STATIC VERSUS DYNAMIC TRAVEL TIME SKIMS FOR DISAGGREGATE MODE CHOICE MODEL ESTIMATION

Ramachandran Balakrishna, Srinivasan Sundaram and Howard Slavin
Caliper Corporation

Introduction

The traditional four-step planning process involves the application of static methods to peak and off-peak time periods, with a matrix of average travel time skims being calculated for each period. Each of these time periods can span several hours. Even when separate AM and PM planning models are developed, crucial within-period congestion dynamics may be ignored owing to the known limitations of the static modeling approach.

Mode choice models form a key step in the above four-step framework, splitting a single production-attraction (PA) or origin-destination (OD) matrix into a set of mode-specific demand matrices that can then be assigned onto the network links available to each mode. Mode splits are currently captured through discrete choice models popularly cast in the Logit and Nested Logit frameworks, and require estimation from disaggregate or survey data. These data almost always have to be augmented with the latest network skims such as travel times, in order to evaluate the (dis)utility of various modes. Static travel time skims are routinely utilized for this purpose, ignoring the impact of departure time on expected congestion and hence mode choice.

With the advent of time-dependent models based on Dynamic Traffic Assignment (DTA), it is now possible to estimate mode choice models using dynamic skims in place of the static skims generated by the planning process. Skims dependent on trip departure time are expected to represent a more realistic picture of trip-maker perceptions about expected network conditions. This paper explores the impact of such a substitution on the estimated parameters of a mode choice model based on real survey data.

Methodology

Without loss of generality, we focus our attention on the Nested Logit (NL) model, a framework that is rapidly gaining popularity in mode choice practice. The basic premise of NL is that an individual's choice is a result of a hierarchy of decisions, the levels being represented schematically as a tree:

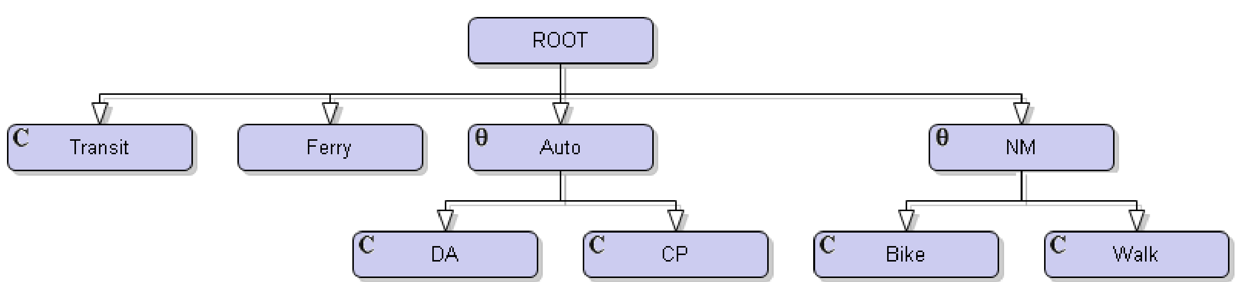

Figure 1. Nested Logit Mode Choice Model Tree

The example in Figure 1 indicates that trip makers consider a choice set of six modes: drive alone (DA), carpool (CP), bike, walk, transit and ferry. However, the actual choice is derived in two stages. In stage one, the trip maker decides between auto, non-motorized (NM), transit and ferry. If auto is preferred, then a second choice between DA and CP is executed. On the other hand, if NM is preferred, then a corresponding choice between bike and walk modes is executed. These modes are therefore selected conditional on a higher-level group or nest (auto or NM) being selected first.

The application of such a model requires the specification and estimation of utility equations for each of the modes, and the estimation of structural parameters (denoted as theta in the literature) for each of the nests (Auto and NM, in our example above). The theory behind NL model development is covered in depth in the literature (e.g. Ben-Akiva and Lerman (1985)).

The utility equations contain variables that are essential to explaining trip makers' mode choice decisions in the study region. In this context, network skims such as mode-specific travel times, transit wait times, fares and tolls are particularly relevant, and may be functions of the level of congestion. Further, these skims often do vary significantly over even small time intervals of a few minutes, an effect ignored by the static approach.

The methodology proposed for this study is as follows:

- Develop static travel time skims from the four-step planning model for the period of study.
- Estimate a NL mode choice model using these skims as inputs, together with a mode choice survey dataset.
- Perform a DTA on the same dataset.
- Calculate dynamic travel time skims for each combination of OD pair and departure time interval encountered in the survey. Tag these values into a new field in the survey dataset.
- Re-estimate the mode choice model after substituting the new dynamic skims in place of the static variable.
- Compare the estimation results from the two exercises.

