Environmental Considerations in Freight Transport Choice Modelling

A review of environmental considerations in the choice modelling process regarding freight transportation mode

by

O.L. Olberts

Literature report
Delft University of Technology,
Department of Transport, Engineering and Logistics

Student number: 4633822
Project duration: September 6, 2018 – January 10, 2019
Project supervisor: Dr. B. Atasoy, TU Delft
Specialization: Transport Engineering and Logistics

Report number: 2019.TEL.8313

Title: Environmental considerations in freight transport choice modelling

Author: O.L. Olberts

Title (in Dutch) Milieuoverwegingen bij de modellerings van vrachttransport

Assignment: Literature assignment

Confidential: no

Initiator (university): Dr. B. Atasoy

Supervisor: Dr. B. Atasoy

Date: January 10, 2019

This report consists of 33 pages and 0 appendices. It may only be reproduced literally and as a whole. For commercial purposes only with written authorization of Delft University of Technology. Requests for consult are only taken into consideration under the condition that the applicant denies all legal rights on liabilities concerning the contents of the advice.
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Figures</td>
<td>vii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>ix</td>
</tr>
<tr>
<td>Summary</td>
<td>xi</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2 Freight transport</td>
<td>3</td>
</tr>
<tr>
<td>2.1 Freight mode choice</td>
<td>3</td>
</tr>
<tr>
<td>2.2 Driving factors for optimisation</td>
<td>4</td>
</tr>
<tr>
<td>3 Choice Models</td>
<td>7</td>
</tr>
<tr>
<td>3.1 Choice problem</td>
<td>7</td>
</tr>
<tr>
<td>3.2 Decision maker</td>
<td>8</td>
</tr>
<tr>
<td>3.3 Alternatives</td>
<td>8</td>
</tr>
<tr>
<td>3.4 Attributes</td>
<td>8</td>
</tr>
<tr>
<td>3.5 Decision rule</td>
<td>8</td>
</tr>
<tr>
<td>3.6 Mathematical models</td>
<td>9</td>
</tr>
<tr>
<td>3.6.1 Probit</td>
<td>10</td>
</tr>
<tr>
<td>3.6.2 Logit</td>
<td>10</td>
</tr>
<tr>
<td>3.6.3 Multinomial Logit (MNL)</td>
<td>10</td>
</tr>
<tr>
<td>3.6.4 Mixed Logit (MXL)</td>
<td>10</td>
</tr>
<tr>
<td>3.6.5 Nested Logit (NL)</td>
<td>10</td>
</tr>
<tr>
<td>3.7 Dataset</td>
<td>10</td>
</tr>
<tr>
<td>3.8 Aggregation</td>
<td>11</td>
</tr>
<tr>
<td>4 State of Research</td>
<td>13</td>
</tr>
<tr>
<td>4.1 Modes</td>
<td>13</td>
</tr>
<tr>
<td>4.1.1 Road</td>
<td>13</td>
</tr>
<tr>
<td>4.1.2 Rail</td>
<td>13</td>
</tr>
<tr>
<td>4.1.3 Waterways</td>
<td>15</td>
</tr>
<tr>
<td>4.1.4 Intermodal</td>
<td>15</td>
</tr>
<tr>
<td>4.2 Environmental considerations</td>
<td>15</td>
</tr>
<tr>
<td>5 Conclusion</td>
<td>17</td>
</tr>
<tr>
<td>Bibliography</td>
<td>19</td>
</tr>
</tbody>
</table>
## List of Figures

2.1 Modal split for inland freight transportation in the EU between 1995 and 2015 in billion tonne-kilometers. ([9]) ................................................................. 4

2.2 Modal split for inland freight transportation in the EU between 1995 and 2015 in percentage per billion tonne-kilometers. ([9]) ............................... 5

3.1 Two-level decision tree example for a nested logit model. .......................... 11

3.2 The probability at the average representative utility underestimates the average probability. [35] ................................................................. 11

3.3 Individuals’ $a$ and $b$ probability responses are relatively small when compared to the probability response at the average representative utility. [35] .......... 11
List of Tables

1.1 Air emissions for road, rail and waterway transport in grams/tonne-km for a number of green house gasses ([25]). .......................................................... 2

4.1 Summarised characteristics of researched case study papers. The region wherein the research took place is divided into the United States, Asia and the European Union. Mathematical model used: Probit (P), Logit (L), Multinomial Logit (MNL), Mixed Logit (MXL), Nested Logit (NL) or other (O). For the data is indicated whether it was acquired using Stated Preference (SP), Revealed Preference (RP) or Modelled (M). The alternatives and attributes considered are marked with an 'x' per paper. .......................................................... 14
Summary

The growth in global trade has an impact on the environment, whether it is from noise pollution or from emissions contributing to air pollution and climate change. This encourages research into the consciousness of environmental impacts under freight shippers and thus the research question of this report reads: What is the effect of environmental considerations on freight choice in transportation choice models? To answer this research question, the report has been divided into four sections.

