Irish Car Availability Logit Modelling

Author 1
Dr William Brazil*
NTA & TCD
William.brazil@nationaltransport.ie
**Corresponding author*

Author 2
Dr James Fox
Jacobs

Author 3
Wen Zhang
NTA

Abstract

The availability of a car when making a trip is one of the most important determinants of mode choice. This paper outlines the estimation of a trip level Car Availability model based upon nearly 29 thousand trips from the NTA's National Household Travel survey, accounting for trip purpose, origin area type, and area level car ownership statistics.

The determination of Car Availability on a trip basis is a key step when modelling mode and destination choice. Unlike car ownership, that may be estimated as an attribute of an individual or a zone, Car Availability is a trip level characteristic, describing the origin end availability of a car at the point the trip commences.

This paper outlines the results of a study undertaken to model trip level Car Availability for multiple trip purposes and area types based upon NHTS data. Utilising a logit choice modelling approach, this research examines the relationship between the origin Census Small Area (CSA), the purpose of trip being taken, the car ownership characteristics of the area, and the availability of car as a mode at trip level.

Additionally, this paper presents the application of the estimated model parameters to 18,641 small areas in Ireland, producing trip specific Car Availability proportions for each CSA for the census year of 2016. Geo-spatial analysis and mapping of Car Availability with respect to trip purpose is also presented, highlighting differing levels of trip specific Car Availability across the Irish state.

Results demonstrate significant differences in levels of Car Availability based upon the number of households within a CSA with no car, the area type of the CSA where the trip originated, and the purpose for which the trip was undertaken.

Keywords: Car Availability, Demand Modelling, Logit Models.

Introduction

Car Availability, or Car Ownership, is well established as one of the main determinants of mode choice within a transport context. Put simply, when the option to drive is available, for all but the shortest trips, it will tend to be the preferred mode. For the purposes of this work, Car Availability is defined as a trip being car available at the origin (where the trip is generated), so even if there is no parking present at the destination, the car is considered to have been available for the trip. This is due in part to the need to define Car Availability before the destination and mode choices, stages 2 and 3 of the traditional 4 stage model, are undertaken by the RMS.

Car Availability differs from Car Ownership, while being highly correlated. Specifically, Car Availability is an instantaneous characteristic of the trip being undertaken, rather than the trip maker. For example, while a household may possess a car, it may be in use by another family at the time the trip maker is required to travel, rendering the trip Non Car Available. Note that many transport models define 'Car Availability' as a

function of Car Ownership and (e.g.) the number of adults, however these models do not a measure of Car Availability that is defined at the trip level.

Whereas Car Ownership is a revealed and measureable quantity, Car Availability is a discrete rather than probabilistic trip level attribute that is determined in two means in transport surveys:

- Was the trip undertaken by car, as either a passenger or driver
- If not a car trip, does the survey respondent state there was car available for the trip

When applied to a Census Small Area, estimated levels of Car Availability refer to the number of trips that are estimated to be Car Available rather than the Car Availability level of each trip. For example if a CSA produces 200 trips and is estimated to be 65% Car Available, 130 of those trip will enter a mode choice set including car, and 70 will enter a choice set restricted to walking, cycling, and public transport. This practice of estimating at trip level and applying at area level is common in areas such as mode choice and trip generation.

Background

The logit model outlined in this paper is designed to be one in a series of models within the Irish regional Modelling System, so it will receive information on Car Ownership, itself derived from socio-economic variables, from preceding models. While this makes it slightly different from the area level Car Ownership models, as it is primarily concerned with trip level car availability, it is informative to consider previous work in the broader area of access to vehicle that has used similarity logistic and probit modelling approaches.

A number of studies have utilised approaches similar to the one presented in this paper, looking at Car Availability in Switzerland Beige et al [1] used a binary logit model, while in the area of Car Ownership ordered logit Models [2], mixed logit models [3], and binary probit models [4] have all been used. When reviewing factors impacting access to cars, a considerable number of explanatory variables have been utilised, including respondent specific attributes such as age and gender educational level [2], employment attributes [5], and pro-driving attitudes [6]. In addition, studies have also included area specific variables such as population and employment density land use mix, and distance to the CBD [7], as well as walking and transit accessibility [3].

