and *BLP Joint* specifications, the hypothetical researcher uses the instruments $Z^{BLP} = [constant \;\; x_{jmt} \;\; w_{jmt} \;\; x^2_{jmt} \;\; w^2_{jmt}]$.

3.6 Details on the Counterfactual Analyses

In the previous subsection, we provided the details for the estimation procedure. We assume that the aforementioned hypothetical researcher conveys all these estimations, but s/he performs the counterfactual analyses based on the estimates of **only joint BLP estimation**. Scilicet, we assume that s/he wants to collect the estimates of the other three approaches, but wants to convey the counterfactual analyses based on the estimates of joint mixed logit specification. These counterfactual analyses consist of getting the new pricing equilibrium, new profits, and new total consumer surplus when the ”missing” brand in the first market and first time period combination is introduced into this specific market-time period duo.

Before embarking on describing the general structure of the counterfactual analyses, there are some points worth keeping in mind: First, we assume that the hypothetical researcher actually knows the observed product characteristics and observed cost shifters for the ”missing” product.[33]We can interpret this situation as follows: The researcher wants to know what happens in the first market and first time period combination when the first brand, which was missing from this market-time period combination, enters into it with some *specific* observed characteristics and observed cost shifters, which are the true numerical values of x_{111} and w_{111}. Second, one can notice that to be able to introduce this ”missing” product into $m = 1$ & $t = 1$, the imagined researcher needs to somehow infer the value of the unobserved product characteristics of this brand. The point is, the value of ξ_{111} could be informed from several different means. As we mentioned before, this idea constitutes the main premise of this thesis.

3.6.1 Different Means for Getting The Unobserved Product Characteristics of the Missing Brand

[33]Recall that in our modified specification, the first brand in the first good at the first time period is missing. That was the point of doing the data manipulation in subsection 3.3.

Now, we explain by which means the researcher could infer the unobserved product characteristics of the missing product. The researcher has several potential ways to proceed, with some examples being:

- $\underline{\xi^1}$: This notation represents the counterfactual scenario in which the unobserved characteristics of the missing good is inferred from the same good present in other markets and time periods except $m = 1$ & $t = 1$ (i.e. there are 199 distinct values to infer the unobserved product characteristics:[34]The researcher calculates the profit of the first firm, profits of the other (i.e. rival) firms and the total consumer surplus separately for all those 199 values and then, s/he takes their averages). In that sense, we can say that the researcher gets the conditional expectation of both profits and total consumer surplus based on these different ξ picks. Note that the researcher gets those ξ values from the joint BLP estimation.[35]

- $\underline{\xi^2}$: This notation represents the counterfactual scenario in which the unobserved characteristics of the missing good is inferred from the same brand present in other markets, but **only** from time period 1 (i.e. there are 19 distinct values to infer the unobserved product characteristics: Analogous to the previous method, the researcher calculates the profit of the first firm, profits of the other (i.e. rival) firms and the total consumer surplus separately for all these 19 values and then, s/he takes their averages). Note that the researcher again gets those ξ values from the joint BLP estimation.

- $\underline{\mu_{\xi^1}}$: This notation represents the counterfactual scenario in which the unobserved product characteristics of the missing good is inferred from the **mean** of ξ^1: Instead of calculating the profits and total consumer surplus separately and taking their averages, the researcher calculates those at once by inserting the mean of ξ^1.

- $\underline{\mu_{\xi^2}}$: This notation represents the counterfactual scenario in which the unobserved product characteristics of the missing good is inferred from the **mean** of ξ^2: Instead of calculating the profits and total consumer surplus separately and taking their averages, the researcher calculates those at once by inserting the mean of ξ^2.

Additionally, the researcher can make use of the correlation structure of the ξs across markets and time periods, and s/he can perform a regression estimation to infer the aforementioned unobserved products characteristics of the missing good:

- $\underline{\widehat{\xi}_{Naive}}$: This notation represents the counterfactual scenario in which the researcher uses the ξ values s/he obtained from joint BLP estimation and s/he regresses them

[34]Since we have $M = 20$ markets and $T = 10$ time periods.

[35]The researcher gets those values as the demand-side error term.

on brand-specific dummies, market dummies, and time dummies to get the estimates of the mean and standard deviation of brand-specific components, market-fixed effects and time-fixed effects.[36] After s/he gets those estimates, s/he "naively" uses the estimate of the brand-dummy of the first product in lieu of the unobserved product characteristics of the missing good. Viewed in this way, the regression specification providing the predicted unobserved product characteristics is given as:

$$\widehat{\xi_{jmt}} = \widehat{x_j} + \widehat{m_m} + \widehat{t_t}^{37} \tag{22}$$

- $\underline{\widehat{\xi}_{Draws}}$: This notation represents the counterfactual scenario in which the researcher uses the ξ values s/he obtained from the joint BLP estimation and s/he regresses them on brand-dummies, market-dummies, and time-dummies to get the estimates of the mean and standard deviation of brand-specific components, market-fixed effects, and time-fixed effects. After s/he gets those estimates, s/he uses the estimate of the brand-dummy of the first product and the standard deviation of its residual to generate 100 draws and uses them separately to calculate new profits and total consumer surplus, and then takes their averages.

