Measuring Complementarity among Automobiles in a Multiple-Discrete Choice Model

Naoki Wakamori
University of Pennsylvania

Abstract

In several markets, such as the automobile and personal computer markets, consumers face multiple-discrete choices. In such cases, they are able to choose multiple products from differentiated products, which could potentially have some substitutabilities or complementarities. This paper presents a multiple-discrete choice model that allows decision makers to choose at most two differentiated products, which could have some complementarities or substitutabilities. Utilizing a newly collected micro-level household automobile ownership data and macro-level market share data in Japan, I use the model to estimate automobile demand in Japan.

1 Introduction

The automobile industry attracts not only industrial organization economists but also macro economists,[1] because of its size which can potentially have a large impact on economies. Automobiles are a differentiated durable product and a non-trivial fraction of households own more than one automobile, like other differentiated durable goods such as televisions, personal computers, and video recorders. Among these examples, it seems that there is no complementarity between multiple televisions, nor multiple video recorders. However, according to the newly collected data in a Japanese household level panel, *Keio Household Panel Survey*, there is an interesting pattern among owners of automobiles; regardless of their income, some constant fraction of households purchase one *Kei*-car[2] and one Normal-size car (hereafter referred to as Normal-car), even though they can afford to purchase two Normal cars. This fact suggests that there are some complementarities among the automobile consumption within one household. Therefore, the common assumption in the literature that each household (or decision maker) can purchase only one product might be problematic. Thus, the main goal of this paper is to present a estimable model that allows consumers to purchase more than one automobile, and to use this model to measure the welfare effect when the Japanese government changes the current automobile tax policy.

Estimating demand functions is one of the central issues for empirical economists, since it enables us to study the sources of market power as in **?** and **?**, measure the welfare effect from new products as in **?**, study trade policy as in **?**, and answer other policy related questions. One of the most common approaches in this literature is characteristics approach, developed by **?** and **?**, which considers products as bundles of characteristics, and assume that consumers maximize their utility derived from these product characteristics. Among them, the random coefficients discrete choice models, developed by **?** and **?**, is one of the most attractive and convenient approaches, because it does not require micro-level data and allows flexible substitution patterns. Due to these advantages, the random coefficient models are widely applied to estimate the differentiated products in various industries, such as **?** for the ready-to-eat cereal industry, **?** for the automobile industry, **?** for the Yellow Pages, and so on. The existing literature, however, is limited to analyzing a single discrete choice, i.e., decision makers can only choose one alternative from the choice set, because of difficulties in identification and computation.

There are several papers that tackle this problem. One of the most successful approach was developed by **?** where he studies the complementarity among paper version of newspapers and online newspapers. Basically, there are three

[1] For example, **?** study the effect of subsidies for scrapping cars in France, based on the dynamic stochastic general equilibrium model with automobile consumption.

[2] It is also called as light vehicle. Notice that the Japanese *Kei*-car is much smaller than the American compact car and subcompact car. It is close to British supermini car. The definition of a Japanese *Kei*-car is that the displacement of the automobile is less than 660 cc.

approaches[3] in the literature before **?**, but each approach needs to assume two differentiated products *ex-ante* are either complements, substitutes, or independent in an ad-hoc way. His model, however, is quite flexible in the sense that two differentiated products could be complements, substitutes, or independent. Although his approach is highly admirable, it does not use the characteristics approach, and thus it would not give me flexible substitution patterns if I used it directly for my demand estimation of automobiles. Therefore, in this paper, I estimate a random coefficient model with **?**'s method so that the model has both flexible substitution patterns and complementarities, using both micro-level consumer decision data and macro-level market share data.

In my model, I assume that the households consisting of one member (hereafter referred to as single-person households) purchase at most one automobile, while multiple-person households purchase at most two automobiles. I also assume that no households discard their automobiles within two years of purchase.[4] Using these assumptions, I can transform cross sectional data into a quasi-dynamic structure, and thus each household can be categorized by its ownership status of automobiles at the time of making decisions.[5] Moreover, I specify the complementarity term as a function not of each brand of automobiles, but by classification of automobiles. With this procedure, I can reduce the computational complexity and easily achieve identification. In particular, the households that eventually purchase two automobiles are useful for identifying the complementarity terms, while the households that only ever purchase one automobile are useful for identifying the coefficients of evaluations for each product characteristic.

This empirical study is also related to the literature on combining micro- and macro-level data when both types of datasets are available. In many occasions, empirical economists face some difficulties in having individual-level data. That is why the **?** method is very convenient because it enables us to estimate the demand functions from only macro-level market share data. However, I have both levels of data, and would like to utilize both sets of information. As **?** suggest and as applied by **?**, I construct the objective function from micro-level data and maximize it subject to the moment condition from macro-level data. In that way, I exploit both datasets.

This technical innovation enables us to assess the welfare effect of hypothetical

---

[3]The first approach was developed by **?** where he studies the return of computerization. In his model, companies face multiple-discrete choices as the decision makers, and they can purchase multiple units as well as multiple-brands at the same time. However, his model implicitly assumes that two products are substitutes. For the second approach, see **?** as an example. In their model, the goods must be independent and there is no interaction between two goods. The last approach is extending the choice set to the set of all possible bundles, and estimate it via logit or nested logit. See, **?**. This approach implicitly assumes that two products are complements.

[4]I checked the plausibility of this assumption by utilizing the panel data, and it suggests this assumption is not so problematic.

[5]For example, single households are classified under three categories: having no car, having one *Kei*-car, or having one Normal car. I also let households in the latter two states choose to purchase one of the products or nothing.

changes to the automobile tax system. The Japanese government has implemented tax advantages for purchasing new *Kei*-cars, and now there is a discussion over whether the government should abolish these advantages or not. Several Japanese manufactures of *Kei*-cars insist that the demand for *Kei*-cars would dramatically decrease if these advantages were abolished. The demand for *Kei*-cars, however, might not decrease so sharply if there exist some complementarities among *Kei*-cars and Normal-cars. Thus, in order to assess the welfare effect accurately, this empirical study is important. As it turns out, my current preliminary results show that there exist some complementarities among *Kei*-cars and Normal-cars.

The rest of this paper is structured as follows. In section 2, I show some statistics from the micro data as well as develop my model. I then describe the data that I use in this empirical study in section 3. Section 4 explains the detail of my estimation and section 5 gives estimation results. Section 6 concludes.

2 The Model

2.1 Motivating Facts

Most of the existing literature of estimating demand functions in automobile industries assumes that each household (or individual) purchases at most one automobile.[6] Newly collected Japanese household-level micro data recalls into question the plausibility of this assumption.

