Estimation of demand models for long-distance cross-border travel

Ida Kristoffersson*[1], Chengxi Liu[1]

[1] Dr., VTI Swedish National Road and Transport Research Institute, Box 55685,
10215 Stockholm, Sweden

SHORT SUMMARY

Although long-distance cross-border travel contributes significantly to global emissions from the transport sector, transport models for this type of travel are scarce. In this study, a disaggregated travel demand forecasting model is estimated using Swedish national travel survey data 2011-2016 along with detailed supply data from European road, train, and ferry networks and a World-wide air network, aiming at forecasting Swede's long-distance travel abroad. Mode choice, destination choice and trip generation are modelled by traditional Nested Logit models and Multinomial Logit models. Results show that values of time of long-distance cross-border travel derived from the model estimation are in general higher than values of time of long-distance domestic travel. Furthermore, elasticity 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 cross-border travel, especially for business trips.

Keywords: Discrete choice modelling; Transportation network modelling; Long-distance cross-border travel; Mode choice; Destination choice; Trip generation

1. INTRODUCTION

Long-distance cross-border travel differs from regional and national travel in many respects, such as what determines traveller trip generation, mode, and destination choice. Due to 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 cross-border trips is in general lower than the number of regional and national trips. Passenger-kilometres travelled by mode is important, especially since it is related to CO2 emissions from transport, for which ambitious reduction targets have been set both at the EU and national levels. Travel demand forecast models are an important part of large-scale modelling to provide accurate input to cost-benefit analyses 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.

One of the few existing demand models of long-distance cross-border travel is Trans-tools, which is a transport model for both passenger and freight transport in 42 European countries. The demand model for passenger transport is described in Rich and Mabit (2012). 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 fixed origin-destination (OD) matrices. Pieters et al. (2012) describe an effort to develop sub-models for border crossing traffic in the Dutch national model. Somewhat more common are 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. Due to the aggregate nature of these models, it is not possible to calculate e.g., cross-elasticities between modes. Examples of direct demand models include Divisekera (2010) for Australia, Santana-Jiménez and Hernández (2011) for Canary Islands and Li et al. (2017) for China. There are also some direct-demand models that focus on a certain mode, especially air travel, and predict e.g., number of air trips to certain airports (Gelhausen et al., 2018; Kim & Shin, 2016; Suh & Ryerson, 2019).

The lack of disaggregated travel demand models for long-distance cross-border travel can be a problem in practice when certain investments or policy measures might have a substantial impact on cross-border travel demand and cross-elasticities are of interest. One such example is high-speed train that connects large cities across countries. Witlox et al. (2022) determine a number of existing bottlenecks for European rail, such as the train travel time not being fast enough and too many interchanges. An analysis of the ability of policy measures and investments to remove these bottlenecks would benefit from travel demand models for long-distance cross-border travel.

2. METHODOLOGY

The travel demand models are formulated using classic discrete choice theory and logit formulations (McFadden, 1974). There are two sub-models per trip purpose (private/business): one nested logit model for mode and destination choice and one multinomial logit model for trip generation.

For the mode and destination choice model, the utility equation for an alternative (mode *i* and destination *j*) are formulated as:

$$U_{i,j} = ASC_i + \gamma_i I + \beta_i L_{i,j} + \delta_i D_j + \phi log(A_j) + \varphi_i + \varepsilon_{i,j} \qquad (1)$$

In Equation (1), ASC_i is the alternative specific constant for mode i. I is the vector of individual socio-economic 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, e.g., GDP and number of hotel beds per resident. A_j is a destination attraction variable (size variable) that represents the attractiveness in terms of size and quantity of each destination zone, for which a non-linear log formulation is used, see (Daly, 1982), and φ_i refers to the error term at the mode level. Thus, the alternatives with the same mode *i* will share the same error term φ_i and therefore those alternatives are not independent of each other. $\varepsilon_{i,j}$ refers to the error term that is unique and independent for each alternative. The mode and destination choice model with the utility function described in Equation (1) then is a nested Logit model where 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. The model structure that in estimation yields a logsum parameter which is within the range of 0 and 1 is the preferred structure.

