GRIPS Discussion Paper 15-08

Discrete choice models for multicategory goods

Yukihiro Kidokoro

June 2015

National Graduate Institute for Policy Studies
7-22-1 Roppongi, Minato-ku,
Tokyo, Japan 106-8677

Discrete choice models for multicategory goods

Yukihiro Kidokoro[1]

National Graduate Institute for Policy Studies

7-22-1 Roppongi, Minato-ku, Tokyo, 106-8677, Japan

E-mail address: kidokoro@grips.ac.jp

Abstract:

Although discrete choice models are well suited to describing the demand structure of differentiated goods, two important problems remain unsolved in their application. First, the total demand for a choice set is exogenously fixed. Second, multiple categories of goods cannot be handled in an unrestrictive way. In this paper, we address these flaws by formulating a complete utility maximization problem that is consistent with discrete choice models and derive the implications for applied research. We then apply the results to the ketchup and mayonnaise markets and investigate the differences arising from the consideration of multiple categories of goods.

Keywords: Discrete choice, multiple categories, endogenous demand, product differentiation, logit

JEL classification code: C35, D11, D12

[1] I would like to thank Charles Blackorby, Jan K. Brueckner, David Donaldson, Chiaki Hara, Tatsuo Hatta, Yoshitsugu Kanemoto, Se-il Mun, Kenneth A. Small, and Jacques-François Thisse for helpful suggestions and discussions. I am solely responsible for any remaining errors and omissions. This work was supported by JSPS KAKENHI Grant Number 26590037.

1. Introduction

Discrete choice models are well suited to describing the demand structure of differentiated goods. However, two important problems remain unsolved. First, the total demand for a choice set is exogenously fixed. For instance, consider the choice of ketchup brands A and B. In conventional discrete choice models, given the assumption that a consumer selects only one alternative, a consumer selects “brand A,” “brand B,” or “no purchase”. Such a model can then readily handle a change in the demand for each brand of ketchup, but provides no explanation as to why a consumer selects only one alternative. This is a serious flaw of discrete choice models. In fact, McFadden (1999: p. 273) and Nevo (2000: fn. 14; 2001: fn. 13, 2011) have repeatedly pointed out this limitation.[3] This problem is especially serious given the application of discrete choice models to daily consumables, including ready-to-eat cereal (Nevo, 2001) and canned tuna (Nevo and Hatzitaskos, 2006), whose demand is more variable than that for durable goods such as housing (Earnhart, 2002) and automobiles (Berry et al., 1995; Goldberg, 1995; Petrin, 2002).

Second, conventional discrete choice models have a limited ability to describe the unrestricted substitutionary or complementary relationship between goods. In relation to the first abovementioned flaw of discrete choice models, each

[3] Nevo (2001: fn. 13) states: “A comment is in place about the realism of the assumption that consumers choose no more than one brand. Many households buy more than one brand of cereal in each supermarket trip but most people consume only one brand of cereal at a time, which is the relevant fact for this modeling assumption. Nevertheless, if one is still unwilling to accept that this is a negligible phenomenon, this model can be viewed as an approximation to the true choice model.” Nevo (2000: fn. 14, 2011) makes a similar point. McFadden (1999: p. 273) likewise raises the possibility that an alternative can be interpreted “…as a ‘portfolio’ of decisions made in sequence, or as one of the multiple decisions.”

alternative is necessarily a substitute because of assumed unitary choice. This substitutionary relationship between alternatives continues to hold for the case of multiple categories of goods.[4] For example, while the nested logit model is capable of representing multiple categories of goods, it has some limitations: i) the relationship between each category is the logit, and ii) the total demand for a choice set is unity. In reality, a consumer may consume multiple units of multiple brands of multiple categories. For instance, one unit of brand A ketchup, two units of brand B ketchup, two units of brand C mayonnaise, and one unit of brand D mayonnaise. Accordingly, a decrease in the price of brand A ketchup would affect the demand for brands C and D mayonnaise, as well as the demand for brands A and B ketchup, and therefore change the total demand for both ketchup and mayonnaise. However, standard discrete choice models cannot generally represent such a relationship.

The purpose of this paper is to: i) develop a model resolving these flaws of discrete choice models, ii) examine the properties the derived model implies, and iii) apply the model to an actual situation in which a consumer selects many goods from multiple categories. First, we formulate the consumer's deterministic utility maximization problem, which yields a consistent result with the discrete choice models. In other words, we reformulate the discrete choice models using a standard deterministic utility maximization framework to allow for a more general

[4] Gentzkow (2007) analyzes the complementary relationship between the paper and online versions of a newspaper by including the choice of "buying both paper and online versions of newspaper." In this setup, each alternative is a substitute, but the paper version and the online version of the newspaper can be complementary. This also implies that the maximum demand for the newspaper is unity. However, while the assumption of unitary demand may appear innocuous with a newspaper, we cannot generally extend this to other goods where the consumer can purchase multiple units.

demand structure between the various categories of goods. We begin our analysis using a simple logit model, but then extend it to more general discrete choice models, such as the generalized extreme value (GEV) and mixed logit models.[5]

Second, we derive the elasticities of the demand functions implied by our model and the method for the calculation of welfare change. We can decompose these elasticities into the elasticities for each category of goods and those of the choice probability. The former contribute to describing the complementary relationship between goods, even though they are disregarded in conventional discrete choice models of unitary demand for a choice set. For the calculation of welfare change, we extend the results in Small and Rosen (1981) to allow for the variable total demand of goods and multiple categories of goods.

Finally, to illustrate the application of our analysis, we estimate the demand functions for the ketchup and mayonnaise markets, examine their characteristics, and calculate the welfare change under a hypothetical scenario using point-of-sale (POS) data from supermarket checkouts. The estimated results suggest that the goods in both markets are complementary and not well represented by conventional discrete choice models designed for a single category of goods with fixed unitary demand for a choice set.

Before proceeding, we briefly relate our paper to the broader literature. First, the analysis in this paper relates to Anderson et al. (1988, 1992: Ch. 3) and Verboven (1996), both of which formulate a utility maximization problem consistent with discrete choice models. The former derives a direct utility function that corresponds to the logit model, whereas the latter does so for the nested logit model. Our theoretical analysis differs from these studies in two respects. The first is that our framework can incorporate multiple categories of goods, and

[5] Our analysis is quite general, in that Dagsvik (1995) and McFadden and Train (2000) show that these models can approximate any random utility model.

accordingly, easily applies to empirical demand analyses, as we illustrate using actual ketchup and mayonnaise data. The second is that we formulate a utility maximization problem that corresponds to the GEV and mixed GEV models, which are more general models than the logit and nested logit models.

Second, our analysis also relates to the so-called “discrete/continuous” models in the literature, which enables us to select multiple goods within the framework of discrete choice models (Dubin and McFadden (1984), Hanemann (1984), Hendel (1999), Dube (2004), and Bhat (2005, 2008)). However, these studies do not consider multiple categories of goods. Importantly, the analysis in this paper can consider multiple categories of goods without assuming an a priori substitutionary or complementary relationship between categories. Song and Chintagunta (2007) and Mehta and Ma (2012) apply the approach in Hanemann (1984) to a choice across multiple categories of goods. However, in both of these analyses, each brand in a category is a perfect substitute, such that by assumption, the choice is of only one brand. Our analysis is free of this restriction in that a consumer may select as many brands as they like from within each category, and at the same time, may consume as many units of each brand as they like.

Third, an effective way to describe a system of demand is to apply the Almost Ideal Demand System (AIDS) developed by Deaton and Muellbauer (1980), whereby we consider AIDS as the first-order approximation of any demand model. The theoretical model derived in this paper is consistent with the demand models from logit, GEV, and mixed GEV models. Thus, our approach is of a narrower fit than AIDS from the viewpoint of functional form. However, this approach has a clear merit in that our model can better represent product differentiation with fewer parameters to estimate, and is more easily applied to actual data. Nonetheless, as with AIDS, our model can handle any number of categories, address arbitrary relationships between categories, and enable unrestricted choice within a category (multiple choices of brands and multiple choices of the units of each brand).

The remainder of the paper is structured as follows. In Section 2, as the simplest case, we formulate a utility maximization problem in which a consumer chooses goods from one category of differentiated goods, the result for which is consistent with the logit model. We then examine the elasticities of the derived demand functions and the corresponding method for welfare estimation. In Section 3, we generalize the analyses in Section 2 to accommodate multiple categories of differentiated goods, while maintaining focus on the logit model for a consumer. In Section 4, we consider an additional restriction reflecting the actual situation of multiple consumers. Section 5 extends our analysis to the GEV and mixed GEV models. Section 6 empirically applies our framework to an analysis of the markets for ketchup and mayonnaise. Section 7 concludes the paper.

2. The logit model with flexible total choice for a single category

For simplicity, we begin by focusing on the simplest case in which a single consumer selects goods from one category of differentiated goods, with the selection results represented by the logit model. We consider the consumer's "complete" utility maximization problem in which the consumer selects as many goods as they like without fixing the total demand for a choice set or including a "no purchase" alternative, and yet yields results consistent with the logit model. Note that a standard approach in discrete choice models considers "partial" utility maximization where, by assumption, the consumer selects only one alternative (including the "no purchase" alternative) given that the total demand for a choice set is fixed at unity. Appendix A lists the variables and parameters.

2.1. Indirect utility function

Our purpose here is to construct a utility maximization model, which yields a demand function for differentiated good j that is consistent with the logit model

(1) $\quad x_j(p_1,...,p_M,y) = x_g(p_1,...,p_M,y)s_j(p_1,...,p_M,y)$,

where $x_g(p_1,...,p_M,y)$ is the total demand for a category and $s_j(p_1,...,p_M,y)$ is the choice probability of the logit model. Thus, the demand function for differentiated good j, (1), has the form of the logit model, but endogenous total demand for a category. We place two assumptions on (1). The first is that the total demand for a category, x_g, depends only on prices through the aggregated price index, $P(p_1,...,p_M,y)$. This property is useful for the actual estimation of the system of demand functions, where we can separate the total demand for a category and the choice probability. The second assumption is that the choice probability of differentiated good j is a decreasing function of price. The conventional logit model satisfies this property. Under these two assumptions, (1) is modified to

(2) $\quad x_j(p_1,...,p_M,y) = x_g(P(p_1,...,p_M,y),y)s_j(p_1,...,p_M,y)$,

where $\dfrac{\partial s_j}{\partial p_j} < 0$.

The necessary and sufficient conditions for the demand function for differentiated good j to have the form of (2) is that the indirect utility function has the form of

(3) $\quad v(p_1,...,p_M,y) = v(P(p_1,...,p_M,y),y)$,

where

(4) $\quad P(p_1,...,p_M,y) \equiv \sum_{j'=1}^{M} \exp(\alpha_{j'}(y) - \beta(y)p_{j'})$ $\;(\beta(y) > 0)$.

We easily check the sufficiency of this result using Roy's identity. Applying Roy's identity to (3), we have

$$(5) \quad x_g(P(p_1,...,p_M,y),y) = \frac{\beta(y)\dfrac{\partial v}{\partial P}P}{\dfrac{\partial v}{\partial y}} \text{ and}$$

$$(6) \quad s_j(p_1,...,p_M,y) = \frac{\exp(\alpha_j(y) - \beta(y)p_j)}{\sum_{j'=1}^{M}\exp(\alpha_{j'}(y) - \beta(y)p_{j'})},$$

which are consistent with (2). In (5), the total demand for a category, $x_g(P(p_1,...,p_M,y),y)$, is nonnegative from $\frac{\partial v}{\partial p_j} \le 0$ and $\frac{\partial v}{\partial y} > 0$, which follow from the properties of the indirect utility function. We prove the necessity by solving the system of differential equations obtained from (2). See Appendix B for details.

Three points warrant further explanation. First, a consumer may consume as many goods as they like, because we include no assumption regarding the total demand for a choice set, contrary to the standard approach where it is fixed at unity. A consumer may also consume multiple brands of differentiated goods within the one category because the demand for differentiated good j is the total demand for a category multiplied by the choice probability of good j. The case of a corner solution is included, such that when $\alpha_j(y) - \beta_j(y)p_j$ approaches negative infinity, the choice probability of good j is zero.

Second, the argument for each exponential function in the aggregated price index, (4), must be: i) linear in price and ii) have the same coefficient for price, so that we derive the logit demand representation, (2), in which the total demand for a category depends only on the aggregated price index, (4). When the first condition is unmet, we cannot obtain the common total demand for a category, (5), that is, the logit demand representation, (2), no longer holds. When the second condition is unmet, we lose the property that the total demand for a category, (5), depends only on the aggregated price index.

