**Online Supplement to:**

**E-Scooter Sharing and Bikesharing Systems: An Individual-Level Analysis of Factors Affecting First Use and Use Frequency**

By Gina Blazanin, Aupal Mondal, Katherine E. Asmussen, and Chandra R. Bhat (corresp. author)

**Census Population Statistics and Discussion for the Current Study**

Table S1 displays the descriptive statistics found on the surveyed individuals. These statistics include individual traits, household traits, and built environment factors, all answered through the U.S. Census or retrieved by combining multiple geocoded sources. Since this survey took place in Austin, Texas, the survey statistics can be compared with the Census data for the Austin-Round Rock Metropolitan Area (U.S. Census Bureau, 2018). There are a few key demographics with major differences to note between the Austin-Round Rock Census data and the survey used: age, gender, education level, and income showed skew in our sample.

Age showed skew for all age groups in the sample, as the sample displayed a much younger population than what is true for Austin. The youngest age bracket was overrepresented with 54% of respondents in the survey being 18 to 24 years old, yet only 9.8% of the Census population showed this age range. Again, 21.1% of the sample is within the 25- to 39-year-old group, while the census shows 39.4% of adults to be in this range. The older age groups are also skewed: adults 25 to 39 years old are 21.1% of the sample (39.4% of the census), adults 40 to 54 years old are 12.3% of the sample (29.3% of the census), and adults aged 55 and over are 12.6% of the sample (21.5% of the census). In terms of gender, differences were noted as follows: 68% of the sample were women, while only 48.6% of the census were women. While this skew is quite high, the relationships between ESS/BSS use within each gender can still be accurately represented using this survey. Another socio-demographic, education level, showed skew towards highly educated individuals. The number of respondents who have completed some undergraduate courses (without obtaining a degree) was 37.0%, which is a large proportion of the population compared to the census amount of 27.7%. Education levels are difficult to compare, as many students use their home residence as their census location, so many students can be underrepresented in censuses completed. Even so, this shows that, even though younger college age adults may be left out of the census data, they had a strong appearance on this survey (U.S. Census Bureau, 2018).  Finally, the survey showed clear skew in income levels where the sample has a much lower annual income level than that of the census. In fact, 38.1% of survey respondents said their household annual income is under $50,000, which relates with the 29.4% from the census. Additionally, only 14.1% of respondents said their household earned over $150,000 per year, while the census shows 21.9% of people in the Austin-Round Rock region to earn the same amount.

These skews in the sample when compared to the census likely appeared through a variety of factors such as interest in the survey topic, any financial incentives applied, the method of distribution, and the method of response. Younger adults are more likely to participate in a survey that is of interest to them, has a financial incentive, and is easy to find and participate in, which aligns with the skew to the younger population in the sample. However, even though this skew, and others, are present in the data, analysis of the survey can still be completed accurately. While not all populations are accurately represented in the survey, each of the relationships between the different demographics are still accurately displayed.  The findings from these models are applicable to the relationships of the associated population groups in the entire population of adults.

