Asking the Wizard-of-Oz: how experiencing AVs affects preferences towards their use for feeder trips in public transport.

Francisco J. Bahamonde-Birke*
Sociale Geografie en Planologie
Universiteit Utrecht
e-mail: bahamondebirke@gmail.com; f.j.bahamondebirke@uu.nl

Jan Weschke
Institut für Verkehrsforschung
Deutsches Zentrum für Luft- und Raumfahrt
e-mail: jan.weschke@dlr.de

Kay Gade
Institut für Verkehrsforschung
Deutsches Zentrum für Luft- und Raumfahrt
e-mail: kay.gade@dlr.de

Evanthia Kazagli
Haute Ecole d'Ingénierie et de Gestion du Canton de Vaud (HEIG-VD)
Haute école spécialisée de Suisse occidentale (HES-SO)
e-mail: evanthia.kazagli@heig-vd.ch

1. INTRODUCTION

The introduction of autonomous vehicles (AVs) is expected to have a disruptive impact on the transportation systems (Fagnant and Kockelmann, 2015; Milakis et al., 2018; Fraedrich et al, 2019, among many others). One of the most significant changes that AVs are likely to induce is enabling/facilitating the provision of highly flexible on-demand services, as the absence of a driver will result in a substantial reduction of the operational costs, increasing the desirability of such services (Correia and v. Arem, 2016; Bösch et al., 2018; Bahamonde-Birke et al., 2018). The advance of on-demand services, in turn, is expected to significantly affect the way mobility requirements are perceived by end-users and has the potential to revolutionize transport and urban systems (Fagnant and Kockelmann, 2015; Bahamonde-Birke et al., 2018, etc.).

However, most of the aforementioned issues refer to a given point of time, when the introduction of AVs and the development of on-demand services has achieved a stationary status. The extensive adoption of both AVs and on-demand transportation is likely to take a long time. It is precisely this timeframe, when transportation and city planners will have the best chance to shape the transport systems of the future and to introduce regulation in order to steer transportation systems towards sustainability and social welfare optima (Smith, 2012; Bahamonde Birke et al., 2018).

One of the possibilities associated with on-demand services and AVs that has been widely discussed is the integration of them into the provision of public transport (Davidson and Spinuola, 2015; Yap et al., 2016; Kolarova et al., 2019, etc.). This would ease the implementation of regulation, while, at the same time, it would avoid that on-demand services cannibalize mass-transportation modes (which are the most efficient modes in terms of social costs) leading to the well-known Vicious Circle of the Decline in Public Transportation and Increased Urban Congestion (Ortúzar and Willumsen, 2011).

The present study aims precisely at evaluating the possibilities that AVs and on-demand services will offer to improve the provision of public transport. Basically, we considered the current preferences of individuals towards both autonomous buses and on-demand services in the context of modal choice preferences. For this purpose, a stated-preferences (SP) experiment was developed, in which respondents were asked to state which transportation mode they would prefer (including on-demand autonomous buses) to carry out a recent trip. Then, the participants of the study had the possibility to take part in a Wizard-of-Oz experiment (Kelley, 1984), in which they were able to use autonomous buses during a period of one month. After the completion of the Wizard-of-Oz experiment, the participants were faced again with a similar SP-experiment in order to consider how experiencing autonomous bus services have affected their preferences.

The present paper reports the results of the aforementioned study. Section 2 presents the details of the experimental setting, including both the SP-experiments as well as the Wizard-of-Oz experiment. Section 3 introduces the methodological framework used to evaluate the results, while Section 4 presents and discusses the main findings of the study. Finally, Section 5 reports the conclusions of the study.

