

# **A COMPARISON OF METHODS FOR TRANSFERRING LOGIT MODELS OF GAP- ACCEPTANCE BEHAVIOUR**

*Rossi Riccardo - University of Padova, e-mail: riccardo.rossi@unipd.it*

*Meneguzzer Claudio - University of Padova, e-mail: claudio.meneguzzer@unipd.it*

*Gastaldi Massimiliano - University of Padova, e-mail: massimiliano.gastaldi@unipd.it*

## **ABSTRACT**

Alternative methods for transferring Logit models of gap-acceptance behaviour are implemented and compared in this study using experimental data collected at two priority intersections. The effectiveness of model transfer is evaluated on the basis of several indicators proposed in the literature regarding travel demand applications of transferability, and the results are compared to those obtained by locally estimating the model in the application context. The main conclusion is that the accuracy of the transferred models is generally similar to that of the locally estimated ones, and that the method known as *Combined Transfer Estimation* performs best among the tested approaches. These results are not significantly affected by the size of the sample of observations used for model transfer.

*Keywords: priority intersections, gap-acceptance behaviour, model transfer.*

## **INTRODUCTION**

Gap-acceptance behaviour is an important determinant of the operational performance of priority intersections. Since this type of behaviour is strongly affected by several site-specific characteristics, one possible approach is to specify and estimate gap-acceptance models with specific reference to each application context. However, this approach is clearly expensive in terms of data collection costs. As an alternative, it seems reasonable that a gap-acceptance model specified and estimated in one given context could be transferred to similar contexts and/or used in the same context after a significant number of years.

This issue, which is generally known as spatial and/or temporal model transferability, has been studied by several authors with reference to travel demand models (for example, mode choice models), but in the literature there appear to be no applications specifically regarding gap-acceptance models. The purpose of this paper is to describe a comparative analysis of alternative methods of model transfer applied to a Logit model of gap-acceptance behaviour.

The performance of these methods is evaluated on the basis of several indicators of transfer effectiveness.

## **LOGIT MODELS OF GAP-ACCEPTANCE BEHAVIOUR**

The gap-acceptance problem considered in this paper refers to the situation in which a driver, starting from the secondary approach of a priority intersection, wants to perform a crossing or merging maneuver into a primary road. Essentially, this requires the choice between two mutually exclusive alternative actions: to accept or reject a gap (or lag<sup>1</sup>) of a given time size in the primary traffic stream. Evidently, such a choice is the result of a decision process affected both by driver characteristics (for example, driving experience, sex and age; see Wennel and Cooper 1981, Teply et al. 1997a,b) and characteristics of the choice situation (for example, gap/lag size, waiting time and speed of vehicles on the primary road, see Adebisi and Sama 1989, Polus et al. 1996). Thus, as shown by several previous studies, gap-acceptance behaviour varies among drivers and, for the same driver, over time. Probabilistic discrete choice models and, in particular, Logit models are considered to be appropriate for modelling the choice behaviour under examination. Indeed, several applications of the Logit model to the representation of gap-acceptance behaviour can be found in the literature; see, for example, Cassidy et al. (1995), Teply et al. (1997a,b), and Maze (1981).

## **MODEL TRANSFERABILITY**

In general terms model transferability (spatial and/or temporal) refers to situations in which a model specified and estimated in a given original (estimation) context is subsequently transferred and applied to another (application) context. The basic idea is that the parameters estimated in the original context can be used to improve the accuracy of parameter estimation in the application context.

The main reasons that support the idea of model transfer are:

- it reduces the efforts in model development (the same structure of the model previously identified is used);
- it reduces or eliminates the need for a large data collection in the application context.

Model transferability has been widely studied in the past with reference to trip generation models (Agyemang-Duah and Hall 1997), mode choice models (Atherton and Ben-Akiva 1976, Badoe and Miller 1995, Koppelman, Kuah and Wilmot 1985, Koppelman and Wilmot 1982), and four-step models (Karasmaa 2007). These authors studied the effectiveness of both full model transfer (direct transfer from the estimation to the application context without

---

<sup>1</sup> A lag is the time interval between the arrival of a vehicle at the stopline of the secondary road and the passage of the first vehicle on the main road.

updating the parameters) and updating the original model using a small dataset for the application context.

## **Model transferability and updating methods**

Obviously, the simplest way to transfer an original model to another context is to use directly the model without updating the transferred coefficients (*direct transfer*); this implies the assumption that the choice behaviour in the application context is exactly the same as in the estimation context. In a study of transferability of trip generation models, Agyemang-Duah and Hall (1997) report that satisfactory results may be obtained by direct transfer, based on a comparison of predicted shares in the application context with the observed choices in the same context; this is in agreement with the findings previously reported in Atherton and Ben-Akiva (1976) regarding choice models transfer.