First, an introduction and overview of common freight choices is given, including: route, carrier, time period and mode choice. Since the found differences in air pollution between transport modes are significant, the mode used to transport cargo is believed to have a great impact on environmental pollution. Hence, mode choice is chosen as the focus for the rest of the report.

Second, mode choice is discussed in more detail, where freight transportation worldwide can be split up into five modes: road, rail, air, waterways and inter-modal. The modal split for inland freight transportation shows road transport to be dominant with around 75%, followed by rail (20%) and inland waterways (5%). The percentage of cargo transported by air is negligible in terms of weight ∗ distance. However, since it is highly suitable for small and relatively expensive products, air transport does have a share of over 35% of global trade by value. Mode choice models can benefit both companies and policy makers. Shippers can have different driving factors for mode optimisation, where costs, transit time, reliability and flexibility are frequently used. Even though policy makers or (local) governments are not concerned with the transportation process directly, choice models can help to adjust policies in order to regulate emissions or to persuade shippers in using an underused mode compared to the available capacity.

Third, the essence of choice modelling theory is covered. A researcher targets a specific decision maker and scopes down a complementary set of alternatives. The decision maker and alternatives have their own set of attributes, which act as variables in the mathematical description of a choice model. Common mathematical models used are: Pobit, Logit, Multinomial Logit (MNL), Mixed Logit (MXL) and Nested Logit (NL). Data for the models can be obtained by Stated Preference (SP) from surveys, Revealed Preference (RP) from transport databases or a combination of both.

Last, to answer the research question a literature research among 15 case study choice model papers concerning freight transport has been conducted to give a state of the art of choice modelling. Four attributes have a high occurrence rate: cost, transit time, reliability and frequency. A large portion of researches name congestion on road networks as one of the reasons to investigate mode choice behaviour. Cost factors seemed to be the most influential attribute in general and also when environmental attributes are considered, cost is the underlying motivation for companies. Policy changes in the form of taxes or penalties for the use of non eco-friendly transportation are expected to be effective enough in persuading companies to rethink their transport method and consider emissions as an important factor.

Research papers occasionally acknowledge the appearance of environmental consequences, but see this more as a side effect or obstacle for possible improvements on other factors and it is largely omitted in their discussions. In the current state of research, environmental impacts are an insignificant factor, but validity of this fact is inconclusive, since research papers generally do not include environmental attributes. As environmental awareness is growing and technology advances, mode choice behaviour may change and environmental aspects of mode choice will need to be looked in more when it comes to freight choice modelling.
1

Introduction

Global freight transport has grown considerably in the past decades, and this trend is expected to continue. This growing global transport comes paired with a growing impact on the environment, whether it is from noise pollution or emissions contributing to air pollution and climate change. The International Maritime Organisation (IMO) has reported that the transport sector produces 27.7% of human carbon emissions (IMO (2009)) and Marie Sawadogo and Roy (2011) state that, for their case study of the Amsterdam-Barcelona corridor, the transport sector is the most widespread source of noise in all countries.

These facts are reason for concern and encourage research into the consciousness of environmental impacts under freight shippers around the world. The research question of this literature report reads:

What is the effect of environmental considerations on freight choice in transportation choice models?

To answer this question a literature research is done among a number of case study choice model papers concerning freight transport. The remainder of this introduction will introduce the types of freight choice. Chapter 1 focuses on freight mode choice in particular. Chapter 2 gives a basic theoretical course on choice models in general. Finally, an impression of the current state of research concerning choice modelling of inland freight transport modes and their attributes is given in Chapter 3.

Multiple examples of choices made in freight transportation can be listed. Route choice for example is concerned with the route taken from point A to point B. At first glance one would say that the shortest path is the optimal path as it would save fuel and time. There are however more factors that can be taken into account. In case of truck transportation a driver may encounter toll roads and, depending on the company, the driver or it’s employer turns up for the costs, which in turn makes the outcome dependent of the decision maker. Also, a route might be chosen in such a way that certain locations have to be visited i.e. to pick up or to drop off parts of the cargo (an example of a travelling salesman problem). Furthermore, emissions can play a role when certain vehicles are not allowed to enter low-emission zones.

A transport problem might also be concerned with choice of a transport carrier. A carrier is a firm that is responsible for the movement of cargo on behalf of a customer. Which carrier to pick might be influenced by the total price for the movement from point A to B. Piyush Tiwari and Doi (2003) found that the distance between shipper and port, distance to destination, port congesting and shipping line fleet size (indicating factors like transport time, reliability, service and flexibility) are of great importance.

