A COMPARISON OF DISAGREGGATE MODE CHOICE MODELS IN DIFFERENTE ENVIRONMENTS

J. de Ortúzar and J.E. Fernández
Department of Transport Engineering,
Pontificia Universidad Católica de Chile,
Casilla 114-D, Santiago, Chile

1. INTRODUCTION

One of the more important debates among transport scientists and practicioners is whether models estimated for a given situation can be used with confidence (after minor adjustments) in a different one; ie. the stability or transferability issue, both temporal and/or accross different cultures.

In this paper the same model structure (except for the number of alternatives) is estimated using two different data sets. The first one gathered in 1975 for a suburban corridor in Leeds, England, and the second one gathered in 1981 for an urban corridor in Santiago, the capital of Chile. In both cases alternatives range from car alone to rail (underground in Santiago) with several combinations like park-and-ride (P & R), kiss-and-ride (K & R) and feeder bus services. Both sets comprise roughly 600 observations and present almost the same average household car ownership value. This confers additional value to the comparative analysis done as car ownership (usually a crucial factor in explaining mode choice) can be considered a constant; therefore any differences found in the estimated models may be considered a product of the different economic and cultural environments of the two areas.

The model to be examined has the following multinomial logit form (Mc Fadden, 1974):

$$P_j = \frac{\exp(\overline{U}_j)}{\sum_k \exp(\overline{U}_k)} \qquad (1)$$

where : P_j = probability of choosing option j, j = 1,...., N

$\overline{U}_j$ = measurable utility of alternative j

$= \sum_i \Theta_i Z_{ji}$

Θ_i = parameters to be estimated

Z_{ji} = attribute i of alternative j.

The attributes considered are travel times and costs, and socio-economic variables of each individual's household.

The paper is organised as follows. In section 2 both data sets and settings are briefly described. Section 3 summarises the results obtained using English data, while section 4 does the same for the Chilean case. Finally, section 5 joins everything together and presents our conclusions.

2. DESCRIPTION AND COMPARISON OF THE DATA SETS.

Apart from the fact that England and Chile are very different themselves (a developed and a developing country respectively) the corridors and the data used were also of a somewhat different nature, as can be seen in the summary given in Table 1. For example, Santiago is a big capital city of 4 million inhabitants with a rather low density which makes it very extended; while Leeds is a medium-sized provincial city (of half a million inhabitants) which constitudes the core of the West Yorkshire conurbation of some 1 million inhabitants. Other differences, especially regarding the modes available, train frequencies, fares and operating costs are shown in the table. The corridors are depicted in Figures 1 and 2.

The two samples, although of similar size, also differ markedly because important pieces of information, such as income and data about location and cost of parking, were missing from the 1975 Leeds survey (for a good description of the data and its origins, see Ortúzar, 1980). The 1981 Santiago survey, on the other hand, was especially designed to produce appropriate data for calibrating and testing discrete choice models (Ortúzar and Donoso, 1983).

As we mentioned above, both samples present almost the same value of household car ownership. This is due to the fact that the area analyzed in the Chilean case comprises several residential zones inhabited by people belonging to the highest income strata in the country, whose socio-economic characteristics resemble those of people living in an industrialized nation.

In both cases alternatives were assigned objectively, based on the same availability criteria (ie. maximum walking distances, car ownership, etc.). However, in the Santiago case we also had information on reported alternatives so we could compare the impact in model coefficients of using either type of information in the calibration stage (see Ortúzar, Donoso and Hutt, 1983).

Table 2 shows the frequency of choice and availability of each alternative in both samples. We will discuss the table fully in section 5.

3. RESULTS USING ENGLISH DATA.

Several MNL models were estimated (Ortúzar, 1980; 1983), but we will show here only the one judged to be more adequate. Firstly though, it is appropriate to briefly present the variables considered in the specification searches, and the statistical measures used as one of the ways of comparing performance. This is done in Tables 3 and 4. As it can be seen from the first, we distinguished socio-economic (SE) variables, mostly relating to the individual's household; level of service variables (LOS), relating to the trip characteristics by each mode available, and mode-specific contants which would tend to capture the effects of unobserved variables, measurement errors, etc. One of the main problems we faced was which SE variables to enter in each mode utilities. We proceeded for this in a step-wise fashion eliminating non-significant 1/ or incorrectly specified 2/ variables from a general starting specification with all SE and LOS variables according to a-priori beliefs. Table 5 presents the best MNL model found and the best 'null' model, the constants only model. As it can be seen, several of the variables originally considered were eliminated.

1/ With a 't-ratio' smaller than the critical 95% confidence level value of 1.96.

2/ For example with a wrong sign.

