REPRESENTING HOUSEHOLD VEHICLE HOLDING DURATION WITH HETEROGENEOUS DISTRIBUTIONS BASED ON LATENT CLASS APPROACH

Masashi KUWANO
Assistant Professor
Graduate School of Engineering,
Hiroshima University
Kagamiyama 1-4-1, Higashi-Hiroshima,
739-8527 Japan
Fax: +81-82-424-7825
E-mail: kuwano@hiroshima-u.ac.jp

Junyi ZHANG
Associate Professor
Graduate School for International Development and Cooperation
Hiroshima University
Kagamiyama 1-5-1, Higashi-Hiroshima,
739-8529 Japan
Fax: +81-82-424-6919
E-mail: zjy@hirosima-u.ac.jp

Akimasa FUJIWARA
Professor
Graduate School for International Development and Cooperation
Hiroshima University
Kagamiyama 1-5-1, Higashi-Hiroshima,
739-8529 Japan
Fax: +81-82-424-6921
E-mail: afujiw@hirosima-u.ac.jp

Abstract: It is expected that vehicle holding durations differ substantially across households. Some households may keep their vehicles as long as possible, while others may replace some of their vehicles after holding several years. Even though existing studies have incorporated such heterogeneity using duration models, one-peak distribution is usually assumed. In reality, this assumption can be easily violated. In this sense, the heterogeneity of duration distributions has not been satisfactorily represented in existing literature. To represent the heterogeneity in the household vehicle holding duration, the paper proposes to apply a latent class modeling approach to simultaneously incorporate different duration distributions. Such modeling approach could deal with multi-peak distribution of vehicle holding duration. Using a data collected in several Japanese cities in 2006, this paper confirms the model effectiveness from both model performance and applicability. It is found that household characteristics are important factors to explain the latent classes.

Key Words: Vehicle holding duration, Heterogeneity, Multi-peak distribution

1. INTRODUCTION

In many countries, energy consumptions and pollutants from transport sector have been increasing considerably in the past decade. In Japan, for instance, fuel consumptions and CO₂ emissions in transport sector significantly increased, and as a result, under the same increase rate in the future, emission levels are expected to rapidly increase by 40% from the year 1990 to the year 2010. Considering the growing concern for global warming, this alarming rate of increase calls for measures to enhance energy efficiency and to reduce emissions in the entire transportation systems. Therefore, policy makers are interested in the improvement of transportation efficiency and consequently, they support the promotion of new technologies
and smart change of travel behavior. To understand the behavioral responses of consumers to the policies, many studies have been conducted about car ownership and use behavior. To clarify the effective measures of reducing the emissions from vehicles, it is necessary to systematically deal with car ownership behavior. It is already known that car ownership behavior is usually defined according to the stages of “car purchase”, “car usage”, “duration of car ownership”, and “renewal and destruction” (De Jong, 2004). It is expected that all of these elements interact with each other over time. In this study, we only focus on the household’s vehicle holding duration since it is an important factor to understand the consumer car ownership behavior and promote low emission vehicle, but has not been satisfactorily modeled in the existing literature.