Shippers may also be interested in the time period wherein transportation takes place. This may be to avoid congestion during peak hours or to fine tune a Lean but strict company time schedule. Gerard de Jong (2016) studied the effect of time-period choice for road freight transport and have found that receivers of goods are to a large extend unwilling to move from
peak hours even if congestion is at its maximum and would react much more when they are taxed for the use of these time periods.

A very common area of research, both in freight as in passenger transport, is mode choice. A mode choice problem revolves around the question of which transportation mode or combination of transportation modes to choose from.

The impact of these types of freight choices on the environment can be different. Since the differences in air pollution between transport modes as shown in Table 1.1 are significant, the mode used to transport cargo is believed to have a great impact on environmental pollution. Hence, mode choice will be the focus in the rest of this literature report.

<table>
<thead>
<tr>
<th>Pollutant</th>
<th>Road</th>
<th>Rail</th>
<th>Waterways</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO</td>
<td>0.25-2.40</td>
<td>0.02-0.15</td>
<td>0.018-0.2</td>
</tr>
<tr>
<td>CO₂</td>
<td>127-451</td>
<td>41-102</td>
<td>30-40</td>
</tr>
<tr>
<td>NOₓ</td>
<td>1.85-5.65</td>
<td>0.20-1.01</td>
<td>0.26-0.58</td>
</tr>
<tr>
<td>SO₂</td>
<td>0.10-0.43</td>
<td>0.07-0.18</td>
<td>0.02-0.05</td>
</tr>
<tr>
<td>Particulates</td>
<td>0.04-0.9</td>
<td>0.01-0.08</td>
<td>0.02-0.04</td>
</tr>
</tbody>
</table>

Table 1.1: Air emissions for road, rail and waterway transport in grams/tonne-km for a number of greenhouse gasses (OECD (1997)).
2

Freight transport

2.1. Freight mode choice

Freight transportation worldwide can be split up into five modes: road, rail, air, waterways and inter-modal. Pipelines might be a sixth mode, but is only suitable for large amounts of liquids and gas and thus is omitted in this research.

Figure 2.2 shows the modal split for inland freight transportation in the EU between 1995 and 2015. This graph show a steady and significant rise in total freight transport and a less dramatic change in modal split over the years.

With around 75% of all inland cargo transported, truck transportation has a clear dominance over all other alternatives. In countries where inland waterways and railroads are scarce, a road network usually do exist and might be the only feasible option for cargo delivery. Furthermore, truck transport is important since it comes into view for other transportation modes as trucks control the 'last mile' in cargo delivery (Chang and Thai (2017)). Compared to other transport modes, road transport has as few advantages: relatively small capital costs, high relative speed and flexibility of route choice (J.P. Rodrigue and Slack (2013)). On the other hand, road transport has relatively high maintenance costs (J.P. Rodrigue and Slack (2013)) plus more air and noise pollution and congestion, per container per mile (Chang and Thai (2017)).

Rail, with a modal split of 20% is the second most common transportation mode. Rail transport is suitable for transporting large quantities of bulk cargo over long distances in a relatively short time. In that perspective it combines two advantages of road and waterway shipping. Rail transport fails to be a door-to-door solution and does not have the same level of flexibility as road transport. This disadvantage also holds for air and waterway transport.

Air transport is the least used transport mode in terms of weight * distance. While air transport is fast compared to the other types of transport, it is also the costliest and for low-value not feasible. Air freight is also not suitable for large products or bulk cargo due to the limited a size of the unit load devices (ULD) frequently used. Air transport does have a share of over 35% of global trade by value (IATA (2018)), which makes it a specialised mode of transport for small and relatively expensive products.

Figure 2.2 shows the modal split for inland waterways. Another, more frequently used term is Short Sea Shipping (SSS). SSS can generally be defined as any waterway-based transport that does not involve crossing an ocean. This therefore includes shipments from nation to nation as well as transport from port to port within the same nation. The benefits of SSS include the handling of large shipments, reduction of road congestion and energy efficiency. However, from the shipper’s perspective SSS is relatively slow and complex (Chang and Thai (2017)).

An alternative to the above mentioned modes is intermodal transport. Intermodal transportation involves a combination of the preceding modes, where the handling of the goods themselves or the change of their transportation unit type is omitted (Munima and Haralambides (2018)). For this mode type it is required that the cargo is loaded in transport units (e.g.
2. Freight transport

2.2. Driving factors for optimisation

A few driving factors behind the appropriate choices in transport have already been mentioned. Transit time is generally defined as door-to-door travel time, but can be defined in various ways. Some studies treat transit time as the average lead time or the travel time. Even though additional handling or inventory activities are usually not been taken into the calculation (Train (2003)), some studies concerning intermodal transport do take these operations into account (B.A. Chiara and Spione (2008)).

