Measures to Promote Green Cars: Evaluation at the Car Variant Level *

Taiju Kitano†
Graduate School of International Management, Aoyama Gakuin University

May 2016

Abstract

Automobile firms commonly offer multiple variants for each of their car models. Heterogeneity at the variant level is an important element to be considered when assessing attribute-based policy interventions, such as tax incentives and subsidy for green cars, because of substantial variant-level heterogeneity in the attributes within a model. This paper presents a discrete choice model of product differentiation at the variant level, and estimates the structural parameters of the econometric model using data at different levels of aggregation: model-level sales and variant-level prices and attributes. Using these estimates, this paper examines the policies to promote green cars in Japan.

Keywords: Discrete choice model; Green cars; Car variants; Tax incentives; Subsidy

JEL Classification: F18; L62; Q56

*This paper is conducted as part of the project "Basic Research for a New Industrial Policy," undertaken at the Research Institute of Economy, Trade and Industry (RIETI). The author would like to thank Yasusada Murata, and the seminar participants at Sophia University, RIETI, Prefectural University of Kumamoto, Aoyama Gakuin University, and University of Tokyo, and conference participants at the 5th Meeting on Applied Economics and Data Analysis for their helpful comments. All remaining errors are my own.

†4-4-25 Shibuya, Shibuya-ku, Tokyo 150-8366, JAPAN. E-mail: kitano@gsim.aoyama.ac.jp

1 Introduction

Recent empirical studies examining the existing policies in automobile markets have employed discrete choice analyses to estimate the multi-product demand function. The demand estimates allow researchers to assess the impact of policy interventions by conducting counterfactual simulation under certain supply-side assumptions. The dataset used in this type of discrete choice analysis comprises market-level information: data on the sales, prices, and characteristics for each alternative in a choice set. In automobile markets, an alternative in the choice set is usually a car model (nameplate) because the sales data are available at the model level. The problem here is that the prices and charateristics data are available at a finer level of aggregation, namely the variant level. Thus, a proper method to associate the model level data with the variant level data is necessary to construct the database for the discrete choice analyses.

Since the variant-level differences in prices and attributes are substantial, the method used to construct the database is an important element when assessing attribute-based policy interventions, such as tax incentives and subsidies for the promotion of green cars. The standard method used to construct the database is to identify a base variant for each model and match its price and characteristics with its sales at the model level (e.g., **?**; **?**; **?**). This is the case of *polar weighting*, as the weight of the base variant is set to be one. The polar weighting is problematic when assessing attribute-based policy interventions because the outcome of counterfactual simulation depends solely on the characteritics of the base variants. In the case of Japanese automobile market, the amount of tax reduction and subsidies can be different across the variants of a model. In fact, a substantial number of models exist whose baseline variants are out of (within) the policy target and other variants are within (out of) it. With respect to these models, the effects of policy determined by the characteristics of the base variants, although the policy had different effects on the variants other than the base. Apart from the polar weighting, a few studies employ another method: matching the average prices and attributes of the variants with their model-level sales (e.g., **?**). This is the case of *equal weighting*, as the share of each variant of a model is assumed to be the same. The equal weighting is clearly a challenge because this rules out the possibility that the policy shifts the demand between the variants in the policy target and those out of it. Another problem is that most car models often have a luxury variant, which is much more expensive than the other variants, and thus is rarely chosen.[1]

[1]The sole exception is **?**, who estimates the variant-level demand using the variant-level sales data. However, the relevant data are rarely available, as in the Japanese automobile market. As subsequently explained, this study therefore proposes a method to analyze the variant-level demand without the variant-level sales data.

This study develops a method to estimate variant-level demand using the data at different levels of aggregation, in order to overcome the problems regarding the construction of the database. The model introduced in this study is based on a two-level nested logit model, where consumers choose a car model in the first level and one of its variants in the second level. I then derive the model-level, that is, nest-level demand function, in which the variant-level prices and characteristics information are incorporated as a logit-inclusive value. The key assumption in this econometric analysis is that an unobserved demand characteristic or shock is common among the variants of a car model. Then, the parameters of a model-level demand function can be estimated based on the moment condition on the unobserved characteristics, as in **?** and **?**. Note that the variant-level demand can be obtained with the parameter estimates. Therefore, the model-level prices and characteristics can be derived as the weighted average over the variants of the model: this contrasts with the polar and equal weighting employed in the previous studies. In addition, the model allows me to address the substitution among variants of a model induced by the measures to promote green cars. In this study, I conduct counterfactual simulation to assess the effects of the measures.

This paper contributes to the literature that explores ways to estimate demand based on the discrete choice models under various data availability conditions. **?** introduces a method to incorporate the information of the pattern of car purchases across demographics in the estimation of demand using the discrete choice model, to obtain more precise demand estimates. Similarly, **?** proposes a method to use both individual-level and market-level data to identify demand parameters in the discrete choice model. While these studies indicate the ways to use richer information, **?** propose a method to estimate the demand system under data limitation where market-level quantity data are unavailable but market share ranks data are available. This study presents a method to estimate demand under a different type of data limitation, wherein the price and sales data are available at different levels of aggregation.