Hazard-based duration model, which is more effective in capturing the dynamic process of vehicle ownership than traditional discrete choice analysis, is generally used to analyze the household vehicle holding duration (Gibert, 1992; Hensher, 1994; De Jong, 1996; Yamamoto and Kitamura, 2000). De Jong (1996), for example, developed a vehicle holding duration model as a part of a modeling system consisting of vehicle holding duration, vehicle type choice and usage. Yamamoto and Kitamura (2000) developed the models of actual and intended vehicle holding durations. They adopted a non-parametric approach (i.e., mass point approach) in model estimation, which requires no assumption about the distribution of the error components. On the other hand, one of the most important behavioral elements in hazard based duration model is heterogeneity (Lee and Timmermans, 2007). The problem of heterogeneity is that household-specific covariates are intended to incorporate observation-specific effects. But if the model specification is incomplete and if systematic individual differences in the duration distribution remain after the observation-specific effects are accounted for, then inference based on the improperly specified model is likely to be problematic. In the case of household vehicle holding durations, it can be expected that vehicle holding durations differ substantially among households. Some households may keep their vehicles as long as possible, while other households may replace some of their vehicles after holding several years. Parts of the variations in vehicle holding durations are due to variations across households in their intentions about vehicle holding. In this sense, heterogeneity exists in household vehicle holding duration. Even though some of the existing studies have incorporated either the observed or unobserved heterogeneity into the duration model, one-peak distribution is usually assumed. Such assumption is made simply because all of the widely applied distributions (e.g., normal, log-normal, Weibul, and Gamma distributions) have one-peak. In reality, this assumption can be easily violated. In this sense, the heterogeneity of duration distributions has not been satisfactorily represented in the existing literature. To incorporate such heterogeneity in the household vehicle holding duration model, the paper suggests applying latent class modeling approach. It would have the additional advantage that the relationship between socio-demographic variables and latent classes (i.e., different distributions) could be specified. Theoretically, latent class modeling approach can be applied to represent the same type of distribution, but with different distribution parameters, or different types of distributions. This is done by integrating latent class modeling approach under the framework of hazard based duration model. The proposed model will be examined using a data collected in Japan in 2006.

In light of these considerations, we first describe a latent class hazard model incorporating heterogeneous distributions (Section 2) and then introduce the data used in this study (Section 3). This is followed by model estimation and discussion in Section 4. Finally, the paper is concluded and some future research issues are also mentioned in Section 5.
2. MODEL DEVELOPMENT

2.1 Hazard Based Duration Model

Hazard based duration analysis has been used extensively in other research fields such as econometrics, biostatistics, medical sciences and industrial engineering. In the area of transportation, we can find some literature with regard to the time between vehicle purchases (e.g., Gibert, 1992; Hensher, 1994; De Jong, 1996; Yamamoto and Kitamura, 2000), the time between activities (e.g., Niemeier and Morita, 1996), and the time between vehicle accidents (e.g., Mannering, 1991, 1993). In this study, the duration that a vehicle is owned by a household is examined using hazard-based duration models. In the duration model, time $T$ is a continuous random variable. It measures the duration of staying at some state, for instance ownership of some vehicle. Suppose that $T$ has a continuous probability distribution $f(t)$, where $t$ is a realization of $T$. The cumulative probability $F(t)$ is given as,

$$ F(t) = \int_0^t f(s)ds = \Pr[T \leq t] $$

(1)

where, $\Pr[\cdot]$ denotes the probability that a household gives up owning the vehicle before time $t$ elapses since the time when the vehicle was purchased. The hazard function can be written as a function of the cumulative distribution function $F(t)$ and the corresponding density function $f(t)$ of the variable $t$.

$$ h(t) = \frac{f(t)}{1 - F(t)} $$

(2)

Another important construct in hazard-based models is the survivor function $S(t)$, which gives the probability of survival up to time $t$. It is related to the cumulative distribution function as

$$ S(t) = \Pr(T \geq t) = 1 - \Pr(T \leq t) = 1 - F(t). $$

(3)

Since $f(t) = -dS(t)/dt$, the hazard can also be written as

$$ h(t) = -\frac{d\log S(t)}{dt}. $$

(4)

If the hazard is known, the survivor function can be found through

$$ S(t) = \exp(-\int_0^t h(u)du). $$

(5)

And then, the density function of $t$ is expressed by

$$ f(t) = h(t)\exp(-\int_0^t h(t)dt). $$

(6)
If the distribution \( f(t) \) is known, then \( S(t) \) and \( h(t) \) can be uniquely derived. A number of distributions have been proposed and examined in earlier studies. This study examines the following five distributions for the vehicle holding duration model, i.e., 1) Exponential, 2) Weibull, 3) Gamma, 4) Log-logistic, and 5) Log-normal distributions, which density functions are shown as follows:

1) Exponential
\[
f(t \mid X) = \exp(-\beta X) \exp\left\{-t \exp(-\beta X)\right\}
\]

2) Weibull
\[
f(t \mid X) = \gamma^{t^{-1}} \exp(-\gamma t \exp(-\gamma X))
\]