The effectiveness of model transfer can be improved using an updating procedure based on a small sample of choice observations in the application context. In this case the problem is essentially the determination of the size of the sample that guarantees an adequate model transfer/update. Updating procedures have the obvious advantage of reducing data collection costs in the application context as compared to a full re-estimation of the model. A review of the existing literature (for example, Agyemang-Duah and Hall 1997, Atherton and Ben-Akiva 1976, Badoe and Miller 1995, Koppelman, Kuah and Wilmot 1985, Koppelman and Wilmot 1982, Karasmaa 2007) indicates generally good performance and applicability of updating procedures.

In particular, Badoe and Miller (1995) studied transferability of mode choice models under the following hypotheses:

- a disaggregate multinomial Logit choice model is to be transferred from an estimation to an application context
- the estimation context model parameters are known (by a calibration process using choice observations collected in the same context)
- the original model and the model to be transferred have the same specification
- a small set of data from the application context is available.

They varied the size of the application context data sample (randomly extracted from the full sample), and tested the effectiveness of the model transfer by comparing the choices predicted by the transferred model with those observed in the application context. In a similar way, Karasmaa (2007) tested different updating methods with reference to four-step model transferability, and used a “bootstrap” method to extract the samples from the full application context dataset. Both studies have focussed on the performance of four updating methods:

- Bayesian Updating (Atherton and Ben-Akiva 1976);
- Transfer Scaling (Gunn, Ben-Akiva and Bradley 1985, Koppelman, Kuah and Wilmot 1985);

- Combined Transfer Estimation (Ben-Akiva and Bolduc 1987);
- Joint Context Estimation (Ben-Akiva and Morikawa 1990).

The main factors affecting the effectiveness of the models updated through these methods are the level of specification of the original model, the size of the sample in the application context and the availability or not of the estimation dataset (in particular, the Joint Context Estimation method needs the datasets of both estimation and application contexts). In general terms, as emphasized in Badoe and Miller (1995), updating a model estimated in a given context through the use of a small data sample in another context significantly improves the model's transferability to this context, as compared to direct transfer.

The primary aim of our study is to assess the results of a comparative analysis of alternative transfer methods applied to a Logit model of gap-acceptance behaviour.

### Model transferability measures

The evaluation of the effectiveness of transferred models is performed through a set of measures classified as follows (Koppelman and Wilmot 1982):

- Tests of model parameter equality
- Tests of disaggregate prediction
- Tests of aggregate prediction

The first class includes tests of equality between model parameters; the tested hypothesis is that the choice process can be described by a common model. The limit of this approach is that such tests are symmetric, while transferability is usually a "directional" property. The second class of tests includes indicators useful for measuring the capability of the transferred model to describe the individual choices observed in the application context. These indicators are computed using log-likelihood measures (Fig. 1).

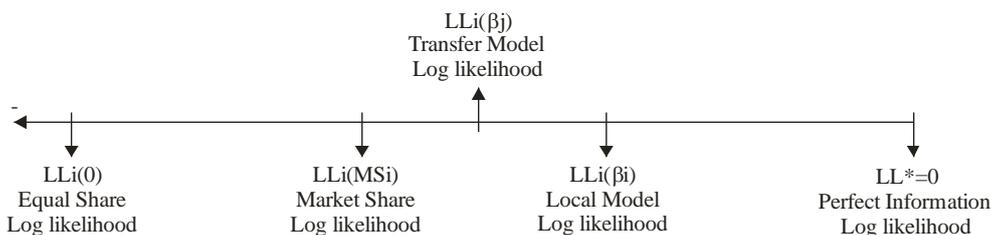


Fig. 1 - Log-likelihood measures for estimated and transferred models. (Source: Koppelman and Wilmot 1982).

In Figure 1 the subscript  $j$  refers to the original (estimation) context and the subscript  $i$  to the application context. Starting from the difference (assumed as a natural measure of transferability) between the value of the log-likelihood function computed for the model transferred from the estimation context and updated with the application context data, and the log-likelihood computed for the model locally estimated in the application context,

$\{-[LL_i(\beta_j) - LL_i(\beta_i)]\}$ , the authors formulate three specific indicators of disaggregate transferability:

*Transferability Test Statistic*

$$TTS_i(\beta_j) = -2[LL_i(\beta_j) - LL_i(\beta_i)] \quad (1)$$

*Transfer Index*

$$TI_i(\beta_j) = [LL_i(\beta_j) - LL_i(MS_i)] / [LL_i(\beta_i) - LL_i(MS_i)] \quad (2)$$

*Transfer rho-square*

$$\rho_i^2(\beta_j) = [LL_i(\beta_j) - LL_i(MS_i)] / [LL_i^* - LL_i(MS_i)] = 1 - [LL_i(\beta_j) / LL_i(MS_i)] \quad (3)$$