2. EXPERIMENTAL SETTING

The experimental setting of the study consists of three-steps. First, an SP experiment was developed and conducted in order to capture mode choice preferences *ex-ante*. On a second step, a Wizard-of-Oz experiment was developed and conducted. Finally, the SP experiment was repeated to capture preferences *ex-post*. Both experiments were carried out in two cities in Germany, namely Berlin (big-sized city; ca. 3.7 MM inhabitants) and Braunschweig (medium-sized city; ca. 0.28 MM inhabitants). The dataset consists of 65 individuals, 14 from the city of Braunschweig while the others are from Berlin. All individuals answered the experiment *ex-ante*, while only 33 of them answered the experiment *ex-post*, from which twelve are from Braunschweig and 21 from Berlin. In each wave, each individual provided answers to twelve choice situations. In total, the experiment yielded 1176 observations (780 *ex-ante* and 396 *ex-post*). Given the short duration of the experiment, it was considered that the outside conditions remained stable and no control group was analyzed.

The SP experiment represents a modal choice situation. The setting of the modal choice is supposed to represent the last feeder trip from home to a mass transportation mode station. Feeder trips were selected as they most closely resemble the setting, in which the Wizard-of-Oz experiment with autonomous buses was conducted. For the purposes of the feeder trip, individuals were offered five transportation alternatives, namely:

a) Conventional public transport (bus), including the attributes walking time to the bus stop, waiting time at the bus stop, and travel time in the vehicle.
b) Walking, including the attribute walking time to the mass transportation mode station.
c) Biking, including the attribute travel time as well as walking time representing the additional walking time required to park/secure the bike at the mass transportation mode station.
d) Autonomous on-demand bus, including the attributes walking time to the pick-up point (which may be zero if the person is picked-up at home), waiting time (either at home or at the pick-up point), travel time in vehicle, as well as information regarding whether further passengers were on-board or has to be picked up after and whether it was possible to reserve a given pick-up time in advance (in this case the users still faced a waiting time, which only applies if they do not reserve in advance).

Regarding the cost attribute, it was considered that the price of conventional public transport (bus) was included in the price of mass transportation mode. Consequentially, neither conventional public transport, nor walking or biking resulted in an increase of the costs faced by the user. The autonomous on-demand bus, in turn, considered a premium to be paid on top of the price of the mass transportation mode ticket. It ranged between 0€ and 1€. Finally, every choice situation included five alternatives, as two alternatives offering autonomous on-demand bus services were included. The reason behind it is that autonomous on-demand bus services included more attributes than conventional services and, consequentially, more options were required to depict this higher degree of internal variability.

The attribute selection is based on qualitative analysis, which was specially conducted in the context of the experiments (Stark et al., 2019). The SP design was considered with help of personal interviews (which included the collection of choices to estimate prior-model) and the final design was defined by maximizing the D-efficiency relying on the aforementioned priors (Rose and Bliemer, 2009).

The Wizard-of-Oz experiment offered the participants the possibility of experiencing autonomous bus services during a period of one month. The services were not actually autonomous and the buses indeed had a human driver. The users, however, were completely unaware that the buses were not actually being driven autonomously. The buses were completely designed and customized the portrait the experience of autonomous driving. Similarly, the drivers received special training to guarantee that the vehicles be driven in the fashion AVs are expected to perform.

The experiment in Braunschweig was conducted in August 2019 and it considered the deployment of one bus. The experiment in Berlin was conducted in November 2019 and considered the deployment of three buses. Given those limitations, the experiments were confined to a small geographical area.

3. METHODOLOGICAL APPROACH

Under the assumption that individuals *q* are rational decision-makers belonging to a population Q $(q \in Q)$, it can be postulated that individuals facing a set $A(q)$ of available alternatives for each individual *q*, which is a subset $A(q) \in A$ of the full set of alternatives *A*, will choose the alternative *i* that maximizes their perceived utility. In accordance with Random Utility Theory (Thurstone, 1927; McFadden, 1974), it is possible to depict the utility of the different alternatives in the choice set as the sum of a representative component and an error term, which, under the assumption of additive linearity, leads to the following expression (Ortúzar and Willumsen, 2011):

