# How accessibility shapes the landscape of mobility tool ownership and use



**Author(s):**
Loder, Allister

**Publication date:**
2016-07

**Permanent link:**
https://doi.org/10.3929/ethz-b-000250468



**Originally published in:**
Arbeitsberichte Verkehrs- und Raumplanung 1135

# Patterns of Mobility Tool Ownership

**Allister Loder**

**Kay W. Axhausen**

**Working paper** **1135**

**Institute for Transport Planning and Systems** **July 2016**

Institut für Verkehrsplanung und Transportsysteme
Institute for Transport Planning and Systems

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Working paper 1135

# Patterns of Mobility Tool Ownership

Allister Loder
IVT
ETH Zürich

CH-8093 Zürich
phone: +41-44-633 62 58
fax: +41-44-633 10 57
allister.loder@ivt.baug.ethz.ch

Kay W. Axhausen
IVT
ETH Zürich

CH-8093 Zürich
phone: +41-44-633 39 43
fax: +41-44-633 10 57
axhausen@ivt.baug.ethz.ch

July 2016

## Abstract

We propose the multivariate probit model with sample selection for travel demand modeling to simultaneously capture correlated and mutually exclusive choices. Our modeling approach is suitable for increasingly diverse, but complementary, means of transport and new kinds of travel demand. In contrast to multinomial logit models, each single choice is modeled, instead of each single element in the choice set.

The model is empirically applied to data from the 2010 Swiss transportation micro-census. In Switzerland, residents can augment their choice of car ownership by purchasing a public transportation 'season ticket', with either nation-wide or local coverage (mutually exclusive). We capture the influence of land use by accessibility at the municipality level, finding that, in less accessible regions, car ownership is unavoidable, whereas with better accessibility, car use can be substituted through season tickets. This model allows planners and policy makers to quantify the long-term effects of changes in accessibility on the choices of cars and season tickets.

## Keywords
Travel behavior, multivariate probit, mobility tool, car ownership, season-ticket ownership

## Preferred citation style

Loder, A. and K. W. Axhausen (2016) Patterns of Mobility Tool Ownership, *Working paper*, **1135**, Institute for Transport Planning and Systems (IVT), ETH Zurich, Zurich.

# 1 Introduction

In Switzerland, the choice of mobility tool ownership bundles a choice of car ownership and a public transport season ticket subscription. With the latter option, residents can choose between the nation-wide season ticket Generalabonnement (GA) and local season tickets with limited spatial coverage. In this choice set, some choice combinations are complements (car and local season ticket), others might be more substitutive (car and GA), whereas others are mutually exclusive (GA and local season ticket). Traditional discrete choice models, like the multinomial logit (MNL) or nested logit (NL), have limitations in capturing all these aspects Hensher et al. (2015).

With alternatives in the MNL corresponding to all possible outcome combinations, one can argue it violates the assumption of independent irrelevant alternatives to have the unobserved prevalence for car ownership in more than one outcome. Deploying a NL instead allows nesting of alternatives, e.g. the prevalence for car ownership, but requires a choice hierarchy that does not reflect true choice environment Ben Akiva and Lerman (1985). Comprehensive discrete choice analyses of mobility tool ownership are scarce in the literature, with a significant tremendous lack on the season ticket side. Notably, cars and season tickets ownership interactions show substitution patterns Simma and Axhausen (2001); Scott and Axhausen (2006); Kowald et al. (2016). A comprehensive overview of car ownership models is available by de Jong et al. (2004). However, no ownership model incorporates different and mutually exclusive season ticket types while simultaneously considering car ownership.

In this paper, we present a novel discrete choice model for the travel demand modeling field that allows for complementary choices and simultaneous, mutually exclusive choices. The model builds on the multivariate probit model Jenkins et al. (2006) and incorporates a Heckman-like sample selection Heckman (1976, 1979). Our proposed model adds new insights to the travel demand literature and can easily be extended to new transport modes like car sharing Becker et al. (2016); Schmid et al. (2016)

In addition to socioeconomic attributes, land-use is a strong determinant of mobility tool ownership Ewing and Cervero (2010). Because the choice of mobility tool ownership is a long-term decision, we are interested in long-term factors. Whereas travel times are responsible for short-term mode choice, the concept of accessibility is relevant for long-term decisions Metz (2008). Accessibility links together the number of accessible opportunities and the costs of reaching them Hansen (1959). Accessibility is also a measure of generalized travel cost Weis and Axhausen (2009). A high degree of accessibility at a location corresponds to a low level of generalized travel cost. We use the concept of accessibility as a measure of land-use, as well as generalized cost of travel for mobility tool ownership choice.

The remainder of this paper is organized as follows; the next section gives the novice reader a brief overview of mobility tool ownership and accessibility literature. The next section describes available data; we then propose our statistical model, present our results and discuss our findings.

## 2 Literature

Understanding patterns of mobility tool ownership, especially of cars, is relevant for many areas of research, including travel demand forecasting, estimating environmental impacts and tax income Train (1986); de Jong et al. (2004). While modeling car ownership has attracted much research interest with a variety of methods, modeling its interaction with season ticket ownership has not.

At disaggregated level, mobility tool ownership is modeled by discrete choice methods Ben Akiva and Lerman (1985). MNL and ordered logit models are used to model the number of cars at a household level, e.g. Potoglou and Kanaroglou (2008). Households' choices of number of cars and annual mileage can be jointly addressed by multiple disctrete-continous extreme value models, e.g. Tanner and Bolduc (2014), in which both choices are transformed into a microeconomic model where utility is maximized, given a budget constraint.

Season ticket ownership modeling is rare in literature, especially with discrete choice methods. On an aggregated scale for Madrid García-Ferrer et al. (2006), Spain and Freiburg FitzRoy and Smith (1998), Germany, a significant positive effect of season ticket introduction on public transport usage is found. Arguably, season ticket ownership is rarely modeled directly, as this choice can also modeled by mode choice models, e.g. Hensher and Rose (2007).

At household level, joint modeling of season ticket and car ownership can be carried out with bivariate ordered probit models Scott and Axhausen (2006). On an individual scale, structural equation modeling, also encorporating mileage, can be used Simma and Axhausen (2001). An MNL can be used for revealed preference Kowald et al. (2016) and stated preference data. Neglecting interactions, both mobility tools can independently model using univariate logit or probit models. Various findings suggest the substitutive nature of cars and season tickets.

