Technologies for Safe & Efficient Transportation

THE NATIONAL USDOT UNIVERSITY TRANSPORTATION CENTER FOR SAFETY

Carnegie Mellon University | UNIVERSITY of PENNSYLVANIA

Impact of Personal Attitudes on Propensity to Use Autonomous Vehicles for Intercity Travel

FINAL RESEARCH REPORT

Megan Ryerson (PI), Ivan Tereshchenko

Contract No. DTRT12GUTG11

DISCLAIMER

Problem:

The autonomous vehicles are about to become a reality. The researchers estimate the benefits from each autonomous vehicle to be between $2000 and $4500 per vehicles. The societal benefits include higher travel time savings, reduced congestion, fuel consumption, lower rate of traffic accidents. However, as with any disruptive technology, autonomous vehicles bring a number of difficult challenges, such as the need to prepare the transportation system for the new technology. Research by T-SET professor Erick Guerra shows that planning organizations and local governments – those who plan for large scale investment into the existing infrastructure in the United States – are struggling account for the introduction of new intercity transportation technology, such as autonomous vehicles, in their plans. This is particularly true for the planning of *intercity* transportation, as the planning for intercity transportation is done in a fragmented, modally siloed way. Thus, the potential impact of driverless vehicles on the structure of the intercity transportation system is therefore unknown, yet positioned to be significant.

One of the primary objectives of this study was to create a model of traveler behavior for intercity travel that takes the information about personal attitudes and norms into account. A large body of literature show that such models should include socio-demographic variables, such as age, gender and income, as well as personal attitudes towards travel and life, such as feeling of dependence on cars, concerns about safety and flexibility of schedules. The researchers use the general framework of the Theory of Planned Behavior (TPB) to account for these variables. The main idea behind this theory is that the behavior can be explained by the personal intentions, which in turn can be explained by a set of attitudes, subjective norms and behavioral controls. By studying the attitudes, norms and controls we can create accurate models of intentions and future behavior. The main research problem of this study is to analyze the factors that influence the propensity to use autonomous vehicles in the future. The study aims at incorporating the information about personal attitudes into the model of future travel behavior. This is one of the first studies that applies TPB to modelling the intercity travel behavior.

Approach:

In this study, we collected the data on personal attitudes and travel behavior in the Northeast corridor. Each attitude was measured by a series of indicators. These indicators were later used to calculate the latent attitudes. The latent attitudes served as the predictors for the intended behavior – the use of driverless cars instead of other intercity ground modes when they become available. Finally, we looked at the relationship between the latent attitudes and socio-demographic characteristics of the respondents. The socio-demographic variables can later be used as proxies for attitude variable in probabilistic simulations of mode choice in intercity travel.

Methodology:

In this section we will describe the questionnaire, sample design, and the variables in the data set. This section was completed through Dr. Ryerson's partnership with Resource Systems Group (RSG).

Survey design

The questionnaire consisted of several sections. It included introduction and screener sections and information about the most recent trip in the study corridor. The second major part of the survey was the stated-preference survey for seven hypothetical trips in the Northeast corridor. The questionnaire also contained questions about general attitudes and values, scenario-specific questions and profiling questions about the socio-demographic characteristics of the respondents. This study deals exclusively with the part of the survey related to the respondents' travel attitudes, behavior and socio-demographic characteristics.

The section of the survey that asked about the latest trip in the study corridor included questions about the origin and destination, mode, purpose, and duration of the trip. The questionnaire asked about the size of the travel party and its composition. The survey included questions about the hypothetical situations related to the trips, such as the preferred transportation modes in case the actual mode was not available.

The information about the respondent's attitudes included attitudes towards community and environment, automobiles, interacting with others. These questions related the people's general attitudes towards life and travel. Another set of attitude questions asked people to focus on the latest actual trip in the study corridor. The respondents were asked to provide information about their thoughts on rail cost, time, and flexibility for this trip, thoughts about privacy and safety, flexibility of air, bus and rail, feelings about the stress experienced on aircraft, buses, or trains compared to driving, crime and unpleasant behavior on intercity transportation, subjective norms, trip feasibility and appeal. Finally, the questionnaire included questions about ideal home location, technology ownership, use and dependency. The profiling section of the survey contained questions about the age, gender, race, education, income, employment status, household characteristics, vehicle ownership, and home location.

The sample was designed for the minimum sample size of 2,000 respondents. All surveys were taken online in two regions - North-East corridor (NEC) that includes Washington DC, Philadelphia, New York City, and Boston metropolitan areas, and North-West corridor that contains Portland, Seattle, and Vancouver metropolitan areas. The survey had to contain a minimum of 1,500 respondents in NEC. It had to be a good mix of female and male respondents. It was intended to have an equal representation of all age and income groups.