It should be noted that NL model estimation is not a globally convex optimization problem, unlike multinomial logit (MNL) model estimation (Brownstone and Small, 1989; Daganzo and Kusnic, 1992; Koppelman and Bhat, 2006; Balakrishna and Sundaram, 2009; Balakrishna et al., 2009). Consequently, the starting values of the nest coefficients can potentially impact the final estimated coefficients and parameters. Care should therefore be taken to perform a systematic search on the space of the nesting coefficients to ensure that the best model(s) in each experiment are identified. The remainder of this paper focuses on a case study demonstrating the proposed methodology.

Case study: Victoria, British Columbia

A mode choice survey dataset from Victoria, British Columbia in Canada was employed for the numerical analysis. Home based work (HBW) trips during the AM peak (7:00-9:00) were chosen, and the model structure in Figure 1 was estimated. The utility equations were specified as follows (Table 1):

Coefficient	DA	CP	Bike	Walk	Transit	Ferry
ASC_DA	1					
ASC_CP		1				
ASC_Bike			1			
ASC_Walk				1		
ASC_Transit					1	
ASC_Ferry						Base
B_Distance			Dist	Dist		
B_Ped			Ped	Ped	Ped	
B_IVTT	Auto_IVTT	Auto_IVTT			Bus_IVTT	
B_Park	Park_Cost	Park_Cost				

Table 1. Utility Equations for Peak HBW Trips

Alternative Specific Constants (ASCs) were defined for all modes except the ferry, which served as the base for comparison. The walk distance (Dist) was specified for the non-motorized modes and Ped (a dummy variable set to 1 if the trip was within any of three dense urban districts) was used to capture a propensity for transit and non-motorized travel. Auto in-vehicle travel time (Auto_IVTT) and parking cost (Park_Cost) completed the utility equations for the two drive modes.

Static skims and other mode choice model inputs were available from the existing TransCAD planning model developed for the region. These data were used with TransCAD's built-in NL model estimation procedure to develop the HBW mode choice model for the AM peak period. The starting values of the nest coefficients for the Auto and NM nests were exhaustively varied between 0 and 1 in steps of 0.05 to ascertain a set of feasible models. The best model identified through this approach is summarized in Table 2 below.

Parameter	Estimate	t statistic
B_Ped	0.6391	4.54
B_IVTT	-0.0049	-0.43
B_Distance	-0.2053	-5.97
B_Park	-0.2376	-10.07
ASC (Transit)	4.8620	4.67
ASC(DA)	7.3248	7.17
ASC(CP)	5.8049	4.38
ASC(Bike)	5.9994	5.66
ASC(Walk)	5.8280	5.35
Theta(Auto)	0.5973 (starting = 0.6)	-3.13
Theta(NM)	0.3659 (starting = 0.2)	-8.87
Log-Likelihood at Zero		-2161.33
Log-Likelihood at End		-1140.67
Rho^2		0.4722
Adjusted Rho^2		0.4671

Table 2. Mode Choice Model Estimated with Static AM Peak Travel Time Skims

The planner's DTA in TransCAD version 5.0 was selected to test the hypothesis that dynamic skims can significantly alter the model parameter estimates. This DTA implementation is an adaptation of the work contained in Janson (1991) and Janson and Robles (1995). The DTA model is formulated as a constrained optimization problem whose solution closely satisfies a temporal extension of Wardrop's first principle, i.e. all used paths between a given origin-destination (OD) pair for the same departure time have the same and minimum experienced travel time. The solution algorithm contains two levels of iterative processes. An outer process solves for a consistent node-time-arrival matrix that governs the dynamic propagation of OD flows in the network (and can be roughly viewed as a temporal extension of the link-path incidence matrix in a static assignment problem); the inner process solves for a user equilibrium for a given node-time-arrival matrix. When both iterative processes converge, dynamic user equilibrium is reached for a node-time-arrival matrix consistent with actual link travel times.

The DTA procedure in TransCAD extends Janson's algorithm in several respects. One significant departure is a correction that ensures First-In-First-Out (FIFO) flows and more accurate estimation of travel times. There are also some changes in the algorithm and convergence checks that result in more consistent calculations. The TransCAD DTA introduces spillback as in Janson (2001), but uses a different set of adjustments. Lastly, an option is provided to use stochastic user equilibrium (SUE) instead of deterministic user equilibrium (UE) algorithm as the assignment method. For SUE, the Method of Successive Averages (MSA) algorithm is used to solve the upper subproblem.

The DTA in TransCAD can operate on existing planning models with minimal additional effort, thus providing a dynamic alternative to model very large regional networks. It is therefore a candidate for generating the dynamic skims for mode choice model estimation in the Victoria region. The static AM peak demand, specified as two hourly matrices, was profiled into 15-minute slices and the DTA procedure was run to convergence at a relative gap of 0.001.

