University of Glasgow

School of Social & Political Sciences

Postgraduate Coursework Coversheet

After you have submitted your electronic copy through the course Moodle, complete this coversheet in full and attach to the front of your paper copy. Refer to your course handbook for submission deadlines and presentation requirements.

Please attach to each copy of your work

Student Number:	2	5	1	6	2	0	2

Course Title: ________Transport Planning Method_________

Essay/Assignment Title:_Compared to automobile car users, are public transport commuters more or less likely to buy an autonomous vehicle? A case study of Seattle_

Course Co-ordinator:_________Dr. Jinhyun Hong__________

Date of Submission:________22 March 2021___________

Word Count: __3,119 words (excluding references)__

Confirmation of electronic copy submission (read & tick in box below):

Please note: You must upload the final version of your work to the course Moodle before handing in your printed copy. The Course Administrator will check Moodle shortly after the deadline and if your electronic copy is missing, normal late penalties will apply until you provide it.

I confirm that I have uploaded my electronic copy to the course Moodle before handing in this paper copy	√

Compared to automobile car users, are public transport commuters more or less likely to buy an autonomous vehicle? A case study of Seattle

1. Introduction

1.1 Research background

Rapid development of smart technology has been taking place in the transportation system in recent years along with the release of more autonomous vehicles (AVs), an innovative transport option and has been used in many places around the world (Bagloee et al., 2016). The passenger transportation system consists of public transport and private automobiles, and AVs can be counted as a 'third' travel mode. However, AVs are still not a widespread transport mode, but this novel technology can have a devastating influence on traditional transportation system (Jing et al., 2019; Fagnant & Kockelman, 2015) and influence people's travel mode choice preference, which can affect urban traffic planning and demand. AVs can help save people's travel cost (i.e., parking cost, insurance fee), provide accessibility and comfort (i.e., the release in both physical and mental activities associated with driving) to those who cannot or do not want to drive (Gandia et al., 2018; Bagloee et al., 2016). Additionally, AV technology can increase road safety and reduce car accidents caused by human errors (i.e., alcoholism, fatigue and unskilled driving). Despite AVs not being a formal or widespread transportation mode, investigations into the favourability of autonomous vehicles for public transport and private car users can be a useful reference for the AV industry and urban traffic planners.

Most existing research on AVs can be separated into macro and micro. On the macro-level, research mainly focuses on policy, law and ethical issues (Salatiello and Felver, 2017). For example, Fagnant and Kockelman (2015) suggested that, to reduce uncertainty on the impacts and interactions between AVs and the traditional transport system after implementation, the government should set up a framework to introduce national standards in liability, security and data privacy for AVs. In terms of the micro-aspect, topics such as factors influencing an individual's choice on AV (Yuen et al., 2020) are more popular. These studies in AV preference mainly use sociodemographic, travel mode preference and technology sensitivity as exploratory variables (main independent variables) in the model. However, most studies did not consider on categorising traditional transportation users into public transport users and private car users then exploring their AV purchasing preference.

Thus, based on the difference in technology acceptance whilst considering the heterogeneity of different traveller preferences, taking Seattle as the research area, the purpose of this paper is to explore who are more likely to purchase an AV, public transport users or private car users. A Binary Logit model is conducted as the research model. This model will combine individual socio-demographic characteristics, trip pattern, technology sensitivity of public transport and private car users.

1.2 Literature review

The purpose of this paper is to explore the preference on purchasing an AV between public transport and private car users. In this section, we will review the two main relevant research fields: 1) AV impacts to the traditional transportation system, and 2) factors influencing people to use AV.

1.2.1 The impacts of AV to transportation system

Human mobility is largely provided by traditional transport (public transport) system, people can arrive at their destination conveniently by public transport when the travel distance is not walkable (Lam et al., 2016). Most public transport route, such as bus, streetcar, urban railway and so on, is immobilised and irreversible. Nevertheless, taxis and private cars can provide flexible and customised route based on passengers' need. AVs have been experimented with and undergone the transport system for several decades, they can meet most of the public and private transport requirements and influence the traditional transportation system.

