



Munich Personal RePEc Archive

Airport and Access Mode Choice : A Generalized Nested Logit Model Approach

Gelhausen, Marc Christopher and Wilken, Dieter

German Aerospace Center (DLR), Air Transport and Airport
Research

2006

Online at <https://mpra.ub.uni-muenchen.de/4311/>
MPRA Paper No. 4311, posted 01 Aug 2007 UTC

Gelhausen, Wilken

ATRS 2006 World Conference, Nagoya, Japan

Airport and Access Mode Choice

- A Generalized Nested Logit Model Approach –

Marc Ch. Gelhausen, Dieter Wilken
German Aerospace Center (DLR)
Air Transport and Airport Research
Cologne, Germany
Tel.: ++49 2203 601 2463
E-mail: Marc.Gelhausen@dlr.de

The authors want to thank Peter Berster for his helpful contributions in providing statistical information on airport and access mode choice on the basis of the Air Traveller Survey of the German Airport Association (ADV) of 2003.

ABSTRACT

This paper presents a “Generalized Nested Logit Model for a Combined Airport and Access Mode Choice”. The paper starts with a description of the airport system in Germany and some empirical findings concerning access mode and airport choice. The second part of this paper deals with the estimation of the aforementioned model.

The aim is to develop a generally applicable airport and access mode choice model. Thereby it is possible to analyse future scenarios not existing today. Kohonen’s Self-Organizing-Maps are used to identify different airport categories. Based on these airport categories the aforementioned nested logit model is estimated.

KEYWORDS: Access mode choice, airport choice, airport cluster, air traveller survey, grouping of alternatives, independence from irrelevant alternatives, Kohonen’s Self-Organizing Maps, nested logit model, neural networks

1. Introduction

Modelling airport choice has been a subject of interest to air transport scientists and airport planners already for a while. The authors have reported of a market segment specific model approach to airport choice in Germany in a paper entitled “Airport Choice in Germany – New Empirical Evidence of the German Air Traveller Survey 2003 –“ presented at the ATRS 2005 Conference in Rio de Janeiro, Brazil. In continuation of the analysis of airport choice, based on

the evidence coming from the data of the survey mentioned, this paper deals with the attempt to model the combined airport and access mode choice in Germany by market segment.

Why model the combined airport and access mode choice? Our hypothesis was that an airport with good rail accessibility attracts passengers from a greater influential area, because they can reach the airport within a relatively short time by means of fast trains. This would at the same time mean that the modal split of the feeder travel flows would have higher portions of fast train usage.

As we have seen from the data of the German Air Traveller Survey 2003 potential air travellers will prefer to choose an airport with good train service accessibility, given a good offer of flight services at that airport. Indeed, an airport with fast train access attracts passengers from other cities and urban areas served by fast inter-city trains, and these passengers will prefer these services when travelling to the airport in order to start the air journey proper there. A model describing the combined choice of access mode and airport may be a tool therefore for railway companies to forecast patronage of trains serving directly airports. In planning routes of trains serving airports the German railway company Deutsche Bahn has a great interest to study route alternatives and thus optimise the routing and scheduling process for future airport trains. The development of the model dealt with in this paper stems from this interest, testing of the model for practical planning purposes of the Deutsche Bahn is about to commence.

The prime objective of the paper is to report on the methodological progress in devising and calibrating a so called generalized nested logit model of a combined access mode and airport choice for 7 different market segments of the total air travel demand in Germany based upon the doctoral thesis of one of the authors, Marc Gelhausen. In supporting the model approach we will describe shortly the actual situation of airport choice in Germany, that is describe the different roles that airports play with respect to attracting passengers due to varying levels of rail infrastructure at German airports, the fast train utilisation in feeder travel to and from airports and describe some general characteristics of airport and access mode choice in Germany.

The outline of the paper is as follows:

- Airports in Germany: Functions and modal access
- Access mode and airport choice: Some findings in Germany
- Grouping of alternatives in models of discrete choice
- Database and grouping of clusters: Kohonen's self-organizing maps
- Preparing the data set for model estimation and estimation results
- Conclusions and further research needs

2. Airports in Germany: Functions and Modal Access

Germany, a country with a population of about 82 million people and a size of nearly 360 000 km², has a rather dense network of classified airports. There are first of all 17 international airports (with the Berlin airport system consisting of 3 single airports counting as one) which – together with some 10 regional airports – serve primarily the public air transport system with scheduled and non-scheduled services on domestic and international traffic relations. In 2005, the

international airports handled a traffic volume of 164 million passengers enplaned and deplaned and of about 2.2 million air transport movements (ATM's) on mainly scheduled services.

Commercial air traffic in Germany is rather concentrated on a few airports, Frankfurt alone accounts with 52 million passengers for almost one third of the total passenger volume of Germany. The six airports with the highest traffic volumes, Frankfurt, München, Düsseldorf, Berlin, Hamburg and Stuttgart handle about 133 million passengers, which make up for more than 80 % of the total traffic. Only these airports have already now or face in the near future capacity problems, the other 11 international and the 5 regional airports, which carry less than 20 % of the traffic, have sufficient capacity surplus.

While most of the airports offer a mixture of scheduled, charter and tourism services and other non-scheduled services, there are several airports that rely more or less on low cost carriers. Hahn is an airport almost exclusively used by one low cost carrier, namely Ryanair; although this airport is located in rural areas and far from any agglomeration Ryanair succeeded within 4 years to generate a traffic volume of 3 million passengers in 2005. Other airports with high portions of low cost carrier traffic are Köln/Bonn, Berlin-Tegel and –Schönefeld, München, Stuttgart and Düsseldorf. Frankfurt as an airport with severe capacity problems does not provide any low cost carrier services, one reason being the scarcity of available slots.

The two biggest airports, Frankfurt (52 million pass.) and München (29 million pass.) function as important airports for the origin-destination traffic as well as Lufthansa-hubs for corresponding traffic. In Frankfurt, more than 50 % of all passengers are transfer passengers taking advantage of the hub function, while in München this portion is smaller (35 %), however, growing from year to year with the still growing hub function of the airport.

Access conditions to international German airports vary with their function and traffic importance. Clearly, all airports have more or less good road and motorway connections, and a growing number is equipped with rail terminals for urban transit or regional trains (Bremen, Hannover, Berlin-Tempelhof and –Schönefeld, Dresden, Leipzig/Halle, Düsseldorf, Köln/Bonn, Frankfurt, Stuttgart, Nürnberg and München). There are thus less international airports still lacking direct rail access, they are Hamburg, Berlin-Tegel, Hahn, Dortmund, Münster/Osnabrück, Erfurt and Saarbrücken. Among these, Berlin-Tegel and Hamburg are important airport carrying 11.5 resp. 10.6 million passengers in 2005, the other airports have much lower traffic volumes. All airports except Bremen are served by scheduled bus lines, maintaining regional access to the urban area of the airport and inter-modal connections, in particular to main rail stations.

From an integrated transport planning standpoint it is interesting to connect airports directly with inter-city trains so that short distance flights may be substituted by high speed trains and the modal split of trips complementary to air trips be changed to a higher degree of rail usage. Only three airports in Germany are multi-modal centers in a sense that they have next to the air terminal one or two rail terminals which serve regional trains as well as form part of the inter-regional rail network. Köln/Bonn has the train station next to the terminal early since 2004.

An example par excellence is Frankfurt airport with a terminal for trains serving the regional network and a separate rail terminal (next to the terminal for regional trains) with high speed inter-city trains serving the northern and southern networks of Deutsche Bahn in Germany. There are several high speed trains (so called ICE-trains, the highest train category of Deutsche Bahn)

per hour departing from and arriving in Frankfurt airport with destinations in a great number of urban areas in Germany. Since 2003 a new high speed line has been in service between the Rhein/Main and the Rhein/Ruhr Region, interconnecting Frankfurt airport not only with Köln and cities further north, but also directly with Köln/Bonn airport since 2004, and the same in the other direction, the urban area of Köln is not only connected with the Frankfurt region but also directly with Frankfurt airport. This of course includes a direct rail link between the two airports.

Lufthansa is using reserved seats in these trains in order to feed long distance flights from Frankfurt with passengers from the Köln region, thus saving costs for otherwise providing short haul flights between Köln/Bonn and Frankfurt. The travel time by train is with one hour almost as short as the flight time, there is, however, a much denser frequency of services by train. Lufthansa has therefore created check-in facilities at Köln main station so that passengers can check through their luggage from Köln to the final destination and vice versa. This rail-air service can be regarded as an example for further inter-modal connections primarily connecting cities without direct access to the air transport system with hub airports by means of high speed trains, which serve as a “feeder flight” to typical flight connections with European and inter-continental destinations from the hub airport.

The two other airports with direct inter-city train services are Düsseldorf and Leipzig/Halle. Düsseldorf is the third biggest airport in Germany with a traffic volume of 15.5 million passengers in 2005. Because of capacity problems Düsseldorf has not succeeded to build up hub functions. The airport is located at the Western edge of the Rhein-Ruhr District, an industrialized region with about 10 million people, predominantly living in urban areas. This catchment area is the biggest of its kind in Germany, Düsseldorf has to share it, however, with two other airports, that is Dortmund (1.7 million passengers) and Köln/Bonn (9.4 million passengers). Leipzig/Halle on the other side is a relatively small airport in East-Germany, which handled in 2005 about 2 million passengers. In contrast, the second biggest airport in Germany, München, is connected only to the urban rail network, the next inter-city rail terminal in München is more than 45 Minutes travel time away.

3. Access Mode and Airport Choice: Some Findings in Germany

The aforementioned paper (Wilken et al. 2005) has reported in extenso of the findings of the German air traveller survey 2003 regarding airport choice behaviour. The main conclusions have been, that – on the basis of the regional structure given by 97 Spatial Planning Regions (SPR's) in Germany (see also chapter 4) with an average population of 850 thousand inhabitants each– air travellers of any region choose on average between eight airports, with a standard deviation of two airports. Interestingly, the demand of the SPR Düsseldorf, in which Düsseldorf airport is located, is using 14 airports for commencing a flight; no other region in Germany is served by so many airports. On the other hand, the minimum number of airports used by the demand of one region – and thus the highest concentration of selecting airports in the revealed choice behaviour - is 3 airports. There is not a single zone the demand of which is served by just one airport. We can conclude that although there is a concentration of demand in selecting the nearest or most attractive airport -in fact, 67 % of the demand of an airport region chooses that airport -, in a decentralized airport environment like in Germany air travellers of a given region are selecting between a great number of airports - actually eight - for departing for an air journey.