The rest of this paper is organized as follows. Section 2 describes the fuel economy standard and tax system in Japan, the key elements to understand the measures to promote green cars in the country. Section 3 presents the data used in this analysis and explains the variation in attributes that causes the effects of measures to differ across the variants of a model. Section 4 introduces a discrete choice model of product differentiation at the car variant level. Section 5 explains the estimation procedure and provides the estimation results. Section 6 reports the simulation results for policy assessment. Section 7 concludes the paper.

2 Green cars and tax systems

There are two types of green cars in Japan. One is next generation cars which are environmental friendly and based on fuels other than gasoline and includes clean diesel, plug-in hybrid, electronic, fuel cell and natural gas vehicles.

3 Fuel economy standards and tax systems

Fuel economy standards are key to understanding the automobile tax system and tax incentives introduced in Japan. I first introduce the fuel economy standards prevalent during the study period, April 2012 to March 2014.

From Table 1, the Japanese fuel economy standards are based on car weights and fuel economies. The standards during the study period are based on the 2015 target, which specifies the average fuel economy level the newly sold cars of every car manufacturer should exceed until March 2015. Thus, the standards act as the Corporate Average Fuel Economy regulation of the United States, although there is no explicit penalty for the violation of standards in the Japanese market. Instead, the government provides car manufacturers with incentives to meet the standards through the automobile tax systems and tax incentives and subsidy measures, as will be discussed in the following sections.

3.1 Automobile related taxes

A variety of automobile related taxes exists in Japan. At the time of purchase, car users have to pay a 5% acquisition tax, in addition to the 5% consumption tax.[2] During the ownership of cars, users have to pay a tonnage tax and automobile/mini-vehicle taxes on a yearly basis. The amount of tonnage tax for each car depends on the car weight and changes over time, as shown in Table 2. Following the change in tonnage tax system from May 1, 2012, fuel efficiency of cars has become the factor determining the amount of tax: the amount of tax for vehicles complying with the 2015 fuel efficiency standards was 2500 JPY ($\approx$ 25 USD)/year, while that for other vehicles was 4100 JPY/year.

The tonnage tax is assessed at the time of purchase of new cars and at every car inspection. For each occasion, the car owners have to pay tax for the period until the next car inspection. For example, car owners have to pay tonnage tax for three years at the time of

[2] To be more precise, acquisition tax is imposed on the tax base for each car model; this is usually around 90% of the new car prices. Thus, the real tax rate is 4.5%. In addition, this tax is exempted if the tax base is less than .5 million JPY; this exemption is virtually irrelevant to this study because none of the new car prices is below this tax base. The exemption matters for cheaper used cars, which are out of the scope of this study.

purchase, because the first car inspection is three years after purchase; similarly, at the time of first inspection, they have to pay tax for two years because the second car inspection is two years after the first inspection.

The amount of automobile tax depends on the vehicles' engine capacity; a mini-vehicle tax is imposed if the car is categorized as a mini vehicle; that is, cars with (1) engine displacement of $\leq$ 660cc and with (2) length $\leq$ 3.4m, width $\leq$ 1.48m, and height $\leq$ 2.0m. As shown in Table 2, mini-vehicle owners receive favorable treatment in terms of tax payment compared to owners of standard-sized cars (>660cc): the minimum automobile tax is 29500 JPY, whereas the tax for mini-vehicles is only 7200 JPY.

Note that the original tax system already shows a favorable treatment for green cars; however, the tax incentives explained in the following section further strengthen the motivation to purchase green cars.

3.2 Tax incentives

In order to promote fuel-efficient vehicles, the Japanese government has employed tax incentives, as summarized in Table 3. The first column of the table shows that green cars (electric vehicles and gas and hybrid vehicles complying with the fuel-efficiency and emission standards) are eligible for tax reduction. During the study period, most of the car models meet the emission standards; hence, the fuel-economy standards are the key to qualifying for tax exemption or reduction.

3.3 Subsidies for green cars

In addition to tax incentives, the Japanese government introduced a subsidy measure to promote green cars in April 2012. The measure was initially scheduled to end by February 2013, but it ended in September 2012 because of exhaustion of the budget for subsidy: the budget was set at 270 bil. JPY. The subsidy was provided to consumers who purchased a car complying with the 2015 fuel economy standards, or the 2010 fuel economy standards +20%. The amount of subsidy per unit is 100,000 JPY for standard-sized cars and 70,000 JPY for mini-vehicles.

4 Data and variant-level heterogeneity

4.1 Data

The data are on a monthly basis, covering the period from April 2012 to March 2014. I collect the variant-level prices and characteristics from *Goo-net*, a used car website operated by PROTO Corporation.[3] The prices given are the listed prices, although it is known that car dealers usually give a discount to their customers through price negotiations. Unfortunately, the transaction prices data are not available and hence I use the listed prices in the following analyses. The model-level data of sold vehicles are obtained from *Jidousha Touroku Tokei Jouhou: Shinsha-hen* (*New Car Registration Statistics*) published monthly by the Japan Automobile Dealers Association. This paper focuses on the Japanese car models that account for the majority of the market.

