Asymptotics for differentiated product demand/supply systems with many markets in the presence of national micro moments

Yuichiro KANAZAWA[*]

International Christian University

1 Introduction

Industrial organization literature is concerned with the structure of industries in the economy and the behavior of firms and individuals in these industries. A dramatic shift in the 1980s toward what Bresnahan (1989) coined the "New Empirical Industrial Organization (NEIO)," which tries to takes advantage of the fact that individual industries are sufficiently distinct, and industry details are sufficiently important.

In the NEIO, products are regarded as bundles of characteristics, and preferences are defined on those characteristics, so each consumer chooses a bundle that maximizes its utility. Consumers have different preferences for different characteristics, and hence make different choices, however. In other words, consumers are heterogeneous.

Simulation is used to obtain aggregate demand from the heterogeneous consumers' choices in the following manner: first we draw vectors of consumer characteristics from the distribution of those characteristics; second we determine the choice probability that each of the drawn consumers would make for a given value of the parameter; now we aggregate those choice probabilities into a prediction for aggregate demand conditional on the parameter vector; finally we employ a search routine that finds the value of that parameter which makes these aggregate probabilities as close as possible to the observed market shares.

The theoretical and econometric groundwork for product characteristic based demand systems dates back to the work of Lancaster (1971) [9] and McFadden (1974, 1981) [10],[11]. Applications of the Lancaster/McFadden framework, however, increased significantly after Berry, Levinsohn, and Pakes (1995, henceforth BLP(1995) [2]) showed how to circumvent a problem that had made it difficult to apply the early generation of characteristic-based models in I.O. contexts.

The early generation of models did not allow for unobserved product characteristics. Consumer goods typically are differentiated in many ways. As a result even if the econometricians measured all the relevant characteristics, they could not expect to obtain precise estimates of their impacts. One solution proposed in Berry (1994) [1] is to put in the "important" differentiating characteristics and an unobservable, say ξ, which picks up the aggregate effect of the multitude of characteristics that are being omitted.

Often the econometricians find that there is not enough information in product level demand data to estimate the entire distribution of preferences with sufficient precision. This should not be surprising given that they are trying to estimate a whole distribution of preferences from just aggregate choice probabilities. The literature has added information in two ways: One is to add an equilibrium assumption and work out its implications for the estimation of demand parameters, the other is to add more data.

It is not surprising that when the pricing system is added to the demand system, the precision of the demand parameters estimates tends to improve (see, BLP (1995) [2]). Almost all of it has assumed static or myopic profit maximization, and that one side of the transaction has the power to set prices while the other can only decide whether and what to buy conditional on those prices. However, models in marketing science started looking one period ahead (see, Che, et al. (2007) [6], Kamai and Kanazawa (2016) [8]) with interactions between manufacturers and a retailer factored in.

On adding more data, there are a number of types of micro data that might be available: Surveys that match individual characteristics to a product chosen by the individual (point-of-sales data); Surveys providing information on the proprietary consumer's second choice (Berry, Levinsohn, and Pakes (2004) [3]); Alternatively, the econometricians may have access to summary statistics that provide information on the joint distribution of consumer and product characteristics (Petrin 2002 [17]). Petrin (2002) proposes a technique for obtaining more precise estimates of demand and supply curves when the econometricians are constrained to market-level data. The technique allows them to augment market share data with information relating the average demographics of consumers to the observable characteristics–Myojo and Kanazawa (2012) called "discriminating attributes"–of products they purchase that determines a subset of products in the market. Petrin (2002) states that "[t]his extra information plays the same role as consumer-level data..."(p.705, [17]).

Berry, Linton, and Pakes (2004) [4] provides asymptotic theory of the estimate of the demand system objective function as the number of products increases in one (national) market. The paper shows that, provided one accounts for simulation and sampling error in the estimate of the objective function, standard approximations to the limit distribution work (see, for e.g. Pakes and Pollard, 1989 [16]). Myojo and Kanazawa (2012) [12] provides a sharper asymptotic variance- covariance of the estimate of the demand than Berry, Linton, and Pakes (2004) by adding the pricing equation and micro moment objective functions as the number of products increases in one (national) market.

For durable goods like automobiles, the number of models increased from somewhere in the 150's in 1980s to in the 260's in 2018. Asymptotic theories of Berry, Linton, and Pakes (2004) and Myojo and Kanazawa (2012) presuppose such a market. On the other hand, for many non-durable consumer good products, the number of product offering is

limited because of the limited shelf space at the retail outlets. In Nevo (2001), for example, there are 1124 markets and 24 products.[1] Therefore a different type of asymptotics, the one in the number of markets, is needed.

Freyberger (2015) [7] provides asymptotically normal *biased* parameter estimate (that depends on the number m of markets, the fixed number R of simulation draws from the distribution of consumers to calculate the observed market share of the demand system objective function) for a fixed number of products as the number of markets increases. It also provides how the leading asymptotic bias terms can be eliminated by using an *analytic bias correction* method.

Asymptotic bias is generated because, for each market, its participating households are self-selected and unique in its own way. As a result simply increasing the number of such markets will not do. One general estimation idea when you have estimates in many comparable but heterogeneous subgroups within a population is to combine the individual estimates, each unbiased, to manufacture an overall unbiased estimate, using the local variances and overall variance between subgroup means to select the best linear estimator. An approach for the current problem along this idea is to alter the number of simulation draws from the fixed R in Freyberger (2015) to market-dependent R_m to reflect the population. This approach should work in theory, but a proper choice of R_m presupposes we know the market-by-market variation of the consumer preferences, the exact knowledge we are trying to estimate. In reality, different portions of the population may still be over- or under-represented as we increase the number m of markets.

If we wish to have a data-driven method as an alternative to analytic (asymptotic) bias correction proposed by Freyberger (2015), however, there is another idea we can employ (for non-durable product markets with a limited number of product offering), however. That is, for these markets, information that encompasses many regions are available, and we can take advantage of such information to adjust the bias. In this manuscript, we show that we can pursue the second idea, namely, we can achieve the data-driven asymptotic bias correction by incorporating 1) the pricing (profit maximizing) equation for *national* suppliers and 2) the *national* micro moment regarding consumers as the number m of markets increases.

2   Method

We assume the same finite set of products, $j = 1, \ldots, J$, is available in each regional market.[2] We also assume each consumer only participates in one market and chooses one product including "outside" good that maximizes his/her utility within that market. The price and the unobserved product characteristic of the same product may as well vary from one market to the other because of the differences in demographics among markets.

