Issue 24(4), 2024, pp. 62-88
ISSN: 1567-7141
TU Delft OPEN

EJTIR

Research paper
https://ejtir.tudelft.nl
DOI: 10.59490/ejtir.2024.24.4.6696

Estimation of demand models for long-distance international travel – Key determinants for Swedes' travel abroad

Ida Kristoffersson[1], Chengxi Liu[2]

[1] corresponding author ida.kristoffersson@vti.se , Traffic Analysis and Logistics Department, VTI Swedish National Road and Transport Research Institute, Sweden; https://orcid.org/0000-0002-3738-9318
[2] Traffic Analysis and Logistics Department, VTI Swedish National Road and Transport Research Institute, Sweden; https://orcid.org/0000-0001-6966-9077

Keywords
Long-distance international travel
Mode and destination choice
Large-scale transport model

Publishing history
Submitted: 26 August 2022
Revised date(s): 15 March 2023, 18 August 2023, 17 December 2023, 20 May 2024
Accepted: 06 September 2024
Published: 29 November 2024

Cite as
Kristoffersson, I., & Liu., C. (2024). Estimation of demand models for long-distance international travel – Key determinants for Swedes' travel abroad. *European Journal of Transport and Infrastructure Research, 24*(4), 62-88.

Abstract

Although long-distance international travel contributes significantly to global emissions from the transport sector, disaggregated travel demand forecasting models on long-distance international travel are scarce. Large infrastructure investments such as high-speed rail may have a profound impact on long-distance international travel demand and thus need to be evaluated using such models. In this study, a disaggregated travel demand forecasting model is estimated using Swedish national travel survey data from 2011-2016 along with detailed supply data from European road, train, and ferry networks and a World-wide air network, aiming at forecasting Swedes' long-distance travel abroad. Mode choice, destination choice and trip generation are modelled by traditional Nested Logit models and Multinomial Logit models. The model is segmented by purpose (private or business) and for private trips also by number of nights away. The model estimation results reveal effects of individual socio-economic attributes, level-of-service attributes, and destination characteristics. Marginal effect estimates of level-of-service attributes for train suggest that infrastructure investments in high-speed rail network may have a profound effect on demand for long-distance international travel, especially for business trips.

1 Introduction

Long-distance international travel differs from regional and national travel in many respects, such as what determines traveller trip generation, mode, and destination choice. Because of the long distances of these trips, they usually contribute significantly to a country's total passenger kilometres travelled, even though the number of long-distance international trips is generally lower than the number of regional and national trips. Passenger kilometres travelled by mode is important, especially because it is related to CO2 emissions from transport, for which ambitious reduction targets have been set both at the EU level (European Environment Agency, 2021) and national levels (Swedish Ministry of the Environment, 2021). Furthermore, there has for several years been an increased focus in Europe on a Single European Railway Area and long-distance cross-border passenger rail travel (European Commission, 2021). Travel demand forecast models are an important part of large-scale modelling to provide accurate inputs for cost-benefit analyses (CBA) of large infrastructure investments or policy measures. The major advantage of these forecast models is that planned but not implemented investments and policies can be tested in the models and effects analysed. Most operational travel demand forecast models, such as national travel demand models (Beser and Algers, 2002; Daly, 2007; Rohr et al., 2013), include components for trip generation, destination choice, and mode choice at the individual or population segment level. Despite the above, literature on travel demand models that estimate the key determinants of long-distance international travel is still scarce. The definition of long-distance international in the literature also varies, but in many cases, it refers to cross-border trips that are longer than 100 km.

One of the few existing demand models of long-distance international travel is Trans-Tools, which is a transport model for both passenger and freight transport in 42 European countries. Rich and Mabit (2012) described the demand model for passenger transport. In Trans-tools, there are separate demand models for business trips, vacation trips, and other private trips. The model includes five modes: car as driver, car as passenger, bus, train, and air. It is based on data from the DATELINE travel survey conducted in 2000. The networks (car, train, and air) and their level of service attributes are described in Rich et al. (2009). A model called Trust (TRT Trasporti e Territorio, 2018) was developed as a follow-up to the Trans-tools model; however, in Trust, there is no demand model; instead, demand is treated as a fixed origin-destination (OD) matrix. Pieters et al. (2012) describe an effort to develop sub-models for cross-border traffic in the Dutch national model. Somewhat more common than large-scale demand models of international travel are the so-called direct-demand models, especially concerning tourist travel. These models typically calculate the total number of tourists travelling to/from a destination zone as a function of e.g., GDP and population. Owing to the aggregate nature of these models, it is not possible to calculate the cross-elasticities between the modes. Examples of direct demand models include Divisekera (2010) for Australia, Santana-Jiménez and Hernández (2011) for the Canary Islands, and Li et al. (2017) for China. Some direct-demand models focus on a certain mode, especially air travel, and predict the number of air trips to certain airports (Gelhausen et al., 2018; Kim and Shin, 2016; Suh and Ryerson, 2019). Dargay and Clark (2012), on the other hand, modeled total travel distance per individual for long-distance trips independent of mode, and Janzen et al. (2018) utilized mobile phone data to validate the number of long-distance trips, showing that the number of long-distance trips is often underestimated in travel surveys. There are also studies that focus on coach as a mode for long-distance travel and investigate which factors increase its competitiveness relative to other modes (Van Acker et al., 2020) and how to adapt the coach service to make it attractive specifically for business travellers (Lannoo et al., 2018).

The lack of travel demand models for long-distance international travel can be a problem in practice when certain investments or policy measures have a substantial impact on international travel demand. One such example is high-speed rail that connects large cities across countries. Witlox et al. (2022) determined a number of existing bottlenecks for European rail, such as train travel time not being fast enough and too many interchanges. An analysis of the ability of policy measures and investments to eliminate these bottlenecks would benefit from travel demand models for long-distance international travel. For Sweden, the plans of high-speed rail linking Stockholm, Gothenburg, and Malmö (with easy access to Copenhagen) have the potential to alter the current picture of domestic travel as well as international travel. To systematically quantify the impact of investments such as high-speed rail on travel demand in CBA, national demand models need to be expanded to include demand for long-distance international travel.

An effort is made in this study by estimating demand models for Swedes' long-distance travel (trip distance of 100 km or longer). The current national model for passenger transport in Sweden, *Sampers*, previously included a module for international trips (Beser and Algers, 2002) but that module is no longer in use. In this study, models were estimated for private and business trips, where private trips include work trips[1], study trips, shopping trips, vacation trips (the largest category), visit friends/relatives, recreational trips, and trips for religious purposes. Three models for private trips are estimated depending on nights away, showing the importance of the behavioural difference in mode and destination choice depending on how many nights the traveller is away. The estimated demand models include i) socio-economic variables such as income, age, sex, and children in the household; ii) level-of-service variables such as travel time, travel cost, and waiting time calculated from detailed European car, train, and ferry networks, as well as a worldwide network for air travel; and iii) destination attraction variables such as population, vacation zone, and GDP. The estimated demand models were further investigated by analysing the models' value of travel time and by calculating model elasticities when changing cost, travel time, and waiting time for the train mode.

The remainder of this paper is organized as follows. Section 2 describes the travel survey data and the European/Worldwide network used to generate the level of service attributes. Section 3 presents details of the model formulation. Section 4 includes the model estimation results for mode and destination choices, trip generation, and high-speed rail scenario elasticities. In Section 5, the values of the travel time are discussed. Section 6 concludes.

2 Data

2.1 *Travel demand data*

The travel demand data consists of 5174 observations of long-distance (one-way distance of 100 km or longer) international trips[2] from the Swedish national travel survey for the years 2011-2016 (Trafikanalys, 2017). The respondents in the Swedish national travel survey were asked which trips longer than 100 km they had made during the last month, and which trips longer than 300 km they had made during the last three months. There was one more national travel survey conducted in 2019, but in the 2019 survey, only trips from the measurement day were asked, which resulted in very few long-distance international trips. Therefore, the 2019 survey was not included in this study. Of the 5174 observations, 194 observations had missing values in numbers of nights away, 41 observations had missing values in travel mode, and 433 observations had missing values in trip origin and destination zone codes. Subtracting these observations leads to 4506 observations. Further analysis shows that 228 observations have possibly wrongly coded chosen travel choices, for example, one observation shows travel from Stockholm to Canary Island in Spain by car. This is checked by skimming the network described in Section 2.2. Finally, 4278 observations were used in the model estimation.

---

[1] Work trips are commute trips to work which are paid for by the traveller, unlike business trips which are trips paid for by the employer.
[2] The observations are *tours* from home to the destination and back home again, but we use the word trip in this paper for convenience. Trips that have start location within Sweden, while their destination is outside Sweden are categorised as international trips. We model the trip from Sweden to destinations abroad in the model estimation. In a possible future implementation of the model, it will be assumed that the homebound trip will be made by the same mode as the outbound trip.

After the data cleaning process described above, the trip data consists of 3561 (83%) private trips and 717 (17%) business trips. Out of the private trips, 324 (9%) are daytrips, 1348 (38%) are trips with 1-5 nights away, and 1889 (53%) are trips with six or more nights away. Figure 1 shows the modal shares observed for private and business trips, where private trips are divided by the number of nights away. The modal shares for private trips differ significantly depending on the number of nights away, which is the motivation for testing model segmentation across this variable. Figure 1 shows that car trips dominate for private day trips, car and air trips are of approximately equal size for private trips 1-5 nights away, and air is the dominant mode for private trips 6+ nights away and for business trips.

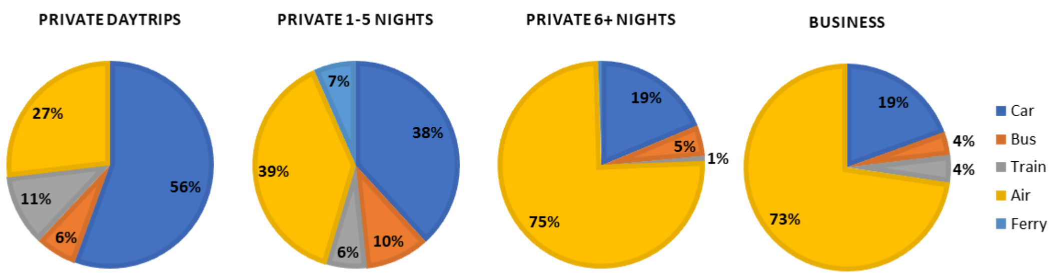


Figure 1. Modal shares for Swedes' long-distance international trips in the travel behaviour survey from 2011-2016.