Most research has suggested that people will change their existing travel mode to AVs, Emberger and Pfaffenbichler (2020) simulated using a transport interaction mode (MARS, Metropolitan Activity Relocation Simulator) to examine the potential influence of AVs on the public transport system, and the simulation results showed that AVs have explicit impacts on road capacity (-2%) and the number of private car users (+22%). Moreover, AVs show a sensitive impact which is uncertain on public transport system (Salatiello and Felver, 2018). Booth et al. (2019)'s regression model of AV impacts on active travel (public transport, walk etc) implied that around 50% public transport users would like to use AV instead of traditional transport. However, if more public transport users use or own an AV, it will bring worse results and negative influence on the traditional transport system. Traffic congestion and air pollution issues can be obvious disadvantages for having increases in AV ownership. AV technology can make up the disadvantages of traditional transport, and have unique advantages, however, it also has some influence on traditional transport system, which might influence people's travel mode choice in the future.

1.2.2 The factors influencing people to use AV

Many researchers have explored the factors that influence people's travel mode choice and AV using preference. Booth et al. (2019) identified demographic and other transport relevant factors that are associated with the likelihood of using AVs for commuting. Jing et al. (2016) conducted an online questionnaire based on the hypothesis that determine the impact of AVs or SAVs on people's travel mode choice include attitude, knowledge, perceived risk, subject norm to AV/SAV. The results show that travellers' knowledge and attitude on technology can be AV mode selection triggers. Dai et al. (2021) also implied that people's initial experience, such as the understanding level of AV and the education level, can influence AV preference. Many active travel and public transport users may change their existing travel mode with AVs, and this is more likely to happen among younger people, people who have a positive preference and attitude towards AV, and public transport regular users. Thus, younger people and people living in urban area or city centre tend to be more sensitive to technology (Zhang and Guhathakurta, 2018; Hong and Thakuriah, 2018).

From the literature discussion above, this paper will focus on the socioeconomic, technology sensitive factors as explanatory variables based on the survey data set and aim to explore the AV preference of public transport and private car users.

2. Method

2.1 Data

This report used the Household Travel Survey data (2019) from Puget Sound Regional Council (PSRC), the planning organization that develops policies about transportation, economy, and land use that encompassed the four-county (King, Kitsap, Pierce, and Snohomish) in Seattle. The PSRC Travel Study was conducted by face-to-face and smartphone app (rSurvey and rMove), included 82 cities and towns with a total population of about four million people and comprised approximately 1,548,788 households between April and June in 2019. The investigated regions show in figure 1.

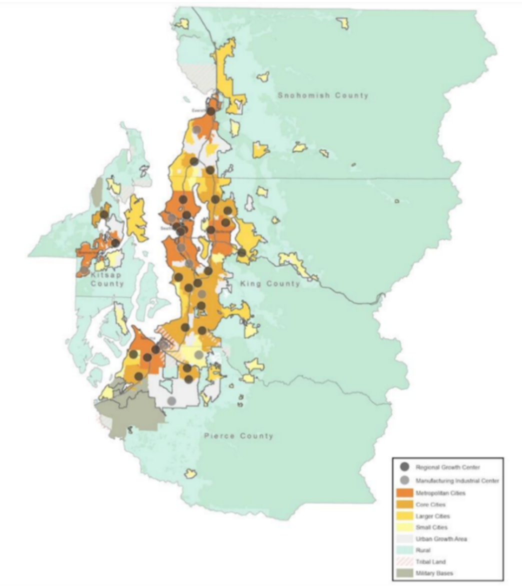

Fig. 1. Map of study area (PSRC, 2019)

The survey contains participants' demographic information including age, gender, number of workers, household size, education level, annual income, location of the participant's according to metropolitan area, etc. Smartphone question such as 'what phone you are using', for those participants who did not choose 'I do not have a smartphone' will answer the sub-question 'do you have a phone that used for over (include) 4 years?'. Commute mode is the key independent variable for this report, whose question includes active travel (walk, skateboard, wheelchair etc), public transportation (urban railway, streetcar, bus transit etc), and driving cars (driving alone). In this research, we discarded the active travel and re-sorted the other values as public transport users and car users. In addition, the dependent variable, buying an AV or not, is an ordinal variable by using interest scale options, and the values of which are combined and assigned as binary (value 1: interested in buying an AV = 'somewhat interested' + 'very interested', value 0: not interested in buying an AV = 'not interested at all' + 'somewhat uninterested'). The final observation size is 1712, including 744 private car users and 968 public transport users. The detailed statistical description is shown in table 1.