---

[1] We study non-durable goods such as ready-to-eat cereals because serious policy implications abound for markets of such non-durable goods: For instance, Nevo(2001) [14] claims that "Previous researchers have concluded that the ready-to-eat cereal industry is a classic example of an industry with nearly collusive pricing behavior and intense non-price competition"(p.307).

[2] This simplifying assumption can be relaxed so that the set of products available can differ somewhat by region.

2.1 Data

We suppose the following information is available: 1) observed exogenous product characteristics $\boldsymbol{x}_j = (x_{2j}, \ldots, x_{Kj})^T$, $(K-1) \times 1$ vector, other than price and the market-by-market price p_j^m of products j, $j = 1, \ldots, J$; 2) observed market-by-market share vector s_j^{m,n_m}, of products j, $j = 1, \ldots, J$, $m = 1, \ldots, M$, in market m estimated from a sample of size n_m, independent with respect to m; 3) An estimate η_{kq}^N of k-th, $k = 1, \ldots, K$, observed demographic characteristic of consumers who choose q-th discriminating attribute. This estimate is constructed from a national sample of N consumers, independent of the sample of size n_j^m.[3]

We also suppose there is a national database (e.g. the aforementioned CE-PUMD, for instance) of consumer i, $i = 1, \ldots, I$ with their $K \times 1$ observed demographics $\boldsymbol{\nu}_i^m = (\nu_{i1}^m, \ldots, \nu_{iK}^m)^T \in \Re^K$ whose true distribution is P^0. This national database, we suppose, identifies the market each consumer participates by his/her region/market indicator m, $m = 1, \ldots, M$. We can draw T consumers to match the η_{kq}^N above. With this national database with region indicator, we can construct region-by-region databases indexed by m. We can randomly draw R_m demographics[4] from this regional database to match the s_j^{m,n_m} whose observed demographics are now indexed by m so that $K \times 1$ observed demographics $\boldsymbol{\nu}_i^m = (\nu_{i1}^m, \ldots, \nu_{iK}^m)^T \in \Re^K$ whose true distribution P^{0m}.

2.2 Demand side

Reflecting the heterogeneity of consumers, we define the model calculated probability σ_{ij}^m of consumer i choosing product j in market m to be obtained from the random coefficient utility of consumer i with observed demographics $\boldsymbol{\nu}_i^m$ in market m for product j as

$$U_{ij}^m = \boldsymbol{x}_j^{mT} \cdot \boldsymbol{\beta} + \sum_{k=1}^K \pi_k x_{jk} \nu_{ik}^m + \xi_j^m + \epsilon_{ij}^m, \tag{1}$$

with the observed $K \times 1$ product characteristics vector $\boldsymbol{x}_j^m = (p_j^m, \boldsymbol{x}_j^T)^T$ that may include a constant or product dummies, the demand-side $K \times 1$ parameter vector $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_K)$ associated with the observed product characteristics, $K \times 1$ parameter vector $\boldsymbol{\Pi} = (\pi_1, \ldots, \pi_K)$ associated with the observed demographics, unobservable ξ_j^m likely to be correlated with price p_j^m, and with unobservable idiosyncratic taste ϵ_{ij}^m assumed to be i.i.d. type-I extreme value.

The conditional probability σ_{ij}^m of choosing product j in market m obtained from (1)

$$\sigma_{ij}^m = \frac{\exp\left(\boldsymbol{x}_j^{mT} \cdot \boldsymbol{\beta} + \sum_{k=1}^K \pi_k x_{jk} \nu_{ik} + \xi_j^m\right)}{1 + \sum_{j=1}^J \exp\left(\boldsymbol{x}_j^{mT} \cdot \boldsymbol{\beta} + \sum_{k=1}^K \pi_k x_{jk} \nu_{ik} + \xi_j^m\right)} \tag{2}$$

---

[3] We do not believe that the independent-source requirement is so restrictive or unrealistic. We could sample individuals from the Consumer Expenditure Survey Public-Use Microdata (CE-PUMD) to simulate market shares of products, whereas the additional micro moment could be obtained from publicly available sources independent of the CE-PUMD because what we require is the aggregate estimates and not the sample itself—most likely to be proprietary—from which these aggregate estimates are computed.

[4] These two draws T and are R_m are assumed independent.

is a map from consumer i's tastes $\boldsymbol{\nu}_i^m$, a demand-side parameter $2K \times 1$ vector $\boldsymbol{\theta}_d = (\boldsymbol{\beta}^T, \boldsymbol{\Pi}^T)^T \in \Re^{2K}$, and the set of characteristics of all products $\boldsymbol{p}^m = (p_1^m, \ldots, p_J^m)^T$, a $J \times 1$ vector, $\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_J)$, a $(K-1) \times J$ matrix, $\boldsymbol{\xi}^m = (\xi_1^m, \ldots, \xi_J^m)^T$, a $J \times 1$ vector, and is thus denoted as $\sigma_{ij}^m(\boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{\xi}^m, \boldsymbol{\nu}_i^m, \boldsymbol{\theta}_d)$.

The BLP framework generates the vector of the model-calculated $J \times 1$ market share vector $\boldsymbol{\sigma}^m = (\sigma_1^m, \ldots, \sigma_J^m)^T$ in market m by aggregating over the individual choice probability in (2) with the true distribution P^{0m} of the consumer tastes ν_i^m as

$$\sigma_j^m(\boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{\xi}^m, \boldsymbol{\theta}_d, P^{0m}) = \int \frac{\exp\left(\boldsymbol{x}_j^{mT} \cdot \boldsymbol{\beta} + \sum_{k=1}^K \pi_k x_{jk} \nu_{ik} + \xi_j^m\right)}{1 + \sum_{j=1} \exp\left(\boldsymbol{x}_j^{mT} \cdot \boldsymbol{\beta} + \sum_{k=1}^K \pi_k x_{jk} \nu_{ik} + \xi_j^m\right)} dP^{0m}. \quad (3)$$

We assume the set of exogenous product characteristics $\boldsymbol{x}_j = (x_{2j}, \ldots, x_{Kj})^T$, $j = 1, \ldots, J$, is a random sample from the underlying population of product characteristics. We also assume that, given $\boldsymbol{x}_j$'s, ξ_j^m, $j = 1, \ldots, J$ are a random sample of unobserved characteristics from the portion of the population representing market m. Thus, $\boldsymbol{x}_j$ are assumed independent across j, and ξ_j^m are independent conditionally on $\boldsymbol{x}_j$, while $\boldsymbol{p}^m = (p_1^m, \ldots, p_J^m)^T$ are not, in general, independent across j since they are endogenously determined in market m as functions of others' and its own product characteristics.