3.6.2 Solving for the Counterfactual Equilibrium Prices

Above, we explained the methods that the researcher will employ when s/he needs to infer the unobserved product characteristics of the missing product, as well as how s/he solves for the new profits and total consumer surplus based on these ξ choices. Here, we present how the researcher solves for the counterfactual equilibrium prices.

Since we are interested in the specific market-time period combination in which the first brand is missing, we focus on $m = 1$ & $t = 1$ combination. By following the same logic that was presented in subsection 3.4, yet shifting our focus on $m = 1$ & $t = 1$, we modify the equation (20) as follows:

$$p_{j11} - w_{j11}\cdot\gamma - \omega_{j11} = \frac{\frac{1}{I}\cdot\sum_i^I \frac{exp(x_{j11}\cdot\beta - \alpha\cdot p_{j11} + \xi_{j11} - p_{j11}\cdot\sigma_\alpha\cdot v_{i\alpha})}{1+\sum_{\mathfrak{z}\in J_{11}} exp(x_{\mathfrak{z}11}\cdot\beta - \alpha\cdot p_{\mathfrak{z}11} + \xi_{\mathfrak{z}11} - p_{\mathfrak{z}11}\cdot\sigma_\alpha\cdot v_{i\alpha})}}{\frac{1}{I}\cdot\sum_i^I (\alpha+\sigma_\alpha\cdot v_{i\alpha})\cdot\frac{exp(x_{j11}\cdot\beta - \alpha\cdot p_{j11} + \xi_{j11} - p_{j11}\cdot\sigma_\alpha\cdot v_{i\alpha})}{1+\sum_{\mathfrak{z}\in J_{11}} exp(x_{\mathfrak{z}11}\cdot\beta - \alpha\cdot p_{\mathfrak{z}11} + \xi_{\mathfrak{z}11} - p_{\mathfrak{z}11}\cdot\sigma_\alpha\cdot v_{i\alpha})}\cdot(1 - \frac{exp(x_{j11}\cdot\beta - \alpha\cdot p_{j11} + \xi_{j11} - p_{j11}\cdot\sigma_\alpha\cdot v_{i\alpha})}{1+\sum_{\mathfrak{z}\in J_{11}} exp(x_{\mathfrak{z}11}\cdot\beta - \alpha\cdot p_{\mathfrak{z}11} + \xi_{\mathfrak{z}11} - p_{\mathfrak{z}11}\cdot\sigma_\alpha\cdot v_{i\alpha})})}^{38} \tag{23}$$

[36] Note that for this method and also for $\widehat{\xi}_{Draws}$, we use J product dummies, M-1 market dummies, and T-1 time dummies, excluding the first market and first time dummies, also excluding the constant term from this regression specification. This means the researcher infers the unobserved product characteristics of the missing product from the estimate of the first brand dummy alone.

[37] The regression is performed for $j \in \{1,2,3,4,5\}$; $m \in \{2,3,4,\ldots,19,20\}$; and $t \in \{2,3,4,\ldots,9,10\}$ dummies. The estimation method for this case is inferred from the equation (18). Also, note that the same regression is also employed for the $\widehat{\xi}_{Draws}$ method.

[38] Notice that this time the counterfactual prices are generated by using I, instead of using ns whose value was equal to 50 and which was used for generating the prices in subsection 3.4. The reasoning is, we use ns for the initial generation of prices, and I for the estimation and counterfactual exercises.

Notice that instead of having to solve J_{mt} nonlinear system of equations for each market and time period combination, we only need to solve J_{11} equations for our counterfactual exercises. In other words, we only need to solve for the counterfactual prices concerning $m = 1$ & $t = 1$ combo.[39] As mentioned in subsection **3.6.1**, the researcher uses the specified methods to infer the value of ξ_{111}: After collecting these values, the researcher simply puts them in equation (23) distinctively and get the new counterfactual equilibrium prices for each possible inference of ξ_{111}.

3.6.3 Calculating the New Consumer Surplus

In our setting, ex-ante total consumer surplus in $m = 1$ & $t = 1$ is given by:

$$CS^{Baseline} = \sum_{i}^{I} \frac{1}{\widehat{\alpha}_i} \cdot ln\Big(\sum_{j=1}^{4} exp(x_{j11} \cdot \widehat{\beta} - \widehat{\alpha} \cdot p_{j11} + \xi_{j11} - p_{j11} \cdot \widehat{\sigma_\alpha} \cdot v_{i\alpha})\Big) + C^{40} \tag{24}$$

Now it is time to explain how the researcher gathers the ex-post consumer surplus, i.e. when the missing good is introduced. Without loss of generality, suppose that the researcher wants to get the total consumer surplus when s/he prefers the ξ^1 method we described in subsection 3.6.1. To ensure this, s/he uses the following expression:

$$\begin{aligned} CS^{\xi^1} = \sum_{i}^{I} \frac{1}{\widehat{\alpha}_i} \cdot ln\Big(& exp(x_{111} \cdot \widehat{\beta} - \widehat{\alpha} \cdot p_{111} + \xi^1_{111} - p_{j11} \cdot \widehat{\sigma_\alpha} \cdot v_{i\alpha}) + \\ & \sum_{j=2}^{4} exp(x_{j11} \cdot \widehat{\beta} - \widehat{\alpha} \cdot p_{j11} + \xi_{j11} - p_{j11} \cdot \widehat{\sigma_\alpha} \cdot v_{i\alpha})\Big) + C \end{aligned} \tag{25}$$

Note that in equation (25), the term ξ^1_{111} corresponds to the values representing the unobserved product characteristics of the previously missing good, all distinctively evaluated from the method ξ^1. The same approach will also be applied for all the other methods that we specified in subsection 3.6.1.