Table 1: Automobile Ownership in 2004

family	Number of Automobiles in a Household								
size	No Car		1 Car		2 Cars		3 Cars		Total
1	173	(50.9%)	162	(47.7%)	3	(0.9%)	2	(0.6%)	340
2	151	(20.0%)	408	(54.0%)	163	(21.6%)	34	(4.5%)	756
3	115	(12.4%)	522	(56.1%)	208	(22.4%)	85	(9.1%)	930
4	112	(10.2%)	547	(49.9%)	312	(28.4%)	126	(11.5%)	1097
5	45	(9.1%)	214	(43.3%)	149	(30.2%)	86	(17.4%)	494
6+	21	(5.4%)	124	(32.0%)	102	(26.3%)	141	(36.3%)	388
Total	617	(15.4%)	1977	(49.4%)	937	(23.4%)	474	(11.8%)	4005

Table **??** presents the relationship between the family size and the Japanese automobile ownership of households in 2004. The data shows that single-person households generally have one automobile at most, and half of them purchase nothing. The households with more than one automobile are almost a negligible fraction.

[6] Of course, there are some exceptions such as **?**. As I noted before, their assumption implicitly assumes that products are complements.

Table 2: Relationship between Income and Body-Size with one automobile in 2004

Annual HH income	Normal Car		*Kei* Car		Total
- $20,000	118	(75.2%)	39	(24.8%)	157
$20,000 - $30,000	181	(82.7%)	38	(17.4%)	219
$30,000 - $40,000	230	(85.5%)	39	(14.5%)	269
$40,000 - $50,000	186	(87.3%)	27	(12.7%)	213
$50,000 - $60,000	189	(87.9%)	26	(12.1%)	215
$60,000 - $70,000	134	(92.4%)	11	(7.6%)	145
$70,000 - $80,000	110	(93.2%)	8	(6.8%)	118
$80,000 - $90,000	112	(94.1%)	7	(5.9%)	119
$90,000 - $100,000	81	(95.3%)	4	(4.7%)	85
$100,000 - $110,000	66	(95.7%)	3	(4.4%)	69
$110,000 - $120,000	35	(97.2%)	1	(2.8%)	36
$120,000 - $130,000	33	(94.3%)	2	(5.7%)	35
$130,000 - $150,000	39	(97.5%)	1	(2.5%)	40
$150,000 -	37	(100.0%)	0	(0.0%)	37

On the other hand, if the households consist of more than one member, then they are likely to have at least one automobile. The larger the household size grows, the larger the number of automobiles owned by that household. This fact might imply that the common assumption in discrete choice literature that each household (individual) owns only one automobile is not plausible. However this observation is not enough to claim so. The reason is not so complicated: within one household they might choose their automobiles independently. Consider the case that one household consisting of two members, say husband and wife, purchases two automobiles. If their choices are independent, that is, there is no correlation between the two automobiles, the common assumption might not be problematic. If their choices however show some correlations, then the common assumption could be problematic. For example, they might have a large car as a first car for traveling, and a small car as a second car for commuting or daily shopping. Table **??** and **??** show the relationship between income and body-size of automobile, depending on the number of automobiles owned by a household.

According to Table **??**, as a household's annual income increase, it is more likely to have a Normal-car. The households who own a Kei car are basically low income households. Thus, for these households, Normal-cars and *Kei*-cars seem to be substitute goods. Table **??** however shows an interesting pattern that non-trivial fractions of the households with high income still purchase at least one *Kei*-car. Moreover, interestingly, the fraction of households with one Normal-car and one *Kei*-car does not decrease as the household income increases, unlike the decrease of the fraction of two *Kei*-car holders. This suggests the possibility that there exist some complementarities among automobile consumptions. If there are no complementarities, the fraction of households choosing one normal car and *Kei*-car should decrease as the household income increases.

The above discussion leads to the conclusion that applying the common assumption of choosing only one alternative in a differentiated market might be problematic for the demands function estimation in Japanese automobile market. Therefore in this paper I suggest a model which allows households to choose more than one automobile, taking account into the complementarities/substitutabilities among these automobiles.

Table 3: Relationship between Income and Body-Size: the number of Household with owning two/three automobiles in 2004

Annual HH Income	Household with 2 cars (N,N)		(N,K)		(K,K)		Sub-total	Household with 3 cars (N,N,N)		(N,N,K)		(N,K,K)		(K,K,K)		Sub-Total
- $20,000	26	(74.3%)	7	(20.0%)	2	(5.7%)	35	5	(45.5%)	4	(36.4%)	1	(9.1%)	1	(9.1%)	11
$20,000 - $30,000	50	(63.3%)	25	(31.7%)	4	(5.1%)	79	15	(60.0%)	4	(16.0%)	6	(24.0%)	0	(0.0%)	25
$30,000 - $40,000	54	(54.0%)	42	(42.0%)	4	(4.0%)	100	19	(46.3%)	17	(41.5%)	5	(12.2%)	0	(0.0%)	41
$40,000 - $50,000	70	(54.7%)	56	(43.8%)	2	(1.6%)	128	23	(54.8%)	14	(33.3%)	5	(11.9%)	0	(0.0%)	42
$50,000 - $60,000	39	(46.4%)	44	(52.4%)	1	(1.2%)	84	26	(59.1%)	13	(29.6%)	2	(4.6%)	3	(6.8%)	44
$60,000 - $70,000	51	(56.0%)	39	(42.9%)	1	(1.1%)	91	13	(33.3%)	25	(64.1%)	1	(2.6%)	0	(0.0%)	39
$70,000 - $80,000	54	(66.7%)	27	(33.3%)	0	(0.0%)	81	7	(26.9%)	14	(53.9%)	5	(19.2%)	0	(0.0%)	26
$80,000 - $90,000	28	(50.9%)	27	(49.1%)	0	(0.0%)	55	30	(68.2%)	9	(20.5%)	5	(11.4%)	0	(0.0%)	44
$90,000 - $100,000	32	(72.7%)	12	(27.3%)	0	(0.0%)	44	21	(47.7%)	22	(50.0%)	1	(2.3%)	0	(0.0%)	44
$100,000 - $110,000	25	(67.6%)	12	(32.4%)	0	(0.0%)	37	11	(47.8%)	9	(39.1%)	2	(8.7%)	1	(4.4%)	23
$110,000 - $120,000	15	(60.0%)	10	(40.0%)	0	(0.0%)	25	12	(46.2%)	12	(46.2%)	2	(7.7%)	0	(0.0%)	26
$120,000 - $130,000	13	(72.2%)	5	(27.8%)	0	(0.0%)	18	5	(62.5%)	3	(37.5%)	0	(0.0%)	0	(0.0%)	8
$130,000 - $150,000	24	(75.0%)	8	(25.0%)	0	(0.0%)	32	17	(51.5%)	13	(39.4%)	2	(6.1%)	1	(3.0%)	33
$150,000 -	17	(56.7%)	12	(40.0%)	1	(3.3%)	30	12	(75.0%)	3	(18.8%)	1	(6.3%)	0	(0.0%)	16

Remark: K denotes the *Kei*-car and N denotes the Normal-car in the bracket in the second row.
For example, (N,K,K) means that the household owns two *Kei*-cars and one Normal-car.

