Joint Modeling of Household Vehicle Holding Duration and Use with a Copula-based Multivariate Survival Model

Masashi KUWANO, Akimasa FUJIWARA, Junyi ZHANG, and Makoto TSUKAI
Graduate School of Engineering, Hiroshima University,
1-4-1 Kagamiyama, Higashi-Hiroshima, 739-8527, Japan
Phone/Fax: +81-82-424-7825
E-mail: kuwano@hiroshima-u.ac.jp

ABSTRACT

Vehicle ownership behavior consists of the correlated stages of “vehicle purchase (i.e. vehicle type choice)”, “vehicle use (i.e. annual vehicle kilometer of travel)”, and “duration of vehicle ownership”. Some studies have attempted to describe the complicated vehicle ownership behavior. For example, Bhat and Sen (2006) applied discrete-continuous models to deal with the choice of vehicle type and use of vehicle simultaneously. However, little has still been done to incorporate the correlation between vehicle holding duration and vehicle use. Recently, in Japan, it is observed that a rapid increase in fuel price has caused the decrease of vehicle kilometers traveled, the increase of vehicle replacement with more fuel efficient engine, and the decrease of total household car ownership. It is no doubt that the change of vehicle operation cost affects both on the vehicle use and duration of vehicle ownership at the same time. Therefore, vehicle ownership behavior analysis which assumes the independence across the stages may lead to misunderstanding of the behavior. In other words, the decision makings of vehicle holding duration and vehicle use should be argued in a unique model system.

Focusing on household behavior of vehicle holding duration and use, this study tries to propose a Copula-based Multivariate Survival (CMS) model to capture the interdependence between these two behaviors. “Copula” is a device or function that to approximate a stochastic dependence relationship among random variables with pre-specified marginal distributions. The copula function provides a simple and powerful approach to give a closed-form analytic expression for the joint probability, and substantial flexibility in the correlated random variables, which may not even have the same marginal distributions. Moreover, the copulas can be applied to the multivariate models to integrate various characteristics of correlation structures. In this study, we attempt to apply the widely used copulas, including Clayton, Gumbel, Frank, and Normal copulas, and explore how these copulas could fit the data within the context of this study.

Such the observations indicate that kilometers of the vehicle use and durations of the holding would be dependent, but the behavior is not homogeneous among the household. We assume that household vehicle transaction behavior is substantially heterogeneous across households. Some households may change their vehicles with the consideration of holding duration, while others may replace some of their vehicles by considering vehicle use. The existing studies have incorporated the household heterogeneity using hazard based duration models that assumes
one-peak distribution or uni-dimensional decision-making structure. In reality, such assumption could be easily violated. In order to represent the heterogeneity in the household vehicle transaction behavior mentioned above, this study further proposes to apply a latent class modeling approach to simultaneously incorporate different distributions and different dependence structures between duration of vehicle ownership and vehicle use. This approach can deal with multi-peak distribution and multi-dimensional decision making mechanisms of vehicle transaction behavior. In the latent class modeling approach, the latent class membership probability that a household belongs to a latent class with a specific sample distribution and dependence structure between holding duration and use is specified at first, and the likelihood of household choice behavior conditional on different distributions and dependence structures is weighted by the membership probabilities. In addition, Bayesian approach is applied to repeatedly update the latent class membership probabilities during the optimization process, and EM (Expectation-Maximization) steps are alternatively applied up to reach the convergence in likelihood. Theoretically, the resultant models can be used to incorporate various types of household transaction behaviors.

An empirical analysis is carried out based on the web-survey data obtained at Chugoku region in Japan, 2006. Model estimation results confirm the effectiveness of the suggested model from both model performance and applicability comparing with the conventional models. To decide the optimal number of latent classes for the household vehicle transaction behavior, the models are estimated using the EM algorithm. As a result, two latent classes were selected based on the maximum likelihood estimation, and it is further shown that introducing the latent classes is effective to describe the household vehicle transaction behavior. The estimation results about membership probabilities clarified that household characteristics and vehicle attributes are significant factors to explain the heterogeneity of household affecting the vehicle use and the holding duration. Moreover, comparisons in the performance with the conventional model, ignoring heterogeneity could lead to biased estimates of household transaction behavior. The copula parameters showed that vehicle holding duration and vehicle use are significantly correlated with each other. As the dependence structure between holding duration and vehicle use, Clayton copula is appropriate for one, but Gumbel copula is appropriate for the other latent class. This result shows that types of dependence structures between vehicle holding duration and use are different among the household. Finally, simulation analyses in fuel price showed that the proposed model can appropriately capture the recent household vehicle ownership behavior, for example, an increase in fuel price cause the prompt decrease of both vehicle holding duration and use.

REFERENCES