$$U_q = X_q \cdot \beta + \varepsilon_q \qquad [2.1],$$

where U_q is a (J_q X 1) vector of utilities associated with the J_q alternatives in the choice-set of each individual *q*; X_q is a (J_q X *M*) matrix standing for *M* observed attributes of the J_q alternatives and characteristics of the individuals (whose rows represent a given alternative and its columns stand for the different attributes constituting this option[1]). β is a (*M* X 1) vector of parameters to be estimated (whose rows are associated with the different elements of X_q). Finally, ε_q is a (J_q X 1) vector representing the error; if it is assumed for the error terms to be independent EV1 distributed with same mean (for all alternatives) and diagonal homoscedastic covariance (J_q X J_q) matrix (Σ_ε), the choice probabilities will be given by a Multinomial Logit model (Domencich and McFadden, 1975; MNL).

Nevertheless, the assumption of independence may not hold, when the observations arise from panel or pseudo-panel data, as the observations associated with the same individual are likely to be correlated. For the purposes of this short paper, due to space constraints, this limitation will be ignored, but it will be lifted in further research. Consequentially, all observations have been treated independently. However, observations arising post exposition (*ex-post*) have been accordingly identified. Similarly, observations arising from the cities of Braunschweig (BS) and Berlin (B) have been properly identified.

4. RESULTS AND DISCUSSION

Table 1 reports the model estimates. In accordance with eq. [1] additive linearity and i.i.d. EV1 error terms were assumed. The first column of Table 1 indicates the variable associated with the estimators (either variables presented in the SP-experiment or associated with the individual/wave or interactions), the second is indicative for the utility function affected by the variable, while the third, fourth and fifth stand for the expected value of the estimator, its standard deviation and the t-test against zero, respectively. The model reports parameters to be found statistically significant at a significance level of 5%,

---

[1] If a given attribute is considered to be generic for all alternatives, a given column would exhibit values for all rows; otherwise (alternative-specific parameters) the column would contain only one element different from zero. The latter always applies when dealing with characteristics of the individuals (or variables not exhibiting variability across individuals; Ortúzar and Willumsen, 2011).

but it also includes parameters found to be insignificant but that are deemed relevant for the analysis. Similarly, differences between parameters *ex-ante* and *ex-post* are only included when found to be statistically significantly different from zero. Different parameters for the cities of Braunschweig (BS) and Berlin (B) are also reported when the differences are statistically significant, or when deemed relevant for the analysis. No further parameter was found to be statistically significantly different from zero. No alternative specific parameters were considered due to the size of the sample, which may have led to overfitting.