The third class of transferability measures evaluates the transferred model in terms of its ability to produce aggregate predictions. Koppelman and Wilmot (1982), in the context of an analysis of the transferability of modal choice models, have introduced a relative error measure:

$$REM_{mg} = (\hat{N}_{mg} - N_{mg}) / \hat{N}_{mg} \quad (4)$$

where

$\hat{N}_{mg}$  is the number of individuals from group  $g$  predicted to choose alternative  $m$

and

$N_{mg}$  is the number of individuals from group  $g$  observed to choose alternative  $m$

and using this measure have proposed three indicators of aggregate transferability:

*the weighted root mean square error measure*

$$RMSE = \left( \frac{\sum_{m,g} \hat{N}_{mg} REM_{mg}^2}{\sum_{m,g} \hat{N}_{mg}} \right)^{1/2} \quad (5)$$

*the aggregate prediction statistic*

$$APS = \sum_{m,g} \hat{N}_{mg} REM_{mg}^2 = \sum_{m,g} (\hat{N}_{mg} - N_{mg})^2 / \hat{N}_{mg} \quad (6)$$

For a more detailed description and interpretation of these measures, see Koppelman and Wilmot (1982).

## DESCRIPTION OF THE ESTIMATION AND APPLICATION CONTEXTS

As observed previously, Logit applications to gap-acceptance behaviour modeling are described in several papers, but studies about transferability of these models are not reported in the literature to the authors' knowledge. In this paper, we carry out a transferability analysis of Logit models of gap-acceptance for right-turning vehicles from the minor street of priority intersections.

### Experimental data

The field data used for the transferability analysis are gap-acceptance observations collected at two three-leg intersections, indicated by "E" (Estimation context) and "A" (Application context) in the remainder of the paper (see Figure 2). All observations relate to the right turn movement from a minor street controlled by "Stop" sign. Both intersections are located outside urban areas, and have a different geometric layout in terms of the angle between the two intersecting roads. The observations were collected in 1999 at intersection "E" and in 2009 at intersection "A"; therefore, both a spatial and a temporal transfer are performed in our study.

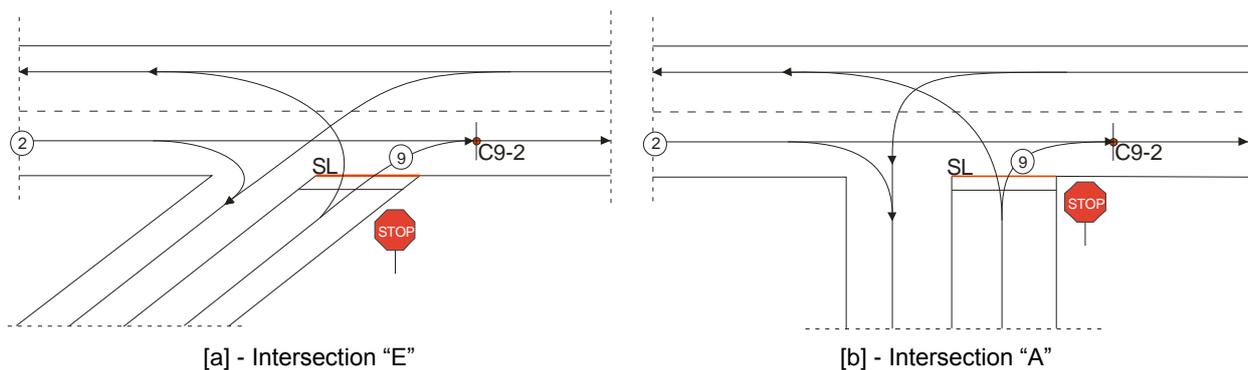


Fig. 2 – Layout of the analyzed intersections.

The experimental observations were collected during peak-hour periods through video camera recorder. The videos were processed using an application software that allows the user to record the secondary vehicle arrival and departure at the stop line (SL in Fig. 2), the primary vehicle arrival at the conflict point (C9-2) together with the vehicle category (car, van, truck, etc.). The data were organised in a database and then processed through a software procedure that allows to extract the following information associated to each driver's decision:

- Type of time interval (lag or gap)
- Interval size
- Waiting time of the secondary street vehicle at the stop line
- Class of secondary street vehicle

- Class of primary street vehicle closing the interval
- Driver decision (interval acceptance or rejection)

A summary of the data used in the transferability analysis is shown in Tab. 1.

Tab. 1 - Intersections E (Estimation) and A (Application). Sample sizes of observed decisions.

