Integrating train access and egress choices with mode and destination choice in the Dutch National Model System

Jasper Willigers, Significance
Frank Hofman, Rijkswaterstaat, Ministry of Infrastructure and Water Management

With large-scale strategic transport models there is a tendency to include an increasing number of aspects of travel behaviour into a single choice structure. This study for the Dutch National Model System has evaluated the possibilities to (further) integrate train access and egress choices with the model for mode, destination and time of day choice.

The Dutch National Model System, LMS, is a strategic model, used for long-term road and public transport travel forecasts. It consists of a growth model to produce origin-destination travel flows by mode and time-of-day, which iterates with an assignment model for road traffic assignment. The core of the growth model is a tour based choice structure that models tour frequency, mode choice, destination choice, time-of-day choice, train access and egress mode choice and train station choice. This is formulated as a nested logit model. The nesting structure varies across travel purposes, but consistently tour frequency is always at the "top" of the nesting structure and train access and egress choices are at the "bottom".

For the previous version of the LMS the growth model was sequentially estimated in three steps: tour frequency (with a stop-repeat structure), a combined model of mode, destination and time of day choice for car travellers and a model of train access and egress mode and train station choice. Within each step, full information maximum likelihood was used and different nesting options were possible per travel purpose.

Sequential estimation of a nested logit model does have a number of drawbacks (Daly, 1987). A resulting compromise was an inconsistency in the cost function of train travel compared to the other transport modes. A coefficient for a generalized cost function was estimated in the train access and egress mode and station choice model and only impacted on mode and destination choices via the train access/egress choice logsums. The train cost elasticities of the resulting model were criticized to be low compared to values found in literature.

The study for the new version of the LMS initially aimed at a simultaneous estimation of mode and destination choices with train access/egress and station choices via full information maximum likelihood. This estimation uses a combination of two survey data sets: three years of the MON Dutch national travel survey and the Dutch Railways customer survey (Klimaat V), the latter only for access/egress mode and station choices. During the project such an extensive model estimation (over 280,000 choice alternatives, 5 nesting levels) appeared to be possible, but impractical. Using the ALOGIT software with 43,000 observations such an estimation took three months to converge.

Instead of a completely simultaneous estimation a sequential estimation is still needed. To account for the drawbacks of the previous version of the LMS, some adaptations are made to the estimations. Train access mode and egress mode choices are estimated simultaneously with mode and destination choice, whereas station choice is estimated in a separate estimation. By assuming no cost sensitivity for station choice, travel cost is included at the access/egress mode choice level. This structure still allows for consistent estimation of the cost coefficient across all modes.

Station choice does depend on train travel time. A change compared to the previous model version is the replacement of average train travel time per station pair by a generalized journey time with rooftop calculation (e.g. Wheat and Wardman, 2017). This will lead to a more consistent reaction of the model outcomes with new train time table data. As a disadvantage it is not possible to integrally estimate the frequency and interchange penalties, therefore exogenous values for these penalties are used.

Examination of the estimation results reveals two problems with the estimated model. Firstly, the model does not well replicate the observed tour length distribution for the train. More specifically, very short train trips (under 10 km) are considerably overestimated and longer train trips are underestimated. A second issue is concerning the amount of travellers who travel via the largest stations, mostly the central stations of the larger cities in The Netherlands. For these stations the number of travellers tends to be underestimated. As the cause of these two mismatches has not been clarified, it is solved by adding constants for distance classes for the train destination choice alternatives and by adding constants for the largest stations in the station choice model.

The re-estimation of the model system has improved several other aspects, including more consistent cost sensitivities across modes. Regarding cost elasticity and value of time of train travel the model shows a substantial improvement compared to the previous model version.

References:

Daly, A., 1987. Estimating “tree” logit models. Transportation Research Part B: Methodological 21, 251–267.

Wheat, P. and Wardman, M., 2017. Effects of timetable related service quality on rail demand. Transportation Research Part A 95 (2017), 96-108.