A classification of the long-distance travel market

van Goeverden, Kees; van Nes, Rob; van Arem, Bart

Publication date
2019

Document Version
Final published version

Published in
Moving towards more sustainable mobility and transport through smart systems

Citation (APA)
van Goeverden, K., van Nes, R., & van Arem, B. (2019). A classification of the long-distance travel market. In F. Witlox (Ed.), *Moving towards more sustainable mobility and transport through smart systems: Proceedings of the BIVEC-GIBET Transport Research Days 2019* (pp. 241-255)

Paper published in “Proceedings of the BIVEC-GIBET Transport Research Days 2019, Moving towards more sustainable mobility and transport through smart systems”, pp. 241-255
Editor: Frank Witlox

A classification of the long-distance travel market

Kees van Goeverden[1]
Rob van Nes[2]
Bart van Arem[3]

Abstract: *Long-distance travel contributes significantly to climate change. One of the mitigation options is a shift to more sustainable modes. An efficient policy on modal choice demands knowledge on which market segments are promising for a mode. The paper describes a method for breaking down the travel market into segments that are homogeneous with respect to the appropriateness of a mode for the travellers; the method is applied to all typical long-distance modes. The appropriateness of a mode depends on the demanded standards and the extent the standards of a mode meets the demand. A number of variables that define the appropriateness are identified, and by crossing the most important variables a large number of small, elementary market segments are defined. These are building stones for the mode-specific larger segments with a certain standard of a mode. Based on the explanatory power of a variable for modal choice and the standard of a mode for journeys in the variable category, a standard score of each mode is calculated for each elementary segment. Segments where the score of a mode compared to the score of the best performing alternative is similar (that is: within defined limits) are clustered into one of five segments with a certain standard of the mode. Some results: the proportion of long-distance journeys where a mode has at least a comparable standard is 79% for the car, 60% for the train, and 30% for the airplane. Expressed in mileage, the proportions are 40%, 33%, and 75% respectively.*

Keywords*:* "market segment", "long-distance travel", "Western Europe", "modal standard".

1. <u>Introduction</u>

The transport sector is faced with a number of problems. Currently, climate change receives a lot of attention. The discussion of transport problems is in passenger transport focussed on daily travelling. Considering the contribution to climate change, the incidentally made long-distance journeys should be taken into account as well. This long-distance travel market accounts for a large and increasing share of the GHG emissions of passenger transport (Aamaas et al, 2013, Van Goeverden et al, 2016). The EU addresses the problem and aims at a shift to the most energy efficient modes, particularly the train (EC, 2011).
An efficient policy on modal choice demands knowledge on the travel market: which market segments are promising for the separate modes. The paper addresses this question by describing a method for breaking down the long-distance travel market into segments that are rather homogeneous with respect to the appropriateness of a mode and applying this for each typical long-distance mode. A high appropriateness of a mode makes it promising for attracting passengers from other modes. The result is some mode-specific classifications of the long-distance travel market and the current travel volumes by mode and market segment. This is both helpful for defining policy measures that aim at influencing the modal split, and can be used for an exploration of the market of a new mode, like the Hyperloop. However, the usefulness for the latter can be limited because some characteristics of the new mode may ask for a somewhat adapted classification.

[1] Delft University of Technology, Transport & Planning Department, c.d.vangoeverden@tudelft.nl
[2] Delft University of Technology, Transport & Planning Department
[3] Delft University of Technology, Transport & Planning Department

Long-distance journeys are defined in the paper as journeys to destinations ≥100 km away as the crow flies. This is in Europe a generally accepted standard and matches with the data used for the analyses. The major data source for the analysis are the microdata of the DATELINE project. DATELINE was a survey on long-distance travelling by EU residents, carried out in 2001 and 2002 in the 5th Framework Programme of the EU. The survey covered all 15 EU-countries at that time and Switzerland (DATELINE, 2001). We made an update of these data to 2017 using statistics of Eurostat on the development of aviation and travel of tourists since 2001.

The breakdown analysis has two phases. First, the market is broken down into a large number of small segments with homogeneous standards for all long-distance modes. Second, all small segments with a similar standard for a mode are clustered into a larger segment that includes all trips where the mode has a certain standard. The small segments will be indicated as ‘elementary’ segments; these are building stones for the definition of the larger mode-specific segments with homogenous standards for the modes.

The breakdown into the elementary segments starts with an analysis of the variables that affect the appropriateness of the modes; this is described in Section 2. In Section 3, a selection is made of variables that in the paper are used for definition of the segments, and relevant categories of the variables are defined. The elementary segments are defined by crossing the categories of all selected variables. The clustering into the larger segments is described in Section 4. Five segments are considered: the standard of a mode is inferior, poor, common, good or superior when compared with the best performing alternative mode. The section includes a test of the plausibility of the resulting breakdown by comparing the market shares of a mode in each of the five defined segments; the share should significantly increase when moving to the next segment with higher standards. The conclusions are summarized in Section 5.

2. <u>Variables affecting the appropriateness of long-distance travel modes</u>

The appropriateness of a mode for a journey depends on the demanded travel standards by the traveller and the extent the modal standards meet the demand. The better a mode meets the demands, the higher is the appropriateness and the more promising a mode will be. Next the demanded standards and the supplied standards by the separate modes will successively be discussed. Three attributes for the standards of travelling are considered: time, comfort, and price.

Demanded standards

We consider six variables that influence the demanded standards. These are listed in Table 1. A positive correlation with the demand is indicated by ‘+’, a negative correlation by ‘-‘, and no clear correlation by ‘0’.

Table 1 : Demanded attribute standards

Variable/Attribute	Time	Comfort	Price
Trip duration	0	+	0
Duration of the activity	-	0	0
Time constraints	+	0	0
Physical condition traveller	0	-	0
Income	0	0	-
Luggage	0	+	0

A longer trip duration increases the demands on comfort. A longer duration of the activity might make a short travel time less important because the time spent on travelling is relatively small in the whole trip. The opposite is true in the case of time constraints. The physical condition of the traveller is negatively related to the comfort demand; particularly persons whose physical condition is poor will make high demands. A higher income increases the willingness to pay a higher price. Taking along luggage increases the demands on comfort. The comfort demands may regard different comfort aspects. Comfortable seating will be demanded in the case of a long trip duration and by travellers with a poor physical condition. The latter as well as people with a lot of luggage will demand for no (or a low number of) transfers, which we consider as another aspect of comfort. People with luggage will additionally demand sufficient space for the luggage.