Costs can be defined, similarly to transit time, to be the total costs from door-to-door with all possible cost aggregated. This total can include: fuel, storage, terminal handling costs, taxes, etc.

Reliability can also have different definitions depending on the research in question, but usually it relates to an aspect of time. Ravibabu (2013) names four types: a standard deviation of transit time, variation of transit time, maximum period of time and the percentage of shipments arriving in time. Reliability can also relate to safety aspects of cargo (e.g. theft, loss or damage).

The ability to adapt, the flexibility, of a transport mode might also be of importance. In intermodal transport for instance, shipping containers increase flexibility since they allow for faster transshipments and also for a wide verity of routes.

An obvious party that benefits from accurate mode choice models are companies that transport freight. 'Companies' can be an ambiguous term, since companies who have freight to be transported might hire a specific other company to actually take care of the transportation process itself. Thus, a party that is concerned with the shipping process and which benefits from mode choice research are best to be referred to as shippers.

Besides companies being interested in the optimal mode choice for their freight, accurate choice models can also be helpful for (local) governments or policy makers. Even though this party is not concerned with the transportation process directly, such models can clarify the trends and needs of users of the infrastructure. Knowing what factors are important when it comes to mode choice, can help policy makers to adjust policies in order to persuade freight
2.2. Driving factors for optimisation

companies in using a mode which might be underused compared to the available capacity. Policy makers might also try to regulate emissions and therefore favour the use of a specific transport mode.
This chapter is devoted to a summarised course on the theory behind choice models. There are various kinds of choice models, but the general framework consists of a few aspects that will be discussed here.

3.1. Choice problem
Making choices is a part of everyday life, whether it is irrational or rational. The framework of a choice in general can be described as a process including the following steps (Ben-Akiva and Lerman (1985)):

- Defining the choice problem
- Generation of alternatives
- Evaluation of attributes of the alternatives
- Choice between the alternatives
- Final implementation

The outcome of a choice problem need not be the result of an explicit decision-making process. Choices, most often by individuals, can be made out of habit or intuition, an assumed form of behaviour or imitation of an expert or leader (Ben-Akiva and Lerman (1985)). These forms of uninformed choice behaviour are not quantifiable and therefore not suited for choice modelling. Corporate choices are meant to be taken rationally and with the aim to maximise financial profit or to minimise financial loss, either on a short- or long term basis.

The steps concerning an informed choice are to be made successive and specify the theory of a rational choice model as a process wherein the the following elements should be defined (Ben-Akiva and Lerman (1985)):

- Decision maker
- Alternatives
- Attributes of the alternatives
- Decision rule

These elements will be covered in more detail in the following chapters.
3.2. Decision maker
A decision maker can generally be described as an ‘actor’ (Ben-Akiva and Lerman (1985)). An actor can represent a single individual or, in the most abstract context of a firm or organisation, a group of individuals as a single decision maker. Decision makers usually have different characteristics and behave differently as a result of operation in different environments, company size, type of freight, etc.

Some decisions can be made by taking the decision of competing actors in the same decision-system into account. This type of decision making is most often seen in the corporate environment and enters the realm of game theory, rather than decision theory.

3.3. Alternatives
A choice problem is comprised of a set of two or more alternatives. From the universal set of alternatives, the decision maker considers a subset, a choice set, of alternatives which are both feasible and within the scope of knowledge (Ben-Akiva and Lerman (1985)).

Choice sets can be divided into two types: Continuous or discontinuous. Continuous choice sets usually appear in consumer theory where non-negative continuous variables allow calculus based demand functions. Discrete choice modelling is usually concerned with a discontinuous choice set whereby a decision rule is used to state the preference of the decision maker (Ben-Akiva and Lerman (1985)).

3.4. Attributes
The alternative related factors of interest when choosing between a set of alternatives are called the attributes. The attributes of a certain alternative describe the characteristics of that alternative. A decision rule determines the attractiveness of an alternative based on the values of its attributes. For attributes to be processed by a mathematical model, they are assigned to values, which can either be ordinal (indicating a position in an ordered set) or cardinal (countable values). Decision makers are generally not alike, and therefore have specific attributes as well.

3.5. Decision rule
To rationally select the best alternative from a choice set, the decision maker needs to apply a decision rule. Decision rules can generally be classified in four categories (Ben-Akiva and Lerman (1985)).

1. Dominance: If an alternative is better for one or more attributes and no worse for all other attributes, it is said to be dominant with respect to the other alternatives. This decision rule type is virtually as simple as it gets and thus easily applicable. A problem with this type of decision rule is that it does not result in a unique outcome for most situations. Therefore, this rule type can best be used to rule out sub-optimal alternatives.