Maneuver	Intersection	Observation period length (hours)	Type of interval	Total number of decisions (acceptances and rejections)
Right turn from minor road	E	6	Gap	1.448
			Lag	892
			Total	2.340
	A	4,5	Gap	1.265
			Lag	639
			Total	1.904

## ESTIMATION OF THE ORIGINAL LOGIT MODEL

Several Logit models of gap-acceptance behaviour were specified and tested in this study, but only some of these resulted statistically significant; among these, the model including interval size and type (the latter represented by a dummy) as explanatory variables, and indicated as E\_s\_lg, was selected (Tab. 2 and 3). The Gauss® program was used for model estimation.

Tab. 2 - Intersection "E". Model diagnostic tests.

Model	$\rho^2$	$\rho_c^2$	Average probability of chosen alternative	Percent right
E_s_lg	0,742	0,740	89,44%	92,35%

Tab. 3 - Intersection "E". Estimation results: model parameters and corresponding *Student's t-statistic* (within brackets).

Parameters	Model "E_s_lg"
Alternative specific constant (acceptance)	-6,26 (-22,5)
Interval size	1,0 (22,4)
Interval type	2,0 (10,4)

This model specification appears to be consistent with results reported in the literature (for example, Ashworth and Bottom 1977, Solberg and Oppenlander 1965, Wagner 1965), which suggest a significant effect of the above variables on driver gap-acceptance behaviour. The estimated model (8) indicates that the acceptance probability increases (as expected) with the interval size and that, for the same interval size, the gap-acceptance probability is lower than the lag-acceptance probability. This result can be explained if we consider the specific geometric layout of the "E" intersection, which allows vehicles making a right turn from the minor road to enter the intersection without stopping, provided a lag of acceptable size is

available. This is obviously not possible if the accepted interval is a gap, because in this case the vehicle has to stop before completing the maneuver.

$$P_{E\_s\_lg, acceptance} = \frac{1}{1 + e^{-(6,26+1,0 \cdot s+2,0 \cdot lg)}} \quad (8)$$

where:

s : time interval size (seconds)

$$lg = \begin{cases} 1 & \text{if interval type} = lag \\ 0 & \text{if interval type} = gap \end{cases}$$

Some further observations on these results are as follows:

- the model (8) shows a good capability to represent the observed driver behaviour (see the statistics in Tab. 2);
- the value of the alternative-specific constant (acceptance) is dominant compared to the other parameter values;
- the contribution of the interval type attribute is statically significant and relatively important.
- In the following transferability analysis, model  $E\_s\_lg$  has been chosen as the original (estimation) model.

## TRANSFERABILITY ANALYSIS

### Application context sampling

The full application context dataset consists of 1.904 decisions, collected in three peak periods, each one hour and half long. The model transferring/updating procedures have been carried out with reference to six subsets of the application context full dataset, extracted from it considering a partition based on the three peak-hour periods ( $p$ - $hp$ ); each larger sample data subset ( $sds$ ) contains all the decisions included in the smaller ones. In order to evaluate the effect of the different combinations of the three  $p$ - $hp$  on  $sds$  composition and then on the model transfer effectiveness, the sample datasets shown in Tab. 4 have been considered. The  $sds$ -123 is the full dataset of the application context.

Tab. 4 - Application context. Sample data subsets (*sds*) drawn from peak-hour periods (*p-hp*).

Nr.	Sample dataset	p-hp1	p-hp2	p-hp3	number of decisions
1	sds-1	x			647
2	sds-2		x		628
3	sds-3			x	629
4	sds-12	x	x		1.275
5	sds-13	x		x	1.276
6	sds-23		x	x	1.257
7	sds-123	x	x	x	1.904

This sampling criterion is justified from an application point of view; in fact, it seems plausible that, in an operational analysis of a priority intersection, the primary need of the analyst is to implement an effective gap-acceptance model (starting from an original model already estimated in another context) using the smallest possible number of new observations. Realistically, these observations will refer to the same (peak-hour) time period.

### **Model transfer and updating methods**

In this work *Direct Transfer*, *Transfer Scaling (TS)*, *Bayesian Updating (BU)* and *Combined Transfer Estimation (CTE)* have been compared. Since the Joint Context Estimation method requires the availability of the full dataset of the original context, it has been excluded from the comparison because, typically, in an operational analysis of gap-acceptance behaviour such dataset is unlikely to be known.

## **EFFECTIVENESS OF THE TRANSFERRED MODELS**

As previously indicated, the Logit gap-acceptance model  $E_s_{lg}$  has been used as original model. First, a direct transfer, that is a transfer without any updating of coefficients, has been performed (model *dtm-123*). Then, different transfers of the original model to the application context with six sample datasets have been carried out. The effectiveness of the transferred models (*TS*, *BU*, *CTE*) has been evaluated with respect to the corresponding models estimated in the application context (*em*). The estimated and transferred model parameters for various application context sample datasets and methods of transfer are shown in Table 5.