In addition to income as a strong determinate of car ownership Goodwin et al. (2004), land-use at household residential location is a key determinant of mobility tool ownership Ewing and Cervero (2010). Land-use factors can generally be classified by density, diversity, design, destination accessibility and distance to transit Cervero and Kockelman (1997). Accessibility has the greatest reported effect size Ewing and Cervero (2010) and is, at the same time, a measure of the generalized travel cost Weis and Axhausen (2009).

Accessibility is an abstract spatial concept that links opportunities in other zones to the generalized travel cost of reaching these opportunities, in terms of travel time Hansen (1959). Accessibility $A$ at location $i$ is defined by $A_i = \sum_{\forall j \in N, j \neq i} O_j exp(\beta c_{ij})$ , with $O_j$ representing being the opportunities at locations $j$ and $c_{ij}$ the generalized travel cost from $i$ to $j$. $\beta$ is the distance decay weighting parameter. Depending on the situation, various measures for opportunities can be used, e.g. number of employed Hansen (1959), population Killer et al. (2013), and housing and retail Crozet et al. (2012). With longer travel usually less favored, the distance decay parameter $\beta$ weights travel time and makes more distant opportunities less attractive.

Accessibility is a generalization of the population-over-distance relationship Hansen (1959). When computed as the log-sum term of a destination-mode choice model, it can capture - besides travel time - monetary cost,comfort and reliability Ben Akiva and Lerman (1985). Accessibility can also be seen as an interface between (urban) economy and (transport) geography Crozet et al. (2012). As travel speed and travel time are crucial elements in initializing trade Krugman (1993), the benefits of increasing accessibility levels can be quantified as a positive relationship with productivity, thus generating positive externalities Venables (2007).

# 3 Data

## 3.1 Socio-economic

Data on mobility tool ownership and accompanying socio-demographic information is provided by the Swiss transportation micro-census Swiss Federal Statistical Office (BFS) (2012). For modeling mobility tool ownership, we use information on car availability and code “car sometimes available” as “car never available”. We generate a variable ticket if the survey person has any season ticket subscription. For everyone with a season ticket subscription, we generate a variable ticket type that equals 1 if the surveyed person has a GA (German abbreviation for a full, nation-wide season ticket) and zero otherwise, i.e. if one owns a local season ticket.

For explanatory variables, we select age, employment status (1 if employed), university degree (1 if university (of applied sciences) degree), daily traveled distance by public transport and having a secondary residence. At the household level, we recode the stated gross monthly household income classes into a continuous scale by assigning the midpoint value of each class to the household. As 24% of all households did not report on their income, we impute the income with an ordered logit model. For each household that did not report their income, we assign the sum of the product of probability multiplied by midpoint income class value (results available on request.).

The Swiss Federal Office of Spatial Development (ARE) established a 5-level scale to classify households' local access to public transport ranging from E to A (best) Swiss Federal Office of Spatial Development (ARE) (2011). The scale uses a household's distance to transit stops and the level of service there. For each observation, this variable is available through the Swiss transportation micro-census. Finally, we include a dummy variable equal to one for all observations undertaken in Switzerland's large cities, e.g. Zurich or Geneva, based on the ARE spatial typology definition Swiss Federal Office of Spatial Development (ARE) et al. (2011). We removed all observed cases younger than 18 years and those who are only mobile with outside support. All case observations without a driver's license are coded as having a car never available.

## 3.2 Accessibility

Four Hansen measures of accessibility are available at a municipality level for Switzerland Axhausen et al. (2015): (1) to population by private transport (PRV_POP), (2) to employment by private transport (PRV_EMP), (3) to population by public transport (PRV_POP), and (4) to employment by public transport (PRV_EMP). With the zoning system based at municipality level, these accessibility measures are macroscopic. Travel times were obtained from the Swiss national transport model and employment and population numbers were taken from the Swiss Federal Office of Statistics for 2010.

The four accessibility measures show a high degree of correlation, as shown in Table 1(a). This is intuitive, because large parts of infrastructure are shared and all municipalities offer working and living opportunities. These correlations lead to multicollinearity problems and the full information provided cannot be extracted. Therefore, we carry out a principal component analysis (PCA) with all four accessibility variables Jolliffe (2002). Our rationale is not to narrowly reduce dimensions, but to remove multicollinearity; thus, we skirt the rules proposed by Jolliffe Jolliffe (2002) on how many factors should be extracted.

Table 1(b) gives the summary statistics of the PCA. The statistics in Table 1(c) and 1(d) allow us to interpret the factors' meaning. The first factor extracts more than 90% of variance; based on the correlations, we see it as general levels of accessibility. The second factor explains 7.6% of variance and describes better access by public transport. The third factor explains 0.3% and

Table 1: Principal component analysis of accessibility measures

(a) Correlations of accessibility measures

| | POP_PRV | POP_PUB | EMP_PRV | EMP_PUB |
|---|---|---|---|---|
| POP_PRV | 1 | 0.741 | 0.983 | 0.754 |
| POP_PUB | – | 1 | 0.703 | 0.996 |
| EMP_PRV | – | – | 1 | 0.725 |
| EMP_PUB | – | – | – | 1 |

(b) PCA summary statistics

| | Factor 1 | Factor 2 | Factor 3 | Factor 4 |
|---|---|---|---|---|
| Standard deviation | 2.175 | 0.626 | 0.126 | 0.043 |
| Proportion of Variance | 0.92 | 0.076 | 0.003 | 0.0003 |
| Cumulative Proportion | 0.92 | 0.996 | 0.999 | 1 |

(c) Factor loadings

| | Factor 1 | Factor 2 | Factor 3 | Factor 4 |
|---|---|---|---|---|
| POP_PRV | 0.523 | −0.405 | −0.720 | 0.210 |
| POP_PUB | 0.448 | 0.539 | −0.180 | −0.690 |
| EMP_PRV | 0.556 | −0.505 | 0.629 | −0.198 |
| EMP_PUB | 0.466 | 0.538 | 0.229 | 0.663 |