4.2 Variant-level heterogeneity

The data from *Goo-net* show substantial heterogeneity across the variants of a model. The number of variants in April 2013 is 1443, while the number of models is 147; thus, every model has about 10 variants on average. Table 4 shows the variant-level heterogeneity in prices and attributes for the 147 models. The first row of the table indicates the substantial price differences: the standard deviation in prices is 0.326 mil. JPY, and the difference between the maximum and minimum prices is 0.906 mil. JPY, which indicates a 40% difference. The variant-level heterogeneity in car characteristics is moderate compared to that in prices, but the difference between the minimum and maximum is large for some characteristics. In particular, a large difference exists in the fuel economy of a model, indicating that the promotion measures can have different effects across the variants of a model.

4.2.1 Case of Subaru IMPREZA

The problem of associating model-level sales data with variant level prices and attributes data can be easily understood from examples of particular models. I here introduce the case of IMPREZA supplied by Subaru. According to the sales data, IMPREZA is one model; for exapmle, its sales is 5262 units in April 2013. On the other hand, according to the *Goo-net*, the data source of prices and attributes, there are four "models" of IMPREZA: IMPREZA, IMPREZA G4, IMPREZA XV, and IMPREZA Sports, and each of them has variants. In total, these models contain 12 variants. Figure * shows the situation of IMPREZA with the information on prices and some key attorubtes of the variants. The price difference is 8000

[3] "http://www.goo-net.com/catalog/" (in Japanese).

USD at maximum. The weights and fuel economies are also different across variants, which implies that the effects of the environmental policies can be different across models as the fuel economy standard is based on the weight and the fuel economies of the cars. In fact, in the case of IMPREZA, 8 variants out of 12 comply with fuel economy standards, while the rest of them does not. In particular, the base variant, namely the cheapest variant does not meet the fuel economy standard and is out of the environmental policies. Therefore, If the polar weighting method is applied as in the previous literature, the effects of policies on IMPREZA sales are predicted to be negative as the counter-factual experiment based on the structural econometric models assumes that IMPREZA is out of the policy target. This is clearly problematic because the policies should have positive impacts on the variants complying with the fuel economy standard in reality, though this possiblity is excluded through the contruction of the database.

4.2.2 Variant-level heterogeneity in fuel efficiencies at the market level

The case of IMPREZA is not specific one. This problem is observed at the market level. To examine the variant-level heterogeneity in the effects of tax incentives and subsidies, I first compute the share of the variants meeting the 2015 Fuel Economy Standards for each car model. The share of a model j is calculated as follows.

$$S_j^{FS} = \frac{N_j^{FS}}{N_j},$$

where N_j^{FS} is the number of variants meeting the standards within the model j, and N_j is the number of variants of the model j. In addition to this, I also compute the shares for the cases of the fuel economy standards +10% and +20% because the tax incentive measures are based on these values, as shown in Table 3. They are denoted as S_j^{FS+10} and S_j^{FS+20}, respectively. Then, I construct the distributions of the shares, S_j^{FS}, S_j^{FS+10} and S_j^{FS+20}, over the models. The left bars in Figure 1 show the proportional frequencies of the range of S_j^{FS} specified in the values of horizontal axis, while the middle and right bars shows those of S_j^{FS+10} and S_j^{FS+20}, respectively. The figure shows that a substantial fraction of models take the values of S_j^{FS}, S_j^{FS+10} and S_j^{FS+20} between 0 and 0.1 or between 0.9 and 1. This indicates that there are a number of models whose variants are similar in the sense that most variants meet or do not meet the fuel economy standards within the models. For these models, the effects of the attribute-based policy intervetions, namely the tax incentives and subsidy for the green cars, are common across the variants.

Note that there are a non-negligible share of models whose variants are dissimilar in terms of the qualification of the fuel economy standards: with respect to S_j^{FS}, 33% of the models lie in the range from 0.1 to 0.9. The variant-level heterogeneity indicates that the effects of the attribute-based policy interventions are different across variants of the models. Thus, the method to associate the model-level sales with the variant-level prices and characteristics must be chosen precisely when assessing such policy interventions, as the case of IMPREZA indicates.

4.2.3 Case of Mazda AXELA

There is a problem on the assessment on the next generation cars. The other example is AXELA marketed by Mazda. As is the case of IMPREZA, AXELA is one model in the sales data, while it has several models in *Goo-net*: AXELA, AXELA Sports and AXELA Hybrid each of which has multiple variants. A problem here is that there are variants with different engine types: gasoline, hybrid and clean diesel engines. Since the sales data contains one value of the AXELA sales, the sales of clean diesel cars, i.e. one of the next generation cars, can not be identified from the data. Under the polar weighting, AXELA sales are not counted as the sales of clean diesel cars because the cheapeset variant is not equipped with the clean diesel engine. This is clearly problematic and thus the variant-level analyses are necessary in order to assess the impacts of policies on the next generation cars.