2. Satisfaction: This decision rule uses the concept of a minimal level of service set by the decision maker. A threshold value can be set for every attribute. If this threshold is not met, the alternative can be excluded. A satisfaction based decision rule also has the weak spot of not necessarily leading to a unique outcome and thus can also best be used as an elimination tool.

3. Lexicographic: A decision maker ranks the attributes by importance and chooses the alternative that is most appealing for the most important attribute. With this decision rule, there are two cases that can occur. In the first case, the attribute are valued by an intrinsic quality and all alternatives with that quality are kept. In the second case a staged process is required. In the first stage the most important attribute is taken into account and the most sub optimal alternative is eliminated. In the second stage, the second most important attribute determines which alternative is eliminated. These steps are to be repeated until an optimal alternative is found.
4. Utility: This type of decision rule uses a combination of all attributes to construct an attractiveness index for every alternative. Comparability of the attributes to a common standard is a vital criterion. The attractiveness of the vector of attributes can be reduced to a scalar, the index of attractiveness, or utility. The alternative with the highest utility is chosen to be the preferred option. The function correlating the vector of attributes to the utility index is called the utility function. Depending on the choice problem, this utility function can be specified for ordinal or cardinal utilities. The calculation (for the researchers’ perspective) of the appropriate weights of the utility function is explained in the following section.

3.6. Mathematical models
In choice model research there are a number of mathematical models that occur frequently:

- Probit
- Logit
- Multinomial Logit (MNL)
- Mixed Logit (MXL)
- Nested Logit (NL)

The models will be discussed to some extend in this chapter, however, for greater detail the review of these models by Kenneth Train in *Discrete Choice Methods with Simulation* or by Ben Akiva in *Discrete Choice analysis - Theory and Application to Travel Demand* can be consulted.

In choice modelling, a decision maker \( n \), can choose between a set of \( J \) alternatives \( C_j \). The decision maker can construct a utility level \( U_{n,j} \) for every alternative \( j \) based on the alternatives’ attributes \( x_{n,j} \) (Train (2003)). From this set the decision maker chooses the alternative with the highest utility.

Note that the attribute weights for a specific alternative are also influenced by the decision maker itself. As stated before, decision makers have specific characteristics. These characteristics are embodied as a set of variables in a vector \( s_n \).

On the other hand there is the researcher. The researcher can observe the same alternatives’ attributes \( x_{n,j} \) and the decision maker’s attributes \( s_n \). The researcher is then able to construct a "representative utility" \( V_{n,j} = V(x_{n,j}, s_n) \), that tries to relate all available attributes to the decision maker’s utility (Train (2003)). The form of \( V \) is linear in its parameters, where \( \beta = [\beta_1, \beta_2, ..., \beta_k] \) is given as the vector of unknown parameters, such that:

\[
V_{n,j} = \beta^T \cdot V(x_{n,j}, s_n) \tag{3.1}
\]

The key point here is that the researcher is not aware of the form of the decision maker’s utility function. therefore a general difference between true-utility and research-utility can arise: \( V_{n,j} \neq U_{n,j} \). To account for the unknown factors, the utility function can be modelled as \( U_{n,j} = V_{n,j} + \epsilon_{n,j} \). Where all \( \epsilon_{n,j} \) together act as the unobserved portion of utility (Train (2003)). Since these terms are unknown to the researcher, they are initially treated as random.

The distribution, or rather density function \( f(\epsilon_{n,j}) \) of the unobserved factors is used to determine the probability of choosing alternative \( i \). The form of \( f(\epsilon_{n,j}) \) specifies the differences in various choice models (Train (2003)). A general form of the probability function can be written as:

\[
P_{n,i} = \text{Prob}(\epsilon_{n,j} - \epsilon_{n,i} < V_{n,j} - V_{n,i} \forall j \neq i) \tag{3.2}
\]

This equation can be expressed using the density function as:

\[
P_{n,i} = \int I(\epsilon_{n,j} - \epsilon_{n,i} < V_{n,j} - V_{n,i} \forall j \neq i) f(\epsilon_n) d\epsilon_n, \tag{3.3}
\]

where \( I \) equals 1 when the expression in parentheses is true and equals 0 when false (Train (2003)). Equation 3.3 is generally applicable. A common type of model is the binary choice model, where the choice set consists at maximum of two alternatives. A choice set containing more than two alternatives is known as a multinomial choice model.
### 3.6.1. Probit
A probit model is based on a normal distribution of the density function: a multivariate normal distribution. This distribution enables to account for correlations between alternatives. The normal distribution can also arise for a single alternative over time when choices are made sequentially. The use of a normal distribution may seem limited. However, since the normal distribution is an important distribution in social sciences and is well established (IESS (2017)), the model is flexible for choices with some form of correlation between alternatives.