Supplied standards by travel mode

The extent to which the service level of a mode meets the demand depends on a number of variables. The variables are derived from the differences in the standards by travel mode for components of the three attributes (Table 2). Five long-distance modes are considered: airplane, train, public bus, private bus, and car. Public buses are express bus services that run according to a timetable, private buses are buses that are chartered by a group for dedicated use. The ferry, that sometimes accounts for a significant part of the distance travelled, is not considered as a separate mode. Generally the ferry is used as a complementary mode to land modes. However, the need to take a ferry for a part of the trip affects the standards of the land modes, and therefore the need to cross a sea barrier will be included as one of the variables in the analysis.

Table 2 : Standards by travel mode and attribute component

Attribute	Component	Airplane	Train	Bus public	Bus private	Car
Time	Cruising speed	++	0/+	0	0	0
	Leaving/ approaching	--	0/+	0	0	0
	Space accessibility	-	0	0	+	++
	Time availability	0/-	0/-	0/-	+	++/--
	Alternative time use	0	+	0	0	-
Time/comfort	Transfers	-	-	-	0	+
Comfort	Room	0	+	0	0	0
Price		+/-	0/-	+	+	+/-

The cruising speed indicates the speed between the acceleration and deceleration phases, or between the LTO phases in air travel. The cruising speed of airplanes is much higher than that of the other modes. The cruising speeds of the land modes are in the same order of magnitude, except for the faster high speed train services.

Leaving and approaching time is high for the airplane and marginal for the other modes. The exception is leaving or penetrating urban areas by the road modes; this can take a substantial amount of time and give these modes a lower quality than the train with respect to this time component.

Space accessibility regards the vicinity of modal access points, like airports and (major) railway stations. Space accessibility influences access and egress times. The accessibility is generally low for the air mode and high for the private modes, in particular for cars that often can be parked close to the origin and destination addresses.

Time availability is defined by the operating period, the service frequency, and the admittance to the system. System admittance is limited if seat reservation is obligatory and persons are not admitted to fully booked services. In that case the experienced frequency is reduced to the frequency of services that still have seats available. Time availability can even be zero if a mode is not available at all. Generally, time availability is limited for modes with scheduled services (airplane, train, and public bus); the service frequency is rather low on many long-distance connections, and seat reservation is frequently obligatory. The two private modes

have a higher time availability. However, sometimes a car is not available for a journey. Then the time availability of the car is zero and inferior to that of the other modes.
The train performs best with respect to alternative time use. Trains are the most spacious vehicles and have the best possibilities to offer facilities that take a lot of room, like dining and sleeping facilities. In particular the provision of good sleeping accommodation in overnight trains gives this mode a strong position for trips on longer distances (at least 8 hours by train).
Passengers with modes that have a scheduled service sometimes have to make transfers during the journey, both transfers between the main mode and the access/egress modes, and transfers between vehicles of the main mode. Transfers usually enlarge the travel time, and lower the comfort of travelling.
Comfort is additionally related to the available room, and this aspect makes the train the most comfortable mode. Room for seating can be somewhat larger, and there are good possibilities for stretching one’s legs.
A final attribute that affects the competitiveness of the modes is the cost of travelling. The airplane is sometimes expensive and sometimes extremely cheap (low cost carriers), the train is generally expensive, the bus is generally cheap. The price of the car depends on the number of travellers. For someone travelling alone, the car may be the most expensive mode, but if four persons travel together, the car could be the cheapest of all modes, especially when the traveller considers only the marginal travel costs.
The analysis of the different characteristics of the modes gives rise to define next variables:

- Distance. The large differences of the modal standards with respect to the different time components implies that the distance is an important explaining factor for the travel time by mode.
- The need to leave or penetrate urban areas, particularly urban centres. This affects the leaving/approaching time.
- The need to cross a sea barrier. This was explained before Table 2.
- The number of travellers that participate in a journey. This affects the price (per person) of the car.
- Car availability. This is obviously relevant for the car being an alternative or not.

Table 3 displays the appropriateness of the different modes for categories of the mentioned variables. The demand variables from Table 1 are added to the table as well, except for the trip duration. The relation of this variable with the modal standards is rather complicated. Trip duration affects the demand on comfort and depends on the distance (which is included in the table) and the actual mode choice. The demand is relatively low for the fastest mode.

Table 3 : Appropriateness of travel modes for trip/traveller categories

Variable	Category	Airplane	Train	Bus public	Bus private	Car	Fastest mode
Distance	Short	--	0	0	+	+	
	Medium	0	0	0	0	0	
	Long	++	0/-	-	-	-	
Location origin or destination	Central urban	0	+	+	+	-	
	Suburban	0	0	0	0	0	
	Rural	-	-	-	+	+	
Sea barrier	Yes	0	-	-	-	-	
	No	0	0	0	0	0	
Number of travellers	Small	0	0	0	--	-	
	Family	0	0	0	--	+	
	Group	0	0	0	+	+	
Car available	Yes	0	0	0	0	0	
	No	0	0	0	0	--	
Duration of the activity	Short						+
	Long						0
Time constraints	Yes						+
	No						0
Physical condition traveller	Good	0	0	0	0	0	
	Poor	0/-	+/-	0/-	0	0/+	
Income	Low	+/-	-	+	0	0	
	Medium, high	0	0	-	0	0	
Luggage	Little	0	0	0	0	0	
	Much	0	0	0	+	+	
	Large	-	0	-	-	-	

The distance is predominantly relevant for the competitiveness of the airplane. The airplane is inferior to the other modes on rather short distances due to the high ‘fixed’ time costs (leaving/approaching and access/egress). On the other hand, on very long distances the airplane is superior because of its much higher cruising speed. The train performs somewhat better than the other land modes on very long distance because of the good alternative time use; e.g. there is no need to stop at a restaurant if there is a dining car. On short distances the private modes perform better because of the better space accessibility.