4 RESULTS

In this part, we present the readers the results of the estimation and of the counterfactual exercises that the hypothetical researcher obtains, following the model specification that we described in section **2** and the general methodology that we introduced in section **3**.

[39] However, note that now the value J_{11} equals to 5 instead of 4, since we introduce the previously missing product $j = 1$ here.

[40] Where the estimates $\widehat{\beta}, \widehat{\alpha}$ and $\widehat{\sigma_\alpha}$ are assumed to be obtained from the joint BLP estimation.

4.1 Estimation Results

Recall that we discussed the estimation strategy and its details in subsection 3.5, with the technical elaborations given in subsection 6.3. The general estimation results for all four predefined specifications are given in Table 1, along with the true values of the parameters.[41]

Looking at the results, we see that the estimate on the constant term and the estimate of the price sensitivity parameter in *OLS Logit* estimation are far away from their true values, whereas the other estimates seem quite plausible for this procedure.

Shifting our focus to *2SLS Logit* estimation, we see some improvements on both the estimates on the constant term and for the price sensitivity parameter. Nevertheless, the amelioration on these estimates doesn't seem tremendous: Both the estimates for β_0 and α are still quite far out from their true values.

Now, we begin to analyze the results for both BLP estimations. Starting from the estimates we get from only evaluating the demand side, we see that the estimates are pretty much close to their true values: Not only the estimates for β_0 and α substantially improved compared to both plain logit estimations, but we also see the estimation is able to capture the disturbance parameter on the price quite well.

Lastly, we check the estimates for joint BLP estimation. We see that the estimated values of β_0 and α get deteriorated a bit compared to the previous method, but the norm of this deterioration seems to be negligible. On the other hand, the estimate for the deviation term for price sensitivity improved slightly, compared to its demand-only counterpart. We also see that the supply side estimates are pretty near to their true values. Now, the researcher is ready to perform the prementioned counterfactual exercises with the parameter estimates of the joint BLP estimation at hand.[42]

4.2 Counterfactual Analyses Results

The results of the counterfactual simulation exercises are given in Table 2, whose columns consist of:

- ”Firm of Interest”: This column represents the profits firm 1 gets in each counterfactual scenario at $m = 1$ & $t = 1$.[43]Since firm 1 was nonexistent at $m = 1$ & $t = 1$, the baseline profits of this firm equals to 0.
- ”Rival Firms”: This column represents the **sum** of the profits of the other firms (i.e. other brands) at $m = 1$ & $t = 1$.

[41] Recall that in subsection 3.1, we fixed the true parameter values.

[42] Also, note that all of these estimates (excluding the estimate of β_2 in *BLP Demand* estimation) are statistically significant at 1% level.

[43] Recall that firm 1 produces brand 1, as a single-product firm.

Table 1: ESTIMATION RESULTS

	True Values	OLS Logit	2SLS Logit	BLP Demand	BLP Joint
β_0	5.0000	1.9148	2.5126	4.9588	4.7409
		(0.2153)***	(0.4016)***	(0.0140)***	(0.0133)***
β_1	3.0000	2.5798	2.5926	2.8070	2.7805
		(0.0516)***	(0.0643)***	(0.0247)***	(0.0227)***
β_2	0.5000	0.4371	0.4424	0.4896	0.4839
		(0.0477)***	(0.0592)***	(0.1416)**	(0.1303)***
β_3	0.5000	0.5772	0.5754	0.5605	0.5621
		(0.0503)***	(0.0624)***	(0.1236)***	(0.1122)***
$-\alpha$	-2.0000	-1.1598	-1.2604	-1.8517	-1.7917
		(0.0353)***	(0.0667)***	(0.0374)***	(0.0352)***
σ_α	0.5000			0.5811	0.5331
				(0.1193)***	(0.1183)***
γ_0	5.0000				5.0284
					(0.0014)***
γ_1	0.5000				0.5480
					(0.0129)***
γ_2	0.5000				0.4981
					(0.0142)***
γ_3	0.5000				0.5674
					(0.0125)***

*** $\iff p < 0.01$; ** $\iff p < 0.05$

- "Total Consumer Surplus": This column represents the **total** consumer surplus in baseline and in each counterfactual scenario at $m = 1$ & $t = 1$.[44]

In subsection **3.6.1**, recollect that we defined some methods to proximate the unobserved product characteristics of the missing brand, which were $\xi^1, \mu_{\xi^1}, \xi^2, \mu_{\xi^2}, \widehat{\xi}_{Naive}$, and $\widehat{\xi}_{Draws}$. These different methods constitute the rows of Table 2, with the addition of two extra factors:

- Baseline: This represents the ex-ante structure, i.e. when brand 1 was missing from $m = 1$ & $t = 1$ combination.
- ξ_{REAL}: This situation represents the counterfactual in which the researcher is assumed to use the **true** values of unobserved characteristics **only for brand 1**, i.e. for the newly introduced existing product. The unobserved product characteristics of the other (rival) goods are left the same (i.e. for the rival firms, the researcher still uses the ξ estimates s/he gets from the joint BLP estimation). The performance of the aforementioned methods for inferring the value of ξ_{111} will be decided based on the closeness of the results to the values that the true ξ_{111} concludes.