Regarding the chosen destinations in the data material, travellers choose destinations further from Sweden for trips with more nights away. Figure 2 shows the share of trips with a destination within Nordic countries[3], a destination within Europe outside Nordic countries[4], and a destination outside Europe, for private daytrips, private 1-5 nights, private 6+ nights, and business trips. For day trips, destinations within Nordic countries dominate. For to 1-5 nights number of trips to Nordic countries and other EU countries is similar. For 6+ nights, destinations to other EU countries constitute the majority, while destinations outside the EU account for a significant share. For business trips, destinations to Nordic countries and other EU countries are the majority, while a considerable number of trips still have destinations outside the EU.

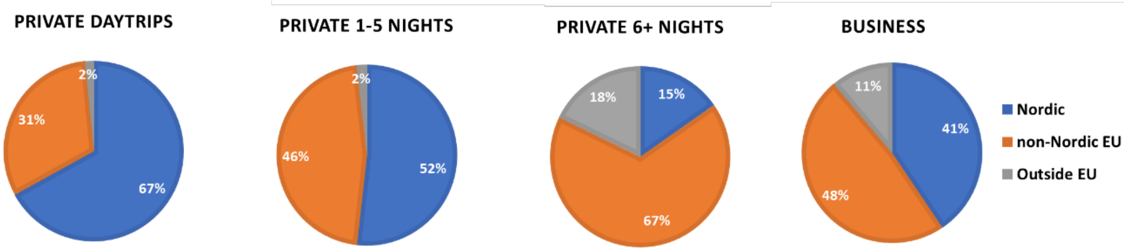


Figure 2. Destination shares for Swedes' long-distance international trips in the travel behaviour survey from 2011-2016.

To estimate trip generation, the choice not to make a long-distance international trip must also be included. Individuals who do not make any long-distance international trips (they have made domestic trips as they are registered in the travel survey data) are therefore added to the data to estimate trip generation. We use this larger trip dataset to model the choice between no trip (only domestic travel), daytrip, 1-5 nights, and 6+ nights for private international travel, and the choice between no trip (only domestic travel) and trip for international business travel. No further

[3] The Nordic countries in this study refer to Denmark, Norway, Finland, and Iceland.
[4] Switzerland is included in the category of "non-Nordic EU" even though Switzerland is not part of European Union. The non-EU Balkan countries such as Serbia are included in the category "Outside EU".

segmentation of the number of nights away for international business travel was made because of the lack of data in some segments.

Table 1. Socio-economic characteristics of the estimation sample.

	Private						Business	
	daytrip		1-5 nights		6+ nights			
Socio-demographics	N	%	N	%	N	%	N	%
Age								
Age <18	34	10%	106	8%	265	14%	0	0%
Age 19-30	42	13%	185	14%	267	14%	46	6%
Age 31-64	148	46%	756	56%	937	50%	644	90%
Age >64	100	31%	300	22%	420	22%	27	4%
Gender								
Male	188	58%	696	52%	919	49%	558	78%
Female	136	42%	651	48%	970	51%	159	22%
Household income								
HHInc<=25 TEUR	25	8%	77	6%	123	7%	8	1%
HHInc<=70 TEUR	153	47%	545	40%	745	39%	167	23%
HHInc>70 TEUR	91	28%	415	31%	532	28%	463	65%
HHInc missing	55	17%	310	23%	489	26%	79	11%
Car ownership								
No car	16	5%	154	11%	185	10%	48	7%
One car	173	53%	643	48%	937	50%	258	36%
Two cars	107	33%	463	34%	637	34%	352	49%
More than two	28	9%	87%	7%	130	6%	59	8%
Housing								
Living in Villa	236	73%	930	69%	1270	67%	538	75%
Not in Villa	88	27%	417	31%	619	33%	179	25%
Children in household								
Have 0-6 years old	41	13%	168	12%	205	11%	164	23%
Have 6+ years old	63	19%	304	23%	554	29%	142	34%

The travel survey data used in this study were collected between 2011-2016. International travel during 2020-2021 has of course, in absolute numbers, been largely affected by the Covid-19 pandemic and the restrictions that have been introduced. However, this does not necessarily mean that the underlying travel preferences of travellers have changed. In contrast, it is likely that people would like to travel if their circumstances were different. It is difficult to say today if there will be long-lasting travel behavioural changes caused by the pandemic, such as changes to travel time and travel cost sensitivities, which could make the demand model estimations conducted in this study less representative. Similarly, using survey data from to 2011-2016 also implies that preferences are assumed to be unaffected by flight shame and generational effects and that any self-selection effect is constant. Previous studies have shown that travel preferences are relatively stable over time and that the main sources of errors in travel forecasts are not related to changes in preferences; rather, they are related to incorrect assumptions about input data for the forecast year, such as assumptions on population income and fuel prices (Andersson et al., 2017). Furthermore, Eliasson (2022) showed that the average daily travel time per person has been surprisingly stable over time, although communication technology and transport network speed have increased significantly. This implies that travellers today travel longer distances to access more opportunities; that is, travel time savings have been exchanged for increased access. Thus, improvements in digital communication during the pandemic are unlikely to lead to a reduction in the time spent travelling in the long term.

2.2 *Level of service and destination attraction data*

One of the major tasks of this work is to develop digital European-wide/worldwide networks for major travel modes so that level-of-service data can be generated from these networks. Level-of-service data were generated at the zone level using the transport modelling software TransCad (https://www.caliper.com/tcovu.htm). The zonal system in the long-distance model component of the Swedish national travel demand model is used for zones within Sweden, whereas the NUTS5 zone system is used to represent Europe. Outside Europe, nations are represented by zones. In total, four networks were developed: car/bus, train, air, and ferry. Networks for car/bus, train, and ferry are European-wide while the network for air is worldwide. These networks are described in detail below.

The road network in the Swedish national travel demand model is used as the base network and constitutes the network for the Swedish territory. Road network data for the rest of Europe were extracted from OpenStreetMap (www.openstreetmap.org) in 2020 and were then added to the road network in Sweden. Only motorways and primary roads extracted from OpenStreetMap were added to keep the network at a feasible size. The complete network is illustrated in Figure 3. For each link, the free-flow speed, number of lanes, and indicator of one-way roads are available attributes. Note that there are two free-flow speeds: one for the car and the other for the bus. In this study, a bus is treated as a car at a lower speed. This simplification is mainly owing to the difficulties in collecting long-distance bus line schedule data. Thus, the long-distance bus alternative modelled is more similar to a charter bus. The shortest travel time path is skimmed in TransCad to generate level-of-service zone matrices for car and buses using the free-flow speed. The procedure generates travel time and cost matrices for car and bus respectively. It is assumed that cost per kilometre is 0.18 Euro per km for private car, and 0.088 Euro per km for bus. The kilometre cost for car is divided by the party size, which we know from travel survey data. The kilometre cost for bus is calculated by assuming that the kilometre cost of bus is similar to the kilometre cost of train in the low season.

[5] https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-20-092

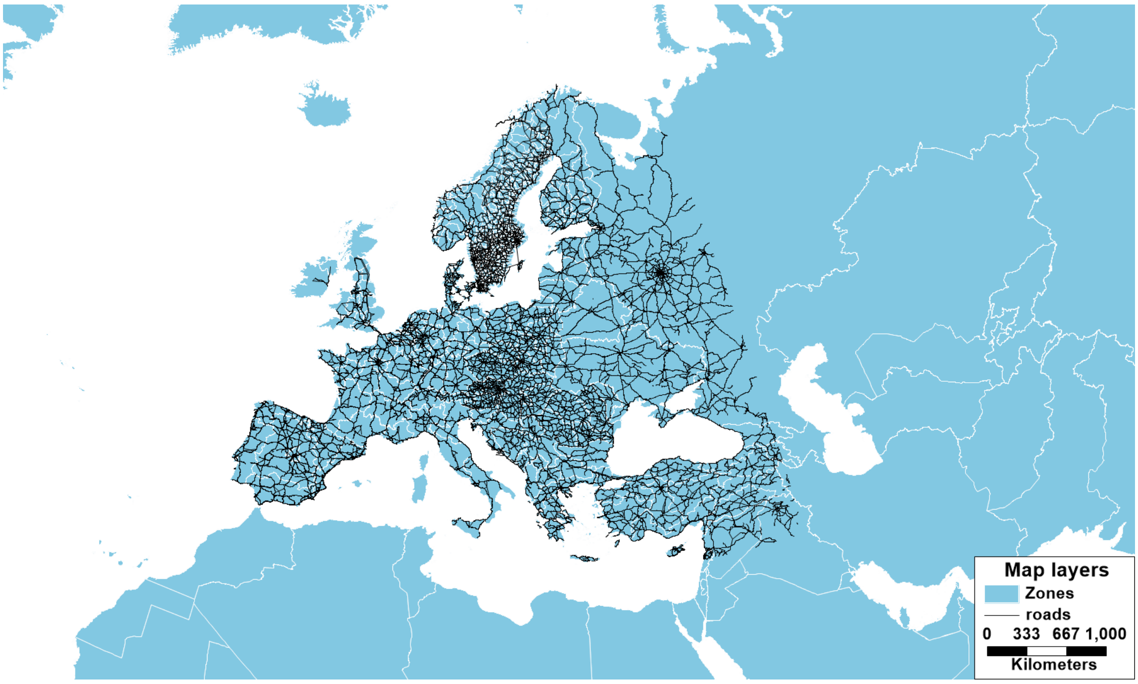


Figure 3. European wide road network used for generating level-of-service attributes.

Similar to the road network, the train network within Sweden was taken from the Swedish national travel demand model. The train network for the rest of Europe was then manually coded in TransCad, as shown in Figure 4. The travel time, travel cost, and frequency of some lines were obtained from the DB trip-planning tool (www.bahn.de) in 2020. For lines in which no ticket price data are available, a regression model is developed to impute the missing values, where the travel cost is modelled as a function of distance. To route on the train network, the road network is connected to the train network as a network for access/egress to train stations. This implicitly assumes that access and egress to train stations are carried out by car or taxi rather than by local public transport. However, this limitation is difficult to overcome because of the difficulties in collecting local public transport network data at a European level. Skimming is performed by shortest path routing, which minimizes the following generalized cost (min):

$$GC = InVehTime + 1.5 \times WaitingTime + 3 \times AccessEgressTime + 5 \times NumberTransfer \quad (1)$$

The skimmed level-of-service attributes for train include in-vehicle time, access/egress time, waiting time within and outside Sweden, and travel cost[6].

[6] Travel cost includes travel cost of access/egress by car, using a kilometre cost of 0.18 Euro/km.