Third, we may be interested in the form of the direct utility function that corresponds to the indirect utility function, (3). Unfortunately, we cannot derive a direct utility function that perfectly corresponds to the indirect utility function, (3), in closed form. However, applying Anderson et al. (1988, 1992: Ch. 3), we can provide an example of the direct utility function:

$$(7)\quad u = z + \psi\left(\frac{a}{\beta}\sum_{j'=1}^{M} x_{j'}\right) + \frac{1}{\beta}\sum_{j'=1}^{M}\left[\alpha_{j'} - \ln\left(\frac{x_{j'}}{\sum_{j'=1}^{M} x_{j'}}\right)\right]x_{j'},$$

where a is a parameter and β does not depend on y.[6] This direct utility function yields the indirect utility function of

$$(8)\quad v = \psi\left(\psi'^{-1}\left(-\frac{\ln P}{a}\right)\right) + \frac{\ln P}{a}\psi'^{-1}\left(-\frac{\ln P}{a}\right) + y,$$

which is a special case of (3).

2.2. Elasticities

We focus here on the properties of the demand functions by deriving the price elasticities. From (2), (5), and (6), the respective own- and cross-price elasticities are

[6] Eq. (7) is a modified direct utility function of Eq. (3.26) in Anderson et al. (1992: Ch. 3). Their direct utility function is not completely consistent with the demand function for the logit model because the choice probability depends on the demand function for all differentiated goods.

(9) $$\frac{\partial x_j}{\partial p_j}\frac{p_j}{x_j} = \theta_j - \beta(y)(1-s_j)p_j$$

and

(10) $$\frac{\partial x_j}{\partial p_k}\frac{p_k}{x_j} = \theta_k + \beta(y)s_k p_k ,$$

where $k = 1,\ldots,M$, $k \neq j$, and $\theta_j \equiv \frac{\partial x_g(P(p_1,\ldots,p_M,y),y)}{\partial p_j}\frac{p_j}{x_g(P(p_1,\ldots,p_M,y),y)}$ is the price elasticity of the consumer's total demand for a choice set.

In this regard, three points warrant explanation. First, the own- and cross-price elasticities have a common structure for the elasticity of the consumer's total demand for a choice set plus that of the choice probability, as pointed out by Taplin (1982) and Oum et al. (1992) in the context of transport demand.[7] In the ordinary logit model where the total demand for a choice set is fixed, we disregard the elasticity of total demand. Consequently, the elasticity derived is not the elasticity of demand, but rather the elasticity of the choice probability. Thus, it is highly probable that the elasticity derived in the logit model with fixed total demand for a choice set differs significantly from the true elasticity when the change in the total demand for a choice set is large.

Second, the cross-price elasticity of (10) demonstrates that the two differentiated goods can be gross complements as Anderson et al. (1992, Eq. 3.41) suggests, although they are necessarily gross substitutes when the total demand for a choice set is fixed.[8] Recall that in the logit model with fixed total demand

[7] In Taplin (1982) and Oum et al. (1992), the elasticity of the consumer's total demand for the category of differentiated goods and that of the choice probability correspond to the elasticity of demand for aggregate traffic and the mode choice elasticity, respectively.

[8] See Mas-Colell et al. (1995: p. 70) for the definition of gross substitutes and complements.

for a choice set, a decrease in the price of brand A necessarily reduces the demand for brand B. That is, all differentiated goods must be gross substitutes. However, this relationship breaks down if we consider the change in total demand. In that case, a decrease in the price of brand A has the effect of increasing the total demand for a choice set as well as diverting demand from brand B to brand A. If the former effect outweighs the latter, the demand for brand B increases.

Third, the above two points suggest problems for a typical logit model with variable total demand, which makes the demand for each good variable by introducing a "no purchase" alternative within the fixed total demand for a choice set. One problem is that the share of the "no purchase" alternative tends to be very high, at least over a short period if, as Nevo (2001) suggests, the total demand for a choice set is fixed at the number of residents. This is because a consumer is less likely to purchase a category of goods when the period in question is shorter. This feature is stronger for more durable and expensive goods like cars and houses. In such a case, the share of each good relative to the number of residents becomes very small, which makes the own- and cross-price elasticities of choice probability, $-\beta(y)(1-s_j)p_j$ and $\beta(y)s_k p_k$ in (9) and (10), very large and very small, respectively. Unfortunately, we cannot determine whether the too-large share of the "no purchase" alternative will over- or underestimate the true elasticities because the true elasticities include the elasticities of the consumer's total demand for a choice set, θ_j and θ_k. The other point is that if the total demand for a category varies because of the "no purchase" alternative, the cross-price elasticities are always positive because the elasticity of total demand for a choice set is zero. Accordingly, this setup is also unable to represent complementary relationships between differentiated goods.

2.3. Welfare analysis

We now consider the method for welfare analysis. Hereafter, the superscripts WO and W denote without and with policy. The indirect utility function, (3), is perfectly compatible with standard deterministic utility maximization. Thus, the compensating variation is the sum of the areas to the left of the compensated demand functions, as is usual. That is, the compensating variation, cv, can be calculated by

$$
\begin{aligned}
cv &= \int_{p_j^W}^{p_j^{WO}} \left(\sum_{j'=1}^{M} \frac{\partial e(p_1,\dots,p_M, v^{WO})}{\partial p_{j'}} \frac{\partial p_{j'}}{\partial p_j} \right) dp_j \\
&= \int_{p_j^W}^{p_j^{WO}} \left(\sum_{j'=1}^{M} x_{j'}^c(p_1,\dots,p_M, v^{WO}) \frac{\partial p_{j'}}{\partial p_j} \right) dp_j ,
\end{aligned}
\tag{11}
$$

where $e(p_1,\dots,p_M,v)$ is the expenditure function, which is derived from the indirect utility function of (3), and x_j^c is the compensated demand function for good j.[9] In (11), we consider the possibility that a change in the price of differentiated good j affects the price of differentiated good j'.

An interesting question is whether we can calculate the compensating variation only from the change in the aggregated price index, (4), utilizing the property that the total demand for a category depends solely on the aggregated price index. Generally, the aggregated price index is a function of income as well as prices, and thus we cannot represent the expenditure function as a function of the aggregated price index. This means that we generally cannot derive the compensating variation given the change in the aggregated price index. However, in a special case where the aggregated price index is independent of income, we

[9] The same result applies to the equivalent variation (ev) if v^{WO} is substituted for v^W.

can calculate the compensating variation with the change in the aggregated price index.

Suppose that $\alpha_j(y) = \alpha_j$ and $\beta(y) = \beta$ hold in the aggregated price index, (4), that is,

$$(12)\quad P(p_1,..., p_M, y) = P(p_1,..., p_M) = \sum_{j'=1}^{M} \exp(\alpha_{j'} - \beta p_{j'}) .$$

In this case, we obtain the compensated demand function for differentiated good j:

$$(13)\quad x_j^c(p_1,..., p_M, v) = x_g^c(P(p_1,..., p_M), v) s_j(p_1,..., p_M, y) ,$$

which has the form of the compensated total demand for a category multiplied by the choice probability. In (13),

$$(14)\quad x_g^c(P(p_1,..., p_M), v) = (-\beta) \frac{\partial e(P(p_1,..., p_M), v^{WO})}{\partial P} P \geq 0 ,$$

where the nonnegativity of $x_g^c(P(p_1,..., p_M), v)$ is derived from the properties of the indirect utility function of $\frac{\partial v}{\partial p_j} \leq 0$ and $\frac{\partial v}{\partial y} > 0$.

Under the assumption for (12), we have

$$(15)\quad cv = -\frac{1}{\beta} \int_{P^W}^{P^{WO}} \left(\frac{x_g^c(P(p_1,..., p_M), v)}{P} \right) dP ,$$

which demonstrates that the compensating variation can be derived from the compensated total demand function for a category using the change in the aggregated price index. Moreover, if we define the log-sum variable as

$$(16)\quad LS(p_1,..., p_M, y) \equiv \ln P(p_1,..., p_M, y) = \ln \sum_{j'=1}^{M} \exp(\alpha_{j'}(y) - \beta(y) p_{j'}) ,$$

Eq. (15) is rewritten as

$$(17)\quad cv = -\frac{1}{\beta} \int_{LS^W}^{LS^{WO}} x_g^c(LS(p_1,..., p_M), v) dLS .$$

This result is an extension of Small and Rosen (1981), which demonstrates that the compensating variation is the change in the log-sum term multiplied by fixed total demand with exogenously fixed total demand. The case of fixed total demand corresponds to a limiting case of (17) in which the total demand function is vertical. Although a standard interpretation of the log-sum term is as a measure of expected consumer surplus,[10] (17) demonstrates that it is more accurately the aggregated price index, including the possibility of a change in the total demand for a choice set.

3. The logit model with flexible total choice for multiple categories

Given the analysis in Section 2, it is now easy to formulate a utility maximization problem corresponding to a logit model including multiple categories of goods. In what follows, we provide the form of the corresponding indirect utility function, the elasticities, and the method for welfare estimation.

3.1. Indirect utility function

Suppose that differentiated good j in category h ($h = 1,...,H$) is represented by the demand function for category h, $x_{hg}(P_1(\mathbf{p}_1, y),..., P_H(\mathbf{p}_H, y), y)$, multiplied by the logit choice probability within category h, $s_{hj}(\mathbf{p}_h, y)$, that is,

(18) $x_{hj}(p_1,..., p_M, y) = x_{hg}(P_1(\mathbf{p}_1, y),..., P_H(\mathbf{p}_H, y), y)s_{hj}(\mathbf{p}_h, y)$,

where $\mathbf{p}_h$ is the price vector for the differentiated goods in category h and $P_h(\mathbf{p}_h, y)$ is the aggregated price index for category h. As in Section 2, we

[10] See, for example, Train (2009). Domencich and Macfadden (1975) and Ben-Akiva and Lerman (1985) respectively interpret the log-sum term as an inclusive value and the measure of accessibility in relation to transport demand modeling.

restrict our attention to the form of the total demand for category h that depends only on the aggregated price indices for category h, $P_h(\mathbf{p}_h, y)$.

The necessary and sufficient conditions for the demand function for differentiated good j to have the form of (18) is that the indirect utility function has the form of

(19) $v(p_1,..., p_M, y) = v(P_1(\mathbf{p}_1, y),..., P_H(\mathbf{p}_H, y), y)$,

where

(20) $P_h(\mathbf{p}_h) \equiv \sum_{k \in h} \exp(\alpha_{hk}(y) - \beta_h(y) p_{hk})$ ($\beta_h(y) > 0$).

For sufficiency, applying Roy's identity to (19) is sufficient to derive

(21) $x_{hg}(P_1(\mathbf{p}_1, y),..., P_H(\mathbf{p}_H, y), y) = \dfrac{\beta_h(y) \dfrac{\partial v}{\partial P_h} P_h}{\dfrac{\partial v}{\partial y}} \geq 0$ [11] and

(22) $s_{hj}(\mathbf{p}_h, y) = \dfrac{\exp(\alpha_{hj}(y) - \beta_h(y) p_{hj})}{\sum_{j' \in h} \exp(\alpha_{hj'}(y) - \beta_h(y) p_{hj'})}$,

which are consistent with (18). The proof of necessity is a simple extension of that in Section 2, and thus omitted.

The indirect utility function, (19), is an extension of (3), taking into account multiple categories of goods. Now that the total demand function for category h depends not only on the aggregated price index for category h but also on those for any other categories in an unrestrictive way, we can represent any substitutionary or complementary relationships between the categories. This feature is a major advantage of our framework because, as we will see in Section

[11] The nonnegativity of the total demand function for category h follows from the properties of the indirect utility function, as in Section 2.

6, it enables us to analyze the relationship between categories without any premises. This contrasts with the standard analysis for multiple categories of goods using the nested logit model, for instance, in which the relationship between categories is restricted to the logit model.

3.2. Elasticities

We summarize how we modify the results for the elasticities in Section 2.2. The differences from the elasticities in Section 2.2 are: i) the price elasticities of the total demand for category h and the choice probability for the category within category h differ by category, and ii) the cross-price elasticity for the goods within the same category differs from that across categories.

First, in terms of the own-price elasticity, we derive

$$(23)\quad \frac{\partial x_{hj}}{\partial p_{hj}} \frac{p_{hj}}{x_{hj}} = \theta_{hj} - \beta_h (1 - s_{hj}) p_{hj} ,$$

where $\theta_{hj} \equiv \dfrac{\partial x_{hg}(P_1(\mathbf{p}_1, y), ..., P_H(\mathbf{p}_H, y), y)}{\partial p_{hj}} \dfrac{p_{hj}}{x_{hg}(P_1(\mathbf{p}_1, y), ..., P_H(\mathbf{p}_H, y), y)}$.

This is a natural extension of (9), taking into account multiple categories. Second, when differentiated goods j and k are in the same category, the cross-elasticity is

$$(24)\quad \frac{\partial x_{hj}}{\partial p_{hk}} \frac{p_{hk}}{x_{hj}} = \theta_{hk} + \beta_h s_{hk} p_{hk} ,$$

which is the same as (10), except that the elasticity of the total demand for category h and the choice probability within category h differ by category. When the goods belong to different categories, that is, differentiated good j belongs to category h and differentiated good k belongs to category h' ($h' = 1, ..., H$ and $h' \neq h$), the corresponding cross-price elasticity is

(25) $\dfrac{\partial x_{hj}}{\partial p_{h'k}} \dfrac{p_{h'k}}{x_{hj}} = \theta_{hk}$,

which shows that the cross-price elasticity depends only on the price elasticity of the total demand for category h. This is because the choice probability of a differentiated good depends only on the prices of differentiated goods within the same category, and is unaffected by the prices of goods in other categories.