Table 2: COUNTERFACTUAL ANALYSES RESULTS

	Firm of Interest	Rival Firms	Total Consumer Surplus
Baseline	0	0.2211	0
ξ^1	0.1094	0.1765	31.0129
ξ^2	0.0707	0.1867	23.4244
μ_{ξ^1}	0.0790	0.1804	27.4912
μ_{ξ^2}	0.0546	0.1906	20.8248
$\widehat{\xi}_{Naive}$	0.0715	0.1834	25.5261
$\widehat{\xi}_{Draws}$	0.0714	0.1835	25.4845
ξ_{REAL}	0.1370	0.1610	40.6047

Profits and Consumer Welfare for Different Unobserved Product Attributes

[44] Note that we normalized the consumer surplus of the baseline specification to 0 and wrote each counterfactual scenarios' total consumer surplus relative to the baseline specification. The actual value of the baseline total consumer surplus equals to -83.0855, meaning that the constant term C appearing in equation (24) is negative.

By looking at Table 2, we see that the entrant firm cannibalizes from all other four incumbents for all approaches, after the missing product is introduced into $m = 1$ & $t = 1$ market-time period combination, which is reasonable to expect. Additionally, the total consumer surplus of that market-time period composition goes up for each method. Since the consumers have a wider variety of choices compared to before, this situation also makes sense. Now, it is time to evaluate the performances of the different methods being employed for the pick of ξ_{111} compared to the true case:

- **In terms of profits:** In Table 2, we see that the closest profit values with respect to the profits resulting with ξ_{REAL} approach are given by the ξ^1 method, both for the profits of the entrant firm and incumbents. The worst performance is due to μ_{ξ^2} approach. This makes sense since the profit and welfare calculations are nonlinear in nature, and because of that, performing these counterfactual exercises with a single data point in terms of ξ is counter-intuitive. However, we also see that μ_{ξ^1} approach -which is analogous in nature compared to μ_{ξ^2} approach- results in closer counterfactual profits to the true case, compared to the regression approaches and ξ^2 approach. This is actually engrossing, but it might be the case that choosing the potential unobserved product characteristics of the brand 1 only for $t = 1$ doesn't work well in our setting and one should infer these values not only across markets, but across time periods as well. Last, the resulting counterfactual profits for both regression approaches are also not appealing. This last issue might be due to the fact that we don't observe a fair amount of markets and time periods to be able to get closed-to-true estimates for the regression dummies.

- **In terms of total consumer surplus:** We see the exact pattern in terms of sensitivity that we see for the profits: ξ^1 approach gives us the closest total consumer surplus compared to the true case, whereas the results of the μ_{ξ^2} approach lie farthest from the true counterfactual total consumer surplus value, followed by ξ^2 approach. The performances of both regression specification methods are -again- very close to each other. In terms of analyzing these results, we can apply the same thought process that we applied in terms of the values and sensitivity of the counterfactual profits.

All in all, we see that in our general specification, inferring the unobserved product attributes of the missing product from the same good across markets and time periods gives us the closest counterfactual analyses results compared to the true case, which signals that the conjectural researcher should follow the ξ^1 approach for the counterfactual exercises in this setting. Nevertheless, we should again emphasize that the results could be unreliable due to having only one drawn sample.

The important thing to notice is that all of these different approaches actually result in different numerical final values regarding the counterfactual prices and total consumer surplus, which indeed proves the main point of this thesis.

5 CONCLUSIONS

Since its introduction to the empirical industrial organization literature, the BLP framework (Berry et al. (1995)) became one of the most popular tools in executing several policy simulations, mainly in the context of the new mergers and the introduction of the new products into the industry of interest. One collective issue regarding these policy simulations is that the empirical industrial organization researchers usually make simplistic assumptions concerning the nature of the unobserved product characteristics of the brands of interest, while performing these counterfactual analyses. These assumptions, in return, could have non-negligible consequences on the counterfactual prices, profits, and consumer welfare of these aforementioned policy analyses. In this thesis, we intend to investigate this issue in the context of the introduction of the existing product, which can be considered as a simple and intuitive baseline regarding the conjecture of this work. In other words, we aim to perform a sensitivity check on the counterfactual prices, profits, and consumer welfare, in respect of the different choices concerning the unobserved product characteristics of the brand of interest.

To be able to implement this idea, we first fix the numerical values of the required model primitives for our described model. By following these true values of our model primitives, we synthetically construct a dataset that acts upon the data of some industry and which provides the numerical values on the observed product attributes (x_{jmt}s), observed cost shifters (w_{jmt}s), market-level equilibrium prices (p_{jmt}s) and market-level shares (S_{jmt}s). Then, we assume a conjectural researcher, having access to these data, estimates the preconstructed model with various estimation strategies. By using the estimates of the joint BLP framework, s/he performs the prementioned simulation exercises in the context of the pre-existing good introduction into a specific market-time period combination.