We have seen as a result of an airport choice model estimation that airport choice varies with the type of air journey. A market segmentation according to the trip purpose (business, holiday and short stay personal) and trip distance (domestic, European and intercontinental) has shown that overall the two main factors influencing airport choice are the travel time to the airport and the existence of a direct flight connection to the destination airport, with the access travel time having the strongest influence on airport choice. For long haul flights the number of destinations served plays a role, too. The shorter the flight the more passengers choose an airport with a shorter access time. Business travellers are more time-sensitive than private travellers. The existence of a direct flight connection is more important for private travellers than for business travellers and the shorter the flight distance the more important this attribute becomes for international flights.

One question in the German air traveller survey 2003 referred to the mode of transport used to reach the airport. The prespecified options and their modal share over all airports were

- private car driver: 18 %
- car passenger: 34 %
- rental car driver: 4 %
- taxi passenger: 19 %
- bus passenger: 9 %
- regional transit passenger: 11 %
- train passenger: 5 %

Three quarters of all air passengers in Germany reached the airport by car, however, only 18 % by driving a private car, most of the travellers coming to the airport by car have been car passengers. Only one quarter of all travellers came by public mode, most of them by public transit. Since access mode conditions vary from airport to airport the overall modal split can give only a general picture of mode choice, each airport has a specific modal distribution depending on the spectrum of access mode services and the importance of the airport in terms of supply of air services. In addition, individual preferences of travellers depend on personal factors, like car availability, locational conditions, like relative and absolute accessibility of airports, the type of journey, of modal supply conditions and air service quality of each airport.

To demonstrate the modal variation between airports we can select a few rather unique examples, Frankfurt as an airport with a rich supply of regional and inter-city trains, Hahn as a low cost carrier airport relying solely on road (car and bus) access and Hamburg as an airport lying in an urban agglomeration without public rail transit, however, good regional bus services.

The public mode portion in Frankfurt lies with 32 % clearly above average (25 %), most of the passengers arriving at Frankfurt airport by means of a public mode have chosen a train. In fact, 18 % arrive by intercity train, thus coming from journey origins lying in greater distances away from the airport than typically found. Indeed, Frankfurt has the greatest influential area of any airport in Germany, also caused to a great deal by the great number of flight destinations and frequencies offered at that airport (490 thousand ATM's in 2005). This modal concentration on trains verifies empirically the working hypothesis that fast train services will influence travellers' behaviour regarding the choice set "access mode and airport", in this case in favour of the intercity train to Frankfurt airport.

In contrast, passengers to Hahn airport, flying primarily for non business reasons on low cost services, use in nine out of ten cases the private car, a taxi or rented car, only 11 % arrive by bus. Although the modal choice is limited to car and bus (offered only on a small number of routes) people come from places partly far away in order to catch low price flights. For the most part, people have to cover long ways to Hahn airport, since this airport is lying in a somewhat remote location, the nearest urban area being more than 100 kms away. Would Hahn airport be interconnected with the intercity rail network patronage would be much higher, since people prefer to take the train on longer distances. This will be shown when we look into the modal split of a sample of access relations to and from Frankfurt.

Although Hamburg airport is located in an urban area of about 1.7 million people the site is not connected yet with the otherwise good rail transit network. Busses serve the airport, their modal share is therefore twice as high as on average. Nevertheless, the modal share of the car (private car, rented car, taxi) is with over 80 % higher than normally found at airports in Germany. One reason is the high patronage of taxis: Distances within the Hamburg region to the airport are comparatively short, the portion of passenger arriving by taxi is therefore with over 30 % relatively high, in fact, this portion is the highest modal share at German airports.

As has been stated access mode choice varies with the trip purpose, too. Business travellers rely much more on taxi and rented car when going to the airport, 42 % of them use these modes, whereas private travellers are brought by private car to a much greater deal (45 %) to the airport, a taxi and rented car is chosen only in 15 % of all trips.

For obvious reasons, passengers arriving at an airport by flight choose more than passengers commencing their air journey there, taxi and rented car in order to reach their final destination (35 % versus 16 %), passengers beginning the journey prefer to take a private car more often (23 % as opposed to 3 % of arriving passengers).

Urban Area	Share of Passengers Using Intercity Trains (%)	Distance to Frankfurt (kms)
Hamburg	83	495
Bremen	57	445
Hannover	68	350
Berlin	36	545
Dortmund	57	225
Düsseldorf	70	220
Köln	68	180
Leipzig	82	385
Stuttgart	50	205
Nürnberg	44	225
München	66	390

Tab. 3.1: Share of Passengers Using Intercity Trains to Frankfurt Airport

As has been shown Frankfurt airport attracts the highest portion of passengers arriving by intercity trains (18 %). This average shades the high portions of train passengers coming from other cities, which are typically much better served by intercity trains than rural areas. All big urban areas in Germany are more or less served and interconnected by intercity trains, and since these areas of population and economic concentrations are at the same time important demand

generators, we find in most of them airports as well. It might be interesting therefore to analyse the modal share of passengers coming from these urban areas and choosing the train to Frankfurt airport to either begin or finish the flight. Based on the German air traveller of 2003 survey we can derive these data for a selection of 12 urban areas (see table 3.1).

As can be seen passengers coming from other urban areas and choosing Frankfurt as airport take in most cases the intercity train, the train share varies between 36 % for Berlin passengers and 83 % for Hamburg passengers. The variation may be partly caused by differences in train speed, train speeds between Hannover, Düsseldorf and Köln, and Stuttgart and München and Frankfurt are much higher than on the other routes. The combined preference for fast trains and Frankfurt as departure airport suggests to simulate this behaviour in a single model of access mode and airport choice.

4. Grouping of Alternatives in Discrete Choice Models

The fundamental hypothesis of discrete choice models is the assumption of individual utility maximization. Alternatives are evaluated by means of an utility function and the one with the highest utility is supposed to be chosen. From an external point of view the utility of an alternative for a specific individual is a random variable, so that the utility U_i for alternative i is composed of a deterministic component V_i and a random component ε_i (Maier et al. 1990, p. 100):

$$(4.01) \quad U_i = V_i + \varepsilon_i$$

The random component of the utility function is introduced for various reasons, i.e. a lack of observability of the relevant attributes of the alternatives or their incomplete measurability (Maier et al. 1990, pp. 98f.).

From an external point of view, only evidence in terms of the probability of an alternative being the one with the highest utility can be given, because of the random component in the utility function. Specific discrete choice models differ in terms of their assumptions of the random component. The most prominent member of this class of models is the logit-model with independently and identically distributed random components. The choice probability of an alternative i is computed as (Train 2003, p. 40):

$$(4.02) \quad P(a_i = a_{opt}) = \frac{e^{\mu V_i}}{\sum_j e^{\mu V_j}}$$

As a consequence of the independently and identically distributed random components of the utility functions the ratio of two choice probabilities is only dependent on the utility of those two alternatives (Ben-Akiva et al. 1985, p. 108):

$$(4.03) \quad \frac{P(a_i = a_{opt})}{P(a_j = a_{opt})} = \frac{\sum_k e^{\mu V_k}}{\sum_k e^{\mu V_k}} \cdot \frac{e^{\mu V_i}}{e^{\mu V_j}} = \frac{e^{\mu V_i}}{e^{\mu V_j}}$$

This property of the logit-model is called “Independence from Irrelevant Alternatives” (IIA) and it is both a weakness and a strength of the model. Due to the distribution assumptions of the random component of the utility function it is not possible to model correlations among alternatives owing to unobserved factors. A major advantage of the IIA-property is the possibility to estimate the model parameters, excluding alternative-specific variables, on a subset of the alternatives (McFadden 1974, p. 113; McFadden 1978, pp. 87ff.; Ortuzar et al. 2001, pp. 227f.; Train 2003, pp. 52f.) and the possibility of an evaluation of new alternatives without the need to re-estimate alternative-unspecific model parameters (Domencich et al. 1975, p. 69f.). The problem of estimating alternative-specific variables from a subset of alternatives will be discussed below.

The nested logit-model relaxes the IIA-restriction to some extent without losing the closed-form expression of the choice probabilities. For this purpose the random component in (4.01) is split up into a part ε_i^a , which varies over all alternatives i and a part ε_k^c , which is identical for all alternatives of a nest k (Maier et al. 1990, p. 154f.):

$$(4.04) \quad U_i = V_i + \varepsilon_i^a + \varepsilon_k^c$$

It is possible to model correlations due to unobserved factors among subsets of the alternatives, so that the choice set is partitioned into clusters with highly correlated alternatives. (4.05) is an example of a covariance matrix for 4 alternatives partitioned into two clusters with the first two belonging to cluster one and the last two assigned to cluster two.

$$(4.05) \quad \Omega = \begin{bmatrix} \sigma_{11}^2(\mu_1^c) & \sigma_{12}^2(\varepsilon_1^c) & 0 & 0 \\ \sigma_{21}^2(\varepsilon_1^c) & \sigma_{22}^2(\mu_1^c) & 0 & 0 \\ 0 & 0 & \sigma_{33}^2(\mu_2^c) & \sigma_{34}^2(\varepsilon_2^c) \\ 0 & 0 & \sigma_{43}^2(\varepsilon_2^c) & \sigma_{44}^2(\mu_2^c) \end{bmatrix}$$

Each cluster k is characterized by an individual scale parameter μ_k^c and an identical non-negative covariance for all alternatives i within a cluster k . Alternatives of different clusters are assumed not to be correlated.

For technical reasons the choice probabilities $P(a_i = a_{opt})$ are decomposed into an unconditional choice probability $P(c_k = c_{opt})$ that cluster k is chosen, and a conditional choice probability $P(a_i = a_{opt} | a_i \in c_k)$, that alternative i from cluster k is chosen (Maier et al. 1990, p. 156):

$$(4.06) \quad P(a_i = a_{opt}) = P(a_i = a_{opt} | a_i \in c_k) * P(c_k = c_{opt})$$

The conditional choice probabilities comply with the logit-model and the choice set is restricted to the alternatives of the appropriate nest. The choice probability of a nest k is determined by its maximum utility V_k^c (Maier et al. 1990, p. 157):

$$(4.07) \quad V_k^c = \frac{1}{\mu} \ln \sum_{i \in k} e^{\mu V_i}$$

The choice probability of an alternative i in nest k can be written as (Maier et al. 1990, p. 158):

$$(4.08) \quad P(a_i = a_{opt}) = \frac{e^{\mu V_i}}{\sum_{j \in k} e^{\mu V_j}} * \frac{e^{\mu_k^c V_k^c}}{\sum_l e^{\mu_l^c V_l^c}}$$

The hierarchical structure of (4.08) does not imply a sequential decision process. An extension to more than two levels is possible (see i.e. Ben-Akiva et al. 1985, p. 291ff.).