(d) Correlations of factors and accessibility measures

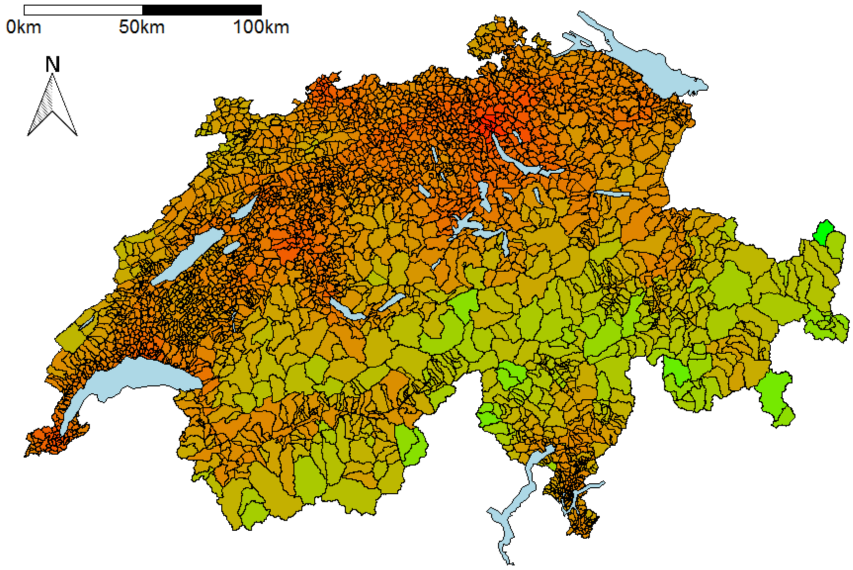
| | Factor 1 | Factor 2 | Factor 3 | Factor 4 |
|---|---|---|---|---|
| POP_PRV | 0.973 | −0.217 | -0.078 | 0.008 |
| POP_PUB | 0.944 | 0.327 | −0.022 | −0.028 |
| EMP_PRV | 0.966 | −0.252 | 0.063 | −0.007 |
| EMP_PUB | 0.948 | 0.315 | 0.027 | 0.026 |

describes better access to workplaces. The extracted variance of the fourth factor is minuscule; since it not really relevant to our analysis, we omit this variable. In Figure 1, we show spatial distribution of the first factor across Switzerland.

## 3.3 Dataset

After removing all immobile and all younger than 18 cases, we obtain a dataset with 52,476 observations. From all observations, 20.96% have neither a car nor a season ticket, 14.08%

Figure 1: Spatial illustration of first accessibility component describing general accessibility levels: the more red, the greater the general accessibility levels. Map by Swiss Federal Statistical Office (BFS) (2015)




have a season ticket, 56.93% have a car and 8.08% have both. From all season ticket owners, 36.04% have a local season ticket, 27.66% have a GA, 20.65% have a local season ticket and a car, and 15.65% have a GA and a car. Our sample distributes across the ARE spatial typology Swiss Federal Office of Spatial Development (ARE) et al. (2011) with 32.5% living in city centers, 47.73% in the agglomeration, 0.7% in isolated towns and the remaining 19.05% in the countryside. In the city centers, 31.93% have at least a season ticket and 46.67% have a car available. In the countryside, 11.23% have a season ticket and 72.75% have a car available. In Table 2, we present sample summary statistics.

As we include households' access to public transport and accessibility at municipality level, we check for correlations between both measures. We find a correlation with the first component of 0.59, with the first and third components in the city of 0.56 and 0.45, respectively, with the first component in the agglomeration of 0.42, and with the third component in isolated towns of 0.46. All in all, these correlations are high, but we do not expect a large bias in the estimates

Table 2: Principal component analysis of accessibility measures

| | Mean | Std. deviation | Minimum | Maximum |
|---|---|---|---|---|
| Car available | 0.64 | 0.48 | 0.00 | 1.00 |
| Season ticket owner | 0.23 | 0.42 | 0.00 | 1.00 |
| GA holder | 0.43 | 0.49 | 0.00 | 1.00 |
| Age | 47.43 | 17.77 | 18.00 | 99.00 |
| Age squared | 25.65 | 18.07 | 3.24 | 98.01 |
| Male | 0.49 | 0.50 | 0.00 | 1.00 |
| Working | 0.68 | 0.46 | 0.00 | 1.00 |
| University level education | 0.17 | 0.38 | 0.00 | 1.00 |
| Log of monthly household income | 8.83 | 0.55 | 7.31 | 9.90 |
| Local access to public transport: Level E | 0.26 | 0.44 | 0.00 | 1.00 |
| Local access to public transport: Level D | 0.26 | 0.44 | 0.00 | 1.00 |
| Local access to public transport: Level C | 0.20 | 0.40 | 0.00 | 1.00 |
| Local access to public transport: Level B | 0.16 | 0.37 | 0.00 | 1.00 |
| Local access to public transport: Level A | 0.13 | 0.33 | 0.00 | 1.00 |
| 1. Factor: General levels of accessibility | 1.73 | 1.90 | -11.61 | 5.81 |
| 2. Factor: Better access by public transport | 0.01 | 0.68 | -1.98 | 2.45 |
| 3. Factor: Better access to employment | 0.04 | 0.14 | -0.49 | 0.51 |
| Center of agglomeration region | 0.29 | 0.45 | 0.00 | 1.00 |
| Secondary residence | 0.08 | 0.27 | 0.00 | 1.00 |
| Log of monthly household income | 8.83 | 0.55 | 7.31 | 9.90 |
| Self-reported distance [km] | 29.67 | 62.72 | 0.00 | 949.34 |
| Observations | 52476 | | | |

# 4 STATISTICAL MODEL

As mentioned in the Swiss context, residents can choose between three different mobility tools: a car, a nation-wide season ticket and a local season ticket. When incorporating the exclusivity of both season tickets, residents are faced with a six-outcome choice set:

1. Nothing
2. Car and no season-ticket
3. Car and local season-ticket
4. Car and GA
5. No car and local season-ticket
6. No car and GA

Standard approaches to these kinds of choice sets are: MNL, e.g. Kowald et al. (2016), nested logit, e.g. Hensher and Rose (2007), or structural equation model, e.g. Simma and Axhausen (2001). However, testing for independence of alternatives in an MNL with a Hausman specification test suggests that alternatives are not independent. Therefore, we instead propose applying a multivariate probit-based model with two principal advantages; each mobility tool is modeled specifically with one equation and it captures correlations in choices, i.e. indicating whether they are substitutes or complements. To capture the exclusive nature of season tickets, we use a Heckman-like sample selection Heckman (1976, 1979), in which the first level determines whether an observation has any kind of season ticket and the second level determines the type of season ticket.