Our results indicate that the means for inferring the unobserved product characteristics for the brand of interest actually matters: We see that in our simulation experiments, different choices on the nature of these unobserved attributes lead to different outcomes. Therefore, the results of our counterfactual analyses point out the importance of the selection of these factors in the simulation exercises, in the context of empirical industrial organization framework: Instead of making simplified assumptions, the applied researchers in this field should be on the alert while inferring the nature of the unobserved product attributes in their policy experiments.

However, the findings of this thesis bear some caveats as well: As we pointed out in subsection 3.2, the way we approach the simulation exercises is not ideal: Optimally, one should simulate many datasets -instead of just one-, and get the averages for the estimates and the counterfactual prices, profits, and consumer welfare, to be able to execute a reliable performance check on the general process. Besides, the total number of markets and time periods could also be increased to get better variation across these segments, in the interest of ameliorating the findings of this study.[45] From these points of view, possible future research on this particular topic would encompass performing the analogous analyses with many datasets and getting the averages of the outcomes, and augmenting the total number of the markets and time periods.

References

[1] Ackerberg, D., Benkard, C. L., Berry, S., & Pakes, A. (2007). Econometric tools for analyzing market outcomes. *Handbook of econometrics*, *6*, 4171-4276.

[2] Aksoy-Pierson, M., Allon, G., & Federgruen, A. (2013). Price competition under mixed multinomial logit demand functions. *Management Science*, *59*(8), 1817-1835.

[3] Amemiya, T. (1977). The maximum likelihood and the nonlinear three-stage least squares estimator in the general nonlinear simultaneous equation model. *Econometrica: Journal of the Econometric Society*, 955-968.

[4] Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 242-262.

[5] Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, 841-890.

[6] Berry, S., & Pakes, A. (2007). The pure characteristics demand model. *International Economic Review*, *48*(4), 1193-1225.

[7] Chamberlain, G. (1987). Asymptotic efficiency in estimation with conditional moment restrictions. *Journal of econometrics*, *34*(3), 305-334.

[8] Conlon, C., & Gortmaker, J. (2020). Best practices for differentiated products demand estimation with pyblp. *The RAND Journal of Economics*, *51*(4), 1108-1161.

[9] Dubé, J. P., Fox, J. T., & Su, C. L. (2012). Improving the numerical performance of static and dynamic aggregate discrete choice random coefficients demand estimation. *Econometrica*, *80*(5), 2231-2267.

[10] Igami, M. (2017). Estimating the innovator's dilemma: Structural analysis of creative destruction in the hard disk drive industry, 1981–1998. *Journal of Political Economy*, *125*(3), 798-847.

[11] Knittel, C. R., & Metaxoglou, K. (2014). Estimation of random-coefficient demand models: two empiricists' perspective. *Review of Economics and Statistics*, *96*(1), 34-59.

[12] McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior.

[45] The reason we stick to $M = 20$ and $T = 10$ specification is purely due to convergence issues in our MATLAB code.

[13] McFadden, D. (1974). The measurement of urban travel demand. *Journal of public economics*, *3*(4), 303-328.

[14] Nevo, A. (2000x). A practitioner's guide to estimation of random-coefficients logit models of demand. *Journal of economics & management strategy*, *9*(4), 513-548.

[15] Nevo, A. (2000y). Mergers with differentiated products: The case of the ready-to-eat cereal industry. *The RAND Journal of Economics*, 395-421.

[16] Olesiński, B. (2020). The Analysis of the Tobacco Product Bans Using a Random Coefficients Logit Model. *Central European Journal of Economic Modelling and Econometrics*, 113-144.

[17] Petrin, A. (2002). Quantifying the benefits of new products: The case of the minivan. *Journal of political Economy*, *110*(4), 705-729.

[18] Powell, M. J. (1970). A new algorithm for unconstrained optimization. In *Nonlinear programming* (pp. 31-65). Academic Press.

6 APPENDICES

Herein, we provide an itemized list regarding the model primitives, along with a comprehensive review concerning the technical details on solving the equilibrium prices and on estimation.

6.1 Primitives of The Model

The model portrayed in this work contains the following elements:

- $m \in \{1, 2, \ldots, M\}$ markets,
- $t \in \{1, 2, \ldots, T\}$ time periods,
- $f \in \{1, 2, \ldots, F\}$ *single-product* firms, existing across markets and time periods,
- $i \in \{1, 2, \ldots, I_{11}; 1, 2, \ldots, I_{12}; \ldots, 1, 2, \ldots, I_{1T}; 1, 2, \ldots, I_{21}; \ldots; 1, 2, \ldots, I_{mt}; \ldots; 1, 2, \ldots, I_{MT}\}$ consumers within each market and time period,
- $j \in \{1, 2, \ldots, J_{11}; 1, 2, \ldots, J_{12}; \ldots, 1, 2, \ldots, J_{1T}; 1, 2, \ldots; J_{21}; \ldots; 1, 2, \ldots, J_{mt}; \ldots; 1, 2, \ldots, J_{MT}\}$ products within each market and time period,
- S_{jmt}: "observed" market-level shares,
- s_{jmt}: "model" market-level shares,
- x_{jmt}: observed product attributes,
- ξ_{jmt}: unobserved product characteristics,
- p_{jmt}: observed prices,
- ϵ_{ijmt}: idiosyncratic demand-side taste shocks,
- w_{jmt}: observed cost shifters,
- ω_{jmt}: unobserved costs,
- $mc_{jmt} = w_{jmt} \cdot \gamma + \omega_{jmt}$: marginal costs,
- v_i: unobserved consumer characteristics,
- θ_1: linear demand parameters $\iff \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$,
- α: common price sensitivity parameter,