Tab. 5 – Estimated and transferred model parameters for various application context sample datasets and transfer methods.

Sample data set	Type of model	Transfer method	Model	Estimated parameters		
				s	lg	constant
sds-1	estimated	-	em-1	1,79	1,81	-8,25
		TS	TS-1	1,58	3,16	-8,01
	transferred	BU	BU-1	1,00	1,75	-5,95
		CTE	CTE-1	1,77	1,81	-8,20
sds-2	estimated	-	em-2	1,62	1,59	-7,67
		TS	TS-2	1,40	2,80	-7,36
	transferred	BU	BU-2	1,01	1,81	-6,04
		CTE	CTE-2	1,60	1,60	-7,62
sds-3	estimated	-	em-3	1,40	1,30	-6,60
		TS	TS-3	1,21	2,42	-6,36
	transferred	BU	BU-3	0,99	1,67	-5,81
		CTE	CTE-3	1,39	1,31	-6,58
sds-12	estimated	-	em-12	1,70	1,69	-7,94
		TS	TS-12	1,48	2,96	-7,66
	transferred	BU	BU-12	1,03	1,67	-5,91
		CTE	CTE-12	1,69	1,69	-7,91
sds-13	estimated	-	em-13	1,58	1,53	-7,36
		TS	TS-13	1,38	2,76	-7,11
	transferred	BU	BU-13	1,01	1,58	-5,77
		CTE	CTE-13	1,57	1,53	-7,34
sds-23	estimated	-	em-23	1,50	1,44	-7,10
		TS	TS-23	1,30	2,60	-6,83
	transferred	BU	BU-23	1,01	1,62	-5,82
		CTE	CTE-23	1,49	1,44	-7,08
sds-123	estimated	-	em-123	1,59	1,55	-7,45

We observe that transferred and locally estimated models have the same parameter signs for all sample datasets.

Looking at the model parameters in more detail, it appears evident that the CTE models are always the closest to the corresponding estimated models. The TS model parameters are not so close to those of the corresponding estimated models, and this is especially true for the *lg* coefficient. The BU model parameters are significantly different from those of the estimated models.

These results are confirmed by the acceptance probability curves for both locally estimated and transferred models, that are separately shown for gaps and lags in Figures 3 and 4 (the curves labelled “em-123” represent the models estimated locally with the full dataset of the application context).

For gaps (Figure 3), these curves suggest that the BU and TS transferred models consistently underestimate the acceptance probability as compared to the locally estimated models (with a higher bias for BU models) regardless of the sample size. These differences

are particularly evident for gaps in the range from 3,5 to 7-9 seconds, where typically the gap-acceptance behaviour is more uncertain.

In the case of lags (Figure 4), the main difference is that the TS models tend to overestimate the acceptance probability as compared to the locally estimated models, while the behaviour of the BU models remains qualitatively the same as before.

In general, we note that the sample size has a limited effect on these trends.

For both gaps and lags, the CTE models produce acceptance probabilities that are very similar to those obtained from the locally estimated models, and this result is essentially independent of both sample size and time interval size. This finding is in agreement with those of Ben Akiva and Bolduc (1987) and Badoe and Miller (1995).