The train and bus modes have generally good standards for trips to or from central urban areas. The major railway stations, bus stations and parking areas for coaches are usually located in or close to such areas. Airports are often located outside the urban areas, and for car drivers riding through large cities is not so convenient and they can be faced with parking problems.

The need to cross an important sea barrier reduces the standards of the land modes.

The number of persons travelling together affects the appropriateness of the car and private bus. An increasing number of travellers decreases the costs per car user significantly, and it increases the probability that at least one traveller owns a car and has a driver’s license. Chartering a private bus is only feasible for groups.

Absence of a car makes car use more difficult. Ownership of a driver’s license (by at least 20-25% of the journey participants) is essential for car use.

An increasing duration of the activity lowers the demand on time and the preference for the fastest mode. Time constraints may increase the preference for the fastest mode.

Persons that have a poor physical condition demand good seating comfort (which is highest in the train), comfortable boarding and alighting (the car might here perform best, but the accessibility of other modes is being improved), and no or a limited number of transfers (making the private modes preferable).

Travellers that have a low income are likely to prefer the cheapest modes, that are low-cost carriers and buses. For those who can afford to pay higher fares, we assume that comfort will be more important than price and that particularly the bus will lose attractiveness.
Transporting a lot of luggage is not convenient for the scheduled public modes that require transfers. An advantage of the spacious train is that it affords transporting luggage with large dimensions, like bicycles. Transport of large luggage by train is still not convenient, but transporting by other modes can be impossible (except for sending separately).

Studied impacts on modal choice
The so far theoretical derivation of the variables that define the appropriateness of modes to traveller demands can be compared with the results of studies on the influence of trip/traveller characteristics on modal choice in long-distance travel. A number of European studies on modal choice in long-distance travel has been conducted. Nearly all of them are limited to the analysis of travel behaviour of residents of one country (in most cases Germany, e.g. Zumkeller et al, 2005, Reichert and Holz-Rau, 2015) and sometimes to relatively short-distance trips by excluding international trips or trips by airplane (e.g., Dargay and Steven, 2012, Arbués, P. et al, 2016). We are aware of just one study that has a European scope (Van Goeverden and Van Arem, 2010).
For the analysis of this paper, we selected two studies: Zumkeller et al (2005), and Van Goeverden and Van Arem (2010). These studies have little limitations, inform about the explanatory power of the variables considered, and are based on a dedicated long-distance travel survey. Zumkeller et al analysed the modal choice of German inhabitants on journeys > 100 km separately for different travel purposes. They excluded commuting and intercontinental journeys. Van Goeverden and Van Arem investigated modal choice of the residents of 16 European countries for journeys 100-1500 km crow-fly with a focus on train choice. They excluded commuting as well. For the current paper, we did an additional analysis on mode choice by Dutch residents using microdata of the SDV-survey on leisure travelling by Dutch residents (Structuuronderzoek Dag- en Verblijfsrecreatie), conducted in 1982. We analysed the journeys with one or more overnight stays to destinations 100-1500 km away (crow-fly) as well as the domestic journeys >= 50 km; the reason for the latter is that information on the destination city is only available for domestic journeys.
Based on the χ^2-values, we indicated the variables as highly influencing, important, of secondary importance but still statistically significant, or not significant. The principle for marking a variable as one stage less influencing is a fall of the χ^2-value in the order of 50%. However, the ratios between the χ^2-values may differ substantially in different studies, and within a study for different modes or different purposes; the indications of the influence (in Table 4) are a kind of average based on the author's judgement. Table 4 lists the most important variables according to the different studies as well as the remaining variables that were defined in the Tables 1 and 3. Some variables that are missing in the studies are represented by variables that are correlated with them. Trip duration is correlated with distance, duration of the activity with the number of overnight stays, time constraints with the travel purpose (time constraints apply more frequently for business journeys), and the physical condition of the traveller with age. We added weights to each variable indicating the explaining power for modal choice and ranked the variables according to decreasing importance. Results on the impact of a sea barrier are not provided in any study and not (strongly) correlated with any other variable that could be used as a proxy; DATELINE data demonstrate that the need to cross a sea barrier reduces the share of land modes by about a factor 4. We assume a weight equal to that of other variables that are indicated as highly important in some studies.

Table 4 : Impacts of variables on modal choice in LD travel

Variable	(1)	(2)	(3)	Weight
Car ownership	HI	HI	HI	100
Number of participants in the journey/group composition	SI	HI	IP	80
Crossing sea barrier	-	-	-	80
Country of residence	-	HI	-	80
Number of stages in the trip	HI	-	-	80
Size of the destination city	-	IP	SI	60
Destination abroad	IP	IP	SI	60
Distance (excl. very long distances)	SI	IP	IP	60
Travel purpose	IP	SI	-	60
Employment status traveller	-	IP	-	60
Age traveller(s)	IP	SI	-	60
Number of luggage items/lodging that needs a lot of luggage	IP	-	SI	60
Gender traveller	SI	SI	-	40
Large luggage (“Sondergepäck”)	SI	-	-	40
Drivers’ license traveller(s)	SI	SI	-	40
Income household	-	-	SI	40
Size of the home city	-	SI	NS	20
Number of overnight stays	-	SI	NS	20

Column headings:

(1): Zumkeller et al. (2005)

(2): Van Goeverden and Van Arem (2010)

(3): Analyses of journeys from the SDV-survey

Content of the table:

HI: highly influencing

IP: important

SI: of secondary importance

NS: not significant

-: not included in the analysis

3. Selection of variables and categories for the definition of the segments

The principle of the breakdown analysis is the definition of elementary segments by crossing the relevant variables and clustering these segments into larger segments with a homogeneous appropriateness of a mode for each travel mode. In Section 2, a large number of relevant variables is defined. Crossing all these variables would produce a huge number of elementary segments (it could be between 10 million and 100 million, depending on the number of defined categories per variable). In order to reduce this number, and partly for other reasons, we left out a number of variables. Next, the selection of variables that are used for the analysis and the definition of categories is described.