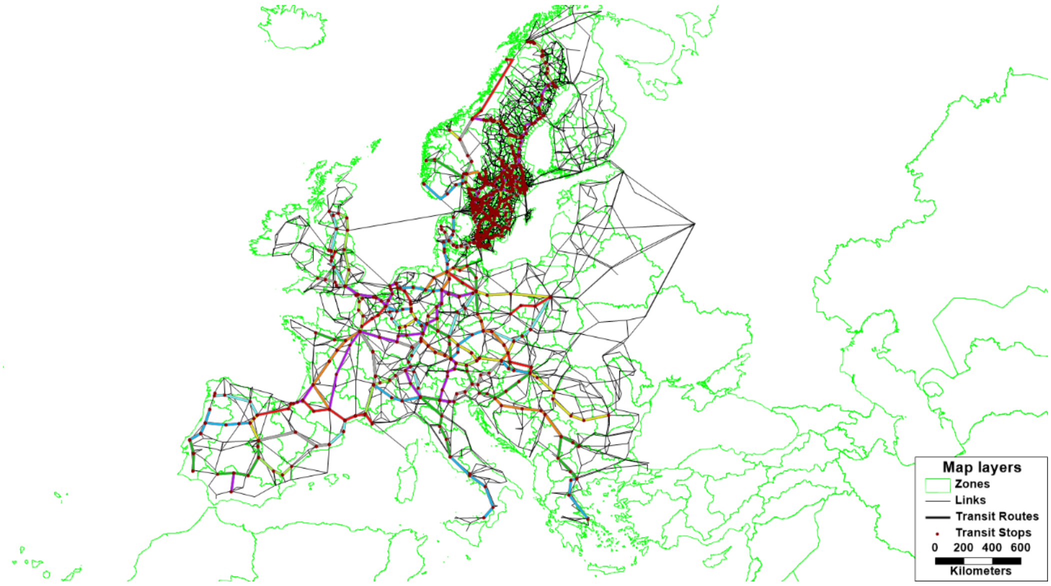


Figure 4. *An overview of the European train network.*

The network for air traffic is shown in Figure 5. The network includes Swedish domestic airlines, international airlines from Sweden to abroad, and international airlines that connect to Swedish airlines. The travel time was calculated assuming an average flying speed of 850 km/h. Price data are extracted from www.travelmarket.se for airlines from major Swedish hubs as well as major hubs in Europe. For airlines for which price data are not available, a regression model is developed to impute the price using lines with price data. The road network was connected to the air network as a network for access and egress. Thus, we assume that trips to and from the airport are conducted by car. In Sweden, around 80% of access/egress travel to airports is conducted by car/taxi (Berglund and Kristoffersson, 2020). The skimmed level service attributes for air include in-vehicle time, access/egress time, and travel cost[7]. Because there are no data available on waiting time at airports, this is not included in the model.

Figure 5. *An overview of the world-wide air network.*

[7] Travel cost includes travel cost of access/egress by private car, using the kilometre cost of 0.18 Euro/km.

The ferry network was extracted from OpenStreetMap, see Figure 6. Given the limited number of boat lines, data on frequency, ticket price, and travel time were manually collected from online sources such as www.directferries.se. The ferry network is connected to the road network. In this study, a ferry trip is defined when the route shows that the travel time on the ferry network is no shorter than that on the road network. This is to prevent trips in which most of the trips are on the road network to be categorized as trips with ferry as the main mode. The skimmed level-service attributes for ferry then include in-vehicle time, access/egress time, and travel cost. Waiting time was initially considered and tested in the model estimation but was dropped because most ferry lines have a low frequency, and it is no longer valid to assume the average waiting time to be half of the headway, as most travellers plan their departure time according to the ferry timetable.

Figure 6. An overview of the ferry lines included in the model.

Apart from the level-of-service data, attributes at the destination zone level were also collected. These variables include GDP per capita, population, employment, and number of hotel beds.

3 Model formulations

Travel demand models are formulated using the classical discrete choice theory and logit formulations (McFadden, 1974). Note that the model formulation differs somewhat from McFadden in that there is a nested logit model for mode and destination choice and a multinomial logit model for trip generation, that is, not all three levels are estimated simultaneously. Figure 7 illustrates the logit tree for private trips, whereas the model for business trips is not segmented by length of stay. It is thus assumed that travellers experience a (dis)utility of the trip, which includes both observable and unobservable parts. The unobservable part is captured by the error term, which is assumed to be Gumble distributed such that one arrives at a logit formulation.

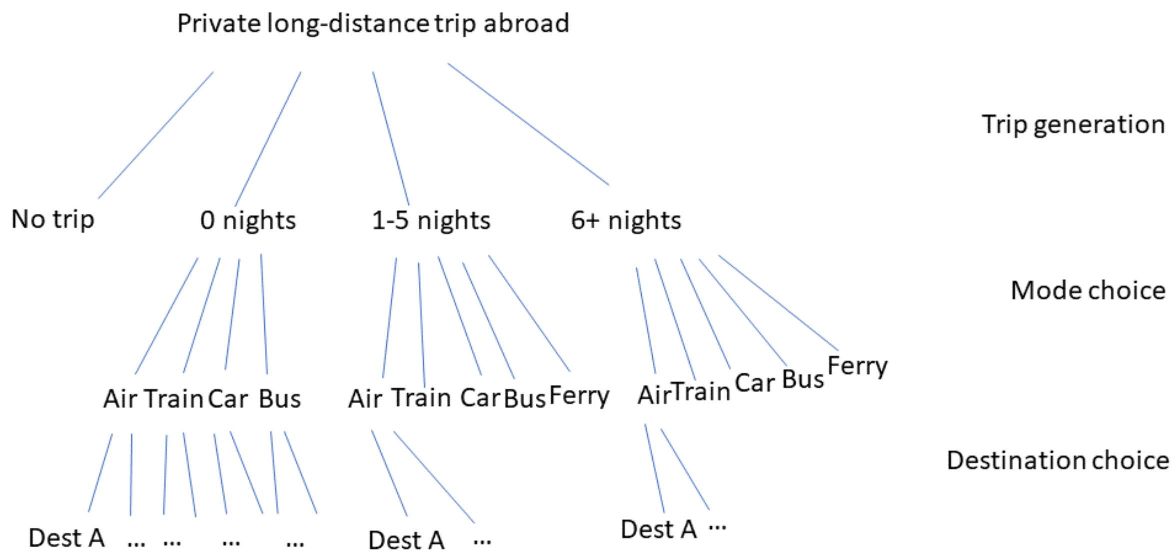


Figure 7. Overview of the logit tree structure for private long-distance international trips from Sweden.

For the mode and destination choice models, the utility equation for an alternative (mode i and destination j) is formulated as

$$U_{i,j} = ASC_i + \gamma_i I + \beta_i L_{i,j} + \delta_i D_j + \phi log(A_j) + \varphi_i + \varepsilon_{i,j} \tag{2}$$

In the above equation, ASC_i is the alternative specific constant for mode i. I is the vector of individual socioeconomic attributes. $L_{i,j}$ refers to the vector of level-of-service attributes for mode i to destination j. D_j is the vector of destination variables per capita, such as GDP and number of hotel beds per resident. A_j is a destination attraction variable (size variable) that represents the attractiveness in terms of the size and quantity of each destination zone, for which a non-linear log formulation is used (Daly, 1982), and φ_i refers to the error term at the mode level. Thus, alternatives with the same mode i will share the same error term φ_i and therefore, these alternatives are not independent of each other. $\varepsilon_{i,j}$ refers to the error term, which is unique and independent of each alternative.

The mode and destination choice models, with the utility function described in Eqs. (2) is a nested logit model, where the mode is on the upper level. The choice of model structure with mode 'above' destination or the other way around is an empirical question which is determined by the data. For consistency, the logsum parameter connecting the two levels in the model should be in the range of 0 to 1. However, in principle both mode 'above' destination and destination 'above' mode can have a logsum parameter in the range of 0 and 1.

For the trip generation model, the utility function for an alternative k is formulated as follows, where k belongs to {no long-distance international trip; daytrip; 1-5 nights, and 6+ nights} for private international travel and {no trip and trip} for international business travel. However, this study does not consider multiple long-distance international travels from the same individual. This is considered to be a possible future improvement.

$$U_k = ASC_k + \theta_k I + \mu_k T + \varphi_k logsum_{modeDestModel} + \varepsilon_k \tag{3}$$

In the above equation, I is again a vector of socio-economic variables, T is a vector of time-period variables such as Christmas, and $logsum_{modeDestModel}$ is the logsum variable calculated from the estimated mode and destination choice model. θ_k, μ_k and φ_k are associated parameter vectors. The trip generation model is a Multinomial Logit model. ε_k is an error term that is unique and independent of each alternative.

Note that the separate estimation of the mode-destination choice model and trip generation model yields a lower estimate of the standard error of the logsum parameter when compared to a joint estimation of the mode-destination choice model and trip generation model. All variables were used in an initial model specification, and insignificant variables (at 5% significance level) were removed stepwise, except key level-of-service variables and alternative specific constants (ASC). Table 2 presents the variables used for the final estimation results.

Table 2. Variable list in the final model.

Variable name	Description
Individual socio-economic variables (I)	
CarHH	Number of cars in the household
Female	Dummy if the traveller is female
Age<18	Dummy if the traveller's age<18
Age31_64	Dummy if the traveller's age is between 31 and 64
Age>64	Dummy if the traveller's age>64
Villa	Dummy if the traveller's home is a villa
HHInc<=25 TEUR	Dummy if the traveller's household income <=25 000 EURO[8]
HHInc<=70 TEUR	Dummy if the traveller's household income <=70 000 EURO
HHInc>70 TEUR	Dummy if the traveller's household income >70 000 EURO
HHIncMiss	Dummy if the traveller's household income is missing.
INDInc<=30 TEUR	Dummy if the traveller's individual income <=30 000 EURO
INDInc>30 TEUR	Dummy if the traveller's individual income >30 000 EURO
IndIncMiss	Dummy if the traveller's household income is missing
SmaChild	Number of small children (0-6 years old) in household
BigChild	Number of big children (7-18 years old) in household
NoWork	Dummy if the trip is not a work trip
Time period variables (T)	
Summer	Dummy if a trip takes place in July or August
Chris	Dummy if a trip takes place within the period 20th December and 10th January
Level-of-service variables (L)	
C_air	Travel cost for air including cost for access/egress.
C_train_p, C_train_b	Travel cost for train including cost for access/egress, for private trip and business trip respectively
C_car	Travel cost for car
C_bus	Travel cost for bus
C_ferry	Travel cost for ferry including cost for access/egress
TT_air	In-vehicle time for air
TT_train	In-vehicle time for train
TT_car	In-vehicle time for car
TT_bus	In-vehicle time for bus
TT_ferry	In-vehicle time for ferry
AC_air	Access and egress time for air
AC_train	Access and egress time for train
AC_ferry	Access and egress time for ferry
TW_train	Waiting time outside Sweden for train
Destination variables (D_j) and destination attraction variables (A_j)	
Pop (A_j)	Population in 10 000
Emp (A_j)	Number of jobs in 100 000
Beds (A_j)	Number of hotel beds in 1000
$GDP_{PerCapita}$ (D_j)	GDP per capita in 100 000 Euro
$Bed_{PerArea}$ (D_j)	Number of hotel beds per 100 residents
HolZone (D_j)	Dummy if the zone is a popular holiday destination[9]

8 Here 1 Euro= 10 SEK.
9 A list of zones that are defined as popular destination zones according to the Swedish tourist destination ranking (https://www.momondo.se/c/year-in-travel/) has been used to define these holiday zones.