3) Gamma
\[
f(t \mid X) = \frac{\gamma^\delta}{\Gamma(1/\delta^2)} \left\{ t^{\gamma\delta-1} \exp(-\gamma \delta X)\right\}^{1/\delta^2} \exp\left\{-\frac{t^{\gamma\delta} \exp(-\gamma \delta X)}{\delta^2}\right\}
\]

4) Log-Logistic
\[
f(t \mid X) = \frac{\gamma^{t^{-1}} \exp(-\beta X)}{1 + t\gamma \exp(-\beta X)}
\]

5) Log-normal
\[
f(t \mid X) = \frac{1}{\sqrt{2\pi \sigma^2}} \exp\left\{-\frac{(\ln t - \beta X)^2}{2\sigma^2}\right\}
\]

Where, \( \beta, \gamma, \delta, \sigma \) are unknown parameters that can be estimated by maximum likelihood method, and \( X \) is the vector of covariates (independent variables).

Generally speaking, different covariates (independent variables) might affect vehicle holding duration. But some of these covariates, for example, household characteristics, may change over time. Pendyala \textit{et al.} (1995) showed that the relationship between car ownership and income is not constant over time, in that the car ownerships of households might vary with their incomes. Therefore, if one is to incorporate the effects of a change in independent variables, then it is necessary to use time-varying covariates. To incorporate these time-varying covariates, let the interval \( 0 \) to \( t \) be divided into \( N \) exhaustive, non-overlapping intervals, \( t_0 < t_1 < \cdots < t_N \), where \( t_0 = 0 \) and \( t_N = t \). The covariates are assumed to stay constant within each of the \( N \) intervals, but may vary from one interval to the next. The survivor function (Eq. (5)) is rewritten as follows:

\[
S(t) = \exp\left(-\int_0^t h(t \mid X(t))\,du\right)
\]

where, \( X(t) \) denotes the time-varying covariate at time \( t \). The time-varying covariates are modeled as a step function, with different values through several intervals between \( t = 0 \) and \( t = t_N \).

\[
X(t) = \begin{cases} 
X_0 & t < t_1 \\
X_1 & t_1 \leq t < t_2 \\
X_2 & t_2 \leq t < t_3 \\
. & . \\
X_N & t_N \leq t 
\end{cases}
\]
The survivor function incorporating the time-varying covariate $X(t)$ is expressed as follows:

$$S(t | X(t)) = \prod_{n=1}^{N} \frac{S(t_n | X_{n-1})}{S(t_n | X_n)} \times S(t | X_n).$$

In this study, vehicle holding duration models are developed for the vehicles using the survey data which investigated the household car ownership from 1996 to 2006. The exact holding durations of previous cars which were purchased and got out from 1996 to 2006 can be observed. When every vehicle is observed at the start and end of holding spell, there is no problem. However, if the vehicle had been bought before the survey, the sample cannot be used since we don’t have the household and main user characteristics before 1996. Furthermore, if the vehicle is under use during the survey period, the exact holding duration is not observed. It is already known that estimating a model without these censored observations would lead to self-selection biases. In this study we incorporate left and right-censored spells to avoid selection biases. To incorporate censored spells, the following likelihood function is adopted.

$$L = \prod_{i=NC} f(t_i | X) \cdot \prod_{i=RC} S(t_i | X) \cdot \prod_{i=L} f(t_i | X) \cdot \prod_{i=LRC} S(t_i | X)$$

where, $NC$ refers to non-censored observations, $RC$ to right-censored observations, $LC$ to left-censored observations, and $LRC$ indicates left and right-censored observations.

### 2.2 Applying Latent Class Modeling Approach

Latent class modeling approach provides an attractive platform for modeling the unobserved heterogeneity. To formulate the latent duration model, it is assumed that there exist $s$ homogeneous latent classes. Let $M_{gs}$ denote the probability (or class membership probability) that household $g$ belongs to latent class $s$ ($s=1,2,\ldots,S$). This probability might be influenced by various factors. This study defines the class membership probability using socio-demographic variables of the households and car attributes. Various functional forms have been proposed to represent the membership probability. This study adopts the most convenient form, i.e., the multinomial logit-type function (e.g., Zenor and Srivastava, 1993; Swait and Sweeney, 2000), and defines the latent class membership probability as follows:

$$M_{gs} = \frac{\exp(\alpha_s + \varphi_s D_g)}{1 + \sum_{s=1}^{S} \exp(\alpha_s + \varphi_s D_g)}$$

where, $\alpha_s$ is the intercept and $\varphi_s$ ($s=1,2,\ldots,S$) is an unknown parameter vector and represents the contribution of the socio-demographic and car attribute variables $D_g$ to the probability of class membership.