In the nested logit-model the IIA-property does only hold for two alternatives of the same cluster:

$$\begin{aligned}
 & \frac{P(a_1 = a_{\text{opt}} | a_1 \in c_1) * P(c_1 = c_{\text{opt}})}{P(a_2 = a_{\text{opt}} | a_2 \in c_1) * P(c_1 = c_{\text{opt}})} \\
 (4.09) \quad &= \frac{\frac{e^{\mu V_1}}{\sum_{j \in c_1} e^{\mu V_j}} * \frac{e^{\mu_1^c V_1^c}}{\sum_1 e^{\mu_1^c V_1^c}}}{\frac{e^{\mu V_2}}{\sum_{j \in c_1} e^{\mu V_j}} * \frac{e^{\mu_1^c V_1^c}}{\sum_1 e^{\mu_1^c V_1^c}}} \\
 &= \frac{e^{\mu V_1}}{e^{\mu V_2}}
 \end{aligned}$$

The ratio of the choice probabilities for two alternatives of different clusters depends on the characteristics of all alternatives of those two clusters:

$$\begin{aligned}
 & \frac{P(a_1 = a_{\text{opt}} | a_1 \in c_1) * P(c_1 = c_{\text{opt}})}{P(a_2 = a_{\text{opt}} | a_2 \in c_2) * P(c_2 = c_{\text{opt}})} \\
 (4.10) \quad &= \frac{\frac{e^{\mu V_1}}{\sum_{j \in c_1} e^{\mu V_j}} * \frac{e^{\mu_1^c V_1^c}}{\sum_1 e^{\mu_1^c V_1^c}}}{\frac{e^{\mu V_2}}{\sum_{j \in c_2} e^{\mu V_j}} * \frac{e^{\mu_2^c V_2^c}}{\sum_1 e^{\mu_2^c V_2^c}}} \\
 &= \frac{\frac{e^{\mu V_1}}{\sum_{j \in c_1} e^{\mu V_j}}}{\frac{e^{\mu V_2}}{\sum_{j \in c_2} e^{\mu V_j}}} * \frac{e^{\mu_1^c V_1^c}}{e^{\mu_2^c V_2^c}}
 \end{aligned}$$

As the nested logit-model lacks the IIA-property for some pairs of alternatives, a model estimation on a subset of the choice set equal to the logit-model is not possible.

If it is feasible to form groups of at least approximatively similar clusters and to assign an identical covariance matrix for all clusters of the same group, an estimation of alternative-unspecific model-parameters equal to the logit-model on a subset of alternatives is possible. Each group of clusters must be represented by at least one member in this subset to enable the estimation of all cluster-specific scale parameters. (4.11) shows a covariance-matrix for six alternatives belonging to three groups, with two alternatives per group. Figure 4.1 shows the dependence between a group and a cluster for this example.

$$(4.11) \quad \Omega = \begin{bmatrix} A & 0 & 0 & 0 & 0 & 0 \\ 0 & B & 0 & 0 & 0 & 0 \\ 0 & 0 & B & 0 & 0 & 0 \\ 0 & 0 & 0 & C & 0 & 0 \\ 0 & 0 & 0 & 0 & A & 0 \\ 0 & 0 & 0 & 0 & 0 & C \end{bmatrix}$$

The letters A, B and C represent the covariance structure of a cluster. Same letters indicate an equal covariance structure for different clusters. Figure 4.1 illustrates the assignment of clusters to groups.

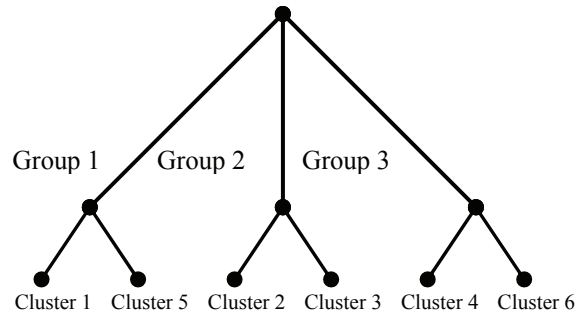


Fig. 4.1: Dependence between Clusters and Groups

If identical alternative-specific model-parameters, especially alternative-specific constants, can be assumed reasonably well for different clusters of the same group, an estimation of all model-parameters is feasible on a subset of all alternatives as described above.

Applying the concept of grouping in the logit-model is possible, however, serves only to estimating alternative-specific variables, as there are no different scale parameters due to independently and identically distributed random components in the utility function.

The main advantage of this approach does not only lie in the reduction of computational costs for very large choice sets, as many econometric software packages limit the maximum number of clusters and alternatives for nested logit estimations, but also in a better way of developing a more generally applicable choice model beyond the alternatives of the estimation dataset, i.e. in the context of scenario analysis.

5. Database and Grouping of Clusters: Kohonen's Self-Organizing Maps

Database

For the considered problem the German Air Traveller Survey 2003 (Berster et al. 2005) contains information on air journeys, especially:

- Journey origin
- Journey destination
- Origin airport
- Access mode to the airport
- Duration of journey
- Traveller information

According to the length and purpose of a journey different market segments are defined:

- Journeys to domestic destinations, subdivided into private and business trip purpose
- Journeys to European destinations for business trip purpose
- Journeys to European destinations for private short stay reasons up to 4 days

- Journeys to European destinations for holiday reasons for 5 days or longer
- Journeys to intercontinental destinations, subdivided into private and business trip purpose

For model estimation only travel demand originated in Germany is considered as air travellers are more free to choose airport and access mode to the airport. Table 5.1 shows the airports of the German Air Traveller Survey 2003 and the available access modes.

In this study alternatives are characterized by:

- Cost
- Time
- Frequency
- Service

	Car	Kiss and Ride	Rental Car	Taxi	Bus	Urban Transit	Train
Berlin	x	x	x	x	x	x	
Bremen	x	x	x	x		x	
Dortmund	x	x	x	x	x		
Dresden	x	x	x	x	x	x	
Düsseldorf	x	x	x	x	x	x	x
Erfurt	x	x	x	x	x		
Frankfurt a. M.	x	x	x	x	x	x	x
Frankfurt Hahn	x	x	x	x	x		
Friedrichshafen	x	x	x	x	x	x	
Hamburg	x	x	x	x	x		
Hannover	x	x	x	x	x	x	
Karlsruhe-Baden	x	x	x	x	x		
Köln/Bonn	x	x	x	x	x		
Leipzig/Halle	x	x	x	x	x		x
Lübeck	x	x	x	x	x		
München	x	x	x	x	x	x	
Münster/Osnabrück	x	x	x	x	x		
Niederrhein	x	x	x	x	x		
Nürnberg	x	x	x	x	x	x	
Paderborn/Lippstadt	x	x	x	x	x		
Saarbrücken	x	x	x	x	x		
Stuttgart	x	x	x	x	x	x	

Tab. 5.1: Airports and Available Access Modes

Only the access mode “car” includes parking at the airport for the duration of the journey. For “kiss and ride” the number of trips is double compared to all other access modes as the car is parked at the trip origin. The “taxi” alternative includes taxis and private bus services operating on demand only. The access mode “bus” contains scheduled public-transit buses. “Urban transit” and “train” are distinguished in terms of the tariff paid. If the tariff of the Deutsche Bahn applies, it is a train, otherwise it is an urban railway.

Access time and access costs are defined for the double trip length between the origin of the journey and origin airport, so that there is no need for an arbitrary allocation of any parking fees at the airport to either the outbound or the return trip. Access frequency is defined as the daily frequency. Its inverse equals the waiting time in the case of a uniformly distributed arrival time. The population density is chosen as a measure for the access time to public transport. The evaluation of the access quality from the access mode terminal to the air terminal is measured binary because of a lack of information about the chosen parking site and air terminal.

Variable (Abbreviation)	Definition
Access Cost (COST)	Cost in € per Person incl. Parking Fees, Double Trip Length
Access Time (TIME)	Time in Minutes, Double Trip Length
Waiting Time (WAIT)	Inverse of the Daily Frequency
Inverse of the Population Density (INVPD)	Inverse of Residents per km ²
Inverse of the Competition on a Direct Flight Connection (COMP)	Inverse of the Number of Alliances and Independent Airlines
Quality of Terminal Access (AAS)	binary (good/bad)
Existence of a Direct Flight Connection (DIRECT)	binary (good/bad)
Frequency of a Direct Flight Connection (DFREQ)	Number Flights per week
Existence of a Low-Cost Connection (LC)	binary (yes/no)
Frequency of a Low-Cost Connection (LCFREQ)	Number Low-Cost Flights per week
Existence of a Charter Flight Connection (CC)	binary (yes/no)
Frequency of a Charter Flight Connection (CCFREQ)	Number Charter Flights per week

Tab. 5.2: Definition of Explanatory Variables

The fare level of a direct flight connection to a specific destination is measured via its competition. It is assumed that a higher degree of competition indicates a lower fare level. For stop-over flights a maximum of competition is reached because of the many possible flights between an origin and a destination. The time advantage of a direct flight connection is measured via its existence, its quality is assessed by means of its weekly flight frequency. To consider price levels, low-cost- and charter-flights are separated additionally. By reasons of a lack of information exact air fares are not considered. Table 5.2 summarizes the explanatory variables and their definitions.

Attributes (Abbreviation)	Definition
Number of Domestic Low-Cost Flights (LCBRD)	Flights per Week
Number of Domestic Charter Flights (CCBRD)	Flights per Week
Number of Domestic Full Service Flights (LBRD)	Flights per Week
Number of European Low-Cost Flights (LCEUR)	Flights per Week
Number of European Charter Flights (CCEUR)	Flights per Week
Number of European Full Service Flights (LEUR)	Flights per Week
Number of Intercontinental Low-Cost Flights (LCINT)	Flights per Week
Number of Intercontinental Charter Flights (CCINT)	Flights per Week
Number of Intercontinental Full Service Flights (LINT)	Flights per Week
Number of Domestic Destinations (NUMBRD)	Number of Destinations
Number of European Destinations (NUMEUR)	Number of Destinations
Number of Intercontinental Destinations (NUMINT)	Number of Destinations

Tab. 5.3: Attributes for Airport Categorization

Airports are categorized from a demand-oriented point of view. Relevant attributes for a categorization concern the structure of their flight plan. Table 5.3 summarizes the attributes utilized and their definitions.