Table 1

Descriptive Statistics for AV interested groups and full samples

Explanatory variables	Full sample		Interested in AV		Not interested in AV	
	Mean	SD	Mean	SD	Mean	SD
Sociodemographic characteristics						
Gender *(male)*	51.17%		59.79%		42.19%	
Age	37.82	0.75	34.88	1.10	39.83	1.06
18-24 years (group0)	5.60%		7.45%		4.29%	
25-34 years (group1)	41.47%		52.12%		30.39%	
35-44 years (group2)	27.69%		25.89%		29.56%	
45-54 years (group3)	14.25%		9.39%		19.31%	
55-64 years (group4)	10.69%		5.15%		16.45%	
Education level						
with a higher education degree	83.64%		88.43%		78.66%	
without a higher education degree	16.36%		11.57%		21.34%	
Income						
< $25,000 (group0)	4.96%		5.38%		4.53%	
$25,000-$49,999 (group1)	12.15%		12.14%		12.16%	
$50,000-$74,999 (group2)	14.95%		12.49%		17.52%	
$75,000-$100,000 (group3)	15.71%		14.55%		15.90%	
> $100,000 (group4)	52.22%		55.44%		45.91%	
License *(have a license)*	96.26%		95.42%		97.14%	
Employment						
Employed full time (group0)	4.15%		2.98%		5.36%	
Employed part time (group1)	8.82%		7.33%		10.37%	
Self-employed (group2)	87.03%		89.69%		84.27%	
Residential location *(urban area)*	53.91%		59.00%		48.63%	
Technology characteristic ('TeC')						
Smartphone age *(over 4 years)*	9.29%		8.02%		10.61%	
Trip characteristic ('TrC')						
Commute mode						
public transport users = 0	43.46%		49.26%		37.43%	
private car users = 1	56.54%		50.74%		62.57%	
Sample size	1712		873		839	

Table 1 presents the descriptive statistics of full sample, people interested in AV and people not interested in AV of the final data set. The final data set contains three categories of variables (Sociodemographic, Technology, and Trip) and several sub-characteristics, and the group 0 of each categorical variable is the baseline for the model. Overall, each category has similar characteristics for each variable and AV interest users. In terms of sociodemographic variables, there are over 51% are male participants, and about 83% people hold a higher education degree (associate degree, bachelor's degree, and master's degree) in total. Interestingly, age group 3 (25-34 years) occupies the most of each sample type, around 41%, 52%, and 30% respectively. Interestingly, the average age of people who are interested in AV is over 4 years younger than the total sample's (about 38 years old). Over 95% of the respondents have a driving license in Seattle, and around half of them living in the urban area. For trip characteristic, approximately 63% of the participants are not interested in buying an

AV, however, there is an over 10% more interest in buying an AV from public transport users than that from private car users (accounting for 49.26% and 37.43% respectively).

2.2 Model and assumptions

In this report, we explore the AV purchasing interest between private car users and public transport users. Results from the AV purchase choice binary logit model estimation are shown in Table 2.

2.2.1 Analytical model

Due to the attribute of the AV buying interest (dependent variable) is binary, thus, the probability of y = 1 (buy an AV) between public transport and private car users can be calculated by a binary logit model (DeMaris, logistical modelling, 1992). In this case, we choose sociodemographic (i.e., age, gender, education, residential location etc), technology (smartphone age), and trip characteristics (commute mode) as the independent variables. The dependent variable represents 'interested in AV' and 'not interested in AV' respectively. Thus, the final expression of binary logit model is specified as:

$$Pr(y_i = 1 \mid X_{Sdi}, X_{TeCi}, X_{TrCi}) = logit^{-1}(\alpha + \beta_{SdC}^{T} X_{SdCi} + \beta_{TeC} X_{TeCi} + \beta_{TrC} X_{TrCi})$$
$$i = 1, 2, 3, \dots, n$$

