



UNDERSTANDING TRAVEL MODE CHOICE BEHAVIOURS IN THE ERA OF ADVANCED TRANSPORT TECHNOLOGIES

Cristian Jara Infante

Student number: 1033334, cjarainfante@student.unimelb.edu.au

***Executive Summary:** Autonomous vehicles, shared vehicles, and autonomous shuttles will be disruptive technologies that within 30 to 40 years would change the way people choose their travel mode. In this paper, A multinomial logit model was generated as a baseline to evaluate the impact of such technologies on different travel modes' characteristics such as travel time, travel cost, waiting time and availability. Autonomous vehicles used as private vehicles have the highest impact on the mode choice behaviour*

1 Introduction

Understanding how people choose the mode of transportation has been a task for transport researchers since the invention of the public transport and each time new technologies were emerging over the years, from faster and more efficient cars to magnetic trains, and electric vehicles. Nowadays, in a high-tech era, different disruptive technologies could have an effect on the transportation industry.

Technologies such as the internet of things and 5G would indirectly affect the travel mode choice by allowing the full development of connected autonomous vehicles (Datta et al., 2016; Shah et al., 2018) or by allowing real-time information by analysing big data with machine learning algorithms (Alsharif et al., 2020). On the other hand, providing real-time information to passengers affects the travel mode choice, since people may use this information to plan their trips minimising in this way their waiting time (Dziekan & Kottenhoff, 2007; Ingvardson et al., 2018).

Besides, in this digitalisation era, Transport Network Companies or best known as (TNCs) are continuously winning more market share in the taxi service, by offering a more convenient, door-to-door and dynamically-priced service (Schwieterman, 2019). However, the on-demand taxi service provided by TNCs is replacing the traditional taxi providers; nevertheless, taxi-sharing or ride-hailing services such as Uber Pool is becoming popular among young, well-educated and higher-income people. Taxi-sharing provides a much cheaper option than traditional taxi services, but with a higher travel time and a longer travelled distance (Danaf et al., 2014; Mohammadzadeh, 2020; Shen et al., 2020)

The most disruptive technology in the transport industry will be autonomous vehicles (AVs), which are called to change the way we choose our mode of transportation (Fagnant & Kockelman, 2015; Morrow et al., 2014; Truong et al., 2017). According to the Society of Automotive Engineers (SAE, 2018) there are six levels of automation from 0, representing zero autonomy, which is the current state of most of the vehicles, to five, representing full automation, where the vehicle performs all driving-related activities. For this paper, full automation level is assumed when evaluating its impact on travel mode choice.

This new technology can be used for different modes of transportation, including AVs for private usage, shared autonomous vehicles replacing a taxi-sharing or as autonomous shuttles used as micro-transit services (Childress et

al., 2015; Krueger et al., 2016; Litman, 2020). AVs will potentially increase travel demand by reducing traffic congestion, accident risks, emissions and reducing parking costs(Heinrichs & Cyganski, 2015).

In this paper, changes in the mode choice behaviour are evaluated for the three uses of AVs, as a private automated vehicle, as a share automated vehicle, and as an autonomous shuttle. This is made by first establishing a baseline travel mode discrete choice model to later change the variables that AVs would impact and evaluate the changes in travel mode share.

2 Methodology

2.1 Base Model

The impact of the new technologies was evaluated using a discrete choice model as a baseline. Several authors have proposed a number of behavioural models, including the Probit, Multinomial Logit (MNL), Nested Logit, Mixed Logit, among others (M. Ben-Akiva & Bierlaire, 1999; M. E. Ben-Akiva & Lerman, 1985; Train, 2009). For this paper, the Multinomial Logit approach has been chosen since it is one of the most trackable and widespread models among researchers.

2.1.1 Theoretical background

Multinomial logit models are based on the random utility maximisation theory, which assumes that the decision-maker behave as a "Homo Economicus" who will choose the alternative that provides him with the maximum benefit (Koppelman & Bhat, 2006). The utility function is defined by (1):

$$U_{in} = V_{in} + \varepsilon_{in} \tag{1}$$

where, U_{in} is the true perceived utility of the alternative i to the individual n
 V_{in} is the deterministic utility estimated of the alternative i to the individual n
 ε_{in} is the random error term due to unobserved utility

And,

$$V_{in} = ASC + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k \tag{2}$$

Where, ASC is the Alternative Specific Constant
 β_k is the parameter estimated to the explanatory variable X_k

The random error terms are assumed independently and identically distributed (i.i.d) following a Gumbel distribution. Therefore the probability that a decision-maker n chooses the alternative i is defined by:

$$P_{in} = \frac{e^{V_{in}}}{\sum_j e^{V_{nj}}} \tag{3}$$

Maximum Likelihood Estimation (MLE) is used to estimate the parameters of the model B_k , where the likelihood of an individual n choosing the alternative i that he actually chooses is defined by:

$$L = \prod_i (P_{in})^{y_{in}} \tag{4}$$

Where, y_{in} represents if individual n chose alternative i , if yes $y_{in} = 1$ and zero otherwise. Therefore, the likelihood of every single individual choosing the alternative that has been observed is:

$$L(\beta) = \prod_{n=1}^N \prod_i (P_{ni})^{y_{ni}} \quad (5)$$

Where, N represents the size of the sample, and β is the vector containing the parameters. To improve estimation and numerical performance the Log-Likelihood function is used.

$$LL(\beta) = \sum_{n=1}^N \sum_i y_{ni} \ln(P_{ni}) \quad (6)$$

Since the complexity of the model is relatively high, a python package called PandasBiogeme (Bierlaire, 2020) is used to estimate the parameters based on the theory described above.

2.1.2 Data Sources

Different sources of data were used to increase the model accuracy by adding explanatory variables. These sources provided different types of explanatory variables such as mode of transportation attributes, including travel time, travel cost, among others. Also, decision-makers' socio-economic characteristics such as economic status, gender, working status, car ownership, among others. Besides variables representing the trip characteristics such as, whether the trip was made on peak time or a weekday. Finally, this paper includes the weather condition on the day of the trip as explanatory variables.

2.1.2.1 Victorian Integrated Survey of Travel and Activity (VISTA)

VISTA is an ongoing study of household travel activity among Greater Melbourne and Geelong (*Victorian Integrated Survey of Travel and Activity*, 2016). The VISTA data from the year 2012 to 2016 was used to build the base model. This survey provides most of the socio-economic explanatory variables and includes the characteristics of the trips.

VISTA includes more than 120,000 trips from 2012-2016; however, during the data cleaning process, some of them were excluded from this study. Trips are referenced to the smallest geographical areas available known as SA1 (Statistical Area Level 1), to keep the privacy of the respondents; nevertheless, a significant part of the trips was performed within one single SA1 which made impossible to estimate the modes' explanatory variables such as travel time, distance or cost using Google Maps, TripGo or PTV Journey Planner.

