Analysis of the Temporal Transferability of Disaggregate Work Trip Mode Choice Models

DANIEL A. BADOE AND ERIC J. MILLER

An empirical study is presented of the long-range temporal transferability properties within a fixed geographic area of disaggregate logit models of work trip mode choice. The study area is the greater Toronto area, Ontario, Canada. The two temporal contexts are 1964 and 1986, with models estimated from 1964 data being used to predict 1986 travel choices. In addition to the very long transfer period (which does not appear to have been previously examined), a major feature of this study is that a wide variety of model specifications, ranging from the simplest possible market share model to a complex market segmentation model, are tested to investigate the relationship between model specification and transferability. Major findings of the study include (*a*) as in most transferability studies, model parameters are not temporally stable; (*b*) pragmatically the transferred models provide considerable useful information about application context travel behavior; (*c*) in general, improved model specification improves the extent of the model's transferability; (*d*) an important exception to Point c is the complex market segment model, which appears to be "overspecified" and, in the face of changing contextual factors during the 22-year period predicts 1986 conditions quite poorly; (*e*) Point c notwithstanding, simple level-of-service models perform very well in terms of their spatially aggregate predictions (which are often of primary practical importance to planners); (*f*) the models that best fit the estimation context (1964) data do not always transfer the best to 1986 conditions; and (*g*) "transfer scaling," in which modal utility constants and scales are updated, can significantly improve model transferability.

An important expected benefit from use of random utility models in transport modeling is transferability, that is, application of a model to a context different from which it was estimated. This expectation is based on the belief, first, that these models better represent the travel decision-making process and, second, that in the estimated model parameters the values associated with the different socioeconomic classes are built in. Hence, once a model is well specified to capture the decision process in one context, it should be applicable in other contexts so long as the basic nature of the decision-making process remains the same.

Consequently, transferability has been a subject of research interest for the following reasons: first, if it is feasible, the costs and time associated with transport decision-making, in a number of instances, can be reduced significantly; and second, it provides direct evidence of how well models that were estimated in one context perform in forecasting free of errors that would arise from having to forecast explanatory variables, thus making a statement about the range of validity of these models. Several empirical studies have been conducted to assess the effectiveness of model transfer from one context to another (*1–7*). Some of these studies have examined model transfer from one spatial context to another (*1,3,5,8,9*), whereas others have examined the temporal transfer of these models (*2,4,10–12*). The temporal dimension of transferability is the focus of this paper.

D. A. Badoe, Joint Program in Transportation, University of Toronto, 42 St. George Street, Toronto, Ontario, M5S 2E4 Canada. E. J. Miller, Department of Civil Engineering, University of Toronto, 35 St. George Street, Toronto, Ontario, M5S 1A4 Canada.

The assessment of transferability in the temporal domain has been mixed. The studies reported elsewhere (*2,4,10,12*), even though in some cases they involved simple specifications [e.g., Hensher and Johnson (*2*) used only level-of-service variables in their models], found disaggregate demand model explanatory variable coefficient estimates to show stability and provide a great degree of useful information in the transfer context and concluded that the developed models were temporally transferable. On the other hand, the transferability studies of Talvitie and Kirshner (*3*) and Train (*11*) reject temporal transferability. Train found the forecast errors from transferring estimated models on a pre-BART (Bay Area Rapid Transit) context to a post-BART context to be large and therefore rejected temporal transferability. Train's study, however, was clouded by problems of introduced new modes; therefore his findings are not entirely surprising. Talvitie and Kirshner assessed transferability on the basis of a statistical test of the set of model parameters from the pre-BART context being equal to the set of post-BART model parameters (this included the modal constant terms, which are context specific). As argued by Ben-Akiva (*13*) and Koppelman and Wilmot (*5*), assessing model transferability only on the basis of the set of model parameters being equal in the two contexts is stringent and unlikely to be met because no model is perfectly specified; as a result, all models are in principle context dependent (*6*). A more pragmatic evaluation of transferability is achieved by assessing the extent of useful information provided in an application context by transferred models (*6,13*). This viewpoint for assessing model transferability is also adopted in this paper.

Basically all the temporal transferability studies to date have been limited to short intervening times between estimation and application contexts, where differences in urban conditions between the two contexts are unlikely to be large. Nevertheless, these models are also applied in long-range forecasting in which significant changes in urban conditions occur. Horowitz (*14*, p.145) writes:

> An issue in temporal transferability that has arisen relatively recently concerns whether random utility travel demand models are likely to be transferable over time in periods of significant macroeconomic structural change, such as appear to be occurring now in some Western countries. There is, at present, no empirical evidence on this issue.

This paper examines the temporal transferability of morning peak period work-trip disaggregate multinomial logit mode choice models in the greater Toronto area (GTA), Ontario, Canada. Three travel modes are considered: automobile drive, public transit, and walk modes. The two urban contexts used in the study have an intervening period of 22 years between them, during which significant urban changes occurred. The validity of the assumptions inherent in the use of cross-sectional random utility models in such long-term temporal transfer has not yet been rigorously tested. From a theoretical perspective, Horowitz (*14*) points out that the ability of random utility models to transfer successfully over time during periods of structural change depends on whether such change entails substantial alterations of people's tastes or whether it consists mainly of changes in the attributes of the alternatives people face. In the former case, Horowitz states that it is unlikely that models can be transferred, whereas there is reason for cautious optimism in the latter case. He goes on to state (*14*, p.145):

> . . . if structural change influences mainly the attributes of available travel alternatives, then disaggregate random utility models can be expected to be transferable if they are free of serious specification error and if their explanatory variables encompass all attributes relevant to the choices of interest whose levels change significantly.