Data

NHTS Records

The primary dependent variable is trip level Car Availability, and this is based upon the NTA's National Household Travel Survey. This survey was undertaken in 2017 and is designed to be nationally represented when trip makers are weighted. The survey comprised of 5,906 households with 10,289 trip makers.

The trips recorded in the National Household Travel Survey (NHTS) include the mode that was taken for the purpose of that trip. In addition, where the mode selected was not car, the respondent was asked to indicate whether or not a car had been available to them for the trip. In this way, it is possible to construct binary Car Availability variables, where a car is considered to be available if the respondent either travelled by car or indicated that there was a car available. This variable is considered to be the choice variable for the purposes of MNL modelling, with Car Available and Non-Car Available being the two discrete options.

In addition, the trip level NHTS records also contained information regarding the purpose of the trip, which was then used as an explanatory variable in the MNL model. In total, 54911 records were included in the initial analysis, however this fell to 29372 when return trips were excluded, as it is assumed for the NTA model that Car Availability is determined upon leaving the home.

SAPS Data

Car Availability is, at its most basic, a function of the fact that there is a car present to undertake a trip. This may be owned by the trip maker or someone in the household. While there may be additional Car Availability such as car sharing schemes or borrowing a car from friends or neighbours, previous NTA modelling has shown that household level Car Ownership is highly correlated to Car Availability.

While the observations of Car Availability are on a trip level, the model is designed to operate at CSA level, therefore levels of Car Availability will be determined by the characteristics of the CSA from which the trip originated, rather than of the trip maker. While this may represent a loss of detail with regard to explanatory power, it enables the parameters that are estimated to be applied at a CSA level in a consistent manner.

The CSO's Small Area Population Statistics (SAPS) (CSO, 2016) provides a large array of CSA level variables based on the Irish census, including the number of households within a CSA, and the number of households that don't own a car. These variables were then used to calculate the CSA level value for no-car households. This was then amended to each record in the NHTS dataset, with each trip therefore having an associated level of CSA no-car proportion.

Table 1 outlines a summary of the no-car households variable across the 18,641 CSAs. While there are CSAs where every household has access to a car, there are no CSAs where all households lack a car.

Mean	Std. Dev	Minimum	Maximum
0.13	0.14	0	0.84

Table 1: No-Car Households Descriptive Statistics

Trip Purpose

In contrast to the previous NCA/CA model used in the NTA's National Demand Forecasting Model, this model only uses one function, and rather than performing separate calculations for each trip purpose, takes trip purpose as dummy variable.

This model considers a number of trip purposes as defined by the NTA's NHTS. This model considers the following purposes:

- Travel to Work
- Travel to Education
- Visiting
- Food Shopping
- Non Food Shopping
- Escort to Work
- Escort to Education
- All Other Trips

Trip purposes are recorded by the survey respondent, and are at trip record level, rather than CSA level.

Area Types

This analysis uses four area type definitions as used by the NTA's planning sheets and these are applied at CSA level:

- City Region
- Large Urban Town
- Small Town
- Rural

Table 2 describes the variables used in this analysis, as well as their source and the level they are expressed at.

Data	Source	Level
Car Availability	NHTS	Trip Level
Trip Purpose	NHTS	Trip Level
Zero Car Households	SAPS 2016	CSA Level
Area Type	NTA Designation	CSA Level

Table 2: Variable Properties

Methods

4.1 Model Form

As the response type was binary in nature, the trip was deemed to be either Car Available or Non-Car Available, and the model outputs were required to be proportional (to allocate trips), the logit model was considered to be the most appropriate for the estimation.