As we can see from (24) and (25), we can unrestrictedly represent a variety of relationships between the multiple categories using the elasticity of the total demand for category h, θ_{hk}. This feature stems from the fact that the total demand for category h, (21), depends on the aggregated price indices for all categories.

3.3. Welfare analysis

Regarding the calculation of the consumer's compensating variation, cv, we again obtain (11). This is because how we categorize the goods does not affect the method for welfare calculation if we analyze a change in welfare based on the compensated demand function for each good.

If the aggregate price index is independent of income, that is, if $\alpha_k(y) = \alpha_k$ and $\beta_h(y) = \beta_h$, the methods using the aggregated price indices and log-sum terms, (15) and (17), are modified to

(26) $cv = \int_{P_h^W(\mathbf{p}_h)}^{P_h^{WO}(\mathbf{p}_h)} \left(\sum_{h'=1}^{H} \left(-\dfrac{x_{h'g}^c(P_1(\mathbf{p}_1),...,P_H(\mathbf{p}_H),v^{WO})}{\beta_{h'} P_{h'}(\mathbf{p}_{h'})} \dfrac{\partial P_{h'}(\mathbf{p}_{h'})}{\partial P_h(\mathbf{p}_h)} \right) \right) dP_h(\mathbf{p}_h)$ and

(27) $cv = \int_{LS_h^W(\mathbf{p}_h)}^{LS_h^{WO}(\mathbf{p}_h)} \left(\sum_{h'=1}^{H} \left(-\dfrac{1}{\beta_h} x_{h'g}^c(LS_1(\mathbf{p}_1),...,LS_H(\mathbf{p}_H),v^{WO}) \dfrac{\partial LS_{h'}(\mathbf{p}_{h'})}{\partial LS_h(\mathbf{p}_h)} \right) \right) dLS_h(\mathbf{p}_h)$,

where

(28) $LS_h(\mathbf{p}_h) \equiv \ln P_h(\mathbf{p}_h) \equiv \ln \sum_{j' \in h} \exp(\alpha_{hj'}(y) - \beta_h(y) p_{j'})$.

Eq. (27) is particularly useful in application, because it demonstrates the relationship between the log-sum terms and compensating variation in the case of multiple categories of goods.

4. Aggregation among multiple consumers

We have thus far considered a single-consumer economy. However, most data are market data, which aggregate the behaviors of many heterogeneous individuals. In this section, we consider the application of our analysis to market data.

We need to aggregate the model from a single- to a multiple-consumer economy. As is well known, each consumer's indirect utility function must take the Gorman form for consistent aggregation.[12] That is, consumer i's indirect utility function has the form

(29) $v^i(p_1,...,p_M,y^i) = A^i(P_1(\mathbf{p}_1),...,P_H(\mathbf{p}_H)) + B(P_1(\mathbf{p}_1),...,P_H(\mathbf{p}_H))y^i$,

where superscript i denotes the index of a consumer. By aggregating (29) over i, the representative indirect utility function is

(30) $V = \sum_i v^i(p_1,...,p_M,y^i) = \sum_i A^i(P_1(\mathbf{p}_1),...,P_H(\mathbf{p}_H)) + B(P_1(\mathbf{p}_1),...,P_H(\mathbf{p}_H))Y$,

where $Y \equiv \sum_i y$.

The difference between the indirect utility function in a single-consumer economy, (19), and that in a multiple-consumer economy, (29), stems from the fact that the aggregated price index must be independent of income. This is necessary because the Gorman form requires that: i) the indirect utility function is linear in income, and ii) it has a common coefficient of income for all consumers.

[12] See Varian (1992) and Mas-Colell et al. (1995) for an explanation of the Gorman form.

The market demand function, which is derived from the indirect utility function of the representative consumer, (30), replaces the total demand for category h, $x_{hg}(P_1(\mathbf{p}_1, y),..., P_H(\mathbf{p}_H, y), y)$ in (21), with the total market demand for category h. Accordingly, we replace the price elasticities of the total demand for categories with those of the total market demand for the categories in (23)–(25).

Because the aggregated price index is independent of income according to the requirements of the Gorman form, each consumer's compensating variation can always be calculated using the aggregated price index or log-sum terms from (26) and (27). We then obtain the aggregated compensating variation by summing each consumer's compensating variation. Alternatively, we can calculate the aggregated compensating variation from (11), (26), and (27) if we replace the compensated demand for good j in (11), $x_{j'}^c(p_1,..., p_M, v^{WO})$, and the total compensated demand for category h in (26) and (27), $x_{h'g}^c(P_1(\mathbf{p}_1),..., P_H(\mathbf{p}_H), v^{WO})$, with their market demand counterparts.

5. Extensions to GEV and mixed logit models

We can generalize the analysis thus far to the GEV model and the mixed logit (or GEV) model. An important point to note is that while the GEV and mixed logit (or GEV) models can remove the independence of irrelevant alternatives (IIA) property, they are still incapable of describing the arbitrary relationship between multiple categories of goods if we retain the assumption of fixed total demand for a choice set. For instance, in the case of the nested logit model, which is a class of GEV model, the relationships between the multiple categories are limited to the logit model, and those within a category must be substitutes. In the mixed logit model, we can represent sophisticated substitutionary relationships free of IIA, but still cannot incorporate the complementary relationship between categories. Appendices C and D provide detailed analyses of the GEV and mixed logit (or GEV) models. Here, we summarize the main differences to the logit model.

First, in the case of the GEV model, we modify the aggregated price indices, log-sum terms, choice probabilities, and elasticities to be consistent with the GEV model. The own- and cross-price elasticities are more complex, but retain the property that the elasticities have the form of the sum of the elasticity for the category and the elasticity for choice probability. Regarding welfare estimation, the results for the logit model continue to hold if we replace the aggregated price indices and log-sum terms in the logit model with those in the GEV model.

Second, in the mixed logit model, each consumer's indirect utility function must take the quasilinear form, which is a special version of the Gorman form, for consistent aggregation in a multiple-consumer economy. This is because each consumer has different parameters, and therefore the aggregated price indices differ between consumers. The Gorman form indirect utility function requires that the coefficient of income is the same for all consumers, so it cannot include the aggregated price index as an argument, that is, it is a constant. This restriction implies that each consumer's indirect utility function, and thus the aggregated indirect utility function, takes the quasilinear form.

6. Application

In this section, we estimate a consumer demand model with multiple categories of goods using actual data and compare the results with those from the conventional logit model with the outside option of "no purchase." Our aim is to shed light on the substitutionary or complementary relationships between categories of goods, which is not possible using the conventional logit model. Thus, for the sake of simplicity, we use the logit model, disregarding the IIA problem.

The data are POS data collected by KSP-SP Co. Ltd from 686 Japanese supermarkets for ketchup and mayonnaise sold in December 2013. The data contain the sales price and quantity sold for each brand of ketchup and mayonnaise, the number of customers who pass through the checkout, and the prefecture where the store is located.

The reasons why we focus on ketchup and mayonnaise are as follows. First, both markets are oligopolistic and are well suited to analysis using the logit model. Selecting those brands with at least 3% of market share by revenue identifies six brands of ketchup and eight brands of mayonnaise, each group having 71% of market share by revenue in total. Appendix E lists the brands.[13] Second, in Japan, it is less likely that ketchup is a substitute for mayonnaise, and vice versa. Intuitively, as they are not substitutes, they do not warrant analysis using the conventional logit model, so they are suited to an analysis using the framework developed in this paper, which allows for complementarity among categories of goods. In fact, as shown later, a good in ketchup and a good in mayonnaise are gross complements.

Table 1 shows some descriptive statistics, including various statistics for prices and quantities counting, for each brand, only those stores selling that brand. The data exhibit two features. First, the number of markets differs significantly by brand. This reflects the fact that the product varieties of ketchup and mayonnaise that each store sells are quite different. Second, the unit market share of each brand becomes very small if our calculations include the outside option of "no purchase." This is because the quantity sold per brand is very small relative to the total number of customers passing through the checkout. We exclude the data for four stores that do not sell any of the six brands of ketchup that are included in the

[13] We treat different package sizes of a given product as separate goods.

study. Accordingly, the sample size is 682. We label ketchup as category 1 and mayonnaise as category 2. All estimations use Stata 11.

6.1. Estimation of the two logit models – one category of goods with the outside option and two categories of goods without the outside option

6.1.1. The logit model for one category of goods with the outside option

We begin by estimating the conventional logit model including the outside option. As usual in the literature, the utility of "no purchase" is normalized at zero. The demand for each brand is

$$(31)\quad x_{hj}^{r} = N^{r}\frac{\exp(\alpha_{hi} - \beta_{h} p_{hj}^{r})}{1+\sum_{j'\in h}\exp(\alpha_{hi} - \beta_{h} p_{hj'}^{r})},$$

where $r = 1,...,682$ is the store number and $h = 1,2$ is the index for the category. When $h = 1$, i.e., the category is ketchup, $j = 1,...,6$ is the index for the brands of ketchup. In the same way, when $h = 2$, i.e., the category is mayonnaise, $j = 1,...,8$ is the index for the brands of mayonnaise. N^{r} denotes the number of consumers who pass through the checkout at store r. Our estimation procedure follows Berry (1994), who "linearizes" the logit model to consider the possibility of endogeneity. We define the number of "no purchase" for category h at store r as x_{h0}^{r}, which is calculated as the difference between the number of consumers who pass through the checkout at store r, N^{r}, and the sum of quantities of category h sold by store r, $\sum_{j} x_{hj}^{r}$. Taking the natural log of the ratio of (31) to the number of "no purchase," we have

$$(32)\quad \log\frac{x_{hj}^{r}}{x_{h0}^{r}} = \alpha_{hi} - \beta_{h} p_{hj}^{r}$$

for both ketchup and mayonnaise.

In estimating (32), we need to take into account the potential endogeneity of p_{hj}^{r}. Given the limitations of our data, the only available instrument is the average price in other stores, as used in Hausman (1996) and Nevo (2001). We then estimate (32) for ketchup and mayonnaise with ordinary least squares (OLS) and two-stage least squares (2SLS), respectively, using the average price in the other stores as the instrument. Table 2 shows the results. As can be seen, all the estimated coefficients are significant at the 1% level, regardless of whether the estimation method is OLS or 2SLS. The result for endogeneity is ambiguous. We reject the null hypothesis that there is no difference in the OLS and 2SLS coefficients for ketchup, but not for mayonnaise. One reason could be that the assumption of Hausman (1996) and Nevo (2001) that the demand shocks are independent across stores might not hold here. This is because our data include several chain stores, where it is highly probable that promotions or advertisements are common to at least some portion of stores. This makes the assumption of independent demand shocks questionable, and therefore undermines the effectiveness of the instrument.

6.1.2. The logit model for two categories of goods without the outside option

We then estimate the logit model that is compatible with multiple categories of goods. Our task is to estimate the market demand function for each good that corresponds to the indirect utility function of the representative consumer, (30). To do this, we need to specify the functional form of (30). Natural candidates are the translog and generalized Leontief functions, both of which are less restrictive and widely used in applied research.

The translog indirect utility function of the representative consumer does not have a realistic property in our model and is represented by

(33)

$$V^r = Y^r + N^r\left[\sum_{h=1}^{2} a_h\left(\ln\sum_{j'\in h}\exp(\alpha_{hj'} - \beta_h p^r_{hj'})\right) + \frac{1}{2}b\prod_{h=1}^{2}\left(\ln\sum_{j'\in h}\exp(\alpha_{hj'} - \beta_h p^r_{hj'})\right)\right],$$

which yields the market demand function of

$$\text{(34)}\quad x^r_{hj} = N^r\left[\beta_h a_h + \frac{1}{2}\beta_h b\ln\sum_{j'\in h^{-1}}\exp(\alpha_{h^{-1}j'} - \beta_{h^{-1}} p^r_{h^{-1}j'})\right]\frac{\exp(\alpha_{hj} - \beta_h p^r_{hj})}{\sum_{j'\in h}\exp(\alpha_{hj'} - \beta_h p^r_{hj'})}.$$

Eq. (34) demonstrates that the demand for one category of goods, that is, the sum of the demand for each good in one category, depends only on the aggregated price index of the other category. For example, the total demand for ketchup is determined not by the aggregated price index of ketchup but by that of mayonnaise. This is why we consider it unrealistic to assume that the indirect utility function of the representative consumer is consistent with the translog form.