In addition, it was found that these traffic/transport applications lose reliability when the SA1 zones are large (>100 square km), because it leads to a significant difference between estimated travel time, distance, cost from the apps and the actual variables experienced by the respondent. Consequently, those differences could lead to building a non-representative model.

2.1.2.2 TripGo API

TripGo Application Programming Interface or TripGo API is a digital tool that allows retrieving data from different transportation modes (SkedGo, 2020). Free developer accounts were used. The information was retrieved by using a short program coded in Python. Five modes of transportations were requested to API, walking, cycling, public Transport, car, taxi. Note that the only inter-modal mode of transportation considered in this study is walk-transit, and no other inter-modal options such as car-transit or bike-transit were considered.

TripGO API does not allow to retrieve data from the past; therefore it is assumed that no significant changes have been made in the Melbourne transport network from 2016 to 2020 that could have considerably impacted on distance or travel time.

The retrieved information for each mode was stored in JSON format for further review. The raw JSON data was parsed using another short program coded in Python for each mode to get the relevant information. For walk and bike modes, the python code chooses the default best trip given by TripGo and get the relevant attributes of that trip. However, it was found that the best trips given by TripGo for car, taxi and transit modes presented inconsistency such as departure times before the start time or not selecting the shortest trips. Therefore the following procedure is used to select the best trip for these modes:

- Choose all the trips where their departure time is equal or greater than the start time from VISTA table 't' (STARTIME)

- From that group of trips, the one with the earliest arrival time is selected.
- From the trip chosen, the program gets its relevant attributes.

2.1.2.3 *Public Transport Victoria (PTV)*

TripGO estimation of frequency and cost of the public transport option has been found to be neither consistent nor reliable. Hence, different Python codes were developed in order to estimate the headway and cost of the public transport mode.

For the cost estimation, updated 2020 full fares, concession fares and conditions were used from the Victorian Fares and Ticketing Conditions and the Victorian Regional Bus Fares Supplement and the PTV API (Public Transport Victoria, 2020a, 2020b). The following procedure was used.

- Obtain the ticket zone (1-13) for start and end stops of all the legs in the trip, either from GTFS, GIS or PTV API.
- If not ticket zone available, verify regional fare.
- If not regional, calculate fare taking into account the lowest and highest ticket zone visited during the trip and whether the trip was made during peak times.
- If regional, calculate fare taking into account the Charging Units, each regional leg is calculated separately.
- Apply concession conditions according to the Victorian Fare and Ticketing Conditions-2020.

On the other hand, PTV timetable and geographic information (GTFS) was used to estimate the headway according to the following procedure:

- First leg's start-stop and end-stop identification from TripGo data
- Services available according to GTFS between those stops for the day and direction of the trip. It can include multiple modes and different lines.
- Select the service with the closest departure time to the start time from TripGo data
- Select the closest service before and after to the selected service above. Those three services may be in different lines or even different modes.
- Calculate the average headway between those three services.

2.1.2.4 *RACV car running costs*

TripGo car cost data was found to be not consistent and reliable; therefore, the cost for car mode was calculated using average running costs (RACV, 2019) according to the following procedure.

- Calculate the average running cost (cents/km) according to the type of vehicles owned by the household according to VISTA data.
- Using travel distance from TripGo calculate the total cost of the trip.

2.1.2.5 *Bureau of Meteorology (BOM)*

The willingness to choose active modes, walking and cycling, as the preferred transportation mode is usually affected by weather conditions such as extreme temperatures, rainfall or wind (Saneinejad et al., 2012; Ton et al., 2019), besides some studies show that bad weather conditions also have a negative impact on the ridership of some public transport modes such as a bus or light rail (Petrović et al., 2020; Tao et al., 2018).

Therefore, to consider the impact of weather conditions on travel mode choice, temperature and rainfall data were acquired from the Bureau of Meteorology (BOM) which offers a free database for daily minimum and maximum temperature as well as daily millimetres of precipitation (Bureau of Meteorology, 2020). This data was obtained from 257 stations within Victoria and for the years 2012 to 2016.

2.1.3 **Choice Set**

To accurately evaluate the impact of AVs, SAVs, TNCs and autonomous shuttle bus, the model includes a choice set of 5 alternatives:

- Walk
- Bike

- Public transport (includes bus, tram and train)
- Car (includes driving a car or as a passenger)
- Taxi

Trips, where the travel mode chosen by the respondent was motorbike, other or non-stated, were excluded from the analysis.

2.1.4 Model Specification

After an iteration process and trying different model specification, the variables and parameters used in the final model are described in the following tables.

2.1.4.1 Specific travel mode attributes

Mode inherent characteristics, such as travel time, cost, co2 emission, among others are included in this category; Table 1 summarises those characteristics considered in the final model.

Most of these attributes were directly acquired from TripGo API; however, transit waiting time, transit cost, and parking cost were separately estimated to improve model accuracy.

Usually, travel mode choice models consider waiting time as $W = 0.5h$ where h is the service headway; assuming that passengers arrive at random times between one service and the next one. However, several studies have found evidence of non-random behaviour of passengers. From simple linear relationships such as $W = 1.79 + 0.14h$ or $W = 2.34 + 0.26h$ (O'Flaherty & Mancan, 1970; Seddon & Day, 1974) setting a threshold where passengers start to not arriving at random times (Bowman & Turnquist, 1981; Fan & Machemehl, 2009). To more complex models proposing different distributions such as Extreme value, Exponential Lognormal, among others (Gong et al., 2016; Reza Amin-Naseri & Baradaran, 2015; Webb et al., 2020).

For this paper, a beta-mixture waiting time is used as proposed by Ingvardson et al. (2018). This approach suggests that passengers waiting time follow a mixture of uniform (random arrivals) and beta distribution where the expected waiting time $E[W]$ is defined by:

$$E[W] = H \times \left(\zeta \frac{\alpha}{\alpha + \beta} + (1 - \zeta) \right) \quad (7)$$

Where, H	Service headway
ζ	Share of passengers following a beta distribution
$1 - \zeta$	Share of passengers following a uniform distribution
α, β	Shape parameters of the beta distribution

The values of ζ, α and β were corrected from the original paper by Ingvardson et al., (2020); however since those values, in the paper, are not continuous but discrete to specific headways of 5, 10, 20, 30 and 60 min; therefore a simple regression was executed, as shown in Figure 1, to generalise the distribution to any other headway value.

Using the headways estimated from PTV timetables, an expected waiting time is calculated. It is important to point out that since waiting time can be divided into open waiting time and hidden waiting time. The open waiting time corresponds to the time the passenger is waiting for a vehicle at the stop; therefore, the calculated waiting time following the beta-mixture distribution corresponds to this category.