The German Air Traveller Survey contains no information about the explanatory variables and attributes described above, so they had to be gathered from different sources. Table 5.4 summarizes the data sources for the variables and attributes.

Variable	Data Sources
Choices	Berster et al. (2003)
Access Time and Access Cost for Car and Kiss and Ride	Kraftfahrt-Bundesamt (2003), ADAC (2003), Microsoft (2003), Deutsche Flughäfen (2003a), INVERMO (2005)
Access Time and Access Cost for Rental Car	Kraftfahrt-Bundesamt (2003), ADAC (2003), Microsoft (2003), Avis (2003), Sixt (2003), INVERMO (2005)
Access Time and Access Cost for Taxi	Taxi (2003), Private Busdienste (2003), Microsoft (2003), INVERMO (2005)
Access Time, Access Cost and Frequency for Bus and Urban Railway	Verkehrsverbände (2003), Die Bahn (2003a)
Access Time, Access Cost and Frequency for Train	Die Bahn (2003b)
Population Density	Statistisches Bundesamt (2003)
Quality of Terminal Access	Deutsche Flughäfen (2003b)
Flight Frequency and Level of Competition on a Direct Flight Connection and the Number of Destinations of an Airport	OAG (2003)

Tab. 5.4: Database

Due to a lack of information on alternative-specific characteristics some assumptions had to be made:

- Basis for the access modes car, kiss and ride and rental car is an averaged car.
- Parking fees are considered for the cheapest feasible category.
- For public access modes the connection with the shortest travel time for each mode of transport is chosen.
- Basis for access cost per person for private transport modes is an average of group size.

The general level of spatial abstraction are so called Spatial Planning Regions. Spatial Planning Regions with an airport and all Spatial Planning Regions neighbouring an airport region are broken down into counties to enable the definition of sensible bus and urban transit connections. Figure 5.1 illustrates the segmentation of Germany into Spatial Planning Regions and counties. Joining counties of the same colour constitute a spatial planning region.

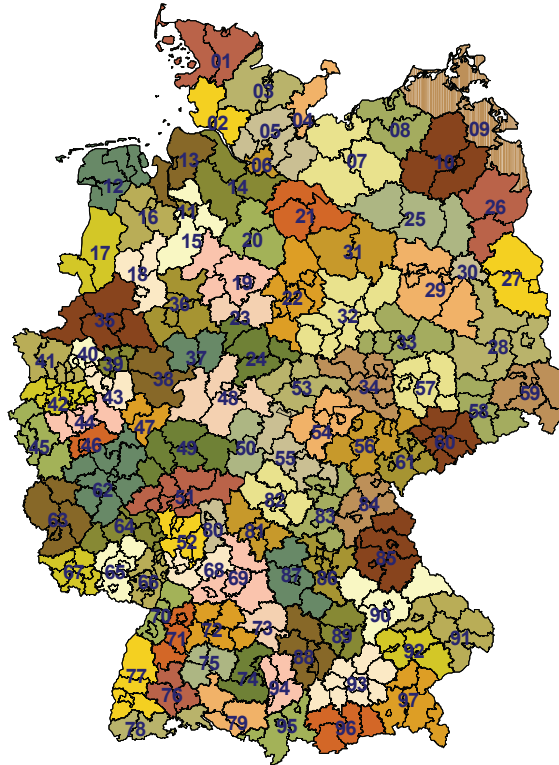


Fig. 5.1: Segmentation of Germany in Spatial Planning Regions and Districts

Grouping of Clusters

Clusters of the same group are characterized by an identical covariance-matrix and alternative-specific parameters, especially alternative-specific constants. As correlations among alternatives and alternative-specific constants represent unobserved factors, a grouping of clusters corresponds to an aggregation in terms of similarity of those unobserved factors. Airport and access mode choice is a two-dimensional choice problem, so that a categorization in respect of both dimensions is necessary, however, as the access mode choice is sufficiently general, only airports are to be categorized.

As mentioned above, airports have been categorized from a demand-oriented point of view by means of the general “picture” of flight services at an airport. The flight service of an airport is measured on the basis of the number of flights per destination type and flight type and the number of different destinations segmented by type of destination. Three types of destinations are defined:

- Domestic
- Europe
- Intercontinental

Flight types are divided into:

- Low-Cost
- Charter

- Full Service

Table 5.3 summarizes the relevant attributes for airport categorization.

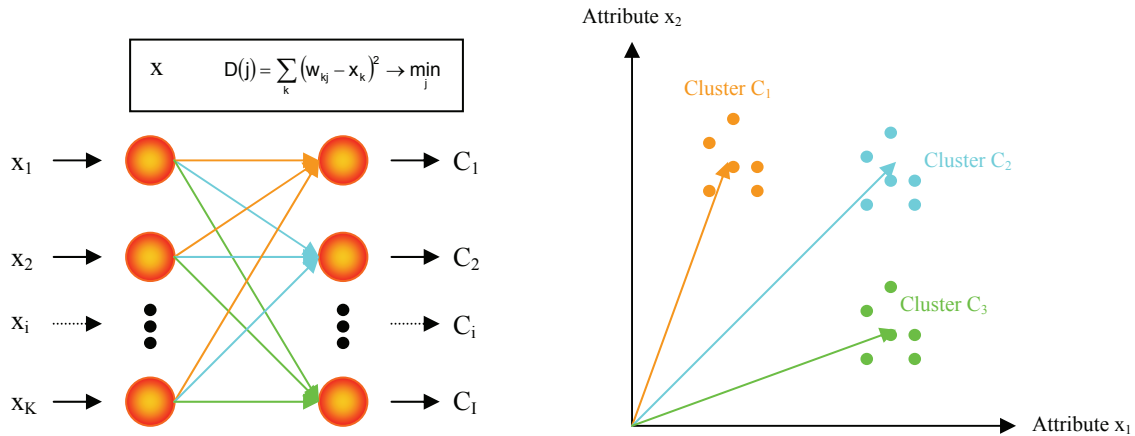


Fig. 5.2: Self-Organizing Map

Clusters are identified with Kohonen’s Self-Organizing Map (see i.e. Kohonen 2001, pp. 109ff.). Figure 5.2 is a schematic illustration of a Self-Organizing Map. The neurons are simple computational units connected by weighted edges. Computations in a neuron are performed according to a simple transfer function. Input neurons correspond to clustering attributes and output neurons represent the clusters. The transfer function of the input neurons is the identical function $f(x) = x$. The output neurons have a “winner-takes-all” transfer function. The neuron with the smallest distance between the input vector and its synaptic weight vector wins the competition and is activated. During learning of the self-organizing map the synaptic weight vector of the output neurons approach the corresponding cluster centroid as the right part of figure 5.2 illustrates.

Table 5.5 shows the parameters for an optimal cluster identification. The self-organizing map was not highly sensitive with regard to parameter variations.

Parameter	Value
Topology of output neurons	Linear
Measure of distance	Euclidean
Neighbourhood function	linear: $2 - 0.002 \cdot \text{Iteration}$
Learning rate	0.01
Number of iterations	10 000
Data normalization	yes, [-1; 1]
Number of input neurons	12
Number of output neurons	3

Tab. 5.5: Parameters of a Self-Organizing Map for Airport Categorization

Three airport categories have been identified. The output neurons are arranged in a linear grid and distances between input vectors and output neurons are measured Euclidean. A linear neighbourhood function is used and the neighbourhood contains all output neurons at the beginning of the learning process. It shrinks to 0 within 1000 iterations. The number of learning

iterations is 10000 and the learning rate is chosen rather small with 0.01. Each element of the input vector is normalized to the interval $[-1; 1]$.

Table 5.6 shows the synaptic weights for the trained self-organizing map. The colour of the columns equals the colour of the synaptic weights in figure 5.2.

Attributes	Airport		
	Category 1	Category 2	Category 3
LCBRD	0.054281	0.026181	-0.973566
CCBRD	0.63343	-0.23698	-0.902359
LBRD	0.820399	-0.16164	-0.810737
LCEUR	-0.814996	-0.248973	-0.717447
CCEUR	0.673964	0.145995	-0.811895
LEUR	0.767974	-0.596754	-0.967617
LCINT	-0.999997	-0.507511	-0.862715
CCINT	0.459986	-0.679604	-0.986041
LINT	0.128171	-0.975403	-0.999997
NUMBRD	0.810002	0.570222	-0.409338
NUMEUR	0.791409	-0.012681	-0.737397
NUMINT	0.314031	-0.817745	-0.991489

Tab. 5.6: Cluster Centroids of Airport Categories

Table 5.7 illustrates the assignment of airports of the German Air Traveller Survey 2003 to identified categories.

Category	Airport (IATA-Code)
AP 1	Frankfurt a. M. (FRA)
AP 1	München (MUC)
AP 2	Düsseldorf (DUS)
AP 2	Hamburg (HAM)
AP 2	Köln/Bonn (CGN)
AP 2	Stuttgart (STR)
AP 3	Bremen (BRE)
AP 3	Dortmund (DTM)
AP 3	Dresden (DRS)
AP 3	Erfurt (ERF)
AP 3	Frankfurt Hahn (HHN)
AP 3	Friedrichshafen (FDH)
AP 3	Hannover (HAJ)
AP 3	Karlsruhe/Baden (FKB)
AP 3	Leipzig/Halle (LEJ)
AP 3	Lübeck (LBC)
AP 3	Münster/Osnabrück (FMO)
AP 3	Niederrhein (NRN)
AP 3	Nürnberg (NUE)
AP 3	Paderborn/Lippstadt (PAD)
AP 3	Saarbrücken (SCN)

Tab. 5.7: Assignment of Airports to Categories

Because the three airports of Berlin were interviewed as one single airport, however, airport characteristics vary substantially, they are not usable for model estimation. For this reason they are not considered in the airport categorization.

Table 5.8 and table 5.9 illustrate some properties of the identified 3 airport categories in per cent and absolute values. The two highest values concerning the flight frequency and the number of different destinations are highlighted in colour.

	LCBRD	CCBRD	LBRD	LCEUR	CCEUR	LEUR	LCINT	CCINT	LINT	NUMBRD	NUMEUR	NUMINT
AP 1	3.18	0.43	20.39	0.87	5.83	55.81	0.00	1.24	12.25	8.31	60.27	31.42
AP 2	8.97	0.58	28.27	11.65	11.76	37.24	0.02	0.71	0.79	16.23	74.62	9.16
AP 3	1.29	0.86	39.22	32.57	15.57	10.05	0.02	0.42	0.00	19.94	78.90	1.16

Tab. 5.8: Structure of Flights per Airport Category (in %)

	LCBRD	CCBRD	LBRD	LCEUR	CCEUR	LEUR	LCINT	CCINT	LINT	NUMBRD	NUMEUR	NUMINT
AP 1	106	16	756	32	225	2138	0	49	517	19	144	83
AP 2	104	7	348	129	153	487	0	11	11	17	80	12
AP 3	3	1	80	47	25	39	0	0	0	6	22	1

Tab. 5.9: Structure of Flights per Airport Category (absolute)

Airports of the first category are hubs. They offer mainly full service flights. Their focus is mainly on European and intercontinental destinations. The number of domestic destinations is relatively low, but they are served with higher frequency. Hub airports offer the highest number of flights and destinations.