In tables 1-5 you can see the actual distribution of the socio-demographic characteristics in the collected sample. The final sample contains 5628 respondents from NEC who completed the survey. The sample is skewed toward people age 45-64. These two age categories account for 42 % of the sample. There are slightly more women than men, 52 % versus 48 %. The most highly represented households have income between $75 and $ 150 thousand, which is significantly higher than the US average of $55 thousand. The lower income spectrum ($0-35,000) accounts for about 10 % of the sample. The geographic representation of different metropolitan areas is roughly the same. The majority (71 %) of people drove in their latest trip in the NEC. Other intercity modes have between 400 and 600 observations, or 7-11 % of the sample.

Table 1. Breakdown of the sample by age.

Age	Count	Share
18-24	258	0.05
25-34	876	0.16
35-44	994	0.18
45-54	1134	0.20
55-64	1211	0.22
65-74	958	0.17
75 or older	197	0.04

Table 2. Breakdown of the sample by gender.

Gender	Count	Share
Female	2929	0.52
Male	2699	0.48

Table 3. Breakdown of the sample by income.

Income	Count	Share
Less than -$25,000	342	0.06
$25,000 - $34,999	293	0.05
$35,000 - $49,999	434	0.08
$50,000 - $74,999	842	0.15
$75,000 - $99,999	1116	0.20
$100,000 - $149,999	1230	0.22
$150,000 - $199,999	583	0.10
$200,000 - $249,999	308	0.05
$250,000 - $299,999	139	0.02
$300,000 - $349,999	77	0.01
$350,000 or more	163	0.03

Table 4. Breakdown of the sample by metropolitan area.

Metro area	Count	Share
Boston	1309	0.23
New York	1878	0.33
Philadelphia	1296	0.23
Washington DC	1145	0.20

Table 5. Breakdown of the sample by mode of the last trip.

Mode	Count	Share
Personal car	3705	0.66
Rental car	292	0.05
Intercity bus	482	0.09
Intercity rail	646	0.11
Airplane	372	0.07
Other	131	0.02

Variables

The dataset contains a wide variety of variables. These include socio-demographic variables, variables that describe personal general and travel attitudes, attitudes and feelings related to a particular trip. We are primarily interested in determining the variables that predict the propensity to use driverless cars. We created theses variables by transforming the attitude variables into factors using principal component analysis. Additionally, we used the available socio-demographic variables, and information related to the use of technology.

The goal of this paper is to examine the factors that influence the propensity to use driverless cars if they become available. The propensity to use driverless cars in the future is measured by one question from the questionnaire in the section of the survey about general personal attitudes. The question was stated in the following form: ”If driverless cars were to become a reality, I would be less likely to travel by rail or bus”. The response was measured on a seven-point scale from ”Strongly agree” to ”Strongly disagree”. Figure 1 below illustrates the distribution of the results.

Approximately a third of all respondents neither agree, nor disagree with the statement. This means that they are uncertain about whether they will use driverless cars in the future. For our purposes we excluded those respondents from the analysis. The remaining six answers were joined into two groups. One group includes the people who agree with the statement, the remaining people disagree with the statement. This binary responses was used as an independent variable in further analysis.

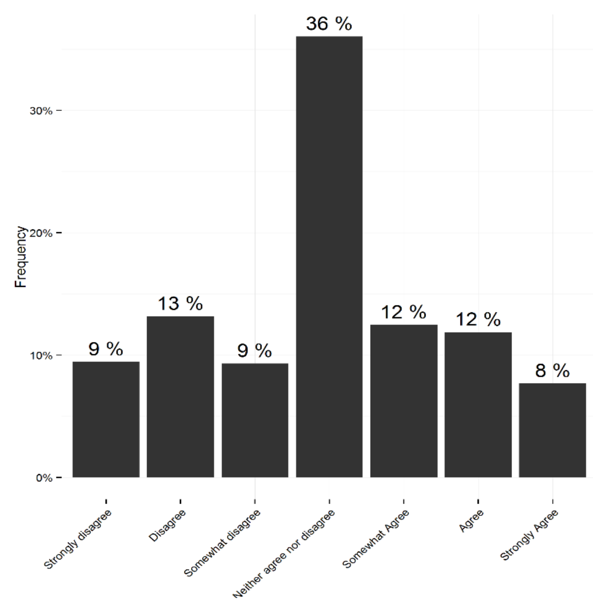

Figure 1. If driverless cars were to become a reality, I would be less likely to travel by rail or bus