### 3.6.2. Logit
A Logit model is based on the assumption that the set of unobserved values \( \varepsilon_{n,i} \) is iid (independent and identically distributed) extreme value for all \( i \). This means that they are uncorrelated and have identical variance for every alternative (Train (2003)). The binary logit probability is given by:

\[
P_n(i) = \frac{e^{\mu V_{ln}}}{e^{\mu V_{ln}} + e^{\mu V_{jn}}},
\]

with \( \mu \) is a positive scale parameter.

### 3.6.3. Multinomial Logit (MNL)
Multinomial models are used, as the name implies, when there are more than two alternatives, as is the case with binary choice models. The choice probability for MNL is given by:

\[
P_n(i) = \frac{e^{\mu V_{ln}}}{\sum_{j \in C_n} e^{\mu V_{jn}}},
\]

By which it can be seen that binary logit is a special case of MNL when \( J_n = 2 \) and \( \mu = 1 \). (Ben-Akiva and Lerman (1985))

### 3.6.4. Mixed Logit (MXL)
The Mixed Logit (MXL) model is an improvement on three limitaitons of the logit model. MXL allowes for random taste variations, correlation in unobserved factors over time and unrestricted patterns (Train (2003)). MXL is basically a weighted average of logit functions for different \( \beta \):

\[
P_n(i) = \int \frac{e^{\mu V_{ln}}}{\sum_{j \in C_n} e^{\mu V_{jn}}} f(\beta) d\beta,
\]

where \( f(\beta) \) is the mixing distribution. Thus, MXL is a choice model wherein the density function can have any distribution and is fully general as it can approximate any discrete (random utility) choice model (Train (2003)).

### 3.6.5. Nested Logit (NL)
Nested logit models are used when the choice sets are divided into subsets, or nests. Each branch of the nested model contains alternatives which correlate to each other. An example of a NL structure is illustrated in Figure 3.1.

### 3.7. Dataset
Thus far, the framework for behavioural choice models is a known process. The modeller can target a specific decision maker and scope down a set of alternatives and their attributes. The last element in the process, decision rule, is where behavioural data is of the essence.

Researchers can acquire data of one of two types: revealed preference (RP) and stated preference (SP).

SP data are collected by means of survey or interviews presenting hypothetical choices to decision makers. The preferred alternative of their choice set is used by the researcher to construct a utility function. SP can be a powerful technique to explore a variety of ideas to a
problem, since SP data can be used for situations that do not physically exist. The downside is that respondents may behave differently to what they state in a hypothetical situation and that it can be difficult to source a large sample size. RP data takes away the hypothetical uncertainty, since it is based on actual choices decision makers have made in the past. RP data may also be collected by the use of surveys or observed through analysis of (large) data sets. Of course, RP cannot be used reliably for hypothetical situations.

Ben-Akiva (1994) has shown that SP and RP can also be successfully combined to use the advantages of both types of datasets, although it is not possible for hypothetical situations.

3.8. Aggregation

Even though choice models are built around individual decision makers, a model designed to reflect the behaviour of a single person is of little use for planning purposes. A researcher is therefore usually interested in an aggregate forecast using disaggregate models. In contrast to linear models, discrete choice models are not suitable for obtaining a reliable aggregate probability by inserting aggregate variable values. An example of this is illustrated in Figure 3.3, where an average is used as a way to aggregate the data. It can be seen that when taking an average representative utility between utility $a$ and $b$, the probability at the average of choosing a certain alternative is underestimating the average probability. Generally, when the individuals’ choice probabilities are low, the probability at the average underestimates the average probability and overestimates when the individuals choice probabilities are high (Train (2003)).

Evaluating the aggregate probability response, or elasticity, can also be problematic as is

![Figure 3.1: Two-level decision tree example for a nested logit model.](image1)

![Figure 3.2: The probability at the average representative utility underestimates the average probability. Train (2003)](image2)

![Figure 3.3: Individuals’ $a$ and $b$ probability responses are relatively small when compared to the probability response at the average representative utility. Train (2003)](image3)
illustrated in Figure 3.2. When the probability derivatives for both individuals is relatively small, the average derivative must also be small. However, at the average representative utility itself, the probability derivative is large. Both these examples show that the aggregation should be done with much care to avoid misleading models.
A literature review among 15 case study modal choice research papers has been conducted to give a state of the art of choice modelling. Aspects covered include: alternatives and attributes used and their geographic differences, mathematical models used, the acquisition of the dataset and the impact of policy changes on mode choice. A summarised table can be found in 4.1.