The generalized Leontief indirect utility function of the representative consumer is free of this defect and is modeled as

$$\text{(35)}\quad V^r = Y^r + N^r\left[\sum_{h=1}^{2} a_h\left(\sum_{j'\in h}\exp(\alpha_{hj'} - \beta_h p^r_{hj'})\right) + b\prod_{h=1}^{2}\left(\sum_{j'\in h}\exp(\alpha_{hj'} - \beta_h p^r_{hj'})\right)^{\frac{1}{2}}\right],$$

from which we derive the market demand function for each good as

(36)

$$x^r_{hj} = N^r\left[a_h\beta_h\left(\sum_{j'\in h}\exp(\alpha_{hj'} - \beta_h p^r_{hj'})\right) + \frac{1}{2}b\beta_h\left(\prod_{h=1}^{2}\left(\sum_{j'\in h}\exp(\alpha_{hj'} - \beta_h p^r_{hj'})\right)^{\frac{1}{2}}\right)\right]\frac{\exp(\alpha_{hj} - \beta_h p^r_{hj})}{\sum_{j'\in h}\exp(\alpha_{hj'} - \beta_h p^r_{hj'})}.$$

Eq. (36) demonstrates that the demand for a category of goods depends on the aggregated price indices of both categories, unlike demand using the translog form. Thus, we adopt the generalized Leontief form as the functional form of the indirect utility function of the representative consumer.

We implement the estimation in two steps. The first step is to estimate the choice probability of the logit model, given the demand for a category. The second step is to calculate the aggregated price indices, based on the parameters derived in the first step, and then estimate the demand for the categories using the aggregated price indices.

The first step is identical to the estimation of the logit model for one category of goods with the outside option in Section 6.1.1. In the same way as (32), we "linearize" the logit model to

$$(37)\quad \log \frac{x_{hj}^{r}}{x_{hj'}^{r}} = (\alpha_{hj} - \alpha_{hj'}) - \beta_h (p_{hj}^{r} - p_{hj'}^{r}) .$$

One difference from (32) is that the denominator is not the demand for the outside option of "no purchase," but rather the demand for another good, $x_{hj'}^{r}$.[14] This treatment reduces the sample size compared with the logit model for one category of goods with the outside option. We again estimate using OLS and 2SLS, using the average price in the other stores as the instrument. The results are shown in Table 3, where the constant for good 1 is assumed to be zero because only the relative size of constants matters in the logit model. All the coefficients are again significant at the 1% level using either OLS or 2SLS. The result for endogeneity is ambiguous. We reject the null hypothesis that there is no difference in the estimated coefficients using OLS and 2SLS for ketchup, but not for mayonnaise. Because we cannot determine which of OLS and 2SLS is a good estimator, we conduct the second-step estimation for both cases.

---

[14] For ketchup, we basically use brand 1 as the demand for another good, $x_{hj'}^{r}$, because it is the most widely sold. When brand 1 is not sold in a market, we use brand 6. If neither brands 1 nor 6 are sold, we use brand 3. For mayonnaise, the procedure is the same, such that we use brand 7, which is the most widely sold, or brand 1 when brand 7 is unavailable.

The second step is to estimate

(38) $$X_h^r = N^r a_h \beta_h P_h^r + \frac{1}{2} N^r b \beta_h \left(\prod_{h=1}^{2} \left(P_h^r \right)^{\frac{1}{2}} \right),$$

where $P_h^r \equiv \sum_{j' \in h} \exp(\alpha_{hj'} - \beta_h p_{hj'}^r)$ is the aggregated price index of category h at store r computed using the first step. We include dummy variables for a_h to distinguish the prefectural location in stores,[15] taking into account the possibility that the demand for ketchup or mayonnaise differs by prefecture, reflecting the differences in local dietary habits. Denoting the coefficient of $\prod_{h=1}^{2} \left(P_h^r \right)^{\frac{1}{2}}$ by $coef_h \equiv \frac{1}{2} N^r b \beta_h$, (35) implies a constraint across categories of

(39) $$coef_1 \beta_2 - coef_2 \beta_1 = 0.$$

We jointly estimate (38) for ketchup and mayonnaise under the constraint of (39). However, P_h^r might correlate with the error term. Thus, three-stage least squares (3SLS) would be a natural candidate for an estimation method. Unfortunately, we do not have a good instrument for the aggregated price index for each category at each store. The only possibility is to use the average aggregated price index in other areas. Denoting the average aggregated price index in other areas by $\overline{P}_h^{-r}$, we have

(40) $$\overline{P}_h^{-r} \equiv \frac{\sum_{r'=1}^{682} P_h^{r'} - P_h^r}{682 - 1} = \frac{\sum_{r'=1}^{682} P_h^{r'}}{681} - \frac{P_h^r}{681}.$$

The variation in $\overline{P}_h^{-r}$ arises from the second term on the right-hand side, because the first term is common to all $\overline{P}_h^{-r}$. The first term is far larger than the second

[15] If a store in located in Tokyo, all the dummy variables are zero.

term, and therefore the average aggregated price index in other areas, $\overline{P}_h^{-r}$, is almost constant. Thus, 3SLS does not work properly here.[16] Therefore, we use seemingly unrelated regressions as the estimation method for the second step. We then have the two estimation results based on the estimation method in the first stage, OLS or 2SLS. Table 4 shows the results.

In Table 4, all the coefficients, except for the dummy variables, are significant at the 1% level. One reason why the dummy variables for prefectures are not significant could be the small sample size for each prefecture. For example, eight prefectures (Yamagata, Yamanashi, Shizuoka, Tokushima, Kagawa, Ehime, Kochi, and Oita) only provide two samples each.

6.2. Elasticities

We examine the price elasticities of the derived demand functions in Tables 5 and 6. Note that we have just three different values of elasticities for each case, the own-price elasticities, the cross-price elasticities within the same category, and the cross-price elasticities across the categories, because the IIA property of the logit model makes the cross-price elasticities within the same category equal. Tables 5 and 6 report the results arising from the logit model with and without the outside option, respectively. Tables 5-1 and 6-1 detail the estimated elasticities in the case of OLS, whereas Tables 5-2 and 6-2 show the values in the case of 2SLS. In all tables, we present the median of the calculated elasticities for each store. Tables 5 and 6 demonstrate that OLS and 2SLS generally yield a similar profile of elasticities.

[16] The Chi-square value for the Hausman test is almost zero.

First, we check the results from the logit model with the outside option in Tables 5-1 and 5-2.[17] Because the demand for each good depends only on the choice probability, the cross-price elasticities of a brand of ketchup (mayonnaise) with respect to the price of mayonnaise (ketchup) are zero. From (9) and (10), the own-price elasticities of the choice probability depend on the market share of other goods within a category, and the cross-price elasticities depend on the market share of the good concerned. As suggested in Table 1, the choice probability for each good is very small if we take into account an outside option. In fact, we calculate the choice probability of the outside option, "no purchase," from Table 1 as 99.4% for ketchup and 98.5% for mayonnaise. This implies that the market share for other brands is very large whereas that for the brand being considered is very small. Thus, the own-price elasticities of the choice probability become very large and the cross-price elasticities become very small. Because the elasticities of choice probability coincide with the overall price elasticities in the logit model with the outside option, the overall own-price elasticities of demand are very large, whereas the overall cross-price elasticities are very small.

Second, we investigate the results from the logit model without the outside option in Tables 6-1 and 6-2. We know that the elasticities of the logit model without the outside option differ from those with the outside option in two ways. First, the choice probability of each good is larger because there is no outside option of "no purchase" within a category. Second, the price elasticities of

[17] In Tables 5 and 6, the elasticity indicated by the "own-price" row and the column " p_{11} " denotes the own-price elasticity of x_{11} to p_{11}. The elasticity at the "cross-price 1" row and column " p_{11} " denotes the within-category cross-price elasticity of $x_{1j''}$ ($j''=2,...,6$) to p_{11}. The elasticity at the "cross-price 2" row and column " p_{11} " denotes the across-category cross-price elasticity of x_{2j} ($j=1,...,8$) to p_{11}. The same applies to the other elasticities.

demand for a category exist. For the own-price elasticities, the former effect works to make the absolute value of the own-price elasticities smaller, whereas the latter effect works in the opposite direction. This is why we do not have a clear relationship concerning the relative size of the own elasticities for with and without the outside option in our example. For the cross-price elasticities within a category, both effects work to make the absolute value of the cross-price elasticities larger. Another point to note is that the cross-price elasticity with respect to the price of brand 1 of ketchup is very large. This reflects the fact that the market share of brand 1 of 47% is much larger than that of the other brands.

For the cross-price elasticities across the categories, the results in Tables 6-1 and 6-2 demonstrate that they are negative. This implies that a brand of ketchup and a brand of mayonnaise are gross complements. Thus, our example illustrates a situation that is inapplicable to the logit model for one category with the outside option.

6.3. Shapes of the demand functions

For a better intuitive understanding, we depict the demand functions. We focus on the change in the price of brand 1 of ketchup as an example. We fix prices other than the price of brand 1 of ketchup at their median values.

Figure 1 depicts the relationship between the price and quantity of brand 1 of ketchup, i.e., typical demand curves. All the demand curves are smoothly downward sloping and the differences are very small. This implies that when considering the relationship between own-price and quantity, it does not make much difference whether we adopt the logit model for one category with the outside option or that for two categories without the outside option.

Figure 2 illustrates the relationship between the price of brand 1 and the quantity of brand 2 of ketchup, i.e., a substitute good within the same category of goods. All curves indicate substitutionary relationships. As shown in Section 6.2, in the logit model for one category with the outside option, the unit market share of the outside option of “no purchase” is very large, and therefore the unit market share of each brand is very small. This makes the cross-elasticities very small. Thus, the demand for brand 2 of ketchup is almost unaffected by the price of brand 1 of ketchup. On the contrary, in both the logit models for ketchup and mayonnaise without an outside option, the cross-elasticities are larger, and the depicted curves are more elastic.

Figure 3 represents the relationship between the price of brand 1 of ketchup and the quantity of brand 1 of mayonnaise, i.e., a brand in the other category of good. In the logit model for one category with the outside option, we disregard the effect on the other category of goods. Thus, we obtain a vertical line concerning their relationship.[18] In the logit model for ketchup and mayonnaise without an outside option, we instead have downward-sloping curves, which illustrates that brand 1 of ketchup and brand 1 of mayonnaise are gross complements.

6.4. Welfare change with the introduction of a new brand

We calculate the change in welfare with the introduction of a new brand using the example of brand 5 of ketchup and brand 6 of mayonnaise, neither of which are sold in Tokyo. We consider a hypothetical situation in which sales commence in Tokyo at the national median price and estimate the change in welfare. The population of Tokyo is set at 13,296,019, based on the official figure in December 2013. The results in Table 7 show that the model significantly affects the estimated welfare change, such that the change in welfare differs by 15% to 33% ,

[18] In the logit model for one category with the outside option, the lines depicting the OLS and 2SLS results overlap because we fix the demands for brands in other categories at their median values.

depending on whether we use the logit model for one category with the outside option or the logit model for two categories without the outside option.

In considering the welfare change, we need to focus on two counteracting effects. The first is on the cross-elasticities within the same category. In the logit model for one category with the outside option, the large market share of the outside option implies small absolute values of the cross-elasticities between brands. Accordingly, a decrease in the price of one brand results in a relatively small decrease in the quantity of other brands within the same category of goods. Because the decrease in welfare associated with a decrease in the quantity of other brands is small, the welfare change tends to be larger in the logit model for one category with the outside option. The second effect is on the cross-elasticities across categories of goods. Because the relationship between brands of ketchup and brands of mayonnaise is complementary, a decrease in the price of brand 5 of ketchup (brand 6 of mayonnaise) increases the demand for all brands of mayonnaise (ketchup). This makes the welfare change smaller in the logit model for one category with the outside option.

We have no clear predetermined results regarding whether the first effect dominates the second effect. In our simulation, the first effect dominates the second effect concerning the decrease in the price of brand 5 of ketchup, whereas the reverse holds regarding the decrease in the price of brand 6 of mayonnaise. This arises because of the differences in the size of the cross-elasticities such that the absolute values of the cross-elasticities are larger for the demand for ketchup in relation to the price of mayonnaise than for the demand for mayonnaise in relation to the price of ketchup. Thus, for a decrease in the price of brand 6 of mayonnaise, the second effect is sufficiently large to offset the first effect.

7. Conclusion

This paper develops a utility maximization model that is consistent with the results of the logit model but allows the incorporation of multiple categories of goods in an unrestrictive way, and examines the characteristics of the derived demand functions and a method for welfare estimation. The analysis extends naturally to the cases of more general GEV or mixed GEV models. Using POS data, we empirically applied the model to the markets for ketchup and mayonnaise to suggest that a brand of ketchup is complementary to a brand of mayonnaise. Our framework is most useful in describing complementary relationships across categories, which are generally intractable in the conventional discrete choice models without restrictive assumptions.

In concluding the paper, we note four points for future research, mainly from the viewpoint of the possibility of more sophisticated empirical analyses. First, our framework is equally applicable to person-level data on purchase behavior and market-level data. The model in this paper allows a consumer to purchase as many brands and goods as they like. If we can estimate the individual demand function using person-level data, we can undertake a more accurate demand analysis.

Second, we exclude behavior on the supply side. A natural direction is to extend our analysis to structural analyses including both demand and supply behaviors. As Chintagunta and Nair (2011) suggest, it is rather difficult to model supply-side behavior for a real-world situation, and incorrect modeling would yield significant misspecification bias. However, it would be meaningful to derive the implications of the introduction of the supply side by, for example, assuming Bertrand competition.