For the trip generation model, the utility function for an alternative *k* is formulated as follows, where *k* belongs to {no long-distance cross-border trip; daytrip; 1-5 nights, and 6+ nights} for private travel, and {no trip and trip} for business travel.

$$U_k = ASC_k + \theta_k I + \mu_k T + \varphi_k logsum_{modeDestModel} + \theta_k \qquad (2)$$

In Equation (2), 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 then a Multinomial Logit model.

3. RESULTS AND DISCUSSION

The travel demand data consists observations of long-distance (one-way distance 100 km or longer) cross-border trips 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 have made during the last month and which trips longer than 300 km they have made during the last three months. There has been one more national travel survey conducted after this, in 2019, but in the 2019 survey only trips from the measurement day were asked for, which resulted in very few long-distance international trips. Therefore, the 2019 survey could not be used in this study. After the data cleaning process, 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 6 or more nights away. The modal shares for private trips differ a lot depending on number of nights away, which is a motivation for testing model segmentation across this variable. Car trips dominate for private daytrips, car and air trips are of about equal size for private trips 1-5 nights away, and air is the dominating mode for private trips 6+ nights away and for business trips.

One of the major tasks of this work was 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 is generated at 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, while NUTS zone system is used to represent Europe. Outside Europe, nations are represented as zones. In total, four networks are developed for car/bus, train, air, and ferry. Networks for car/bus, train, and ferry are European-wide while network for air is worldwide. As an example, the network for train is shown in Figure 1.

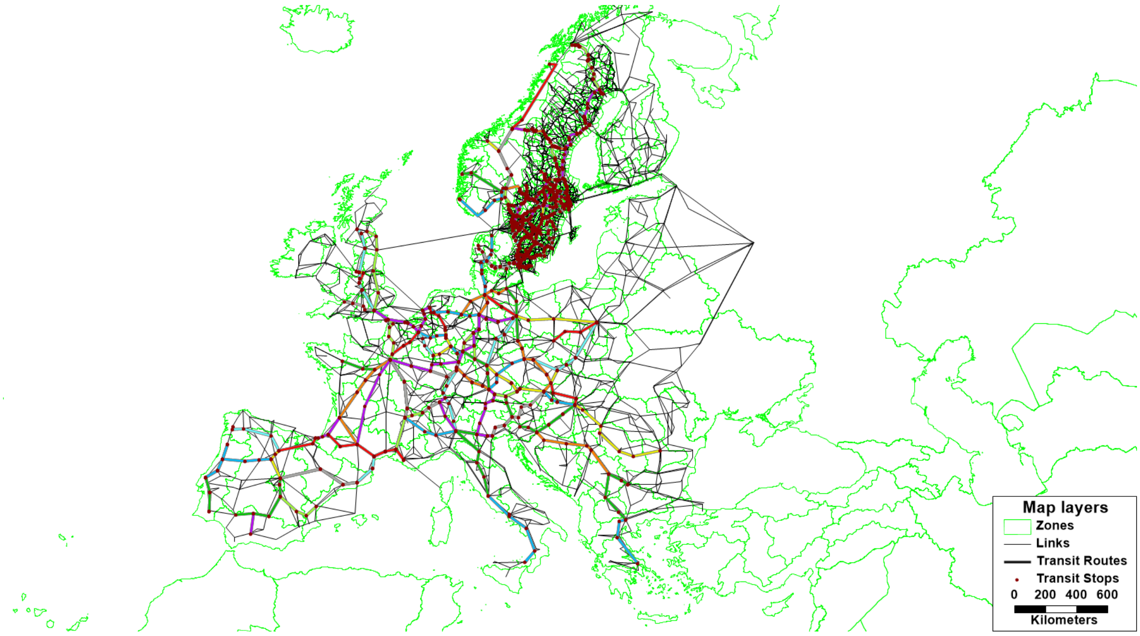

Figure 1: The European train network developed in TransCad.