Multivariate probit models without sample selection are represented in transportation literature, e.g. correlated responses to congestion policies Choo and Mokhtarian (2008), acceptance of road pricing Rentziou et al. (2011) and car and season ticket demand Scott and Axhausen (2006). Inclusion of the Heckman-like sample selection for the bivariate case is frequent, e.g. demand of deductibles in health insurance Van de Ven and Van Praag (1981), to market entry choices of firms Henisz (2000), or firms initial public offer success Gulati and Higgins (2003). However, the multivariate case with more than two equations is seldom seen: only in patterns of consent analysis Jenkins et al. (2006), labor market outcomes Van der Straeten et al. (2003) and low income transitions Cappellari and Jenkins (2004).

## 4.1 Model specification

The proposed model is based largely on previous work by Jenkins et al. and Van der Straeten et al. Jenkins et al. (2006); Van der Straeten et al. (2003). The six outcomes can be modeled by the equations given in Table 3. Equation 3 only applies to those cases where the first equation has an observed outcome of one. The correlations in $P$ are informative: the sample selection can be ignored only if $\rho_{13} = 0$. If $\rho_{13} > 0$, unobserved variables affect selection and outcome in a similar way. Generally speaking, a negative correlation can be seen, as both outcomes are substitutes, whereas a positive correlation corresponds to complements.

## 4.2 Estimator

Based on these equations, we can define the probabilities for each outcome according to Table 4. $\phi_m$ denotes the $m$-variate cumulative normal distribution function and $P_m$ the $m$-dimensional correlation matrix of the error term. With these outcome probabilities, we define the log

Table 3: Model equations

| No. | Outcome | Equation | Observed outcome |
|---|---|---|---|
| (1) | Any season ticket | $Y_1^* = x_1\beta_1 + \varepsilon_1$ | $Y_1 = I(Y_1^* > 0)$ |
| (2) | Car available | $Y_2^* = x_2\beta_2 + \varepsilon_2$ | $Y_2 = I(Y_2^* > 0)$ |
| (3) | GA or local ticket | $Y_3^* = x_3\beta_3 + \varepsilon_3$ | $Y_3 = I(Y_3^* > 0)$ if $Y_1 = 1$, otherwise unobserved |
| (4) | Error terms | $(\varepsilon_1, \varepsilon_2, \varepsilon_3) \sim N(0, P_3)$, $P_3$ with $\rho_{ij} = \rho_{ji}, i \neq j$ and $\rho_{jj} = 1$ | |

likelihood function as

$$log\mathcal{L} = \sum_{i=1}^{N}\sum_{j=1}^{6}\delta_{ij}log(P_{ij})$$

with $P_{ij}$ being the probability for observation $i$ to choose outcome $j$ and $\delta_{ij} = 1$ if observation $i$ has chosen alternative $j$, and else equals zero otherwise. We ensure that the same correlations appear in the bivariate and trivariate correlation matrix by applying the Cholesky decomposition of the $P$ matrix. We define the three-dimensional Cholesky matrix $C$ as

$$C = \begin{bmatrix} c_{11} & 0 & 0 \\ c_{12} & c_{22} & 0 \\ c_{13} & c_{23} & c_{33} \end{bmatrix}$$

and from this, define the bivariate correlation matrix as

$$P_2 = \begin{bmatrix} 1 & c_{12} \\ c_{12} & 1 \end{bmatrix}$$

and the trivariate correlation matrix as

$$P_3 = \begin{bmatrix} 1 & c_{12} & c_{13} \\ c_{12} & 1 & c_{12}c_{13} + c_{23}\sqrt{1 - c_{12}^2} \\ c_{13} & c_{12}c_{13} + c_{23}\sqrt{1 - c_{12}^2} & 1 \end{bmatrix}$$

As the log likelihood function requires solving multidimensional integrals, we use maximum simulated likelihood Greene (2003). We are able to reduce computation time with Halton Halton (1960) and antithetic draws instead of pseudo-random draws. Halton and antithetic reduce variance in the simulator and allow us to achieve the same level of precision with less draws. Both measures introduce negative covariance across draws within and across observations Train (2003). Note that most conclusions about variance reduction are drawn from mixed logit and

Table 4: Outcome probabilities

| No. | Choice | Probability |
|---|---|---|
| (1) | Nothing | $P_1 = \Phi_2(-x_1\beta_1;\ -x_2\beta_2;\ P_2)$ |
| (2) | Car and no season ticket | $P_2 = \Phi_2(-x_1\beta_1;\ x_2\beta_2;\ P_2)$ |
| (3) | Car and local ticket | $P_3 = \Phi_3(x_1\beta_1;\ x_2\beta_2;\ -x_3\beta_3;\ P_3)$ |
| (4) | Car and GA | $P_4 = \Phi_3(x_1\beta_1;\ x_2\beta_2;\ x_3\beta_3;\ P_3)$ |
| (5) | No car and local ticket | $P_5 = \Phi_3(x_1\beta_1;\ -x_2\beta_2;\ -x_3\beta_3;\ P_3)$ |
| (6) | No car and GA | $P_6 = \Phi_3(x_1\beta_1;\ -x_2\beta_2;\ x_3\beta_3;\ P_3)$ |

fewer from multivariate probit Cappellari and Jenkins (2004). Thus, we verified that these options do not meaningfully alter estimates.

Maximum simulated likelihood estimators are consistent, asymptotically normal, efficient and equivalent to maximum likelihood if the number of draws tends faster toward infinity than the square root of observation numbers Train (2003). A sufficient criterion for model identification is that the model has instruments affecting the selection, but without effect on the outcome. Another criterion is to reject a restricted model with all cross equation correlations set to zero Jenkins et al. (2006).

We use 1000 Halton and antithetic draws and evaluate the log likelihood with a Geweke-Hajivassilou-Keane (GHK) based simulator Gourieroux and Monfort (1996) implemented in Stata StataCorp. (2015); Cappellari and Jenkins (2006). For the maximization routine, we use Stata's modified Newton-Raphson algorithm together with the robust option for robust standard errors. An estimation without the robust option results in an almost identical solution.