Third, we need to be more cautious regarding the possibly of endogeneity. In the empirical example in this paper, we use the average prices of other areas as instruments, following Hausman (1996) and Nevo (2001), as they are the only possible instruments given the limitations of our data. The differences between the OLS and 2SLS results are generally minor, partly because the variation in instruments is rather limited. Another way to combat endogeneity is to use the control function proposed by Petrin and Train (2010). However, even under this approach, we need instruments. We need to check the general problem of endogeneity in our framework using other data that enable us to use different instruments.

Finally, the unrealistic implication of the IIA property remains unaddressed in our empirical example because our primary aim is to estimate discrete choice models that allow multiple categories of goods that can be either substitutes or complements. In order to conduct a proper empirical analysis in a more realistic context, we need to apply our method to GEV or mixed GEV models that are free of the IIA property.

References

Anderson, S.P., De Palma, A., Thisse, J.-F., 1988. A representative consumer theory of the logit model. International Economic Review 29, 461–466.

Anderson, S.P., De Palma, A., Thisse, J.-F., 1992. Discrete Choice Theory of Product Differentiation. MIT Press, Cambridge, MA.

Ben-Akiva, M., Francois, B., 1983. μ Homogeneous Generalized Extreme Value Model. MIT Department of Civil Engineering Working Paper.

Ben-Akiva, M., Lerman, S.R., 1985. Discrete Choice Analysis. MIT Press, Cambridge, MA.

Berry, S., 1994. Estimating Discrete-Choice Models of Product Differentiation. Rand Journal of Economics 25, 242–262.

Berry, S., Levinsohn, J., Pakes, A., 1995. Automobile prices in market equilibrium. Econometrica 63, 841–890.

Bhat, C.R., 2005. A multiple discrete-continuous extreme value model: formulation and application to discretionary time-use decisions. Transportation Research Part B 39, 679–707.

Bhat, C.R., 2008. The multiple discrete-continuous extreme value (MDCEV) model: role of utility function parameters, identification considerations, and model extensions. Transportation Research Part B 42, 274–303.

Chintagunta, P.K., Nair, H.S., 2011. Structural workshop paper—discrete-choice models of consumer demand in marketing. Marketing Science 30, 977-996.

Dagsvik, J.K., 1995. How large is the class of generalized extreme value random utility models? Journal of Mathematical Psychology 39, 90–98.

Deaton A., Muellbauer, J., 1980. An almost ideal demand system. American Economic Review 70, 312-326.

Domencich, T. A. and Macfadden, D., 1975. Urban Travel Demand. North-Holland, Amsterdam.

Dube J.-P., 2004. Multiple Discreteness and Product Differentiation: Demand for Carbonated Soft Drinks. Marketing Science 23, 66-81.

Dubin, J.A., McFadden, D., 1984. An econometric analysis of residential electric appliance holdings and consumption. Econometrica 52, 345–362.

Earnhart, D., 2002. Combining revealed and stated data to examining housing decisions using discrete choice analysis. Journal of Urban Economics 51, 143–169.

Gentzkow, M., 2007. Valuing new goods in a model with complementarity: online newspapers. American Economic Review 97, 713–744.

Goldberg, P.K., 1995. Product differentiation and oligopoly in international markets: the case of the U.S. automobile industry. Econometrica 63, 891–951.

Hanemann, W.M., 1984. Discrete/continuous models of consumer demand. Econometrica 52, 541–562.

Hendel, I., 1999. Estimating multiple-discrete choice models: an application to computerization returns. Review of Economic Studies 66, 423–446.

Mas-Colell, A., Whinston, M.D., Green, J.R., 1995. Microeconomic Theory. Oxford University Press, Oxford.

McFadden, D., 1978. Modeling the choice of residential location, in: Karlqvist, A., Lundqvist, L., Snickars, F., Weibull, J.W. (Eds.), Spatial Interaction Theory and Planning Models, North-Holland, Amsterdam, pp. 75–96.

McFadden, D., 1999. Computing willingness-to-pay in random utility models, in: Melvin, J.R., Moore, J.C., Riezman, R. (Eds.), Trade, Theory, and Econometrics: Essays in Honor of J.S. Chipman, Routledge, London.

McFadden, D., Train, K., 2000. Mixed MNL models for discrete response. Journal of Applied Econometrics 15, 447–470.

Mehta, N., Ma, Y., 2012. A multicategory model of consumers' purchase incidence, quantity, and brand choice decisions: methodological issues and implications on promotional decisions. Journal of Marketing Research 50, 435-451.

Nevo, A., 2000. A practitioner's guide to estimation of random-coefficients logit models of demand. Journal of Economics and Management Strategy 9, 513–548.

Nevo, A., 2001. Measuring market power in the ready-to-eat cereal industry. Econometrica 69, 307–342.

Nevo, A., 2011. Empirical models of consumer behavior. Annual Review of Economics 3, 51–75.

Nevo, A., Hatzitaskos, K., 2006. Why does the average price paid fall during high demand periods? URL: http://faculty.wcas.northwestern.edu/~ane686/research/tuna3.pdf, accessed on December 10, 2012.

Oum, T.H., Waters II, W.G., Yong, J.-S., 1992. Concepts of price elasticities of transport demand and recent empirical estimates—an interpretative survey. Journal of Transport Economics and Policy 26, 139–154.

Petrin, A., 2002. Quantifying the benefits of new products: the case of the minivan. Journal of Political Economy 110, 705–729.

Petrin, A. and Train, K., 2010. A control function approach to endogeneity in consumer choice models. Journal of Marketing Research 47, 3-13.

Small, K.A., Rosen, H. S., 1981. Applied welfare economics with discrete choice models. Econometrica 49, 105–130.

Song, I., Chintagunta, P. K., 2007. A discrete-continuous model for multicategory purchase behavior of households. Journal of Marketing Research 44, 595-612.

Taplin, J.H.E., 1982. Inferring ordinary elasticities from choice or mode-split elasticities. Journal of Transport Economics and Policy 16, 55–63.

Train, K., 2009. Discrete choice methods with simulation, second ed. Cambridge University Press, Cambridge, UK.

Varian, H.R., 1992. Microeconomic analysis, third ed. Norton, New York, NY.

Verboven, F., 1996. The nested logit model and representative consumer theory. Economic Letters 50, 57–63.

Zachmanoglou, E.C., Thoe, D.W., 1986. Introduction to partial differential equations with applications, Dover Publications, New York.

Figures and Tables

Figure 1 The demand curve for good 1 in ketchup

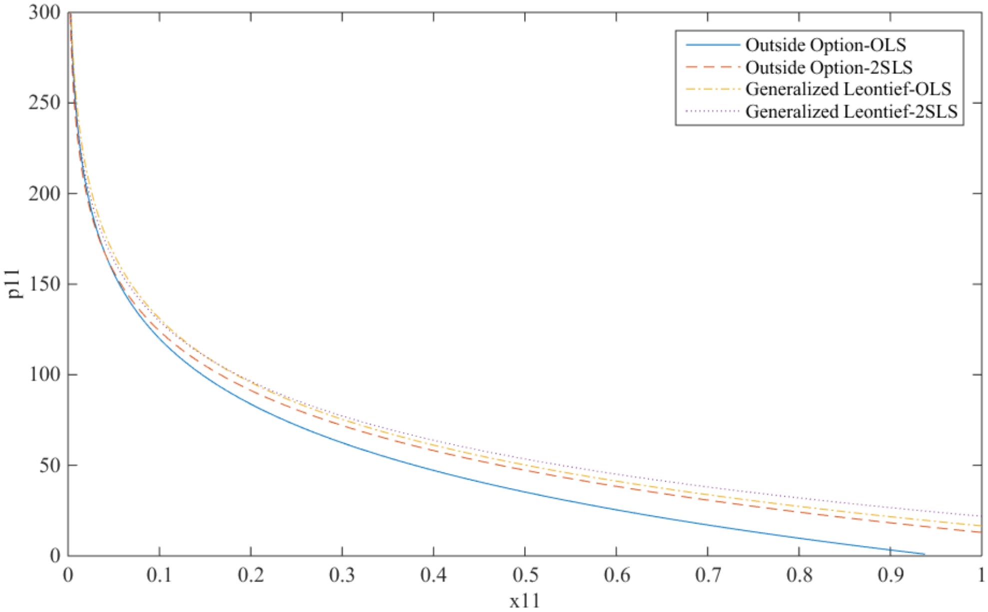

Figure 2 The relationship between the demand of good 2 in ketchup and the price of good 1 in ketchup

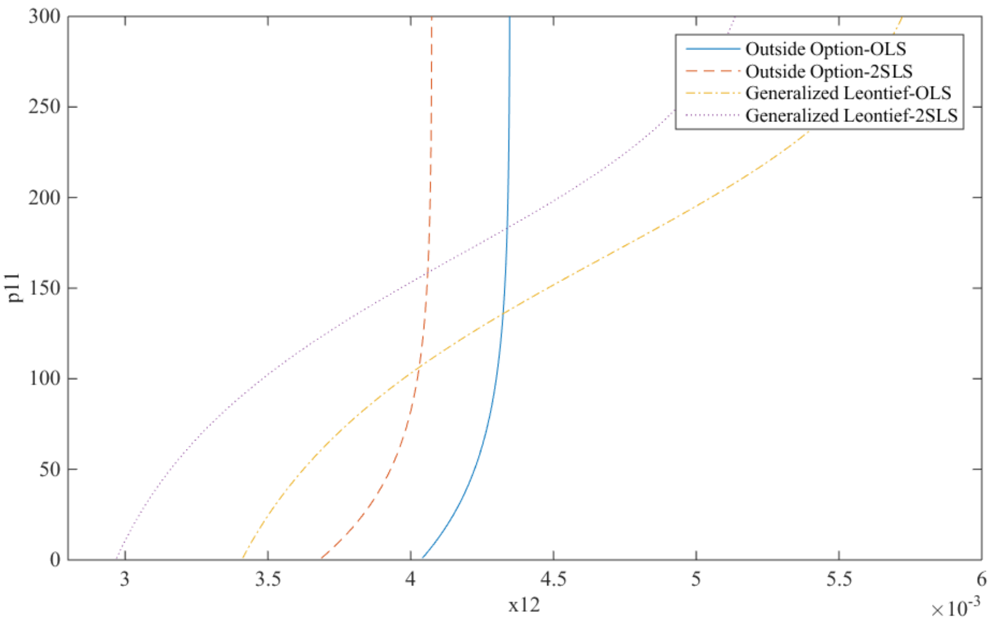

Figure 3 The relationship between the demand of good 1 in mayonnaise and the price of good 1 in ketchup

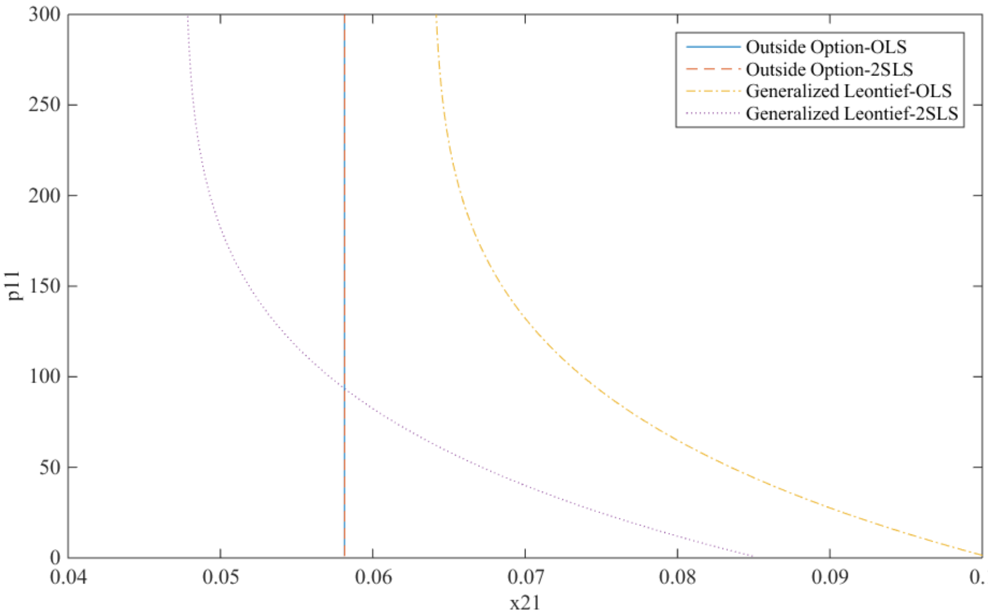

Table 1 Descriptive statistics

		Ketchup						Mayonnaise							
		Good1	Good2	Good3	Good4	Good5	Good6	Good1	Good2	Good3	Good4	Good5	Good6	Good7	Good8
Revenue market share (%)		46.97	5.95	5.94	5.41	3.61	3.15	23.82	15.29	10.15	6.37	5.48	3.41	3.41	3.02
Unit market share (%)	Without outside option	47.31	4.88	6.73	4.86	4.22	2.82	24.09	15.20	11.99	6.50	6.20	4.23	3.77	2.96
	With outside option	0.39	0.04	0.06	0.04	0.03	0.02	0.47	0.30	0.23	0.13	0.12	0.08	0.07	0.06
Number of Market		673	272	301	588	182	540	645	657	663	618	656	197	669	536
Price (yen)	Mean	165	203	170	177	136	189	213	218	194	221	193	165	184	221
	Median	165	195	179	176	134	183	213	213	189	227	189	163	183	226
	Standard deviation	25	26	38	17	13	32	30	34	41	31	32	19	16	28
	Min	100	153	99	141	95	97	134	132	107	126	106	93	122	138
	Max	298	284	300	244	189	260	350	341	333	303	265	203	244	333
Quantities sold at each store	Mean	278	71	88	33	92	21	349	216	169	98	88	201	53	52
	Median	208	42	37	26	73	13	213	213	189	227	189	163	183	226
	Standard deviation	244	102	113	26	70	28	398	223	190	181	82	169	44	68
	Min	9	1	1	1	1	1	2	4	3	1	2	9	4	2
	Max	1714	837	746	212	399	360	3375	1802	1399	1512	408	831	329	576