Table 1 - Model Estimates

Variable	***Equation***	***Model***		
ASC Bike	*Utility Alternative Bike*	**0**	*(fixed)*	*fixed*
ASC Conventional Public Transport (CPT) BS	*Utility Alternative CPT*	**-0.292**	*(0.632)*	*-0.461*
ASC Conventional Public Transport (CPT) B	*Utility Alternative CPT*	**0.135**	*(0.295)*	*0.457*
ASC Walking BS	*Utility Alternative Walking*	**-0.508**	*(0.86)*	*-0.591*
ASC Walking B	*Utility Alternative Walking*	**-0.438**	*(0.491)*	*-0.892*
ASC Autonomous Vehicle BS	*Utility Alternative AV1, AV2*	**0.764**	*(0.683)*	*1.12*
ASC Autonomous Vehicle B	*Utility Alternative AV1, AV2*	**1.35**	*(0.342)*	*3.96*
Premium Price BS	*Utility Alternatives AV1, AV2*	**-2.05**	*(0.283)*	*-7.24*
Premium Price B	*Utility Alternatives AV1, AV2*	**-1.68**	*(0.145)*	*-11.6*
Travel Time BS	*Utility Alternatives Bike, CPT, AV1, AV2*	**-0.175**	*(0.0279)*	*-6.27*
Travel Time B	*Utility Alternatives Bike, CPT, AV1, AV2*	**-0.131**	*(0.017)*	*-7.71*
Walking Time BS	*Utility All Alternatives*	**-0.226**	*(0.0398)*	*-5.68*
Walking Time B	*Utility All Alternatives*	**-0.158**	*(0.0197)*	*-8*
Waiting Time BS	*Utility Alternative CPT, V1, AV2*	**-0.23**	*(0.0836)*	*-2.75*
Waiting Time B	*Utility Alternative CPT, V1, AV2*	**-0.126**	*(0.0341)*	*-3.7*
Waiting Time when reserved BS	*Utility Alternatives AV1, AV2*	**0.0344**	*(0.0401)*	*0.858*
Waiting Time when reserved B	*Utility Alternatives AV1, AV2*	**-0.0128**	*(0.0223)*	*-0.013*
Reservation BS	*Utility Alternatives AV1, AV2*	**-0.517**	*(0.795)*	*-0.651*
Reservation B	*Utility Alternatives AV1, AV2*	**-0.239**	*(0.4)*	*-0.598*
Further Passengers BS	*Utility Alternatives AV1, AV2*	**-0.501**	*(0.381)*	*-1.31*
Further Passengers B	*Utility Alternatives AV1, AV2*	**0.0836**	*(0.197)*	*0.423*
Last Passenger BS	*Utility Alternatives AV1, AV2*	**-0.179**	*(0.357)*	*-0.502*
Last Passenger B	*Utility Alternatives AV1, AV2*	**-0.0705**	*(0.189)*	*-0.373*
Change in ASC Autonomous Vehicle ex-post BS	*Utility Alternatives AV1, AV2*	**1.38**	*(0.676)*	*2.04*
Change in ASC Autonomous Vehicle ex-post B	*Utility Alternatives AV1, AV2*	**1.32**	*(0.566)*	*2.32*
Change in Reservation ex-post BS	*Utility Alternatives AV1, AV2*	**-1.18**	*(0.515)*	*-2.28*
Change in Reservation ex-post B	*Utility Alternatives AV1, AV2*	**0.22**	*(0.34)*	*0.646*
Change in Further Passengers ex-post B	*Utility Alternatives AV1, AV2*	**-1.11**	*(0.441)*	*-2.51*
Change in Last Passenger ex-post BS	*Utility Alternatives AV1, AV2*	**-1.08**	*(0.602)*	*-1.8*
Change in Last Passenger ex-post B	*Utility Alternatives AV1, AV2*	**-1.08**	*(0.506)*	*-2.13*
Log-likelihood			**-1235.049**	

The parameters associated with travel costs as well as with travel, waiting and walking time exhibit the expected signs. Similarly, in line with the hypothesis, waiting and walking time in Braunschweig are perceived more negatively than travel time, while in Berlin walking time is perceived more negatively than the waiting and travel time, which are perceived quite similar in Berlin.

The magnitudes both of the cost and time parameters in Berlin are smaller than in Braunschweig. In comparison between the marginal utility of travel time and the marginal utility of price this yields to a value of time of 5.12 €/hr in Braunschweig and 4.68 €/hr in

Berlin, which is quite comparable and in line with previous studies in Germany. No significant change in these parameters was observed after exposure.

When reserving a given pick-up time was possible, the waiting time was found to be not statistically significant in both cities. In this case, the waiting time reflects how long individuals would have to wait if they did not previously reserve a pick-up time; hence, the insignificance of the parameter means that the individuals were willing to reserve a vehicle and forgot the waiting time. The necessity of having to reserve a vehicle was also found to be statistically insignificant prior to the experiment. After exposure, no changes in the valuation of waiting time when reserving a given pick-up time were identified, but a significant disutility was identified in association with the necessity itself of reserving a vehicle (in order to avoid waiting times) in Braunschweig, while it remained an insignificant factor in Berlin. The disutility of reservation amounts to 7.3 min of waiting time, when reservation is not possible.

Furthermore, it was found that being the last passenger to board the vehicle or traveling with further passengers on board had no significant impact. However, the impact of being the last passenger after exposition is statistically negatively significant in Berlin and in Braunschweig. This finding contradicts initial expectations, as being the last passenger normally diminishes the stress associated with unreliability; however, as in the experiment the travel time was considered to be fixed, it is possible that other considerations may have played a larger role (especially as the effect was only identified post exposure), such as having troubles finding a desired place to sit. In Berlin a significant disutility of traveling with other people was found after the exposure, while there was no significant change observable in Braunschweig.