Where Xi represents sociodemographic, technology usage, and trip pattern based on the variable description in Table 1, n represents the number of observations. commute_mode is a dummy variable which equals to 1 if people use private car, and 0 if people use public transport. The equation can be interpreted that the probability of people who want to buy an AV given by sociodemographic, technology, and trip characteristics. In this case, we assign av_intereste_6 (1: people want to buy and AV; 0: people do not want to buy an AV) as latent variables in this model and the assumptions could be written as:

$$Y_{av\_interest}^{*} = X_{av\_interest}\beta_{av\_interest} + \mathcal{E}$$
$$Y_{av\_interest} = \begin{cases} 1, if\ Y_{av\_interest}^{*} > 0 \\ 0, if\ Y_{av\_interest}^{*} \leq 0 \end{cases}$$

Where Ɛ is the error that distributed according to normal distribution, then we can use Maximum Likelihood Estimation to calculate the distribution of errors.

2.2.2 Model assumption

The logit regression model does not require assumptions that are based on OLS such as LM or GLM regarding linearity, normality, and homoscedasticity, especially due to the categorical attribute of each independent variable. However, there are still some unique assumptions that binary logit model needs to apply. Firstly, there should be binary or ordinal dependent variable (av_interest_6). Secondly, the sample size should be sufficiently large, which means that it should contain enough observations (i.e., N = 500) to support the model and assumptions. Thirdly, the independence of observation is the premise of logit model, each observation of PSRC survey did not come from repeated measurements of the same individual or be related to each other. In addition, the multicollinearity is the basis for each regression model. This assumption means there should be none or little multicollinearity among independent variables. In this case, the independent variables such as commuting mode, income and so on should not have strong correlation between each other. In logit model, the multicollinearity can be tested by Spearman's correlation. The result is shown in fig. 2.

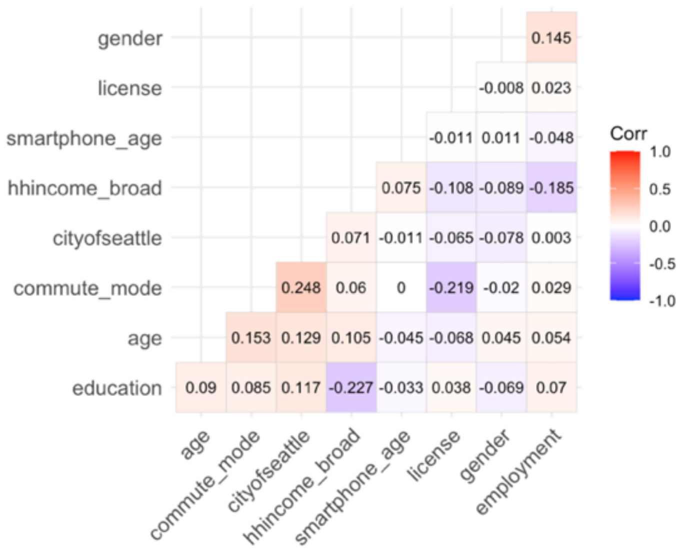

Fig. 2. The correlation matrix between each independent variable

Fig. 2. illustrates the correlation between each independent variable. The deeper colour shows higher correlation, and we can see that the highest correlation is from income and education, accounting for -0.227 (<0.5), which means that there is no significant correlation between income and education. Thus, the multicollinearity between each independent variable can be accepted and ignored.

3. Result

3.1 Model estimation

The binary logit model results present in table 2. Table 2 contains the sociodemographic, Technology, and Trip characteristics of combined models.

Table 2
The results of the relationship between AV purchasing decision and different characteristics for full samples (total sample size: 1712).