Table 2 The logit model with outside option

		Ketchup				Mayonnaise			
		OLS		2SLS		OLS		2SLS	
Price		-0.0194	***	-0.0213	***	-0.0197	***	-0.0197	***
		(0.000502)		(0.000541)		(0.000286)		(0.000297)	
Constant	Good 1	-2.557	***	-2.231	***	-1.397	***	-1.396	***
		(0.0869)		(0.0931)		(0.0654)		(0.0676)	
	Good 2	-3.452	***	-3.052	***	-1.838	***	-1.837	***
		(0.110)		(0.117)		(0.0666)		(0.0689)	
	Good 3	-4.116	***	-3.782	***	-2.584	***	-2.582	***
		(0.0934)		(0.0995)		(0.0601)		(0.0621)	
	Good 4	-4.421	***	-4.072	***	-2.814	***	-2.813	***
		(0.0930)		(0.0996)		(0.0677)		(0.0700)	
	Good 5	-4.324	***	-4.056	***	-3.193	***	-3.192	***
		(0.0843)		(0.0887)		(0.0599)		(0.0619)	
	Good 6	-4.899	***	-4.525	***	-2.84	***	-2.839	***
		(0.0993)		(0.106)		(0.0637)		(0.0650)	
	Good 7					-3.711	***	-3.71	***
						(0.0574)		(0.0593)	
	Good 8					-3.296	***	-3.295	***
						(0.0682)		(0.0705)	
Observations		2,556		2,556		4,641		4,641	
Hausman Statistics			98.178				0.00893		

*** Significant at 1% level

Table 3 The logit model without outside option

		Ketchup				Mayonnaise			
		OLS		2SLS		OLS		2SLS	
Price		-0.0219	***	-0.0233	***	-0.0181	***	-0.0182	***
		(0.000612)		(0.000676)		(0.000326)		(0.000339)	
Constant	Good 2	-0.849	***	-0.801	***	2.28	***	2.283	***
		(0.0552)		(0.0561)		(0.0274)		(0.0275)	
	Good 3	-1.475	***	-1.467	***	1.808	***	1.812	***
		(0.0485)		(0.0485)		(0.0290)		(0.0292)	
	Good 4	-1.854	***	-1.837	***	1.112	***	1.113	***
		(0.0355)		(0.0357)		(0.0268)		(0.0268)	
	Good 5	-1.805	***	-1.843	***	0.83	***	0.834	***
		(0.0646)		(0.0651)		(0.0303)		(0.0304)	
	Good 6	-2.303	***	-2.268	***	0.494	***	0.495	***
		(0.0392)		(0.0398)		(0.0270)		(0.0270)	
	Good 7					0.856	***	0.854	***
						(0.0491)		(0.0491)	
	Good 8					0.333	***	0.337	***
						(0.0320)		(0.0322)	
Observations		1,872		1,872		3,922		3,922	
Hausman Statistics			23.592				1.317		

*** Significant at 1% level

Table 4 Demand functions

Variable		Ketchup OLS		Ketchup 2SLS		Mayonnaise OLS		Mayonnaise 2SLS	
$P_{Ketchup}$		0.09690	***	0.112***	***				
		(0.00857)		(0.0105)					
$P_{Mayonnaise}$						0.0185	***	0.0187	***
						(0.000976)		(0.000980)	
$(P_{Ketchup}P_{Mayonnaise})^{0.5}$		0.01220	***	0.0153***	***	0.0101	***	0.0120	***
		(0.00212)		(0.00235)		(0.00175)		(0.00184)	
Prefecture Dummy	Hokkaido	0.01840	***	0.02430	***	0.00541	***	0.00551	***
		(0.00579)		(0.00715)		(0.00123)		(0.00125)	
	Aomori	-0.0115		-0.0118		-0.00429	**	-0.00426	**
		(0.0140)		(0.0176)		(0.00204)		(0.00207)	
	Iwate	0.00279		0.00536		-0.00216		-0.00217	
		(0.00883)		(0.0110)		(0.00163)		(0.00165)	
	Miyagi	0.0000236		-0.00181		-0.000445		-0.000526	
		(0.00698)		(0.00850)		(0.00148)		(0.00151)	
	Akita	-0.03190	*	-0.0375		-0.00230		-0.00226	
		(0.0181)		(0.0229)		(0.00245)		(0.00249)	
	Yamagata	-0.0122		-0.0110		-0.00360		-0.00357	
		(0.0403)		(0.0514)		(0.00656)		(0.00668)	
	Fukushima	0.06710	***	0.09340	***	0.00674	*	0.00699	*
		(0.0242)		(0.0313)		(0.00352)		(0.00359)	
	Ibaraki	0.000401		0.00555		0.00195		0.00209	
		(0.0217)		(0.0278)		(0.00306)		(0.00311)	
	Tochigi	-0.0160		-0.0189		-0.00451		-0.00458	
		(0.0190)		(0.0237)		(0.00282)		(0.00286)	
	Gunma	-0.00842		-0.00837		0.00393	**	0.00406	**
		(0.00858)		(0.0107)		(0.00163)		(0.00166)	
	Saitama	-0.03600	***	-0.04490	***	-0.000550		-0.000577	
		(0.00670)		(0.00821)		(0.00136)		(0.00138)	
	Chiba	-0.00972		-0.0129		0.00217		0.00219	
		(0.00939)		(0.0115)		(0.00194)		(0.00197)	
	Kanagawa	-0.02530	***	-0.02990	***	-0.000240		-0.000200	
		(0.00740)		(0.00913)		(0.00154)		(0.00157)	
	Niigata	0.0159		0.0227		0.00522	**	0.00535	**
		(0.0117)		(0.0146)		(0.00256)		(0.00260)	
	Toyama	-0.05240	***	-0.06620	***	-0.00628	***	-0.00642	***
		(0.00890)		(0.0107)		(0.00182)		(0.00185)	
	Ishikawa	-0.00979		-0.0145		-0.00513	***	-0.00519	***
		(0.00976)		(0.0118)		(0.00172)		(0.00174)	
	Fukui	-0.0144		-0.0168		-0.000958		-0.000952	
		(0.00978)		(0.0121)		(0.00189)		(0.00192)	
	Yamanashi	0.20700	**	0.25400	**	0.02420	***	0.02480	***
		(0.0845)		(0.105)		(0.00468)		(0.00476)	
	Nagano	-0.04340	***	-0.05210	***	-0.000368		-0.000401	
		(0.00836)		(0.0103)		(0.00218)		(0.00222)	
	Gifu	0.00571		0.00660		0.00208		0.00214	
		(0.0170)		(0.0212)		(0.00221)		(0.00224)	
	Shizuoka	0.00378		0.0109		0.000467		0.000539	
		(0.0583)		(0.0745)		(0.0120)		(0.0122)	

*** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level

Table 4 Continued

Aichi	-0.00826		-0.00874		-0.000464		-0.000395	
	(0.00904)		(0.0113)		(0.00138)		(0.00140)	
Mie	-0.05950	***	-0.07520	***	-0.00552	***	-0.00564	***
	(0.0104)		(0.0125)		(0.00182)		(0.00185)	
Kyoto	-0.0752		-0.0913		0.00177		0.00202	
	(0.0549)		(0.0716)		(0.00487)		(0.00496)	
Osaka	-0.03120	***	-0.03890	***	-0.000466		-0.000468	
	(0.00735)		(0.00898)		(0.00141)		(0.00143)	
Hyogo	-0.01800	***	-0.02170	*	-0.000703		-0.000688	
	(0.0101)		(0.0125)		(0.00184)		(0.00187)	
Nara	-0.0112		-0.0142		0.00265		0.00272	
	(0.0182)		(0.0224)		(0.00336)		(0.00342)	
Wakayama	-0.0178		-0.0189		-0.000238		-0.000201	
	(0.0157)		(0.0197)		(0.00373)		(0.00379)	
Shimane	0.0349		0.0526		0.00614	**	0.00636	**
	(0.0255)		(0.0331)		(0.00281)		(0.00286)	
Okayama	-0.05350	***	-0.06550	***	-0.000816		-0.000854	
	(0.00908)		(0.0111)		(0.00218)		(0.00222)	
Hiroshima	-0.02670	**	-0.03190	**	-0.00534	***	-0.00537	***
	(0.0111)		(0.0139)		(0.00156)		(0.00158)	
Yamaguchi	0.0319*		0.0428*		0.00801	**	0.00821	**
	(0.0177)		(0.0222)		(0.00370)		(0.00376)	
Tokushima	0.0641		0.0852		0.01280	**	0.01310	**
	(0.0461)		(0.0588)		(0.00550)		(0.00559)	
Kagawa	0.0391		0.0534		0.00722		0.00748	
	(0.0440)		(0.0560)		(0.00632)		(0.00644)	
Ehime	0.0170		0.0231		0.00289		0.00300	
	(0.0369)		(0.0465)		(0.00516)		(0.00524)	
Kochi	-0.0203		-0.0216		0.00346		0.00364	
	(0.0441)		(0.0562)		(0.00681)		(0.00693)	
Fukuoka	-0.02250	***	-0.03370	***	-0.00124		-0.00138	
	(0.00567)		(0.00681)		(0.00107)		(0.00109)	
Saga	-0.03290	***	-0.04570	***	0.01370	***	0.01380	***
	(0.0114)		(0.0137)		(0.00181)		(0.00183)	
Nagasaki	-0.00534		-0.0110		0.01730	***	0.01750	***
	(0.0108)		(0.0130)		(0.00183)		(0.00185)	
Kumamoto	-0.00562		-0.0151		0.02080	***	0.02080	***
	(0.0125)		(0.0149)		(0.00196)		(0.00198)	
Oita	-0.0194		-0.03180	**	0.02040	***	0.02050	***
	(0.0135)		(0.0158)		(0.00199)		(0.00201)	
Miyazaki	-0.04800	***	-0.05890	***	0.00145		0.00151	
	(0.0145)		(0.0179)		(0.00236)		(0.00239)	
Kagoshima	-0.0167		-0.0169		0.000723		0.000856	
	(0.0436)		(0.0565)		(0.00382)		(0.00388)	
Observations	682		682		682		682	

*** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level

Table 5-1 Elasticities for the logit model with one category with outside option-OLS

		Price													
		p_{11}	p_{12}	p_{13}	p_{14}	p_{15}	p_{16}	p_{21}	p_{22}	p_{23}	p_{24}	p_{25}	p_{26}	p_{27}	p_{28}
Type of elasticities	Own Price	-3.1798	-3.7756	-3.4645	-3.4081	-2.5989	-3.5346	-4.1722	-4.1934	-3.7100	-4.4630	-3.7207	-3.1959	-3.6175	-4.4566
	Cross Price 1	0.0101	0.0027	0.0018	0.0014	0.0025	0.0008	0.0156	0.0099	0.0068	0.0030	0.0036	0.0075	0.0023	0.0019
	Cross Price 2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Own Price: own price elasticities, Cross Price 1: cross price elasticities within the same category, Cross Price 2: cross price elasticities across the categories

Table 5-2 Elasticities for the logit model with one category with outside option-2SLS

		Price													
		p_{11}	p_{12}	p_{13}	p_{14}	p_{15}	p_{16}	p_{21}	p_{22}	p_{23}	p_{24}	p_{25}	p_{26}	p_{27}	p_{28}
Type of elasticities	Own Price	-3.5038	-4.1602	-3.8175	-3.7554	-2.8637	-3.8947	-4.1738	-4.1951	-3.7115	-4.4648	-3.7222	-3.1971	-3.6189	-4.4583
	Cross Price 1	0.0112	0.0031	0.0019	0.0015	0.0028	0.0009	0.0156	0.0099	0.0068	0.0030	0.0036	0.0075	0.0023	0.0019
	Cross Price 2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 6-1 Elasticities for the logit model with multiple category without outside option-OLS in the first stage