Table 1 presents the estimation results for the mode and destination choice model for business trips. Note that there are four mode-destination choice models estimated: private daytrip, private 1-5 nights, private 6+ nights, and business, but there is only space to show results of one of these in this short paper. The table shows the final model specification for business trips. The initial model specifications were set to include all variables that are relevant and then insignificant variables have been removed gradually. A large number of model specifications were tested before selecting the final version. The t-values in the table show the statistical significance of the parameters in the model. A t-value (absolute value) larger than 1.96 means that there is a 95% probability that the parameter is different from zero, i.e., it has an effect in the model. A few parameters with a lower significance level are kept in the model (shown in red in Table 1). These are either alternative specific constants that would be used as calibration constant in an implementation of the model or important level of service variables.

Table 1: Estimated parameter values of the mode and destination choice model for long-distance cross-border business trips

Parameter name	Explanation	Mode	Parameter value	t-value
ϕ	Log-size	all	0.717	23.70
$\beta_{BedPerArea.noAir}$	Hotel beds per area	car, bus, train	0.059	3.65
$\beta_{GDP.noAir}$	GDP per capita	car, bus, train	3.810	8.59
$\beta_{GDP.air}$	GDP per capita	air	3.196	18.65
$\beta_{TT.car}$	Travel time	car	-0.0080	-11.97
$\beta_{TT.PT}$	In-vehicle time	bus, train, air	-0.0039	6.61
$\beta_{AC.train}$	Access/egress time	train	-0.0783	-3.76
$\beta_{AC.air}$	Access/egress time	air	-0.0093	-6.75
β_{Cost}	Travel cost	all	-0.0038	-4.67
$\beta_{LogCostLowMedInc}$	Log(Travel cost) for low/medium income segment	all	-0.6302	-1.70
$\beta_{CarHH.car}$	Availability of car in household	car	0.372	3.26
ASC_{bus}	Alternative specific constant	bus	-1.767	-5.74
ASC_{train}	Alternative specific constant	train	0.341	0.98
ASC_{air}	Alternative specific constant	air	-0.007	-0.02
$\text{Logsum}_{destination}$	Accessibility to destination	all	0.786	1.71
Number of observations			717	
Number of parameters			15	
Log-likelihood			-3454.2	
Log-likelihood all parameters=0			-4969.4	
McFadden rho			0.305	

Parameters of destination attraction variables are positive, showing that the quantity in terms of number of hotel beds/population/employment has a positive effect in attracting travellers to given destination zones. When it comes to level-of-service variables, all travel time and travel cost parameters are negative as expected. The disutility of travel time for car is in general higher than that for public transport which is expected since travel time on PT can be used for work activities. Furthermore, those with individual income lower than 30 TEUR have a higher cost sensitivity. Looking into the effects of socio-economic variables, number of cars in household is, as expected, a strong factor for choosing car. The logsum parameter is within the range of 0 and 1, indicating that the nested-Logit structure with mode at the upper level is valid. Value of time (VOT) estimates are derived from the estimated parameters of in-vehicle time and travel cost. Results are then compared to the VOT derived from the existing domestic long-distance model. In the business trip segment, VOT for car for long-distance cross-border trips is higher than VOT derived from the domestic long-distance trip model, while a reversed trend is found for public transport modes.

Table 2 shows the estimation results of the trip generation model for business trips. The available alternatives are taking no trip or conducting a business trip. Note that trip generation model estimation results for private trips (conducting no trip, private daytrip, private trip 1-5 nights, or private trip 6+ nights) exist but had to be left out due to space limitation.

Table 2: Estimated parameter values of the trip generation model for long-distance cross-border business trips

Parameter name	Explanation	Alternative	Parameter value	t-value
$\beta_{LowMedInc.0}$	Low/medium income segment	No trip	1.207	5.34
β_{CarHH}	Availability of car in household	Business trip	0.181	4.53
β_{Female}	Traveller is female	Business trip	-1.019	-11.21
β_{Age31_64}	Traveller age 31-64	Business trip	0.729	4.64
$\beta_{Age>64}$	Traveller age >64	Business trip	-1.127	-4.67
$\beta_{HighInc}$	High income segment	Business trip	1.156	8.01
β_{Summer}	Summer time	Business trip	-0.837	-5.68
$\beta_{Christmas}$	Christmas time	Business trip	-0.945	-4.01
ASC	Alternative specific constant	Business trip	-4.740	-23.81
Number of observations			39996	
Number of parameters			9	
Log-likelihood			-3066.4	
Log-likelihood all parameters=0			-27723.1	
McFadden rho			0.889	