Models	**Model A: SdC**		**Model B: SdC + TeC**		**Model C: SdC + TeC + TrC**	
Coefficient	β	SE	β	SE	β	SE
Intercept (α)	-0.046	0.420	-0.085	0.422	-0.114	0.422
Sociodemographic characteristics						
Gender *(male)*	0.756***	0.107	0.758***	0.107	0.762***	0.108
Age *(18-24 years)*						
group1	-0.116	0.235	-0.120	0.235	-0.111	0.236
group2	-0.807***	0.242	-0.808***	0.242	-0.788**	0.242
group3	-1.278***	0.263	-1.275***	0.263	-1.228***	0.264
group4	-1.653***	0.284	-1.646***	0.284	-1.619***	0.284
Education level *(with a degree)*	0.619***	0.152	0.619***	0.152	0.596***	0.151
Income *(< $25,000)*						
group1	-0.068	0.282	-0.088	0.283	-0.090	0.283
group2	-0.351	0.279	-0.364	0.280	-0.353	0.280
group3	-0.283	0.279	-0.304	0.280	-0.276	0.280
group4	0.013	0.259	-0.011	0.260	0.004	0.280
License *(have a license)*	-0.356	0.287	-0.358	0.287	-0.222	0.260
Employment *(full time)*						
group1	-0.160	0.193	-1.500	0.194	-0.149	0.194
group2	-0.385	0.275	-0.386	0.275	-0.352	0.275
Residential location *(urban area)*	0.339**	0.106	0.338**	0.106	0.285**	0.109
Technology characteristic						
Smartphone age *(over 4 years)*			-0.203	0.181	-0.205	0.181
Trip characteristic						
Commute mode						
private car users = 1					-0.258*	0.112
AIC	2165.30		2166.10		2162.70	
Residual deviance	2135.3		2134.1		2128.7	
Log likelihood (LLH)	-1067.67		-1067.04		-1064.37	
Chi-square statistic	237.3		238.6		243.9	
P-value	0		0		0	
McFadden R^2	0.100		0.101		0.103	

*: Statistically significant at 0.95 level of confidence; **: Statistically significant at 0.99 level of confidence; ***: Statistically significant at 0.999 level of confidence.

Using group 0 as the reference, the results show that gender, age, living area, education, and commute mode have significant associations between AV purchasing willingness. Gender and most age groups are significant at 0.001 level. Furthermore, gender, education level, and residential location have a positive relation to the AV buying willingness whereas others are negative. Model c shows that private car users are less like to buy an AV compared to public transport users. In terms of the coefficients, the likelihood of private car users buying an AV is 77.32% lower than public transport users. Similarly, older people are not as willing to buy an AV compared to the young, take age group 1 (18-24 years) as the baseline, the likelihood of buying an AV between group 3 (35-44 years), group 4 (45-54 years) and group 5 (55-64 years) is around 46%, 29%, 20% lower, respectively. On the contrary, people living in the urban area and holds a higher education degree, both estimated by including only statistically significant at the 0.95 level of confidence or greater, are around 1.3 and 1.8 times as likely to buy an AV compared to those who do not.

3.2 Model fit

One of the model goodness standards is log-likelihood (LL). Logistic model use 'maximum likelihood' (DeMaris, logistical modelling, 1992) as evaluation, which means the larger the LL is, the better the model can be. We can see the LL of model C is -1064.37, slightly higher than model A (-1067.67) and model B (-1067.64), thus we can reject the null hypothesis that the regression model will not be influenced by independent variables. Similarly, the chi-square statistic of each model is: 243.9 (model c) > 238.6 (model b) > 237.3 (model a), with the p-value of 0, therefore, we reject the null hypothesis and conclude that there is a significant difference between the observed and the expected frequency, which means model C has the best model fit. Akaike's Information Criterion (AIC) is a concise indicator to evaluate the model fit (Otomo and Akaike, 1972), which can be calculated by: AIC=2k-lnL, where k means the number of parameters and L is likelihood function. The AIC ranges (-∞,0), with smaller value indicates a better goodness. Table 2 shows that model C has the lowest AIC value (2162.7) can be counted as the best model. The last equivalent model goodness parameter is Pseudo R2 (Domencich and McFadden, 1975), unlike LR with OLS estimation, there is no statistics that explain the explanation ability of the estimation value from the given model. The value of Pseudo R2 can be calculated by: where $R^2_{McFadden} = 1 - \frac{\log(LL_{full})}{\log(LL_{null})}$, where LL_{full} represents the likelihood of the fitted model, and LL_{full} represents other models. Apparently, model C shows the highest Pseudo R2 which indicates that model C has a good model fitness. Overall, we can conclude that model C presents the best model fit among the assessment parameters above.