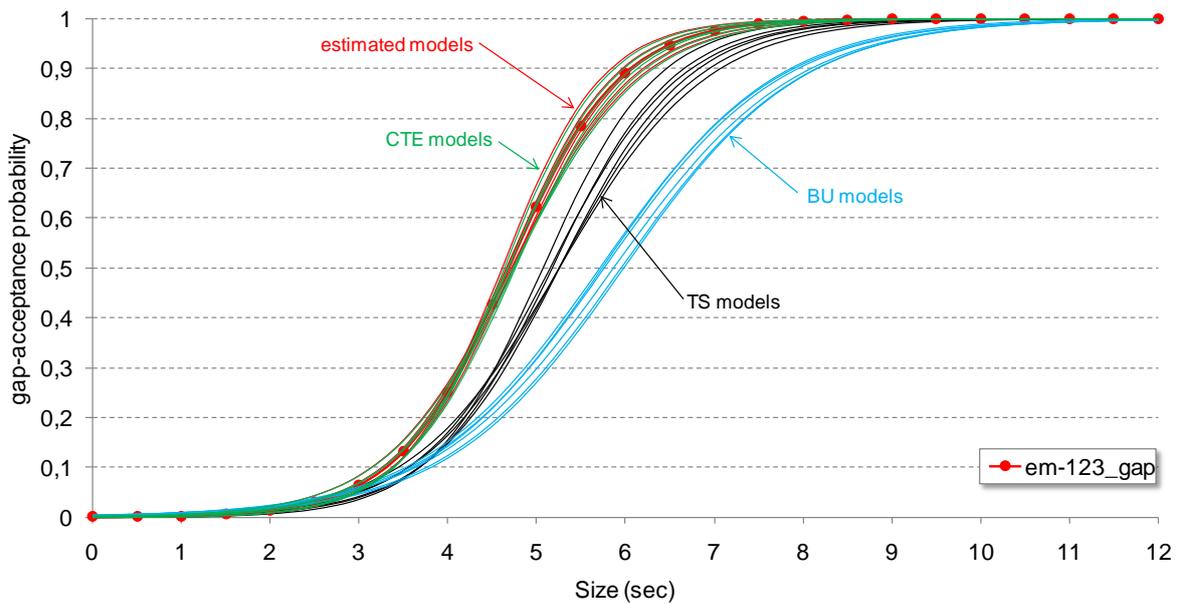


Fig. 3 – Logit models. Probability of gap-acceptance for estimated and transferred models.

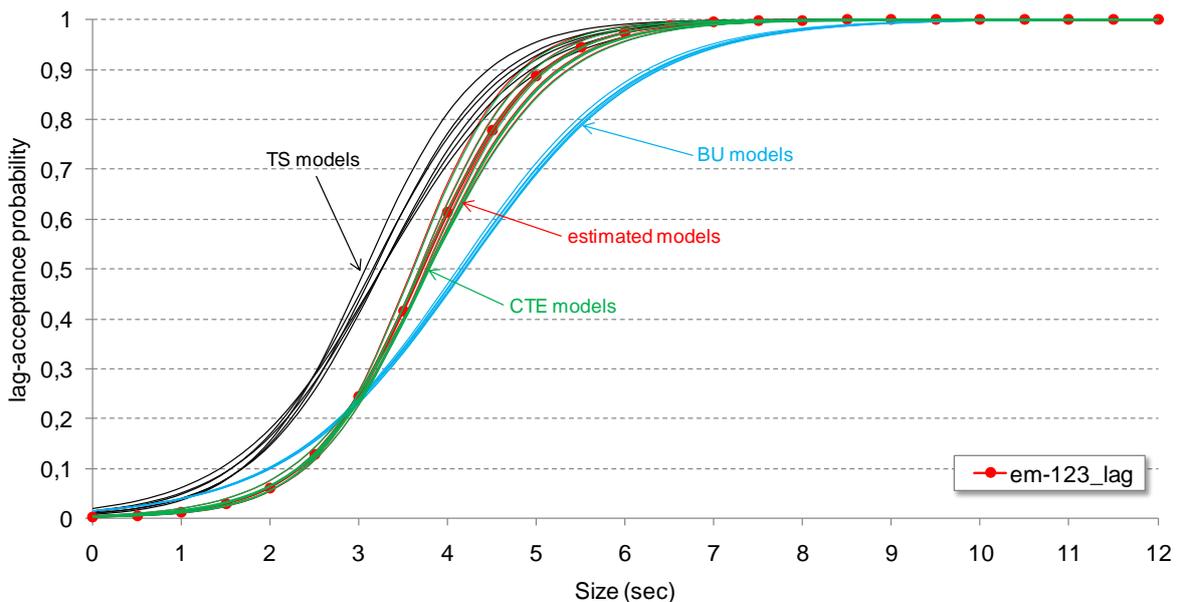


Fig. 4 – Logit models. Probability of lag-acceptance for estimated and transferred models.

In order to evaluate the effectiveness of the transferred models, the indicators previously described in this paper have been used. The results of this analysis are shown in table 6, where the critical chi-square values at 5 percent significance level are reported within brackets.

Tab. 6 – Effectiveness measures of model transfer as a function of sample dataset size (critical chi-square values are within brackets). Comparison between models transferred, directly transferred and estimated with various application context sample datasets.