The other area in which this paper differs from existing empirical temporal transferability studies is in model specification; a single areawide model specification is not assumed, a priori, to be the most appropriate to capture traveler mode choice behavior. Instead, alternate specifications are explored. [It is noted, for example, that Train (*11*) and Koppelman and Wilmot (*9*) tested alternate specifications; however, the tested specifications had the underlying assumption that all travelers placed the same weight on transport system attributes.] Some of these specifications allow for taste differences among defined subgroups in the travel market. This testing of alternate specifications permits an assessment of the relationship between long-term transfer effectiveness and model specification.

The impact reestimation of modal constants and utility scale parameter has on transfer effectiveness is also discussed, thus allowing comments on whether tastes changed over time in response to significant macroeconomic changes.

The next section of this paper describes the two data sets used for the analysis. The section on comparison of urban structure attributes discusses briefly the differences in urban conditions between 1964 and 1986. The section on model specification presents the alternate model specifications investigated. The section on model estimation results presents the statistical estimation results of the estimation context models. The section on evaluation of transferability presents the results and discussion of the various transferability tests conducted. The impact reestimating modal constants or utility scale parameters, or both, has on transfer effectiveness is discussed in the section on updating constants or utility scale parameter. Finally, the conclusions and findings drawn from this study are outlined.

DATA

The two sources of data for this study are the 1964 Metropolitan Toronto and Regions Transport Study (MTARTS) data base and 1986 Transportation Tomorrow Survey (TTS) data base. The 1964 data were collected in a home interview survey conducted in metropolitan Toronto and its neighboring regions. The total usable questionnaires from this survey totalled 24,000, representing 3.3 percent of all households in the survey area. It provides detailed information on trips and personal characteristics of all household members in the sample. The 1986 TTS data, collected in a telephone interview survey conducted throughout the entire GTA, also provide detailed information on trips and personal characteristics of each household member in the sample. The number of usable household questionnaires totalled 67,000, representing 4 percent of all households in the sampling frame. The data sets do not contain identical information. For example, the 1964 survey collected information on occupation of household members and household income, whereas the 1986 survey did not. These data inconsistencies are considered in model specification. Census data obtained from Statistics Canada are used to augment the travel survey data in the brief descriptive comparison of urban conditions.

All level-of-service data required for model development, with the exception of parking costs and transit fares, were generated using computerized representations of the GTA automobile and transit networks maintained within the EMME/2 modeling system.

The 1964 travel data base is used for estimation of models that are to be transferred. The 1986 data base represents the travel context to which the estimated 1964 models are transferred for evaluation of transferability.

Although automobile passenger, automobile access to transit (park and ride or kiss and ride), and (in 1986) commuter rail modes were also observed to be used by workers in the data bases, these modes were excluded from this analysis to reduce modeling complexity with respect to specification, decision structure (e.g., avoidance of nested decision structures associated with access mode choice), and introduction of new modes (the commuter rail service did not exist in 1964).

COMPARISON OF 1964 AND 1986 URBAN CONDITIONS

Table 1 presents figures on the various characteristics of urban structure in the GTA for 1964 and 1986. The population of the GTA grew from about 2.7 million in 1964 to 4.1 million in 1986, representing a 53 percent increase in the 22-year period, whereas the number of households grew from 0.71 million to 1.47 million, a 106 percent increase. Average household size thus declined from 3.7 persons per household to 2.8 persons per household. A predictable outcome was the increase in percentage of single- and two-person households.

The labor force participation rate for females rose from 45 percent in 1971 to 66 percent in 1986, with the corresponding figures for males being 78 and 81 percent, respectively. This contributed to an increase in the proportion of multiple-worker households. The rate of driver license ownership among female workers also rose from 43.4 percent in 1964 to 77.8 percent in 1986. The corresponding figures for males were 88.6 percent and 93.8 percent, respectively. Private car registration in the GTA rose dramatically from 0.54 million in 1964 to 2 million in 1986, representing close to a 300 percent increase. Household car ownership consequently rose from 0.80 cars per household to 1.4 cars per household in the respective years.

The economic base of the GTA also changed, with the service industry superseding the manufacturing sector as the major employment source for GTA residents. Location patterns for these two industry types are different. The service industry is oriented more to the central business district (CBD), whereas the manufacturing industry, which in 1964 was largely located within the bounds of metropolitan Toronto, is primarily located in the subur-

TABLE 1 Comparison of Urban Attributes of GTA in 1964 and 1986

Attribute	Year	
	1964	**1986**
Population (thousands)	2,657	4,063
Average Weekday Travel (thousands)	3,800	8,800
Average Household Size	3.7	2.8
Private Auto Registration (thousands)	542	1,996
Transit Route Kilometres	953	1345
Transit Vehicle Kilometres (millions)	88	189

Source: TTC Annual Report, 1964 and 1986
Canada Statistics, Road Motor Vehicle Registration
1964 MTARTS and 1986 TTS Travel Survey Data

ban areas of neighboring regions to metropolitan Toronto, where space is available and cheap. Decentralization resulted in the percentage of total population residents in the suburban regional municipalities rising from 34 to 47 percent. These spatial trends in employment and residential locations in turn altered trip distribution patterns within the GTA.

The transport system also experienced expansion. However, the balance of investment was in favor of public transport, which increased its output, measured in transit vehicle kilometers, from 88 million in 1964 to 189 million in 1986.

Average weekday travel in 1964 was about 3.8 million trips, whereas in 1986 this was about 8.8 million trips, representing a 158 percent growth in travel. Notwithstanding the decline in average household size, the number of trips made per household rose from 5.50 to 5.85, and the number of trips per person rose from 1.4 to 2.1. Car use increased by 120 percent from 2.2 million trips in 1964 to 4.8 million trips in 1986. However, the passengers carried in these cars increased less than 60 percent from 0.8 million trips to 1.3 million trips in the respective years, resulting in a decline in the car occupancy rate. Use of public transport increased over 90 percent, from 0.7 million daily trips to 1.35 million daily trips. The average work trip length (Euclidean distance) increased from 7.9 km in 1964 to 11.5 km in 1986, the increase being particularly pronounced for trips by car and transit.