3.3 Conclusion

With immediately development of smart transport technology, AV has been considered one transport to energy security and environmental protection. Using SR and RP travel survey data in April and June (2019), this paper estimates the AV ownership probability of public transport users and private car users by conducting binary logit model for AV interest choice component for part and full sample. Overall, the findings of this paper provide theoretical implications in assessing AV ownership willingness based on different variables. The model results clearly emphasise that urban residential location, individual difference (i.e., gender, younger age group people (use 18-24 years age group as the base line), education level, and commuting mode play important roles for buying an AV. Particularly, in terms of commuting mode, the key independent variable, public transport users show a high AV ownership willingness compared to private car users, which shows the potential of AV accessibility in traditional public transportation system.

References

Bagloee, S., Tavana, M., Asadi, M. and Oliver, T., 2016. Autonomous vehicles: challenges, opportunities, and future implications for transportation policies. *Journal of Modern Transportation*, **24**(4), pp. 284-303.

Booth, L., Norman, R. and Pettigrew, S., 2019. The potential implications of autonomous vehicles for active transport. *Journal of Transport & Health*, **15**, p. 100623.

Box, G.E.P. and P.W, Tidwell., 1962. Transformation of the independent variables. *Technometrics*, **4**(4), pp. 531-550.

Dai, J., Li, R. and Liu, Z., 2021. Does initial experience affect consumers' intention to use autonomous vehicles? Evidence from a field experiment in Beijing. *Accident Analysis & Prevention,* **149**, pp. 105778.

DeMaris, A., 1992. *Logit modeling*. 1st ed. Newbury Park, Calif.: SAGE.

Domencich, T. and McFadden, D., 1975. *Urban travel demand* - a behavioral analysis. Amsterdam: North-Holland Publishing Company.

Emberger, G. and Pfaffenbichler, P., 2020. A quantitative analysis of potential impacts of automated vehicles in Austria using a dynamic integrated land use and transport interaction model. *Transport Policy*, **98**, pp. 57-67.

Fagnant, D. and Kockelman, K., 2015. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice*, **77**, pp. 167-181.

Gandia, R., Antonialli, F., Cavazza, B., Neto, A., Lima, D., Sugano, J., Nicolai, I. and Zambalde, A., 2018. Autonomous vehicles: scientometric and bibliometric review. *Transport Reviews*, **39**(1), pp. 9-28.

Hong, J. and Thakuriah, P., 2018. Examining the relationship between different urbanization settings, smartphone use to access the Internet and trip frequencies. *Journal of Transport Geography*, **69**, pp. 11-18.

Jing, P., Huang, H., Ran, B., Zhan, F. and Shi, Y., 2019. Exploring the Factors Affecting Mode Choice Intention of Autonomous Vehicle Based on an Extended Theory of Planned Behavior—A Case Study in China. *Sustainability*, **11**(4), pp. 1155.

Lam, A., Leung, Y. and Chu, X., 2016. Autonomous-Vehicle Public Transportation System: Scheduling and Admission Control. *IEEE Transactions on Intelligent Transportation Systems*, **17**(5), pp. 1210-1226.

Long, J., 2003. *Regression models for categorical and limited dependent variables*. Thousand Oaks: Sage Publications.

Meilland, M., Comport, A. and Rives, P., 2014. Dense Omnidirectional RGB-D Mapping of Large-scale Outdoor Environments for Real-time Localization and Autonomous Navigation. *Journal of Field Robotics*, **32**(4), pp. 474-503.

Millar, R., 2011. *Maximum likelihood estimation and inference [electronic resource] : with examples in R, SAS, and ADMB*. 1st ed. Hoboken, N.J.: Wiley.

Otomo, T., Nakagawa, T. and Akaike, H., 1972. Statistical approach to computer control of cement rotary kilns. *Automatica*, **8**(1), pp. 35-48.

Salatiello, C. and Felver, T., 2017. Current Developments in Autonomous Vehicle Policy in the United States: Federalism's Influence in State and National Regulatory Law and Policy. *Global Jurist*, **18**(1).

Yuen, K., Huyen, D., Wang, X. and Qi, G., 2020. Factors Influencing the Adoption of Shared Autonomous Vehicles. *International Journal of Environmental Research and Public Health*, **17**(13), pp. 4868.

Zhang, W. and Guhathakurta, S., 2018. Residential Location Choice in the Era of Shared Autonomous Vehicles. *Journal of Planning Education and Research*, pp.0739456X1877606.