## 4.3 Explanatory variables

For the car and ticket equation, we use age, gender, household income employment status, university education, local access to public transport, predicted factor scores of the accessibility PCA and a dummy for living in one of the large Swiss cities. For the outcome of a GA or local season ticket, we assume that it is a question of income and choice of residential and workplace location. For the latter, the dataset has no measure available for all observations except commuters, but they make up only 50% of the sample. To circumvent the issue of losing 50% of observations, we use the daily traveled distance by public transport as the best available measure for all observations. This is, of course, endogenous. Further research must find an appropriate instrumental variable to account for residential and workplace choice. We also include a dummy variable in this equation indicating whether the household has a secondary residence.

We use, as instrument required for model identification, the measure of a household's local access to public transport and service level. We also estimated a model with predicted factor scores of the accessibility PCA in the season ticket type equation, but this model could not be identified because we could not reject the restriction of of cross-equation correlations set to zero

# 5 Results

The estimates of the univariate probit model are listed in Table 5 and the multivariate probit with sample selection in Table 6. Wald tests on hypothesis $\rho_{13} = 0$ lead to $\chi^2(1) = 386.5(p < .000)$ and show that the sample selection equation could not be ignored. In addition, Wald tests show that correlations between the two remaining equations are also significantly different from zero; for $\rho_{21} : \chi^2(1) = 386.5(p < .000)$ and for $\rho_{23} : \chi^2(1) = 173.6(p < .000)$. Finally, a Wald test with $\rho_{12} = \rho_{23} = \rho_{31} = 0$ has a $\chi^2(2) = 3268.5(p < .000)$ and shows that this constraint could be rejected and all equations must not be estimated independently.

The estimates show that the relationship between age and season ticket ownership is U-shaped (minimum ownership probability at 56.03 years). In contrast, relationship to the probability of car ownership is reversed (maximum at 57.14 years). Men are more likely to own a car than subscribe to a season ticket. University graduates tend to have a greater likelihood of subscribing to a public season ticket than owning a car. Being employed increases the probability of season ticket subscription and car ownership, with a greater effect on the latter. For household income, we find a strong influence on car ownership, and less on season ticket ownership. For the GA ticket choice, we find a positive income influence.

Better local access to public transport encourages the ownership of season tickets and discourages car ownership. The variables reflecting the first three factors obtained from the PCA with the Hansen accessibility measure at municipality level have a significant influence on the car and season ticket equation. While general levels of accessibility increase the probability of season ticket ownership and decrease the probability of car ownership, better access by public transport only significantly reduces probability of car ownership. Arguably, the lack of significant effect public transport accessibility on season ticket ownership is due to the fact that most public transport travel takes place within a municipality, especially in larger cities. Better access to jobs does favor the season ticket subscription, as opposed to car ownership.

Table 5: Univariate probit estimates

| | Car available | | Season ticket | | GA holder | |
|---|---|---|---|---|---|---|
| Age | 0.098*** | (0.002) | -0.072*** | (0.002) | | |
| Age squared | -0.086*** | (0.002) | 0.065*** | (0.002) | | |
| Male | 0.425*** | (0.015) | -0.115*** | (0.015) | | |
| Working | 0.247*** | (0.019) | 0.066** | (0.021) | | |
| University level education | -0.066** | (0.021) | 0.201*** | (0.021) | | |
| Log of monthly household income | 0.380*** | (0.015) | 0.077*** | (0.016) | 0.111*** | (0.025) |
| Local acc. to pub. transport: Level E | 0.512*** | (0.032) | -0.469*** | (0.033) | | |
| Local acc. to pub. transport: Level D | 0.391*** | (0.029) | -0.354*** | (0.030) | | |
| Local acc. to pub. transport: Level C | 0.297*** | (0.028) | -0.270*** | (0.028) | | |
| Local acc. to pub. transport: Level B | 0.165*** | (0.026) | -0.114*** | (0.026) | | |
| Local acc. to pub. transport: Level A | Reference | | | | | |
| 1. Fac.: General levels of accessibility | -0.027*** | (0.005) | 0.085*** | (0.006) | | |
| 2. Fac.: Better access by pub. transport | -0.077*** | (0.011) | 0.024* | (0.012) | | |
| 3. Fac.: Better access to employment | -0.540*** | (0.057) | 0.691*** | (0.062) | | |
| Center of agglomeration region | -0.231*** | (0.017) | 0.162*** | (0.018) | | |
| Secondary residence | | | | | 0.335*** | (0.051) |
| Self-reported distance [km] | | | | | 0.005*** | (0.000) |
| Constant | -5.922*** | (0.142) | 0.282 | (0.146) | -1.355*** | (0.217) |
| Observations | 52476 | | 52476 | | 11598 | |
| Pseudo $R^2$ | 0.145 | | 0.106 | | 0.049 | |
| *AIC* | 58128.44 | | 50237.75 | | 15516.13 | |
| ll | -29049.22 | | -25103.87 | | -7754.06 | |
| chi2 | 5900.395 | | 3488.387 | | 354.182 | |
| p | 0.000 | | 0.000 | | 0.000 | |

Standard errors in parentheses
$^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

The correlation in errors between the season ticket selection and GA outcome is positive, but shows that part of the reason for purchasing a season ticket is also relevant for choosing the GA. This is intuitive, as the reasons and utility gained through subscribing to one of the two season tickets vary: different geographic coverage and different prices, as well as further residential and workplace location choice effects, might not be observed.

Nonzero cross-equation correlations show that equations must not be estimated independently. Comparing both estimates shows that, in most cases, estimates deviate from each other that lead to different predictions in this model's applications.