MODEL SPECIFICATIONS

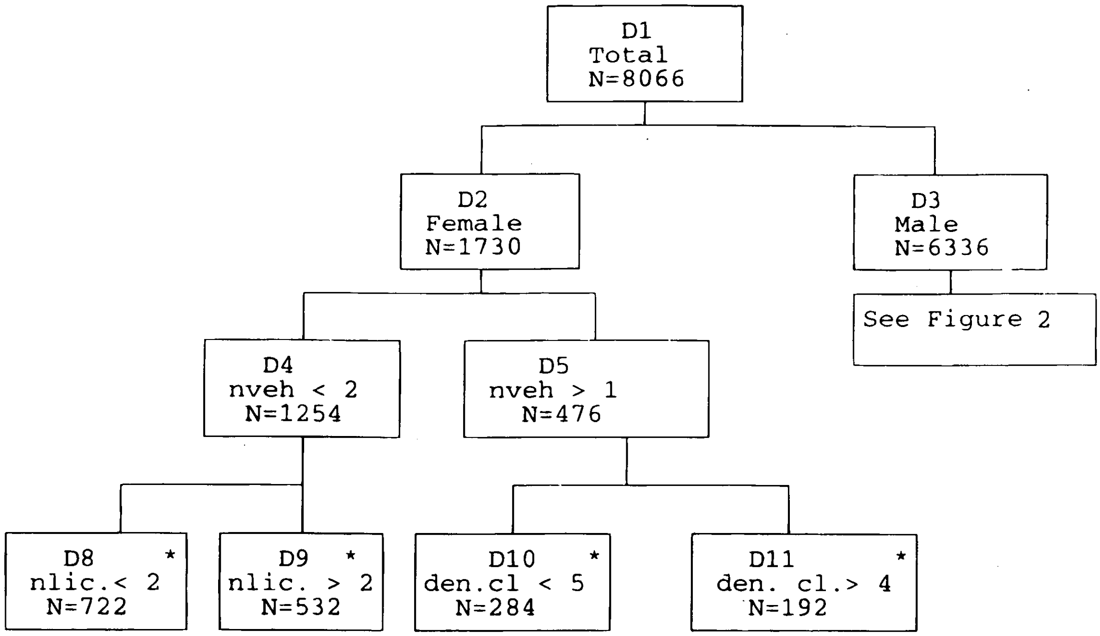
Seven different model specifications are explored. The first is a simple market share model, with the interpretation that each of the considered modes retains its relative share in the forecast context or that no explanatory variables are necessary to explain choice variations in the forecast context. It gives a lower bound on model transfer performance. The second and third models are simple level-of-service models. The first of them treats all the variables, with the exception of in-vehicle costs, as mode specific. The second level-of-service model treats the in-vehicle cost and in-vehicle time of the variables as generic attributes in the automobile drive and transit utilities. Further, it assigns the same importance weight to transit wait time and transit access and egress times. The fourth is termed a fully specified model. In addition to level-of-service attributes, it includes spatial, personal, and household characteristics of the tripmaker.

These four models assume the same coefficient estimates for all travelers in the GTA. The next three model specifications are defined for subgroups of workers that are determined to be relatively internally taste homogeneous. In line with this, the fifth model uses a heuristic segmentation procedure, which essentially consists of applying the automatic interaction detector (*15*) with multinomial logit models to identify 10 multivariately defined market segments with relatively homogeneous tastes. These mutually exclusive segments, which are defined by socioeconomic and spatial variables, are shown in Figures 1 and 2. Simple level-of-service models, similar in specification to the first level-of-service model mentioned, are estimated using data from each subgroup. For a complete description of the segmentation procedure used, see work by Badoe (*16*). Although such an extensive segmentation scheme would not generally be practical in most forecasting applications, it was supportable in this study given the large data sets available. Given this, it was felt that as a research exercise it was worthwhile to explore the impact that multivariate segmentation would have on model performance relative to more conventional nonsegmented or univariate segmentation schemes.

The sixth model takes the first pair of subgroups to emerge from application of the segmentation procedure mentioned earlier to the 1964 data (here the entire sample is stratified by gender to yield subgroups of male and female workers) and estimates models similar in specification to Model 4 mentioned earlier but excluding the gender alternative–specific socioeconomic variables on the two gender worker groups. The seventh model takes the first set of homogeneous subgroups to emerge from application of the segmentation procedure to the 1986 data set. In this case, the worker subgroups were defined according to household automobile ownership level. Models similar in specifications to Model 4 were estimated on the obtained subgroups. The structure of all the models mentioned earlier is the multinomial logit model with three modes: automobile drive, transit, and walk.

MODEL ESTIMATION RESULTS

Table 2 defines the variables included in the models considered. Estimation results for these models are presented in Tables 3 and 4. Most of the models' estimated parameters are statistically well determined and have signs consistent with a priori expectations.



Abbreviations

nveh. number of vehicles available to household.
den.cl. trip-end density class (ranges from 1 to 6, higher class numbers mean greater trip-end density).
nlic. number of persons in household with driver's licence.
orig. work-trip origin.
dest. work-trip destination.

FIGURE 1 Multivariately defined market segments, Part 1; selected samples indicated by asterisk.

However, the multivariately defined market segment models have some parameters of counterintuitive sign. For forecasting purposes, the affected models are reestimated constraining parameters of counterintuitive sign to 0 values.