The model used in this estimation is based upon a logit model with the form:

$$P_{nca} = \frac{e^{\alpha + \beta_x}}{1 + e^{\alpha + \beta_x}} \quad (1)$$

Where:

P_{nca} = the probability of being Car Non-Available at individual trip level

α = the constant term

β = the estimated parameters

x = the value of the independent variables

Therefore, given the estimated probability of being Non-Car Available, the probability of being car available (P_{ca}) is simply 1 minus the probability of Non-Car Available.

$$P_{ca} = 1 - P_{nca} \quad (2)$$

Results

Table 3 presents the results of the final Car Availability estimation. CA is the reference category, therefore all variable coefficients listed relate to NCA. Estimates are also presented in the form of odds ratios to aid with interpretation.

Higher (above 1) odds ratios represent higher levels of Non-Car Availability, or lower levels of Car Availability, with comparison to the reference category in the variable, while lower values (less than 1) represent the inverse.

In terms of model fit, pseudo rho square values of 0.5429 and 0.1659, with respect to zero and constants respectively, indicate what could be considered to be a good model fit [8].

Results demonstrate that levels of Car Availability vary with regard to trip purpose, with the highest levels being recorded for escort to work and the lowest for all the other trips category, With respect to work trips, the reference group, higher levels of Non-Car Availability (and conversely lower levels of Car Availability) are observed for both food and non-food shopping, and especially for education trips. A large proportion of education trips are taken under by individuals under the age of 17 (the legal age to drive in Ireland) and therefore it is not unexpected that levels of Car Availability are lower here, even accounting for parents escorting their children to school.

Coefficient	Estimate	Odds Ratio	T Ratio
Constants			
Car Available	Ref	1.00	Ref
Non-Car Available	-2.132	0.12	-39.8*
Trip Purpose Dummy Variables			
Work	Ref	1.00	Ref
Education	.7514	2.12	11.8*
Visiting	-.1806E-01	0.98	-.2
Food Shopping	.1837	1.20	2.7*
Non Food Shopping	.3825	1.47	7.6*
Escort to Work	-.4285	0.65	-4.6*
Escort to Education	-.1463E-01	0.99	-.1
All Other Trips	1.382	3.98	5.4*
Area Type Dummy Variables			
City Region AT1	Ref	1.00	Ref
Large Urban Town AT2	-1.008	0.36	-19.2*
Small Town AT3	-.7721	0.46	-15.6*
Rural Area AT4	-1.711	0.18	-26.0*
CSA No-Car Households	4.530		34.9*
Observations			29372
Final LL			-9181.7235
$Rho^2(0)$			.5429
$Rho^2(C)$			.1659
Weighted Observations			28978.66
CA	87.34%		25310.00
NCA	12.66%		3668.66

Table 3: Logit Model Results

Model Application

The primary purpose of the NCA/CA modelling is to determine rates of Car Availability on a CSA level, and not just establish the role of the independent variables. To this end, the second part of this project was based around the application of the model parameters to all 18,641 CSAs in Ireland, to produce trip purpose specific NCA/CA values based upon CSA area type and the proportion of households that fall within the no-car category. It should be noted that while the parameters where estimated based on household level observations, with CSA level independent variables, the proportions/percentages reported apply to CSAs, as the model is applied at CSA level.

Census Year Application

Using the estimated parameters, it is possible to estimate the resulting levels of CA/NCA for each CSA in the state, and for each of the trip purposes considered. Therefore, for each of the 18,641 CSAs in the state, a ratio of CA/NCA trips can be estimated for each trip purpose. For the purposes of brevity, only work trips are presented, although such analysis is available for all trip types.

One of the principal issues inherent with the concept of CA/NCA is centred on the fact that this measure is unique to the NHTS, and therefore cannot be validated against an external data source, such as SAPS data. Therefore, one means of providing added reassurance is compare the purpose specific weighted CA/NCA averages recorded in the NHTS, with those calculated for the CSAs based upon the model parameters, taking into account the NHTS has been designed to a nationally representative survey of travel patterns.

