

Online Supplement to

“Accounting for Multi-Dimensional Dependencies Among Decision-Makers Within a Generalized Model Framework: An Application to Understanding Shared Mobility Service Usage Levels”

Data Imputation

Respondents in the 2014 Puget Sound Travel Survey provided their frequency of use (within the past month) of (1) smart-phone apps and (2) websites to obtain travel-related information. This information was elicited using a 7-point scale, which was recoded into a five-point scale to ensure there are enough sample data points within each level of usage:

- (1) Never
- (2) Less than once a week
- (3) One day per week
- (4) Two to four days per week
- (5) More than 4 days per week

Since the levels indicate an increasing degree of use, the two frequency variables from the 2014 sample are treated as ordinal outcomes and separate ordered response probit models are developed using exogenous variables reflecting individual socio-demographic characteristics (age, education level, employment status, income) and other relevant variables (smart-phone ownership and household density expressed in number of households per square mile). The results are consistent with expectations as well as findings reported in the literature (Pew Research Center, 2014). Assuming the effects of the above variables remain constant over the course of a year, the estimated models are applied to the 2015 sample, with coefficients fixed at their 2014 values. The model application provides a probabilistic assignment to the five usage levels for each ordinal frequency variable, thus, resulting in two sets of five probability values for every respondent. The naive approach to impute the 2015 usage levels for smart-phone apps and websites for travel information would be to adopt the levels corresponding to the highest probability values. However, the sample, though considered representative of the population of interest, cannot fully capture behavioral heterogeneity amongst decision makers. To account for this heterogeneity and introduce greater variability in the sample, pseudo-random draws from a uniform distribution (0,1) are taken. The interval (0,1) is split into five blocks (labeled as five levels of the ordinal variable) with thresholds based on the predicted probabilities from the model application step. These thresholds vary amongst respondents due to differences in socio-demographic and other characteristics. The respondent’s imputed level of usage is decided by the block into which the pseudo-random number falls. This procedure results in the imputed levels of usage for smart-phone apps and websites for travel information in 2015. To generate a new composite variable, frequency of using technology-based platforms for travel information, we transform the two imputed ordinal variables into monthly counts using the follow rubric:

- (1) 0 times
- (2) 0.3 times
- (3) 4 times
- (4) 8 times
- (5) 16 times

The monthly counts for app and website usage are summed to obtain monthly instances of using technology-based platforms, and then reconverted to an ordinal scale variable “frequency of using technology-based platforms for travel info” with four levels (five levels are collapsed into four to ensure a more equitable distribution of sample points and improve computational tractability) namely,

- (1) Never
- (2) Less than once a week
- (3) One day per week
- (4) Two or more days per week