Log-likelihood values for these models indicate the multivariately defined segment models to give the best fit to the 1964 data. This is followed by the gender-based models and then the conditional household automobile ownership models, and so forth, ranked according to log-likelihood value. However, when penalty is applied for the number of estimated parameters, as given by the adjusted likelihood ratio index, the gender-based models and the multivariately defined segment models have similar goodness-of-fit.

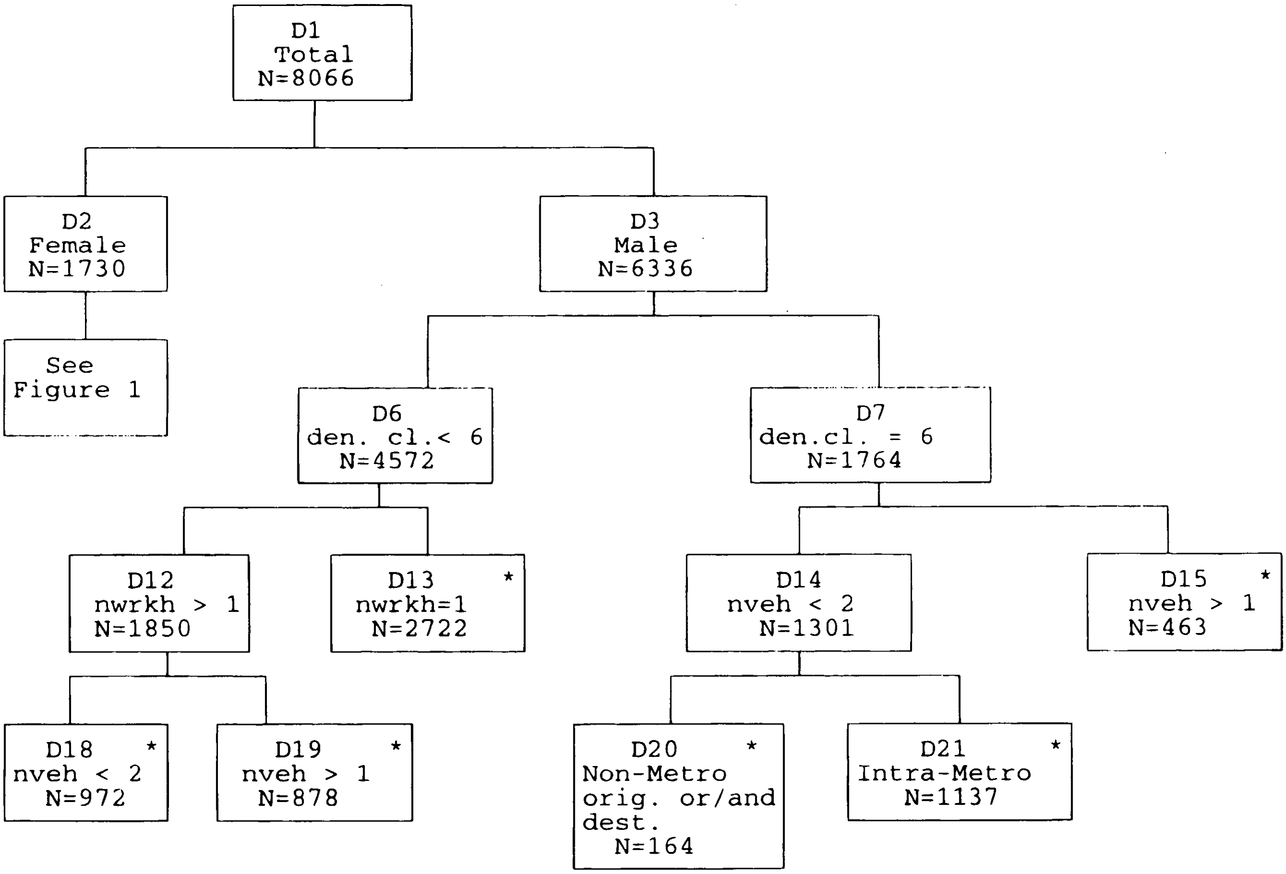
MEASURES OF TRANSFERABILITY

The following are criteria for judging model transferability:

1. Statistical similarity of estimation and application model coefficients: A nested likelihood ratio test is conducted for this purpose.

2. Ability of the transferred model to replicate individual choice in the application context: In absolute terms, performance here is assessed by the transfer log-likelihood value, which indicates relative disaggregate prediction performance of alternate model specifications. Relative performance measures that indicate how well the 1964 models perform in disaggregate prediction relative to locally estimated similarly specified models on the 1986 data set are provided by the transfer index (TI) and transfer goodness-of-fit measures (*5*). The transfer index has a maximum value of 1.0.

3. Ability of the transferred model to replicate observed aggregate shares: Aggregate predictions of mode use are obtained for seven destination regions of the GTA, which comprise the six constituent municipalities of the GTA, with metropolitan Toronto (by far the largest of the six regions) being split in two, for example, the CBD (Planning District 1) and the remaining districts of metropolitan Toronto. Root-mean-square error (RMSE) and mean absolute error (MAE) values, which are absolute measures of predictive accuracy, are computed using the aggregate predictions to assess forecast accuracy. The relative aggregate transfer error (RATE), which is the ratio of the RMSE value from application of the transferred model in the application context to the RMSE value from a locally estimated model on the application context is also computed. Expressions for these error measures can be found elsewhere (*5*). In addition, 95 percent prediction intervals are constructed (*17*) to determine whether the intervals given by each of these models include the observed mode use for each destination region. The rationale for obtaining confidence intervals is that the 1986 forecasts are based on relationships between dependent and independent variables that are not precise but subject to random errors. A point estimate alone would therefore be suggestive of a precise relationship not subject to random errors. Thus, the observed mode use



Abbreviations

nveh.	number of vehicles available to household.
den.cl.	trip-end density class (ranges from 1 to 6, higher class numbers mean greater trip-end density).
nwrkh.	number of workers in household.
orig.	work-trip origin.
dest.	work-trip destination.