2.2 Households' Behavior

Let $i = 1, 2, \cdots, N$ denote the individual household. I divide the households into the two mutually exclusive groups, S and F, by their family size, where S denotes the set of single-person households and F denotes the set of multiple-person households. That is, if household i consists of only one member, then $i \in S$, and otherwise $i \in F$. Assume that each household $i \in S$ owns at most one automobile, and each household $i \in F$ owns at most two automobiles.[7] Let $J = 0, 1, \cdots, J$ denote the product, and $j = 0$ denote the outside option.

2.2.1 Single-Person Household

Each household $i \in S$ solves the following maximization problem;

$$\max_{C,a^1} \quad U_i(C, a^1) = \max_{C,a^1}[u^c(C) + u_i^a(a^1)], \quad \text{s.t.} \quad C + p(a^1) = y_i$$

where u_i^a is the utility from automobile consumptions which could be different for each household even though they have the same automobile, and u^c is the utility from other consumptions. Now, I specify these utility functions as $u^c(C) = \alpha C$ and $u_i^a(j) = \bar{u}_{ij} + \varepsilon_{ij}$, where $\bar{u}_{ij}$ is defined by

$$\bar{u}_{ij} = \boldsymbol{x}_j \tilde{\boldsymbol{\beta}}_i' + \xi_j = \sum_{m=1}^{M} x_{jm} \tilde{\beta}_{im} + \xi_j, \tag{1}$$

with

$$\tilde{\beta}_{im} = \bar{\beta}_m + \sum_{r=1}^{R} z_{ir}^p \beta_{mr}^o + \beta_m^u \nu_{im}, \tag{2}$$

where the $\boldsymbol{x}_j = [x_{j1}, \cdots, x_{jM}]$ and the ξ_j are the observed and unobserved characteristics for *product* j respectively, the $\tilde{\boldsymbol{\beta}}_i = [\tilde{\beta}_{i1}, \cdots, \tilde{\beta}_{iM}]$ represents the household i specific evaluation for each product characteristic, the $\boldsymbol{z}_i^p = [z_{i1}^p, \cdots, z_{iR}^p]$ and ν_i are observed and unobserved *household* attributes, and the ε_{ij} denote idiosyncratic individual preferences, assumed to be independent of the product characteristics and of each other. Moreover, the $\boldsymbol{\beta}^o$ and the $\boldsymbol{\beta}^u$ denote the coefficient for the observable and unobservable households' attributes. One of the merits of this specification is that each household is able to have a different evaluation for each product, depending on the household attributes. For example, a household with a lot of kids might prefer a large seating capacity car to small one. Substituting (**??**) into (**??**) and putting them together with original maximization problem, the utility of household

[7] For simplicity, I assume that households purchase at most two automobiles, since the discussion of complementarities or substitutabilities among three goods is quite difficult and there is no paper that deals with this problem so far as I know.

i choosing j can be given by the following simple equation,

$$
\begin{aligned}
u_{ij} &= \boldsymbol{x}_j \boldsymbol{\beta}_i' + \xi_j + \varepsilon_{ij} + \alpha(y_i - p_j) \\
&= \underbrace{\sum_{m=1}^{M} x_{jm}\bar{\beta}_m + \xi_j}_{\delta_j} + \underbrace{\sum_{m=1}^{M} x_{jm}\left[\sum_{r=1}^{R} z_{ir}^{P}\beta_{mr}^{o} + \beta_m^u \nu_{im}\right] + \alpha(y_i - p_j)}_{\mu(\boldsymbol{x}_j,\boldsymbol{\beta},\boldsymbol{\nu}_i,\boldsymbol{z}_i,y_i,p_j)} + \varepsilon_{ij}.
\end{aligned}
$$

For notational simplicity, let δ_j denote the mean utility from product j which is the same for every household, and $\mu(\boldsymbol{x}_j, \boldsymbol{\beta}, \boldsymbol{\nu}_i, \boldsymbol{z}_i, y_i, p_j)$ denote the remaining part except ε_{ij}. Assuming that ε follows Type I extreme value distribution, the probability of choosing product j conditional on the characteristics of household i, and all products is given by

$$
\Pr(d_i^1 = j | \boldsymbol{z}_i, \boldsymbol{\nu}_i, \boldsymbol{x}, \boldsymbol{\theta}) = \frac{\exp[\delta_j + \mu(\boldsymbol{x}_j, \boldsymbol{\beta}, \boldsymbol{\nu}_i, \boldsymbol{z}_i, y_i, p_j)]}{1 + \sum_{l \in J} \exp[\delta_l + \mu(\boldsymbol{x}_l, \boldsymbol{\beta}, \boldsymbol{\nu}_i, \boldsymbol{z}_i, y_i, p_l)]}, \tag{3}
$$

where $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^o, \boldsymbol{\beta}^u, \delta)$ is the set of estimates, and d_i^g is the household i's choice for the g-th, $g = 1, 2$, automobile.

2.2.2 Multiple-Person Household

Basically, each household $i \in F$ solves the following maximization problem;

$$
\max_{C,\boldsymbol{a}} U_i(C, \boldsymbol{a}) = u^c(C) + u_i^a(\boldsymbol{a}) \qquad \text{s.t.} \;\; C + p(a^1) + p(a^2) = y_i,
$$

with

$$
\begin{aligned}
u^c(C) &= \alpha C, \\
u_i^a(\boldsymbol{a}) &= u_i(a^1) + u_i(a^2) + \Gamma(a^1, a^2, z_i^c) + \varepsilon_{ir},
\end{aligned}
$$

where $u_i(a^l)$ is the same as before, i.e. if $a^l = j$, then $u_i(j) = \boldsymbol{x}_j \boldsymbol{\beta}_i + \xi_j$. This specification is almost the same as the single household problem, except now we have an interaction term, Γ, in their utility function, which will capture substitutabilities and complementarities between two automobiles.

Interaction Term The most desirable way to capture the complementarities and substitutabilities is defining it pairwisely, i.e. we need to define the substitutability/complementarity term for each possible combination. It is, however, almost impossible to estimate them due to difficulties in identification and computation.[8] Therefore, I categorize them into two mutually exclusive sets, the set of *Kei*-cars denoted by $\mathcal{K}$ and the set of Normal-cars denoted by $\mathcal{N}$, following the standard

[8] Suppose in the market there are 100 differentiated products, i.e., $J = 100$. Then, the number of possible combinations for these products should be $4,950$, and thus the element of choice set becomes more than $5,000$, since the decision maker can also choose only one product.