Airports of the second category serve mainly domestic and European destinations with full service flights. The share of European low-cost and charter flights is approximately equal but much smaller than the share of full service flights.

The structure of flights and destinations of airports of the third category is similar to those of the second category, but their focus is shifted more on domestic full service flights and European low-cost and charter traffic. These airports are the smallest in terms of number of flights and destinations.

Table 5.10 shows the standard deviation of the attributes for each attribute. Airports of the first category exhibit the greatest and airports of the third category the smallest heterogeneity.

	LCBRD	CCBRD	LBRD	LCEUR	CCEUR	LEUR	LCINT	CCINT	LINT	NUMBRD	NUMEUR	NUMINT
AP 1	96.00	3.50	75.00	8.50	37.50	279.00	0.00	18.00	396.00	1.00	16.50	43.00
AP 2	77.00	0.00	32.00	164.50	66.50	162.00	0.50	2.00	1.50	2.50	5.50	1.00
AP 3	0.00	0.00	68.00	9.00	21.00	29.50	0.00	0.00	0.00	2.50	10.50	0.00

Tab. 4.10: Standard Deviation of Attributes

6. Preparing the Data Set for Model Estimation and Estimation Results

Preparing the Data Set for Model Estimation

For model estimation the estimation data set is partitioned into several disjoint data subsets. Each data subset contains only a subset of the full set of airport-access mode alternatives, namely one airport of each category and its access modes. Each data subset includes observations of individuals, who have chosen one of the alternatives of the reduced alternative set. By a suitable definition of data subsets, it is possible to estimate a model with the full set of 7 access modes for all 3 airport categories. For this purpose, the inclusion of the airports Frankfurt/Main, Düsseldorf and Leipzig/Halle is necessary, as these are the only airports of their category with an access via train in 2003. The individual data subsets are merged into a single new estimation data set. The number of alternatives is reduced from 122 to 21. By weighting each observation the estimation data set is statistically representative. Figure 6.1 shows the definition of the data subsets. The nearest airport of each category is assigned to each data set marked in different colours. Every subset is named according to its airport of the third category.

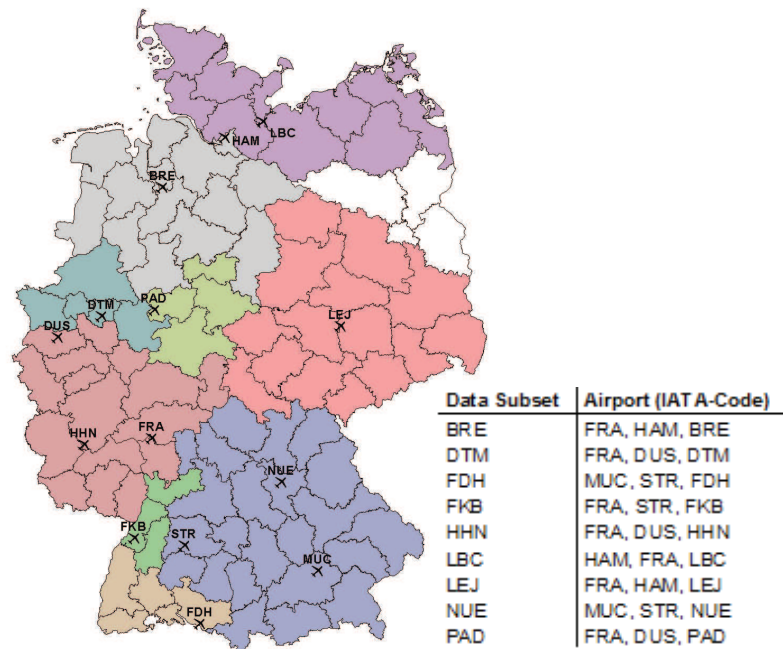


Fig. 6.1: Data Subsets and Assignment of Airports

General Model Estimation and Application

After selecting the airports and access modes for a specific application case, they are assigned to categories with the appropriate model parameters. The model can be applied to any number of airports. An application of the estimated model to other airports and airport/access mode combinations than those of the estimation data set is possible as a result of the grouping of clusters. Figure 6.2 summarizes the general process of the model estimation and its application. The next chapter deals with the estimation of the group-specific model parameters.

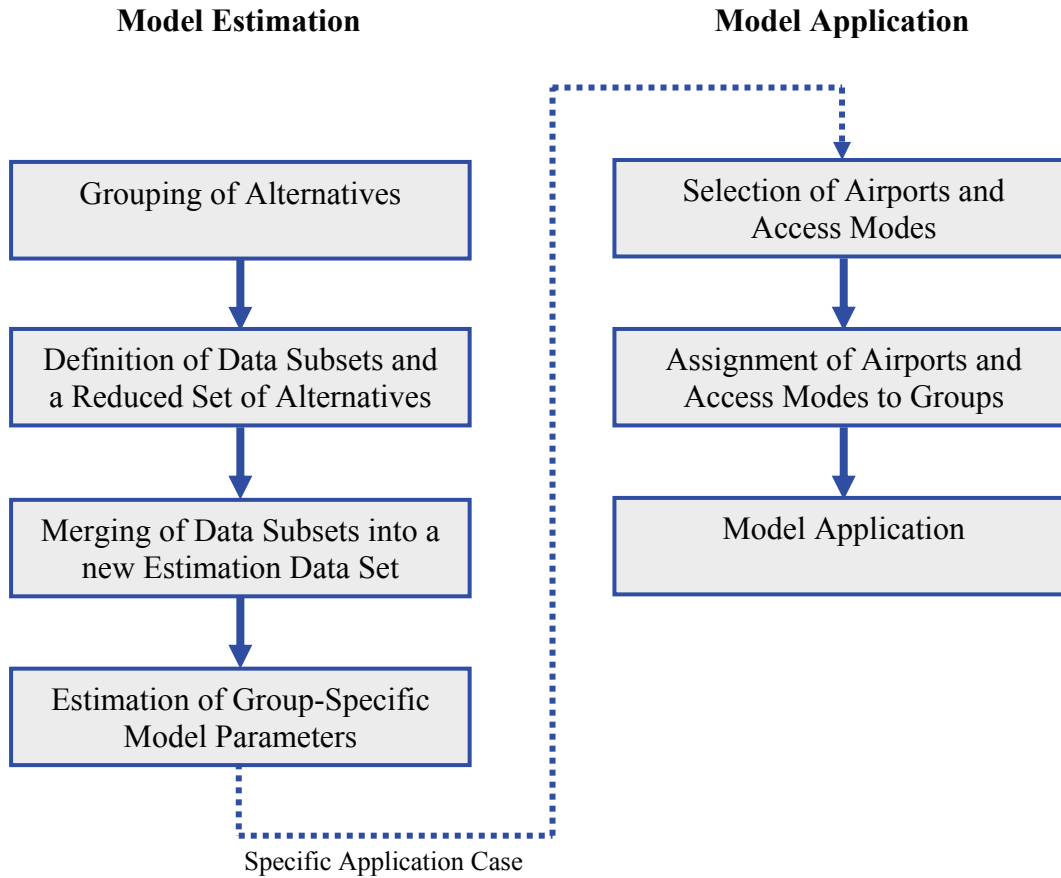


Fig. 6.2: Estimation and Application of Airport and Access Mode Choice Model

Estimation Results

Table 6.1 shows the reduced alternative set. Every alternative is being made up of an airport category and an access mode to the airport. Figure 6.3 illustrates the nesting structure. Every nest consists of one airport category at the top and 7 access modes below subdivided into private (PR) and public (PU) transport.

The deterministic part of the utility function is of a linear form:

$$(6.1) \quad V_i = \text{alt}_i + \sum_k b_k * x_{k,i}$$

where

- alt_i: Alternative-specific constant of alternative i
- b_k: Coefficient of attribute k
- x_{k, i}: Value of attribute k for alternative i

Alternative specific constants are denominated according to their alternative abbreviation. One alternative-specific constant has to be arbitrarily chosen and set to 0. In this study the constant of the alternative AP 3/Train is set to 0. Scale parameters are normalized on the lowest level of the nesting structure to a value of 1.

Alternative	Abbreviation
AP 1/Car	AP1CAR
AP 1/Kiss and Ride	AP1KAR
AP 1/Rental Car	AP1RC
AP 1/Taxi	AP1TAXI
AP 1/Bus	AP1BUS
AP 1/Urban Railway	AP1UR
AP 1/Train	AP1TR
AP 2/Car	AP2CAR
AP 2/Kiss and Ride	AP2KAR
AP 2/Rental Car	AP2RC
AP 2/Taxi	AP2TAXI
AP 2/Bus	AP2BUS
AP 2/Urban Railway	AP2UR
AP 2/Train	AP2TR
AP 3/Car	AP3CAR
AP 3/Kiss and Ride	AP3KAR
AP 3/Rental Car	AP3RC
AP 3/Taxi	AP3TAXI
AP 3/Bus	AP3BUS
AP 3/Urban Railway	AP3UR
AP 3/Train	AP3TR

Tab. 6.1: Reduced Alternative Set

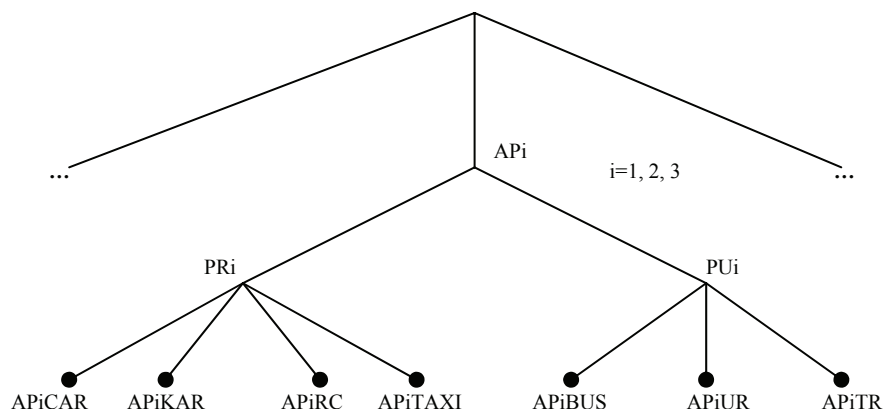


Fig. 6.3: Nesting Structure

Model parameters are estimated via maximum-likelihood and the BFGS-algorithm is used for numerical optimization (see i.e. Greene 2003, pp. 938ff.). The covariance matrix of the estimated parameters is computed by means of the BHHH-estimator (see i.e. Train 2003, pp. 196ff.). The significance of model parameters is evaluated by the t-ratio and p-value. The goodness-of-fit is assessed by means of the pseudo- R^2 . Benchmark is a model without any variables (R^2_{null}) and a market share model (R^2_{const}).