FIGURE 2 Multivariately defined market segments, Part 2; selected samples indicated by asterisk.

value, if the confidence interval bounds it, confirms the appropriateness of the model specification.

RESULTS

Test of Parameter Equality

As in other transferability studies, results of the nested likelihood ratio test of 1964 and 1986 model parameters being statistically equal (Table 5, Column 2) reject the null hypothesis for all model specifications. As discussed earlier, this is not surprising given the errors in the modeling procedure. The emphasis is thus on the more pragmatic measures of transferability assessment reported below.

Disaggregate Measures of Transferability

The transfer log-likelihood values for each model are indicated in the Column 3 of Table 5. The worker mode choice model conditional on household automobile ownership level, with a log-likelihood value of −10,304, is found to give the best disaggregate predictions on the observed 1986 data. This is followed by the gender segment models and then the single areawide fully specified model. The multivariately defined segment models, which gave the best data fit in the estimation context, performed quite poorly yielding a log-likelihood value of −11,366. With this exception, however, improved model specification in general translates into improved disaggregate predictive performance in the application context.

TI values range from 0.132 for the multivariately defined segment models to 0.894 for Level-of-Service Model I, indicating that some of the 1964 models provide a significant component of information obtained from local 1986 models. The less well specified models have higher TI values than the better-specified models. Computed transfer goodness-of-fit measures (Table 5, Column 5), in general, compare favorably with the goodness-of-fit index values given by the locally estimated models on the 1986 data set (Column 6). The negative goodness-of-fit value for the transferred 1964 market share model means its log-likelihood is lower than the log-likelihood given by the 1986 market share model.

TABLE 2 Definition of Variables Specified in Mode Choice Models in Table 3

dauto	= 1 in auto-drive mode utility (modal constant); = 0 otherwise
dwalk	= 1 in walk mode utility (modal constant); = 0 otherwise
aivtt	= auto in-vehicle travel time (min.), enters into auto-drive mode utility; = 0 otherwise
ivtt	= aivtt (auto in-vehicle travel time) in auto-drive mode utility; = tivtt (transit in-vehicle travel time in transit mode utility; = 0 in walk mode utility.
tivtt	= transit in-vehicle travel time (min.) in transit mode utility; = 0 otherwise
twait	= transit wait time (min.) in transit mode utility; = 0 otherwise
twalk	= transit access + egress + transfer time (min.) in transit mode utility; = 0 otherwise
tovtt	= transit out-of-vehicle time (min.) (access+egress+transfer+wait times) in transit mode utility; = 0 otherwise
ivtc	= aivtc (the auto in-vehicle travel costs ($) for auto-drive mode); = 0 in walk mode utility; = tfare (the transit fare) in transit mode utility.
apkcst	= auto daily parking cost ($) in auto-drive mode utility; = 0 otherwise
wdist	= walk distance (km.) in walk mode utility; = 0 otherwise
avplic	= number of vehicles per licensed person in household. Enters into auto-drive mode utility; = 0 otherwise
wcbd	= 1 in walk utility if worker's employment location is in Central Business District; = 0 otherwise
amal	= 1 in auto-drive utility, if worker is male; = 0 otherwise
tcbd	= 1 in transit utility if worker's employment location is in Central Business District; = 0 otherwise
tgend	= 1 in transit mode utility, if worker is female; = 0 otherwise

Aggregate Measures of Transferability

Aggregate transfer measure values based on aggregate predictions given by naively transferred models from the 1964 context are presented in Table 5. Even though RMSE and MAE penalize the prediction error differently, they both indicate that the Level-of-Service Model I, which treats nearly all the system attributes as alternative specific, notwithstanding its simplicity in specification, to yield the best spatial predictions of mode use. Level-of-Service Model II and the worker choice model conditional on automobile ownership level have comparable aggregate forecast performance. The forecast performance of the single areawide fully specified model and the multivariately defined segment models are disappointing given their superior specification to the simple level-of-service models.

RATE values range from 1.0 for the market share model to 5.6 for the market segment models. In general, the better the model specification the higher the RATE value. This is understandable, given the fact that a well-specified model estimated on the 1986 context yields far superior aggregate predictions compared with a poorly specified local 1986 model.

Table 6 is a summary table that shows whether the confidence intervals for predicted mode use by destination region, given by each of the 1964 models, does include the observed mode use. Where the confidence interval given by a particular model type bounds the observed mode use for a destination region, the abbreviation for that model is recorded in the cell described by that mode and destination region. The results indicate that with the exception of the market share model, most models do not yield confidence intervals that include the observed walk mode use for most of the destination regions. This, however, improves for the transit and automobile drive modes. Hamilton is the only region for which the confidence interval given by most of the models bounds the observed mode use for all the three modes. Interestingly this destination region, which is self-contained in terms of trip distribution, underwent comparatively minor change in urban conditions in the 22-year period (*16*). None of the models yields modal confidence intervals that bound the observed for work trips destined to Planning District 1, the downtown district of Toronto. This is disappointing because in any long-range planning mode split of trips to this district would be of considerable interest. The three models here, which are of superior performance compared with the others, are the Level-of-Service Model I, the gender segment models, and the automobile ownership models.