Table 6: Multivariate probit estimates

| Variable | Estimate | Standard error |
|---|---|---|
| Season ticket owner | | |
| Age | -0.065$^{***}$ | (0.002) |
| Age squared | 0.058$^{***}$ | (0.002) |
| Male | -0.135$^{***}$ | (0.015) |
| Working | 0.073$^{***}$ | (0.020) |
| University level education | 0.151$^{***}$ | (0.021) |
| Log of monthly household income | 0.075$^{***}$ | (0.016) |
| Local access to public transport: Level E | -0.474$^{***}$ | (0.032) |
| Local access to public transport: Level D | -0.346$^{***}$ | (0.029) |
| Local access to public transport: Level C | -0.259$^{***}$ | (0.028) |
| Local access to public transport: Level B | -0.097$^{***}$ | (0.026) |
| Local access to public transport: Level A | Reference | |
| 1. Factor: General levels of accessibility | 0.091$^{***}$ | (0.006) |
| 2. Factor: Better access by public transport | -0.002 | (0.012) |
| 3. Factor: Better access to employment | 0.723$^{***}$ | (0.060) |
| Center of agglomeration region | 0.130$^{***}$ | (0.018) |
| Constant | 0.145 | (0.146) |
| Car available | | |
| Age | 0.096$^{***}$ | (0.002) |
| Age squared | -0.084$^{***}$ | (0.002) |
| Male | 0.428$^{***}$ | (0.014) |
| Working | 0.242$^{***}$ | (0.019) |
| University level education | -0.050$^{*}$ | (0.020) |
| Log of monthly household income | 0.380$^{***}$ | (0.015) |
| Local access to public transport: Level E | 0.506$^{***}$ | (0.031) |
| Local access to public transport: Level D | 0.383$^{***}$ | (0.029) |
| Local access to public transport: Level C | 0.288$^{***}$ | (0.028) |
| Local access to public transport: Level B | 0.155$^{***}$ | (0.026) |
| Local access to public transport: Level A | Reference | |
| 1. Factor: General levels of accessibility | -0.029$^{***}$ | (0.005) |
| 2. Factor: Better access by public transport | -0.069$^{***}$ | (0.011) |
| 3. Factor: Better access to employment | -0.546$^{***}$ | (0.056) |
| Center of agglomeration region | -0.222$^{***}$ | (0.017) |
| Constant | -5.864$^{***}$ | (0.142) |
| GA holder | | |
| Secondary residence | 0.304$^{***}$ | (0.044) |
| Log of monthly household income | 0.129$^{***}$ | (0.022) |
| Self-reported distance [km] | 0.005$^{***}$ | (0.000) |
| Cross equation correlations | | |
| r21: Season ticket and car available | -0.454$^{***}$ | (0.009) |
| r31: Season ticket and GA | 0.606$^{***}$ | (0.033) |
| r32: GA and car available | -0.247$^{***}$ | (0.019) |
| Observations | 52476 | |
| ll | -60177.517 | |
| k | 37.000 | |
| chi2 | 3467.822 | |
| p | 0.000 | |

### 5.1 Compare probabilities

We avoid computing marginal effects and conditional marginal effects Greene (2003) because we are interested in effects of continuous predictors. Instead, we compute probability expectancy value at sample mean and varying income, the first and second accessibility components. The GA outcome equation is applied to those cases with success in the selection equation. Figure 2 shows the predicted probabilities for each single outcome, the joint success of car and season ticket and conditional probabilities of car and season ticket ownership by varying income, see 2(a), first and second components, see 2(b) and 2(c), from the accessibility PCA.

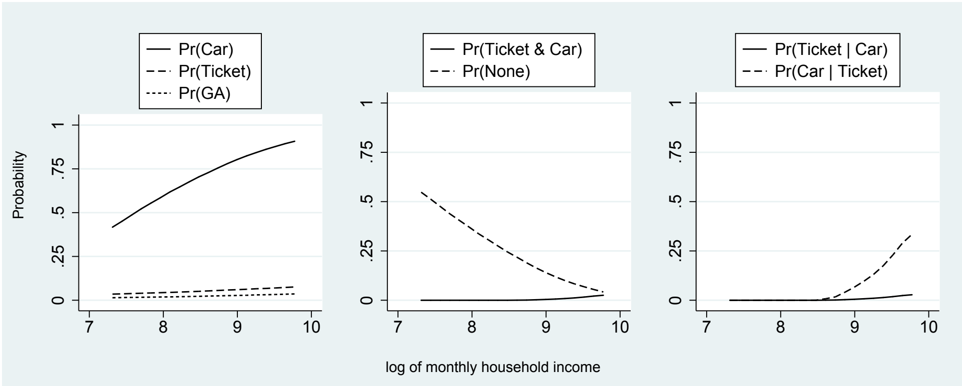
For household income effect, we observe that the probability of car ownership doubles in the considered interval. The probability of having no mobility tool tumbles from 50% almost to 0%, while having both mobility tools available sees only a slight increase. The conditional probability for a season ticket, given car ownership, follows a patter similar to the ownership of both categories. In contrast, given a subscription to a season ticket, ownership of a car becomes more likely with increasing income.

Considering the effect of general accessibility (first component of PCA) shows that, in the least accessible municipalities, car ownership tends to be unavoidable and season ticket ownership is only marginal. Whereas car ownership becomes less and less likely with increasing levels of general accessibility, a subscription to season tickets does not capture the remaining market shares in the same way. The likelihood of subscription grows strongest in the last quartile of general accessibility; non-ownership of a car or season ticket increases, while the probability of having both mobility tools increases marginally. The conditional probability of season ticket ownership given car ownership behaves similarly to the probability of owning both. Interestingly, the probability of car-ownership given season ticket ownership increases. Arguably, this can be attributed to the fact that season ticket ownership becomes relevant only at higher general accessibility levels.

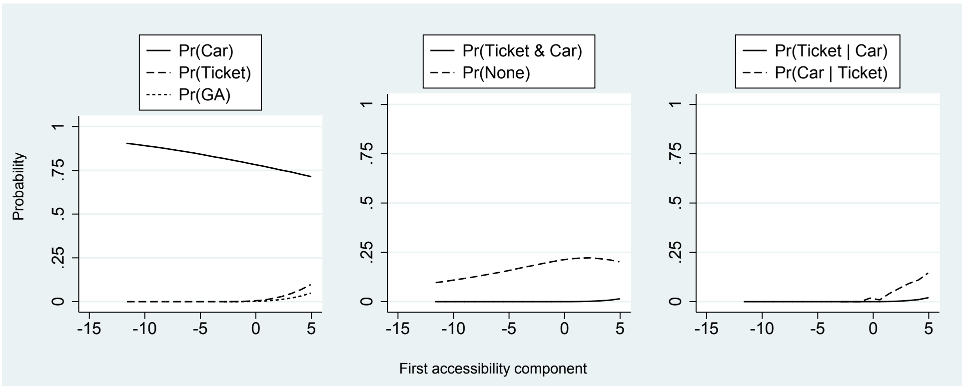
Better accessibility by public transport above and beyond general accessibility decreases the probability of car ownership, but has no effect on season ticket ownership. Having a season ticket and a car becomes less likely with greater accessibility by public transport, while at the same time, probability of having no mobility tool at all increases. Conditional probabilities reveal that the likelihood of subscribing to a season ticket given car ownership does not increase with better access by public transport. However, conditional probability of car ownership decreases given season ticket ownership with better accessibility by public transport.