Japanese classification of automobiles, so that I can reduce the dimensions and computational burden. More explicitly, I assume the following parametric form for Γ;

$$\Gamma(a^1, a^2, z_i^c) = \begin{cases} \Gamma_{KK}, & \text{if } (a^1, a^2) \in \mathcal{K} \times \mathcal{K} \\ \Gamma_{KN}, & \text{if } (a^1, a^2) \in \mathcal{K} \times \mathcal{N} \cup \mathcal{N} \times \mathcal{K} \\ \Gamma_{NN}, & \text{if } (a^1, a^2) \in \mathcal{N} \times \mathcal{N} \\ 0, & \text{otherwise} \end{cases}$$

with

$$\Gamma_r = \gamma_r^1 + \sum_{l=2}^{L} \gamma_r^l z_{il}^c, \quad \text{for} \quad r = KK, NK, NN,$$

where the $z_i^c = [z_{i1}^c, \cdots, z_{iL}^c]$ is the household i's characteristic which does not have common characteristics with z_i^p,[9] and the $\gamma_r = [\gamma_r^1, \cdots, \gamma_r^L]$ is the coefficient for the household characteristics, which includes a constant term.

I want to emphasize why this interaction term, Γ, is able to capture the complementarities or substitutabilities using a simple example.[10] Suppose there are two products, K and N, and consumers can choose at most one unit of each.[11] Thus, the choice set for consumers consists of $\{0, N, K, KN\}$. Let $u'_0, u'_N, u'_K, u'_{KN}$ denote the gross utility from each consumption bundle respectively, and define Γ_{KN} as

$$\Gamma_{KN} = (u'_{KN} - u'_K) - (u'_N - u'_0).$$

The first bracket is the incremental utility of having product N additionally, when the consumer originally has product K. The second bracket is the incremental utility of having product N additionally, when the consumer originally has nothing. Thus, if there are complementarities between product K and N, Γ_{KN} becomes positive, since the first bracket should be greater than the second. On the other hand, if these goods are substitutes, then Γ_{KN} will be negative since the second bracket should be larger than the first one. So, normalizing the utility by u'_0, we can define

$$\begin{aligned} u_0 &= 0, \\ u_n &= \delta_n - p_n, \\ u_k &= \delta_k - p_k, \\ u_{kn} &= u_n + u_k = (\delta_k + \delta_n) - (p_k + p_n) + \Gamma_{KN}. \end{aligned}$$

Denoting $F(\boldsymbol{u})$ as the distribution of $\boldsymbol{u} = (u_k, u_n, u_{kn})$, and assuming that con-

[9] This is necessary for the identification condition.
[10] See also Section I. in **?** to more detailed explanation.
[11] We can immediately extend this model to allow consumers to choose two unit of the same product.

sumers maximize utility, the choice probabilities are given by

$$\begin{aligned} P_k &= \int_{\boldsymbol{u}} \mathbf{1}_{\{u_k \geq 0\}} \mathbf{1}_{\{u_k \geq u_n\}} \mathbf{1}_{\{u_k \geq u_{kn}\}} dF(\boldsymbol{u}), \\ P_n &= \int_{\boldsymbol{u}} \mathbf{1}_{\{u_n \geq 0\}} \mathbf{1}_{\{u_n \geq u_k\}} \mathbf{1}_{\{u_n \geq u_{kn}\}} dF(\boldsymbol{u}), \\ P_{kn} &= \int_{\boldsymbol{u}} \mathbf{1}_{\{u_{kn} \geq 0\}} \mathbf{1}_{\{u_{kn} \geq u_k\}} \mathbf{1}_{\{u_{kn} \geq u_n\}} dF(\boldsymbol{u}). \end{aligned}$$

Therefore, the expected demand per consumer for goods k and n are given by $Q_k = P_k + P_{kn}$ and $Q_n = P_n + P_{kn}$. The standard definitions of complements and substitutes are: Goods n and k are substitutes if $\frac{\partial Q_k}{\partial p_n} > 0$, independent if $\frac{\partial Q_k}{\partial p_n} = 0$, and complements if $\frac{\partial Q_k}{\partial p_n} < 0$. Then **?** shows the following proposition, which enables us to capture the complementarities or substitutabilities by Γ;

Proposition 1 *(Gentzkow 2007) Goods n and k are substitutes if $\Gamma_{KN} < 0$, independent if $\Gamma_{KN} = 0$, and complements if $\Gamma_{KN} > 0$.*

Choice Set for Non-single Household Only some consumers face the above decision problem, as some might already have automobiles. Thus, I need to classify the state of consumers' automobile ownership at the decision time. There are six states: (F1) the households with no car, (F2) the households with one *Kei*-car, (F3) the household with one Normal-car, (F4) the household with two *Kei*-cars, (F5) the household with one *Kei*-car and one Normal-car, and (F6) the households with two Noraml-cars.

Table **??** summarizes these states for the households and choice sets for each category of the households. Within these six categories, the households with two cars, i.e., from (F4) to (F6), will purchase nothing with probability one. The choice set for the household with one car, i.e., (F2) and (F3), should consist of (i) purchase nothing, (ii) purchase one *Kei*-car, and (iii) purchase one Normal-car. The choice set for the household with no car consists of (i) purchase nothing, (ii) purchase one *Kei*-car, (iii) purchase one Normal-car, (iv) purchase two *Kei*-cars, (v) purchase one *Kei*-and one Noraml-car, and (vi) purchase two Noraml-cars.

For the household with one car which is denoted by j', the maximization problem can be modified to a very simple discrete choice problem

$$\max_{a^2 \in J} \; u_i(j') + u_i(a^2) + \Gamma(j', a^2, z_i^c) + \alpha(y_i - p(a^2)) + \varepsilon_{ir}.$$

Thus, since the utility level from the 1st automobile does not affect the choice for the next automobile,[12] the probability of choosing product j can be calculated as

$$\Pr(d_i^2 = j | d_i^1 = j', \boldsymbol{z}_i, \boldsymbol{\nu}_i, \boldsymbol{x}, \boldsymbol{\theta}) = \frac{\exp[\delta_j + \mu_{ij} + \Gamma(j', j, z_i^c)]}{1 + \sum_{l \in J} \exp[\delta_l + \mu_{il} + \Gamma(j', l, z_i^c)]}, \tag{4}$$

[12] Of course, the classification of 1st automobile DOES affect the choice for the second automobile.

which is the almost same as eq. (??), except now it contains the interaction term, Γ.