- σ_α: deviation term regarding prices,
- $\theta_2 = \{\alpha, \sigma_\alpha\}$: nonlinear parameters,
- θ_3: linear supply side cost parameters $\iff \gamma = \begin{pmatrix}\gamma_0\\ \gamma_1\\ \gamma_2\\ \gamma_3\end{pmatrix}$,
- $\theta^{Logit} = \{\theta_1, \alpha\}$: Structural parameters for the plain logit specification,
- $\theta^{RCLogit} = \{\theta_1, \theta_2, \theta_3\}$: Structural parameters for the mixed logit specification.

6.2 Solving for the Equilibrium Prices

Recall that we arrived to the following equilibrium pricing equation for the mixed logit model, at the end of subsection 2.2:

$$p_{jmt} - mc_{jmt} = \frac{s_{jmt}}{\int \alpha_i \cdot s_{ijmt} \cdot (1 - s_{ijmt}) \cdot dP_v(v)} \tag{26}$$

By inferring $\int \alpha_i \cdot s_{ijmt} \cdot (1 - s_{ijmt}) \cdot dP_v(v) \approxeq \frac{1}{I} \cdot \sum_i^I \alpha_i \cdot s_{ijmt} \cdot (1 - s_{ijmt})$, we attain:

$$p_{jmt} - mc_{jmt} = \frac{s_{jmt}}{\frac{1}{I} \cdot \sum_i^I \alpha_i \cdot s_{ijmt} \cdot (1 - s_{ijmt})} \tag{27}$$

Since we know from Nevo (2000x) that individual choice probabilities take the form: $s_{ijmt} = \frac{exp(\delta_{jmt}+\mu_{ijmt})}{1+\sum_{\mathfrak{z}} exp(\delta_{\mathfrak{z}mt}+\mu_{i\mathfrak{z}mt})}$, we insert them into equation (27) and get:

$$p_{jmt} - mc_{jmt} = \frac{s_{jmt}}{\frac{1}{I} \cdot \sum_i^I \alpha_i \cdot \frac{exp(\delta_{jmt}+\mu_{ijmt})}{1+\sum_{\mathfrak{z}\in J_{mt}} exp(\delta_{\mathfrak{z}mt}+\mu_{i\mathfrak{z}mt})} \cdot (1 - \frac{exp(\delta_{jmt}+\mu_{ijmt})}{1+\sum_{\mathfrak{z}\in J_{mt}} exp(\delta_{\mathfrak{z}mt}+\mu_{i\mathfrak{z}mt})})} \tag{28}$$

By also inferring $s_{jmt} = \int \frac{exp(\delta_{jmt}+\mu_{ijmt})}{1+\sum_{\mathfrak{z}\in J_{mt}} exp(\delta_{\mathfrak{z}mt}+\mu_{i\mathfrak{z}mt})} dP_v(v) \approxeq \frac{1}{I} \sum_i^I \frac{exp(\delta_{jmt}+\mu_{ijmt})}{1+\sum_{\mathfrak{z}\in J_{mt}} exp(\delta_{\mathfrak{z}mt}+\mu_{i\mathfrak{z}mt})}$ for the mixed logit framework, we get:

$$p_{jmt} - mc_{jmt} = \frac{\frac{1}{I} \cdot \sum_i^I \frac{exp(\delta_{jmt}+\mu_{ijmt})}{1+\sum_{\mathfrak{z}\in J_{mt}} exp(\delta_{\mathfrak{z}mt}+\mu_{i\mathfrak{z}mt})}}{\frac{1}{I} \cdot \sum_i^I \alpha_i \cdot \frac{exp(\delta_{jmt}+\mu_{ijmt})}{1+\sum_{\mathfrak{z}\in J_{mt}} exp(\delta_{\mathfrak{z}mt}+\mu_{i\mathfrak{z}mt})} \cdot (1 - \frac{exp(\delta_{jmt}+\mu_{ijmt})}{1+\sum_{\mathfrak{z}\in J_{mt}} exp(\delta_{\mathfrak{z}mt}+\mu_{i\mathfrak{z}mt})})} \tag{29}$$