		Price													
		p_{11}	p_{12}	p_{13}	p_{14}	p_{15}	p_{16}	p_{21}	p_{22}	p_{23}	p_{24}	p_{25}	p_{26}	p_{27}	p_{28}
Type of elasticities	Own Price	-3.2223	-4.1862	-3.8399	-3.7705	-2.8875	-3.9736	-3.7719	-3.8188	-3.3715	-4.0900	-3.4152	-2.9199	-3.3188	-4.0907
	Cross Price 1	0.4000	0.0942	0.0852	0.0498	0.0849	0.0252	0.0747	0.0444	0.0306	0.0160	0.0164	0.0310	0.0115	0.0100
	Cross Price 2	-0.1594	-0.0367	-0.0333	-0.0206	-0.0348	-0.0103	-0.2619	-0.1558	-0.1029	-0.0518	-0.0540	-0.0948	-0.0367	-0.0309

Table 6-2 Elasticities for the logit model with multiple category without outside option-2SLS in the first stage

		Price													
		p_{11}	p_{12}	p_{13}	p_{14}	p_{15}	p_{16}	p_{21}	p_{22}	p_{23}	p_{24}	p_{25}	p_{26}	p_{27}	p_{28}
Type of elasticities	Own Price	-3.3625	-4.4410	-4.0711	-3.9935	-3.0619	-4.2205	-3.7896	-3.8382	-3.3894	-4.1131	-3.4348	-2.9354	-3.3374	-4.1138
	Cross Price 1	0.4782	0.1119	0.1012	0.0594	0.1015	0.0301	0.0797	0.0473	0.0325	0.0171	0.0174	0.0330	0.0123	0.0108
	Cross Price 2	-0.1791	-0.0411	-0.0374	-0.0232	-0.0389	-0.0117	-0.3041	-0.1803	-0.1193	-0.0602	-0.0625	-0.1100	-0.0426	-0.0359

Table 7 Welfare change by an introduction of good 5 in ketchup or good 6 in mayonnaise

			Price	
			p_{15}	p_{26}
Change in CV (million yen)	Single category with outside option (a)	OLS	0.6710	1.5816
		IV	0.6107	1.5810
	Multiple categories (b)	OLS	0.6065	2.0428
		IV	0.5217	2.0992
Ratio (%)	b/a	OLS	90	129
		IV	85	133

Appendix A. List of variables and parameters

x_j demand for differentiated good j

p_j price of differentiated good j ($j=1,...,M$)

y a consumer's income

x_g total demand for a choice set

P aggregate price index

s_j choice probability of differentiated good j

v indirect utility function of a consumer

$\alpha_j(y)$ alternative-specific parameter

$\beta(y)$ common coefficient of p_j

u direct utility function of a consumer

a parameter in the direct utility function of (7)

θ_j elasticity of x_g with respect to p_j

cv compensating variation

e expenditure function

x_j^c compensated demand for differentiated good j

x_g^c compensated total demand for a choice set

LS log-sum term for the logit model

P_h aggregate price index for category h

$\mathbf{p}_h$ price vector for differentiated goods that belong to category h ($h=1,...,H$)

LS_h log-sum term for category h

x_{hj} demand for differentiated good j in category h

p_{hj} price of differentiated good j in category h

x_{hg} total demand for category h

s_{hj} choice probability for differentiated good j within category h

β_h a common coefficient of p_k for category h ($k \in h$)

θ_{hj} elasticity of x_{hg} with respect to p_j

x^c_{hg} compensated total demand for category h

V indirect utility function of the representative consumer

Y aggregated income across consumers

Appendix B. The necessary conditions for the demand function for differentiated good to have the form of (2)

From (2) and Roy's Identity, the indirect utility function must satisfy:

$$\text{(B.1)} \quad x_g(p_1,\ldots,p_M,y) = \frac{-\dfrac{\partial v}{\partial p_1}}{\dfrac{\exp(\phi_1(p_1,y))}{\sum_{j'=1}^{M}\exp(\phi_{j'}(p_{j'},y))}\dfrac{\partial v}{\partial y}} = \ldots = \frac{-\dfrac{\partial v}{\partial p_M}}{\dfrac{\exp(\phi_M(p_M,y))}{\sum_{j'=1}^{M}\exp(\phi_{j'}(p_{j'},y))}\dfrac{\partial v}{\partial y}}.$$

We first consider the case of $M = 2$. From the second and last equations of (B.1), we have

$$\text{(B.2)} \quad \frac{-1}{\exp(\phi_1(p_1,y))}\frac{\partial v}{\partial p_1} + \frac{1}{\exp(\phi_2(p_2,y))}\frac{\partial v}{\partial p_2} = 0.$$

The solution of (B.2) can be derived by solving the system of equations (e.g., Zachmanoglou and Thoe (1986)):

$$\text{(B.3)} \quad \frac{dp_1}{\dfrac{1}{-\exp(\phi_1(p_1,y))}} = \frac{dp_2}{\dfrac{1}{\exp(\phi_2(p_2,y))}} = \frac{dy}{0} = \frac{dv}{0}.$$

The third and fourth equations of (B.3) immediately implies $dy = dv = 0$, from which we obtain:

(B.4) $y = C_1$ and

(B.5) $v = C_2$,

where C_1 and C_2 are constants. In the following, C_l denotes constant. The first and second equations of (B.3) is rewritten as

(B.6) $\exp(\phi_1(p_1, y))dp_1 + \exp(\phi_2(p_2, y))dp_2 = 0$.

Solving (B.6) regarding p_1 and p_2, noting that y is constant from (B.4), we obtain:

(B.7) $\sum_{j'=1}^{2} \exp(\phi_{j'}(p_{j'}, y)) \frac{1}{\frac{\partial \phi_{j'}}{\partial p_{j'}}} = C_3$,

where $\frac{\partial \phi_j}{\partial p_j}$ is a constant that does not depend on p_j. This immediately implies that $\frac{\partial \phi_j}{\partial p_j}$ is linear in p_j.

We have the two cases to consider: $\frac{\partial \phi_1}{\partial p_1} \neq \frac{\partial \phi_2}{\partial p_2}$ and $\frac{\partial \phi_1}{\partial p_1} = \frac{\partial \phi_2}{\partial p_2}$. In the case of $\frac{\partial \phi_1}{\partial p_1} \neq \frac{\partial \phi_2}{\partial p_2}$, we have

(B.8) $\phi_j = \alpha_j(y) - \beta_j(y) p_j$.

Substituting (B.8) into (B.7) yields

(B.9) $\sum_{j'=1}^{2} \exp(\alpha_{j'}(y) - \beta_{j'}(y) p_{j'}) \frac{1}{\beta_{j'}(y)} = C_3$.

From (B.4), (B.5), and (B.9), the solution of (B.3) is:

(B.10) $f\left(y, v, \sum_{j'=1}^{2} \exp(\alpha_{j'}(y) - \beta_{j'}(y) p_{j'}) \frac{1}{\beta_{j'}(y)} \right) = 0$.

Solving (B.10) regarding v yields:

(B.11) $v = v\left(\sum_{j'=1}^{2}\exp(\alpha_{j'}(y) - \beta_{j'}(y)p_{j'})\frac{1}{\beta_{j'}(y)}, y\right)$.

Substituting (B.11) into the second equation of (B.1), we derive:

(B.12) $$\frac{-\frac{\partial v}{\partial \tilde{P}}\exp(\alpha_1(y) - \beta_1(y)p_1)}{\frac{\exp(\alpha_1(y) - \beta_1(y)p_1)}{\sum_{j'=1}^{M}\exp(\alpha_{j'}(y) - \beta_{j'}(y)p_{j'})}\frac{\partial v}{\partial y}} = \frac{-\frac{\partial v}{\partial \tilde{P}}\sum_{j'=1}^{M}\exp(\alpha_{j'}(y) - \beta_{j'}(y)p_{j'})}{\frac{\partial v}{\partial y}},$$

where $\tilde{P} \equiv \sum_{j'=1}^{2}\exp(\alpha_{j'}(y) - \beta_{j'}(y)p_{j'})\frac{1}{\beta_{j'}(y)}$ is the aggregated price index. Eq. (B.12) demonstrates that the demand for a category, x_g, depends on $\sum_{j'=1}^{M}\exp(\alpha_{j'}(y) - \beta_{j'}(y)p_{j'})$, as well as the aggregated price index, $\tilde{P}$, and income, y, although the first equation of (B.1) requires that x_g depends only on $\tilde{P}$ and y. (Note that $\sum_{j'=1}^{M}\exp(\alpha_{j'}(y) - \beta_{j'}(y)p_{j'})$ cannot be represented as the function of $\tilde{P}$.) In conclusion, the case of $\frac{\partial \phi_1}{\partial p_1} \neq \frac{\partial \phi_2}{\partial p_2}$ is not consistent with (B.1).

In the case of $\frac{\partial \phi_1}{\partial p_1} = \frac{\partial \phi_2}{\partial p_2}$, $\beta_j(y)$ does not depend on j. Thus, using (B.4), (B.9) can be rewritten as

(B.13) $\sum_{j'=1}^{2}\exp(\alpha_{j'}(y) - \beta(y)p_{j'}) = \beta(y)C_3 = \beta(C_1)C_3 \equiv C_4$.

From (B.4), (B.5), and (B.13), the solution of (B.3) is:

(B.14) $f\left(y, v, \sum_{j'=1}^{2}\exp(\alpha_{j'}(y) - \beta(y)p_{j'})\right) = 0$.

Solving (B.14) regarding v yields (3) and (4) in the case of $M=2$. $\beta(y)>0$ follows from $\frac{\partial s_j}{\partial p_j}<0$.

We then consider the case of $M=3$. Solving (B.2) in the same manner as the case of $M=2$, we have the two cases to consider: $\frac{\partial\phi_1}{\partial p_1}\neq\frac{\partial\phi_2}{\partial p_2}$ and $\frac{\partial\phi_1}{\partial p_1}=\frac{\partial\phi_2}{\partial p_2}$. We can exclude the case of $\frac{\partial\phi_1}{\partial p_1}\neq\frac{\partial\phi_2}{\partial p_2}$, as is the case with $M=2$. Thus, we have $\frac{\partial\phi_1}{\partial p_1}=\frac{\partial\phi_2}{\partial p_2}$, which yields a counterpart of (B.10):

(B.15) $$g\left(y,v,p_3,\frac{1}{\beta(y)}\sum_{j'=1}^{2}\exp(\alpha_{j'}(y)-\beta(y)p_{j'})\right)=0.$$

Solving (B.15) regarding v yields

(B.16) $$v=\tilde{g}\left(y,p_3,\frac{1}{\beta(y)}\sum_{j'=1}^{2}\exp(\alpha_{j'}(y)-\beta(y)p_{j'})\right).$$

In the case of $M=3$, we also have the relationship of

(B.17) $$\frac{-1}{\exp(\alpha_1(y)-\beta(y)p_1)}\frac{\partial v}{\partial p_1}+\frac{1}{\exp(\phi_3(p_3,y))}\frac{\partial v}{\partial p_3}=0$$

Substituting (B.16) into (B.17), we derive

(B.18) $$\frac{\partial\tilde{g}}{\partial\left(\frac{1}{\beta(y)}\sum_{j'=1}^{2}\exp(\alpha_{j'}(y)-\beta(y)p_{j'})\right)}+\frac{1}{\exp(\phi_3(p_3,y))}\frac{\partial\tilde{g}}{\partial p_3}=0.$$

The solution of (B.18) can be derived by solving the system of equations:

(B.19) $$d\left(\frac{1}{\beta(y)}\sum_{j'=1}^{2}\exp(\alpha_{j'}(y)-\beta(y)p_{j'})\right)=\frac{dp_3}{\frac{1}{\exp(\phi_3(p_3,y))}}=\frac{dy}{0}=\frac{dv}{0}.$$

The third and fourth equations of (B.19) imply $dy = dv = 0$. In the same way as (B.4) and (B.5), we obtain:

(B.20) $y = C_5$ and

(B.21) $v = C_6$.

Solving (B.19) regarding the two left equations, noting that y is constant from (B.20), we obtain:

(B.22) $$\frac{1}{\beta(y)}\sum_{j'=1}^{2}\exp(\alpha_{j'}(y)-\beta(y)p_{j'}) - \frac{\exp(\phi_3(p_3, y))}{\frac{\partial\phi_3}{\partial p_3}} = C_7,$$

where $\frac{\partial\phi_3}{\partial p_3}$ is a constant that does not depend on p_3. This immediately implies that $\frac{\partial\phi_3}{\partial p_3}$ is linear in p_j. Noting that $\frac{\partial\phi_1}{\partial p_1} = \frac{\partial\phi_2}{\partial p_2} = \frac{\partial\phi_3}{\partial p_3}$ in the same way as $M = 2$, we can write ϕ_3 as

(B.23) $\phi_3 = \alpha_3(y) - \beta(y)p_3$.