For the households with no car, the estimation procedure is a little bit more complicated than the other cases, and here I utilize the quasi-dynamic structure of *KHPS*. First, I let the household choose only one car, say d_i^1. Then, if $d_i^1 \neq 0$, I again let them choose the second car, d_i^2, given the first car. More conveniently, let $k \in \mathcal{K}$ and $n \in \mathcal{N}$ denote the representative element of each categories, and $\Pr[a^1, a^2] = \Pr[d_i^1 = a^1, d_i^2 = a^2 | \boldsymbol{z}_i, \boldsymbol{\nu}_i, \boldsymbol{x}, \boldsymbol{\theta}]$. Then, the probability distribution for the households with no car becomes

$$\left.\begin{aligned}
\Pr[0,0] &= \Pr(d_i^1 = 0|\boldsymbol{z}_i, \boldsymbol{\nu}_i, \boldsymbol{x}, \boldsymbol{\theta}) \\
\Pr[k,0] &= \Pr(d_i^2 = 0|d_i^1 = k, \boldsymbol{z}_i, \boldsymbol{\nu}_i, \boldsymbol{x}, \boldsymbol{\theta})\ \Pr(d_i^1 = k|\boldsymbol{z}_i, \boldsymbol{\nu}_i, \boldsymbol{x}, \boldsymbol{\theta}) \\
\Pr[n,0] &= \Pr(d_i^2 = 0|d_i^1 = n, \boldsymbol{z}_i, \boldsymbol{\nu}_i, \boldsymbol{x}, \boldsymbol{\theta})\ \Pr(d_i^1 = n|\boldsymbol{z}_i, \boldsymbol{\nu}_i, \boldsymbol{x}, \boldsymbol{\theta}) \\
\Pr[k,k] &= \Pr(d_i^2 = k|d_i^1 = k, \boldsymbol{z}_i, \boldsymbol{\nu}_i, \boldsymbol{x}, \boldsymbol{\theta})\ \Pr(d_i^1 = k|\boldsymbol{z}_i, \boldsymbol{\nu}_i, \boldsymbol{x}, \boldsymbol{\theta}) \\
\Pr[k,n] &= \Pr(d_i^2 = k|d_i^1 = n, \boldsymbol{z}_i, \boldsymbol{\nu}_i, \boldsymbol{x}, \boldsymbol{\theta})\ \Pr(d_i^1 = n|\boldsymbol{z}_i, \boldsymbol{\nu}_i, \boldsymbol{x}, \boldsymbol{\theta}) \\
&\qquad + \Pr(d_i^2 = n|d_i^1 = k, \boldsymbol{z}_i, \boldsymbol{\nu}_i, \boldsymbol{x}, \boldsymbol{\theta})\ \Pr(d_i^1 = k|\boldsymbol{z}_i, \boldsymbol{\nu}_i, \boldsymbol{x}, \boldsymbol{\theta}) \\
\Pr[n,n] &= \Pr(d_i^2 = n|d_i^1 = n, \boldsymbol{z}_i, \boldsymbol{\nu}_i, \boldsymbol{x}, \boldsymbol{\theta})\ \Pr(d_i^1 = n|\boldsymbol{z}_i, \boldsymbol{\nu}_i, \boldsymbol{x}, \boldsymbol{\theta})
\end{aligned}\right\} \tag{5}$$

where $\Pr(d_i^1 = j|\boldsymbol{z}_i, \boldsymbol{\nu}_i, \boldsymbol{x}, \boldsymbol{\theta})$ and $\Pr(d_i^2 = j'|d_i^1 = j, \boldsymbol{z}_i, \boldsymbol{\nu}_i, \boldsymbol{x}, \boldsymbol{\theta})$ are defined by equations (??) and (??).

2.3 Firms' Strategic Pricing Behavior

Since there are only ten manufactures in the Japanese automobile market, it is natural to assume that their ways of setting prices and choosing product lineups are affected by other firms' strategy. Moreover, as Table ?? presents, each firm produces multiple models in a given year. Thus firms need to consider not only other firms' strategy, but also the effect of their pricing strategy on other products they produce, when firms set the prices. In this situation, the profit for firm f, $f = 1, 2, \cdots, F$ can be written as

$$\Pi_f = \sum_{j \in \mathscr{F}_f} (p_j - \mathrm{mc}_j) s_j \tilde{Q}, \qquad \text{with} \qquad \ln(\mathrm{mc}_j) = \boldsymbol{x}_j \boldsymbol{\phi}' + \omega_j,$$

where the $\mathscr{F}_f$ is the set of products which is produced by firm f, the mc_j and s_j denote the marginal costs and the market share of product j, $\tilde{Q}$ is the potential market size, the $\boldsymbol{\phi}$ denotes the cost parameters for the product characteristics, and the ω_j represents the unobservable cost factors. This formulation is able to capture not only the strategic interaction among firms, but also the pricing strategy within one firm. Because of the time constraint, however, I do not include the firms' behavior in my estimation.

3 The Data

For this empirical study I use mainly four data sets; *Keio Household Panel Survey* (hereafter reffered to as *KHPS*) which contains household-level information, *Auto-*

Table 4: List of Automobile Makers and Their Product Lineup

	1997				2004			
	Kei		Normal		Kei		Normal	
Daihatsu	28	(63.6%)	16	(36.4%)	27	(73.0%)	10	(27.0%)
Fuji Heavy	16	(29.1%)	39	(70.9%)	7	(18.9%)	30	(81.1%)
Honda	7	(7.2%)	90	(92.8%)	7	(13.5%)	45	(86.5%)
Isuzu	0	(0.0%)	21	(100.0%)	-		-	
Mazda	4	(6.7%)	56	(93.3%)	7	(9.3%)	68	(90.7%)
Mitsubishi	16	(18.6%)	70	(81.4%)	12	(30.0%)	28	(70.0%)
Mitsuoka	-		-		1	(20.0%)	4	(80.0%)
Nissan	0	(0.0%)	119	(100.0%)	1	(3.5%)	28	(96.6%)
Suzuki	8	(44.4%)	10	(55.6%)	23	(60.5%)	15	(39.5%)
Toyota	0	(0.0%)	126	(100.0%)	0	(0.0%)	67	(100.0%)
Total	79	(12.6%)	547	(87.4%)	85	(22.4%)	295	(77.6%)

motive Guidebook which provides the product-level panel data, *Automobile Ownership Statistics* which gives the aggregate sales number of automobiles in Japan, and *2005 Population Census* for macro-level household characteristics distribution. I describe the characteristics of these datasets in this section.

3.1 Keio Household Panel Survey: Micro Data for Households

KHPS is provided by Keio University, one of the private research universities in Tokyo, Japan. One of the main goals of KHPS is providing the Japanese household level micro panel data in order to promote empirical research about Japan. The sample size of KHPS was approximately 4,000 households with 7,000 individuals from 2004 to 2006.[13] As for automobiles ownership, KHPS inquires in 2004 about: (1) month and year of purchase, (2) maker, brand, and model of automobiles, and (3) whether it was purchased as a new car or a used car, for up to three cars. Unfortunately, it does not inquire into the purchase prices. Since I use the static model in this study, basically I utilize the ownership information from 2004.

3.2 Automotive Guidebook: Micro Data for Products

Automotive Guidebook series are issued by Japan Automobile Manufactures Association (JAMA) every year. I construct the product level panel data from the series of this book, since each year edition provides the set of available models of automobiles and the characteristics for each automobile, such as price, interior/exterior

[13]Starting from 2007, the sample size has been increased by 1,400 households with 2,500. Thus we currently have 5,400 households with 9,500 individuals in total.

dimensions, seating capacity, and displacement.[14] Table ?? shows the average characteristics of automobiles which were sold in 2003.