**Table S1: Individual, Household, and Built Environment Statistics of Entire Sample**

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable** | **Full Study Count (n=1107)** | | **Census** |
| **Count** | **%** | **%** |
| *Individual-Level Characteristics* |  |  |  |
| **Age** |  |  |  |
| 18-24 | 598 | 54.0 | 9.8 |
| 25-39 | 234 | 21.1 | 39.4 |
| 40-54 | 136 | 12.3 | 29.3 |
| 55+ | 139 | 12.6 | 21.5 |
| **Gender** |  |  |  |
| Male | 355 | 32.0 | 50.4 |
| Female | 752 | 68.0 | 49.6 |
| **Race** |  |  |  |
| Non-Hispanic, Non-Latino White | 571 | 51.6 | 51.6 |
| Other | 536 | 48.4 | 48.4 |
| **Education** |  |  |  |
| Completed HS or Less | 149 | 13.5 | 30.3 |
| Completed Some Undergraduate | 411 | 37.0 | 27.7 |
| Completed Undergraduate Degree or some Graduate Courses | 368 | 33.5 | 27.4 |
| Completed Graduate Degree | 179 | 16.0 | 14.7 |
| **Student or Worker Status** |  |  |  |
| Student (part or full time) | 607\* | 54.7 |  |
| Worker (part or full time) | 659\* | 59.7 |  |
| Neither a student nor worker | 115 | 10.5 |  |
| *Household-Level Characteristics* |  |  |  |
| **Annual Income, before taxes** |  |  |  |
| High (Over $150,000) | 157 | 14.1 | 21.9 |
| Medium (between $50,000 and $150,000 | 529 | 47.8 | 48.7 |
| Low (Under $50,000) | 421 | 38.1 | 29.4 |
| **Licensed Driver** |  |  |  |
| Yes | 981 | 88.5 |  |
| No | 126 | 11.5 |  |
| **Number of Vehicles in the Household** |  |  |  |
| 0 Vehicles | 79 | 7.1 |  |
| 1 Vehicle | 266 | 24.0 |  |
| 2 Vehicles | 358 | 32.5 |  |
| 3 Vehicles | 227 | 20.5 |  |
| 4+ Vehicles | 177 | 15.9 |  |
| **Household Size** |  |  |  |
| 1 Person | 255 | 23.0 |  |
| 2 People | 292 | 26.5 |  |
| 3 People | 166 | 15.0 |  |
| 4 People | 243 | 21.9 |  |
| 5+ People | 151 | 13.6 |  |
| **Kids Present in Household** |  |  |  |
| No kids | 913 | 82.5 |  |
| Kids | 194 | 17.5 |  |
| *Built Environment Factors* |  |  |  |
| **Land Use** |  |  |  |
| Urban | 300 | 27.1 |  |
| Suburban | 697 | 63.0 |  |
| Rural | 110 | 9.9 |  |
| **Transit Accessibility (3/4 mile)** |  |  |  |
| Has Access | 427 | 38.5 |  |
| **Population Density** |  |  |  |
| High (over 20 Activity Units) | 114 | 10.2 |  |
| Low (under 20 Activity Units) | 993 | 89.8 |  |

\*274 are both a student and work

**Table S2: Indicator Loadings for Latent Constructs**

|  |  |  |
| --- | --- | --- |
| **Latent Variable** | Coeff. | t-stat |
| ***Safety Concern*** |  |  |
| AVs would make me feel safer on the street as a pedestrian or as a cyclist. | -0.755 | -18.77 |
| I am concerned about the potential failure of AV sensors, equipment, technology, or programs. | 0.441 | 16.18 |
| I would feel comfortable sleeping while traveling in an AV. | -0.917 | -18.43 |
| ***Time Consciousness*** |  |  |
| I am too busy to do many of the things I like to do. | 0.151 | 5.64 |
| I try to make good use of the time I spend traveling. | 0.432 | 10.59 |
| The level of congestion during my daily travel bothers me. | 0.378 | 10.38 |
| ***Green-Lifestyle Propensity*** |  |  |
| The government should raise the gas tax to help reduce the negative impacts of transportation on the environment. | 0.688 | 15.46 |
| I am committed to using a less polluting means of transportation (e.g. walking, biking, and public transit) as much as possible. | 0.746 | 14.74 |

## Mathematical Formulation of the GHDM for the Current Study

Since the main outcome variables consist of two binary outcomes and two ordinal outcomes, the binary outcomes can be modeled as ordinal variables as well (with 1 and 2 as the ordered levels). Given all the indicators are ordinal in nature, the GHDM model is formulated with only ordinal outcomes.