Note that these market shares are still random variables due to the stochastic nature of the product characteristics $\boldsymbol{X}$ and $\boldsymbol{\xi}^m$. If we evaluate $\sigma_j^m(\boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{\xi}^m, \boldsymbol{\theta}_d, P^{0m})$ at $\boldsymbol{\theta}_d^0$, where $\boldsymbol{\theta}_d^0$ is the true parameter value, we have the "conditionally true" market shares s_j^{m0} given the product characteristics $(\boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{\xi}^m)$ in the portion of the population representing market m, i.e.,

$$\boldsymbol{\sigma}^m(\boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{\xi}^m, \boldsymbol{\theta}_d^0, P^{0m}) \equiv \boldsymbol{s}^{0m}. \quad (4)$$

BLP (1995) [2] provides general conditions under which there is a unique solution for the $\boldsymbol{\xi}^m$ for the equation

$$\boldsymbol{\sigma}^m(\boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{\xi}^m, \boldsymbol{\theta}_d, P^m) = \boldsymbol{s}^m \quad (5)$$

for every $(\boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{\theta}_d, P^m)$. If we solve (5), at any $(\boldsymbol{\theta}_d, \boldsymbol{s}^m, P^m) \neq (\boldsymbol{\theta}_d^0, \boldsymbol{s}^{0m}, P^{m0})$, the independence assumption for the resulting $\xi_j(\boldsymbol{X}, \boldsymbol{\theta}_d, \boldsymbol{s}^m, P^m)$ no longer holds. On the other hand, if we solve the identity in (4) with respect to $\boldsymbol{\xi}^m$ under the conditions to guarantee the uniqueness of the $\boldsymbol{\xi}^m$ in (5), we are able to retrieve the original $\xi_j^m(\boldsymbol{X}, \boldsymbol{\theta}_d^0, \boldsymbol{s}^0, P^{m0})$, which we assume are conditionally independent across j given $\boldsymbol{x} = (x_{2j}, \ldots, x_{Kj})^T$.

Since the econometrician do not have access to P^{m0} nor $\boldsymbol{s}^{0m}$, he/she must use the R_m random draws from P^{m0} and the observed share $\boldsymbol{s}^{m,n_m} = (s_1^{m,n_m}, \ldots, s_J^{m,n_m})$ which is usually estimated from a random sample of size n_m. The R_m random draws are assumed independent of the sample of size n_m above. The observed market share s_j^{m,n_m} of product j in market m is thus matched with the model-calculated market share σ_j^m of consumer i choosing product j in market m from m-th reginal database as

$$s_j^{m,n_m} = \frac{1}{R_m} \sum_{i=1}^{R_m} \sigma_{ij}^m(\boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{\xi}^m, \boldsymbol{\theta}_d, P^{R_m, m}), \quad (6)$$

where $P^{R_m,m}$ is the empirical measure that signifies a random sample of R_m is taken from the database indexed by the market index m. We denote the $\boldsymbol{\xi}$ as a solution to the J simultaneous equations (4) with J unknown ξ_j, $j = 1, \ldots, J$ as $\boldsymbol{\xi}(\boldsymbol{\theta}_d, \boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{s}^{m,n_m}, P^{R_m,m})$.

We assume the mean independent conditions to hold for each market $m = 1, \ldots, m$ as

$$E\left(\xi_j^m \middle| \boldsymbol{X}\right) = 0, \tag{7}$$

with probability 1 and to have a finite conditional variance

$$\sup_{1 \leq m \leq M} \max_{1 \leq j \leq J} E\left(\left(\xi_j^m\right)^2 \middle| \boldsymbol{x}_j\right) < \infty,$$

with probability 1.

We assume there is a $J \times M_1$ matrix demand-side instruments $\boldsymbol{Z}_d = (\boldsymbol{z}_1^d, \ldots, \boldsymbol{z}_J^d)^T$ whose j-th component $\boldsymbol{z}_j^d$ can be written further as $\boldsymbol{z}_j^d(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_J) \in \Re^{M_1}$, where $\boldsymbol{z}_j^d(\cdot)$: $\Re^{(K-1) \times J} \to \Re^{M_1}$, for $j = 1, \ldots, J$ such that its the resulting $M_1 \times 1$ vector

$$E\left(\boldsymbol{Z}_d^T \boldsymbol{\xi}^m(\boldsymbol{\theta}_d^0, \boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{s}^{0m}, P^{m0})\right) = \boldsymbol{0}, \tag{8}$$

in view of (7), where $\boldsymbol{0}$ is a $M_1 \times 1$ vector of zeroes, if we assume, for moment, that we know the underlying taste distribution of P^{m0} and we are able to observe the true market share $\boldsymbol{s}^{0m}$. The expectation in (8) is taken with respect $\boldsymbol{\xi}$ conditionally on $\boldsymbol{X}$, and then with respect to $\boldsymbol{X}$.

It should be noted that the demand side instruments $\boldsymbol{z}_j^d$ for product j are assumed to be a function of the exogenous characteristics not only of its own, but of the other products in the market. This is because the instruments by definition must correlate with the product characteristics p_j^m, and these endogenous price are determined by both its own and its competitors' product characteristics.

Then the sample counterpart of this condition employing the solution to (6) is the following $M_1 \times 1$ vector

$$G_M^d\left(\boldsymbol{\theta}_d, \{\boldsymbol{s}^m\}, \{P^{R_m}\}\right) = \frac{1}{M} \sum_{m=1}^{M} \boldsymbol{Z}_d^T \boldsymbol{\xi}_j^m(\boldsymbol{\theta}_d, \boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{s}^{m,n_m}, P^{R_m,m}). \tag{9}$$

2.3 Supply side

The supply side model formulates the pricing equations for the J products marketed. We assume an oligopolistic market where a finite number of suppliers provide multiple products to all M markets. Suppliers ($m = 1, \ldots, F$) producing $\mathcal{J}_f$ are maximizers of profit from the combination of products they offer.

A profit function of supplier f in market m is given for by

$$PR_f^m = \sum_{j=1}^{J} \left(p_j^m - mc_j^m\right) \sigma_j^m(\boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{\xi}^m, \boldsymbol{\theta}_d, P^m) \cdot I^m \tag{10}$$

where $j \in \mathcal{J}_f$ and mc_j^m is the marginal cost of product j in market m and I_m is the size of population in market m. Then, the total profit of the firm f is $\sum_{m=1}^{M} PR_f^m$.