5 Demand

The demand side of the market is modelled in a discrete choice framework. The set of models supplied in a market at time t is $\mathcal{J}_t$, where each car model $j \in \mathcal{J}_t$ has the set of variants $\mathcal{B}_j$. In addition to the option to purchase one from the set of inside option, i.e. $\cup_{j \in \mathcal{J}_t} \mathcal{B}_j$, each consumer can choose an outside option, i.e. not to purchase a new car. Given these possible choices, every consumer chooses an alternative that gives the highest utility.

Consumer i's (indirect) utility obtained from model j with variant $n \in \mathcal{B}_j$ is specified as (hereafter, the time subscript t is suppressed)

$$u_{ij_n} = \delta_{j_n} + \mu_{ij_n} + \epsilon_{ij_n}, \tag{1}$$

where $\delta_{j_n} + \mu_{ij_n}$ is the deterministic part of the utility, which is the function of car attributes and individual characteristics. δ_{j_n} captures the mean evaluation of variant $n \in \mathcal{B}_j$ common to all consumers, and μ_{ij_n} is the individual-specific evaluation of the variant. As is common in the literature, the deterministic part of utility obtained from the outside option is normalized to zero; that is, $\delta_0 + \mu_{i0} = 0$.

δ_{j_n} is further decomposed as:

$$\delta_{j_n} = \delta_j + \Delta\delta_{j_n}, \tag{2}$$

where

$$\delta_j = \sum_{k \in \mathcal{K}_j} x_{jk}\beta_k + \xi_j, \text{ and } \Delta\delta_{j_n} = \sum_{k \in \mathcal{K} \setminus \mathcal{K}_j} x_{j_n k}\beta_k. \tag{3}$$

$\mathcal{K}$ is the set of car characteristics (incl. a constant term) and $\mathcal{K}_j \subseteq \mathcal{K}$ is the set of characteristics common to all the variants of model j. x_{jk} is the value of characteristics common across variants whereas $x_{j_n k}$ is the value of characteristics different across variants. Coefficients β_k are parameters to be estimated. Note that $\mathcal{K}_j$ can differ from model to model; that is, $\mathcal{K}_j \neq \mathcal{K}_{j'}$ for $j' \neq j$. In addition, x_{jk} includes the constant term and thus $\mathcal{K}_j$ never be a null set.

ξ_j represents an unobserved characteristic and a demand shock specific to model j. The key assumption here is that unobserved characteristics are not allowed to vary across variants but has to be common across all variants of the model. This might be a reasonable assumption, in that the variants usually share the same design and demand shock. As explained in the following section, this assumption is necessary for my estimation strategy.

μ_{ij_n} is also decomposed into the parts of utilities common and different across variants; that is,

$$\mu_{ij_n} = \mu_{ij} + \Delta\mu_{ij_n}, \tag{4}$$

where,

$$\mu_{ij} = \sum_{k \in \mathcal{K}_j} x_{jk}\nu_{ik}\sigma_k, \text{ and } \Delta\mu_{ij_n} = \alpha_i p_{j_n} + \sum_{k \in \mathcal{K} \setminus \mathcal{K}_j} x_{j_n k}\nu_{ik}\sigma_k. \tag{5}$$

Note that $\alpha_i \equiv \alpha / y_i$ is the price sensitivity of consumer i inversely proportional to consumer i's income y_i. α is a parameter to be estimated and p_{j_n} is the expenditure on the purchase of variant $n \in \mathcal{B}_j$; this consists of the price and automobile-related taxes, including the tax incentives. ν_{ik} is consumer i's specific taste of characteristic k, which is assumed to follow a standard normal distribution. σ_k is a parameter to be estimated; this captures the standard deviation on the individual-specific taste of characteristic k.

ϵ_{ij_n} represents consumer i's idiosyncratic taste of variant n of model j and is assumed to follow a Generalized Extreme Value (GEV) leading to the following choice probability:

$$s_{ij_n} = s_{ij} \cdot s_{in|j}. \tag{6}$$

Here, s_{ij_n} is the probability of consumer i choosing variant n of model j, which is a product of the choice probability of model j,

$$s_{ij} = \frac{e^{\delta_j + \mu_{ij} + \lambda I_{ij}}}{1 + \sum_{l \in \mathcal{J}} e^{\delta_l + \mu_{il} + \lambda I_{il}}}, \tag{7}$$

and the choice probability of variant n conditional on choosing model j,

$$s_{in|j} = \frac{e^{(\Delta\delta_{jn} + \Delta\mu_{ijn})/\lambda}}{\sum_{m \in \mathcal{B}_j} e^{(\Delta\delta_{jm} + \Delta\mu_{ijm})/\lambda}}. \tag{8}$$

Note that I_{ij} is a logit-inclusive value specified as

$$I_{ij} = \ln\left(\sum_{n \in \mathcal{B}_j} e^{(\Delta\delta_{jn} + \Delta\mu_{ijn})/\lambda}\right). \tag{9}$$

λ is a parameter to be estimated, and, as shown in **?**, has to lie between 0 and 1 to be consistent with the utility maximization problem. As λ approaches 0, the dependency across variants becomes stronger, and, in the limit, that is, $\lambda \to 0$, the model converges to the elimination-by-aspect model of **?**. On the other hand, as λ approaches 1, the dependency becomes weaker and the model reduces to the single-level logit model at $\lambda = 1$. I will statistically test whether the estimate of λ is located within the interval.