Work Trips

The application of the estimated model to the SAPS data produced the following mean values. With a mean value of 87.75% it is observed that the majority of trips within an Irish commuting context can be considered to be highly available to cars, however a significant range of values is observed, with an estimated lower value of 10.95%. Table 4 highlights the results of the application of the model to work trips.

Mean	Std. Dev	Minimum	Maximum
87.75%	13.89%	10.95%	97.90%

Table 4: Estimated Work Trips Descriptive Statistics

Figure 1 outlines the distribution of CSAs in terms of the estimated levels of Car Availability for work trips originating in each area. It is clear that the majority of CSAs can be considered to be highly car available, with over half of the 18,641 areas being estimated to be more than 90% car available.

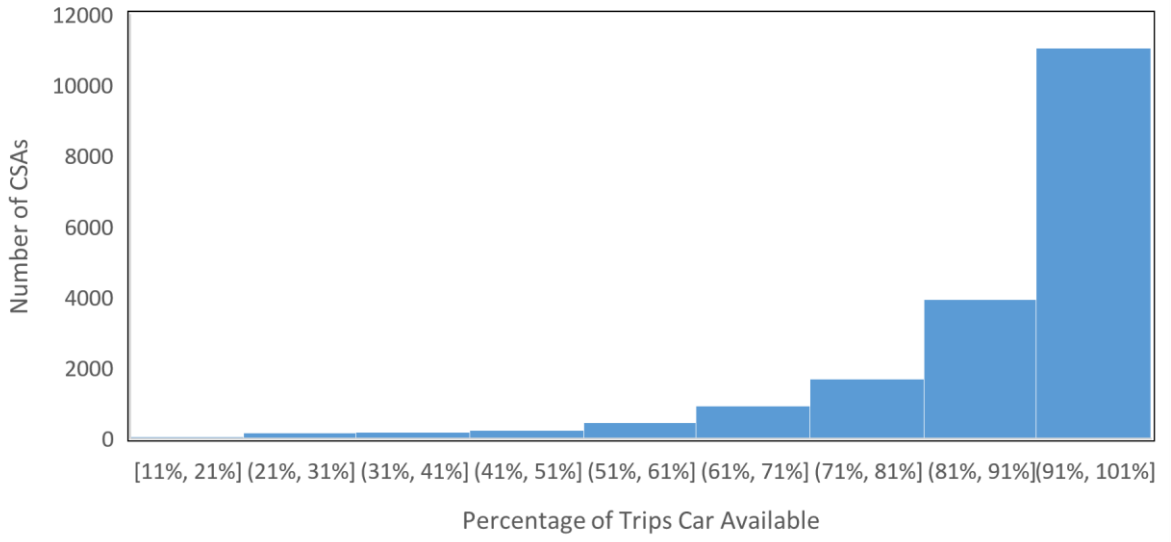

Figure 1: CSAs by Car Availability Levels

Figure 2 outlines the relationship between the proportion of households in a given CSA without a car available and the percentage of work trips that are estimated to be car available. This figure clearly demostrates the inverse relationship between the two variables, as well as the impact of the categorical area types on the model, and finally also the underlying logistic regression function (the figure contains four distinct lines, as this relates to the relationship between Car Availability and No_Car Households for each of the distintch area types as expected with dummy variables). It should be remembered that as this model is being applied for only work trips, the trip purpose variable is not included in the application of the model, and therefore is not reflected in the figure. With regard to the area types, in general, have both the highest levels of car availability, and the lowest levels of no-car households, as would be expected given the more dispersed settlement patterns in such areas and the lower frequency public transport services that are provided.