Hidden waiting time corresponds to the duration of time between the time the person wants to travel and the time the person start its trip. Therefore hidden waiting time is calculated as follows:

$$HWT = Max((ST - DT) - AT - OWT, 0) \quad (8)$$

Where, ST Is the time the person started the trip according to VISTA (hh:min:sec)

- DT Is the time the first leg service departs from the stop according to TripGo (hh:min:sec)
- AT Walking time from start point to first leg stop according to TripGo (mins)
- OWT Open waiting time (min)

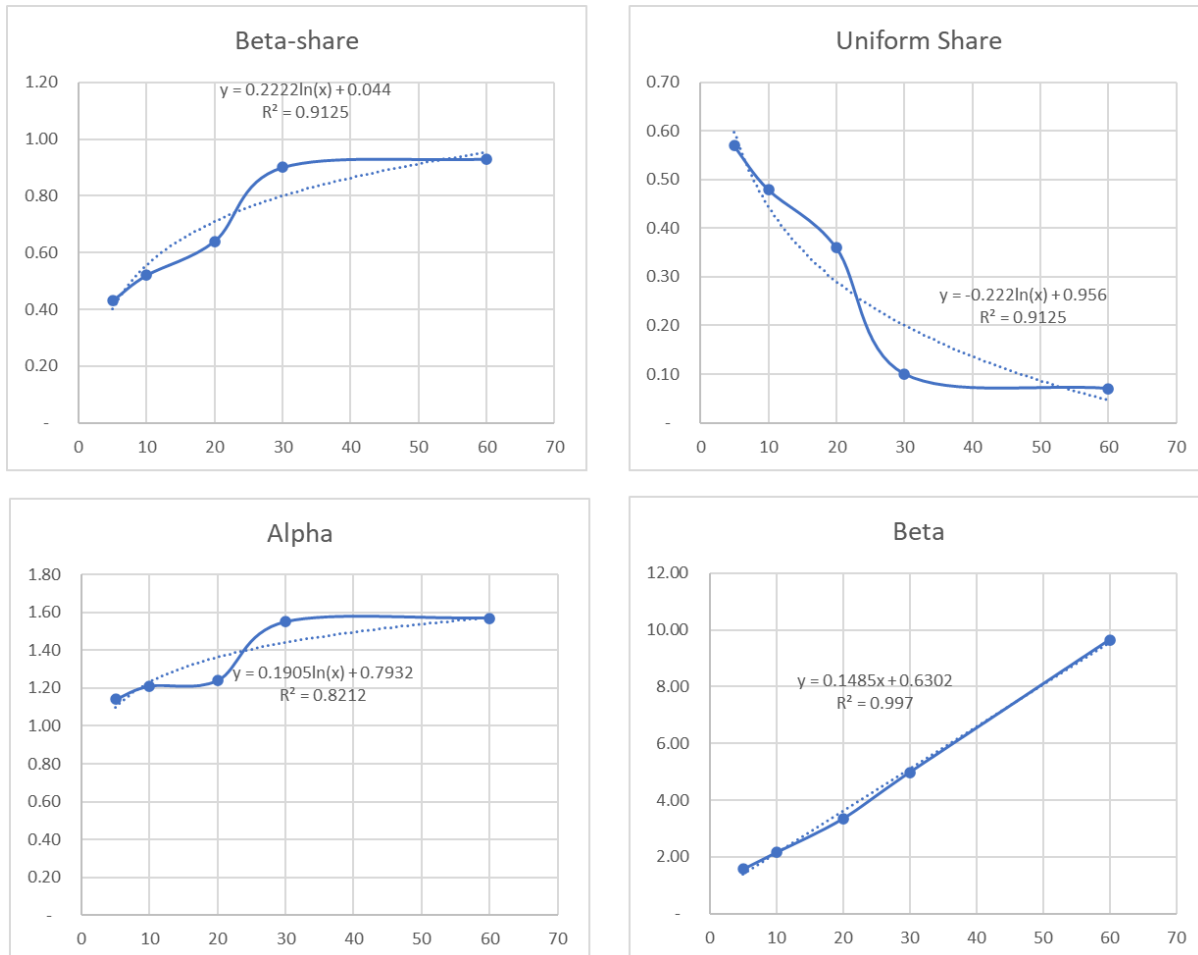


Figure 1 Regression for beta and uniform share and shape parameters of the beta-mixture distribution.

Table 1 Specific travel mode attributes

Variable Name	Parameter	Unit	Description	Source
walk_cal	B_SM_CAL	(cal)	Trip's calories consumption when walking	TripGo API
walk_tt	B_WALK_TT	(min)	Travel time when walking	TripGo API
bike_cal	B_SM_CAL	(cal)	Trip's calories consumption when using a bike	TripGo API
bike_tt	B_BIKE_TT	(min)	Travel time by walking when using a bike	TripGo API
bike_friendly	B_BIKE_FRIENDLY	(%)	Percentage of bicycle-friendly road (with bicycle lane)	TripGo API
transit_onboardtime	B_PT_OBT	(min)	In vehicle time for public transport	TripGo API
transit_openwaittime	B_PT_OPEN_WT	(min)	Average time in minutes between passenger arrival to first leg stop and service departure.	PTV*
transit_hidwaittime	B_PT_HIDDEN_WT	(min)	Time in minutes between the time that a passenger wants to travel and the time a passenger can start the trip due to the frequency of the service.	PTV*

transit_walktime	B_PT_WALK_TIME	(min)	Access or Egress time to the public transport service assuming walking only.	TripGo API
transit_transfers	B_PT_TRANSFERS	(unit)	Number of transfers	TripGo API
transit_transferwait	B_PT_TRASNF_WAIT	(min)	Waiting time between transfers	TripGo API
transit_cost	B_COST	(AUD)	Cost in AUD for using the public transport system	PTV*
transit_co2	B_CO2	CO2	Co2 emissions when travelling in public transport	TripGo API
taxi_tt	B_TAXI_TT	(min)	Travel time when travelling in a taxi (Includes waiting time)**	TripGo API
taxi_tcost	B_COST	(AUD)	Cost in AUD when travelling in a taxi	TripGo API
taxi_co2	B_CO2	(kg)	Co2 emissions when travelling in a taxi	TripGo API
car_tt	B_CAR_TT	(min)	Travel time when travelling in a car	TripGo API
car_tcost	B_COST	(AUD)	Cost in AUD when travelling in a car (includes Parking Cost)***	TripGo API
car_taxi_co2	B_CO2	(kg)	Co2 emissions when travelling in a car	TripGo API

* Estimation using data from PTV. The detail of the calculation procedure is described above

** Waiting time estimation from TripGo around (8 min) when no initial walking is needed.

*** Parking Cost is an estimation detailed above.

2.1.4.2 Socio-Economic attributes

Many socio-economic attributes are available in VISTA database such as household income, age-group, working and studying status. However, any model estimated with Biogeme in which those attributes were included, resulted in positive values for travel cost, which is not a valid representation of the reality. Since travel cost is one of the most critical attributes that the new technologies would affect (increasing or decreasing), therefore, just some of those attributes were included. Table 2 summarises all the attributes considered to the iterative process of running models, the ones with an assigned parameter are the ones used for the final model.