3.1 Selected variables

Starting from the variables listed in Table 4, we left out a number of variables. These are:

- All person characteristics (except for driver’s license); the majority of long-distance journeys are made by several people travelling together that may have different characteristics.
- Number of stages in a trip; this variable is likely more the result of modal choice than an explanatory factor; travelling by a public mode on longer distances usually impels the traveller to use other modes for access and egress, unlike travelling by a private mode.
- The country of residence and the domestic/international character of the journey; the influence of these variables is not related to the system characteristics of the modes

but to differences in historical developments that created differences in particularly train and bus standards in different countries as well as between domestic and international services.
- Trip purpose and number of overnight stays; these variables are not related to the standards of specific modes, but to those of the fastest mode.
- The two luggage variables and income; these variables are missing in the data that are used for assessing the volumes.

The variables “car ownership” and “driver’s license” are combined into one new variable that we call “car availability”. The definition of car availability in the paper is car ownership of the household combined with at least one traveller that has a driver’s license. Summarizing, next six variables are used for the segment definition:
- Car availability.
- Number of participants in the journey.
- Crossing important sea barrier.
- Distance.
- Size of the destination city.
- Size of the home city.

3.2 Definition of categories and category limits

In this section, the categories of the selected variables will be specified. Additionally, some attention will be paid to the treatment of some missing values.

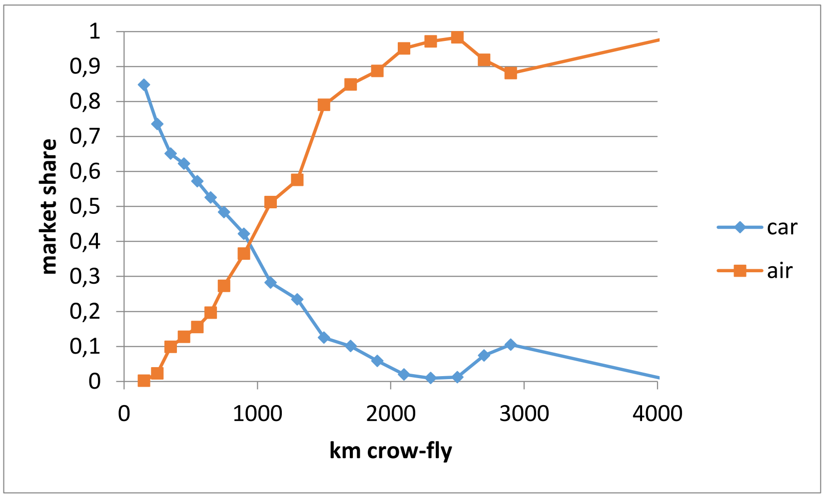
Car availability
This variable has two categories: yes or no. A car is available if a car is owned by the household and at least one traveller has a driver’s license.

Number of participants in the journey
The study of van Goeverden and van Arem (2010) demonstrates that single travellers have outstanding choice behaviour; they are much more inclined to choose the train or bus, and much less inclined to travel by car. For that reason, the category “small” in Table 3 will be split up into two categories: one traveller and two travellers. The defined limit between family and group is 15 persons. A family includes 3-14 travellers, a group ≥15 travellers. Groups should be large enough to make chartering a private bus feasible.

Sea barriers
An important sea barrier on the route is assumed if there is a distance of at least 20 km between an isle and the mainland. No barriers are assumed if there is a bridge or tunnel, or if a direct train service is provided using a ferry. A barrier is also assumed in the case a sea barrier can only be avoided by making a long detour. Examples are journeys from Cornwall to Bretagne, and from South Finland to Central Sweden.

Distance
The definition of distance categories is predominantly relevant with respect to the competitiveness between the land modes and the airplane. As indicated in Table 3, the airplane is not competitive on short distances because of too long travel times, the other modes are not competitive on very long distances for the same reason. The definition of the limits between the distance categories is based on observed modal market shares.
The Figures 1a and 1b show how the market shares of airplane, car, train and bus are correlated with distance. The figures are based on DATELINE data. These data give not the opportunity to distinguish between public and private bus trips. For a good illustration of the

correlation, the figures exclude domestic journeys and journeys that need to cross a sea barrier. The modal shares by distance class differ largely between domestic and international journeys, while the proportion of domestic journeys is strongly correlated with distance. The modal shares for journeys that cross the sea differ significantly from those for other journeys as well (as reported in Section 2).
The curves of the airplane and car in Figure 1a are continuously rising or falling up to their maximum and minimum at ca 2200 km (with remarkably a small increase of the car market share between 2500 and 3000 km). The steep increase of the air curve starts at 200 km.
The curves of the train and bus show an increase to 200 km, a more or less stable (though fluctuating) course to ca 900 km (train) and 1500 km (bus), and a decline to ca 1800 km. The train curve shows a clear dip for the distance class 500-600 km which marks presumably the division between the markets for day and night trains. High-speed trains that could fill the gap generally do not serve international connections.
Based on these observations, we define three distance classes: 100-200 km with rising shares of train and bus and hardly no travel by airplane, 200-1200 km where train and bus have their highest shares, 1200-2000 km with declining shares for all land modes to nearly zero, and >2000 km which is the typical market for the airplane.



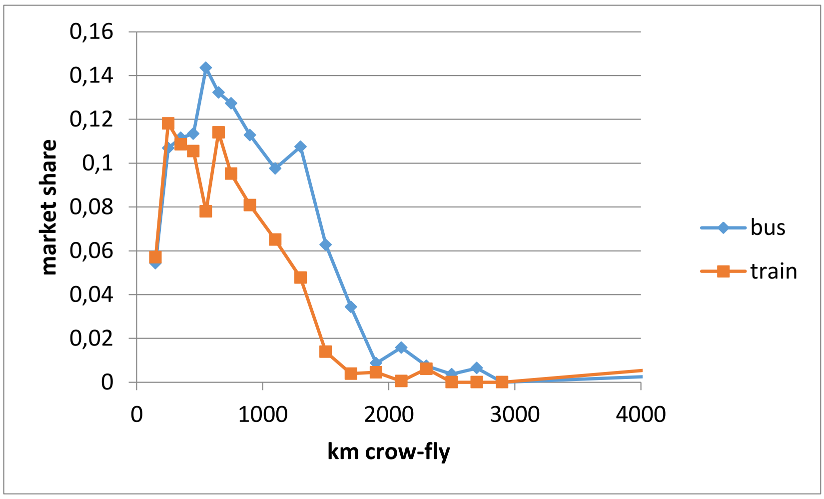


Figure 1a (above) and 1b (below) : The market shares of car, airplane, bus, and train by distance for international journeys without a sea barrier in Europe; source: DATELINE survey

Location of origin or destination
In Table 3 three types of origin or destination are defined: central urban, suburban, and rural. A location is defined as “central urban” if it is in a city that has at least 200.000 inhabitants

and is the main city in the agglomeration. The other cities/municipalities in agglomerations that have a core city with at least 200.000 inhabitants are defined as “suburban”. In order to get rather comparable agglomeration figures all over Europe, we defined the agglomerations by viewing maps (Google earth).