Finally, the ASC of the alternative AV was found to significantly increase after exposure. This means that the individuals were more likely to use the alternative AV, after having taking part in the experiment. This is true for both cities while there is no statistically significant difference between them.

5. CONCLUSIONS

The study reports the results regarding preferences towards the use of on-demand autonomous buses for the purposes of feeder trips in the cities of Braunschweig and Berlin. The results show that experiencing autonomous bus services positively impacts the propensity of using such services in the future. Furthermore, the results indicate that the experience increased the negative disutility associated with having to reserve a pick-up time a priori, which may have been neglected *ex-ante*, given the hypothetic nature of SP-experiments (but it was highlighted by experiencing the alternative). Also being the last passenger to board the vehicles becomes negatively significant in the propensity of using the new services in both cities. A similar effect is observed for traveling with further passengers for the city of Berlin. Both the disutility of travel time (including waiting and walking time) and of the price remain constant after exposure. Further research will be aimed at providing a better depiction of the error terms and of the characteristics of the individuals.

REFERENCES

Bösch, P. M., Becker, F., Becker, H., & Axhausen, K. W. (2018). Cost-based analysis of autonomous mobility services. *Transport Policy*, *64*, 76-91.

Correia, G.H.d.A., & van Arem, B. (2016). Solving the User Optimum Privately Owned Automated Vehicles Assignment Problem (UO-POAVAP): A model to explore the impacts of self-driving vehicles on urban mobility. *Transportation Research Part B: Methodological* **87**, 64-88.

Davidson, P., & Spinoulas, A. (2015, July). Autonomous vehicles: what could this mean for the future of transport. In *Australian Institute of Traffic Planning and Management (AITPM) National Conference, Brisbane, Queensland*.

Domencich, T., & McFadden, D. (1975). *Urban Travel Demand – A Behavioural Analysis*. North Holland, Amsterdam.

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.

Fraedrich, E., Heinrichs, D., Bahamonde-Birke, F. J., & Cyganski, R. (2019). Autonomous driving, the built environment and policy implications. *Transportation research part A: policy and practice*, *122*, 162-172.

Kelley, J. F. (1984). An iterative design methodology for user-friendly natural language office information applications. *ACM Transactions on Information Systems (TOIS)*, *2*(1), 26-41.

Kolarova, V., Steck, F., & Bahamonde-Birke, F. J. (2019). Assessing the effect of autonomous driving on value of travel time savings: A comparison between current and future preferences. *Transportation Research Part A: Policy and Practice*, *129*, 155-169.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behaviour. In Zarembka, P. (ed.), *Frontiers in Econometrics*, 105-142. Academic Press, New York.

Milakis, D., Kroesen, M., & van Wee, B. (2018). Implications of automated vehicles for accessibility and location choices: Evidence from an expert-based experiment. *Journal of Transport Geography*, *68*, 142-148.

Ortúzar, J. de D., & Willumsen, L.G (2011). *Modelling Transport*. Fourth Edition, John Wiley and Sons, Chichester.

Rose, J. M., & Bliemer, M. C. (2009). Constructing efficient stated choice experimental designs. *Transport Reviews*, *29*(5), 587-617.

Smith, B. W. (2012). Managing autonomous transportation demand. *Santa Clara L. Rev.*, *52*, 1401.

Stark, K., Gade, K., & Heinrichs, D. (2019). What Does the Future of Automated Driving Mean for Public Transportation?. *Transportation Research Record*, 2673(2), 85-93.

Thurstone, L.L. (1927). A law of comparative judgment. *Psychological Review* **34**, 273 286.

Yap, M. D., Correia, G., & Van Arem, B. (2016). Preferences of travellers for using automated vehicles as last mile public transport of multimodal train trips. *Transportation research part a: policy and practice*, *94*, 1-16.