Table 2 Socio-economic Attributes

Variable Name	Parameters	Description	Source
eq_inc_rank		Rank calculated dividing every equivalent economic income by its average	VISTA*
Gender	B_PT_GENDER B_CAR_GENDER B_TAXI_GENDER	1 for male, 0 otherwise	VISTA
bef_boomers		1 if the person was born before 1943, 0 otherwise	VISTA
boomers		1 if the person was born on or after 1943 and before 1961, 0 otherwise	VISTA
genX		1 if person was born on or after 1961 and before 1982, 0 otherwise	VISTA
genY		1 if the person was born on or after 1982 and before 2001	VISTA
stu_status		1, if the person is studying, 0 otherwise	VISTA
work_status		1, if the person is working wither casual, parttime or fulltime, 0 otherwise	VISTA
veh_per_person	B_VEH_PER_PERSON	The average number of vehicles in the household, not considering motorbikes or scooter	VISTA
adult_bikes	B_ADULT_BIKES	Number of adult bikes in the household	VISTA
car_license	B_CAR_LICENCE	1 if the person has a driving license, 0 otherwise	VISTA

* Estimation using data from PTV. The detail of the calculation procedure is described above

2.1.4.3 Trip Characteristics

Some trip characteristics are available in VISTA database such as, whether the person travels in a workday or during pick ours, and the purpose of the trip. However, those were not included in the final model since their parameters were not relevant for this study. Table 3 summarises the trip characteristics used during the running of models, but none of them was included in the final model.

Table 3 Trip Characteristics

Variable Name	Parameters	Description	Source
travel_period		1 if travelling during peak periods, 0 otherwise	VISTA
trip_purp_work		1 if the purpose of the trip was work-related, 0 otherwise	VISTA
trip_purp_edu		1 if the purpose of the trip was education-related, 0 otherwise	VISTA
trip_purp_pick		1 if the purpose of the trip was to pick up or drop off something or someone, 0 otherwise	VISTA
week_day		1 if the trip happened from Monday to Friday, 0 otherwise	VISTA

2.1.4.4 Weather Conditions

Several studies have shown the impact of weather conditions on the travel mode choice, mainly on active modes such as walking, or cycling (Saneinejad et al., 2012; Ton et al., 2019). Besides, extreme temperatures or rainfall has been shown to affect the choice of public transport as the preferred travel mode (Petrović et al., 2020; Tao et al., 2018).

Since BOM provides just data for daily min and max temperature for free, a temperature for a departure time of each trip was estimated with the method proposed by Parton & Logan in 1981 by using the 5 closest stations to the departure latitude and longitude, considering the elevation obtain from Google Elevation (You et al., 2008). The algorithms are coded in python.

The maximum and minimum temperature for each trip at the departure point is estimated using the inverse distance weighting method described by You et al.

$$x = \frac{\sum_{i=1}^5 [y_i w_i]}{\sum_{i=1}^5 w_i} \quad (9)$$

Where, x is the predicted temperature at the departure point

y_i is the min or max temperature in the i th closest station to the departure point from BOM

w_i is the inverse of the distance from the i th closes station to the departure point from BOM

The temperature on the departure point for a given time during the day and night is estimated by 1 and 2 respectively.

$$T_i = (T_x - T_N) \sin\left(\frac{\pi m}{Y + 2a}\right) + T_N \quad (10)$$

$$T_i = T_N + (T_s - T_N) \exp(-bn/Z) \quad (11)$$

Where, T_i is the temperature on the departure point at the i th hour

T_x is the maximum temperature on the departure point calculated with (9)

T_N is the minimum temperature on the departure point calculated with (9)

T_s is the temperature at sunset, calculated with (10)

m is the number of hours after the minimum temperature occurs until the sunset (assumed that min temperature occurs at sunrise)

n is the number of hours after sunset until the time of the minimum temperature

a is the lag coefficient for the maximum temperature (assumed to be 1.5)

b is the nighttime temperature coefficient (assumed to be 4)

Y is day length in hours

Variable Name	Parameters	Description	Source
temp_coeff	B_SM_DAY_TEMP		BOM*
rainfall	B_SM_RAINFALL	mm of precipitation during the day of the trip	BOM*
bad_weahter	B_PT_BAD_WEATHER	Sum of temp_coeff and rainfall	BOM*

* Temperature and rainfall for the location and time of the start of the trip were estimated using BOM data, following the procedure described above.

2.2 Impact of New technologies

Four sets of scenarios have been designed to explore how new technologies could affect the travel mode shares. First, an isolated case for each use of AVs: Private, shared, and transit. And a fourth scenario with a combined effect. Each scenario assumes the total replacement of traditional vehicles with the new technology

For the first scenario, an average operating cost of 1\$/km, which is higher than the average 0.77 \$/km from the model, is assumed (Litman, 2020). A factor of reduction 0.65 is assumed for in-vehicle travel time due to the total automation of the fleet and a reduced parking cost to 50% (Childress et al., 2015).

For the second scenario, Litman suggests that shared autonomous vehicles will probably have an average operating cost of 0.7\$ per vehicle per km and a 0.4\$ per passenger per km when using as a taxi. Besides, since travel time in shared taxis is higher than regular taxis, due to picking up and dropping off people on the way to the traveller's destination, travel time is increased in 16.7 % (Young et al., 2020).

For the third scenario, Litman suggests that shared autonomous vehicles can be used as micro-transit services; therefore, the same operating cost is assumed per vehicle per km. In addition, since micro-transit services operate at high frequencies, the hidden waiting time would tend to disappear, and the open waiting time would tend to decrease. A 50% reduction in open waiting time is considered and a 100% of reduction in hidden waiting time is considered.

For the fourth scenario, we assumed that all the private vehicles are autonomous vehicles, all the taxi services run under a ride-share option and all the public transport has switch to micro-transit services

Table 4 Scenario descriptions

Cost	Autonomous Vehicle	Shared Autonomous Vehicle (Taxi)	Shared Autonomous Ride (Transit)
Scenario 1	- Operating average cost of 1 \$/km - Time travel factor of 0.65	No changes	No changes

Scenario 2	No changes	- Increase travel time by 16.7%	No changes
		- Cost of \$0.4 per km per passenger	
Scenario 3	No changes	No changes	- Cost of \$0.4 per km per passenger
			- 0 hidden time
			- 50% open waiting time reduction
Scenario 4	- Operating average cost of 1 \$/km	- Increase travel time by 16.7%	
	- Time travel factor of 0.65	- Cost of \$0.4 per km per passenger	

3 Results and Discussion

3.1 Best Base Model

Using Biogeme to estimate the parameters, many models were run with different specifications in an iterative process, the model with the most significant parameters and higher log-likelihood was chosen. The final model has 85,962 trips, with a goodness of fit value (rho-squared) of 0.661 and a final log-likelihood of -45,552.29. The estimation report can be found in Table 5. The utility functions used to estimate this model, and the values of its estimated parameters are presented below.