Missing values of car availability and number of travellers

The variables ‘car availability’ and ‘number of travellers’ are missing in the data for a substantial number of journeys. Car availability is missing in the case surveyed persons have no car available and travel together with non-surveyed persons for who information about car ownership and driver’s license is missing. The number of travellers is missing for commuting journeys. For both variables, we did not treat missing values as missing, but assumed a category based on observed market shares of the train; the market share of this mode is strongly associated with both car availability and number of travellers.

In the case persons that have *no car available* travel together with non-surveyed persons, we assume that no car is available for the journey. The reason is that the modal share of the train for this group is similar to that of others who indicate that they have no car available, and quite different from those who have a car available (see Table 5). Presumably, if people of different households travel together, those who have no car available often join others who have no car available as well.

In the case of *commuting journeys*, we assume that people travel alone. The share of the train is for commuters comparable to that for other travellers that travel alone (also shown in Table 5). The DATELINE-data suggest that anyway a large majority of the commuters travels alone. Commuters by car were asked to report whether they travelled as a driver or as a passenger, and they indicated travelling as a driver for 94% of the car trips.

Table 5 : Market share of the train by car availability, number of travellers, and main purpose; source: DATELINE survey

Distance	**Car availability**	**All purposes except for commuting**				**Commuting**
		1 traveller	**2 travellers**	**3-14 travellers**	**≥15 travellers**	
100-200 km	Yes	15%	7%	5%	4%	21%
	No	74%	55%	6%	-	39%
	No?*	-	37%	19%	3%	-
200-1200 km	Yes	24%	9%	8%	6%	22%
	No	52%	27%	47%	-	90%
	No?*	-	28%	23%	20%	-
1200-2000 km	Yes	1%	1%	1%	0%	0%
	No	1%	3%	0%	-	-
	No?*	-	5%	3%	0%	-

*: No car available for surveyed traveller(s), and car availability missing for accompanying traveller(s)

There are two other problems with respect to the number of travellers. The first problem is that sometimes respondents might indicate that they travel alone or with a small group while the mode choice is made for a large group. This can happen in the case of organized group travel for persons that individually sign up for the trip and are not familiar with the other group members (before travelling).

A second serious problem is missing information about the group size for journeys of DATELINE respondents living in countries where the survey was on person level. Therefore, the analyses are limited to the countries where the survey was on household level. These are Austria, Flanders, Germany, Ireland, Italy, Luxemburg, the Netherlands, Sweden and the United Kingdom. Most of the Mediterranean countries are excluded, particularly those where the bus is the dominant public transport mode for long-distance trips. They will not be fully representative for Europe, but will represent Western Europe fairly well.

3.3 Overview and size of the elementary segments

Crossing the defined categories of the selected six variables produces the 184 elementary segments that are indicated in Table 6. If there is a sea barrier or if the distance is longer than 1200 km, we did not distinguish between the locations of origin or destination. We assume that the impact of the sea barrier or very long distance is dominant compared to that of the location types. The table shows the size of the segments in terms of journey numbers in 2017 for residents living in the 9 studied countries. The calculation of the journey numbers is based on the updated data of DATELINE.

Table 6 : Number of journeys by elementary segment in 2017 (millions)

			Car available				No car available				Total
	Number of travellers		One	Two	3-14	≥15	One	Two	3-14	≥15	
Distance	Origin	Destination									
100-200 km	Core city	Core city	40,3	16,1	11,7	0,0	2,5	2,1	0,6	0,0	73,3
		Suburb	6,3	1,3	0,7	0,0	0,3	0,3	0,0	0,0	8,8
		Rural	65,2	32,3	41,7	0,9	5,0	3,9	8,2	1,1	158,3
	Suburb	Core city	21,3	8,3	7,9	0,0	2,8	0,0	0,0	0,0	40,3
		Suburb	1,7	0,8	1,9	0,0	0,1	0,5	0,0	0,0	5,0
		Rural	27,2	15,2	35,5	2,7	0,9	1,7	0,7	0,2	84,0
	Rural	Core city	86,0	51,4	61,5	2,4	4,1	3,8	2,0	0,4	211,6
		Suburb	13,5	4,0	7,4	0,2	1,8	0,2	0,5	0,3	27,9
		Rural	142,6	125,6	151,0	5,7	8,2	4,2	5,4	2,2	444,9
	Sea barrier >20 km		0,8	0,6	1,1	0,1	0,1	0,1	0,1	0,0	3,0
200-1200 km	Core city	Core city	24,8	10,1	11,3	0,3	4,7	3,0	1,3	0,1	55,6
		Suburb	1,3	0,6	0,7	0,0	0,3	0,4	0,1	0,0	3,4
		Rural	15,3	20,9	29,2	0,9	2,9	4,0	2,7	0,5	76,5
	Suburb	Core city	9,1	4,5	5,7	0,2	0,6	0,7	0,2	0,0	20,9
		Suburb	2,1	0,3	0,3	0,2	0,1	0,1	0,0	0,0	3,1
		Rural	10,7	9,1	14,9	0,4	1,0	1,2	1,5	0,1	38,9
	Rural	Core city	52,3	36,5	40,3	1,3	4,3	2,6	1,8	0,3	139,3
		Suburb	3,2	2,9	3,5	0,0	0,1	0,2	0,0	0,0	10,0
		Rural	57,1	64,0	99,7	3,0	4,6	3,0	2,9	0,9	235,3
	Sea barrier >20 km		5,6	7,3	9,2	0,6	1,3	0,7	0,9	0,0	25,7
1200-2000 km	No sea barrier >20 km		5,4	12,1	17,5	0,6	1,2	2,5	1,5	0,2	41,0
	Sea barrier >20 km		1,5	8,1	14,5	0,2	0,6	0,6	1,2	0,1	26,8
>2000 km	All		11,9	35,7	37,4	1,1	3,0	3,7	3,2	0,3	96,3
Total			605,2	467,7	604,6	20,9	50,4	39,4	34,7	6,8	1829,8