Needless to say, class membership probabilities must meet the condition $\sum_{s} M_{gs} = 1$. Under latent class modeling framework, each class has the same form of the likelihood for vehicle
holding duration shown in equation (11). To differentiate from equation (11), the likelihood $L_{gs}$ for class $s$ is re-written in equation (14). Then the household’s logarithm likelihood function can be defined by summing up the products of the class membership probability ($M_{gs}$) and the class likelihood for vehicle holding duration over all the classes. This is summarized in equation (14). Here, parameters $\alpha_s$ and $\gamma_s$ are fixed to zero for one reference class in order to estimate the probabilities of other classes.

$$L_{gs} = \prod_{i \in NC} f_{gs}(t \mid X) \cdot \prod_{i \in SC} S_{gs}(t \mid X) \cdot \prod_{i \in LC} \frac{f_{gs}(t \mid X)}{S_{gs}(v \mid X)} \cdot \prod_{i \in LRC} S_{gs}(v \mid X)$$

$$\ln L = \sum_g \ln \left( \sum_s \left( M_{gs} L_{gs} \right) \right)$$

where, $L_{gs}$ is the likelihood of household vehicle holding duration under latent class $s$.

The above latent class hazard model is estimated using the Expectation-Maximum Likelihood (EM) algorithm. The EM algorithm is a general method for the analysis of missing data, and models with latent variables incorporating heterogeneity, but it has been rarely used in transportation. Zenor (1993) shows that estimates obtained from the EM algorithm represent the best latent partitioning for any desired number of segments in the context of logit model. Dempster et al. (1997) show that the EM algorithm constantly increases the value of the likelihood iteration by iteration, provided that the likelihood function is convex and unimodal.

Although the EM algorithm has been successfully applied in a variety of contexts, it is time-consuming to reach the convergence in many cases. In contrast, the EM algorithm for the latent class model turns out to be relatively easy (Green, 2001). One issue is that as normal non-linear optimization, estimates obtained from the EM algorithm are sensitive to starting values (Laird, 1978). But, in general, there is no optimization algorithm that can guarantee to reach a global maximum. McLachlan and Krishnan (1997) and Cox and Lakes (1984) give an extensive discussion of EM algorithm.

To implement the EM algorithm, let $R_{gs} = 1$ if household $g$ belongs to latent class $s$, and $R_{gs} = 0$ otherwise. $R_{gs}$ is treated as the unobservable or missing data to be estimated. If $R_{gs}$ becomes observable, the following log-likelihood function can be used to estimate the model instead of equation (14).

$$\ln L(k) = \sum_g \sum_s R_{gs}(k) \ln M_{gs}(k) + \sum_g \sum_s R_{gs}(k) \ln L_{gs}(k)$$

As shown in equation (15), the log-likelihood function of equation (14) is factorized into two parts that can be maximized separately. The EM algorithm comprises two steps. The expectation (E) step requires the calculation for the current conditional expectation of unobserved data $R_{gs}(k)$. Using Bayes’ Theorem, the E step replaces the $R_{gs}(k)$ at the $k$th iteration using the posterior probability $r_{gs}(k-1)$, which is calculated using the parameters estimated at the $k$-1th iteration.
The maximization (M) step maximizes the resulting conditional log-likelihood with the estimated posterior probabilities treated as known. Both the E-step and M-step are repeated until convergence.

3. DATA

The data set used in this study is from a revealed preference survey about household vehicle ownership behavior is used here. The data was collected in October 2006 from the households living in Tyugoku area (the largest city is Hiroshima city) in Japan. All the recruited household were asked to answer the questions about household and individual attributes, attributes of owned passenger cars in past 10 years (i.e. from 1996-2006).