Tables 6.2 - 6.8 show the estimated model parameters, t-ratios and p-value for the 7 market segments as defined above. Alternative-specific constants and scale parameters are separated by a dashed line.

Variable	Coefficient	Standard Deviation	t-ratio	p-value
COST	-0.0263035	7.47E-05	-352.091	2.89E-15
TIME	-0.0081889	3.65E-05	-224.172	2.89E-15
WAIT	-28.8061	0.0521136	-552.755	2.89E-15
INVPD	-187.86	2.74598	-68.4127	2.89E-15
COMP	-0.158635	0.0204772	-7.74689	9.33E-15
AAS	0.920627	0.0109263	84.2575	2.89E-15
DIRECT	2.29637	0.0252162	91.0672	2.89E-15
DFREQ	0.00682913	0.00016972	40.2374	2.89E-15
AP1CAR	-0.89308	0.0299652	-29.8039	2.89E-15
AP1KAR	-0.935515	0.0312753	-29.9123	2.89E-15
AP1RC	-4.1011	0.0360866	-113.646	2.89E-15
AP1TAXI	-1.66527	0.0317124	-52.5116	2.89E-15
AP1BUS	-0.0749869	0.0448874	-1.67055	0.0948097
AP1UR	0.671661	0.0431181	15.5772	2.89E-15
AP1TR	-0.289548	0.0422378	-6.85519	7.12E-12
AP2CAR	-1.42599	0.0497169	-28.6823	2.89E-15
AP2KAR	-0.969869	0.0508523	-19.0723	2.89E-15
AP2RC	-4.31713	0.0554302	-77.884	2.89E-15
AP2TAXI	-1.66024	0.0511273	-32.4727	2.89E-15
AP2BUS	-2.0108	0.0755529	-26.6145	2.89E-15
AP2UR	-0.561955	0.0722517	-7.77775	7.33E-15
AP2TR	-0.628393	0.0717579	-8.75712	2.89E-15
AP3CAR	-2.32656	0.0266369	-87.3434	2.89E-15
AP3KAR	-2.28413	0.0265816	-85.9291	2.89E-15
AP3RC	-4.56071	0.0611955	-74.527	2.89E-15
AP3TAXI	-3.28287	0.0273826	-119.889	2.89E-15
AP3BUS	-5.74305	0.150649	-38.1219	2.89E-15
AP3UR	-2.56922	0.0464991	-55.2532	2.89E-15
PR1	1.07092	0.0100494	106.566	2.89E-15
PU1	0.745385	0.00715937	104.113	2.89E-15
PR2	0.492518	0.00595683	82.6813	2.89E-15
PU2	0.390636	0.00358923	108.835	2.89E-15
PR3	0.817955	0.0174313	46.9245	2.89E-15
PU3	0.428619	0.0104805	40.8967	2.89E-15
AP1	1.81029	0.0161987	111.755	2.89E-15
AP2	2.10174	0.0240208	87.4967	2.89E-15
AP3	2.35248	0.0467621	50.3075	2.89E-15
			R2(null)	57.41%
			R2(const)	43.82%

Tab. 6.2: Nested Logit-Model for Domestic Private Travel (BRD P)

Variable	Coefficient	Standard Deviation	t-ratio	p-value
COST	-0.0204609	1.08E-05	-1900.36	2.89E-15
TIME	-0.0152572	2.79E-05	-546.331	2.89E-15
WAIT	-18.935	0.0524438	-361.053	2.89E-15
INVPD	-21.8829	1.08584	-20.1529	2.89E-15
AAS	1.12781	0.00482371	233.805	2.89E-15
DIRECT	3.64119	0.0137238	265.318	2.89E-15
DFREQ	0.00601159	8.99E-05	66.8909	2.89E-15
AP1CAR	0.821324	0.0249217	32.9562	2.89E-15
AP1KAR	-0.205879	0.0254374	-8.09355	2.89E-15
AP1RC	-1.86138	0.0256406	-72.5952	2.89E-15
AP1TAXI	-0.3315	0.0251872	-13.1615	2.89E-15
AP1BUS	-1.47598	0.0298635	-49.4241	2.89E-15
AP1UR	-0.361618	0.0277497	-13.0315	2.89E-15
AP1TR	-1.53084	0.0277493	-55.1667	2.89E-15
AP2CAR	0.448667	0.0240099	18.6868	2.89E-15
AP2KAR	-1.03968	0.0243685	-42.6648	2.89E-15
AP2RC	-1.5527	0.024637	-63.023	2.89E-15
AP2TAXI	-0.475198	0.0243418	-19.5219	2.89E-15
AP2BUS	-1.74549	0.0306954	-56.8649	2.89E-15
AP2UR	-0.554791	0.0284689	-19.4876	2.89E-15
AP2TR	-0.771201	0.0283786	-27.1755	2.89E-15
AP3CAR	-0.625039	0.0221069	-28.2735	2.89E-15
AP3KAR	-1.73868	0.0222633	-78.0963	2.89E-15
AP3RC	-2.23438	0.025964	-86.0567	2.89E-15
AP3TAXI	-1.82039	0.0224969	-80.9173	2.89E-15
AP3BUS	-3.74058	0.0331825	-112.728	2.89E-15
AP3UR	-2.3761	0.0182418	-130.256	2.89E-15
PR1	1.02375	0.00561628	182.283	2.89E-15
PU1	0.978059	0.00470008	208.094	2.89E-15
PR2	1.00829	0.0054788	184.035	2.89E-15
PU2	0.992109	0.00421163	235.564	2.89E-15
PR3	1.00988	0.011452	88.1839	2.89E-15
PU3	0.999286	0.00799378	125.008	2.89E-15
AP1	1.01119	0.00545905	185.231	2.89E-15
AP2	1.00887	0.00552003	182.766	2.89E-15
AP3	1.01164	0.011702	86.45	2.89E-15
			R2(null)	54.10%
			R2(const)	40.47%

Tab. 6.3: Nested Logit-Model for Domestic Business Travel (BRD B)

Variable	Coefficient	Standard Deviation	t-ratio	p-value
COST	-0.0199987	6.35E-05	-315.076	2.89E-15
TIME	-0.0061063	3.08E-05	-197.958	2.89E-15
WAIT	-8.33078	0.101522	-82.0589	2.89E-15
INVPD	-215.876	3.45959	-62.3992	2.89E-15
COMP	-1.22176	0.0143873	-84.9193	2.89E-15
AAS	0.20336	0.0105667	19.2453	2.89E-15
DIRECT	3.63327	0.0204966	177.262	2.89E-15
DFREQ	0.0104684	0.00020263	51.6641	2.89E-15
LC	0.0863075	0.0103855	8.31037	2.89E-15
LCFREQ	0.0631856	0.00061005	103.575	2.89E-15
AP1CAR	-0.498688	0.0666011	-7.48768	7.02E-14
AP1KAR	0.318789	0.0674283	4.72781	2.27E-06
AP1RC	-3.33871	0.0706322	-47.269	2.89E-15
AP1TAXI	-0.435522	0.06765	-6.43788	1.21E-10
AP1BUS	0.210693	0.0906689	2.32377	0.020138
AP1UR	1.50982	0.0897749	16.8179	2.89E-15
AP1TR	0.122875	0.0904775	1.35807	0.174442
AP2CAR	-0.303182	0.0680182	-4.45737	8.30E-06
AP2KAR	0.278229	0.0686423	4.05333	5.05E-05
AP2RC	-3.171	0.0716133	-44.2795	2.89E-15
AP2TAXI	-0.0993231	0.0688372	-1.44287	0.149057
AP2BUS	0.65932	0.0990006	6.65975	2.74E-11
AP2UR	1.27978	0.0981204	13.043	2.89E-15
AP2TR	0.98543	0.0983198	10.0227	2.89E-15
AP3CAR	0.40639	0.0634284	6.40707	1.48E-10
AP3KAR	0.538874	0.0643244	8.37744	2.89E-15
AP3RC	-3.70737	0.0712379	-52.0421	2.89E-15
AP3TAXI	-0.131292	0.0646538	-2.0307	0.0422853
AP3BUS	0.528475	0.127801	4.13513	3.55E-05
AP3UR	0.71755	0.126304	5.68113	1.34E-08
PR1	0.764486	0.0087763	87.1079	2.89E-15
PU1	0.593257	0.00626677	94.6671	2.89E-15
PR2	0.767123	0.00715629	107.196	2.89E-15
PU2	0.543582	0.00583578	93.1464	2.89E-15
PR3	0.821821	0.00996985	82.4306	2.89E-15
PU3	0.395656	0.00806925	49.0325	2.89E-15
AP1	1.80601	0.0199672	90.4489	2.89E-15
AP2	1.76862	0.0162451	108.871	2.89E-15
AP3	1.74828	0.0226854	77.0664	2.89E-15
			R2(null)	52.40%
			R2(const)	41.94%

Tab. 6.4: Nested Logit-Model for European Private Short Stay Travel (EUR S)