Substituting (B.23) into (B.22) yields

(B.24) $$\sum_{j'=1}^{3}\exp(\alpha_{j'}(y)-\beta(y)p_{j'}) = \beta(y)C_7 = \beta(C_5)C_7 \equiv C_8.$$

From (B.20), (B.21), and (B.24), the solution of (B.19) is:

(B.25) $$h\left(y, v, \sum_{j'=1}^{3}\exp(\alpha_{j'}(y)-\beta(y)p_{j'})\right) = 0$$

Solving (B.25) regarding v yields (3) and (4) in the case of $M = 3$. The same method applies to the proof for the case of $M \geq 4$.

Appendix C. The GEV model

From McFadden (1978: Theorem 1), the GEV model can be described by using the function $F(\mu_1, \cdots, \mu_M)$, where $\mu_j \equiv \exp(\phi_j(p_j, y))$.

(GEV-1) $F(\mu_1, \cdots, \mu_M)$ is nonnegative.

(GEV-2) $F(\mu_1, \cdots, \mu_M)$ is homogeneous of degree ς.[19]

(GEV-3) $\lim_{z_j \to \infty} F(\mu_1, \cdots, \mu_M) = \infty$.

(GEV-4) The κ-th order partial derivative of $F(\mu_1, \cdots, \mu_M)$ with respect to any combination of distinct μ_j is nonnegative if κ is odd and nonpositive if κ is even. That is, $\dfrac{\partial F}{\partial \mu_j} \geq 0$ for all j, $\dfrac{\partial^2 F}{\partial \mu_j \partial \mu_{j'}} \leq 0$ for all $j' = 1, \ldots, M$ and $j' \neq j$, $\dfrac{\partial^3 F}{\partial \mu_j \partial \mu_{j'} \partial \mu_{j''}} \geq 0$ for any distinct j, j', and j'' ($j'' = 1, \ldots, M$), and so on for higher-order derivatives.

Under assumptions (GEV-1) to (GEV-4), from McFadden (1978: Theorem 1), the choice probability for differentiated good j is:

(C.1) $$s_{GEVj}(p_1, \ldots, p_M, y) = \frac{\dfrac{\partial F}{\partial \mu_j} \mu_j}{\varsigma F}.$$

Extending the analysis in Sections 2 to the GEV model is straightforward. The points to note are as follows.

i) The aggregated price index in the case of one category is modified from (4) to:

(C.2) $$P_{GEV} \equiv F(\exp(\alpha_1(y) - \beta(y) p_1), \cdots, \exp(\alpha_M(y) - \beta(y) p_M)).$$

---

[19] McFadden (1978: Theorem 1) assumed homogeneity of degree one. Ben-Akiva and Francois (1983) demonstrate that H can be homogeneous of degree ς. See also Ben-Akiva and Lerman (1985: p. 126).

ii) The choice probability for differentiated good j is modified from (6) to:

$$\text{(C.3)} \quad s_{GEVj}(p_1,..., p_M, y) = \frac{\dfrac{\partial F}{\partial \exp(\alpha_j(y) - \beta(y) p_j)} \exp(\alpha_j(y) - \beta(y) p_j)}{\varsigma F}.$$

The argument of exponential function is $\alpha_j(y) - \beta(y) p_j$, which is linear in price and has the same coefficient for price.

iii) As is the same with the logit model, we cannot derive the direct utility function that perfectly corresponds to the GEV model in a closed form. Verboven (1996) derives an example of direct utility function that corresponds to the nested logit model, which is a special class of the GEV model.

iv) Regarding the elasticities, the essence remains unchanged: for the GEV model, we add the elasticities of total demand to the standard own- and cross-price elasticities when the total demand is endogenous. The own-price elasticity of differentiated good j is modified from (9) to:

$$\text{(C.4)} \quad \frac{\partial x_j}{\partial p_j} \frac{p_j}{x_j} = \theta_j - \beta(1 - \varsigma s_{GEVj}) p_j + \eta_{jj},$$

where $\eta_{jj} \equiv \dfrac{\partial \left(\dfrac{\partial F}{\partial \mu_j} \right)}{\partial p_j} \dfrac{p_j}{\left(\dfrac{\partial F}{\partial \mu_j} \right)}$ is the elasticity of $\dfrac{\partial F}{\partial \mu_j}$ with respect to the price of differentiated good j, p_j. The cross-price elasticity is modified from (10) to:

$$\text{(C.5)} \quad \frac{\partial x_j}{\partial p_k} \frac{p_k}{x_j} = \theta_k + \beta \varsigma s_{GEVk} p_k + \eta_{jk},$$

where $\eta_{jk} \equiv \frac{\partial\left(\frac{\partial F}{\partial \mu_j}\right)}{\partial p_k} \frac{p_k}{\left(\frac{\partial F}{\partial \mu_j}\right)} = -\beta(y)\mu_k \frac{\partial^2 F}{\partial \mu_j \partial \mu_k} \frac{p_k}{\left(\frac{\partial F}{\partial \mu_j}\right)} \geq 0$. The cross-price elasticity can be negative when the elasticity of total demand is taken into account. Note that the cross-elasticity is always positive, if the elasticity of total demand for a choice set is not taken into account.

v) Welfare analysis is the same as that for the logit model, except that the aggregated price index and the log-sum term is modified from (4) and (16) to (C.2) and

(C.6) $LS_{GEV} \equiv \ln P_{GEV} = \ln F(\exp(\alpha_1(y) - \beta(y)p_1), \cdots, \exp(\alpha_M(y) - \beta(y)p_M))$.

vi) Extending to the case of multiple categories of differentiated goods and many consumers in the GEV model is totally analogous to the analysis of the logit model in Sections 3 and 4.

Appendix D. The mixed logit model

Following Train (2009: pp. 134–137), the mixed logit model is a model that has the following choice probability:

(D.1) $s_{MLj} = \int_{\gamma} s_j(p_1, ..., p_M, y(\gamma), \gamma) f(\gamma) d\gamma$,

where γ is the consumer's parameter, $y(\gamma)$ is the consumer's income that has the parameter of γ, $f(\gamma)$ is a probability density function of γ, and $s_j(p_1, ..., p_M, y(\gamma), \gamma)$ is the logit-type choice probability for a consumer with the parameter γ.

In (D.1), we allow for heterogeneity not only in consumer income, $y(\gamma)$, but also in the parameter, γ.

Extending this definition to include endogenous total demand, we define the mixed logit model as the model in which the expected demand function of differentiated good j has the form of:

(D.2)
$$\begin{aligned} E(x_j) &= \int_\gamma x_g(p_1,..., p_M, y(\gamma),\gamma) s_j(p_1,..., p_M, y(\gamma),\gamma) f(\gamma) d\gamma \\ &= \int_\gamma x_j(p_1,..., p_M, y(\gamma),\gamma) f(\gamma) d\gamma, \end{aligned}$$

where $E(x_j)$ is the consumer's expected demand function for differentiated good j.

Applying the analysis in Sections 2 to the mixed logit model is straightforward. The points to note are as follows.

i) A consumer's expected indirect utility function has the form of:

(D.3) $$E(v) = \int_\gamma v(P(p_1,..., p_M, y(\gamma),\gamma), y(\gamma),\gamma) f(\gamma) d\gamma,$$

where

(D.4) $$P(p_1,..., p_M, y(\gamma),\gamma) \equiv \sum_{j'=1}^{M} \exp(\alpha_{j'}(y(\gamma),\gamma) - \beta(y(\gamma),\gamma) p_{j'}).$$

ii) The expected demand function of differentiated good j has the form of (D.2), where:

(D.5)

$$x_g(p_1,..., p_M, y(\gamma),\gamma) = x_g(P(p_1,..., p_M, y(\gamma),\gamma), y(\gamma),\gamma) = \frac{\beta(y(\gamma),\gamma) \dfrac{\partial v}{\partial P} P}{\dfrac{\partial v}{\partial y(\gamma)}}$$

and

(D.6) $$s_j(p_1,..., p_M, y(\gamma),\gamma) = \frac{\exp(\alpha_j(y(\gamma),\gamma) - \beta(y(\gamma),\gamma) p_j)}{\sum_{j'=1}^{M} \exp(\alpha_{j'}(y(\gamma),\gamma) - \beta(y(\gamma),\gamma) p_{j'})}.$$

As is the case with the logit model, (D.6) implies that the argument of exponential function is linear in price and has the same coefficient for price.

iii) The consumer's expected own- and cross-price elasticities with respect to the price of differentiated good j, p_j, are respectively:

$$\text{(D.7)}\quad \frac{\partial E(x_j)}{\partial p_j}\frac{p_j}{E(x_j)} = \int_\gamma \left\{ \frac{x_g s_j}{E(x_j)}\left(\theta_j - \beta(y(\gamma),\gamma)(1-s_j)p_j\right)\right\} f(\gamma)d\gamma ,$$

and

$$\text{(D.8)}\quad \frac{\partial E(x_j)}{\partial p_k}\frac{p_k}{E(x_j)} = \int_\gamma \left\{ \frac{x_g s_j}{E(x_j)}\left(\theta_k + \beta(y(\gamma),\gamma)s_k p_k\right)\right\} f(\gamma)d\gamma .$$

When the total demand for a choice set is endogenous, we add the price elasticity of total demand, as in the logit model. This can make the cross-price elasticity positive in the mixed logit model. A change in total demand is necessary to represent the gross-complementary relationship between differentiated goods, although the mixed logit model can deal with complex substitutionary patterns. For example, the IIA property, which is a feature of the logit model, does not hold because the cross-price elasticities in (D.7) depend on $\frac{x_g s_j}{E(x_j)}$, which differs across differentiated goods.

iv) The consumer's expected compensating variation, $E(cv)$, generally can be written as:

$$\text{(D.9)}\quad E(cv) = \int_\gamma \left(\int_{p_j^W}^{p_j^{WO}} \sum_{j'=1}^{M} \left(x_{j'}^c(p_1,\ldots,p_M, v^{WO}(\gamma),\gamma)\frac{\partial p_{j'}}{\partial p_j}\right) dp \right) f(\gamma)d\gamma .$$

If the aggregate price index, (D.5), is independent of income, we can calculate the consumer's expected compensating variation using the aggregated price index or the log-sum term:

(D.10)

$$E(cv)=\int_{\gamma}\left(-\frac{1}{\beta(y(\gamma),\gamma)}\int_{P_{ML}^{W}(p_1,...,p_M,\gamma)}^{P_{ML}^{WO}(p_1,...,p_M,\gamma)}\frac{x_g^c(P(p_1,...,p_M,\gamma),v^{WO}(\gamma),\gamma)}{P(p_1,...,p_M,\gamma)}dP(p_1,...,p_M,\gamma)\right)f(\gamma)d\gamma$$

(D.11)

$$E(cv)=\int_{\gamma}\left(-\frac{1}{\beta(y(\gamma),\gamma)}\int_{LS_{ML}^{W}(p_1,...,p_M,\gamma)}^{LS_{ML}^{WO}(p_1,...,p_M,\gamma)}x_g^c(P(p_1,...,p_M,\gamma),v^{WO}(\gamma),\gamma)dLS(p_1,...,p_M,\gamma)\right)f(\gamma)d\gamma$$

.

Eqs. (D.9)-(D.11) respectively are the expected values of the second equation of (11), (15), and (17) concerning γ.

v) We can straightforwardly extend the above analysis to cases of multiple categories of differentiated goods.

vi) In the mixed logit model with multiple consumers, we derive a different result from the logit model regarding the form of the expected indirect utility function. As is the same with the analysis for the logit model in Section 4, the aggregated price index must be independent of income to take the Gorman form in the case of a multiple-consumer economy. In addition, the coefficient of income needs to be common among consumers; this implies that the coefficient of income cannot depend on the aggregated price index, which is a function of γ. These restrictions make consumer i's expected indirect utility function the quasilinear form, which is a special form of the Gorman form:

(D.12) $E(v^i)=\int_{\gamma^i}\left(A^i(P^i(p_1,...,p_M,\gamma^i),\gamma^i)+\overline{B}y^i(\gamma^i)\right)f(\gamma^i)d\gamma^i$,

where $\overline{B}$ is a constant. As a result, the compensating variation, the equivalent variation, and a change in consumers' surpluses coincide.

vii) The above analysis on the mixed logit model can be generalized to the analysis on the mixed GEV model, applying Appendix B.

Appendix E. The brand name of the goods

Ketchup

Good 1: Kagome tomato ketchup tube 500g

Good 2: Delmonte tomato ketchup value bottle 800g

Good 3: Delmonte tomato ketchup tube 500g

Good 4: Kagome tomato ketchup tube 300g

Good 5: CGC tomato ketchup 500g

Good 6: Heintz tomato ketchup reverse bottle 460g

Mayonnaise

Good 1: Kewpie mayonnaise 450g

Good 2: Kewpie Kewpie half 400g

Good 3: Ajinomoto pure select mayonnaise 400g

Good 4: Kewpie mayonnaise 350g

Good 5: Ajinomoto pure select koku-uma 65% calorie cut 360g

Good 6: CGC mayonnaise 500g

Good 7: Kewpie tartar sauce 155g

Good 8: Kewpie Kewpie half 300g