Our main objective was to describe the relationship between the propensity to use driverless cars when they become available and people’s attitudes towards life and intercity travel. The dataset contains fifty attitude variables that measured different general attitudes and attitudes towards intercity travel on a seven-point scale. In order to use this information for our purposes, we reduced the dimensionality of the data using principal component analysis. Overall, we were able to identify nine principal components that account for the majority of information about attitudes that were included in the original dataset. The resulting factors correspond to the passengers' safety and flexibility concerns, discomfort while travelling on trains, fears, associated with travelling with strangers, independence from using automobiles, social approval to travel by trains, and the need to use electronic devices during travel.

The attitude variables can be divided into two groups. The first group of variables corresponds to general attitudes towards life and travel, such as enjoying being around other people, or liking urban environment more that suburban. The second group of the variables relates to attitudes associated with each respondent's latest trip in North-East Corridor. Because of this difference, we conducted two separate PCAs in order to compute two sets of factors.

The original dataset joined the variables into small groups of 3-5 similar attitudes. Each attitude was measured on a seven-point scale. Then those attitudes would serve as the inputs for latent factors calculated using confirmatory factor analysis. The latent variables would serve as inputs for the structural equation model. Instead of following this approach, we decided to conduct the principal component analysis that would find all of the correlated variables and substitute them with factor scores. The principal component analysis was conducted using psych package for R programming language.

We split the principal component analysis into two parts. Even though the first set of variables is similar to the second set, there is a crucial difference. The first set of variables describes the general attitudes towards travel and life. The second set relates to the specific situations that the respondents might face, for example home-based trips, or travelling to Boston by train. Splitting the principal component analysis into two parts yields better outcomes. The resulting principal components reduce more variance in case of a two-part analysis, compared to the PCA for the combined pool of 50 variables.

The results that we present in this report only include the variables that correlate significantly with at least one principal component. The variables that do not correlate with the principal components were excluded from the analysis and are not presented in the final results. During the process of constructing principal components we excluded twelve variables out of the analysis. The list of excluded variables includes the following attitudes:

Statistical methods

The principal component analysis was conducted using a polychoric correlation matrix, based on the Kendall rank correlation coefficient. Kendall rank correlation coefficient is a broadly used alternative to Pearson correlation that can be used for ordinal data.

Let $(x_1,y_1),...,(x_n,y_n)$ be the set of observations (x_i,y_i). If two consecutive observations satisfy the properties $x_k > x_i$ and $y_k > y_i$, or $x_k < x_i$ and $y_k < y_i$, the pair is called concordant. If $x_k > x_i$ and $y_k < y_i$, or $x_k < x_i$ and $y_k > y_i$, then the pair is disconcordant. The following formula can be used to calculate
Kendall's τ coefficient:

$$\tau = \frac{number\ of\ concordant\ pairs - number\ of\ discordant\ pairs}{\frac{1}{2} \times n(n-1)}$$

The coefficient describes the strength of a monotonic relationship (correlation) between a pair of variables. It can be used to determine the correlation between ordinal variables, or to describe non-linear relationships. Kendall's τ varies between -1 and 1, the sign determines the direction of the relationship. The correlation matrix based on τ coefficient can be used to construct the factors and determine factor loadings.

Finally, after preparing the necessary data we created several logistic regression model. The dependent variable in all of those models was the propensity to use driverless vehicles instead of buses and trains, if they become available. We tested the models with all of the attitude variables and some of the socio-demographic variables, such as age, gender, and income. The analysis showed that the attitude variables already account for much of the impact of socio-demographic variables. In the final section of the study we review the relationship between attitudes and the socio-demographic variables.

Findings:

Principal component analysis

We extracted factor loadings for the variables that describe the respondents' general attitudes towards life and travel. Table 6 only contains the highest factor loadings for each factor. The first factor component is highly correlated with the following attitudes: enjoying being with people, liking to live in a walkable neighborhood, and believing in saving the environment using joint actions of the society. We called this component 'community engagement'. The second factor correlates most strongly with the attitudes such as feeling less dependent on cars than parents, liking to drive, and liking freedom and independence. It is negatively correlated with these variables. Higher values of the principal component are associated with lower dependence on automobiles. We called this component

’independence from cars’. The third factor has the highest correlation with just one variable - feeling uncomfortable on trains and buses with other people. It also has a slight negative correlation with the variables associated with the community engagement component. We called this component ’fear of strangers’ The fourth factor loads the most on two attitudes: feeling the importance of performing tasks on laptops during travel, and importance of receiving electronic messages. We called this variable ’importance of electronic devices’.