From table 4.1 can be seen that the number of alternatives considered in the researches range between two (binary choice) to five: road, rail, waterways, inter-modal and air. The binary models tend to focus on road- versus rail- transport or road- versus intermodal transport. The number of attributes used ranges significantly more than the number of alternatives used. Table 4 shows the occurrence of attributes across all investigated researches. Four attributes have a high occurrence rate: cost, transit time, reliability and frequency. This is largely in accordance to a more elaborate literature study conducted by Cullinane and Toy (2000), where the characteristics of the goods was ranked as the fourth most occurring attribute.

4.1. Modes

4.1.1. Road

A clear interest in truck transport can be seen for every research. Truck freight transport is, especially for inland freight, the most common transportation mode and forms therefore a baseline for research. A large portion of researches name congestion on road networks as one of the reasons to investigate mode choice behaviour. For example Derakhshan and Shah (2013a) and Ravibabu (2013) see the congestion as a cause for longer transit times, whereas Chang and Thai (2017) focuses more on the induced air- and noise pollution. Also due to this mode split dominance, truck transportation simply cannot be left out of hypothetical situations and thus occurs frequently in stated preference surveys as well. Furthermore, road transport is in almost all cases an available option, since road networks tend to be more dense than rail or inland waterway networks. As Chang and Thai (2017) acknowledges there are scenarios where trucking can be the only available transportation mode due to the absence of proper railway and waterway infrastructure.

4.1.2. Rail

After road transport, rail freight transportation is the second most used alternative in modelling. Rail transport seems to be relatively disliked in the South East Asia region. In the Delhi-Bombay corridor [Shinghal and Fowkes (2002)], a cost reduction of 15-30% over road transportation is required to persuade shippers to make use of rail services, even under the assumption it is able to match the service quality offered by road transportation. Also in the
South East Asia region, Chang and Thai (2017) finds that higher transport costs shift preferences towards SSS and lower costs shift preferences towards road services, leaving rail transport relatively disliked. Even in other regions, Asian goods repel rail network. Derakhshan and Shah (2013a) found that higher transport costs shift preferences towards SSS and lower costs shift preferences towards road services, leaving rail transport relatively disliked. Even in other regions, Asian goods repel rail network.

Table 4.1: Summarised characteristics of researched case study papers. The region wherein the research took place is divided into the United States, Asia and the European Union. Mathematical model used: Probit (P), Logit (L), Multinomial Logit (MNL), Nested Logit (NL), or other (O). For the data it is indicated whether it was acquired using Stated Preference (SP), Revealed Preference (RP) or Modelled (M). The alternatives and attributes considered are marked with an ‘x’ per paper.

<table>
<thead>
<tr>
<th>Paper</th>
<th>Region</th>
<th>Model</th>
<th>Data</th>
<th>Road</th>
<th>Rail</th>
<th>Waterway</th>
<th>Intermodal</th>
<th>Air</th>
<th>Cost</th>
<th>Transit Time</th>
<th>Reliability</th>
<th>Frequency</th>
<th>Flexibility</th>
<th>Environmental</th>
<th>Distance</th>
<th>Cargo Weight</th>
<th>Cargo Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Derakhshan and Shah (2013a)</td>
<td>US</td>
<td>P &amp; L</td>
<td>SP</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Derakhshan and Shah (2013b)</td>
<td>US</td>
<td>L</td>
<td>SP</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shen and Wang (2012)</td>
<td>US</td>
<td>L</td>
<td>SP</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shinghal and Fowkes (2002)</td>
<td>Asia</td>
<td>L</td>
<td>SP</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Munajat Tri Nugroho and de Jong (2016)</td>
<td>Asia</td>
<td>MXL</td>
<td>SP</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mary R. Brooks and Sammons (2012)</td>
<td>Asia</td>
<td>MXL</td>
<td>SP</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ravibabu (2013)</td>
<td>Asia</td>
<td>NL</td>
<td>RP</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chang and Thai (2017)</td>
<td>Asia</td>
<td>-</td>
<td>RP</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moschovou and Giannopoulos (2012)</td>
<td>EU</td>
<td>L</td>
<td>RP</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J. Rich and Hansen (2009)</td>
<td>EU</td>
<td>NL</td>
<td>RP</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ana-Isabel Arencibia and Román (2015)</td>
<td>EU</td>
<td>MNL</td>
<td>SP</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maria Feo (2011)</td>
<td>EU</td>
<td>MXL</td>
<td>SP</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marie Sawadogo and Roy (2011)</td>
<td>EU</td>
<td>0</td>
<td>M</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reis (2014)</td>
<td>EU</td>
<td>0</td>
<td>SP</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Michal Šimeček (2016)</td>
<td>EU</td>
<td>-</td>
<td>SP</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
akhshanan and Shah (2013b) found that for the inland handling of international freight in the US, importers of Asian goods tend to use the more expensive road transport and do not utilise cheaper available modes like rail. The difference in prices of goods at origin and destination is given as a possible reason for this counter intuitive behaviour, but a solid explanation is still missing for the time being.