Writing the terms δ_{jmt} and μ_{ijmt} explicitly, we obtain:

$$p_{jmt}-mc_{jmt}=\frac{\frac{1}{I}\cdot\sum_i^I \frac{exp(x_{jmt}\cdot\beta-\alpha\cdot p_{jmt}+\xi_{jmt}-p_{jmt}\cdot\sigma_\alpha\cdot v_{i\alpha})}{1+\sum_{\mathfrak{z}\in J_{mt}} exp(x_{\mathfrak{z}mt}\cdot\beta-\alpha\cdot p_{\mathfrak{z}mt}+\xi_{\mathfrak{z}mt}-p_{\mathfrak{z}mt}\cdot\sigma_\alpha\cdot v_{i\alpha})}}{\frac{1}{I}\cdot\sum_i^I \alpha_i\cdot\frac{exp(x_{jmt}\cdot\beta-\alpha\cdot p_{jmt}+\xi_{jmt}-p_{jmt}\cdot\sigma_\alpha\cdot v_{i\alpha})}{1+\sum_{\mathfrak{z}\in J_{mt}} exp(x_{\mathfrak{z}mt}\cdot\beta-\alpha\cdot p_{\mathfrak{z}mt}+\xi_{\mathfrak{z}mt}-p_{\mathfrak{z}mt}\cdot\sigma_\alpha\cdot v_{i\alpha})}\cdot(1-\frac{exp(x_{jmt}\cdot\beta-\alpha\cdot p_{jmt}+\xi_{jmt}-p_{jmt}\cdot\sigma_\alpha\cdot v_{i\alpha})}{1+\sum_{\mathfrak{z}\in J_{mt}} exp(x_{\mathfrak{z}mt}\cdot\beta-\alpha\cdot p_{\mathfrak{z}mt}+\xi_{\mathfrak{z}mt}-p_{\mathfrak{z}mt}\cdot\sigma_\alpha\cdot v_{i\alpha})})} \tag{30}$$

Implying $\alpha_i = \alpha + \sigma_\alpha \cdot v_{i\alpha}$, we acquire:

$$p_{jmt}-mc_{jmt}=\frac{\frac{1}{I}\cdot\sum_i^I \frac{exp(x_{jmt}\cdot\beta-\alpha\cdot p_{jmt}+\xi_{jmt}-p_{jmt}\cdot\sigma_\alpha\cdot v_{i\alpha})}{1+\sum_{\mathfrak{z}\in J_{mt}} exp(x_{\mathfrak{z}mt}\cdot\beta-\alpha\cdot p_{\mathfrak{z}mt}+\xi_{\mathfrak{z}mt}-p_{\mathfrak{z}mt}\cdot\sigma_\alpha\cdot v_{i\alpha})}}{\frac{1}{I}\cdot\sum_i^I (\alpha+\sigma_\alpha\cdot v_{i\alpha})\cdot\frac{exp(x_{jmt}\cdot\beta-\alpha\cdot p_{jmt}+\xi_{jmt}-p_{jmt}\cdot\sigma_\alpha\cdot v_{i\alpha})}{1+\sum_{\mathfrak{z}\in J_{mt}} exp(x_{\mathfrak{z}mt}\cdot\beta-\alpha\cdot p_{\mathfrak{z}mt}+\xi_{\mathfrak{z}mt}-p_{\mathfrak{z}mt}\cdot\sigma_\alpha\cdot v_{i\alpha})}\cdot(1-\frac{exp(x_{jmt}\cdot\beta-\alpha\cdot p_{jmt}+\xi_{jmt}-p_{jmt}\cdot\sigma_\alpha\cdot v_{i\alpha})}{1+\sum_{\mathfrak{z}\in J_{mt}} exp(x_{\mathfrak{z}mt}\cdot\beta-\alpha\cdot p_{\mathfrak{z}mt}+\xi_{\mathfrak{z}mt}-p_{\mathfrak{z}mt}\cdot\sigma_\alpha\cdot v_{i\alpha})})} \tag{31}$$

Lastly, writing the marginal cost expression explicitly, i.e. $mc_{jmt} = w_{jmt} \cdot \gamma + \omega_{jmt}$, we finally get:

$$p_{jmt}-w_{jmt}\cdot\gamma-\omega_{jmt}=\frac{\frac{1}{I}\cdot\sum_i^I \frac{exp(x_{jmt}\cdot\beta-\alpha\cdot p_{jmt}+\xi_{jmt}-p_{jmt}\cdot\sigma_\alpha\cdot v_{i\alpha})}{1+\sum_{\mathfrak{z}\in J_{mt}} exp(x_{\mathfrak{z}mt}\cdot\beta-\alpha\cdot p_{\mathfrak{z}mt}+\xi_{\mathfrak{z}mt}-p_{\mathfrak{z}mt}\cdot\sigma_\alpha\cdot v_{i\alpha})}}{\frac{1}{I}\cdot\sum_i^I (\alpha+\sigma_\alpha\cdot v_{i\alpha})\cdot\frac{exp(x_{jmt}\cdot\beta-\alpha\cdot p_{jmt}+\xi_{jmt}-p_{jmt}\cdot\sigma_\alpha\cdot v_{i\alpha})}{1+\sum_{\mathfrak{z}\in J_{mt}} exp(x_{\mathfrak{z}mt}\cdot\beta-\alpha\cdot p_{\mathfrak{z}mt}+\xi_{\mathfrak{z}mt}-p_{\mathfrak{z}mt}\cdot\sigma_\alpha\cdot v_{i\alpha})}\cdot(1-\frac{exp(x_{jmt}\cdot\beta-\alpha\cdot p_{jmt}+\xi_{jmt}-p_{jmt}\cdot\sigma_\alpha\cdot v_{i\alpha})}{1+\sum_{\mathfrak{z}\in J_{mt}} exp(x_{\mathfrak{z}mt}\cdot\beta-\alpha\cdot p_{\mathfrak{z}mt}+\xi_{\mathfrak{z}mt}-p_{\mathfrak{z}mt}\cdot\sigma_\alpha\cdot v_{i\alpha})})} \tag{32}$$

This general procedure gives us the required systems of equations in order to solve for the equilibrium prices.