It is found that low income is an important explanatory factor that contributes to not conducting any long-distance cross-border trips, which is expected. High income is a positive factor for conducting business trips. Pensioners (age >64) and female travellers are less likely to conduct business trips. Number of cars in the household is positively correlated with the likelihood of conducting business trips. It is as expected that there are fewer business trips in summer and Christmas.

Elasticities for the level-of-service attributes for train are derived to provide a first look into magnitudes of the impacts. Elasticities are calculated using sample enumeration. The elasticity shows the unit percentage change of the likelihood given a unit percentage change of a level-of-service attribute. The following scenarios are adopted for the elasticity calculations: 10% increase in travel cost by train and 10% decrease in train in-vehicle time. The results are presented in Table 3.

Table 3: Elasticity results of changes in 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

The elasticity of increased train travel cost is -0.475, which is similar to that of private trips. The business elasticity of decreased train in-vehicle time is much higher than that of private trips, 1.069, suggesting that business travellers are more inclined to take high-speed trains due to the travel time saving. Since the logsum variable is not significant and not included in the trip generation model for business trips, changes in level-of-service variables will not result in a change in the overall likelihood of business trip generation.

4. CONCLUSIONS

Long-distance international travel, although low in number of trips compared to regional travel, contributes significantly to total distance travelled and thus externalities from the transport sector. Despite the abundant literature on 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. 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 trip generation, mode and destination choice are modelled in Multinomial Logit models and Nested Logit models respectively. Swedish national travel survey data is used as observations of long-distance cross-border travel. European networks for road, train, and ferry and a world-wide network for air are developed at a reasonable level of detail. Models for private and business trips are developed where the ones for private trips are further segmented by number of nights away. The estimation results reveal the effects of individual socio-economic variables, level-of-service attributes, and destination variables. Income and access to car in household are found important explanatory factors in trip generation models for business trips. The derived VOT suggest that VOT for long-distance cross-border travel may differ significantly from VOT for domestic long-distance travel.

Elasticities of level-of-service attributes for trains are also derived to provide a first impression of high-speed rail scenarios. The most elastic attribute for private long-distance cross-border trips is travel cost, while for business long-distance cross-border trips it is in-vehicle time. The induced demand, i.e., those who previously did not conduct a long-distance cross-border trip and now travel by train due to the improved train service is however found to be neglectable.

ACKNOWLEDGEMENTS

The research has been conducted within the research project “Forecast models for international travel” funded by the Swedish Transport Administration under grant TRV 2019/98241.

REFERENCES

Divisekera, S. (2010). Economics of leisure and non-leisure tourist demand: A study of domestic demand for Australian tourism. *Tourism Economics*, *16*(1), 117–136.

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.

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.

Li, X., Pan, B., Law, R., & Huang, X. (2017). Forecasting tourism demand with composite search index. *Tourism Management, 59*, 57–66.

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*(2).

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*. https://www.semanticscholar.org/paper/Report-on-Scenario%2C-Traffic-Forecast-and-Analysis-Rich-Broecker/eb43addc54da5362319169437188db7c09260214

Rich, J., & Mabit, S. L. (2012). A long-distance travel demand model for Europe. *European Journal of Transport and Infrastructure Research*, *12*(1), 1–20. http://orbit.dtu.dk/fedora/objects/orbit:42232/datastreams/file_6461773/content

Santana-Jiménez, Y., & Hernández, J. M. (2011). Estimating the effect of overcrowding on tourist attraction: The case of Canary Islands. *Tourism Management*, *32*(2), 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.

TRT Trasporti e Territorio. (2018). *Description of the Trust model*. http://www.trt.it/wp/wp-content/uploads/2016/09/TRUST-model-detailed-description-1.pdf

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), 1–12.