Table 5: Mean Product Characteristics for 1997 and 2003

		Kei		Normal	
		Mean	S.D.	Mean	S.D.
Exterior	Length	3379	100.6	4421	380.6
Dimension	Width	1475	0.000	1739	70.68
(mm)	Height	1602	129.1	1564	172.8
Interior	Length	1735	226.4	2016	452.6
Dimension	Width	1257	33.74	1455	77.85
(mm)	Height	1253	73.92	1217	70.11
Weight	(kg)	840.2	86.18	1388	293.1
Capacity	(person)	3.906	0.4260	5.359	1.322
Mileage		18.92	3.126	13.07	3.842
Displacement	(cc)	658.1	0.8368	2149	732.4
Max. Power	(PS/rpm)	57.14	6.614	162.4	57.84
Price	(¥)	1,113,424	232,013	2,475,661	1,387,203
# of obs.		85	(22.4%)	295	(77.6%)

3.3 Automobile Ownership Statistics: Macro Data for Car Sales

Automobile Ownership Statistics provide the number of automobiles which were sold in a given year, under the supervision of Ministry of Land, Infrastructure, Transportation, and Tourism. Because all Japanese automobiles need to be registered to the government and there are only ten manufactures in Japan, the exact number of automobiles that were sold in a given year is available. As for the sales of used automobiles, however, it is difficult to know the exact number of automobile sales since there are so many companies which deal with used cars and it is difficult to collect and aggregate these decentralized markets information. I therefore predict the used car sales based on this *Automobile Ownership Statistics* and *KHPS*.[15]

4 The Estimation

If there is no unobservable term, $\boldsymbol{\xi}$, in the utility function, then the estimation can be done by a straight forward way, such as maximum likelihood, so that we can match the market shares for each product to those observed in the data. In my

[14] The entire list of these characteristics is summarized in Table ?? in Appendix.
[15] I explain how to predict it in the estimation section.

model, however, there is an unobservable term, $\boldsymbol{\xi}$, in the utility function. There are two ways reconciling this problem; (1) put a distributional assumption on $\boldsymbol{\xi}$ so that we can integrate out the ξ from the probability, or (2) put an orthogonality assumption between $\boldsymbol{X}$ and $\boldsymbol{\xi}$ so that we can utilize the orthogonality condition for the estimation. I apply the latter strategy which is developed by **?** and commonly used by other papers such as **?** and **?**.

Moreover, although **?** have only macro-level market share data not micro-level data, I have both micro-level decision data and macro-level market share data. In this situation, as **?** suggests and **?** follows, I construct the GMM objective function from both micro- and macro-level data as moment conditions. Intuitively, I minimize the set of moments conditions from micro-level data subject to the a moment condition from macro-level data is equal to zero.

4.1 Objective Function and Estimation Procedure

In this empirical study, I match three "sets" of predicted moments to their data analogues: (1) the market share of the J products, (2) the covariance of the observed consumer attributes $\boldsymbol{z}_i^p$, with the observed product characteristics, $\boldsymbol{x}_j$ which is chosen by the households, and (3) the covariance of the observed product characteristics for household with two cars. In this section, I define these sets of moments, explaining the algorithm and procedure of my estimation.

The first set of moments, the market shares of the J products, can be derived from the macro data. As **?** suggests, make a initial guess for δ_j^0 by taking the difference between log of market share of outside option and market share of product j;

$$\delta_j^0 = \ln(s_0^D) - \ln(s_j^D), \quad \text{for } \forall j \neq 0$$

where the s_j^D denotes the actual observed market share for product j. Then, given this $\boldsymbol{\delta}^0$ and $\boldsymbol{\theta}$, I calculate the market shares for each product which are given by

$$\begin{aligned}
s_0^p(\boldsymbol{\delta}|\boldsymbol{\theta}) &= N_{S_1}P_{S_1}(0) + N_{F_1}P_{F_1}(0) + N_{F_2}P_{F_2}(0) + N_{F_3}P_{F_3}(0) \\
&\qquad + N_{S_2} + N_{S_3} + N_{F_4} + N_{F_4} + N_{F_5} + N_{F_6}, \\
s_1^p(\boldsymbol{\delta}|\boldsymbol{\theta}) &= N_{S_1}P_{S_1}(1) + N_{F_1}[P_{F_1}(1) + 2P_{F_1}(3) + P_{F_1}(4)] \\
&\qquad + N_{F_2}P_{F_2}(2) + N_{F_3} + P_{F_3}(2), \\
s_2^p(\boldsymbol{\delta}|\boldsymbol{\theta}) &= N_{S_1}P_{S_1}(2) + N_{F_1}[P_{F_1}(2) + P_{F_1}(4) + 2P_{F_1}(5)] \\
&\qquad + N_{F_2}P_{F_2}(2) + N_{F_3} + P_{F_3}(2),
\end{aligned}$$

where the N_{S_h} and N_{F_h} represent the number of single and non-single households which is in state h, and $P_{S_h}(d)$ and $P_{F_h}(d)$ denote the probability that single and non-single households in state h choose d, which is defined in Table **??**. In order to make it clear, consider the $s_1^p(\boldsymbol{\delta}|\boldsymbol{\theta})$. The market share for product $j = 1$ is the sum of the number of single household in state 1 that purchase product 1, the number of non-single households in state 2 and 3 that purchase product 1, the number of

Table 6: Choice Set for Households

ownership status at Dec. 2000	# of obs. in KHPS	Available Choice (d)					
		0	1	2	3	4	5
		(0,0)	(K,0)	(N, 0)	(K, K)	(N, K)	(N, N)
Single							
(1) no car	72.5%	√	√	√			
(2) one K	2.8%	√					
(3) one N	24.7%	√					
Family							
(1) no car	37.5%	√	√	√	√	√	√
(2) (K, 0)	6.4%	√	√	√			
(3) (N, 0)	43.5%	√	√	√			
(4) (K, K)	0.2%	√					
(5) (K, N)	5.5%	√					
(6) (N, N)	6.8%	√					

non-single households in state 1 that purchase only one product 1, $d_i = 1, 4$, and the doubled number of non-single households in state 1 purchasing two product 1.

Then, using new predicted market shares, I obtain an updated $\boldsymbol{\delta}^{t+1}$ by the following formula:

$$\delta_j^{t+1} = \delta_j^t + \ln(s_j^D) - \ln(s_j^p(\boldsymbol{\delta}|\boldsymbol{\theta})),$$

Using a contraction mapping, repeat this algorithm until the difference between $\boldsymbol{\delta}^{t+1}$ and $\boldsymbol{\delta}^t$ becomes a smaller value than a tolerance level. Going through this algorithm, I finally obtain the first set of moments:

$$G^1(\boldsymbol{\theta}) = \sum_{j=0}^{J}(s_j^D - s_j^P),$$

which is defined by the difference of the observed market shares and the predicted market shares.