Since consumer only participates in his/her home market, maximizing the profit function in each regional market m separately leads to maximizing the total profit for each firm.

Assuming BertrandNash pricing provides the first-order condition for the product j of the manufacturer m as

$$\sigma_j^m(\boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{\xi}^m, \boldsymbol{\theta}_d, P^m) + \sum_{h \in \mathcal{J}_f} (p_h^m - mc_h^m) \frac{\delta \sigma_h^m(\boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{\xi}^m, \boldsymbol{\theta}_d, P^m)}{\delta p_j^m}, \tag{11}$$

This equation can be expressed in matrix form as

$$\boldsymbol{\sigma}^m(\boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{\xi}^m, \boldsymbol{\theta}_d, P^m) + \boldsymbol{\Delta}^m (\boldsymbol{p}^m - \boldsymbol{mc}^m) = \boldsymbol{0}, \tag{12}$$

where $\boldsymbol{\Delta}^m$ is the $J \times J$ non-singular matrix whose (j, h) element is defined as

$$\Delta_{jh}^m = \begin{cases} \delta \sigma_h^m(\boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{\xi}^m, \boldsymbol{\theta}_d, P^m)/\delta p_j^m & \text{if the products } j \text{ and } h \text{ are produced by the same firm;} \\ 0, & \text{otherwise} \end{cases}$$

We solve (12) with respect to $\boldsymbol{mc}^m$ as

$$\boldsymbol{mc}^m = \boldsymbol{p}^m + (\boldsymbol{\Delta}^m)^{-1} \boldsymbol{\sigma}^m(\boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{\xi}^m, \boldsymbol{\theta}_d, P^m). \tag{13}$$

Suppose that the observed cost shifters $\boldsymbol{w}_j$ consist of exogenous $\boldsymbol{w}_{1j} \in \Re^{L_1}$ as well as endogenous $\boldsymbol{w}_{2j} \in \Re^{L_2}$, and thus we write $\boldsymbol{w}_j = (\boldsymbol{w}_{1j}^T, \boldsymbol{w}_{2j}^T)^T$, a $(L_1 + L_2) \times 1$ vector, and $\boldsymbol{W} = (\boldsymbol{w}_1, \dots, \boldsymbol{w}_J)^T$, a $J \times (L_1 + L_2)$ matrix. The exogenous cost shifters include not only the cost variables determined outside the market under consideration (e.g., factor price), but also the product design characteristics suppliers cannot immediately change in response to fluctuation in demand.

The cost variables determined by the market equilibrium (e.g., production scale) are treated as endogenous cost shifters. As in the formulation of $\boldsymbol{x}_j$ on the demand side, we assume that the set of exogenous cost shifters $\boldsymbol{w}_{1j}$, $j = 1, \dots, J$, is a random sample from the underlying population of cost shifters. We also assume as in the formulation of ξ_j^m, given $\boldsymbol{w}_{1j}$'s, ω_j^m, $j = 1, \dots, J$, is a random sample of the unobserved cost shifter from the portion of the population representing market m. Thus, $\boldsymbol{w}_j$ are assumed independent across j, and ω_j^m are independent conditionally on $\boldsymbol{w}_j$, although $\boldsymbol{w}_{2j}$ are, in general, not independent across j as they are determined in the market as functions of cost shifters of other products.

We define the marginal cost mc_j^m as an implicit function of the observed cost shifters $\boldsymbol{w}_j$ and the unobserved cost shifter ω_j^m in market m as

$$g(mc_j^m) = \boldsymbol{w}_j^T \boldsymbol{\theta}_c + \omega_j^m, \tag{14}$$

where $g()$ is a monotonic function and $\boldsymbol{\theta}_c \in \Theta_c$ is a vector of cost parameters. Although the choice of $g(\cdot)$ depends on the application, we assume that $g(\cdot)$ is continuously differentiable with a finite derivative for all realizable values of cost. Define $\boldsymbol{g}(\boldsymbol{x}) \equiv (g(x_1), \dots, g(x_J))^T$. Then the $J \times 1$ cost-side unobserved cost shifter vector obtains as

$$\boldsymbol{\omega}^m(\boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{\xi}^m, \boldsymbol{\theta}_d, P^m) = \boldsymbol{g}\left(\boldsymbol{p}^m + (\boldsymbol{\Delta}^m)^{-1} \boldsymbol{\sigma}^m(\boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{\xi}^m, \boldsymbol{\theta}_d, P^m)\right) - \boldsymbol{W}\boldsymbol{\theta}_c. \tag{15}$$

Similar to the demand-side unobservables, the unobserved cost shifter ω_j^m is assumed to be mean independent of the exogenous cost shifters $\boldsymbol{W}_1 = (\boldsymbol{w}_{11}, ..., \boldsymbol{w}_{1J})$, and satisfy

$$E\left(\omega_j^m \middle| \boldsymbol{W}_1\right) = 0, \tag{16}$$

with probability 1 and to have a finite conditional variance

$$\sup_{1 \leq m \leq M} \max_{1 \leq j \leq J} E\left(\left(\omega_j^m\right)^2 \middle| \boldsymbol{w}_{1j}\right) < \infty,$$

with probability 1.

Similar to the demand side, we assume there is a $J \times M_2$ matrix supply-side instruments $\boldsymbol{Z}_c = (\boldsymbol{z}_1^c, \ldots, \boldsymbol{z}_J^c)^T$ whose j-th component $\boldsymbol{z}_j^c$ can be written further as $\boldsymbol{z}_j^c(\boldsymbol{w}_{11}, \ldots, \boldsymbol{w}_{1J}) \in \Re^{M_2}$, where $\boldsymbol{z}_j^c(\cdot) : \Re^{(L_1) \times J} \to \Re^{M_2}$, for $j = 1, \ldots, J$ such that its the resulting $M_2 \times 1$ vector

$$E\left(\boldsymbol{Z}_c^T \boldsymbol{\omega}^m(\boldsymbol{\theta}_d^0, \boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{s}^{0m}, P^{m0})\right) = \boldsymbol{0}, \tag{17}$$

in view of (16), where $\boldsymbol{0}$ is a $M_2 \times 1$ vector of zeroes, if we assume, for moment, that we know the underlying taste distribution of P^{m0} and we are able to observe the true market share $\boldsymbol{s}^{0m}$. The expectation in (17) is taken with respect $\boldsymbol{\omega}$ conditionally on $\boldsymbol{W}$, and then with respect to $\boldsymbol{W}$.