Consider the case of an individual . Let  be the index of the latent constructs and let  be the value of the latent variable *l* for the individual *q*.  is expressed as a function of its explanatory variables as,

,

where  is a column vector of the explanatory variables of latent variable *l* and is a vector of its coefficients.  is the unexplained error term and is assumed to follow a standard normal distribution. Equation (1) can be expressed in the matrix form as,

,

where is a column vector of all the latent variables, is a matrix formed by vertically stacking the vectors  and  is formed by vertically stacking .  follows a multivariate normal distribution centered at the origin and having a correlation matrix of , i.e., , where  is a vector of zeros. The variance of all the elements in  is fixed as unity because it is not possible to uniquely identify a scale for the latent variables. Equation (2) constitutes the SEM component of the framework.

Let denote the index of the outcome variables (including the indicator variables). Letbe the underlying continuous measure associated with the outcome variable. Then,

,

where  denotes the ordinal category assumed by  and  denotes the lower boundary of the *k*th discrete interval of the continous measure associated with the *j*th outcome.  for all *j* and all *k*. Since  may take any value in , we fix the value of and  for all *j*. Since the location of the thresholds on the real-line is not uniquely identifiable, we also set .  is expressed as a function of its explanatory variables as,

,

where is a vector of explanatory variables for the continuous measure including a constant,  is a column vector of the coefficients associated with , and  is the vector of coefficients of the latent variables for outcome *j*.  is a stochastic error term that captures the effect of unobserved variables on .  is assumed to follow a standard normal distribution. Jointly, the continuous measures of the *J* outcome variables may be expressed as,

,

where  and  are the vectors formed by vertically stacking and , respectively, of the *J* dependent variables.  is a matrix formed by vertically stacking the vectors  and  is a matrix formed by vertically stacking .  follows a multivariate normal distribution centered at the origin with an identity matrix as the covariance matrix (independent error terms). . We assume the terms in  to be independent because it is not possible to uniquely identify all the correlations between the elements in and all the correlations between the elements in . Further, because of the ordinal nature of the outcome variables, the scale of  cannot be uniquely identified. Therefore, the variances of all elements in  is fixed to one. The reader is referred to Bhat (2015) for further nuances regarding the identification of coefficients in the GHDM framework.

Substituting Equation (2) in Equation (5),  can be expressed in the reduced form as

,

.

In the right side of Equation (7),  and  are random vectors that follow the multivariate normal distribution and the other variables are constants. Therefore,  also follows the multivariate normal distribution with a mean of  (all the elements of  and  have a mean of zero) and a covariance matrix of .

.

The parameters that are to be estimated are the elements of , strictly upper triangular elements of **Γ**, elements of ***β***, elements of ***d*** and  for all *j* and  (though no  is to be estimated for the binary outcomes, since *k* takes only the values 1 or 2 for these binary outcomes). Let ***θ*** be a vector of all the parameters that need to be estimated. The maximum likelihood approach can be used for estimating these parameters. The likelihood of the *q*th observation will be,

,

where,  denotes the probability density of a *J* dimensional multivariate normal distribution centered at the origin with a covariance matrix **Σ** at the point Since a closed form expression does not exist for this integral and evaluation using simulation techniques can be time consuming, we used the One-variate Univariate Screening technique proposed by Bhat (2018) for approximating this integral. The estimation of parameters was carried out using the *maxlik* library in the GAUSS matrix programming language.

**References**

Bhat, C.R., 2015. A new generalized heterogeneous data model (GHDM) to jointly model mixed types of dependent variables. *Transportation Research Part B* 79, 50-77.

Bhat, C.R., 2018. New matrix-based methods for the analytic evaluation of the multivariate cumulative normal distribution function. *Transportation Research Part B* 109, 238-256.

U.S. Census Bureau. 2018. American Community Survey 1-year Estimates. Census Reporter profile page for Austin-Round Rock, TX metro area. Available at <https://censusreporter.org/profiles/31000US12420-austin-round-rock-tx-metroarea/> [Accessed by: 2/25/20]