Application context	Model	TTS	TI	$\rho^2$	RMSE	APS
sds1	em-1	5,236 (7,81)	0,997	0,731	0,025	1,169 (5,99)
	TS-1	60,557 (7,81)	0,966	0,708	0,023	0,982 (5,99)
	BU-1	135,25 (7,81)	0,924	0,678	0,161	49,293 (5,99)
	CTE-1	4,264 (7,81)	0,998	0,731	0,029	1,648 (5,99)
sds2	em-2	0,993 (7,81)	0,999	0,733	0,039	2,864 (5,99)
	TS-2	54,645 (7,81)	0,969	0,711	0,056	5,909 (5,99)
	BU-2	138,62 (7,81)	0,922	0,676	0,164	50,902 (5,99)
	CTE-2	1,412 (7,81)	0,999	0,733	0,041	3,223 (5,99)
sds3	em-3	5,928 (7,81)	0,997	0,731	0,041	3,223 (5,99)
	TS-3	60,265 (7,81)	0,966	0,708	0,055	5,664 (5,99)
	BU-3	124,89 (7,81)	0,930	0,682	0,153	44,660 (5,99)
	CTE-3	6,398 (7,81)	0,996	0,731	0,042	3,410 (5,99)
sds12	em-12	1,561 (7,81)	0,999	0,733	0,034	2,212 (5,99)
	TS-12	55,216 (7,81)	0,969	0,710	0,047	4,122 (5,99)
	BU-12	104,25 (7,81)	0,941	0,690	0,139	36,878 (5,99)
	CTE-12	1,379 (7,81)	0,999	0,733	0,034	2,212 (5,99)
sds13	em-13	0,147 (7,81)	1,000	0,733	0,032	1,919 (5,99)
	TS-13	53,943 (7,81)	0,970	0,711	0,044	3,718 (5,99)
	BU-13	104,35 (7,81)	0,941	0,690	0,139	36,878 (5,99)
	CTE-13	0,101 (7,81)	1,000	0,733	0,034	2,212 (5,99)
sds23	em-23	1,447 (7,81)	0,999	0,733	0,041	3,223 (5,99)
	TS-23	55,219 (7,81)	0,969	0,710	0,053	5,426 (5,99)
	BU-23	110,02 (7,81)	0,938	0,688	0,143	38,928 (5,99)
	CTE-23	1,907 (7,81)	0,999	0,732	0,044	3,604 (5,99)
sds123	em-123	-	-	-	0,035	2,367 (7,81)
	dtm-123	185,49 (7,81)	0,896	0,657	0,182	62,985 (7,81)

With reference to the TTS measure, none of the TS and BU models is significant at the 0.05 level: the hypothesis of model transferability between the original and the application context is rejected regardless of the sample dataset size in the application context; BU models are always worse than the corresponding TS models despite their greater computational complexity.

For all the locally estimated models (*em*) and for all the CTE models the TTS measures are below the corresponding critical chi-square values, indicating a capability of these models to represent driver behaviour comparable with that of the model estimated on the application context full dataset (*sds123*).

The TI measure varies from 0,966 to 0,970 for the TS models, from 0,922 to 0,941 for the BU models, from 0,996 to 1,0 for the CTE models and from 0,997 to 1,0 for the locally estimated models. The TI values estimated for the transferred models appear to be high if compared with those reported in the literature (Koppelman and Wilmot 1982, Karasmaa 2007); this result (unlike the TTS tests for TS and BU models) suggests that all the transferred models effectively represent the choice behaviour observed in the application context. Nevertheless, the better models are those estimated in the application context and the CTE models.

The results obtained from the rho-square measure are qualitatively similar to those described for TI. In particular, we note that for all the TS, CTE and *em* models the rho-square measure is always over 0,7, while all the BU models appear to be less effective.

The values of RMSE and APS provide a measure of the aggregate prediction ability of the models; three vehicular classes (cars, vans and trucks) have been used in the computation of these indicators. We note that, with respect to both measures, the CTE and TS models show a performance similar to that of the corresponding locally estimated models, while the BU models are the worst regardless of the sample size.

The APS measures are significant at the 0.05 level for all the TS and CTE models, and for all the locally estimated models, but not for the BU models.

Finally we note that, as expected, the directly transferred model (*dtm*) shows the worst performance with respect to all indicators.

## CONCLUSIONS

The transferability of Logit models of gap-acceptance behaviour has been evaluated in the present study using the Transfer Scaling, Bayesian Updating and Combined Transfer Estimation methods. The performance of different transferred models, characterized by varying size of the sample dataset used for updating, has been analyzed in comparison to corresponding models estimated locally in the application context. Several indicators for the evaluation of model transferability have been computed.

Our analysis shows that transferred models (excluding Bayesian Updating models and directly transferred model) are as effective in representing gap/lag acceptance behaviour as the locally calibrated models. Moreover, at least for the specific problem and case study considered, the above results seem to be essentially independent of the size of the sample dataset used for model transfer/estimation. In practice, with a small sample of observations in the application context (1,5 hours of data) it seems possible to obtain good transferred models (especially using Combined Transfer Estimation), that are able to represent gap/lag acceptance behaviour as well as the model estimated locally on a 4,5 hour sample dataset.

The use of transferred instead of locally estimated models may be preferable even for the same sample size: as stated in Karasmaa (2007), the variances of the parameters of the former are generally lower than those of the corresponding locally calibrated models.