The model-level share function can be derived by integrating the individual choice probability in Eq.(7) over the distribution on $\boldsymbol{\nu}_i = (\nu_{ik})_{k \in \mathcal{K}}$ and y_i. ν_i is assumed to follow a standard normal distribution, whereas y_i follows the empirical income distribution obtained from *Kokumin Seikatsu Kiso Chosa* (*Comprehensive Survey of Living Conditions of the People on Health and Welfare*) released annually by the Ministry of Health, Labor and Welfare. Now, the share of the model, j, can be calculated as

$$s_j = \int_y \int_\nu s_{ij} dF_\nu(\boldsymbol{\nu}) dF_y(y), \tag{10}$$

where $F_\nu(\cdot)$ is the cumulative standard normal distribution and $F_y(\cdot)$ is the cumulative empirical income distribution.

The variant-level share function can be derived in a similar manner: the share of variant $n \in \mathcal{B}_j$ is

$$s_{j_n} = \int_y \int_\nu s_{ij_n} dF_\nu(\boldsymbol{\nu}) dF_y(y). \tag{11}$$

In order to estimate the demand-side parameters, I focus on the model-level share function because the variant-level sales data are unavailable. As explained in the following section, it is possible to apply a BLP-type contraction mapping method based on the model-level share function and estimate the parameters from the moment condition on ξ_j.

6 Estimation

6.1 Simple case: no random coefficient

For simplicity, I first explain the case of no random coefficient—common price sensitivity, that is, $\mu_{ij} = \alpha p_{jn}$, and no individual heterogeneity on car characteristics, that is, $\sigma_k = 0$ for all k. Now, the individual share function in Eq.(**??**) becomes the market share function, namely, $s_{ij} = s_j$, and the following equation can be derived:

$$\ln(s_j) - \ln(s_0) = \sum_{k\in\mathcal{K}_j} x_{jk}\beta_k + \lambda \ln\left(\sum_{n\in\mathcal{B}_j} e^{\left(\alpha p_{jn} + \sum_{k\in\mathcal{K}\setminus\mathcal{K}_j}\beta_k x_{j_n k}\right)/\lambda}\right) + \xi_j. \tag{12}$$

As the above expression clearly shows, the equation is linear in unobserved characteristics, ξ_j, and thus the parameters in the utility function $\boldsymbol{\theta} = (\alpha, (\beta_k)_{k\in\mathcal{K}}, \lambda)$ can be estimated using the non-linear estimation method. If ξ_j is uncorrelated with the variables in the equation, the set of parameters can be estimated using non-linear least squares. However, as commonly discussed in the literature, the unobserved characteristics are likely to be correlated with the prices, $p_{j_n}, n \in \mathcal{B}_j$; therefore, certain moment conditions on ξ_j are needed to estimate the parameters.

6.1.1 Moment condition

The model is estimated on the basis of a moment assumption on ξ_{jt} representing the unobserved demand shock and characteristics. A problem here is that ξ_{jt} should be correlated with p_{jt} because the positive unobservable characteristics or demand shocks induce higher prices. In this paper, I use the set of instruments based on the moment condition $E[\xi_{jt}|\mathbf{x}_{1t}, \ldots, \mathbf{x}_{\#J_t t}] = 0$ for all j; this is often used in the literature.

Note that this study deals with rich information on the characteristics because each model usually has multiple variants. Using this variation in characteristics in a model, I set the mean and standard deviation across the variants in the model as the instruments for estimation. These variables are valid instruments because the mean and standard deviation of the characteristics are correlated with the prices, while these are uncorrelated with the error term under the moment assumption.

Using this set of instruments, I implement the two-step efficient generalized method of moments (GMM) estimation proposed by **?**. Now, the non-linear search becomes

$$\min_{\boldsymbol{\theta}} \mathbf{g}(\boldsymbol{\theta})' W \mathbf{g}(\boldsymbol{\theta}), \tag{13}$$

where $\mathbf{g}(\boldsymbol{\theta}) = \boldsymbol{\xi}(\boldsymbol{\theta})'\mathbf{Z}/N$ and $\mathbf{W}$ is the weighting matrix. The choice of initial weighting matrix is $(\mathbf{Z}'\mathbf{Z})^{-1}/N$, where $\mathbf{Z}$ is the IV matrix and N is the number of observations. The efficient weighting matrix is computed from the estimation results in the first stage: $Z'\hat{\boldsymbol{\xi}}\hat{\boldsymbol{\xi}}Z/N$, where $\hat{\boldsymbol{\xi}}$ is the vector of residuals obtained in the first stage.