We also note that some of the exogenous product characteristics $\boldsymbol{x}_j$ affect the price of the product because they affect manufacturing cost. Thus, those $\boldsymbol{x}_j$ may be included among the exogenous cost shifters $\boldsymbol{w}_{1j}$ if they are uncorrelated with the unobservable cost shifter ω_j^m for $j = 1, \ldots, J$.

Then the sample counterpart of this condition is the following $M_2 \times 1$ vector

$$G_M^c\left(\boldsymbol{\theta}, \{\boldsymbol{s}^m\}, \{P^{R_m}\}\right) = \frac{1}{M} \sum_{m=1}^M \boldsymbol{Z}_c^T \boldsymbol{\omega}_j^m(\boldsymbol{\theta}, \boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{s}^{m,n_m}, P^{R_m,m}), \tag{18}$$

where $\boldsymbol{\theta} = (\boldsymbol{\theta}_d, \boldsymbol{\theta}_c)$.

2.4   Micro moments

Market summaries such as the average demographics of consumers who purchased a specific type of product are publicly available, even if detailed individual-level data such as purchasing histories are not. A *discriminating attribute* is an observable product characteristic of products that determines a subset of products in the market, those products that possess the attribute. We denote the set of products with discriminating attribute q as Q_q, and the consumer's choice as C_i. We will write the event "consumer i chooses discriminating attribute q" as $C_i \in Q_q$. A consumer may choose multiple discriminating attributes. We assume that there is a finite number of discriminating attributes $q = 1, \ldots, N_p$, and the market share of each discriminating attribute is positive, i.e., $\Pr\left\{C_i \in Q_q \middle| \boldsymbol{X}, \boldsymbol{\xi}(\boldsymbol{\theta}_d^0, \boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{s}^{0m}, P^{m0})\right\} > 0$ for all q in $1, \ldots, N_p$.

We assume that the joint density of demographics $\boldsymbol{\nu}_i^m$ is of bounded support, $m = 1, \ldots, M$. In market m, consumer i's k-th observed demographic ν_{ik}^m, $k = 1, \ldots, K$ is averaged over all consumers choosing discriminating attribute q in the population to obtain the conditional expectation $\eta_{kq}^{0m} = E\left\{\nu_{ik}^m | C_i \in Q_q, \boldsymbol{X}, \boldsymbol{\xi}(\boldsymbol{\theta}_d^0, \boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{s}^{0m}, P^{m0})\right\}$. An example of this conditional expectation would be the expected value of family size of consumers in the subpopulation P^{m0} corresponding to the market m who purchased a box of ready-to-eat cereal with high sugar content. We assume that η_{kq}^{0m} has a finite mean and variance for all J, i.e., $E_{\boldsymbol{X},\boldsymbol{\xi}}\left\{\eta_{kq}^{0m}\right\} < \infty$ and $V_{\boldsymbol{X},\boldsymbol{\xi}}\left\{\eta_{kq}^{0m}\right\} < \infty$ for $k = 1, \ldots, K, q = 1, \ldots, N_p, m = 1, \ldots, M$.

It may be possible to match the the sample counterpart of η_{kq}^{0m} market-by-market with the sample counterpart of $E\left\{\nu_{ik}^m | C_i \in Q_q, \boldsymbol{X}, \boldsymbol{\xi}(\boldsymbol{\theta}_d^0, \boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{s}^{0m}, P^{m0})\right\}$ whenever a finer data set of consumers' market-by-market average choice of discriminating attribute q is available. If so, finite sample performance of the estimated parameter would be better.

If such a data set is not readily available as we assume here, however, we proceed as follows: Notice, as $m \to \infty$, we have the sample counterpart of $\sum_{m=1}^M \eta_{kq}^{0m}/M$ to converge in probability to η_{kq}^0 for any k and g by virtue of the weak law of large numbers. Similarly the sample counterpart of $\sum_{m=1}^M E\left\{\nu_{ik}^m | C_i \in Q_q, \boldsymbol{X}, \boldsymbol{\xi}(\boldsymbol{\theta}_d^0, \boldsymbol{p}^m, \boldsymbol{X}, \boldsymbol{s}^{0m}, P^{m0})\right\}/M$ to converge in probability to $E\left\{\nu_{ik} | C_i \in Q_q, \boldsymbol{X}, \boldsymbol{\xi}(\boldsymbol{\theta}_d^0, \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{s}^0, P^0)\right\}$ where $\boldsymbol{p} = (\boldsymbol{p}^1, \ldots, \boldsymbol{p}^m)$.

Let $\Pr\left\{d\nu_{ik} | C_i \in Q_q, \boldsymbol{X}, \boldsymbol{\xi}(\boldsymbol{\theta}_d, \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{s}^0, P^0)\right\}$ be the conditional density of consumer i's demographics ν_{ik} given his/her choice of discriminating attribute q and product characteristics $(\boldsymbol{p}, \boldsymbol{X}, \boldsymbol{\xi}(\boldsymbol{\theta}_d, \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{s}^0, P^0))$. Since the conditional expectation η_{dq}^0 can be written as

$$
\begin{aligned}
&E\left\{\nu_{id} | C_i \in Q_q, \boldsymbol{X}, \boldsymbol{\xi}(\boldsymbol{\theta}_d, \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{s}^0, P^0)\right\} \\
=\ & \int \nu_{id} \Pr\left\{d\nu_{id} | C_i \in Q_q, \boldsymbol{X}, \boldsymbol{\xi}(\boldsymbol{\theta}_d, \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{s}^0, P^0)\right\} \\
=\ & \frac{\int \nu_{id} \Pr\left\{C_i \in Q_q | \boldsymbol{X}, \boldsymbol{\xi}(\boldsymbol{\theta}_d, \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{s}^0, P^0), \nu_{id}\right\} P^0(d\nu_{id})}{\Pr\left\{C_i \in Q_q | \boldsymbol{X}, \boldsymbol{\xi}(\boldsymbol{\theta}_d, \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{s}^0, P^0)\right\}} \\
=\ & \frac{\int \nu_{id} \Pr\left\{C_i \in Q_q | \boldsymbol{X}, \boldsymbol{\xi}(\boldsymbol{\theta}_d, \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{s}^0, P^0), \boldsymbol{\nu}_i\right\} P^0(d\boldsymbol{\nu}_i)}{\Pr\left\{C_i \in \mathcal{Q}_q | \boldsymbol{X}, \boldsymbol{\xi}(\boldsymbol{\theta}_d, \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{s}^0, P^0)\right\}} \\
=\ & \int \nu_{id} \frac{\sum_{j \in Q_q} \sigma_{ij}(\boldsymbol{X}, \boldsymbol{\xi}(\boldsymbol{\theta}_d, \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{s}^0, P^0), \boldsymbol{\nu}_i; \boldsymbol{\theta}_d)}{\sum_{j \in \mathcal{Q}_q} \sigma_j(\boldsymbol{X}, \boldsymbol{\xi}(\boldsymbol{\theta}_d, \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{s}^0, P^0), \boldsymbol{\theta}_d, P^0)} P^0(d\boldsymbol{\nu}_i),
\end{aligned}
$$