Variable name	Description
Baltic (D_j)	Dummy if the zone is on Baltic Sea coast

In the destination and mode choice model estimation, the parameters associated with the level-of-service variables are initially estimated per mode, except for the cost parameters. For instance, one parameter is estimated for in-vehicle time for trains, whereas the other is estimated for in-vehicle time for air. However, the estimation yields poor (wrong sign and insignificant) results in some model segments because there are few train trips to central Europe in the travel survey material. Thus, a common parameter was estimated for in-vehicle time for all public transport modes (bus, train, air, and ferry) and a common parameter for access/egress time for air and ferry after testing various model specifications. A more detailed discussion and sensitivity analysis of this issue are presented in Section 4.4.

Another challenge is that destinations far away, such as Thailand, are attractive for Swedish tourists who seek a warmer climate in the winter, implying that travellers can put up with a long air travel time to reach a warmer destination. These destinations are often only accessible by air because of their long distances. Thus, the parameters of the destination variables are estimated for air and separated from those of the other modes.

4 Estimation results

4.1 *Mode and destination choice models*

Table 3 presents the estimation results for the mode and destination choice model. The models were estimated using tailored code written by the authors in MATLAB. Note that there are four mode-destination choice models estimated for private day trips, private 1-5 nights, private 6+ nights, and business trips. The table shows the final model. The initial model specifications were set to include all variables that are relevant and then insignificant variables (at 5% level of significance) have been removed gradually. Several model specifications were tested before the final version was selected. The t-values in the table indicate the statistical significance of the model parameters. A t-value (absolute value) larger than 1.96 means that we can say with 95% confidence that the parameter is different from zero, i.e., it has an effect in the model. A few parameters with lower levels of significance were retained in the model (shown in red in Table 3). These are either alternative specific constants that are used as calibration constants in the implementation of the model or important level-of-service variables. Apart from addition, the variable "Villa" in the utility function of alternatives with Car is also maintained, which is significant at 10% level of significance.

Table 3. Estimation results for mode and destination choice models for Swedes' long-distance international travel segmented based on trip purpose and number of nights away.

Parameter name	Variable name	Mode	Parameter value	t-value	Parameter value	t-value	Parameter value	t-value	Parameter value	t-value
			Private daytrip		Private 1-5 nights		Private 6+ nights		Business	
Destination variables										
ϕ	Log(Pop)	All					0.441	21.44		
ϕ	Log(Beds)	All	1.000	12.88	0.428	14.77				
ϕ	Log(Emp)	All							0.717	23.70
$\beta_{BedPerArea.noAir}$	BedPerArea	Car, bus,					0.092	14.48	0.059	3.65

Parameter name	Variable name	Mode	Parameter value	t-value	Parameter value	t-value	Parameter value	t-value	Parameter value	t-value
		train, ferry								
$\beta_{BedPerArea.air}$	BedPerArea	Air					0.040	21.83		
$\beta_{Hol.noAir}$	HolZone	Car, bus, train, ferry			0.295	3.06	0.257	2.12		
$\beta_{Hol.air}$	HolZone	Air			1.705	16.76	2.645	40.38		
$\beta_{HolNoWork}$	HolZone*NoWork	All	0.923	6.21						
$\beta_{GDP.noAir}$	GDPPerCapita	Car, bus, train, ferry			1.567	7.57	1.658	6.73	3.810	8.59
$\beta_{GDP.air}$	GDPPerCapita	Air			2.017	10.35			3.196	18.65
$\beta_{Baltic.ferry}$	Baltic	Ferry			3.601	3.57				
Level of service variables										
$\beta_{TT.car}$	TT_car	Car	-0.010	-8.29	-0.006	-20.36	-0.003	-16.31	-0.0080	-11.97
$\beta_{TT.PT}$	TT_bus; TT_train; TT_air; TT_ferry	Bus, train, air, ferry [10]	-0.0020	-1.55	-0.0015	-6.44	-0.00048	-3.19	-0.0039	6.61
$\beta_{AC.train}$	AC_train	Train	3* β_{TT_PT}	Fixed	-0.0027	-3.17	3*$\beta_{TT.PT}$	Fixed	-0.0783	-3.76
$\beta_{AC.airFerry}$	AC_air; AC_ferry	Air, ferry [11]	2* β_{TT_PT}	Fixed			2*$\beta_{TT.PT}$	Fixed	-0.0093	-6.75
$\beta_{TW.train}$	TW_train	Train			-0.033	-6.77	-0.011	-3.75		
$\beta_{logCost}$	Log(C_XX)	All	-1.482	-4.12						
β_{Cost}	C_XX	All			-0.007	-8.66			-0.0038	-4.67
$\beta_{logCostLowMedInc}$	Log(C_XX) * HHInc<=70TEUR	All			-0.369	-2.60				
$\beta_{logCostYoung}$	Log(C_XX) * Age<18	All			-0.006	-2.33				
$\beta_{CostLowMedInc}$	C_XX * HHInc<=70TEUR	All					-0.0018	-5.30		
$\beta_{CostHighInc}$	C_XX * HHInc>70TEUR	All					-0.0015	-4.21		
$\beta_{CostIncMiss}$	C_XX * HHIncMiss	All					-0.0017	-4.40		

[10] Ferry is not available as mode for private daytrips and business trips.
[11] Ferry is not available as mode for private daytrips and business trips.

Parameter name	Variable name	Mode	Parameter value	t-value	Parameter value	t-value	Parameter value	t-value	Parameter value	t-value
$\beta_{LogCostLowMedIn}$	Log(C_XX) * INDInc <=30 TEUR	All							-0.6302	-1.70
Socio-economic variables										
$\beta_{CarHH.car}$	CarHH	Car	0.429	2.63	0.656	8.27	0.265	3.45	0.372	3.26
$\beta_{Female.car}$	Female	Car	-0.794	-3.12	-0.451	-3.68	-0.489	-4.05		
$\beta_{Villa.car}$	Villa	Car					0.264	1.77		
$\beta_{Age>64.train}$	Age>64	Train	1.053	2.77						
$\beta_{Age>64.bus}$	Age>64	Bus			0.469	2.33	1.102	5.04		
$\beta_{Age>64.air}$	Age>64	Air			-0.858	-5.34				
Alternative specific constants										
ASCbus	/	Bus	-1.776	-4.96	-1.726	-9.12	-2.511	-7.35	-1.767	-5.74
ASCtrain	/	Train	-0.920	-2.21	-0.470	-2.52	-2.716	-9.35	0.341	0.98
ASCair	/	Air	-1.351	-3.93	-1.497	-6.24	-0.363	-0.70	-0.007	-0.02
ASCferry	/	Ferry			-0.937	-5.07	-1.314	-2.56		
Logsum										
Logsumdestination	/	All	0.441	8.16[12]	0.779	3.34	0.674	2.05	0.786	1.71
Model information										
Number of observations			324		1348		1889		717	
Number of observations choosing car			180		515		354		140	
Number of observations choosing bus			21		136		91		26	
Number of observations choosing train			36		86		16		30	
Number of observations choosing air			87		521		1419		521	
Number of observations choosing ferry					90		9			
Number of parameters			12		22		21		15	
Log-likelihood			-1224.5		-6469.3		-9813.1		-3454.2	
Log-likelihood all parameters=0			-1574.7		-9086.3		-13520.0		-4969.4	
McFadden rho			0.222		0.288		0.274		0.305	

The logsum parameters are all within the range of 0 and 1, indicating that the nested logit structure with mode at the upper level is valid. The models with the alternative nesting structure, that is, destination over mode in the nested logit model structure, are also tested, and the results with the same model parameter setting are presented in the appendix. When comparing the final log-likelihood, the mode over destination nesting structure results in a higher log-likelihood in models of private 1-5 nights, private 6+ nights, and the model of business, thus confirming the validity of adopting the mode over destination nesting structure. For the private daytrip model, the destination over mode nesting structure results in a higher log-likelihood, but the logsum parameter turns out to be higher than 1, although it is not statistically different from 1.

The number of hotel beds is the destination attraction variable in the models of private day trips and private 1-5 nights, while population and employment are used in the model of private 6+ nights and business trips, respectively. The parameters of destination attraction variables are positive, showing that the quantity in terms of the number of hotel beds, population, and employment has a positive effect on attracting travellers to given destination zones. The number of hotel beds per 100 residents was also included in the models of private 6+ nights and business

[12] T-values for logsum variables refer to test of parameter value being equal to 1.

trips in a linear form. If a destination zone is a popular holiday destination (often in warmer climates), it has a stronger effect on air travellers than on travellers using other modes. This captures the fact that Swedish tourists travel far away from warmer countries, such as Thailand, in winter. It also makes sense that this effect is only prominent in the models of private 1-5 nights and 6+ nights, as these trips are most likely tourist trips. Popular holiday destination zones also have a positive effect on attracting private day trips if they are not work trips. GDP is also an important attraction factor for private to 1-5 nights, 6+ nights and business trips. The Baltic dummy is introduced to capture the fact that ferry trips from Sweden are likely to have destinations on the Baltic coast, but this is only found on private 1-5 nights since this is the only segment with sufficient observations of ferry trips.