Variable	Coefficient	Standard Deviation	t-ratio	p-value
COST	-0.0173617	2.08E-05	-835.813	2.89E-15
TIME	-0.00857067	1.13E-05	-759.386	2.89E-15
WAIT	-4.40982	0.0215587	-204.549	2.89E-15
INVPD	-235.641	1.1008	-214.064	2.89E-15
COMP	-1.13258	0.00417551	-271.244	2.89E-15
AAS	0.46823	0.00313156	149.52	2.89E-15
DIRECT	3.31697	0.00579373	572.511	2.89E-15
DFREQ	0.0153856	7.51E-05	204.84	2.89E-15
LC	0.563633	0.00232754	242.158	2.89E-15
AP1CAR	-0.783801	0.0163485	-47.9432	2.89E-15
AP1KAR	1.19964	0.0166094	72.2267	2.89E-15
AP1RC	-3.24672	0.0176445	-184.008	2.89E-15
AP1TAXI	-0.153202	0.0166854	-9.1818	2.89E-15
AP1BUS	0.46742	0.0277141	16.8658	2.89E-15
AP1UR	1.96562	0.0272271	72.1935	2.89E-15
AP1TR	0.850638	0.027015	31.4876	2.89E-15
AP2CAR	-1.02568	0.0149567	-68.5768	2.89E-15
AP2KAR	0.903728	0.0152148	59.398	2.89E-15
AP2RC	-3.10541	0.0159476	-194.726	2.89E-15
AP2TAXI	-0.187646	0.0152637	-12.2936	2.89E-15
AP2BUS	-1.32489	0.0236498	-56.0211	2.89E-15
AP2UR	-0.154352	0.0227366	-6.7887	1.13E-11
AP2TR	-0.359231	0.0226828	-15.8371	2.89E-15
AP3CAR	-0.377357	0.0132672	-28.4428	2.89E-15
AP3KAR	0.315622	0.0135114	23.3597	2.89E-15
AP3RC	-4.37193	0.0182017	-240.194	2.89E-15
AP3TAXI	-0.628438	0.013613	-46.1644	2.89E-15
AP3BUS	-1.77275	0.0123277	-143.803	2.89E-15
AP3UR	-1.44559	0.00937011	-154.277	2.89E-15
PR1	0.61189	0.00189196	323.417	2.89E-15
PU1	0.3847	0.00150032	256.412	2.89E-15
PR2	0.570138	0.0018957	300.753	2.89E-15
PU2	0.437515	0.0014318	305.569	2.89E-15
PR3	0.610065	0.00342601	178.069	2.89E-15
PU3	0.551239	0.00290076	190.033	2.89E-15
AP1	1.65075	0.0049926	330.639	2.89E-15
AP2	1.92646	0.00606395	317.691	2.89E-15
AP3	1.99236	0.0108685	183.315	2.89E-15
			R2(null)	52.29%
			R2(const)	38.22%

Tab. 6.5: Nested Logit-Model for European Holiday Travel (EUR H)

Variable	Coefficient	Standard Deviation	t-ratio	p-value
COST	-0.0216885	2.66E-05	-816.759	2.89E-15
TIME	-0.00795957	1.99E-05	-399.792	2.89E-15
WAIT	-9.94709	0.0352918	-281.853	2.89E-15
COMP	-0.182127	0.00715126	-25.4678	2.89E-15
AAS	0.504623	0.00472046	106.901	2.89E-15
DIRECT	1.43564	0.00850917	168.717	2.89E-15
DFREQ	0.0177437	0.00010425	170.208	2.89E-15
LC	0.275153	0.00504501	54.5396	2.89E-15
LCFREQ	0.0761092	0.00037252	204.307	2.89E-15
AP1CAR	0.72216	0.0296247	24.3769	2.89E-15
AP1KAR	0.233292	0.0300636	7.75995	8.44E-15
AP1RC	-0.661771	0.0301596	-21.9423	2.89E-15
AP1TAXI	0.750386	0.030056	24.9663	2.89E-15
AP1BUS	-0.436805	0.0640814	-6.8164	9.33E-12
AP1UR	1.33854	0.063386	21.1173	2.89E-15
AP1TR	-0.0557889	0.0635451	-0.877942	0.379975
AP2CAR	0.393121	0.0291205	13.4998	2.89E-15
AP2KAR	-0.260475	0.0294758	-8.83691	2.89E-15
AP2RC	-0.671533	0.0296515	-22.6475	2.89E-15
AP2TAXI	0.415442	0.029515	14.0756	2.89E-15
AP2BUS	-1.76693	0.0359288	-49.1786	2.89E-15
AP2UR	-0.855622	0.0343798	-24.8873	2.89E-15
AP2TR	-0.848627	0.0343025	-24.7395	2.89E-15
AP3CAR	-0.300282	0.0223921	-13.4102	2.89E-15
AP3KAR	-0.698722	0.0227567	-30.7041	2.89E-15
AP3RC	-1.05248	0.0239982	-43.8567	2.89E-15
AP3TAXI	-0.609462	0.0226451	-26.9137	2.89E-15
AP3BUS	-2.26991	0.0401428	-56.5459	2.89E-15
AP3UR	-1.49274	0.0246333	-60.5983	2.89E-15
PR1	0.808397	0.00380609	212.396	2.89E-15
PU1	0.386155	0.00263013	146.82	2.89E-15
PR2	0.783306	0.00371673	210.751	2.89E-15
PU2	0.708662	0.00269609	262.848	2.89E-15
PR3	0.937914	0.0123815	75.7514	2.89E-15
PU3	0.805435	0.0108905	73.9574	2.89E-15
AP1	1.61072	0.00814231	197.821	2.89E-15
AP2	1.67197	0.0073826	226.474	2.89E-15
AP3	1.77295	0.0232875	76.1333	2.89E-15
			R2(null)	48.58%
			R2(const)	35.96%

Tab. 6.6: Nested Logit-Model for European Business Travel (EUR B)

Variable	Coefficient	Standard Deviation	t-ratio	p-value
COST	-0.0138527	2.31E-05	-600.751	2.89E-15
TIME	-0.00541014	1.71E-05	-316.804	2.89E-15
WAIT	-18.7546	7.06E-05	-265589	2.89E-15
INVPD	-25.6109	1.1622	-22.0365	2.89E-15
AAS	0.840462	0.00491188	171.108	2.89E-15
DIRECT	1.85847	0.00516084	360.109	2.89E-15
AP1CAR	-1.67803	0.0043471	-386.012	2.89E-15
AP1KAR	-0.675255	0.00641839	-105.206	2.89E-15
AP1RC	-4.52249	0.0104444	-433.006	2.89E-15
AP1TAXI	-2.24118	0.00699765	-320.276	2.89E-15
AP1BUS	-2.76277	0.0150412	-183.68	2.89E-15
AP1UR	-0.567135	0.00827126	-68.5669	2.89E-15
AP1TR	-0.628369	0.00965685	-65.0698	2.89E-15
AP2CAR	-2.55593	0.00563923	-453.241	2.89E-15
AP2KAR	-0.781095	0.0063191	-123.609	2.89E-15
AP2RC	-5.48899	0.0179425	-305.921	2.89E-15
AP2TAXI	-1.9829	0.00663292	-298.949	2.89E-15
AP2BUS	-1.93506	0.0254801	-75.9441	2.89E-15
AP2UR	-1.75212	0.0212681	-82.3822	2.89E-15
AP2TR	-48.5491	8.61E+10	-5.64E-10	1
AP3CAR	-2.09268	0.00426141	-491.077	2.89E-15
AP3KAR	-0.470189	0.00543666	-86.485	2.89E-15
AP3RC	-3.52639	0.00769235	-458.428	2.89E-15
AP3TAXI	-1.13561	0.00554722	-204.716	2.89E-15
AP3BUS	-1.95589	0.00957575	-204.254	2.89E-15
AP3UR	-0.418627	0.00539374	-77.6136	2.89E-15
PR1	1.13266	0.00734164	154.278	2.89E-15
PU1	0.983045	0.00675649	145.496	2.89E-15
PR2	1.06067	0.0131951	80.3838	2.89E-15
PU2	0.927296	0.0110789	83.6991	2.89E-15
PR3	0.813943	0.00281214	289.44	2.89E-15
PU3	0.137029	0.00165706	82.6942	2.89E-15
AP1	1.10489	0.00678013	162.959	2.89E-15
AP2	1.19742	0.0144386	82.9317	2.89E-15
AP3	1.23031	0.00474654	259.201	2.89E-15
			R2(null)	48.89%
			R2(const)	32.86%

Tab. 6.7: Nested Logit-Model for Intercontinental Private Travel (INT P)

Variable	Coefficient	Standard Deviation	t-ratio	p-value
COST	-0.00936472	1.59E-05	-589.728	2.89E-15
TIME	-0.00535887	3.15E-05	-170.349	2.89E-15
WAIT	-35.7591	0.0277649	-1287.92	2.89E-15
INVPD	-32.2589	2.8701	-11.2397	2.89E-15
AAS	0.382595	0.012889	29.6838	2.89E-15
DIRECT	0.439344	0.00441956	99.4091	2.89E-15
AP1CAR	-0.059388	0.0754859	-0.786742	0.431433
AP1KAR	1.17409	0.0772982	15.1891	2.89E-15
AP1RC	-0.823745	0.0767846	-10.728	2.89E-15
AP1TAXI	1.05928	0.076873	13.7796	2.89E-15
AP1BUS	2.01162	0.233108	8.62957	2.89E-15
AP1UR	2.67192	0.232672	11.4836	2.89E-15
AP1TR	1.3506	0.232603	5.80647	6.38E-09
AP2CAR	-1.04963	0.102518	-10.2385	2.89E-15
AP2KAR	0.0612584	0.103547	0.591601	0.554118
AP2RC	-2.32606	0.103863	-22.3954	2.89E-15
AP2TAXI	-0.229266	0.103265	-2.22016	0.0264076
AP2BUS	-1.54098	0.174892	-8.81106	2.89E-15
AP2UR	-0.460972	0.169567	-2.71853	0.00655733
AP2TR	-0.625187	0.1686	-3.70811	0.00020881
AP3CAR	-2.00291	0.098986	-20.2342	2.89E-15
AP3KAR	-1.11849	0.0987287	-11.329	2.89E-15
AP3RC	-3.06497	0.10039	-30.5306	2.89E-15
AP3TAXI	-1.18451	0.0991565	-11.9459	2.89E-15
AP3BUS	-3.09884	0.0707277	-43.8137	2.89E-15
AP3UR	-1.9117	0.0408988	-46.7422	2.89E-15
PR1	1.03073	0.00684748	150.526	2.89E-15
PU1	0.32899	0.00387138	84.9801	2.89E-15
PR2	1.3532	0.0265898	50.8917	2.89E-15
PU2	0.832438	0.0120304	69.1943	2.89E-15
PR3	0.91783	0.0320818	28.6091	2.89E-15
PU3	0.718249	0.0410799	17.4842	2.89E-15
AP1	2.10553	0.0154688	136.115	2.89E-15
AP2	1.16102	0.0217542	53.3699	2.89E-15
AP3	1.73837	0.0551256	31.5348	2.89E-15
			R2(null)	47.46%
			R2(const)	28.30%

Tab. 6.8: Nested Logit-Model for Intercontinental Business Travel (INT B)

The generic variables are highly significant. Table 6.9 summarizes the generic variables, scale parameters, goodness-of-fit-measures and the likelihood-ratio test statistic split up by market segment. The restricted version of the likelihood-ratio test is a logit-model. The last row in table 6.8 shows the critical values for a significance level of 0.5%. Due to an arbitrary fixation of the scale parameters on the lowest level of the nesting structure, model parameters can only be interpreted comparatively.