The second set of moments is derived from the micro data. Having obtained $\boldsymbol{\delta}$, it is easy to calculate the decision probability for each household in micro *KHPS* using the household characteristics via equations (**??**) and (**??**). Now, I know $\boldsymbol{z}_i$ exactly, so I do not need to integrate them out, though I still need to integrate $\boldsymbol{\nu}_i$ out. After obtaining these probabilities, I construct the covariance of the observed consumer attributes $\boldsymbol{z}_i^p$, with the observed product characteristics, $\boldsymbol{x}_j$ which is chosen by the households

$$G^2(\boldsymbol{\theta}) = \sum_{i \in B}\left[z_i \left\{\sum_j (\mathbf{1}_{\{d_i=j\}} - \Pr[d_i = j|\boldsymbol{x}, \boldsymbol{z}_i, \boldsymbol{\theta}])\boldsymbol{x}_j\right\}\right],$$

where B denotes the set of person who purchase one product in *KHPS*. This set of moment conditions is very useful to identify $\boldsymbol{\beta}^o$, since it enable us to predict what kind of household's attributes contributes to make them purchase a particular product.

Finally, I set the third set of moments as the covariance of the observed product characteristics for household with two cars, given that the households own two automobiles eventually. Conceptually, it should be $\mathrm{E}[x_1^D x_2^D] - \mathrm{E}[x_1^P x_2^P]$ where the x_l^P and x_l^D denote the product characteristics of the model prediction and actual data, respectively. More precisely, I can obtain it as

$$
\begin{aligned}
G^3(\boldsymbol{\theta}) \quad = \quad & \sum_i \left[\left\{ \sum_j (\mathbf{1}_{\{d_i^1 = j\}} - \Pr[d_i^1 = j | \boldsymbol{x}, \boldsymbol{z}_i, \boldsymbol{\theta}]) \boldsymbol{x}_j \right\} \right] \\
& \qquad \left[\left\{ \sum_{j'} (\mathbf{1}_{\{d_i^2 = j'\}} - \Pr[d_i^2 = j' | \boldsymbol{x}, \boldsymbol{z}_i, \boldsymbol{\theta}]) \boldsymbol{x}_{j'} \right\} \right].
\end{aligned}
$$

This moment conditions are particularly important for identifying the coefficient for complementarity term, $\boldsymbol{\gamma}_r$.

4.2 Details of Estimation - Data Handling

4.2.1 Potential Market Share

As **?** notes, the potential market size is one of the big issues in this **?** style random coefficient model, because the potential market size is crucial for the market share of outside option. As **?** dealt with this problem and **?** suggested, the most common way of setting the potential market size is to use the number of households in the market. However, in this study, I allow the households to choose more than one alternatives. Thus, I set the potential market share is the sum of the number of single households, $14,457,000$, and the doubled number of non-single households, $2 * 34,606,000$.

4.2.2 Decision Periods for Households

I assume that the households in *KHPS* do not discard their automobiles within two years of their purchase. Assuming this, I can increase the number of samples which I can use for this empirical analysis. There are some possibilities that they discard their automobiles within two years of their purchase. However, when I utilize 2005 dataset of *KHPS*, it shows that the duration of new car is about seven years and the frequency of discarding their cars before three years is very low.

4.2.3 Total Sales for Secondary Market

As I discuss before, the total sales for secondary market is not available in Japan. Thus, I need to predict the total sales numbers based on *Automobile Ownership*

Statistics and *KHPS*. The first and third column of Table **??** suggest that *KHPS* mimics *Automobile Ownership Statistics* data quite nicely in the sense that the ratio of *Kei*- and Noraml-cars is the almost same. Thus, I believe that the ratio of New car and Used car within each group (*Kei*- and Normal-cars) in *KHPS* is the same in Macro data. That is to say, the number of used Normal-car can be calculated by $3,894,000 \times \frac{226}{320} =$. Similarly, the number of used Kei-car can be derived as $1,554,000 \times \frac{81}{117} = 1,062,000$. For 2002, I also use the same strategy for predicting the sales volume in secondary market, and the results are summarized in Table **??**.

Table 7: Derivation of Used Automobiles Sales from Micro Data in 2003

	Micro		Macro	
	New	Used	New	Used
Normal	**320**	**226**	**3,894,000**	*2,629,000*
	(73.2%)	(73.6%)	(71.5%)	
Kei	**117**	**81**	**1,554,000**	*1,062,000*
	(26.8%)	(26.4%)	(28.5%)	

Table 8: Derivation of Used Automobiles Sales from Micro Data in 2002

	Micro		Macro	
	New	Used	New	Used
Normal	**313**	**196**	**3,874,000**	*2,364,000*
	(75.2%)	(77.2%)	(71.1%)	
Kei	**103**	**58**	**1,572,000**	*884,000*
	(24.8%)	(22.8%)	(28.9%)	

4.2.4 Predicting Correlation between Household Characteristics

When I integrate out households' characteristics from choice probabilities which are given by equation (**??**), (**??**) and (**??**), I need to take into account the correlation between household's characteristics. For instance, if the age of household head is less than 25 years old, then they are not likely to have kids.

In this study, I use the following the eleven household characteristics in total; for evaluation for product characteristics, z_i^p, I use family size, sex of the household head, age of household head, log of income, and for complementarity term, z_i^c, I use number of worker in the household, a dummy for having kids or not, a dummy for having old people or not, dummy whether wife works or not, distance to the nearest station, living city's population dummy. Unfortunately, Census data does not give us these correlations between households characteristics, I need to predict some of them based on the micro data, *KHPS*.

For single households, I only use (1) family size, which should be equal 1 by definition, (2) the sex of household head, (3) the age of household head, and (4) income. Census data provides the correlation between (2) and (3) as it is described in Table in Appendix, and thus there is no problem. Since for these household, income data is not available. Thus, I make a prediction for their income by their characteristics, i.e., sex, age and age squared. Here, I utilize the micro data for picking up coefficients for these characteristics.

For non-single households, according to *KHPS*, there is no significant correlation between the household size and living place population, or distance to the nearest station. Thus, I assume that I assume that there does not exist correlation between them. However, other characteristics are highly correlated with the household size, so using *KHPS*, I generate the household characteristics. As Table comparison shows, however, the frequencies of each household size in *KHPS* and Census data are quite different. Thus, I need to adjust each prediction by taking frequency given each household size. For example, when I derive the correlation between the household size and number of worker in the household, first I look at the frequency of the number of worker in *KHPS* given household size.