- Household attributes: number of household members, number of owned passenger cars, residential characteristics etc.
- Individual attributes: age, gender, driving license, occupation, car use behavior, daily activity participation etc.
- Events of household: moving house, marriage, birth of baby, purchase of vehicle etc.
- Vehicle attributes: make, engine displacement, manufacture year, total travel distance of both current and previous vehicles etc.

As a result, we collected the questionnaires from 500 households with cars. As shown in Figure 1, the sample used in this study comprises 1,433 vehicles, among which 309 vehicles (21.6% of the sample) were replaced or disposed from 1996 to 2006. On the other hand, 345 vehicles were purchased before 1996 (left-censored data). We can not observe the household characteristics, vehicle attributes and other influential factors before 1996. Remaining 779 vehicles (54.4% of the sample) were purchased after 1996, but were still held by household when the survey was conducted (right-censored data).
Figure 2 shows the duration of vehicle holding which is replaced or disposed from 1996 to 2006. The average of duration is 6.60 years (the national average is 6.39 years (Source: Automobile Inspection and Registration Association)). As shown Figure 2, the majority of durations are 3 or 4 years. This might be caused by the fact that there is vehicle safety inspection system in Japan, and people often give up holding their vehicles before the inspection. Although a new vehicle has to receive an inspection in its third year, it is usually done every two years.

Figure 2 Vehicle holding duration in samples

Figures 3a) and b) show the relationships between the timing of vehicle disposing and household events. Here, k stands for the year when a member in household gets married or holds a driving license. From these figures, it turns out that car disposing behavior has strong relations with the events in household. This means that it is necessary to reflect the influences of these time-varying covariates in the duration model because the change of household characteristics may influence the vehicle holding duration. Moreover, it is considered that household vehicle holding duration differs according to the household characteristics.
4. FINDINGS FROM MODEL ESTIMATION

The explanatory variables used in the models of this study are summarized in Table 1. As a first step, we explored which underlying distribution best fits the duration of household vehicle holding. We examined the five distributions: 1) exponential, 2) Weibull, 3) gamma, 4) log-normal, and 5) log-logistic distributions. The estimations were run without the consideration of latent classes. It is found that the hazard model using the gamma distribution provided the best goodness-of-fit, following by Weibull, log-normal, log-logistic, and exponential distributions. Since Gamma distribution includes Weibull distribution as a special case, and the estimated shape parameter of the Gamma distribution is not statistically significantly different from 1, this study applied Weibull distribution to analyze the household vehicle holding durations.

To decide the optimal number of latent classes for the household vehicle holding durations, the models were estimated with respect to 1, 2 and 3 latent classes using the EM algorithm. As a result, two latent classes were selected based on Bayesian Information Criterion (BIC) value. Presented in Table 2 are the results of the maximum likelihood estimation of household vehicle holding duration model. All the parameters of latent membership function are significant. The results imply that assuming the latent classes are effective to describe the household vehicle holding duration. The positive sign indicates that the household tends to belong to the latent class 1, and negative sign to the latent class 2. The parameter of single-member household dummy variable has positive sign, but the parameters of the number of children and vehicles have negative signs. In addition, considering the parameter of residential duration gives positive sign, large family tends to belong to the latent class 2. Moreover, household of latent class 2 seems to be located at the adult stage and live a richer lifestyle because the parameters of household incomes and vehicle price are negative, although the parameter of used vehicle dummy variable is positive. The latent classes could be labeled “younger household” and “older household” from these results.

The parameters of main user’s age have positive sign in both classes. This means that older main users tend to own their vehicles longer. The parameters of used vehicle have negative signs. This result implies the duration of used vehicle holding is shorter than other types of vehicles. This seems logical. Moreover, durations of the households with older members are shorter since the parameters of number of children and license holders have negative signs and significant in latent class 2. And, the negative sign of income parameter indicates that the...
high income households tend to dispose their vehicle much earlier than other households. On the other hand, the parameter of dummy variable of commute vehicle is negative in latent class 2.