From the parameters estimated, it can be concluded that the value of time for each mode is not the same being the highest the open waiting time which its value is ten times higher than the on-vehicle transit travel time. Weather conditions do have a negative impact on the utility function for active modes but apparently the bad_weather variable, which is a combination of temperature and rainfall, for public transport has a positive impact on its utility function. This could be interpreted as that on days with extreme weather conditions, people without a vehicle swift from active modes to public transport for being more convenient.

$$\begin{aligned}
 U_WAK &= ASC_WALK + \\
 &\quad B_WALK_TT * walk_tt + \\
 &\quad B_SM_CAL * walk_cal + \\
 &\quad B_SM_DAY_TEMP * temp_coeff + \\
 &\quad B_SM_RAINFALL * RAINFALL
 \end{aligned}$$

$$\begin{aligned}
 U_TAXI &= ASC_TAXI + \\
 &\quad B_TAXI_TT * taxi_tt + \\
 &\quad B_CO2 * taxi_co2 + \\
 &\quad B_COST * taxi_tcost + \\
 &\quad B_TAXI_GENDER * gender
 \end{aligned}$$

$$\begin{aligned}
 U_CAR &= ASC_CAR + \\
 &\quad B_CAR_TT * car_tt + \\
 &\quad B_CO2 * car_co2 + \\
 &\quad B_COST * car_tcost + \\
 &\quad B_CAR_GENDER * gender +
 \end{aligned}$$

$$\begin{aligned}
 U_BIKE &= ASC_BIKE + \\
 &\quad B_BIKE_TT * bike_tt + \\
 &\quad B_SM_CAL * bike_cal + \\
 &\quad B_BIKE_FRIENDLY * bike_friendly + \\
 &\quad B_SM_DAY_TEMP * temp_coeff + \\
 &\quad B_SM_RAINFALL * RAINFALL + \\
 &\quad B_ADULT_BIKES * adult_bikes
 \end{aligned}$$

$$\begin{aligned}
 U_PT &= ASC_PT + \\
 &\quad B_PT_OBT * transit_onboardtime + \\
 &\quad B_PT_OPEN_WT * transit_openwaittime + \\
 &\quad B_PT_HIDDEN_WT * transit_hidwaittime + \\
 &\quad B_PT_WALK_TIME * transit_walktime + \\
 &\quad B_PT_TRANSFERS * transit_transfers + \\
 &\quad B_PT_TRANSF_WAIT * transit_transferwait + \\
 &\quad B_CO2 * transit_co2 + \\
 &\quad B_COST * transit_cost +
 \end{aligned}$$

B_VEH_PER_PERSON*veh_per_person +\
 B_CAR_LICENCE*car_licence

B_PT_GENDER*gender+\
 B_PT_BAD_WEATHER*bad_weather

Table 5 Estimation report for best model

Number of estimated parameters:	27
Sample size:	85,962
Excluded observations:	-
Init log likelihood:	- 134,393
Final log likelihood:	- 45,552
Likelihood ratio test for the init. model:	177,682
Rho-square for the init. model:	0.661
Rho-square-bar for the init. model:	0.661
Akaike Information Criterion:	91,158.58
Bayesian Information Criterion:	91,411.35
Final gradient norm:	0.03
Database readings:	3,818
Iterations:	3,409
Nbr of threads:	8.00

Table 6 Parameters values, standard deviation, t-test and p-test

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_BIKE	-5.27	0.093	-56.6	0	0.124	-42.4	0
ASC_CAR	-0.861	0.0357	-24.1	0	0.0445	-19.4	0
ASC_PT	-1.55	0.0665	-23.4	0	0.077	-20.2	0
ASC_TAXI	-5.78	0.127	-45.6	0	0.123	-47	0
B_ADULT_BIKES	0.427	0.0121	35.4	0	0.0113	37.8	0
B_BIKE_FRIENDLY	0.662	0.128	5.17	2.32E-07	0.159	4.16	3.18E-05
B_BIKE_TT	-0.0351	0.0221	-1.59	0.112	0.036	-0.974	0.33
B_CAR_GENDER	-0.225	0.022	-10.2	0	0.0219	-10.3	0
B_CAR_LICENCE	0.295	0.0255	11.6	0	0.0261	11.3	0
B_CAR_TT	-0.0771	0.00137	-56.5	0	0.00149	-51.7	0
B_CO2	0.0769	0.0179	4.29	1.81E-05	0.0195	3.94	8.02E-05
B_COST	-0.0123	0.000834	-14.7	0	0.000825	-14.9	0
B_PT_BAD_WEATHER	0.00861	0.00314	2.75	0.00601	0.00313	2.75	0.00587
B_PT_GENDER	-0.215	0.0367	-5.85	4.84E-09	0.0362	-5.93	3.10E-09
B_PT_HIDDEN_WT	-0.00139	0.000818	-1.7	0.0898	0.00126	-1.1	0.272
B_PT_OBT	-0.0238	0.00151	-15.8	0	0.0016	-14.9	0

B_PT_OPEN_WT	-0.24	0.0081	-29.6	0	0.00872	-27.5	0
B_PT_TRANSFERS	-0.413	0.0413	-10	0	0.101	-4.08	4.47E-05
B_PT_TRANSF_WAIT	-0.0131	0.0047	-2.79	0.00532	0.0156	-0.842	0.4
B_PT_WALK_TIME	-0.0322	0.00163	-19.8	0	0.00177	-18.2	0
B_SM_CAL	-0.00558	0.00407	-1.37	0.17	0.00654	-0.853	0.394
B_SM_DAY_TEMP	-0.0202	0.00395	-5.1	3.31E-07	0.00394	-5.12	3.13E-07
B_SM_RAINFALL	-0.0177	0.00301	-5.88	4.11E-09	0.00297	-5.97	2.37E-09
B_TAXI_GENDER	-0.324	0.123	-2.63	0.00855	0.122	-2.66	0.00789
B_TAXI_TT	-0.0633	0.00421	-15.1	0	0.00387	-16.4	0
B_VEH_PER_PERSON	0.259	0.0314	8.24	2.22E-16	0.0316	8.19	2.22E-16
B_WALK_TT	-0.0617	0.01	-6.15	7.63E-10	0.0162	-3.82	0.000134

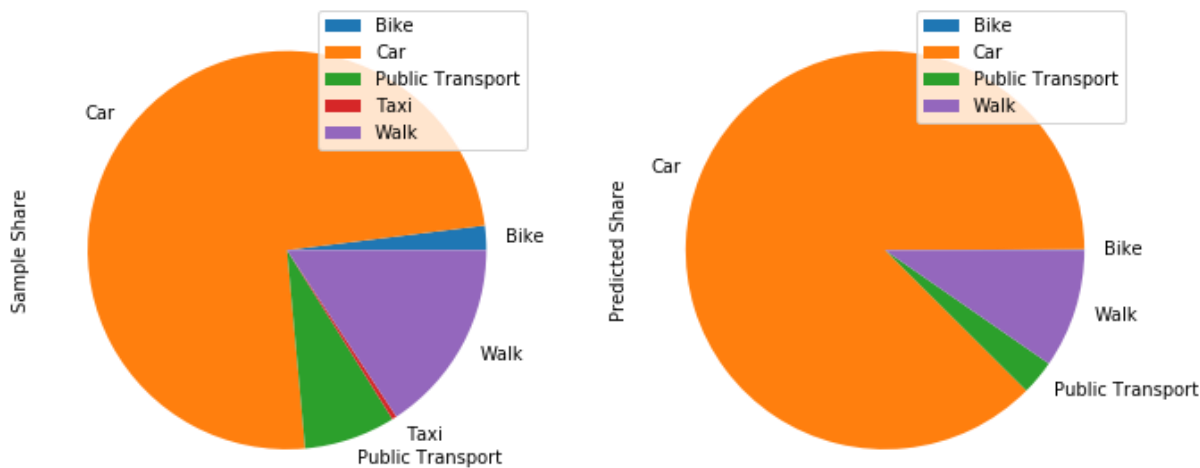
3.1.1 Predicted travel mode share with the best model