4. Definition of the (large) mode-specific segments

The elementary segments shown in Table 4 are the building stones for the large segments of a mode. The large segments are built by clustering all elementary segments that have a similar appropriateness of the mode. Five large segments are distinguished: the mode is inferior, poor, common, good or superior when compared with the best performing alternative mode. The method of attaching one of the five standards of a mode to an elementary segment is as follows:

1. If a mode is indicated by ‘--' in Table 3 or another mode is indicated by ‘++’, the mode is assessed as inferior.
2. If a mode is the only mode indicated by ‘++’ in Table 3, the mode is assessed as superior.
3. In the other cases, an appropriateness score is calculated for each elementary segment and the assignment to one of the five standards is based on the difference of the score

from the scores of both the best performing and the worst performing alternatives. This will be explained next.

The assignment of the standard of a mode (that is not inferior or superior) to the elementary segments could principally be based on the market shares of a mode in these segments. However, an accurate estimation of modal shares requires a lot of observations (preferably several hundreds), and for many elementary segments, the observation number is too small. The total number of independent observations behind the figures of Table 8 is 32.000, representing the total of 1.830 million journeys.

We calculated scores by summing for each variable category the product of the variable weight and the modal appropriateness (Equation 1).

$$S_m^i = \sum_{v=1}^{j} W_v A_m^{c(v)} \qquad (1)$$

where

S_m^i : score indicating the appropriateness of mode m in elementary segment i

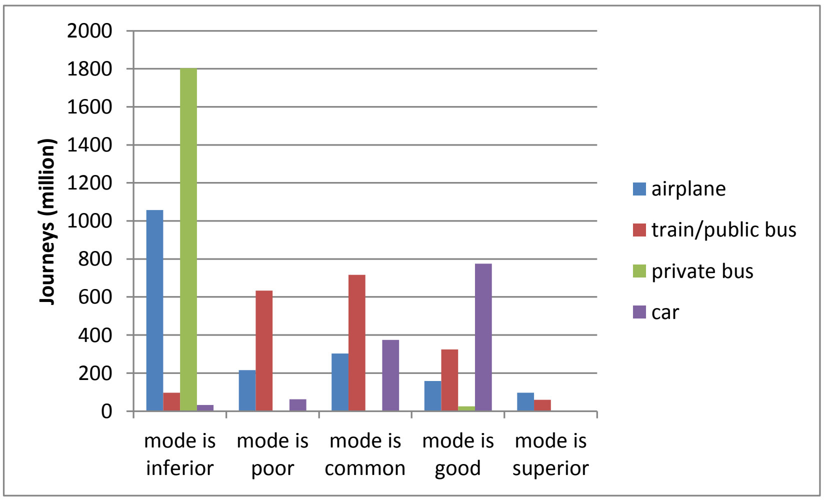
W_v : weight of variable v (according to Table 4)

$A_m^{c(v)}$: appropriateness of mode m for category c of variable v (valid for segment i)

The score $S_m^{c(v)}$ is calculated from the standards indicated in Table 3. The standards are transformed into numbers: ‘--' into -2, ‘-‘ into -1, ‘0’ into 0, ‘+’ into +1, and ‘++’ into +2. Then, the standard of the mode for an elementary segment is assessed by comparing the score of the mode with the scores of both the best and the worst performing modes. The largest differences from the scores of the best/worst performing modes are between 200 and 300 (depending on the mode). The standard of a mode is defined by next algorithm: if the score is more than 150 below that of the best performing mode, the standard is assessed as poor; if it is between 75 and 150, the standard is assessed as common; otherwise, the standard is assessed as either good or common, depending on the difference with the worst performing mode; if this difference exceeds 75 and the worst performing mode is not assessed as inferior, the standard is assessed as good, otherwise it is assessed as common.

The resulting standards of a mode are indicated in the tables in the appendix. The standards for train and the public bus are presented in one table. Their standards are similar, except for comfort (the train is more comfortable) and price (the bus is cheaper). The choice between the two modes will predominantly be explained by income: those who can afford to pay the higher prices will travel by train, others by bus. This is supported by Dargay and Steven (2012) who analysed long-distance travel within Britain and estimate a very low income elasticity for the bus (0,10, compared to 0,83 for the train). However, income is not selected as a variable for the segment definition. Furthermore, the ‘superior’ standard is not defined for both the private bus and the car in any elementary segment. These modes are assumed to be never superior to the best performing alternative.

The figures 2a and 2b show the volumes of the modal segments, in journey numbers and mileage respectively. The presented volumes are the volumes for all modes in a segment with a certain standard for the indicated mode. For example, the airplane is inferior for ca 1 billion journeys, and ca 500 billion kilometres are travelled for journeys where the train/public bus standard is common. Here we remark, that the breakdown for the private bus is predominantly based on the reported number of persons travelling together by DATELINE respondents. As we indicated in Section 3.2, this will likely underestimate the journeys with group travel and hence the market for the private bus. The total journey numbers will then be lower in the segment where the private bus is inferior and higher in the other segments.