When it comes to level-of-service variables, all travel time and cost parameters are negative, as expected. The disutility of travel time for cars is generally higher than that of public transport, which is expected (except for air travel) because travel time on, for example, the train can be used for recreation or work activities. Parameters for access/egress time were fixed in relation to in-vehicle time in the models of private day trips and private 1-5 nights, as model results when they were not fixed gave non-intuitive results (often the parameter values were too high compared to in-vehicle time). Waiting time outside Sweden by train has a strong negative effect. The disutility of waiting time outside Sweden is more than ten times larger than the disutility of in-vehicle time. This simply reflects the fact that train travellers are unwilling to transfer trains abroad. Different cost formulations are tested in the models. The cost parameters were differentiated by sociodemographic variables to represent the possible heterogeneity in cost sensitivity. The model for private six +nights identifies different cost parameters for travellers with household income >700tkr and <=700tkr, which shows that travellers with household income <=700tkr have a higher cost sensitivity. However, the likelihood ratio test shows that the difference between the income groups is not statistically significant. For private 1-5 nights, those with household income lower than 70 TEUR and age <18 have a higher cost sensitivity in the segment of private 1-5 nights. Those with individual incomes lower than 30 TEUR have a higher cost sensitivity in the segment of business trips.

Looking into the effects of socioeconomic variables, the number of cars in a household is, as expected, a strong factor for choosing cars in all trip segments. Female travellers are less likely to travel abroad in all private segments. Those living in Villa are more likely to travel by car abroad, but only in the private six + nights segment. Pensioners are found to take trains more often for private day trips, while taking buses more often on private 1-5 nights and 6+ nights. Pensioners are less likely to take air on private to 1-5 nights.

4.2 *Trip generation*

Two trip generation models were estimated for private and business trips. These models take the form of Multinomial Logit models. For the model of private trips, the available alternatives are the segments of the number of nights used in mode and destination choice models: taking no trips, private daytrips, private 1-5 nights, or private 6+ nights. For the business trip model, the available alternatives take no trips or conduct business trips. Individuals who have made more than one trip during the survey period (last three months) are treated as different individuals in the trip-generation model. Owing to the limited sample size, the models do not consider the segmentation of the number of trips per traveller. Table 4 presents the estimation results.

Table 4. Results of trip generation model estimation for private and business trips respectively.

Parameter name	Variable name	Number of nights away	Parameter	t-value	Parameter	t-value
			Private trips		Business trips	
Socio-economic variables and time/period variables						
$\beta_{LowInc.0}$	HHInc<=25TEUR	No trip	0.671	9.29		

Parameter name	Variable name	Number of nights away	Parameter	t-value	Parameter	t-value
			Private trips		Business trips	
$\beta_{LowMedInc.0}$	INDInc<=30TEUR	No trip			1.207	5.34
$\beta_{IncMiss.1}$	HHIncMiss	daytrip	-0.885	-6.69		
$\beta_{Age<18.2}$	Age<18	1-5 nights	-5.026	-4.68		
$\beta_{HighInc.2}$	HHInc>70TEUR	1-5 nights	0.587	8.99		
$\beta_{IncMiss.2}$	HHIncMiss	1-5 nights	-0.418	-5.97		
$\beta_{Summer.2}$	Summer	1-5 nights	0.137	1.79		
$\beta_{Chris.2}$	Chris	1-5 nights	-0.298	-2.17		
$\beta_{SmaChild.2}$	SmaChild	1-5 nights	-0.144	-2.96		
$\beta_{BigChild.2}$	BigChild	1-5 nights	-0.119	-2.72		
$\beta_{Female.3}$	Female	6+ nights	0.162	2.97		
$\beta_{Age>64.3}$	Age>64	6+ nights	-0.154	-2.44		
$\beta_{IncMiss.3}$	HHIncMiss	6+ nights	-0.422	-7.33		
$\beta_{Summer.3}$	Summer	6+ nights	0.687	11.85		
$\beta_{Chris.3}$	Chris	6+ nights	0.485	5.39		
$\beta_{SmaChild.3}$	SmaChild	6+ nights	-0.196	-4.31		
$\beta_{BigChild.3}$	BigChild	6+ nights	0.060	1.72		
$\beta_{CarHH.4}$	CarHH	Business trips			0.181	4.53
$\beta_{Female.4}$	Female	Business trips			-1.019	-11.21
$\beta_{Age31\_64.4}$	Age31_64	Business trips			0.729	4.64
$\beta_{Age>64.4}$	Age>64	Business trips			-1.127	-4.67
$\beta_{HighInc.4}$	INDInc>30 TEUR	Business trips			1.156	8.01
$\beta_{Summer.4}$	Summer	Business trips			-0.837	-5.68
$\beta_{Chris.4}$	Chris	Business trips			-0.945	-4.01
Alternative specific constant						
ASC_1	/	daytrip	-5.09	-31.64		
ASC_2	/	1-5 nights	-3.22	-69.53		
ASC_3	/	6+ nights	-9.17	-5.70		
ASC_4	/	Business trips			-4.740	-23.81
Logsum variables from mode-destination choice models						
$Logsum_{daytrip}$	/	daytrip	0.353	4.94[13]		
$Logsum_{6+Night}$	/	6+ nights	1.000	3.80		
Model information						
Number of observations			45559		39996	
Number of observations that choose no trip			41843		39267	
Number of observations that choose daytrip			397			
Number of observations that choose 1-5 nights			1425			
Number of observations that choose 6+ nights			1894			
Number of observations that choose business trips					729	
Number of estimated parameters			21		9	
Log-likelihood			-16082.1		-3066.4	
Log-likelihood when all parameters=0			-63158.2		-27723.1	
McFadden rho			0.745		0.889	
Adjusted McFadden rho			0.745		0.889	

The main explanatory variables considered in the trip generation models are socio-economic and time-period variables. The logsum variables from the mode and destination choice models were included to capture the effect of accessibility on the decision to conduct long-distance international trips. Household income is used in the model of private trips, whereas individual income is used for business trips. It was found that low income is an important explanatory factor that contributes to not conducting long-distance international trips, which is expected. High income is a positive factor for conducting private 1-5 nights trips and business trips. It is perhaps a bit surprising that

[13] T-values for logsum variables refer to test of parameter value being equal to 0.

no significant impact of high income on private 6+ nights trips was found. Private 6+ night trips are most likely holiday trips to destinations outside Europe, such as Thailand, where commodity prices are affordable, even for medium- and low-income households. Teenagers (age<18) are less likely to conduct private 1-5 nights trips, while pensioners (age>64) are less likely to conduct private 6+ nights trips and logically conduct business trips. Female travellers are more likely to conduct private 6+ nights trips and are less likely to conduct business trips. The number of cars in a household was positively correlated with the likelihood of conducting business trips. The number of children in a household is another important explanatory factor in trip generation models. Travellers with small children (0-6 years old) are less likely to conduct private 1-5 nights and private 6+ nights trips, while travellers with large children (>6 years old) are less likely to conduct private 1-5 nights trips, but more likely to conduct private 6+ nights trips. As expected, there are more private six + nights trips but fewer business trips in summer and Christmas, as these are the time periods when travellers from Sweden have long holidays or visit their homelands.

The logsum variables calculated from the mode-destination choice models for private day trips and 6+ night trips were found to be significant, suggesting that better accessibility is associated with a higher likelihood of conducting these trips. However, accessibility was not found to contribute to the likelihood of conducting private 1-5 nights trips or business trips. In McFadden's three-level nested logit model, which is estimated sequentially, the logsum variables act as the nesting coefficient that shows whether there is a nested structure between the trip generation and the lower-level (mode-destination choice), and a statistical difference of the logsum variable against 1 would suggest that such a nested structure exists. In this study, the logsum variable in the day trip model is significantly different from 1 (t-value is -9.05), whereas the logsum variable in the 6+ nights model is not significantly different from 1 (t-value is -0.04).

4.3 *Changes in level of service attributes for train and resulting elasticities*

The estimated models are planned to be implemented in the *Sampers* model, among other things, to evaluate the potential impacts of high-speed rail on the demand for long-distance international travel. Thus, the elasticities for the level-of-service attributes for trains are derived to provide a first look at the magnitudes of the impacts. The elasticity shows the unit percentage change of the likelihood for a given mode, given a unit percentage change of a level of service attribute for train. The following scenarios were adopted for the elasticity calculations: 10% increase in travel cost by train, 10% decrease in train in-vehicle time, and 10% decrease in waiting time outside Sweden. Elasticity is defined in (4) as the percentage difference in probability of each mode in the scenario case and baseline, divided by the percentage change in the attribute, i.e., $\frac{v_{scenario}-v_{baseline}}{v_{baseline}} = 10\%$. For each observation in the estimation data (number of observations in Table 3), the model was applied to calculate the choice probabilities for each observation. The average elasticity of each observation is reported.

$$Elasticity = \frac{\left(\frac{L_{scenario} - L_{baseline}}{L_{baseline}}\right)}{\left(\frac{v_{scenario} - v_{baseline}}{v_{baseline}}\right)} \quad (4)$$

For private trips, the elasticities are calculated for each number of night segments, and then the aggregated elasticities are reported. The results are presented in Tables 5 and 6.

Table 5. Elasticity estimates of level-of-service attributes for train for private trips.

		Car	Bus	Train	Air	Ferry	Total
Baseline	Likelihood	2.80%	0.60%	0.33%	4.45%	0.16%	8.34%
10% increase travel cost by train	Likelihood	2.80%	0.60%	0.32%	4.46%	0.16%	8.34%
	Elasticity	0.025	0.028	-0.479	0.012	0.033	-0.002

10% decrease in train in-vehicle time	Likelihood	2.79%	0.60%	0.34%	4.45%	0.16%	8.34%
	Elasticity	-0.014	-0.017	0.291	-0.008	-0.025	0.001
10% decrease in waiting time outside Sweden	Likelihood	2.79%	0.60%	0.34%	4.45%	0.16%	8.34%
	Elasticity	-0.012	-0.017	0.302	-0.007	-0.023	0.003

Table 6. Elasticity estimates of level-of-service attributes for train for business trips.

		Car	Bus	Train	Air	Total
Baseline	Likelihood	0.468%	0.076%	0.068%	1.210%	1.823%
10% increase travel cost by train	Likelihood	0.469%	0.076%	0.065%	1.213%	1.823%
	Elasticity	0.021	0.020	-0.475	0.017	0.000
10% decrease in train in-vehicle time	Likelihood	0.466%	0.076%	0.075%	1.205%	1.823%
	Elasticity	-0.038	-0.041	1.069	-0.043	0.000
10% decrease in waiting time outside Sweden	Likelihood	0.468%	0.076%	0.068%	1.210%	1.823%
	Elasticity	0.000	0.000	0.000	0.000	0.000

In the baseline scenario, the overall likelihood of having a long-distance international private trip was 8.34%, where nearly half, 4.45%) by air. Only 0.33% is by train. A 10% increase in the travel cost by train reduces the likelihood of taking train to 0.32%, corresponding to an elasticity of -0.479, whereas the cross elasticity ranges from 0.025 to 0.033. The elasticity of the decreased train in-vehicle time was 0.291, whereas that of the decreased train waiting time outside Sweden was 0.302. The overall elasticity of conducting a private trip was small, range between -0.002 and 0.003. This is plausible because many private trips are holiday trips to popular tourist resorts where the driver underlying the "decision to travel" is likely not an improvement in accessibility.