4.1.3. Waterways
Preferrence of waterway over other transport modes is largely hindered by the limited by the frequency of the service. Maria Feo (2011) talks about the transit time for waterway transport and came to the conclusion that an increase in travelling speed would go paired with an increase in fuel consumption. This would cause unwanted effects in both economic terms and in regard to energy consumption.

4.1.4. Intermodal
Reis (2014) discusses the evidently false inferiority of intermodal transport in their model regarding the price, transit time, reliability and flexibility and concludes that other factors (implicitly suggesting environmental and other attributes) need to be taken into account in follow-up studies. Interestingly, no paper that reviewed intermodal transport has included environmental attributes.

4.2. Environmental considerations
Table 4.1 shows the occurrence of a number of attributes among the reviewed papers, some more important than others. Cost factors seemed to be the most influential attribute in general and also when environmental attributes are considered, cost is the underlying motivation for companies.

For the costs, usually the total transit/transport cost is taken, but occasionally this attribute is broken down or reduced to a certain stage of the transportation. Derakhshan and Shah (2013a) use fuel costs of truck and rail and Munajat Tri Nugroho and de Jong (2016) take the port costs into account. Cost is generally seen in many researches as the most weighted factor in mode choice with a negative sign in the models. The effect of a higher cost does not necessarily mean a mode shift to a single transportation alternative.

Even though cost is an attribute on its own, it appears as an underlying factor in other attributes as well. For example, shippers can charge more to customers when transit times are shortened, which in turn justifies higher costs to some extend. In short, shippers end up balancing gain for other attributes to the increase in total cost.

The size and price of the shipment also plays a role in the importance of transportation costs. Munajat Tri Nugroho and de Jong (2016) finds empirically that inland mode cost for cargo up to two TEUs is of lesser importance to decision makers.

Increasing or decreasing the costs for certain alternatives, albeit through taxes, penalties or benefits, is practically the only instrument of policy makers to direct a general mode shift. Despite this, Maria Feo (2011) and Ana-Isabel Arencibia and Román (2015) have commented on the effect of road pricing schemes and bonus initiatives and are surefooted that they are effective measures in terms of expected modal shift. An interesting fact of these findings is that a cost increase of road transport creates a slightly higher probability on choosing intermodal transport compared to lowering the costs of intermodal transport itself.

Even though concerns about air pollution around the world are increasing, CO₂- and or other greenhouse gas emissions seem to be an unimportant issue for the transport sector when is comes to mode choice. However, policy changes in the form of taxes or penalties for the use of non eco-friendly transportation would persuade companies to rethink their transport method and consider CO₂ emissions as an important factor [Chang and Thai (2017)]. Other research state that bigger companies (more than 10 TEUs per month) are already more aware of emissions during transport and do take this factor into consideration [Munajat Tri Nugroho and de Jong (2016)].
Conclusion

Despite the literature addressing the consequences of noise pollution, air pollution and climate change, the influence of environmental pollution in the mode choice modelling process remains largely unexplored. Research papers occasionally acknowledge the appearance of environmental consequences, but see this more as a side effect or obstacle for possible improvements on other factors and it is largely omitted in their discussions.

Cost in freight transport seems to be the deciding factor with shippers balancing gain for other attributes to the increase in total cost. This fact, together with the two findings on the positive impact of government pricing schemes and bonus initiatives, implies environmental attributes are possibly overlooked in choice modelling.

In the current state of research, environmental impacts are an insignificant factor, but validity of this fact is inconclusive, since research papers generally do not include environmental attributes. As environmental awareness among the public is growing, shippers might recognise and consider this global environmental issue. This, together with technological advances, may change mode choice behaviour.

An increase of real-time data being available regarding cargo space and arrival time, can make freight forwarders more aware of goods available for transport in the relative vicinity of empty cargo space. This information decreases the amount of empty runs and in turn increases the effective transport capacity and decreases both fuel consumption and costs. Real-time data can also be forwarded to customers as an extra service and further influence mode preference of shippers.

As the number of hybrid and fully electric vehicles will increase, road transport will become a more environmental friendly alternative. Electric vehicles lower noise, pollution, emission levels and are more flexible in cities with zero-emission zones. In the transition phase from fossil fuel to electric vehicles a clear distinction in freight choice modelling has to be made between fossil fuel based transportation and (partial) electric transportation.

With the transport sector having a significant contribution to pollution, mode choice will need to be looked in more when it comes to freight choice. Modellers should be more than invited to examine the environmental consideration of the decision maker and include environmental attributes in their freight choice models.


Ahmad Derakhshan and Muhammad Zaly Shah. Analysis of inland mode choice decision for imported waterborne cargo from new york & new jersey port. Trasporti Europei (Online), 55, 2013b.