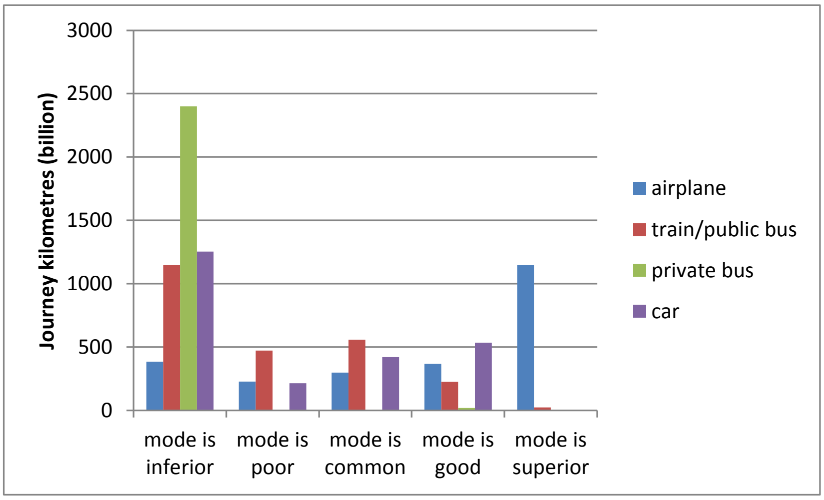


Figure 2a (above) and 2b (below) : The volumes of the segments in terms of journey numbers and mileage in 2017; source: updated DATELINE survey

The figures show that, if we ignore the private bus, most journeys are made in segments where the airplane is inferior, the train is poor or common, and the car is good. The mileage picture is quite different. Most kilometres are made in segments where the airplane is superior and the train and car are inferior. If a ‘promising’ market for a mode is defined as the market where the standard is common or better than that of the best performing alternative, we find that 79% of the journeys are promising for the car, 60% for the train and public bus, and 30% for the airplane. Expressed in the mileage of the markets, the shares are 40% for the car, 33% for the train/public bus, and 75% for the airplane.
One might expect that the market share of a mode correlates with its standard. Calculation of the market shares of a mode for all defined segments can then be used for a check of the breakdown result for this mode. Figure 3 shows the market shares for the airplane, train, and car in the segments where the mode has a certain standard. The shares for the public and private buses could not be produced because the data do not specify these two bus types. The figure shows that indeed a strong correlation exists for all three modes. We remind the reader that no superior standard for the car was defined.

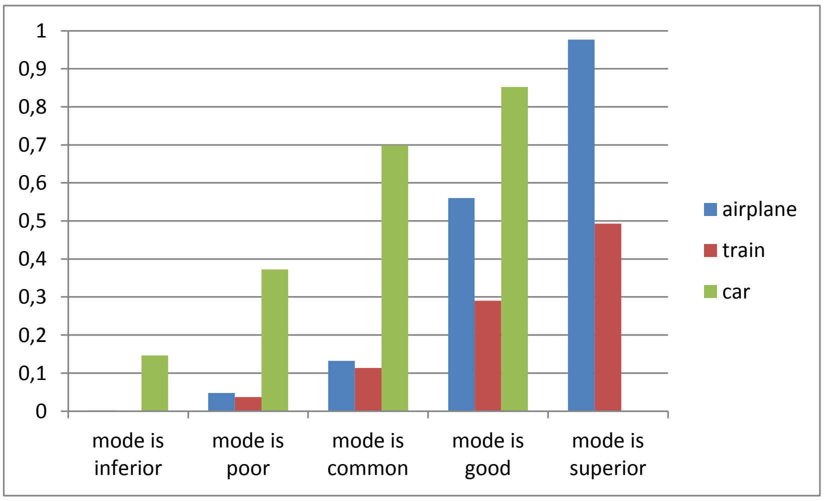


Figure 3 : The market shares by standard segment; source: updated DATELINE survey

One might wonder why the share of the train is only ca 50% for the journeys where this mode is superior. However, in fact, the train is not superior, but the public land modes (train and public bus) together are superior. The market share for train and bus (public and private) is 66% in this segment.

5. Conclusion

The paper presents a rather straightforward method for a breakdown of the long-distance travel market into segments with a certain level of appropriateness of a mode. The starting point is the notion that the appropriateness of a mode for a journey depends on the demanded travel standards by the traveller and the extent the standards of a mode meet the demand. The standard of a mode is assumed to be defined by time, comfort, and price. Variables that define either the demanded or the supplied standards are identified, and a selection of them are used for the definition of the segments. In the selection process all the variables that define the demanded standards were excluded for various reasons. The final breakdown is therefore fully based on the variables that define the supplied standards of the different long-distance modes. These include the most influencing variables with respect to modal choice. The breakdown can be used for the definition of policies on modal shifts or the exploration of the market of a new transport mode.

6. References

Aamaas, B., Borken-Kleefeld, J., Peters, G.P. (2013) The climate impact of travel behavior: A German case study with illustrative mitigation options. *Environmental Science & Policy 33*, 273-282

Arbués, Pelayo, Baños, José F., Mayor, Matías, Suárez, Patricia (2016) Determinants of ground transport modal choice in long-distance trips in Spain. *Transportation Research Part A. 84*, 131-143

Dargay, J. M., Clark, S. (2012) The determinants of long distance travel in Great Britain, *Transportation Research Part A 46*, 576-587

DATELINE Project (2001). *Deliverable 2, Final Survey Design for a European Travel Survey on Long-distance Mobility*.

EC (2011). *White paper, Roadmap to a single European Transport Area – Towards a competitive and resource efficient transport system*. European Commission, Brussels, Belgium.

Reichert, Alexander, Holz-Rau, Christian (2015) Mode use in long-distance travel, *The journal of transport and land use 8* (2), 87-105

Van Goeverden, C.D., and Van Arem, B. (2010). *Background factors explaining train choice in European long-distance travelling*. Paper presented at the 12th World Conference on Transportation Research, Lisbon, Portugal.

Van Goeverden, K., Van Arem, B., and Van Nes, R. (2016). Volume and GHG emissions of long-distance travelling by Western Europeans. *Transportation Research Part D 45*, 28-47

Zumkeller, D., Manz, W., Last, J., and Chlond, B. (2005). *Die intermodale Vernetzung van Personenverkehrsmitteln unter Berücksichtigung der Nutzerbedürfnisse (INVERMO), Schlussbericht*, Institut für Verkehrswesen, Universität Karlsruhe, Germany.

Appendix: Modal standards of the elementary segments

The Tables 7 to 10 show the standards of the airplane, train/public bus, private bus and car for the elementary segments. The standards are indicated by a number: 1 is inferior, 2 is poor, 3 is common, 4 is good, and 5 is superior.