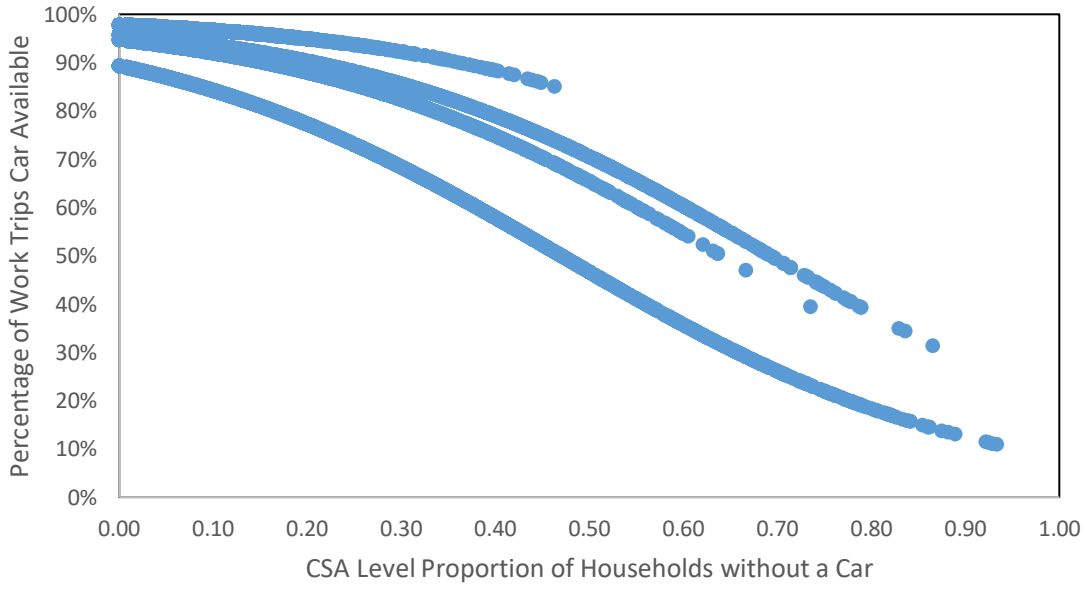

Figure 2: Car Availability vs. No-Car Households

Discussion and Conclusions

This paper outlines research taken to estimate levels of trip purpose specific Car Availability at Census Small Area level within Ireland. Car Availability is an important factor when undertaking mode choice modelling, especially in a car dominated country such as Ireland. This work was undertaken using a combination of an extensive travel dataset, in combination with car ownership data from the Irish Census of 2016. Using a binary logit model, estimation results show clear and significant differences in Car Availability based upon factors such as area type and the number of no-car households in the area. In addition, the model estimation reveals significant differences in Car Availability with respect to trip purpose, suggesting that there may be intra-household dynamics at play, in terms of when and where access to a car is prioritised. This model has been designed with the intention of being integrated into a wider mode and destination choice modelling structure, and further work will examine its potential as part of the implementation of nested logit modelling.

References

[1] Beige, S. and K. Axhausen (2012) Interdependencies between turning points in life and long-term mobility decisions. Transportation, 39, 857-872.

[2] Xinja, C., P. Næss, F. Wolday (2019) Examining the effects of the built environment on auto ownership in twot Norwegian urban regions. Transportation Research Part D, 67, 464-474.

[3] Oakil, A.T.M., D. Ettema, T. Arentze, H. Timmermans (2014) Changing household car ownership level and life cycle events: an action in anticipation or an action on occurrence. Transportation 41, 889-904.

[4] Goetzke, F. and R. Weinberger (2012) Separating contextual from endogenous effects in automobile ownership models.

[5] Anowar, S. N. Eluru and L. Mirandar-Moreno (2014) Alternatlve Modelling Approaches Used for Examining Automobile Ownership: A Comprehensive Review.

[6] Cao, J. and X. Cao (2014) The Impacts of LRT, Neighbourhood Characteristics, and Self-Selection in Auto Ownership: Evidence from Minneapolis-St Paul.

[7]Ding, C., Y. Wang, J. Yang, C. Lui and Y. Lin (2016) Spatial heterogenous impact of built environment on household auto ownership levels: evidence from analysis at traffic anlaysis zone scales. Transportation Letters, 8:1, 26-34.

[8] Hensher, D., Rose, J., & Greene, W. (2015). Applied Choice Analysis (2nd ed.). Cambridge: Cambridge University Press. doi:10.1017/CBO9781316136232