Figure 2: Predicted probabilities

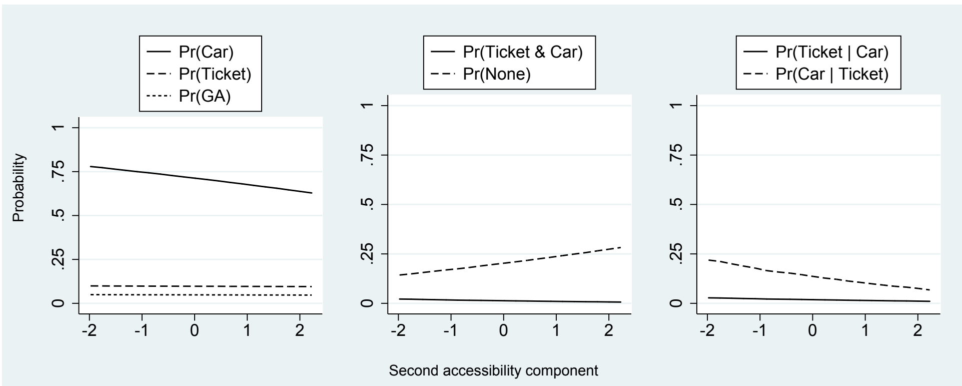
(a) Effects of income




(b) Effects of general levels of accessibility




(c) Effects of better access by public transport




# 6 Discussion

The results show that our proposed multivariate probit with sample selection model gives more consistent estimates than the univariate models. Correlations of unobserved variables show that car and the 'any season ticket' choice are substitutes and that the GA option is an even stronger substitute for a car. The correlation between the season ticket equation and the GA equation shows that reasons for subscribing to one of the tickets are similar, but not identical.

We acknowledge that distance traveled daily by public transport is highly endogenous in this context, but we see it as the best available variable for capturing households' residential and workplace decisions. Driver's license ownership is not included; license ownership as a selection equation for car ownership could be considered, but one could also argue that the costs of license acquisition are negligible compared to a car and a license can be acquired if needed.

It would be interesting to compare the accuracy of our proposed model to standard approaches, like the MNL, to see whether differences occur. This is relevant because the maximum simulated likelihood routine for large samples requires significantly greater computation time compared to MNL models.

At the moment, the estimation is carried out at nation-wide, or macro level. Further estimations could zoom into specific smaller areas and use accessibility values at a grid scale instead of municipality level. On the private mode side, car ownership can be split up into categories, e.g. displacement, or kind of cars, to also model the interactions between car type and season ticket ownership

# 7 Conclusions and outlook

Our findings contribute to existing literature in two ways. First, we introduce a discrete choice modeling technique to travel behavior literature: multivariate probit with sample selection. We show that this technique results in consistent estimates when correlated and mutually exclusive outcomes are considered simultaneously. Second, we provide a comprehensive mobility tool ownership model with an explanatory land-use accessibility variable. We find that differences in accessibility clearly explain differences in mobility tool ownership levels. Our model allows us to quantify accessibility changes' multimodal effects on mobility tools choice. This is especially interesting for transport planners and policy makers in an urban environment, with existing private and public modes.

Future research can extend our model to capture more correlated choices on the private transport side, e.g. car sharing membership or new ride services; these additional estimates will provide new insights across the range of mobility tools. Furthermore, our modeling technique could be applied to stated preference data, including price variation. In contrast to MNL and NL, the obtained estimates allow computation of conditional probabilities as a price function. Finally, our model could also be extended to model time series data and observe mode choice transitions.

# 8 ACKNOWLEDGEMENTS

This work was supported by ETH Research Grant ETH-04 15-1. The authors are grateful for Nathalie Picard's comments, which contributed significantly to this paper. The authors are also thankful to Karen Ettlin for valuable comments. An earlier version of this paper was presented at the 2016 discrete choice modeling workshop, held by Michel Bierlaire at EPFL Lausanne.

# 9 References

Axhausen, K. W., R. Neuenschwander, R. Fuhrer, P. Walker, G. Sarlas and T. Bischof (2015) Gesamtwir tschaftliche Effekte des öffentlichen Verkehrs mit besonderer Berücksichtigung der Verdichtungs- und Agglomerationseffekte, *Schlussbericht*, SBB Fonds für Forschung, Bern und Zürich.

Becker, H., F. Ciari and K. W. Axhausen (2016) Comparing Car-Sharing Schemes in Switzerland: User Groups and Usage Patterns, *Presented at 95th Annual Meeting of the Transportation Research Board, Washington, D.C.*

Ben Akiva, M. E. and S. R. Lerman (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press series in transportation studies, MIT Press, Cambridge.

Cappellari, L. and S. P. Jenkins (2004) Modelling low income transitions, *Journal of Applied Econometrics*, **19** (5) 593–610.

Cappellari, L. and S. P. Jenkins (2006) Calculation of multivariate normal probabilities by simulation, with applications to maximum simulated likelihood estimation, *Stata Journal*, **6** (2) 156–189.

Cervero, R. and K. Kockelman (1997) Travel demand and the 3Ds: Density, diversity, and design, *Transportation Research Part D: Transport and Environment*, **2** (3) 199–219.

Choo, S. and P. L. Mokhtarian (2008) How do people respond to congestion mitigation policies? A multivariate probit model of the individual consideration of three travel-related strategy bundles, *Transportation*, **35** (2) 145–163.