When it comes to business trips, the elasticity of the increased train travel cost is -0.475, which is similar to that of private trips. The elasticity of the decreased train in-vehicle time is much higher than that of private trips (1.069), suggesting that business travellers are more inclined to use high-speed trains because of travel time savings. Decreasing waiting time outside Sweden has no effect on business travellers because the corresponding variable is not significant and is not included in the mode-destination choice model for business trips. Because the logsum variable is not significant and is not included in the trip generation model for business trips, changes in the level-of-service variables will not result in a change in the overall likelihood of business trip generation.

The elasticity value estimated in this study is in general comparable with what is found in Rich and Mabit (2012) which is one of few papers that estimated elasticity values for long-distance international travel. In that study, the elasticities of train travel cost were -1.000/-0.623[14] for private[15] trips, -0.459/-0.289 for holiday trips, and -0.511/-0.306 for business trips. The elasticities of train in-vehicle time in that study were -0.378/-0.453 for private trips, -0.186/-0.185 for holiday trips, and -0.293/-0.275 for business trips. Note that the scenario of increased travel time was tested in this study. The main difference from the results of this study is that the elasticities of train in-vehicle time for business trips are remarkably higher in this study than in Rich and Mabit (2012). However, it is lower than the elasticity reported in the Swedish domestic long-distance demand model, 1.59 (WSP Analysis & Strategy, 2011).

4.4 *Sensitivity analysis of a generic parameter of travel time for all modes*

[14] The values correspond to elasticities for trips with distance below and above 600 km.
[15] Private trips are here private trips other than holiday trips, which is a different segmentation of the model compared to the estimations presented in this paper.

In this section, a sensitivity analysis is included, which presents the results of having a generic parameter of travel time for all modes. One of the challenges in our mode-destination models is the difficulty in obtaining a negative parameter of travel time for air. The reason behind this is the fact that many Swedish people travel to other destinations such as Thailand for summer holidays, which makes it difficult for the model to distinguish destination attractions and travel time because they are correlated due to the geographical location of Sweden in the world. This leads to difficulty in obtaining a reasonable VOT for air. In the previous sections of this paper, the parameter of air travel time was set to the same as that of bus, train, and ferry. However, this significantly lowers the VOT of air compared with the VOT of a car, which makes the VOT estimations less comparable to other existing empirical investigations. Thus, a sensitivity analysis was conducted where only one generic parameter of travel time for all modes was considered; thus, only a generic VOT was obtained, which was not differentiated by modes. Table 7 shows the estimation results. Estimation results with one generic parameter of travel time for all modes.

Table 7. Estimation results with one generic parameter of travel time for all modes.

Parameter name	Variable name	Mode	Parameter value	t-value	Parameter value	t-value	Parameter value	t-value	Parameter value	t-value
			Private daytrip		Private nights 1-5		Private nights 6+		Business	
Destination variables										
ϕ	Log(Pop)	All					0.434	21.33		
ϕ	Log(Beds)	All	0.954	12.75	0.395	13.68				
ϕ	Log(Emp)	All							0.751	24.85
$\beta_{BedPerArea.noAir}$	BedPerArea	Car, bus, train, ferry					0.096	15.08	0.060	3.68
$\beta_{BedPerArea.air}$	BedPerArea	Air					0.039	21.41		
$\beta_{Hol.noAir}$	HolZone	Car, bus, train, ferry			0.315	3.24	0.108	0.91		
$\beta_{Hol.air}$	HolZone	Air			1.801	17.75	2.640	21.44		
$\beta_{HolNoWork}$	HolZone*NoWork	All	1.073	7.35						
$\beta_{GDP.noAir}$	GDPPerCapita	Car, bus, train, ferry			1.690	8.46	1.955	9.11	4.103	10.50
$\beta_{GDP.air}$	GDPPerCapita	Air			2.016	10.54			3.281	19.03
$\beta_{Baltic.ferry}$	Baltic	Ferry			3.785	3.79				
Level of service variables										
β_{TT}	TT_car; TT_bus; TT_train; TT_air; TT_ferry	All[16]	-0.0071	-6.66	-0.0039	-17.98	-0.0017	-13.02	-0.0066	-12.98

[16] Ferry is not available as mode for private daytrips and business trips.

Parameter name	Variable name	Mode	Parameter value	t-value	Parameter value	t-value	Parameter value	t-value	Parameter value	t-value
$\beta_{AC.train}$	AC_train	Train	3* β_{TT_PT}	Fixed	-0.0022	-2.47	3*$\beta_{TT.PT}$	Fixed	-0.072	-3.34
$\beta_{AC.airFerry}$	AC_air; AC_ferry	Air, ferry [17]	2* β_{TT_PT}	Fixed			2*$\beta_{TT.PT}$	Fixed	-0.010	-7.25
$\beta_{TW.train}$	TW_train	Train			-0.029	-5.73	-0.008	-2.74		
$\beta_{logCost}$	Log(C_XX)	All	-0.958	-2.80						
β_{Cost}	C_XX	All			-0.005	-6.35			-0.0014	-1.74
$\beta_{logCostLowMedInc}$	Log(C_XX) * HHInc<=70TEUR	All			-0.076	-0.51				
$\beta_{logCostYoung}$	Log(C_XX) * Age<18	All			-0.004	-1.70				
$\beta_{CostLowMedInc}$	C_XX * HHInc<=70TEUR	All					-0.0002	-0.66		
$\beta_{CostHighInc}$	C_XX * HHInc>70TEUR	All								
$\beta_{CostInCMiss}$	C_XX * HHIncMiss	All								
$\beta_{LogCostLowMedInc}$	Log(C_XX) * INDInc <=30 TEUR	All							-0.655	-1.32
Socio-economic variables										
$\beta_{CarHH.car}$	CarHH	Car	0.417	2.54	0.562	7.35	0.265	3.46	0.372	3.29
$\beta_{Female.car}$	Female	Car	-0.836	-3.25	-0.454	-3.78	-0.488	-4.05		
$\beta_{Villa.car}$	Villa	Car					0.274	1.84		
$\beta_{Age>64.train}$	Age>64	Train	0.862	2.22						
$\beta_{Age>64.bus}$	Age>64	Bus			0.449	2.23	1.107	5.06		
$\beta_{Age>64.air}$	Age>64	Air			-0.971	-6.06				
Alternative specific constants										
ASCbus	/	Bus	-0.624	-1.63	-0.100	-0.53	-0.954	-4.19	-0.481	-1.56
ASCtrain	/	Train	0.670	1.35	0.895	3.51	-1.430	-3.25	0.1.347	3.00
ASCair	/	Air	-1.346	-4.00	-2.551	-7.67	0.225	0.53	0.342	1.06
ASCferry	/	Ferry			0.343	1.94	0.775	0.84		
Logsum										
Logsumdestination	/	All	0.702	3.23 [18]	1.234	-2.10	0.811	0.90	0.857	1.06
Model information										
Number of observations			324		1348		1889		717	
Number of observations choosing car			180		515		354		140	
Number of observations choosing bus			21		136		91		26	
Number of observations choosing train			36		86		16		30	
Number of observations choosing air			87		521		1419		521	
Number of observations choosing ferry					90		9			

[17] Ferry is not available as mode for private daytrips and business trips.
[18] T-values for logsum variables refer to test of parameter value being equal to 1.

Parameter name	Variable name	Mode	Parameter value	t-value	Parameter value	t-value	Parameter value	t-value	Parameter value	t-value
Number of parameters			11		21		18		15	
Log-likelihood			-1236.1		-6563.5		-9886.7		-3471.4	
Log-likelihood all parameters=0			-1574.7		-9086.3		-13520.0		-4969.4	
McFadden rho			0.215		0.278		0.269		0.301	
Adjusted McFadden rho			0.208		0.275		0.267		0.299	
P-value for Likelihood Ratio Test compared to Table 3			1.46e-6		0.00		0.00		4.49e-9	