Reestimation of Modal Constants and Utility Scale

Modal constants are in principle context specific because they capture those aspects of the choice process for which the included model explanatory attributes do not account. Hence, their transferability from one context to another is expected to be weak. Thus, these con-

TABLE 3 1964 Model Parameter Estimates

Variables	Market Share	Level of Service (I)	Level of Service (II)	Fully Specified	Male Model	Female Model	0 Veh. Model	1 Veh. Model	2+ Veh. Model
dauto	1.348	0.090 *	-0.133 *	-1.266	-0.583	-2.298		-0.978	-1.636
dwalk	-1.139	0.924	0.626	1.592	1.050	1.927	0.661 *	1.423	1.597
aivtt		-0.031		-0.009	-0.009	-0.014		-0.015	
ivtt			-0.037						
tivtt		-0.043		-0.029	-0.036	-0.011 *		-0.033	-0.030
twait		-0.205		-0.202	-0.209	-0.202	-0.284	-0.182	-0.216
twalk		-0.046		-0.026	-0.037		-0.051 *	-0.032	-0.003 *
tovtt			-0.123						
wdist		-1.961	-1.918	-1.884	-1.675	-2.347	-2.253	-1.758	-1.810
ivtc		-0.389	-0.040 *	-0.388	-0.386	-0.468	-2.190	-0.278	-0.695
pkcst		-0.333	-0.314	-0.282	-0.273	-0.332		-0.317	-0.199
avplic					1.540	3.487		1.523	2.739
amal								0.672	0.649
tcbd					1.252	1.014	0.293 *	1.061	1.559
tgend							0.546	0.973	0.734
wcbd					0.773	0.984		0.773	1.304
No. of Obs.	8066	8066	8066	8066	6336	1730	640	5150	2276
Log-Likelihood at Zero	-5929.6	-5929.6	-5929.6	-5929.6	-4677.6	-1251.9	-434.7	-3836.7	-1658.2
Log-Likelihood at Conv.	-3847.3	-2839.4	-2883.4	-2590.5	-1990.7	-566.1	-165.1	-1785.3	-614.8
Adjusted Likelihood Ratio Index	0.3511	0.5204	0.5134	0.5625	0.5737	0.5452	0.6159	0.5336	0.6274

veh. - number of vehicles available to household.

Note: Parameter estimates with asterisk (*) sign are insignificant at the 5% level.

TABLE 4 Parameter Estimates of Multivariately Defined Market Segment Models

Sample	Parameters								
	dauto	dwalk	aivtt	tivtt	twait	twalk	wdist	ivtc	apkcst
D8	1.280	2.508	-0.051	-0.017*	-0.244	0.012*	-2.666	-0.109*	-0.308
D9	-0.019*	2.563	-0.012*	0.003*	-0.199	0.036*	-2.297	-0.159*	-0.951
D10	-0.951*	0.201*	-0.108	-0.098	-0.175	-0.115*	-2.675	-1.387	-0.282*
D11	-0.233*	0.415*	0.002*	-0.030*	-0.189	0.082*	-1.527	-1.878	-0.281
D13	1.208	1.254	-0.040	-0.038	-0.168	-0.066*	-1.518	0.039*	-0.073*
D15	-1.978	-1.881*	0.006*	-0.039	-0.498	-0.065*	-1.587	-0.916	-0.326
D18	-1.092	-0.463*	-0.023	-0.039	-0.235	-0.121	-1.739	-0.926	-0.089*
D19	2.884	3.862	-0.048	-0.025*	-0.163	0.049*	-2.139	-0.363*	-0.172*
D20	6.333*	5.952*	-0.097*	-0.000*	0.051*	0.038*	-1.171	1.804*	-0.216*
D21	-1.069	1.740	-0.005*	-0.057	-0.093	-0.017*	-2.376	-0.429	-0.337

Note: Parameter estimates with asterisk sign are insignificant at the 5% level.

stant terms are reestimated while the remaining utility function parameters are transferred to explore how well these models would have performed free of these purely contextual parameters. In another scenario, the remaining utility function parameters are rescaled. Rescaling is equivalent to reestimating the variances of the distributions of the random utility components. The necessary mathematics for this can be found elsewhere (*7*). The intent of the analysis here is to investigate whether shifts in constants or scale, or both, are responsible for the models not yielding better transfer performance.

Evaluating transfer assessment measures reported in Table 7 after reestimating the modal constants using information from the application context indicates that this results in considerable improvement in model predictive performance. Across models, the log-likelihood values increase significantly compared with their naive

TABLE 5 Results from Transferring 1964 Models to 1986 Application Context

1964 Model Type	Nested Likelihood Ratio	Absolute and Relative Disaggregate Transfer Measures				Aggregate Transfer Measures		
		$(LL_{86}(\theta_{64})$	TI	$Rho(\theta_{64})$	$Rho(\theta_{86})$	$RMSE_{64}$	MAE	RATE
Market Share	45	-15115		-0.008	0.000	0.35	0.28	1.01
Level of Service (I)	158	-11352	0.894	0.243	0.272	0.09	0.06	1.82
Level of Service (II)	107	-11378	0.891	0.241	0.271	0.12	0.08	2.21
Fully Specified	371	-10787	0.759	0.281	0.370	0.18	0.10	5.21
Auto Ownership	347	-10304	0.787	0.259	0.329	0.13	0.07	4.09
Gender Segments	306	-10494	0.789	0.280	0.355	0.17	0.09	4.92
Market Segments	367	-11366	0.132	0.024	0.179	0.16	0.10	5.57

$$TI = \frac{LL_{86}(\theta_{64}) - LL_{86}(c_{86})}{LL_{86}(\theta_{86}) - LL_{86}(c_{86})}$$

where TI is the Transfer Index Computed with 1986 context Log-Likelihood given by 1986 Market Share model (c_{86}) as Base; and $LL_{86}(\theta_i)$ is the 1986 context log-likelihood using parameters θ_i from year i ($i \in \{64,86\}$)

$Rho(\theta_i)$ is the Likelihood Ratio Index computed with $LL(c_{86})$ using model parameters θ_i from Year *i* ($i \in \{64,86\}$).

$$RMSE = \left[\frac{\sum_{mg} \frac{(\hat{N}_{mg} - N_{mg})^2}{\hat{N}_{mg}}}{\sum_{mg} \hat{N}_{mg}} \right]^{\frac{1}{2}}$$

where N_{mg} and N_{mg} are the number of persons observed and predicted to choose alternative m from group g respectively.

$$RATE = \frac{RMSE_{64}}{RMSE_{86}}$$

where $RMSE_i$ represents the 1986 context root-mean-square error computed with estimated model parameters from year i (i could be 64 or 86).