Variable	BRD P	BRD B	EUR S	EUR H	EUR B	INT P	INT B
COST	-0.0263035	-0.0204609	-0.0199987	-0.0173617	-0.0216885	-0.0138527	-0.00936472
TIME	-0.0081889	-0.0152572	-0.0061063	-0.00857067	-0.00795957	-0.00541014	-0.00535887
WAIT	-28.8061	-18.935	-8.33078	-4.40982	-9.94709	-18.7546	-35.7591
INVPD	-187.86	-21.8829	-215.876	-235.641	x	-25.6109	-32.2589
COMP	-0.158635	x	-1.22176	-1.13258	-0.182127	x	x
AAS	0.920627	1.12781	0.20336	0.46823	0.504623	0.840462	0.382595
DIRECT	2.29637	3.64119	3.63327	3.31697	1.43564	1.85847	0.439344
DFREQ	0.00682913	0.00601159	0.0104684	0.0153856	0.0177437	x	x
LC	x	x	0.0863075	0.563633	0.275153	x	x
LCFREQ	x	x	0.0631856	x	0.0761092	x	x
PR1	1.07092	1.02375	0.764486	0.61189	0.808397	1.13266	1.03073
PU1	0.745385	0.978059	0.593257	0.3847	0.386155	0.983045	0.32899
PR2	0.492518	1.00829	0.767123	0.570138	0.783306	1.06067	1.3532
PU2	0.390636	0.992109	0.543582	0.437515	0.708662	0.927296	0.832438
PR3	0.817955	1.00988	0.821821	0.610065	0.937914	0.813943	0.91783
PU3	0.428619	0.999286	0.395656	0.551239	0.805435	0.137029	0.718249
AP1	1.81029	1.01119	1.80601	1.65075	1.61072	1.10489	2.10553
AP2	2.10174	1.00887	1.76862	1.92646	1.67197	1.19742	1.16102
AP3	2.35248	1.01164	1.74828	1.99236	1.77295	1.23031	1.73837
pseudo-R ² (null) in %	57.41	54.10	52.40	52.29	48.58	48.89	47.46
pseudo-R ² (const) in %	43.82	40.47	41.94	38.22	35.96	32.86	28.30
LR (MNL)	82414	8740	43774	349740	311756	599974	131576
$\alpha=0.5\%$	25.19	23.59	23.59	23.59	23.59	23.59	23.59

Tab. 6.9: Overview Estimation Results per Market Segment

In many cases the scale parameters do not decline with ascending level in the nesting structure as necessary in a nested logit-model. This is not exceptional in empirical studies (vgl. z.B. Ben-Akiva et al. 1985, S. 316; Hensher et al. 2002, S. 8f; Pels et al. 2003, S. 77). According to the likelihood-ratio test statistic, the nested logit approach is highly significant as compared with a simpler logit model illustrated by table 5.9. Table 5.10 shows the t-ratio for the scale parameters under the null-hypothesis that the true values are 1, so that the nested logit-model equals a logit approach.

Scale Parameter	t-ratio						
	BRD P	BRD B	EUR S	EUR H	EUR B	INT P	INT B
PR1	7.057	4.229	-26.835	-205.136	-50.341	18.070	4.488
PU1	-35.564	-4.668	-64.905	-410.113	-233.390	-2.509	-173.326
PR2	-85.193	1.513	-32.542	-226.756	-58.302	4.598	13.283
PU2	-169.776	-1.874	-78.210	-392.852	-108.059	-6.562	-13.928
PR3	-10.444	0.863	-17.872	-113.816	-5.014	-66.162	-2.561
PU3	-54.518	-0.089	-74.895	-154.705	-17.866	-520.784	-6.859
AP1	50.022	2.050	40.367	130.343	75.006	15.470	71.468
AP2	45.866	1.607	47.314	152.782	91.021	13.673	7.402
AP3	28.923	0.995	32.985	91.306	33.192	48.522	13.394

Critical t-ratio ($\alpha=5\%$): 1.96

Tab. 6.10: Overview of t-ratios for the Scale Parameters

Apart from some scale parameters of the model for the market segment of domestic business travel, the null-hypothesis can be rejected at a significance level of 5%. Parameters, whose null-hypothesis can be rejected at the 5% significance level, are highlighted.

The increase of the values of the scale parameters from the middle to the top level of the nesting structure is rather strong. This indicates that an increase in utility of one access mode to an airport increases the choice probability of that airport independently from the access mode. The underlying correlation structure among the alternatives doesn't fully comply with a nested logit approach. However, the nested logit model can be considered as a better approximation to the problem structure than a simpler logit approach. Both the likelihood ratio test and the t-test show the significance of the nested logit approach as compared with a logit model.

7. Conclusions and further research needs

This paper is subdivided into an empirical and a model estimation part. To better understand the real world behind the abstract model world, the first part of the paper describes the functional airport and model access situation in Germany. It has been shown that passengers of any origin region consider several airports – on average eight – for their air journey itinerary and that the choice of the departure airport is influenced by the spectrum and quality of access modes. Intercity trains with direct services to and from airport terminals have a positive influence on both the catchment area of airports and the modal split by favouring these fast “feeder” trains. The paper then presents an airport and access mode choice model based on a nested logit approach. Due to a categorization of airports this model is applicable to any number of airports and combinations of airports and access modes and is suitable especially for the evaluation of future scenarios of access mode and airport combinations not existing today.

Model quality is evaluated via the pseudo- R^2 . Values of about 0.3 of the pseudo- R^2 with a market share model as benchmark already indicate a decent model quality (Hensher et al. 2005, p. 338f.). It was possible to estimate models of at least decent quality for all market segments. The estimated nested logit models for domestic private and business travel and the two European private travel market segments show a very good fit while the models of European business travel and the two intercontinental travel market segments are only of decent quality.

Many of the scale parameters do not decline with ascending level in the nesting structure. This indicates a correlation structure among the alternatives beyond the assumptions of a nested logit approach, so that the estimated models can be considered as an approximation only.

In his doctoral thesis one of the authors, Marc Gelhausen, is developing a new type of discrete choice model based on the General Extreme Value Model framework and on artificial feedforward neural networks, which allows to model more complex correlations among the alternatives and to apply of the concept of grouping as described above. The usage of artificial neural networks enables to model a utility function of arbitrary functional form without a parametric specification. Usually a linear utility function is used for reasons of simplicity and a lack of knowledge about the functional form of the utility function. This approach might deliver interesting results especially for the market segments of European business travel and Intercontinental travel, as the nested logit models for these market segments are of decent quality only.

References

- ADAC (2003). ADAC Fahrzeugtechnik Sonderdruck – ADAC Autokosten 2003.
- Avis (2003). Mietwagenpreise 2003. <http://www.avis.de>.
- Ben-Akiva, M.; Lerman, S. (1985). *Discrete Choice Analysis: Theory and Applications to Travel Demand*. Cambridge.
- Berster, P.; Focke, H.; Pabst, H.; Wilken, D. (2005). *Gemeinsame Auswertung der ADV-Fluggastbefragung 2003*. Unveröffentlicht, Köln 2005.
- Deutsche Flughäfen (2003a). *Parkpreise an 22 ausgesuchten deutschen Flughäfen*. Verschiedene Orte.
- Deutsche Flughäfen (2003b). *Lagepläne von 22 ausgesuchten deutschen Flughäfen*. Verschiedene Orte.
- Die Bahn (2003a). *Elektronischer Fahrplan Regionalverbindungen*.
- Die Bahn (2003b). *DB ReiseService*.
- Domencich, T. A.; McFadden, D. (1975). *Urban Travel Demand – A Behavioral Analysis*. New York.
- Greene, W. H. (2003). *Econometric Analysis*. New Jersey.
- Hensher, D. A.; Greene, W. H. (2002). *Specification and Estimation of the Nested Logit Model: Alternative Normalisations*. *Transportation Research B* Vol. 36 (2002), pp. 1 – 17.
- INVERMO (2005). *Sonderauswertung zum Projekt INVERMO*, Institut für Verkehrswesen, Universität Karlsruhe. Karlsruhe.
- Kohonen, T. (2001). *Self-Organizing Maps*. Berlin.
- Kraftfahrt-Bundesamt (2003). *Statistische Mitteilungen – Band 2: Kraftfahrzeuge*. Flensburg.
- Maier, G.; Weiss, P. (1990). *Modelle diskreter Entscheidungen – Theorie und Anwendung in den Sozial- und Wirtschaftswissenschaften*. Wien.
- McFadden, D. (1974). *Conditional Logit Analysis of Qualitative Choice Behavior*. In: Zarembka, P.: *Frontiers in Econometrics*, pp. 105 – 142. New York.
- McFadden, D. (1978). *Modelling the Choice of Residential Location*. In: Karlqvist, A.; Lundqvist, L.; Snickars, F.; Weibull, J.: *Spatial Interaction Theory and Planning Models*, pp. 75 – 96. Amsterdam.
- Microsoft (2003). *Microsoft AutoRoute 2003*.
- Official Airline Guide (OAG) (2003). *Market Analysis*. Dunstable.
- Ortuzar, J. de D.; Willumsen, L. G. (2001). *Modelling Transport*. London.
- Pels, E.; Nijkamp, P.; Rietveld, P. (2003). *Access to and Competition between Airports: A Case Study for the San Francisco Bay Area*. *Transportation Research A* Vol. 37 (2003), pp. 71 – 83.
- Private Busdienste (2003). *Preislisten privater Busdienste*. Verschiedene Orte.
- Sixt (2003). *Mietwagenpreise 2003*. <http://www.e-sixt.de>.
- Statistisches Bundesamt (2003). *Bevölkerungsstatistiken des Statistischen Bundesamtes*. Wiesbaden.
- Taxi (2003). *Taxi-Tarifumfrage*. <http://www.ventil-taximagazin.com>.
- Train, K. E. (2003). *Discrete Choice Methods with Simulation*. Cambridge.
- Verkehrsverbünde (2003). *Fahr- und Tarifpläne 2003*. Verschiedene Orte.
- Wilken, D.; Berster, P.; Gelhausen, M. (2005). *Airport Choice in Germany – New Empirical Evidence of the German Air Traveller Survey 2003*. Paper presented at the 9th Air Transport Research Society Conference 2005 in Rio de Janeiro, Brazil.