we can form an identity, which is the basis for additional moment conditions

$$
\eta_{dq}^0 - \int \nu_{id} \frac{\sum_{j \in Q_q} \sigma_{ij}(\boldsymbol{X}, \boldsymbol{\xi}(\boldsymbol{\theta}_d^0, \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{s}^0, P^0), \boldsymbol{\nu}_i; \boldsymbol{\theta}_d^0)}{\sum_{j \in \mathcal{Q}_q} \sigma_j(\boldsymbol{X}, \boldsymbol{\xi}(\boldsymbol{\theta}_d^0, \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{s}^0, P^0), \boldsymbol{\theta}_d^0, P^0)} P^0(d\boldsymbol{\nu}_i) \equiv 0 \tag{19}
$$

for $q = 1, \ldots, N_p, k = 1, \ldots, K$.

Although P^0 is so far assumed known, we typically are not able to calculate the second term on the left hand side of (20) analytically and will have to approximate it by using the empirical distribution P^T of i.i.d. sample $\boldsymbol{\nu}_t, t = 1, \ldots, T$ from the underlying distribution P^0. The corresponding sample moments $\boldsymbol{G}_{J,T}^a(\boldsymbol{\theta}_d, \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{s}^n, P^R, \boldsymbol{\eta}^N)$, where a

on the shoulder stands for additional, are

$$\boldsymbol{G}^a_{J,T}(\boldsymbol{\theta}_d, \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{s}^n, P^R, \boldsymbol{\eta}^N) = \boldsymbol{\eta}^N - \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{\nu}_t \otimes \boldsymbol{\psi}_t(\boldsymbol{\xi}(\boldsymbol{\theta}_d, \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{s}^n, P^R), \boldsymbol{\theta}_d, P^R) \quad (20)$$

where

$$\boldsymbol{\eta}^N = (\eta^N_{11}, \ldots, \eta^N_{1N_p}, \ldots, \eta^N_{K1}, \ldots, \eta^N_{KN_p})^T, \boldsymbol{\psi}_t(\boldsymbol{\xi}(\boldsymbol{\theta}_d, \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{s}^n, P^R), \boldsymbol{\theta}_d, P^R) = \begin{pmatrix} \frac{\sum_{j \in \mathcal{Q}_1} \sigma_{tj}(\boldsymbol{X}, \boldsymbol{\xi}, \boldsymbol{\nu}_t; \boldsymbol{\theta}_d)}{\sum_{j \in \mathcal{Q}_1} \sigma_j(\boldsymbol{X}, \boldsymbol{\xi}, \boldsymbol{\theta}_d, P^R)} \\ \vdots \\ \frac{\sum_{j \in \mathcal{Q}_{N_p}} \sigma_{tj}(\boldsymbol{X}, \boldsymbol{\xi}, \boldsymbol{\nu}_t; \boldsymbol{\theta}_d)}{\sum_{j \in \mathcal{Q}_{N_p}} \sigma_j(\boldsymbol{X}, \boldsymbol{\xi}, \boldsymbol{\theta}_d, P^R)} \end{pmatrix}. (21)$$

The symbol $\otimes$ denotes the Kronecker product. The quantity $\boldsymbol{\psi}_t(\boldsymbol{\xi}, \boldsymbol{\theta}_d, P)$ is the consumer t's model-calculated purchasing probability of products with discriminating attribute q relative to the model-calculated market share of the same products. Note that these additional moments are again conditional on product characteristics $(\boldsymbol{p}, \boldsymbol{X}, \boldsymbol{\xi}(\boldsymbol{\theta}_d, \boldsymbol{s}^n, P^R))$, and thus depend on the indices J and T.

As a result, the actual sample-based objective function we minimize in the GMM estimation is the sum of norm of $G^d_M(\boldsymbol{\theta}_d, \{\boldsymbol{s}^m\}, \{P^{R_m}\})$ in (9), $G^c_M(\boldsymbol{\theta}, \{\boldsymbol{s}^m\}, \{P^{R_m}\})$ in (18), and $\boldsymbol{G}^a_{J,T}(\boldsymbol{\theta}_d, \boldsymbol{s}^n, P^R, \boldsymbol{\eta}^N)$ in (20), that is, the norm of

$$\boldsymbol{G}_{J,T}(\boldsymbol{\theta}, \boldsymbol{s}^n, P^R, \boldsymbol{\eta}^N) = \begin{pmatrix} G^d_M(\boldsymbol{\theta}_d, \{\boldsymbol{s}^m\}, \{P^{R_m}\}) \\ G^c_M(\boldsymbol{\theta}, \{\boldsymbol{s}^m\}, \{P^{R_m}\}) \\ \boldsymbol{G}^a_{M,T}(\boldsymbol{\theta}_d, \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{s}^n, P^R, \boldsymbol{\eta}^N) \end{pmatrix}, \quad (22)$$

a $M_1 + M_2 + KN_p \times 1$ vector.

Notice that the first two moments $\boldsymbol{G}^d_M$ and $\boldsymbol{G}^c_M$ in (22) are the sample moments averaged over products $m = 1, \ldots, M$, although the third moment $\boldsymbol{G}^a_{M,T}$ is averaged over consumers $t = 1, \ldots, T$.