Considering the parameters, estimated with Biogeme, the predicted travel mode choice for each trip is calculated using (3). Each trip in the sample represents a certain number of trips in the population; this is reflected in the VISTA table 't'. Table 1Table 7 summarises the mode amount of trips for each chosen mode and for each predicted mode.

Table 7 Comparison between Chosen and Predicted travel mode share

	Chosen_Trips	Chosen_Share	Predicted	Predicted_Share
Bike	181,294	2.0 %	7,118	0.1%
Car	6,841,363	74.5 %	8,034,365	87.5 %
Public Transport	680,812	7.4 %	257,886	2.8 %
Taxi	35,341	0.4 %	0	0 %
Walk	1,446,495	15.7 %	885,936	9.6 %
	9,185,305	100%		8,678,072

Table 8 Pie chart showing share distribution among the different modes



3.2 Impact of AVs (replacing car option)

Despite the average cost of operating a private car increased from 0.76 to 1.2, reducing travel time to 65 per cent of its original value and reducing parking cost by 50 % have a significant impact on the utility of choosing the car as a

travel mode when AVs replace every car. Travel share of the car mode increased from 87.5 % as predicted in the base model to 91.8%. The increment in the utility of car reduces by more than half the share of the bike and public transport; from 0.1 to 0.04% and form 2.8 to 1.1% respectively. This means that around 60% of the trips, initially assigned to bike and public transport switch to the car mode. However, the walking mode share was reduced by 25% from 9.6 to 7.1%. This small reduction may have a cause in the people living in households with 0 vehicles or without a car license. This scenario could change as the due to as AVs are introduced; teenagers below the license age could be able to ride an AV without supervision.

Table 9 Trips and travel mode share scenario 1

	Chosen Trips	Chosen Share	Base Model Trips	Base Model Share	After Avs Trips	After Avs Share
Bike	181,294	2.0%	7,118	0.1%	3,414	0.04%
Car	6,841,363	74.5%	8,034,365	87.5%	8,434,322	91.8%
Public Transport	680,812	7.4%	257,886	2.8%	96,842	1.1%
Taxi	35,341	0.4%	-	0.0%	-	0.0%
Walk	1,446,495	15.7%	885,936	9.6%	650,726	7.1%
	9,185,305	100%	9,185,305	100%	9,185,304	100%

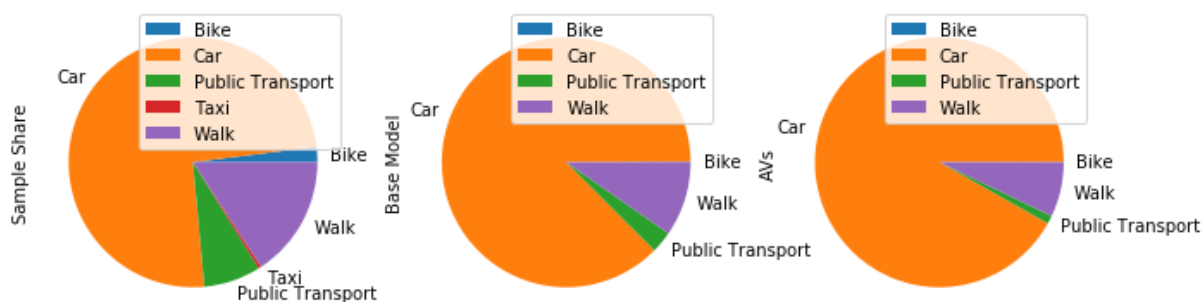


Figure 2 Pie chart showing share distribution among the different modes for scenario 1

3.3 Impact of Shared Autonomous Vehicles (replacing taxi option)

Since the trips in the sample, where the chosen travel mode was a taxi, represents less than 0.5 per cent, the utility function for a taxi from the baseline logit model may not be representative. This is reflected on this results, where the impact of reducing the travel cost from 4.33 \$/km, on average, to 0.4 \$/km per passenger per km according to the literature, is not enough to produce any measurable change on the travel mode share.

3.4 Impact of Shared Autonomous Rides (replacing transit)

Assuming 0.4\$ per person per km for public transport has a mixed impact on the utility. Approximately 72% of the analysed trips have travel distances less than 8.75 km with an average fare of 3.5 \$. This trips would benefit from micro-transit price since the average cost for trips less than 8.75 km would be 1.75\$. However, trips greater than 8.75

km and less than 20km costs on average 3.8\$ due to the concessions, and fixed fares and this trips would not benefit from the micro-transit potential fare.

A reduction of 50% in the open waiting time and 100% in the hidden waiting time for public transport has a significant impact on the perceived utility.

In this scenario, the public transport share shows an increase of 65% going from 2.8% to 4.6% of the total trips.

Table 10 Trips and travel mode share scenario 3

	Chosen Trips	Chosen Share	Base Model Trips	Base Model Share	After SAVs Trips	After SARs Share
Bike	181,294	2.0%	7,118	0.1%	6,599	0.07%
Car	6,841,363	74.5%	8,034,365	87.5%	7,873,058	85.7%
Public Transport	680,812	7.4%	257,886	2.8%	419,712	4.6%
Taxi	35,341	0.4%	-	0.0%	-	0.0%
Walk	1,446,495	15.7%	885,936	9.6%	885,936	9.6%
	9,185,305	100%	9,185,305	100%	9,185,305	100%

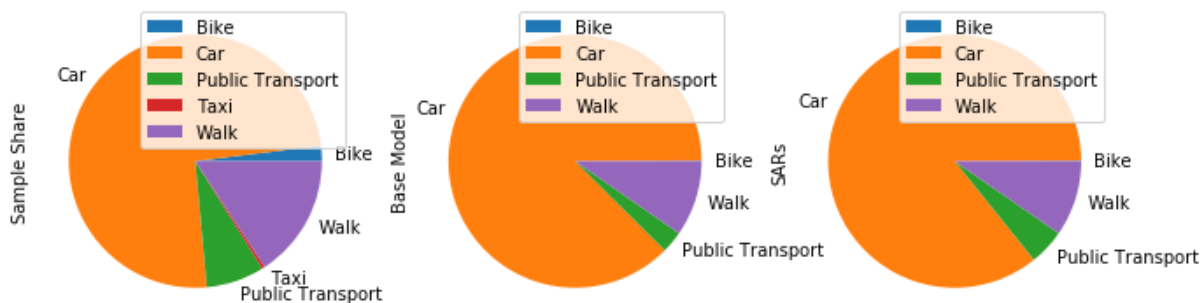


Figure 3 Pie chart showing share distribution among the different modes for scenario 3

3.5 Combined Impact

The combined impact of the autonomous vehicle is presented in Table 11 and Figure 4. Definitely impact of the reduction of 65% in travel time when using AVs as private vehicles prevail. However public transport share is slightly better than the first scenario 1.2% vs 1.1% of the total trips.