6.3 Technical Details on the Estimation Procedures

Here, we talk about how the aforementioned hypothetical researcher performs the estimation for all four specifications:

- For OLS Logit: By directly using the inversion technique invented by Berry (1994), the researcher gets:

$$ln(S_{jmt}) - ln(S_{0mt}) = x_{jmt} \cdot \beta - \alpha \cdot p_{jmt} + \xi_{jmt} \tag{33}$$

  where S_{0mt} indicates the market-level share of the outside good for market m and time period t.

- For 2SLS Logit: The researcher employs a technique which is analogous to the former one. Yet, this time s/he instruments the prices and plug the first-stage estimates of them into equation (33) to convey the second-stage estimation.

- For BLP Demand: Mainly following the guidance of Nevo (2000x), the researcher performs this estimation with 3 stages:

– *Stage 1*: Since in mixed logit specification the estimated market-level shares don't have an analytical formula, the researcher constructs an approximation of them as follows:

$$s_{jmt}(\delta_{.mt};\sigma_\alpha) \approxeq \frac{1}{I}\sum_{i}^{I}\frac{exp(\delta_{jmt}+\mu_{ijmt})}{1+\sum_{\mathfrak{z}\in J_{mt}} exp(\delta_{\mathfrak{z}mt}+\mu_{i\mathfrak{z}mt})} \tag{34}$$

– *Stage 2*: After constructing these shares, s/he employs the contraction mapping proposed by Berry et al. (1995):

$$\delta_{jmt}^{h+1} = \delta_{jmt}^{h} + ln(S_{jmt}) - ln(s_{jmt}(\delta_{.mt};\sigma_\alpha)) \tag{35}$$

When this process is terminated, the researcher gets the error term, which is simply the estimates for the unobserved product characteristics:[46]

$$\xi_{jmt}(\beta,\alpha,\sigma_\alpha) = \delta_{jmt}(S_{jmt};\sigma_\alpha) - x_{jmt}\cdot\beta + \alpha\cdot p_{jmt} \tag{36}$$

– *Stage 3*: After collecting these estimated ξ values, the researcher minimizes the sample analogue of population moments to get the estimates for β, α, and σ_α:

$$\{\widehat{\beta},\widehat{\alpha},\widehat{\sigma_\alpha}\} = argmin_{\beta,\alpha,\sigma_\alpha}(\xi'(\beta,\alpha,\sigma_\alpha)\cdot Z^{BLP}\cdot\Theta^{-1}\cdot Z'^{BLP}\cdot\xi(\beta,\alpha,\sigma_\alpha)) \tag{37}$$

where Θ^{-1} is the weighting matrix needed since $dim(Z^{BLP}) > dim([\beta,\alpha,\sigma_\alpha])$ in our setting, as pointed out by Nevo (2000x).[47]

Note that the above algorithm -which is called the *Nested Fixed Point algorithm*- is formed by an outer loop (in which the GMM is tried to being minimized by the values of the structural parameters of the model) and also by an inner loop (in which the contraction mapping takes place) and it was initially proposed by Berry et al. (1995).

- For BLP Joint: For the demand side of the joint BLP estimation, the researcher employs the same method that s/he employed for *BLP Demand* estimation. However, since the parameters α and σ_α also enter into the supply side specification, the general form of the sample analogue of population moments changes accordingly:[48]

$$\{\widehat{\beta},\widehat{\alpha},\widehat{\sigma_\alpha},\widehat{\gamma}\} = argmin_{\beta,\alpha,\sigma_\alpha,\gamma}(\zeta'(\beta,\alpha,\sigma_\alpha,\gamma)\cdot \Upsilon^{-1}\cdot\zeta(\beta,\alpha,\sigma_\alpha,\gamma)) \tag{38}$$

where $\zeta(\beta,\alpha,\sigma_\alpha,\gamma) = \begin{pmatrix} Z'^{BLP}\cdot\xi(\beta,\alpha,\sigma_\alpha) \\ Z'^{BLP}\cdot\omega(\alpha,\sigma_\alpha,\gamma)\end{pmatrix}$ and $\Upsilon = \begin{pmatrix}\Theta & 0 \\ 0 & \Theta\end{pmatrix}$.

[46] The process is terminated when $|\delta_{jmt}^{H+1} - \delta_{jmt}^{H}| < \tau \ \forall h \geqslant H$, for some threshold value τ, following Nevo (2000x).

[47] Which is the consistent estimate of $E[Z'^{BLP}\cdot\xi\cdot\xi'\cdot Z^{BLP}]$, as pointed out by Nevo (2000x).

[48] For further explanations on the technical details of the BLP estimation, one can refer to Berry et al. (1995), Nevo (2000x), and Conlon et al. (2020).