Crozet, Y., A. Mercier and N. Ovtracht (2012) Accessibility: a key indicator to assess the past and future of urban mobility, in *Accessibility analysis and Transport Planning Challenges for Europe and North America*, 263–279.

de Jong, G., J. Fox, A. Daly, M. Pieters and R. Smit (2004) A comparison of car ownership models, *Transport Reviews*, **24** (4) 379–408.

Ewing, R. and R. Cervero (2010) Travel and the Built Environment, *Journal of the American Planning Association*, **76** (3) 265–294.

FitzRoy, F. and I. Smith (1998) Public transport demand in Freiburg: why did patronage double in a decade?, *Transport Policy*, **5** (3) 163–173.

García-Ferrer, A., M. Bujosa, A. de Juan and P. Poncela (2006) Demand Forecast and Elasticities Estimation of Public Transport, *Journal of Transport Economics and Policy*, **40** (1) 45–67.

Goodwin, P., J. Dargay and M. Hanly (2004) Elasticities of Road Traffic and Fuel Consumption With Respect to Price and Income: A Review, *Transport Reviews*, **24** (3) 275–292.

Gourieroux, C. and A. Monfort (1996) *Simulation-based Econometric Methods*, Oxford University Press, Oxford.

Greene, W. H. (2003) *Econometric Analysis*, Prentice Hall, Pearson Education International, Upper Saddle River.

Gulati, R. and M. C. Higgins (2003) Which ties matter when? The contingent effects of interorganizational partnerships on IPO success, *Strategic Management Journal*, **24** (2) 127–144.

Halton, J. H. (1960) On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals, *Numerische Mathematik*, **2** (1) 84–90.

Hansen, W. G. (1959) How Accessibility Shapes Land Use, *Journal of the American Institute of Planners*, **25** (2) 73–76.

Heckman, J. J. (1976) The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models, *Annals of Economic and Social Measurement*, **5** (4) 475–492.

Heckman, J. J. (1979) Sample Selection Bias as a Specification Error, *Econometrica*, **47** (1) 153–161.

Henisz, W. J. (2000) The Institutional Environment for Multinational Investment, *Journal of Law Economics and Organization*, **16** (2) 334–364.

Hensher, D. A. and J. M. Rose (2007) Development of commuter and non-commuter mode choice models for the assessment of new public transport infrastructure projects: A case study, *Transportation Research Part A: Policy and Practice*, **41** (5) 428–443.

Hensher, D. A., J. M. Rose and W. H. Greene (2015) *Applied Choice Analysis*, Cambridge University Press, Cambridge.

Jenkins, S. P., L. Cappellari, P. Lynn, A. Jäckle and E. Sala (2006) Patterns of consent: Evidence from a general household survey, *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **169** (4) 701–722.

Jolliffe, I. T. (2002) *Principal component analysis*, Springer, New York.

Killer, V., R. Fuhrer, D. Guth, C. Holz-Rau and K. W. Axhausen (2013) Road accessibility of Germany and Switzerland 1970-2007, *Arbeitsberichte Verkehrs- und Raumplanung*, **936**, IVT, ETH Zurich, Zurich.

Kowald, M., B. Kieser, N. Mathys and A. Justen (2016) Determinants of mobility resource ownership in Switzerland : changes between 2000 and 2010, *Transportation*, 1–23.

Krugman, P. R. (1993) *Geography and trade*, MIT Press, Cambridge.

Metz, D. (2008) The Myth of Travel Time Saving, *Transport Reviews*, **28** (3) 321–336.

Potoglou, D. and P. S. Kanaroglou (2008) Modelling car ownership in urban areas: a case study of Hamilton, Canada, *Journal of Transport Geography*, **16** (1) 42–54.

Rentziou, A., C. Milioti, K. Gkritza and M. G. Karlaftis (2011) Urban Road Pricing: Modeling Public Acceptance, *Journal of Urban Planning and Development*, **137** (1) 56–64.

Schmid, B., S. Schmutz and K. W. Axhausen (2016) Explaining Mode Choice, Taste Heterogeneity, and Cost Sensitivity in a Post-Car World, *Presented at 95th Annual Meeting of the Transportation Research Board, Washington, D.C.*

Scott, D. M. and K. W. Axhausen (2006) Household mobility tool ownership: modeling interactions between cars and season tickets, *Transportation*, **33** (4) 311–328.

Simma, A. and K. W. Axhausen (2001) Structures of commitment in mode use: a comparison of Switzerland, Germany and Great Britain, *Transport Policy*, **8** (4) 279–288.

StataCorp. (2015) Stata Statistical Software: Release 14.

Swiss Federal Office of Spatial Development (ARE) (2011) *ÖV-Güteklassen - Berechnungsmethodik ARE*.

Swiss Federal Office of Spatial Development (ARE), Swiss Federal Office of Environment (BAFU) and Swiss Federal Statistical Office (BFS) (2011) *Landschaftstypologie Schweiz*.

Swiss Federal Statistical Office (BFS) (2012) *Mobilität in der Schweiz - Ergebnisse des Mikrozensus Mobilität und Verkehr 2010*, Swiss Federal Statistical Office (BFS), Neuchatel.

Swiss Federal Statistical Office (BFS) (2015) *Generalisierte Gemeindegrenzen der Schweiz*, Eidgenössisches Departement des Innern EDI, Neuchâtel.

Tanner, R. and D. Bolduc (2014) The Multiple Discrete-continuous Extreme Value Model (MDCEV) with Fixed Costs, *Procedia - Social and Behavioral Sciences*, **111**, 390–399.

Train, K. (1986) *Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand*, MIT Press, Cambridge.

Train, K. E. (2003) *Discrete Choice Methods with Simulation*, Cambridge University Press, Cambridge.

Van de Ven, W. P. and B. M. Van Praag (1981) The demand for deductibles in private health insurance, *Journal of Econometrics*, **17** (2) 229–252.

Van der Straeten, K., A. Trannoy, N. Picard and C. Hagneré (2003) L'importance des incitations financières dans l'obtention d'un emploi est-elle surestimée?, *Économie & prévision*, **160** (4) 49–78.

Venables, A. J. (2007) Evaluating Urban Transport Improvements, *Journal of Transport Economics and Policy*, **41** (2) 173–188.

Weis, C. and K. W. Axhausen (2009) Induced travel demand: Evidence from a pseudo panel data based structural equations model, *Research in Transportation Economics*, **25** (1) 8–18.