Considering the shape parameters in the two latent classes 1 and 2, the mean hazard rate of households in latent class 1 was higher than that of households in latent class 2 (the hazard rates of latent classes are illustrated in Figure 4(a)). This means that vehicle holding duration in latent class 1 is shorter than that of households in latent class 2. In other words, younger households have shorter vehicle holding duration.

Figure 4 also shows the hazard rate incorporating no heterogeneity (i.e. only one latent class). Compared with the hazard rate with a homogeneous specification, that of latent class 1 is higher, and that of latent class 2 is lower. However, the hazard rate with a homogeneous specification is located not midway between latent class 1 and class 2, higher position of average rate of latent classes. The result indicates that ignoring heterogeneity may lead to an upward bias of the estimates of hazard dependence over time.
Table 1 Variables used in the analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Household characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>License statement of wife</td>
<td>Dichotomy (or dummy): 1 if wife owns car license; 0 otherwise</td>
</tr>
<tr>
<td>Employed statement of wife</td>
<td>Dichotomy (or dummy): 1 if wife has job; 0 otherwise</td>
</tr>
<tr>
<td>Single-member household</td>
<td>Dichotomy (or dummy): 1 if he/she lives alone; 0 otherwise</td>
</tr>
<tr>
<td>Number of children</td>
<td>Number of children who is under 6 years old in household.</td>
</tr>
<tr>
<td>Number of workers</td>
<td>Number of workers</td>
</tr>
<tr>
<td>Number of vehicles</td>
<td>Number of vehicles in household.</td>
</tr>
<tr>
<td>Number of license holders</td>
<td>Number of license holders</td>
</tr>
<tr>
<td>Income</td>
<td>Household income</td>
</tr>
<tr>
<td>Rent home</td>
<td>Dichotomy (or dummy): 1 if household rents a home; 0 otherwise</td>
</tr>
<tr>
<td>Residential duration</td>
<td>Length of time residing at current home</td>
</tr>
<tr>
<td><strong>Mine user characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>Age of the main user</td>
</tr>
<tr>
<td>Wife</td>
<td>Dichotomy (or dummy): 1 if the main user is wife; 0 otherwise</td>
</tr>
<tr>
<td>Commute</td>
<td>Dichotomy (or dummy): 1 if the main purpose of car use is commute; 0 otherwise</td>
</tr>
<tr>
<td><strong>Vehicle attributes</strong></td>
<td></td>
</tr>
<tr>
<td>Light vehicle</td>
<td>Dichotomy (or dummy): 1 if vehicle type is light vehicle; 0 otherwise</td>
</tr>
<tr>
<td>Compact vehicle</td>
<td>Dichotomy (or dummy): 1 if vehicle type is compact vehicle; 0 otherwise</td>
</tr>
<tr>
<td>Sedan</td>
<td>Dichotomy (or dummy): 1 if vehicle type is sedan; 0 otherwise</td>
</tr>
<tr>
<td>Fuel efficiency</td>
<td>Dichotomy (or dummy): 1 if the fuel efficiency is higher than 15 km/l; 0 otherwise</td>
</tr>
<tr>
<td>Used vehicle</td>
<td>Dichotomy (or dummy): 1 if vehicle was acquired as used car; 0 otherwise</td>
</tr>
<tr>
<td>Vehicle Price</td>
<td>Vehicle Price (10 thousand yen)</td>
</tr>
</tbody>
</table>
### Table 2 Estimation results