Table 11 Trips and travel mode share scenario 4

	Chosen Trips	Chosen Share	Base Model Trips	Base Model Share	Scenario 4 Trips	Scenario 4 Share
Bike	181,294	2.0%	7,118	0.1%	3,414	0.04%
Car	6,841,363	74.5%	8,034,365	87.5%	8,424,457	91.7%

Public Transport	680,812	7.4%	257,886	2.8%	106,707	1.2%
Taxi	35,341	0.4%	-	0.0%	-	0.0%
Walk	1,446,495	15.7%	885,936	9.6%	650,726	7.1%
	9,185,305	100%	9,185,305	100%	9,185,304	100%

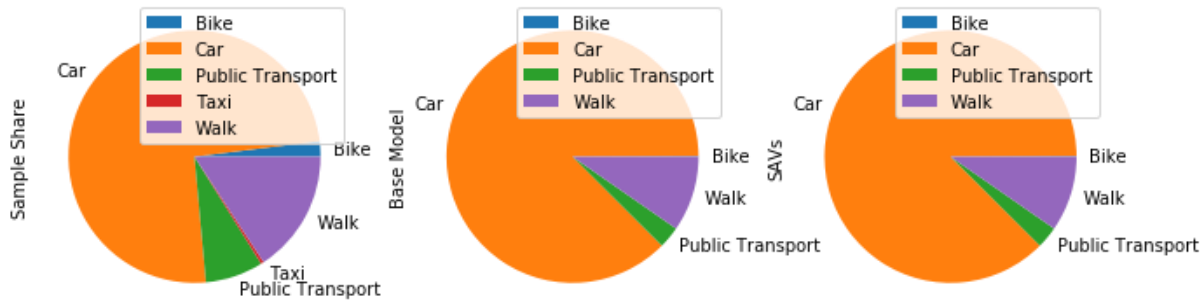


Figure 4 Pie chart showing share distribution among the different modes for scenario 4

4 Conclusions

This study shows the significant impact of AVs used as private vehicles in the future travel mode share, 60% of the trips, initially assigned to bike and public transport modes, would shift to private autonomous vehicles.

Taxi as a mode share could not be adequately evaluated due to the low amount of taxi trips present in the sample; therefore, further research is needed to predict the impact of SAVs accurately. An independent survey with a bigger sample of taxi trips or the use TNC's data to understand the mode share of taxi rides.

The potential impact of shared autonomous rides used as micro-transit modes shows that by reducing 50% of the open waiting time and 100% of the hidden waiting time by offering a high-frequency autonomous micro-transit frequency would double the public transport share in the third scenario.

However, a most likely scenario, where all the modes are combined, and autonomous vehicles are used not only as private vehicles but also as taxis and as micro-transit services, shows that the impact on waiting the waiting time of micro-transit services is not relevant to compare to the reduced travel time of private autonomous vehicles.

It is necessary to develop a strategy and policies to make public transport, in a future ruled by autonomous vehicles, more attractive by optimising travel times or reducing access and egress time by improving the accessibility of the micro-transit service.

5 References

- Alsharif, M. H., Kelechi, A. H., Yahya, K., & Chaudhry, S. A. (2020). Machine Learning Algorithms for Smart Data Analysis in Internet of Things Environment: Taxonomies and Research Trends. *Symmetry*, 12(1), 88. <https://doi.org/10.3390/sym12010088>
- Ben-Akiva, M., & Bierlaire, M. (1999). Discrete Choice Methods and their Applications to Short Term Travel Decisions BT - *Handbook of Transportation Science* (R. W. Hall (Ed.); pp. 5–33). Springer US. https://doi.org/10.1007/978-1-4615-5203-1_2

- Ben-Akiva, M. E., & Lerman, S. R. (1985). *Discrete choice analysis : theory and application to travel demand*. MIT Press.
<https://search.ebscohost.com/login.aspx?direct=true&AuthType=sso&db=cat00006a&AN=melb.b1420353&site=eds-live&scope=site&custid=s2775460>
- Bierlaire, M. (2020). *A short introduction to PandasBiogeme*.
- Bowman, L. A., & Turnquist, M. A. (1981). Service frequency, schedule reliability and passenger wait times at transit stops. *Transportation Research Part A: General*, 15(6), 465–471. [https://doi.org/10.1016/0191-2607\(81\)90114-X](https://doi.org/10.1016/0191-2607(81)90114-X)
- Bureau of Meteorology. (2020). *Climate Data Online*. Commonwealth of Australia.
<http://www.bom.gov.au/climate/data/>
- Childress, S., Nichols, B., Charlton, B., & Coe, S. (2015). Using an Activity-Based Model to Explore the Potential Impacts of Automated Vehicles. *Transportation Research Record*, 2493(1), 99–106.
<https://doi.org/10.3141/2493-11>
- Danaf, M., Abou-Zeid, M., & Kaysi, I. (2014). Modeling travel choices of students at a private, urban university: Insights and policy implications. *Case Studies on Transport Policy*, 2(3), 142–152.
<https://doi.org/10.1016/j.cstp.2014.08.006>
- Datta, S. K., Costa, R. P. F. Da, Härrä, J., & Bonnet, C. (2016). Integrating connected vehicles in Internet of Things ecosystems: Challenges and solutions. 2016 IEEE 17th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM), 1–6. <https://doi.org/10.1109/WoWMoM.2016.7523574>
- Dziekan, K., & Kottenhoff, K. (2007). Dynamic at-stop real-time information displays for public transport: effects on customers. *Transportation Research Part A: Policy and Practice*, 41(6), 489–501.
<https://doi.org/10.1016/j.tra.2006.11.006>
- Fagnant, D. J., & Kockelman, K. (2015). Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice*, 77, 167–181.
<https://doi.org/10.1016/j.tra.2015.04.003>
- Fan, W. (David), & Machemehl, R. B. (2009). Do transit users just wait for buses or wait with strategies?: Some numerical results that transit planners should see. *Transportation Research Record: Journal of the Transportation Research Board*, 2111(1), 169–176. <https://doi.org/10.3141/2111-19>
- Gong, H., Chen, X., Yu, L., & Wu, L. (2016). An application-oriented model of passenger waiting time based on bus departure time intervals. *Transportation Planning and Technology*, 39(4), 424–437.
<https://doi.org/10.1080/03081060.2016.1160583>
- Heinrichs, D., & Cyganski, R. (2015). Automated Driving: How It Could Enter Our Cities and How This Might Affect Our Mobility Decisions. *DisP - The Planning Review*, 51(2), 74–79.
<https://doi.org/10.1080/02513625.2015.1064650>
- Ingvardson, J. B., Nielsen, O. A., Raveau, S., & Nielsen, B. F. (2018). Passenger arrival and waiting time distributions dependent on train service frequency and station characteristics: A smart card data analysis. *Transportation Research Part C: Emerging Technologies*, 90, 292–306.
<https://doi.org/10.1016/j.trc.2018.03.006>
- Ingvardson, J. B., Nielsen, O. A., Raveau, S., & Nielsen, B. F. (2020). Corrigendum to “Passenger arrival and waiting time distributions dependent on train service frequency and station characteristics: A smart card data analysis” [transport. Res. Part C (2018), 292–306]. *Transportation Research Part C: Emerging Technologies*, 120, 102798. <https://doi.org/10.1016/j.trc.2020.102798>