TABLE 6 Models Yielding Confidence Intervals That Include Observed Mode Use

Destination Region	Mode		
	Auto Drive	Transit	Walk
PD1			MC
R.O.M	AO, GEN	LOS I,FS, AO, GEN	
Durham		GEN, MS	MC
York	LOS I, LOS II, AO, GEN, MS	LOS I, MS, LOS II, FS, AO, GEN	
Peel		LOS I, LOS II, AO, GEN, MS	
Halton			MC
Hamilton	LOS I, FS, AO, LOS II, GEN, MS	LOS I, LOS II, FS, AO, GEN, MS	MC, AO, FS, LOS I, GEN, MS

Model Definition

MC - Market Share Model
LOS I - Level of Service Model (I)
LOS II - Level of Service Model (II)
FS - Fully Specified Model
AO - Choice Model Conditional on Auto-Ownership
GEN - Choice Model Conditional on Worker Gender
MS - Multivariately Defined Segment Models

Destination Region Definition

PD1 - Planning District 1
R.O.M - Remaining Planning Districts of Metro Toronto Region

transfer performance, particularly so, for the multivariately defined segment models. TI values correspondingly show an increase across specification.

Aggregate error measure values decline for all models and, consequently, the RATE values also decline very significantly. As in an earlier case, the improvement in aggregate performance is particularly pronounced for the multivariately defined segment models. This is because far more model parameters are reestimated for this model compared with the remaining models. Notwithstanding its big improvement, the multivariately defined segments' models yield a lower transfer log-likelihood than the gender or conditional automobile ownership models. Aggregate error measures are smaller though, but this is largely due to the update of several modal constant terms. From a practical viewpoint, updating such a model compared with the others would require substantially far more data and therefore would be unattractive.

Rescaling the model parameters yields additional significant improvement in transfer log-likelihood values, and hence TI, with the TI values for the simple level-of-service models attaining value close to I (Table 7). This would suggest that the scale parameter changed between the two temporal contexts presumably in response to the significant changes in urban character. The implication of these results is that if the constants or scales, or both, can be updated, then existing models estimated on richer data sets, collected at periods when more resources were available, can be employed in forecasting in present-day contexts. This issue is addressed in more detail in a paper by Badoe and Miller in this Record. TI values after reestimation of the modal constants and utility scale parameter do not attain a value of 1 for the better-specified models, which suggest that in addition to scale and constants changing from one context to the other, the underlying utility function parameters also may have changed.

Overall, both disaggregate and aggregate transferability measures show fairly similar trends in results. That is, in general as specification is improved the absolute and relative disaggregate transfer measures show increase, whereas the aggregate error measures show a decline in magnitude.

SUMMARY AND CONCLUSIONS

This paper examines the transferability of disaggregate demand models for a fixed urban area at two points in time—1964 and 1986—with major differences in urban conditions. It also examines the issue of model specification and transfer effectiveness. Alternate model specifications ranging from simple to complex were estimated on the 1964 work trip data, which represented the estimation context. These models were then naively transferred to the 1986 application context for forecast purposes.

Pure statistical tests of model parameters from the two urban contexts being equal reject the null hypothesis of equality, indicating that model parameters have not remained stable over time. Thus, from a theoretical viewpoint, long-range transferability is rejected. However, from a pragmatic perspective, relative measures of trans-

TABLE 7 Measures of Transferability after Updating Estimation Context Models

1964 Model Type	Transferability Measures after Re-estimation of Constants					Transferability Measures after Re-estimation of Modal Constants and Scale Parameter	
	Transfer Log-Likelihood	Transfer Index	Root Mean Square Error	Mean Absolute Error	Relative Aggregate Transfer Error	Transfer Log-Likelihood	Transfer Index
	$(LL_{86}(\theta_{64})$	$TI(c_{86})$	RMSE	MAE	RATE	$(LL_{86}(\theta_{64})$	$TI(c_{86})$
Market Share	-14996		0.35	0.24	1.00	-14996	
Level of Service (I)	-11229	0.92	0.06	0.03	1.20	-11042	0.97
Level of Service (II)	-11143	0.95	0.07	0.04	1.32	-10992	0.99
Fully Specified	-10323	0.84	0.11	0.07	3.03	-10070	0.89
Auto Ownership	-10039	0.84	0.09	0.06	2.71	-9737	0.91
Gender	-10154	0.85	0.12	0.08	3.32	-9894	0.90
Market Segment	-10451	0.57	0.03	0.01	1.11	-9810	0.88

$TI(c_{86})$ - Transfer Index Computed with 1986 Log-Likelihood of Market Share model as Base.