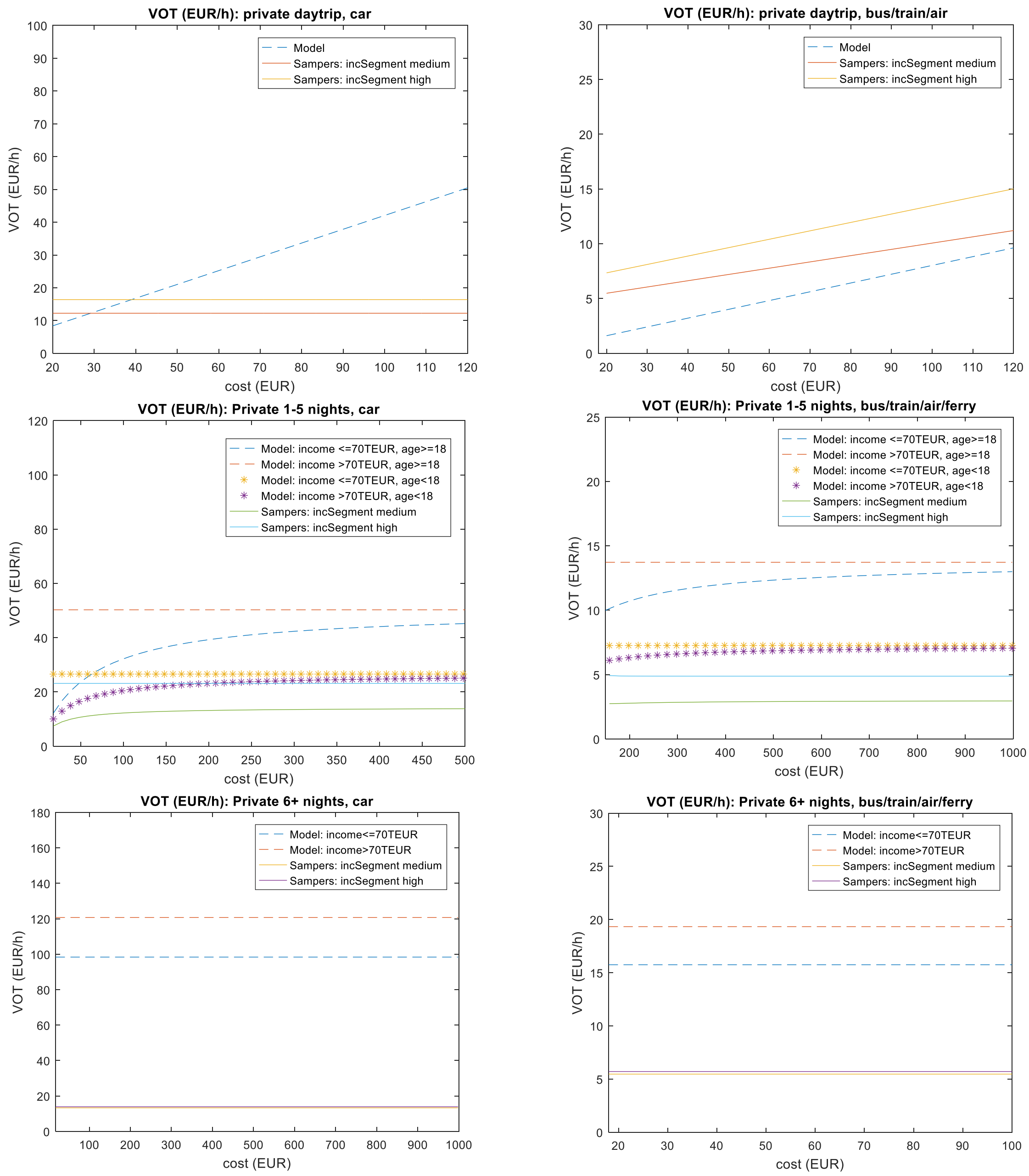
The models presented above can be considered as reduced models compared to those in Table 3. Thus, the likelihood ratio test (LRT) was first investigated (the last row in Table 7). The LRT results indicate that separating the parameters for car and other modes yields a statistically better model fit for all segments tested because all the P-values are small. The estimated generic travel time parameter β_{TT} lies between $\beta_{TT,car}$ and $\beta_{TT,PT}$ in Table 3 for all the segments. Most parameters were relatively unchanged or changed only marginally. One parameter that changes more substantially is the travel cost, especially in the segment of private 6+ days, where no significant cost parameters can be obtained. The alternative specific constants (ASCs) have also changed substantially because fixing the travel time parameters for all modes would indicate different logsum values for different modes; thus, ASCs need to change accordingly. Another issue that arises from the model estimation in this sensitivity analysis is that the logsum parameter is larger than 1 for the segment of private to 1-5 nights. This indicates that the nested structure is no longer valid with the generic travel time parameter for that segment.

5 Discussion

In this section, value of time (VOT) estimates are derived from the estimated parameters of the in-vehicle time and travel cost in each model (Table 3). However, it should be noted that there are issues when using revealed preference (RP) data for VOT estimation. A large European meta-analysis showed that VOT estimates based on RP data were consistently higher than those based on stated preference (SP) data (Wardman et al., 2016). All reasons for this have not yet been established, but some contributors have been identified. For example, Varela et al. (2018) showed that in RP data, the measurement error is often larger in the cost variable than in the travel time variable, leading to attenuation bias in the cost parameter, and thus to a higher VOT.

The VOT results are shown in Figure 8. It is important to note that we cannot derive separate VOT for buses, trains, air, and ferries; rather, we estimate a joint public transport VOT. This is because the in-vehicle time parameters are the same for these modes (β_(TT.PT) in Table 3). An attempt was made to estimate separate in-vehicle time parameters for different public transport modes, but the results showed that parameters of in-vehicle time for air turned out to be positive in models for private 1-5 nights, private 6+ nights, and business. A positive in-vehicle-time parameter is not consistent with the random utility theory and cannot be used. The estimation result could be a special case for Sweden because many Swedes travel to tropical countries for their holidays, and thus the chosen destination is far away, which drives the travel time parameters for air to be positive. Although a holiday zone dummy was introduced in the model to capture the popular destination attraction effect, the correlation of longer air travel time and popular holiday countries persists. The existing domestic long-distance model in the Swedish National Transport Model Sampers also adopted the same setting, that is, estimating a joint in-vehicle time parameter for bus, train, air, and ferry, which facilitates comparison.

As discussed in Section 4.4, one problem associated with the unavailability of estimating a separate travel time parameter for air is that the model cannot reflect the fact that the VOT of air is higher than that of other public transport modes. This strongly affects how the VOT of public transport should be interpreted in different trip segments; since then, it has also been related to the share of air travel in each segment. In Section 4.4, this issue is further investigated by testing a model in which only a generic parameter of travel time for air is estimated, that is, not differentiating modes. Thus, it is not recommended to compare the VOT of public transport with that of cars because of the unavailability of estimating a separate travel time parameter of air. However, a comparison between VOT in international travel and that in domestic travel for a given mode is still relevant.



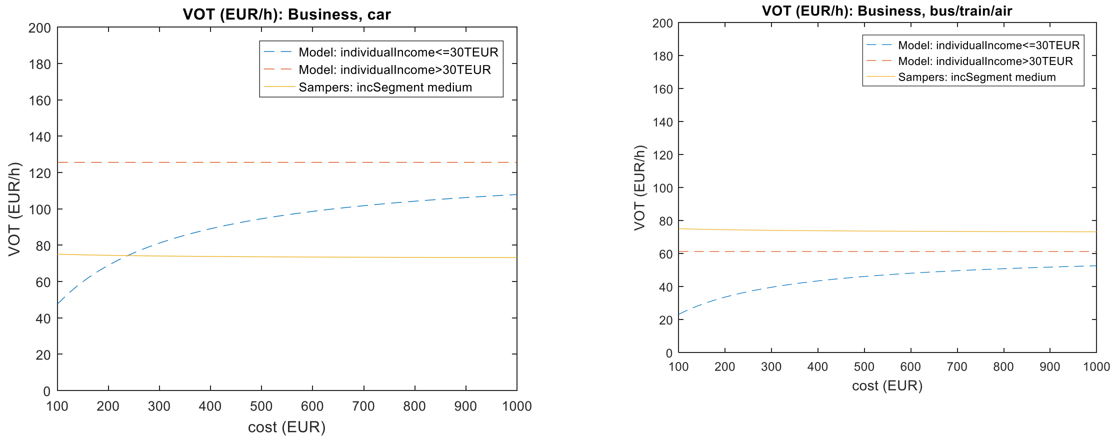


Figure 8. Comparison of value of time between the Swedish domestic long-distance model Sampers and the models developed in this study.

In some of the models (not private 6+ nights), cost has a non-linear (log) formulation, which means that a cost damping effect (Daly, 2010) has been captured, that is, the phenomenon where the sensitivity to cost decreases with distance, which leads to VOT increasing with travel cost.

In the private daytrip segment, the VOT for car in long-distance international trips is higher than that derived from the domestic long-distance trip model when the travel cost is higher than 40 EUR. However, it is difficult to draw the conclusion that VOT for car for long-distance international trips is higher than that for domestic long-distance trips because a linear cost parameter specification is used in *Sampers*. On the other hand, the VOT for public transport for long-distance international trips is lower than that for domestic long-distance trips, where the cost damping effect is captured in both. It should be noted that there are few observations in the private daytrip segment, especially for observations taking public transport; thus, the estimated VOT in this segment needs to be interpreted with caution. In the private 1-5 nights segment, VOT differs depending on the traveller's household income and age. The VOT for long-distance international trips was higher than that for domestic long-distance trips for both car and public transport modes. For the private 6+ nights segment, VOT in long-distance international trips is also higher, and the discrepancy is larger than that found in the private 1-5 nights segment. The VOT of long-distance international trips for cars was much higher than that for domestic long-distance trips. The data show that the average distance of car for the observed trips is 448 km for private daytrips, 594 for private 1-5 nights and 1063 km for private 6+ nights. More than 30% of car trips in private 6+ nights category had a distance longer than 1500 km. Thus, it is relevant to consider why these trips are made by car rather than by more time-efficient modes, such as air. It is likely that other factors caused them to choose car, such as the need to carry heavy luggage or sightseeing along the journey. This is an issue when using RP data for VOT estimation, as there may be many unknown factors of the observed journey that we cannot control for, which may bias the travel time and cost parameters if unknown factors are correlated to travel time and cost.

In the business trip segment, the VOT for cars for long-distance international trips is higher than that for domestic trips, whereas a reversed trend is found for public transport modes. Business travellers are more likely to travel frequently and thus have less burden travelling abroad. This means that there is no substantial difference between international travel and domestic travel for business travellers; thus, the differences in VOT between international and domestic long-distance travel are smaller compared with private travellers.

Compared to the existing literature, Mabit et al. (2013) conducted a stated-preference survey to investigate the VOT of international travel over the Fehmarn Belt. In that study, estimated VOT for car for business travellers is 15.9 Euro/h and 9.5 Euro/h for non-business travellers. These estimates were lower than the values estimated in the present study. The VOT for cars for business travellers ranges between 45 and 130 Euro/h, and VOT for cars for private travellers ranges between 10 and 120 Euro/h, depending on the cost. The estimated VOTs in the bus and rail are approximately 7.5 Euro/h and the VOT in air is 27.9 Euro/h with a rather large variation (1.2 - 119.4 Euro/h) in Mabit et al. (2013). However, the estimates are not differentiated by the trip purpose. This makes the comparison to this study difficult because it differentiates trip purpose and number of nights away, but not public transport modes. Furthermore, the model in Mabit et al. (2013) is a mode-choice model that does not include destination choice. This means that different choice contexts are used and the derived VOT can be sensitive to that; for instance, air travellers to destinations outside Europe cannot be evaluated in the mode choice model context because air is the only viable mode for them. In general, VOT estimates for public transport in this study ranged between 5 and 70 Euro/h. The US long-distance travel model (FHWA, 2018) derives the travel time and cost parameters from the California statewide model, which makes it difficult to compare with the VOT from this study. The UK long-distance model (Rohr et al., 2013), incorporated a combined SP-RP approach, in which the VOT for business trips was successfully identified using SP data (see Table 8). The VOT for car for business trips yields 36.9 and 73.8 Euro/h for low (<30 k£/pa) and medium/high income (30+ k£/pa) groups. These estimates are considerably lower than our estimates, 60-105 and 125 Euro/h for the low (<30+TEUR) and medium/high income (30+TEUR) groups. However, for public transport modes, the VOT from our study ranges between 20-45 Euro/h for the low (<30+TEUR) and 60 Euro/h for the medium/high income (30+TEUR) groups, which is much closer to the estimates in Table 8.