Table 7 : Standards of the airplane

			Car available				No car available			
	Number of travellers		**One**	**Two**	**3-14**	**≥15**	**One**	**Two**	**3-14**	**≥15**
Distance	**Origin**	**Destination**								
100-200 km	Core city	Core city	1	1	1	1	1	1	1	1
		Suburb	1	1	1	1	1	1	1	1
		Rural	1	1	1	1	1	1	1	1
	Suburb	Core city	1	1	1	1	1	1	1	1
		Suburb	1	1	1	1	1	1	1	1
		Rural	1	1	1	1	1	1	1	1
	Rural	Core city	1	1	1	1	1	1	1	1
		Suburb	1	1	1	1	1	1	1	1
		Rural	1	1	1	1	1	1	1	1
	Sea barrier >20 km		1	1	1	1	1	1	1	1
200-1200 km	Core city	Core city	3	3	3	2	3	3	3	2
		Suburb	4	3	3	3	3	3	3	3
		Rural	3	3	2	2	3	3	3	2
	Suburb	Core city	4	3	3	3	3	3	3	3
		Suburb	4	3	3	3	3	3	3	3
		Rural	3	3	2	2	3	3	3	2
	Rural	Core city	4	3	3	2	3	3	3	2
		Suburb	3	3	3	3	3	3	3	3
		Rural	3	2	2	2	3	3	3	2
	Sea barrier >20 km		4	4	4	4	4	4	4	4
1200-2000 km	No sea barrier >20 km		4	4	4	4	4	4	4	4
	Sea barrier >20 km		4	4	4	4	4	4	4	4
>2000 km	All		5	5	5	5	5	5	5	5

Table 8 : Standards of the train and public bus

			Car available				No car available			
	Number of travellers		One	Two	3-14	≥15	One	Two	3-14	≥15
Distance	Origin	Destination								
100-200 km	Core city	Core city	4	4	3	3	5	5	5	3
		Suburb	3	3	3	3	5	5	5	3
		Rural	3	3	2	2	5	5	5	2
	Suburb	Core city	4	3	3	3	5	5	5	3
		Suburb	3	3	3	3	5	5	5	3
		Rural	3	2	2	2	5	5	5	2
	Rural	Core city	4	3	3	2	5	5	5	2
		Suburb	3	3	2	2	5	5	5	2
		Rural	3	2	2	2	5	5	5	2
	Sea barrier >20 km		3	3	3	3	5	5	5	3
200-1200 km	Core city	Core city	4	4	4	3	4	4	4	3
		Suburb	4	3	3	3	3	3	3	3
		Rural	3	3	2	2	3	3	3	2
	Suburb	Core city	4	4	3	3	3	3	3	3
		Suburb	4	3	3	3	3	3	3	3
		Rural	3	3	2	2	3	3	3	2
	Rural	Core city	4	4	3	3	3	3	3	3
		Suburb	3	3	3	3	3	3	3	3
		Rural	3	2	2	2	3	3	3	2
	Sea barrier >20 km		3	3	3	3	3	3	3	3
1200-2000 km	No sea barrier >20 km		3	3	3	3	3	3	3	3
	Sea barrier >20 km		2	2	2	2	2	2	2	2
>2000 km	All		1	1	1	1	1	1	1	1

Table 9 : Standards of the private bus

			Car available				No car available			
	Number of travellers		One	Two	3-14	≥15	One	Two	3-14	≥15
Distance	Origin	Destination								
100-200 km	Core city	Core city	1	1	1	4	1	1	1	4
		Suburb	1	1	1	4	1	1	1	4
		Rural	1	1	1	4	1	1	1	4
	Suburb	Core city	1	1	1	4	1	1	1	4
		Suburb	1	1	1	4	1	1	1	4
		Rural	1	1	1	4	1	1	1	4
	Rural	Core city	1	1	1	4	1	1	1	4
		Suburb	1	1	1	4	1	1	1	4
		Rural	1	1	1	4	1	1	1	4
	Sea barrier >20 km		1	1	1	4	1	1	1	4
200-1200 km	Core city	Core city	1	1	1	4	1	1	1	4
		Suburb	1	1	1	4	1	1	1	4
		Rural	1	1	1	4	1	1	1	4
	Suburb	Core city	1	1	1	4	1	1	1	4
		Suburb	1	1	1	4	1	1	1	4
		Rural	1	1	1	4	1	1	1	4
	Rural	Core city	1	1	1	4	1	1	1	4
		Suburb	1	1	1	4	1	1	1	4
		Rural	1	1	1	4	1	1	1	4
	Sea barrier >20 km		1	1	1	4	1	1	1	4
1200-2000 km	No sea barrier >20 km		1	1	1	3	1	1	1	3
	Sea barrier >20 km		1	1	1	2	1	1	1	2
>2000 km	All		1	1	1	1	1	1	1	1

Table 10 : Standards of the private car

			Car available				No car available			
	Number of travellers		One	Two	3-14	≥15	One	Two	3-14	≥15
Distance	Origin	Destination								
100-200 km	Core city	Core city	2	3	3	2	1	1	1	1
		Suburb	3	3	4	4	1	1	1	1
		Rural	3	4	4	4	1	1	1	1
	Suburb	Core city	3	3	3	3	1	1	1	1
		Suburb	3	3	4	4	1	1	1	1
		Rural	4	4	4	4	1	1	1	1
	Rural	Core city	3	3	3	3	1	1	1	1
		Suburb	3	4	4	4	1	1	1	1
		Rural	4	4	4	4	1	1	1	1
	Sea barrier >20 km		3	3	4	4	1	1	1	1
200-1200 km	Core city	Core city	2	2	3	2	1	1	1	1
		Suburb	3	3	3	3	1	1	1	1
		Rural	3	4	4	4	1	1	1	1
	Suburb	Core city	2	3	3	3	1	1	1	1
		Suburb	3	3	4	4	1	1	1	1
		Rural	3	4	4	4	1	1	1	1
	Rural	Core city	2	3	3	3	1	1	1	1
		Suburb	3	3	4	4	1	1	1	1
		Rural	4	4	4	4	1	1	1	1
	Sea barrier >20 km		2	3	4	4	1	1	1	1
1200-2000 km	No sea barrier >20 km		3	3	3	3	1	1	1	1
	Sea barrier >20 km		2	2	2	2	1	1	1	1
>2000 km	All		1	1	1	1	1	1	1	1