$Rho(\theta_i)$ - Likelihood Ratio Index computed on 1986 Data Using Model Parameters of Year *i*, ($i \in \{64,86\}$).

ferability indicate that the transferred models yield useful information in the application context. TI values show that with the exception of the models of the multivariate segments, transferred models provide at least 76 percent of the log-likelihood provided by locally estimated 1986 models and, with updating, this percent figure rises to 84. RMSE values, which range from 0.09 to 0.18, are comparable to values encountered in the literature for short-range and interurban transferability. RATE values show that use of naively transferred models result in comparatively significant aggregate error, with this error increasing with improved specification. Updating the modal constant terms significantly reduces this aggregate error.

Consistent with the findings of other transferability studies, model transferability is found to improve with improved model specification. However, the best fitting model in the estimation context did not give the best predictive performance on transfer to the application context. The choice models conditional on household automobile ownership level generally give the best model transfer performance, especially when disaggregate measures are used to evaluate transfer performance. However, simple level-of-service specifications appear to be surprisingly robust and performed very well at aggregate levels of typical planning interest. Full market segmentation specifications in the face of changing contextual factors resulted in poor transferability performance. A possible reason for this might be that extensive segmentation resulted in the models being so "trained" to the urban conditions of the estimation contexts that under the major changes that occurred in urban character, the models lost severely in predictive power.

As indicated earlier, reestimation of the modal constants or utility scale translates to significant improvements in model predictive performance, suggesting that these parameters may not have remained stable over time. Where possible, the evidence presented in this work suggests that this subset of parameters (at a minimum the modal constants) be reestimated to enhance transferability.

In sum, the reported results and findings in this work are generally consistent with and support those reported in the literature on short-term temporal transferability analysis.

ACKNOWLEDGMENTS

The work reported in this paper was funded by an operating research grant from the Natural Science and Engineering Research Council of Canada. In addition, funding for a portion of this work was provided by the Ontario Joint Transportation Research Program of the Ministry of Transportation, Ontario, on behalf of the ministry's Travel Demand Research Office. Access to the travel behavior data bases and EMME/2 network modeling system was provided by the University of Toronto Joint Program in Transportation.

REFERENCES

1. Atherton, T. J., and M. E. Ben-Akiva. Transferability and Updating of Disaggregate Travel Demand Models. In *Transportation Research Record 610,* TRB, National Research Council, Washington, D.C., 1976, pp. 12–18.
2. Hensher, D. A., and L. W. Johnson. A Two-Period Analysis of Commuter Mode Choice: The Predictive Capability of Individual Choice Models. *The Logistic and Transportation Review,* Vol. 13, No. 4, 1977, pp. 361–375.
3. Talvitie, A., and D. Kirshner. Specification, Transferability and the Effect of Data Outliers in Modeling the Choice of Mode in Urban Travel. *Transportation,* Vol. 7, No. 3, 1978, pp. 311–332.
4. Silman, L. A. The Time Stability of a Model-Split Model for Tel-Aviv. *Environment and Planning A,* Vol. 13, 1981, pp. 751–762.
5. Koppelman, F. S., and C. G. Wilmot. Transferability Analysis of Disaggregate Choice Models. In *Transportation Research Record 895,* TRB, National Research Council, Washington, D.C., 1982, pp. 18–24.
6. Gunn, H. F., M. E. Ben-Akiva, and M. A. Bradley. Tests of the Scaling Approach to Transferring Disaggregate Travel Demand Models. In *Transportation Research Record 1037,* TRB, National Research Council, Washington, D.C., 1985, pp. 21–30.
7. Koppelman, F. S., G. Kuah, and C. G. Wilmot. Transfer Model Updating with Disaggregate Data. In *Transportation Research Record 1037,* TRB, National Research Council, Washington, D.C., 1985, pp. 102–107.
8. McCoomb, L. A. Analysis of the Transferability of Disaggregate Demand Models Among Ten Canadian Cities. *Transportation Forum,* Vol. 3-1, 1986, pp. 102–107.

9. Koppelman, F. S., and C. G. Wilmot. The Effect of Omission of Variables on Choice Model Transferability. *Transportation Research,* Vol. 20B, No. 3, 1986, pp. 205–213.
10. Parody, T. E. An Analysis of Disaggregate Mode Choice Models in Prediction. In *Transportation Research Record 637,* TRB, National Research Council, Washington, D.C., 1977, pp. 51–57.
11. Train, K. E. A Comparison of the Predictive Ability of Mode Choice Models with Various Levels of Complexity. *Transportation Research,* Vol. 13A, No. 1, 1979, pp. 11–16.
12. McCarthy, P. S. Further Evidence on the Temporal Stability of Disaggregate Travel Demand Models. *Transportation Research,* Vol. 16B, No. 4, 1982, pp. 263–278.
13. Ben-Akiva, M. E. Issues in Transferring and Updating Travel and Updating Travel-Behavior Models. In *New Horizons in Travel-Behaviour Research* (P. R. Stopher, A. H. Meyburg, and W. Brog, eds.), D.C. Heath and Co., Lexington, Mass., 1981.
14. Horowitz, J. L. Random Utility Travel Demand Models. In *Transportation and Mobility in an Era of Transition* (G. R. Jansen, P. Nijkamp, and C. J. Ruijgrok, eds.), Elsevier Science Publishers B.V., Amsterdam, 1985, pp. 141–156.
15. Sonquist, J. A., E. L. Baker, and J. H. Morgan. *Searching for Structure.* Institute for Social Research, University of Michigan, Ann Arbor, 1971.
16. Badoe, D. A. *An Investigation into the Long Range Transferability of Work-Trip Discrete Choice Models,* Ph.D. dissertation. Department of Civil Engineering, University of Toronto, Ontario, Canada, 1994.
17. Daganzo, C. F. The Statistical Interpretation of Predictions with Disaggregate Demand Models. *Transportation Science,* Vol. 13, No. 1, 1979, pp. 1–12.

Publication of this paper sponsored by Committee on Passenger Travel Demand Forecasting.