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Latent Class 1</th>
<th>Latent Class 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>parameter</td>
<td>t-score</td>
</tr>
<tr>
<td><strong>Household characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of children</td>
<td>0.116</td>
<td>1.281</td>
</tr>
<tr>
<td>Number of license holders</td>
<td>-0.088</td>
<td>-0.749</td>
</tr>
<tr>
<td>Number of workers</td>
<td>-0.121</td>
<td>-1.157</td>
</tr>
<tr>
<td>Rent home</td>
<td>-5.33E-03</td>
<td>-0.640</td>
</tr>
<tr>
<td>Household Income</td>
<td>-2.20E-04</td>
<td>-0.738</td>
</tr>
<tr>
<td><strong>Main user characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.060 **</td>
<td>7.876</td>
</tr>
<tr>
<td>Wife</td>
<td>0.252</td>
<td>1.780</td>
</tr>
<tr>
<td>Commute</td>
<td>-0.038</td>
<td>-0.281</td>
</tr>
<tr>
<td><strong>Vehicle attributes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Light Vehicle</td>
<td>0.083</td>
<td>0.466</td>
</tr>
<tr>
<td>Compact Vehicle</td>
<td>-0.092</td>
<td>-0.486</td>
</tr>
<tr>
<td>Sedan</td>
<td>-0.360 *</td>
<td>-2.110</td>
</tr>
<tr>
<td>Fuel Efficiency</td>
<td>-0.382</td>
<td>-1.646</td>
</tr>
<tr>
<td>Used vehicle</td>
<td>0.401 **</td>
<td>3.093</td>
</tr>
<tr>
<td><strong>Shape parameters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>3.081 **</td>
<td>8.168</td>
</tr>
<tr>
<td>Gamma</td>
<td>2.465 **</td>
<td>22.118</td>
</tr>
<tr>
<td><strong>Latent membership probability</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>License statement of wife</td>
<td>0.595 **</td>
<td>4.434</td>
</tr>
<tr>
<td>Employed Statement of wife</td>
<td>-1.303 **</td>
<td>-7.164</td>
</tr>
<tr>
<td>Single-member household</td>
<td>2.472 **</td>
<td>5.224</td>
</tr>
<tr>
<td>Number of children</td>
<td>-0.542 **</td>
<td>-6.484</td>
</tr>
<tr>
<td>Number of vehicles</td>
<td>-0.291 *</td>
<td>-2.544</td>
</tr>
<tr>
<td>Household income</td>
<td>-3.79E-03 **</td>
<td>-13.029</td>
</tr>
<tr>
<td>Residential duration</td>
<td>-1.251 **</td>
<td>-8.670</td>
</tr>
<tr>
<td>Used Vehicle</td>
<td>1.313 **</td>
<td>7.604</td>
</tr>
<tr>
<td>Vehicle Price</td>
<td>-7.09E-03 **</td>
<td>-7.839</td>
</tr>
<tr>
<td>Constant</td>
<td>4.629 **</td>
<td>13.204</td>
</tr>
</tbody>
</table>

| Latent class probability                  | 0.513          | 0.487          |
| Number of samples                         | 1433           |                |
| Log-likelihood at convergence             | -1808.505      |                |

*Note*: significant at the 5% level; **: significant at the 1% level
5. CONCLUSIONS AND FUTURE RESEARCH ISSUES

In this study, we proposed a hazard based duration model which simultaneously incorporates the time-varying covariate, censored observations and household heterogeneity using the latent class modeling approach. The model has been applied to household vehicle holding duration with the data collected in Japanese several cities in 2006. The effectiveness of the model is examined from both model performance and applicability, and it is empirically
confirmed that household heterogeneity are especially influential to the duration of the household vehicle holding. Such modeling approach can flexibly accommodate multiple-peak distribution. The estimation results show that household heterogeneity was related to household characteristics. The results of the empirical study can be summarized as follows. The households which members’ average age is young, own some used cars, income level is relatively low, or number of household member is small tend to dispose their vehicle earlier than other households. On contrary, the durations of households with high incomes, expensive vehicles, 2 or more vehicles, or more members, tend to use their vehicles longer. Moreover, compared with the hazard rate with a homogeneous specification, that of latent class 1 is higher, and that of latent class 2 is lower. The hazard rate with a homogeneous specification is located not midway between latent class 1 and class 2, higher position of average rate of latent classes. The result indicates that ignoring heterogeneity may lead to some biased conclusions.

Recently, it is observed that the number of households with multiple vehicles has been increasing. The durations of these vehicles in the same household are not likely independent with each other. How to represent the interdependency among different vehicles within a household should be explored. Needless to say, a model system is required to incorporate all the elements related to car ownership in order to better help policy makers’ decisions.

REFERENCES

Sons, New York.