6.1.2 Identification of λ

Note that if $\mathcal{B}_j$ is a unit set, parameter λ disappears from Eq.(**??**) and the equation reduces to the estimation equation derived from the standard logit model. This indicates that the presence of λ is the key difference between the model introduced in this paper and the standard model. Thus, I need to examine what variation of the data allows me to identify the parameter λ. To understand the identification issue, consider the simplified case in which all variants have the same price and characteristics; that is, $p_{j_n} = p_j$ for all $n \in \mathcal{B}_j$ and $\mathcal{K} = \mathcal{K}_j$ for all $j \in \mathcal{J}$. Now, the second term on the RHS of Eq.(**??**) becomes $\alpha p_j + \lambda \ln(N_j)$, where N_j is the number of variants of model j. This clearly indicates that λ can be identified in the presence of difference in number of variants over models. Without this simplification, the variant-level difference in prices and characteristics over models would contribute to identifying λ, in addition to the number of variants.

6.2 Random coefficient

I now turn to the estimation of the model in the presence of random coefficients in the utility function. The estimation incorporates a well-known contraction mapping procedure proposed by Berry, Levinsohn, and Pakes (1995).

To apply BLP's contraction mapping, I first define the following:

$$\bar{\delta}_j \equiv \sum_{k \in \bar{\mathcal{K}}} x_{jk}\beta_k + \xi_j \text{ and } \bar{I}_{ij} \equiv \frac{\sum_{k \in \mathcal{K}_j \setminus \bar{\mathcal{K}}} x_{jk}\beta_k + \mu_{ij}}{\lambda} + I_{ij}, \tag{14}$$

where $\bar{\mathcal{K}} = \bigcap_{j \in \mathcal{J}} \mathcal{K}_j$, the set of characteristics common to all variants for every model. Then, the individual choice probability on model j in eq.(**??**) can be rewritten as

$$s_{ij} = \frac{e^{\bar{\delta}_j + \lambda \bar{I}_{ij}}}{1 + \sum_{l \in \mathcal{J}} e^{\bar{\delta}_l + \lambda \bar{I}_{il}}}. \tag{15}$$

Assume that $\boldsymbol{\theta}_1 = (\beta_k)_{k\in\bar{\mathcal{K}}}$, the vector of parameters in $\bar{\delta}_j$, and $\boldsymbol{\theta}_2 = (\alpha, (\beta_k)_{k\in\mathcal{K}\setminus\bar{\mathcal{K}}}, (\sigma_k)_{k\in\mathcal{K}}, \lambda)$, the vector of parameters in $\bar{I}_{ij}$. For an arbitrary $\boldsymbol{\theta}_2$, the common utility part $\bar{\boldsymbol{\delta}} = (\bar{\delta}_j)_{j\in\mathcal{J}}$ that makes $s_j = s_j(\bar{\boldsymbol{\delta}}; \boldsymbol{\theta}_2)$ for all $j \in \mathcal{J}$ can be obtained by computing the following series:

$$\bar{\boldsymbol{\delta}}^{h+1} = \bar{\boldsymbol{\delta}}^{h} + \ln s_j - \ln s_j(\bar{\boldsymbol{\delta}}^{h}; \boldsymbol{\theta}_2), \tag{16}$$

where superscript h indicates the number of iterations. Convergence is achieved if $||\bar{\boldsymbol{\delta}}^{h+1} - \bar{\boldsymbol{\delta}}^{h}||$ becomes smaller than a certain tolerance level.

Likewise the case of no random coefficient, the parameters $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ can be estimated by solving the problem specified in eq. (**??**), though the contraction mapping procedure is incorporated here.

6.3 Estimation results

In this section, I report the estimation results. The results of no random coefficient are shown in Table 6. Table 6 (i) and (ii) give the results for the cases of non-linear least squares (NLS) and non-linear GMM estimation. As mentioned previously, the prices and unobserved characteristics are likely to be positively correlated, thus inducing an upward bias in the estimation of the price coefficient. Along with this argument, Table 6 (i) and (ii) show that the price coefficient α becomes lower after instrumenting. The estimate of λ lies between 0 and 1; this indicates that the estimates are consistent with the random utility maximization problem. Because the estimate of λ is significantly different from 1, I reject the logit model: the substitution between variants within a model is stronger than that between the variants across models. Most of the other coefficients have a reasonable sign; for example, the coefficient of Fuel Cost is negative and significant. The sole exception is the coefficient of Cruise Control, which is expected to be negative; the results show negative estimates, but it is not statistically significant in (ii).

Next I turn to the random coefficient model. The result of GMM estimation is shown in Table 7. Here, Constant and Car Space are allowed to have random coefficients. First, the price coefficient is statistically significant. Second, most of the mean parameters are reasonably estimated. The exception is Cruise Control whose coefficient takes a negative value; however, this is not statistically significant. Third, the estimates of standard deviations are small, but their standard errors are huge. This is clearly problematic because the preference on car space and the tendency to purchase new cars should be different across consumers.