Table 9: Comparison of Consumer Samples

	Micro(*KHPS*)		Macro(Census 2005)	
	#	freq.	# (1,000)	freq.
# of Household Members				
1	296	(7.4%)	14,457	(29.5%)
2	671	(16.8%)	13,024	(26.5%)
3	833	(20.8%)	9,196	(18.7%)
4	1,147	(28.6%)	7,707	(15.7%)
5	562	(14.0%)	2,848	(5.8%)
6+	496	(12.4%)	1,831	(3.7%)
Total	4005		49,063	
Mean Age of Household Head	51.5			
Having Kids	1,698	(42.4%)	18,758	(38.2%)
Having Olds	1,364	(34.1%)	18,562	(39.4%)
Living Area				
14 Biggest Cities	945	(23.6%)	29,413	(23.0%)
Other Cities	2,280	(56.9%)	80,851	(63.3%)
Villages	780	(19.5%)	17,504	(13.7%)

5 Results

5.1 The Estimates and Model Fits

Table **??** displays the method of simulated moments (MSM) estimates of coefficients for product characteristics. I would like to emphasize three points about these results. First of all, family size is an important factor for deciding the capacity. It is quite natural that households of a larger size are more likely to have larger capacity automobiles. However, since these estimates also reflect the evaluation for the secondary car, it is not as large as I expected. Secondly, the sex of the household head also plays an interesting role in choosing automobiles. The dummy variable for the sex of a household head takes 0 if it is a woman, and 1 otherwise. Thus, a household is more likely to choose a Normal car (or a large automobile) if the houseold head is a man. My reasoning is as follows; if a household head is a woman, then that household is most likely to consist of a single mother and her children. Thus, it does not need to have a large capacity automobile. Finally, the age of the household head is a positive effect for purchasing larger automobile. This matches the fact that older people tend to choose larger automobiles, while younger generations tend to choose smaller size automobiles.

Table 10: Estimation Result 1 - Evaluation for Product Characteristics

	Capacity		Fuel Efficiency		Displacement	
	Est.	S.E.	Est.	S.E.	Est.	S.E.
β^u (variance)	0.3312	(0.0523)	0.0285	(0.0036)	0.0128	(0.0013)
β_1^o (family size)	0.1699	(0.0079)	0.0813	(0.0122)	-0.0094	(0.0038)
β_2^o (sex of HH head)	0.2258	(0.0640)	0.0496	(0.0153)	0.0646	(0.0018)
β_3^o (age of HH head)	0.0088	(0.0047)	0.0049	(0.0004)	0.0024	(0.0007)

Table **??** summarises the results for complementarity term. First of all, as the number of workers in a household increases, they shift to purchase two Normal-cars. This fact suggests that people who commute with their cars prefer Normal-cars to *Kei*-cars, because of the margin of safety. The second point is that kids' dummy in Γ_{NN} is negative. This presumably reflects the fact that having children implies having less workers in a household. Thus, these households are less likely to have two Normal cars. Thirdly, the coefficient for the work wife dummy in Γ_{KN} is substantial. That is to say, even though they have a job that might need them to commute by their cars, they prefer *Kei*-cars to Normal-cars. Finally, the households that live in a village area or city which is not among the 14 biggest cities in Japan are more likely to have second cars, and village area residents tends towards choosing *Kei*-cars.

Table 11: Estimation Result 2 - Compelmentarities/Substitutabilities Term

	Γ_{KK}		Γ_{KN}		Γ_{NN}	
	Est.	S.E.	Est.	S.E.	Est.	S.E.
γ_r^1 (constant)	0.0098	(0.0007)	0.0345	(0.0019)	0.1061	(0.0117)
γ_r^2 (# of worker)	0.0042	(0.0017)	0.0083	(0.0027)	0.2718	(0.0137)
γ_r^3 (kids dummy)	0.0764	(0.0053)	0.1124	(0.0072)	-0.2707	(0.0199)
γ_r^4 (olds dummy)	0.0099	(0.0056)	0.0098	(0.0013)	0.0021	(0.0009)
γ_r^5 (work wife)	0.0124	(0.0010)	0.0735	(0.0038)	0.0011	(0.0003)
γ_r^6 (nearest sta.)	0.0119	(0.0012)	0.0186	(0.0029)	0.0410	(0.0410)
γ_r^7 (not big city)	0.0001	(0.00005)	0.0145	(0.0025)	0.0677	(0.0238)
γ_r^8 (village)	0.0039	(0.0006)	0.1441	(0.0075)	0.0329	(0.0023)

5.2 Model Fit

Table **??** shows the model fit. The numbers in the model prediction are derived by multiplying the probability distribution for each choice and the number of observations, given each state. From this table, although it seems that I underestimate the evaluation for the product characteristics, this model mimics the patterns in the actual data quite well. In particular, this model predicts the purchasing behavior of family households with no car pretty well.

Table 12: Model Fit - Household Choice of Data and Prediction (Micro)

HH Size and State	# of ob.		Choice					
			buy nothing	buy K	buy N	buy K,K	buy (K,N)	buy (N,N)
Family (0,0)	1075	Data	458	105	424	9	29	50
		Pred.	**504**	**91**	**401**	**8**	**28**	**43**
Family (K,0)	182	Data	106	11	65			
		Pred.	**117**	**8**	**57**			
Family (N,0)	1248	Data	966	96	186			
		Pred.	**1004**	**82**	**162**			
Single (0)	206	Data	159	10	37			
		Pred.	**166**	**9**	**30**			

6 Conclusion and Future Direction

In this empirical study, I present the framework of how to estimate a random coefficient model with complementarities/substitutabilities term, and show that there exist complementarities among automobile consumptions, by using the Japanese automobile ownership data. In particular, I find there the three driving forces of having one *Kei*-car and one Normal-car: whether the wife works or not, whether there are children or not, and where the family lives.

As a next step, I would like to use these results for assessing the welfare effect of hypothetical changes to automobile tax. The Japanese government has implemented tax advantages for purchasing and owning *Kei*-cars, and now there is a discussion whether the government should abolish these advantages or not. Even though several Japanese manufactures of *Kei*-cars insist that the demand for *Kei*-cars would dramatically decrease if these advantages werer abolished, the results of this empirical study imply that the demand would not drecrease very sharply. Thus, in order to measure the welfare effect accurately, I would like to estimate the full model as the first step.

To that purpose, it is also important to extend this model by taking into account the dynamic aspects of household behavior.[16] However such an extension would bring with it new types of difficulties too, and thus it remains a possible for future research.

Appendix

A3: List of Automobiles' Characteristics

The following Table **??** summaries the available characteristics of automobiles from Automotive Guidebook.

Table 13: Available Characteristics of Automobiles

Exterior dimensions L, W, H (mm)	Wheel base (mm)	Max torque (kgm/p)
Interior dimensions L, W, H (mm)	Treads(F/R) (mm)	Fuel system
Seating capacity (person)	Ground clearance (mm)	Fuel tank capacity
Gross vehicle weight (kg)	Curb vehicle weight (kg)	Transmission
Fuel consumption at 60km/h (km/l)	Min. turing radius (m)	Drive train
Fuel consumption 10-15 modes (km/l)	Engine model	Suspension system
Max. power (net) (kW/rpm)	Cylinders	Braking system
Displacement (cc)	Bore $\times$ Stroke (mm)	Tire size
Price (in Tokyo area)	Compression ratio	

[16]Estimating durable differentiated goods demand function in a dynamic setting has been attracted to many empirical researchers, recently. See **?** and **?**.