Table 6. Factors that describe general attitudes towards life and intercity travel and their factor loadings.

Factor	Attitude	Factor loading
Community Engagement	I enjoy being out and about and observing people	**0.656**
	I like to live in a neighborhood where I can walk to a commercial or village center	**0.595**
	If everyone works together, we could improve the environment	**0.58**
Independence from cars	I feel I am less dependent on cars than my parents are/were	**0.6**
	I need to drive a car to get where I need to go	**-0.678**
	I love the freedom and independence I get from owning one or more cars	**-0.741**
	Rather than owning a car, I would prefer to rent one	**0.698**
Fear of strangers	The idea of being on a train or a bus with people I do not know is uncomfortable	**0.989**
	I don’t mind traveling with people I do not know	**-0.471**
Importance of electronic devices	It would be important to me to receive about my bus or train trip	**0.725**
	Being able to freely use smartphone is important to me	**0.63**

Table 7 includes the second group of factors. The first factor loads most heavily on the variables. Overall, it correlates highly with seven variables that relate to the attitudes about the efficiency and pleasure of taking trains. We called this principal component ’the propensity to travel by train’. The second factor loads heavily on variables that correspond to attitudes about feeling uncomfortable on trains. We called this PC 'discomfort on trains'. The third factor loads heavily on variables that indicate the concerns about the flexibility of airplane, bus, and train schedules. We called this component 'flexibility concerns'. The fourth factor loads on attitudes that show concerns about personal safety on aircraft, buses, and trains. The name of the component is 'personal safety concerns'. Finally, the fifth component loads the highest on variables that indicate the approval for taking a train to Boston from family and friends, therefore we simply called this variable 'social approval for taking a train'.

Table 7. Factors that relate to the latest trip in the North-East Corridor and their factor loadings.

Factor	Attitude	Factor loading
Personal safety concerns	Worry about personal safety/disturbing behavior - Bus	**0.721**
	Worry about personal safety/disturbing behavior - Train	**0.835**
	Worry about personal safety/disturbing behavior - Air	**0.687**
Flexibility concerns	Concerned about flexibility of schedules - Air	**0.707**
	Concerned about flexibility of schedules - Bus	**0.842**
	Concerned about flexibility of schedules - Train	**0.874**
Discomfort on trains	If made trip by train - I would worry about crime or unruly behavior at the train station and on the train	**0.785**
	If made trip by train - I would feel uncomfortable being on the train with strangers	**0.77**
	If made trip by train - I might have to be with people whose behavior I find unpleasant	**0.637**
	If made trip by train - It might be unsafe to make this trip by train	**0.827**
Social approval of taking a train	Most people would approve of my taking the train to Boston.	**0.777**
	My friends take the train when they travel to Boston.	**0.31**
	Everyone would approve of my taking the train to go to Boston.	**0.68**

Factor	Attitude	Factor loading
Propensity to travel by train	Most people who are important in my life would take the train	**0.606**
	If I wanted to, I could easily take the train for this trip	**0.66**
	Rate how possible it would be for you to make this trip via train	**0.586**
	How efficient would it be to take the train for this trip	**0.793**
	How pleasant would it be to take the train for this trip	**0.695**
	I would definitely consider taking the train for this trip	**0.861**
	How likely would it be for you to take the train for this trip	**0.868**

The factor variables that describe general and travel attitudes were found to be the strongest predictors of the propensity to use driverless vehicles when they become available. The logistic regression model shows that the probability of wanting to use driverless cars instead of buses and trains is higher if the person expresses larger concerns about personal safety and discomfort on buses and train, places greater importance on the ability to use electronic devices, has a higher level of community engagement, and is more concerned about the flexibility of bus and train schedules. Increased social approval of travelling by train is negative associated with propensity to use driverless cars instead of buses and train.

The variance of attitude factors accounts for most of the variance associated with socio-demographic characteristics. The people who have fewer cars, live outside of Philadelphia MSA, are younger, and make less money feel more independent from driving. Younger people and people with lower income experience higher levels of discomfort on trains and express larger personal safety concerns while travelling by bus or train. The importance of using electronic devices while travelling is higher for those who have higher dependency on smartphones and tables, and are younger. The concerns about flexibility of schedules are higher for people, who own more vehicles and live in New York or Philadelphia. Finally, social approval of travelling by train is higher for residents of Philadelphia and Washington DC, and people in high-income category.