7 Simulation

Using the estimates in Section 5, I carry out counterfactual simulation to assess the impacts of the tax incentives from April 2012 to March 2014. I do not incorporate the supply side, but simply construct the counterfactual price, $p^c_{j_n}$, by adding the amount of tax reduction to the actual price, p_{j_n}. Given $p^*_{j_n}$, I compute the counterfactual sales based on the demand estimates.

The simulation results are summarized in Table 8. I focus on the effects on green cars, namely, plug-in hybrid and clean diesel cars and cars complying with the 2015 fuel economy standards. The second column of the table shows the actual sales; that is, the sales with policy. Because the data of variant-level sales are unavailable, the values are computed from the demand estimates. As shown in Table 8 (i), subsidies and tax incentives increased the green car sales by 4–4.6% in fiscal years 2012 and 2013. The effects are moderate, but those on clean diesel and plug-in hybrid cars are larger: the policies increased the sales by 7.7–9%, as shown in Table 8 (ii).

8 Conclusion

The variant-level heterogeneity in car markets is substantial; thus, the assessment of attribute-based policy interventions should account for the differences in effects of policy at the variant level. This paper presents a discrete choice model of product differentiation at the variant level and estimates the model's structural parameters using the data at different levels of aggregation: model-level sales and variant-level prices and attributes. From these estimates, I assess the measures to promote green cars in Japan.

The simulation results show that policies increased the sales of green cars by 4–4.6% in fiscal years 2012 and 2013. The increase in sales of plug-in hybrid and clean diesel cars are larger, namely, 7.7–9%. The results indicate that policies have some impacts on the diffusion of green cars.

Weight	2015 Fuel Economy Standards
–600kg	22.5 km/l
601–740kg	21.8 km/l
741–855kg	21 km/l
856–970kg	20.8 km/l
971–1080kg	20.5 km/l
1081–1195kg	18.7 km/l
1196–1310kg	17.2 km/l
1311–1420kg	15.8 km/l
1421– 1530kg	14.4 km/l
1531–1650kg	13.2 km/l
1651–1760kg	12.2 km/l
1761–1870kg	11.1 km/l
1871–1990kg	10.2 km/l
1991–2100kg	9.4 km/l
2101–2270kg	8.7 km/l
2271kg–	7.4 km/l

Note: The calculation of fuel economies are based on the JCO8 mode.

Table 1: Fuel economy standards

	On acquisition		During ownership		
	Acquisition tax	Consumption tax	Tonnage tax	Automobile tax	Mini-vehicle tax
Tax rate/ amount	5% of purchase price	5%	(Before April 30, 2012) - 5000 JPY/0.5t per year (After May 1, 2012) - Vehicles complying with 2015 fuel efficiency standards: 2500 JPY/0.5t per year - Other vehicles: 4100 JPY/0.5t per year	- 1000cc: 29500 JPY/year 1001 - 1500cc: 34500 JPY/year 1501 - 2000cc: 39500 JPY/year 2501 - 3000cc: 45000 JPY/year 3001 - 3500cc: 51000 JPY/year 3501 - 4000cc: 66500 JPY/year 4001 - 4500cc: 76500 JPY/year 4501 - 6000cc: 88000 JPY/year Over 6000cc: 111000 JPY/year	7200 JPY/year

Table 2: Automobile-related taxes in Japan, April 2012–March 2014

Requirement	Acquisition tax	Tonnage tax	Automobile tax
Electric vehicles (incl. fuel cell vehicles); Plug-in hybrid vehicles; Clean diesel vehicles; Natural gas vehicles	Exempt	Exempt at 1st vehicle inspection (3 years); 50% reduction at 2nd inspection (2 years)	50% reduction (1 year)
Gasoline vehicles (incl. Hybrid vehicles) with compliant + 20% compared to 2015 fuel efficiency standards, and with emissions down by 75% from 2005 standards	Exempt	Exempt at 1st vehicle inspection (3 years); 50% reduction at 2nd inspection (2 years)	50% reduction (1 year)
Gasoline vehicles (incl. Hybrid vehicles) with compliant + 10% compared to 2015 fuel efficiency standards, and with emissions down by 75% from 2005 standards	75% reduction	75% reduction at 1st vehicle inspection (3 years)	50% reduction (1 year)
Gasoline vehicles (incl. Hybrid vehicles) with compliant with 2015 fuel efficiency standards, and with emissions down by 75% from 2005 standards	50% reduction	50% reduction at 1st vehicle inspection (3 years)	25% reduction (1 year)

Table 3: Tax incentives, April 2012–March 2014

Variables	Mean	Std. Dev.	Max-Min	(Max - Min)/Min
Price (mil. JPY)	2.941	0.326	0.906	0.398
Car Size(m^3)	11.900	0.128	0.262	0.021
Wheelbase (m)	2.643	0.007	0.013	0.005
Engine Displacement (1000 cc)	1.949	0.085	0.194	0.098
Capacity (l)	53.235	0.842	1.735	0.042
HP (ps)/Weight (kg)	0.104	0.008	0.021	0.244
Weight (1000kg)	1.381	0.045	0.115	0.091
Fuel Economy (km/l)	16.216	1.293	3.211	0.229