Table 8. VOT for business trips derived from UK long-distance model (Rohr et al., 2013).

	Business (<30 k£/pa)	Business (30-50 k£/pa and 50+k£/pa)
VOT car (euro/h)	-0.0104/-0.0002*0.71=36.9	-0.0104/-0.0001*0.71=73.8
VOT rail (euro/h)	-0.0077/-0.0002*0.71=24.7	-0.0077/-0.0001*0.71=54.7
VOT air (euro/h)	-0.0082/-0.0002*0.71=29.2	-0.0082/-0.0001*0.71=58.2

6 Conclusions

Long-distance international travel, although low in number of trips compared to regional travel, contributes significantly to the total distance travelled and thus externalities from the transport sector. Despite the abundant literature analysing tourist demand and long-distance travel, most developed models are direct demand models that focus on a specific mode or specific origin-destination pair. Surprisingly few existing large-scale disaggregated travel demand models include model components for long-distance international trips. The absence of such disaggregated models indicates a lack of ability to calculate modal shift for long-distance international travel for large infrastructure investments, such as high-speed rail.

In this study, a model component for long-distance international travel was developed for Swedish national travel demand model *Sampers*. Trip generation, mode, and destination choice are modelled using multinomial and Nested Logit models, respectively. Swedish national travel survey data were used to observe long-distance international travel. European networks for road, train, and ferry and a worldwide network for air, are developed at a reasonable level of detail. Models for private and business trips were developed, where those for private trips were further segmented by the number of nights away.

The estimation results reveal the effects of individual socioeconomic variables, level-of-service attributes, and destination variables on Swedes' long-distance international travel demand. To capture the effect of Swedish tourist travel to far-away warmer countries, such as Thailand, in the winter, a dummy for holiday zones is introduced. In contrast to common practice, where the parameters of destination variables are the same across travel modes, the holiday dummy is separately estimated for air and other modes. The estimation results show that the parameters of holiday zones for air are higher than their counterparts for the other modes. Income and number of children in households were found to be important explanatory factors in trip generation models. The derived VOT indicates that the VOT for long-distance international travel may differ significantly from the VOT for domestic long-distance travel. However, the models could not obtain a negative travel time parameter for air; therefore, the interpretation of VOT results, especially for air, should be taken with caution. Therefore, a sensitivity analysis was included, which tested the effects of a generic time parameter for all modes. The reason for the positive air-time parameter is presumably because many Swedes travel to tropical countries for their holidays, and thus, the chosen destination is far away. The holiday zone dummy captures this effect to some extent but not fully. In the revealed preference data, it is difficult to separate the effect of popular destination attraction from that of distance given that most popular destinations are far away from the Swedish context. This could also be the case in other Nordic countries.

Furthermore, the estimation results show that the VOT of cars is high for private 1-5 nights trips and private 6+ night trips. In fact, a considerable share, more than 30%) of observations that chose cars had a trip distance longer than 1500 km for private 6+ nights trips. These extremely long car trips are associated with a long travel time but not a proportionally high travel cost (compared to the travel cost by air at the same distance level). It is likely that car was chosen rather than other travel modes because of factors other than time and cost in private 6+ nights trips, which were not captured in the model. This could be related to the use of revealed preference data, where many factors related to mode choice are unknown and various factors are correlated and cannot be separated due to the geographical context, which is a limitation of this study.

The own and cross elasticities of train travel were also derived to provide the first impression of high-speed rail scenarios. The elasticities are generally comparable to the estimates of Rich and Mabit (2012). The most elastic attribute for private long-distance international trips is travel cost, while for business long-distance international trips, it is in-vehicle time. The induced demand, that is, those who previously did not conduct a long-distance international trip and now travel by train owing to the improved train service, is found to be negligible. This can be considered reasonable because holiday trips to popular international tourist resorts are more likely motivated by the need to visit the destination itself rather than improved accessibility.

The model estimation reflects several unique characteristics of Swedish international travellers due to Sweden's geographical location in the world. One such characteristic is that popular holiday destinations are often in warmer climates far away from Sweden and are only accessible by air. Another example is Swedes' travel by ferry to neighbouring Baltic Sea countries. Specific dummy variables are introduced to capture these characteristics, as they are otherwise captured in the level-of-service attributes and lead to wrong/non-intuitive parameter estimates. It also indicates that the model estimations from this study are not directly transferrable to other countries where the geographical location is different. Although the elasticities are generally comparable with the European-wide TRANSTOOLS model (Rich and Mabit, 2012), no further studies have found that produce elasticity/cross-elasticity estimates for international travel and can be used for comparison. Thus, there is a need for further global research on forecasting models for international travel.

Conflict of Interest

The authors declare that they have no affiliations with or involvement in any organization or entity with any financial interest in the subject matter or materials discussed in this manuscript.

Author Contributions

IK contributed to the work in the following manner described by CRediT contributor roles: Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Supervision, Validation, Writing – original draft. CL contributed to the work in the following manner:

Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft.

Acknowledgements

This work was funded by the Swedish Transport Administration under Grant TRV 2019/98241.

References

Andersson, M., Brundell-Freij, K., Eliasson, J., 2017. Validation of aggregate reference forecasts for passenger transport. Transportation Research Part A: Policy and Practice 96, 101–118.

Berglund, S., Kristoffersson, I., 2020. Anslutningsresor: en deskriptiv analys (Connection trips: a descriptive analysis) (No. 2020:3), Working Papers in TransportEconomics.

Beser, M., Algers, S., 2002. SAMPERS—The new Swedish national travel demand forecasting tool, in: National Transport Models. Springer, pp. 101–118.

Daly, A., 2010. Cost damping in travel demand models (RAND Technical Report No. TR-717-DFT).

Daly, A., 2007. National Traffic Forecasting Models in Europe and Elsewhere.

Daly, A., 1982. Estimating choice models containing attraction variables. Transportation Research Part B: Methodological 16, 5–15.

Dargay, J.M., Clark, S., 2012. The determinants of long distance travel in Great Britain. Transportation Research Part A: Policy and Practice 46, 576–587.

Divisekera, S., 2010. Economics of leisure and non-leisure tourist demand: A study of domestic demand for Australian tourism. Tourism Economics 16, 117–136.

Eliasson, J., 2022. Will we travel less after the pandemic? Transportation Research Interdisciplinary Perspectives 13. https://doi.org/10.1016/j.trip.2021.100509

European Commission, 2021. Long-distance cross-border passenger rail services (No. MOVE/2020/OP/0013).

European Environment Agency, 2021. Average carbon dioxide emissions from new passenger cars (based on Monitoring of CO2 emissions from passenger cars – Regulation (EU) 2019/631).

FHWA, 2018. National Long-Distance Passenger Model Documentation - Model Development (No. DTFH61-10- R- 00036).

Gelhausen, M.C., Berster, P., Wilken, D., 2018. A new direct demand model of long-term forecasting air passengers and air transport movements at German airports. Journal of Air Transport Management 71, 140–152.

Janzen, M., Vanhoof, M., Smoreda, Z., Axhausen, K.W., 2018. Closer to the total? Long-distance travel of French mobile phone users. Travel Behaviour and Society 11, 31–42.

Kim, S., Shin, D.H., 2016. Forecasting short-term air passenger demand using big data from search engine queries. Automation in Construction 70, 98–108.

Lannoo, S., Van Acker, V., Kessels, R., Cuervo, D.P., Witlox, F., 2018. Getting business people on the coach: A stated preference experiment for intercity long distance coach travel. Transportation research record 2672, 165–174.

Li, X., Pan, B., Law, R., Huang, X., 2017. Forecasting tourism demand with composite search index. Tourism management 59, 57–66.

Mabit, S.L., Rich, J., Burge, P., Potoglou, D., 2013. Valuation of travel time for international long-distance travel–results from the Fehmarn Belt stated choice experiment. Journal of Transport Geography 33, 153–161.

McFadden, D., 1974. Analysis of Qualitative Choice Behavior. Zarembka, P.(ed.): Frontiers in Econometrics. Academic Press. New York, NY.

Pieters, M., de Jong, G., van der Hoorn, T., 2012. Cross-border car traffic in Dutch mobility models. European Journal of Transport and Infrastructure Research 12. https://doi.org/10.18757/ejtir.2012.12.2.2960

Rich, J., Bröcker, J., Overgård, C.H., Korzenewych, A., Nielsen, O.A., Vuk, G., 2009. Report on scenario, traffic forecast and analysis of traffic on the ten-t, taking into consideration the external dimension of the union: Trans-tools version 2; model and data improvements.

Rich, J., Mabit, S.L., 2012. A long-distance travel demand model for Europe. European Journal of Transport and Infrastructure Research 12, 1–20.

Rohr, C., Fox, J., Daly, A., Patruni, B., Patil, S., Tsang, F., 2013. Modeling long-distance travel in great britain. Transportation Research Record 2344, 144–151.

Santana-Jiménez, Y., Hernández, J.M., 2011. Estimating the effect of overcrowding on tourist attraction: The case of Canary Islands. Tourism Management 32, 415–425.

Suh, D.Y., Ryerson, M.S., 2019. Forecast to grow: Aviation demand forecasting in an era of demand uncertainty and optimism bias. Transportation Research Part E: Logistics and Transportation Review 128, 400–416.

Swedish Ministry of the Environment, 2021. Sweden's climate policy framework.

Trafikanalys, 2017. RVU Sverige - den nationella resvaneundersökningen 2015–2016 The Swedish national travel survey 2015–2016 (No. 2017:13).

TRT Trasporti e Territorio, 2018. Description of the Trust model.

Van Acker, V., Kessels, R., Cuervo, D.P., Lannoo, S., Witlox, F., 2020. Preferences for long-distance coach transport: Evidence from a discrete choice experiment. Transportation Research Part A: Policy and Practice 132, 759–779.

Varela, J.M.L., Börjesson, M., Daly, A., 2018. Quantifying errors in travel time and cost by latent variables. Transportation Research Part B: Methodological 117, 520–541.

Wardman, M., Chintakayala, V.P.K., de Jong, G., 2016. Values of travel time in Europe: Review and meta-analysis. Transportation Research Part A: Policy and Practice 94, 93–111.

Witlox, F., Zwanikken, T., Jehee, L., Donners, B., Veeneman, W., 2022. Changing tracks: identifying and tackling bottlenecks in European rail passenger transport. European Transport Research Review 14, 1–12.

WSP Analysis & Strategy, 2011. Höghastighetståg – modellutveckling (High-speed train - model development).