3 Theoretical Results

Note that in the expression $\boldsymbol{G}^a_{M,T}(\boldsymbol{\theta}_d, \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{s}^n, P^R, \boldsymbol{\eta}^N)$, there exist five distinct randomnesses: one from the draws of the product characteristics $(\boldsymbol{X}, \boldsymbol{\xi}_j, \boldsymbol{w}_{1j}, \boldsymbol{\omega}_j)$, two from the sampling processes of consumers for $\boldsymbol{s}^n$ and $\boldsymbol{\eta}^N$ (not controlled by the econometrician), and two from the empirical distributions $\{P^{R_m}\}$ and P^T (employed by the econometrician). The impact of these randomnesses on the estimate of $\boldsymbol{\theta}$ are decided by the relative sizes of the samples—J, $\{n_m\}$, N, $\{R_m\}$, and T. Carefully operationalizing the sampling and the simulation errors obtain the following result.

Theorem 1 (Consistency and Asymptotic Normality of $\boldsymbol{\theta} = (\boldsymbol{\theta}_d, \boldsymbol{\theta}_c)$) *The estimator $\boldsymbol{\theta}$ that minimizes the objective function $\boldsymbol{G}_{J,T}(\boldsymbol{\theta}, \boldsymbol{s}^n, P^R, \boldsymbol{\eta}^N)^T \boldsymbol{W}_m \boldsymbol{G}_{J,T}(\boldsymbol{\theta}, \boldsymbol{s}^n, P^R, \boldsymbol{\eta}^N)$ is consistent, and asymptotically normal at the rate of $\sqrt{M}$ where $\boldsymbol{W}_m$ is a weighting matrix converging in probability to a positive definite matrix W.*

Due to the limitation of the space, we omit the proof.

4   Simulation Results

Throughout below,

4.1   Demand side setting

The model calculated share σ_{ij}^m is from random coefficient utility of consumer i of demographics ν_i^m in market m for product j as

$$U_{ij}^m = -\alpha p_j^m + \beta_0 x_j + \beta_1 x_j \nu_i^m + \xi_j^m + \epsilon_{ij}^m, \tag{23}$$

with the observed product characteristics $x_j = (p_j^m, x_j)$, the demand-side parameters $\theta_d = (\alpha, \beta_0, \beta_1)$, unobservable ξ_j^m correlated with price p_j^m and with unobservable idiosyncratic taste ϵ_{ij}^m assumed to be i.i.d. extreme value.

We assume there are $J = 20$ products available to consumers of any of m markets. The observed product characteristics x_j and unobserved product characteristics ξ_j^m are random draws from N(3,1) and N(0, 1) respectively. The consumer demographics ν_i^m is random draws from N(0,1) and consumer's idiosyncratic term ϵ_{ij}^m is assumed to be i.i.d. with type I extreme value to derive random-coefficient logit model. The price p_j^m of product j is endogenously determined in each regional market equilibrium and differs from market to market. We set the true demand side parameters as $\alpha = 1.0$, $\beta_0 = 1.0$ and $\beta_1 = 0.5$. These parameter values are common to all the regional markets. We set the total number of consumers in the national market as $I = 10,000$ and assume there exist the same number I/M of consumers in each regional market. We draw R_m consumers from these regional database to calculate ξ_j^m using (9) for $j = 1, \ldots, J$. We use the "conditionally true" market-by-market share s^{0m} using the population $I = 10,000$ instead of the observed share s^{nm} to focus on the effect of supply side and micro moments.

4.2   Supply side setting

We assume there are $F = 5$ oligopolistic suppliers in a national market, each producing the same number $J_f = 4$ of products. As stated above, the same set of products are sold in all regional markets. These suppliers are assumed to have the same cost function of product j in market m as

$$mc_j^m = x_j \gamma + \omega_j^m, \tag{24}$$

where the unobserved cost shifters ω_j^m are random draws from N(0, 1). We set the cost side parameter as $\gamma = 1.5$.

The true market share s^{0m} and the corresponding price $\boldsymbol{p}^m$ are determined at the equilibrium, and thus the values of $\boldsymbol{p}^m$ are obtained by solving (13), that is, a system of J-dimensional nonlinear simultaneous equations, which is solved by an iterative Newton-Raphson algorithm.

# of iternations		$\hat{\beta}_1$	
$R_m = 100$ (fixed)		without micro moments	with micro moments $(T = 1,000)$
Mean	$M = 1$	0.56609	0.565112
	$M = 5$	0.492358	0.496941
	$M = 10$	0.524297	0.506038
	$M = 20$	0.514595	0.494323
	$M = 40$	0.515111	0.498628
S.E.	$M = 1$	0.579877	0.239662
	$M = 5$	0.237166	0.0702588
	$M = 10$	0.172493	0.0696279
	$M = 20$	0.109436	0.0609522
	$M = 40$	0.070198	0.0591414

Table 1: Mean and standard error of β_1 with micro moments

4.3 Instruments

We construct three instruments from $\boldsymbol{X}$ for the product j produced by f: x_j itself, the sum of x_k within the firm f except x_j, and the sum of x_k over the firms other than f, as BLP (1995) proposed. We use these three instruments for demand- as well as supply-side.

4.4 Micro moment setting

We assume that information is available on (a) the expected value of ν_i^m over consumers who choose products priced higher than the average national price and (b) the expected value of ν_i^m over consumers who choose products with x_j greater than the average of x_j. We draw T consumers independent of R_m from the national population database to construct sample analogue of moment in (20). Since we use two discriminating attribute—(a) and (b) above, we have $N_p = 2$. Similarly we compare two—p_j and x_j—product characteristics, so we have $K = 2$. Also there are two instruments on demand- as well as supply-side, $M_1 = M_2 = 2$. Therefore $\boldsymbol{G}_{J,T}(\boldsymbol{\theta}, \boldsymbol{s}^n, P^R, \boldsymbol{\eta}^N)$ is $M_1 + M_2 + KN_p \times 1 = (2 + 2 + 2 \times 2) \times 1 = 8 \times 1$ vector. Since the objective function we use is $\boldsymbol{G}_{J,T}(\boldsymbol{\theta}, \boldsymbol{s}^n, P^R, \boldsymbol{\eta}^N)^T \boldsymbol{W}_m \boldsymbol{G}_{J,T}(\boldsymbol{\theta}, \boldsymbol{s}^n, P^R, \boldsymbol{\eta}^N)$, we need 8×8 weight matrix $\boldsymbol{W}$. For $\boldsymbol{W}$, we used 8×8 identity matrix. Notice that the calculation may not be as efficient relative to when we use the optimal weighting matrix. We set the true parameter values to be our initial values.