- Koppelman, F., & Bhat, C. (2006). A Self Instructing Course in Mode Choice Modelling: Multinomial and Nested Logit Models.
- Krueger, R., Rashidi, T. H., & Rose, J. M. (2016). Preferences for shared autonomous vehicles. *Transportation Research Part C: Emerging Technologies*, 69, 343–355. <https://doi.org/https://doi.org/10.1016/j.trc.2016.06.015>
- Litman, T. (2020). Autonomous vehicle implementation predictions: Implications for transport planning.
- Mohammadzadeh, M. (2020). Exploring tertiary students' travel mode choices in Auckland: Insights and policy implications. *Journal of Transport Geography*, 87, 102788. <https://doi.org/https://doi.org/10.1016/j.jtrangeo.2020.102788>
- Morrow, W. R., Greenblatt, J. B., Sturges, A., Saxena, S., Gopal, A., Millstein, D., Shah, N., & Gilmore, E. A. (2014). Key factors influencing autonomous vehicles' energy and environmental outcome. In *Road vehicle automation* (pp. 127–135). Springer.
- O'Flaherty, C. A., & Mancan, D. O. (1970). Bus passenger waiting times in central areas. *Traffic Engineering & Control*, 11(9), 419–421.
- Parton, W. J., & Logan, J. A. (1981). A model for diurnal variation in soil and air temperature. *Agricultural Meteorology*, 23, 205–216.
- Petrović, D., Ivanović, I., Đorić, V., & Jović, J. (2020). Impact of Weather Conditions on Travel Demand – The Most Common Research Methods and Applied Models. *Promet - Traffic&Transportation*, 32(5 SE-Articles). <https://doi.org/10.7307/ptt.v32i5.3499>
- Public Transport Victoria. (2020a). Victorian Fares and Ticketing Conditions.
- Public Transport Victoria. (2020b). Victorian Regional Bus Fares Supplement.
- RACV. (2019). RACV car running cost calculator. <https://www.racv.com.au/on-the-road/buying-a-car/car-running-costs.html#rates>
- Reza Amin-Naseri, M., & Baradaran, V. (2015). Accurate Estimation of Average Waiting Time in Public Transportation Systems. *Transportation Science*, 49(2), 213–222. <https://search.ebscohost.com/login.aspx?direct=true&AuthType=sso&db=bah&AN=111060809&site=ehost-live&custid=s2775460>
- SAE. (2018). Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. https://www.sae.org/standards/content/j3016_201806/
- Saneinejad, S., Roorda, M. J., & Kennedy, C. (2012). Modelling the impact of weather conditions on active transportation travel behaviour. *Transportation Research Part D*, 17(2), 129–137. <http://10.0.3.248/j.trd.2011.09.005>
- Schwieterman, J. P. (2019). Uber Economics: Evaluating the Monetary and Travel Time Trade-Offs of Transportation Network Companies and Transit Service in Chicago, Illinois. *Transportation Research Record*, 2673(4), 295–304. <https://doi.org/10.1177/0361198119839344>
- Seddon, P. A., & Day, M. P. (1974). Bus passenger waiting times in Greater Manchester. *Traffic Engineering and Control*, 15(9), 442–445. https://scholar.google.cl/scholar?q=o%27flaherty+mangan&hl=en&as_sdt=0,5&scilu=3,89040319167867779:1&scisig=AMstHGQAAAAAVa9F1wq5OdPfq58AWmzO9kSDYAkqf8c#0
- Shah, S. A. A., Ahmed, E., Imran, M., & Zeadally, S. (2018). 5G for Vehicular Communications. *IEEE*

- Shen, H., Zou, B., Lin, J., & Liu, P. (2020). Modeling travel mode choice of young people with differentiated E-hailing ride services in Nanjing China. *Transportation Research Part D: Transport and Environment*, 78, 102216. <https://doi.org/https://doi.org/10.1016/j.trd.2019.102216>
- SkedGo. (2020). TripGo API. <https://skedgo.com/tripgo-api/>
- Tao, S., Corcoran, J., Rowe, F., & Hickman, M. (2018). To travel or not to travel: ‘Weather’ is the question. Modelling the effect of local weather conditions on bus ridership. *Transportation Research Part C: Emerging Technologies*, 86, 147–167. <https://doi.org/https://doi.org/10.1016/j.trc.2017.11.005>
- Ton, D., Duives, D. C., Cats, O., Hoogendoorn-Lanser, S., & Hoogendoorn, S. P. (2019). Cycling or walking? Determinants of mode choice in the Netherlands. *Transportation Research Part A: Policy and Practice*, 123, 7–23. <https://doi.org/https://doi.org/10.1016/j.tra.2018.08.023>
- Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.
- Truong, L. T., De Gruyter, C., Currie, G., & Delbosc, A. (2017). Estimating the trip generation impacts of autonomous vehicles on car travel in Victoria, Australia. *Transportation*, 44(6), 1279–1292. <https://doi.org/10.1007/s11116-017-9802-2>
- Victorian Integrated Survey of Travel and Activity. (2016). <https://transport.vic.gov.au/about/data-and-research/vista/vista-data-and-publications>
- Webb, A., Kumar, P., & Khani, A. (2020). Estimation of passenger waiting time using automatically collected transit data. *Public Transport*, 12(2), 299–311. <https://doi.org/10.1007/s12469-020-00229-x>
- You, J., Hubbard, K. G., & Goddard, S. (2008). Comparison of methods for spatially estimating station temperatures in a quality control system. *International Journal of Climatology*, 28(6), 777–787. <https://doi.org/https://doi.org/10.1002/joc.1571>
- Young, M., Farber, S., & Palm, M. (2020). The true cost of sharing: A detour penalty analysis between UberPool and UberX trips in Toronto. *Transportation Research Part D: Transport and Environment*, 87, 102540. <https://doi.org/https://doi.org/10.1016/j.trd.2020.102540>