Table 4: Variant-level heterogeneity

	(i) NLS		(ii) GMM	
Variables	Coef.	S. E.	Coef.	S. E.
Price	-0.225	0.037	-0.763	0.203
Fuel Cost	-0.256	0.016	-0.280	0.019
Car Space	0.430	0.071	0.635	0.093
Car Size	-0.048	0.040	-0.183	0.062
Engine Displacement	0.141	0.109	1.044	0.345
Horse Power/Weight	7.641	2.049	7.177	2.117
Diameter	-0.035	0.162	-0.154	0.164
4WD	0.913	0.131	1.225	0.159
FR	0.375	0.117	0.725	0.192
Cruise Control	-0.308	0.113	-0.170	0.131
Power Seat	0.023	0.142	0.720	0.257
Stability Control System	0.644	0.074	0.623	0.088
Const	-8.600	0.666	-6.900	0.875
λ	0.554	0.128	0.308	0.187

Note: Monthly dummy variables and brand dummy variables are included in the estimation.

Table 5: Estimation results: no random coefficient

	Mean(β)		Std. Dev.(σ)	
Variables	Coef.	S. E.	Coef.	S. E.
Price	-	-	-1.782	0.766
Fuel Cost	-0.249	0.030	-	-
Car Space	0.406	0.200	0.012	2.074
Car Size	-0.047	0.054	-	-
Engine Displacement	0.186	0.216	-	-
Horse Power/Weight	6.282	2.238	-	-
Diameter	0.078	0.196	-	-
4WD	0.853	0.105	-	-
FR	0.429	0.121	-	-
Cruise Control	-0.283	0.244	-	-
Power Seat	0.043	0.216	-	-
Stability Control System	0.658	0.115	-	-
Const	-8.757	2.912	0.066	6.364
λ	0.597	0.213	-	-

Note: Monthly dummy variables and brand dummy variables are included in the estimation.

Table 6: Estimation results: random coefficient

	(i) Green Cars (1000 unit)			
Year	With Policies	Without Policies	Difference	Rate of Change (%)
2012	2415	2310	105	4.56
2013	2646	2545	102	4.00

	(ii) Clean Diesel and Plug-in Hybrid (1000 unit)			
Year	With Policies	Without Policies	Difference	Rate of Change (%)
2012	23	21	2	8.96
2013	45	42	3	7.65

Note: Green cars are cars complying with 2015 fuel economy standards or clean diesel or plug-in hybrid cars.

Table 7: Effects on Green Cars

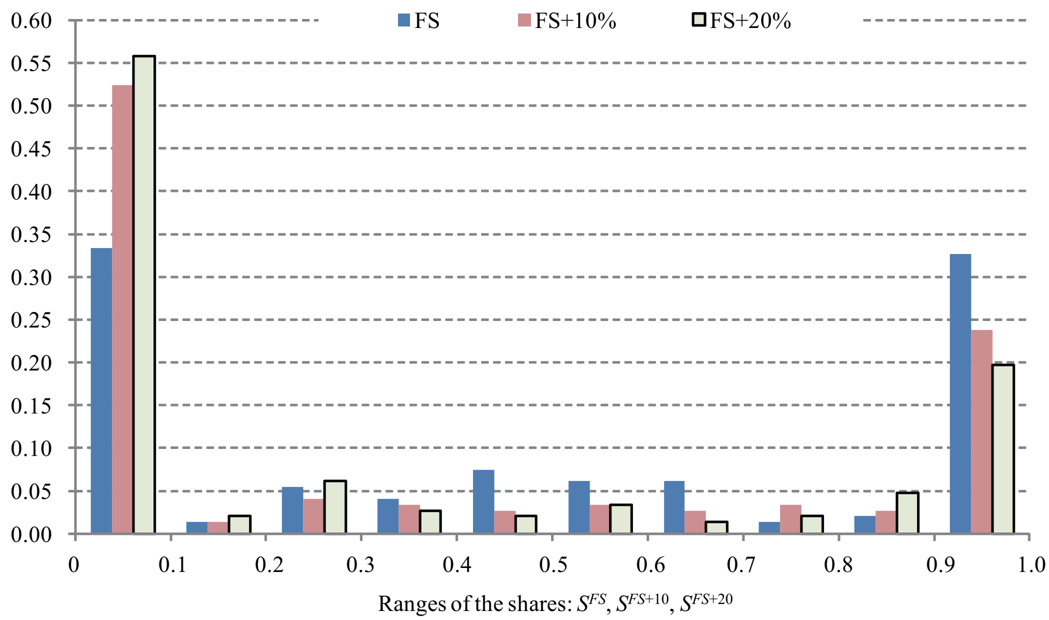

Figure 1: Shares of variants meeting 2015 Fuel Economy Standards