We are most interested in the the consumer heterogeneity (random coefficient) parameter β_1. For the parameter β_1, see Table 1 and Figure 1. We observe the following: (1) The finite sample estimates with micro moments seems asymptotically unbiased, while the estimates without remain biased; (2) The finite sample estimates with micro moments seems more accurate than the estimates without; (3) As expected, the effect of the micro moments wanes as the number m of markets grows. Due to the limitation of the space, we omit the results for the other parameter estimates.

Figure 1: Histogram of β_1 with the micro moments. Solid lines are density estimates and dashed lines are normal curve with the estimated mean and standard error.
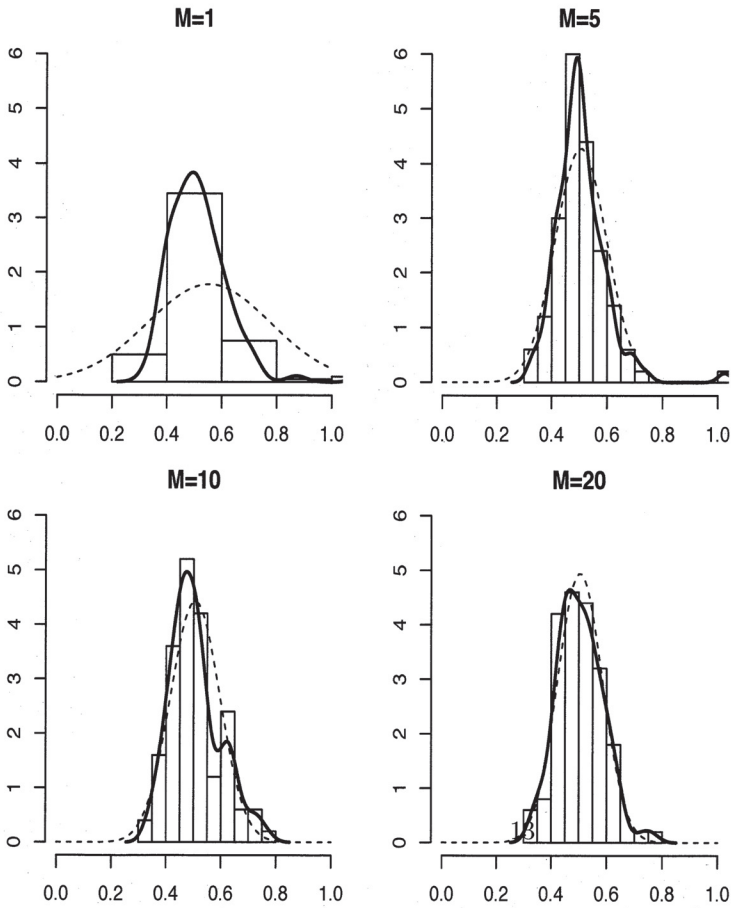
5 Conclusion and Discussion

Overall, we found the following: Improvement in bias correction as expected, but slight standard error improvement is surprising; National micro moments greatly helps us to evaluate the heterogeneity of consumers when the number of market is small (i.e. the information is limited) both in terms of standard error and bias; Contribution from adding micro moment decreases as the number of markets increases as expected. One note of caution for practitioners: With a small sample size T of demographics to construct micro moments, the accuracy of estimates seems to suffer.

References

[1] Berry, S. (1994), "Estimating Discrete Choice Models of Product Differentiation," RAND Journal of Economics, Vol 25(2), 242-262.

[2] Berry, S., J. Levinsohn and A. Pakes (1995), "Automobile Prices in Market Equilibrium," Econometrica, Vol.63, 841-890.

[3] Berry, S., J. Levinsohn and A. Pakes (2004), "Estimating Differentiated Product Demand Systems from a Combination of Micro and Macro Data: The New Car Model," Journal of Political Economy, vol. 112(1), 68-105.

[4] Berry, S., O. Linton and A. Pakes (2004), "Limit Theorems for Estimating the Parameters of Differentiated Product Demand Systems," Review of Economic Studies, Vol.71 613-654.

[5] Bresnahan, T. F. (1989), "Empirical Studies of Industries with Market Power." In Handbook of Industrial Organization, vol. 2, ed. Richard Schmalensee and Robert D. Willig, 101157. Amsterdam: North Holland.

[6] Che, H., Sudhir, K. and Seetharaman, P.B. (2007), "Bounded Rationality in Pricing under State-Dependent Demand: Do Firms Look Ahead, and If So, How Far?" Journal of Marketing Research, Vol. 44 (3), 434-449.

[7] Freyberger, J. (2015), "Asymptotic theory for differentiated products demand models with many markets." Journal of Econometrics, vol. 185, 162-181.

[8] Kamai, T. and Kanazawa (2016), "Is product with a special feature still rewarding? The case of the Japanese yogurt market." Cogent Economics & Finance, Vol. 4(1).

[9] Lancaster, K. (1971), Consumer Demand: A New Approach, Columbia University Press, New York.

[10] McFadden, D. (1974), "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka eds. Frontiers of Econometrics, Academic Press, New York.

[11] McFadden, D. (1981), "Econometric Models of Probabilistic Choice," in C. Manski and D. McFadden, eds. Structural Analysis of Discrete Data with Econometric Applications, MIT Press, Cambridge, MA.

[12] Myojo, S., Kanazawa, Y. (2012), "On Asymptotic Properties of the Parameters of Differentiated Product Demand and Supply Systems When Demographically-Categorized Purchasing Pattern Data are Available,"International Economic Review, Vol.53(3), 887-938.

[13] Nakayama, K. (2014), "Simulation Studies for Asymptotic Properties of the Parameters of Differentiated Product Demand and Supply Systems When the Number of Markets Increases,"University of Tsukuba, Master Thesis.

[14] Nevo, A. (2001) "Measuring Market Power in the Ready-to-Eat Cereal Industry,"Econometrica, vol. 69 (2), 307-342.

[15] Nelder, J., and R. Mead (1965), "A simplex method for function minimization,"Computer Journal 7, 308-313.

[16] Pakes, A. and D. Pollard (1989), "Simulation and the Asymptotics of Optimization Estimators,"Econometrica, vol. 57(5), 1027-1057.

[17] Petrin, A. (2002) "Quantifying the Benets of New Products: The Case of the Minivan,"Journal of Political Economy, Vol.110, 705-729.

[18] Suga, M. (2013), "On Asymptotic Properties of an Estimator for Demand When the Number of Markets Increases,"University of Tsukuba, Master Thesis.

[19] Takeshita, K. (2015), "CAN properties of the random-coefficient model of demand for nondurable consumer goods in the presence of national micro moments: A simulation study,"University of Tsukuba, Master Thesis.