The specific conclusions of our study agree with those of previous research on model transferability; in particular:

- despite its low computational complexity, the Transfer Scaling method using a small sample size in the application context appears reasonably effective in model transfer;

- Combined Transfer Estimation using a small sample size in the application context allows to obtain transferred models as effective as the corresponding locally estimated models;
- the Bayesian Updating method, which does not consider transfer bias explicitly, produces transferred models that are worse than those obtained using the TS and CTE methods;
- the directly transferred model shows the worst performance, indicating that the update of model parameters using a small data set in the application context substantially improves the predictive ability of the model.

Future research should focus on the following issues:

- testing the effect of intersection and site characteristics on the effectiveness of model transfer; this would require a more extensive collection of data for a larger number of intersections. In our application, intersection geometry was significantly different between the original and the application context, and a time transfer of approximately ten years was involved;
- developing and transferring gap-acceptance models that include a larger set of independent variables, in order to test the effect of the level of specification on model transferability.

## REFERENCES

- Adebisi O. and Sama G. N. (1989). Influence of stopped delay on driver gap acceptance behaviour. *Journal of Transportation Engineering*, 115 (3), pp. 305-315.
- Agyemang-Duah K. and Hall F. L. (1997). Spatial transferability of an ordered response model of trip generation. *Transportation Research A*, 31A (5), pp. 389-402.
- Atherton T. J. and Ben-Akiva M. E. (1976). Transferability and Updating of Disaggregate Travel Demand Models. *Transportation Research Record*, 610. Transportation Research Board, National Research Council, Washington, DC, pp. 12-18.
- Ashworth R. and Bottom C.G. (1977). Some observations of driver gap-acceptance behaviour at a priority intersection. *Traffic Engineering and Control*, 18 (12), pp. 569-571.
- Badoe D. A. and Miller E. J. (1995). Comparison of Alternative Methods for Updating Disaggregate Logit Mode Choice Models. *Transportation Research Record*, 1493. Transportation Research Board, National Research Council, Washington, DC, pp. 90-100.
- Ben-Akiva M. E. and Bolduc D. (1987). Approaches to Model Transferability and Updating: the Combined Transfer Estimator. *Transportation Research Record*, 1139. Transportation Research Board, National Research Council, Washington, DC, pp. 1-7.

- Ben-Akiva M. E. and Morikawa T. (1990). Estimation of travel demand models from multiple data sources. Proceedings 11<sup>th</sup> International Symposium on Transportation and Traffic Theory, Yokohama.
- Cassidy M. J., Madanat S. M., Wang M-H and Yang F. (1995). Unsignalized intersection capacity and level of service: revisiting critical gap. *Transportation Research Record*, 1484. Transportation Research Board, National Research Council, Washington, DC, pp. 16-22.
- Gunn H. F., Ben-Akiva M. E. and Bradley M. A. (1985). Tests of the Scaling Approach to Transferring Disaggregate Travel Demand Models. *Transportation Research Record*, 1037. Transportation Research Board, National Research Council, Washington, DC, pp. 21-30.
- Karasmaa N. (2007). Evaluation of transfer methods for spatial travel demand models. *Transportation Research A*, (A41), pp. 411-427.
- Koppelman F. S. and Wilmot C. G. (1982). Transferability Analysis of Disaggregate Choice Models. *Transportation Research Record*, 895. Transportation Research Board, National Research Council, Washington, DC, pp. 18-24.
- Koppelman F. S., Kuah G. K., Wilmot C. G. (1985). Transfer Model Updating with Disaggregate Data. *Transportation Research Record*, 1037. Transportation Research Board, National Research Council, Washington, DC, pp. 102-107.
- Maze T.H. (1981). A probabilistic model of gap acceptance behavior. *Transportation Research Record* 795, Transportation Research Board, National Research Council, Washington, DC, pp. 8-13.
- Polus A., Craus J. and Reshetnik I. (1996). Non-stationary gap acceptance assuming drivers' learning and impatience, *Traffic Engineering and Control*, 37(6), pp. 395-402.
- Solberg P. and Oppenlander J.C. (1965). Lag and gap acceptances at stop-controlled intersections. *Highway Research Record*, 118, pp. 48-67.
- Teply S., Abou-Henaidy M. I. and Hunt J. D. (1997a). Gap acceptance behaviour – aggregate and Logit perspectives: Part 1, *Traffic Engineering and Control*, 9, pp. 474-482.
- Teply S., Abou-Henaidy M. I. and Hunt J. D. (1997b). Gap acceptance behaviour – aggregate and Logit perspectives: Part 2, *Traffic Engineering and Control*, 10, pp. 540-544.
- Wagner F.A. (1965). An evaluation of fundamental driver decisions and reactions at an intersection. *Highway Research Record* 118, pp. 68-84.
- Wennel J. and Cooper D. F. (1981). Vehicle and driver effects on junction gap acceptance. *Traffic Engineering and Control*, 22 (12), pp. 628-632.