Average OD travel times were calculated by weighting the dynamic skims for each departure time interval by the corresponding OD flows. The results revealed that the weighted dynamic skims averaged 32.47 minutes with a standard deviation of 29.31 minutes, the statistics being computed across all OD pairs. In contrast, the static skims averaged 30.93 minutes with a standard deviation of 27.84 minutes. The differences are better illustrated by fitting a straight line passing through the origin, which yielded an R^2 of just 0.27.

The dynamic skims were tagged to the survey data using the trip departure time field, and the HBW mode choice model for the AM peak was re-estimated. A grid search on the nest coefficients was again used. The results of the new estimation are summarized below in Table 3:

Parameter	Estimate	t statistic
B_Ped	0.6291	4.17
B_IVTT	-0.0008	-0.38
B_Distance	-0.2423	-12.37
B_Park	-0.2517	-8.82
ASC (Transit)	4.2160	5.22
ASC(DA)	6.7781	7.66
ASC(CP)	6.7597	7.64
ASC(Bike)	5.7553	6.26
ASC(Walk)	5.7511	6.26
Theta(Auto)	0.0156 (starting = 0.001)	-1349.07
Theta(NM)	0.0089 (starting = 0.001)	-1853.45
Log-Likelihood at Zero		-2161.33
Log-Likelihood at End		-1226.54
Rho^2		0.4325
Adjusted Rho^2		0.4274

Table 3. Mode Choice Model Estimated with Dynamic AM Peak Travel Time Skims

A review of the estimation results in Tables 2 and 3 reveals that many of the coefficients remain similar in the two cases. However, the IVTT coefficient with dynamic skims is almost one-sixth of the value estimated with static skims. Also, the two auto sub-modes are drawn much closer with dynamic skims, reflected in nearly identical ASCs.

The search on the theta domain was also much harder when dynamic skims were used, yielding only one solution for which all coefficients had the expected signs. In contrast, multiple feasible solutions were generated for the case with static skims. The estimated nest coefficients also yield interesting differences. Theta(Auto), for example, is significantly different than unity (Table 3) but is also significantly different than zero with a t statistic of 21.36. Similarly, Theta(NM) is significantly different than zero with a t statistic of 16.68. In the static case (Table 2), the nest coefficients are well away from both zero and unity. This leads to the conclusion that the correlations captured by NL are very different in the two cases, a finding that can have significant impacts in scenario analysis.

Conclusion

The preliminary results from this study indicate that the inclusion of within-period dynamics can impact the outcome of mode choice model estimation. The skims obtained from the static planning model may be inconsistent with actual congestion fluctuations within the study period, thereby introducing errors into the mode choice model estimation process. Further tests are being performed to ascertain the extent of the impacts across a range of time periods and trip purposes. Estimation differences when using static and dynamic skims can necessitate discussions of the impacts on the planning process, which typically does not generate dynamic network skims. Mode and departure time choice dimensions may also have to be combined to account for trip makers' true decision-making processes.

References

1. Balakrishna, R. and S. Sundaram (2009) "Critial Issues in Estimating and Applying Nested Logit Mode Choice Models." Presented at the 12th TRB National Transportation Planning Applications Conference, May 17-21, Houston.
2. Balakrishna, R., S. Sundaram and H. Slavin (2009) "An Empirical Study of Nested Logit Behavioral Model Estimation." Accepted for presentation at the 12th International Conference on Travel Behavior Research, December 13-18, Jaipur, India.
3. Ben-Akiva, M. and S. Lerman (1985) "Discrete Choice Analysis: Theory and Application to Travel Demand." The MIT Press, Cambridge, Massachusetts.
4. Brownstone, D. and K. Small (1989) "Efficient Estimation of Nested Logit Models." Journal of Business & Economic Statistics, Vol. 7, Issue 1, pp. 67-74.
5. Caliper (2009) "TransCAD Transportation Planning Software: Travel Demand Modeling with TransCAD – User's Guide."
6. Daganzo, C. F. and M. Kusnic (1992) "Another Look at the Nested Logit Model." Technical report UCB-ITS-RR-92-2, Institute of Transportation Studies, University of California at Berkeley.
7. Janson, B. N. (1991) "Dynamic Traffic Assignment for Urban Road Networks." *Transportation Research*, **25B**(2/3), 143-161.
8. Janson, B. N. and J. Robles (1995) "Quasi-Continuous Dynamic Traffic Assignment Model." *Transportation Research Record 1493*, 199-206.
9. Koppelman, F. S. and C. Bhat (2006) "A Self-Instructing Course in Mode Choice Modeling: Multinomial and Nested Logit Models." Prepared for U. S. Department of Transportation, Federal Transit Administration.