Characteristic	Garforth-Leeds f/	Las Condes-CBD
i) Of the corridor		
Main city population a/	500,000	4,000,000
Travel pattern	Suburban to urban	Intraurban
Pure modes available	Train, bus, car	Metro, bus, shared taxi, car
Lenght of railway line b/ (Km)	11.5	8
No of railway stations b/	3	11
Train frequency (trains/hour)	5	30
Train fare c/ (Chilean $)	25.5	9
Bus fare d/ (Chilean $)	16.5	10
Av. Car operating cost (Chilean $/Km)	2.9	2.2
ii) Of the sample		
Sample size	724	637
Year of survey	1975	1981
Is there income information?	No	Yes
Is there information about alternatives to the chosen mode?	No	Yes
% of trips going to the CBD	67.7	100
% of households with one or more cars	79.3	88.3
Av. parking charge e/ (Chilean $/hour)	11.5	8

Notes: a/ Leeds and Santiago, respectively; b/ Underground in the Chilean case; c/ Fixed in the Chilean case; d/ Fixed in both cases; e/ Calculated as a weighted (by the corresponding proportions) average of the different amounts payed by the individuals in the sample; f/ In all cases the British cost in £ of 1975 were converted first to 1981 equivalent using a factor of 2.181 (given by the British Embassy in Santiago) and then converted to Chilean $ using the 1981 rate of $ 75 per £.

TABLE: 1: COMPARISON OF THE TWO CORRIDORS AND DATA SETS

We left 124 observations to be used as an 'independent' validation sample, therefore estimating the models with only the 600 remaining data points. The method of chosing the validation sample, and a comparison of its characteristics as opposed to those in the estimation sample can be found in Ortúzar (1980). A description of all the statistics and test employed which are shown in Table 4 can be found in bits and pieces in Mc Fadden (1976), Tardiff (1976) and Gunn and Bates (1982).

4. RESULTS USING CHILEAN DATA

In this case we tried firstly the same variables presented in Table 3 for the specification of the models and we used the same statistical measures shown in Table 4 for the comparisons of model performance.

The first series of calibration runs finished with the same MNL specification obtained for the case of the Leeds corridor (and shown in Table 5) but using the Chilean data. It is important however, to remember that the English data set contained information about only six alternatives while the Chilean one considered a total of nine options (see Table 5 and 6). In addition the

	ENGLISH CASE		CHILEAN CASE	
MODE	FREQUENCE OF CHOICE	AVAILABLE TO	FREQUENCY OF CHOICE	AVAILABLE TO
Car driver	185 (30.8)	313 (52.2)	112 (21.7)	412 (79.7)
Car passenger	100 (16.7)	600 (100.0)	33 (6.4)	464 (89.7)
Shared taxi	-	-	29 (5.6)	461 (89.2)
Train/Metro	21 (3.5)	120 (20.0)	109 (21.1)	158 (30.6)
Bus	263 (43.8)	600 (100.0)	52 (10.0)	379 (73.3)
P & R	10 (1.7)	240 (40.0)	94 (18.2)	313 (60.5)
K & R	21 (3.5)	500 (83.3)	22 (4.3)	330 (63.8)
Shared Taxi-Metro	-	-	27 (5.2)	244 (47.2)
Bus-Metro	-	-	39 (7.5)	173 (33.5)
TOTAL	600 (100.0)		517 (100.0)	

Note: The numbers in parentheses are percentages over the total.

TABLE 2: DISTRIBUTION OF CHOICES IN BOTH SAMPLES

a) Socio-Economic Variables

1. RESID Number of residents in household
2. WORKERS Number of workers in household
3. CAROWN Number of cars in the household
4. CO_1 Car ownership dummy, takes a value of 1 for one car owning households and 0 otherwise
5. CO_2 Car ownership dummy, takes a value of 1 for multiple car owning households and 0 otherwise
6. CBD Destination dummy, takes a value of 1 if trip goes to CBD and 0 otherwise
7. COMP Number of cars/Number of workers with a ceiling of one

b) Level-of-Service-Variables

1. COST Out of pocket and/or operating costs (1981 Chilean $)
2. IVT In vehicle time (min)
3. WAIT Waiting time (min)
4. WALK Walking time, including transfers (min)

c) Mode Specific Constants

One for each mode excluding CAR PASSENGER (which was arbitrarily set to zero)

TABLE 3: EXPLANATORY VARIABLES USED IN THE SPECIFICATION SEARCHES

a) For Estimation Data Set.

L(O) = Log-likelihood at zero, ie. for the equally likely model
L(C) = Log-likelihood of a model with only mode-specific constants
L(Θ) = Log-likelihood at convergence
LR(O) = Likelihood ratio statistic with respect to the equally likely model = $-2\{L(O) - L(\Theta)\}$
LR(C) = Likelihood ratio statistic with respect to the constants only model = $-2\{L(C) - L(\Theta)\}$
ρ^2 = Rho-squared index = $1-L(\Theta)/L(O)$
$\bar{\rho}^2$ = Corrected rho-squared index = $1-L(\Theta)/L(C)$
% Right = Percentage correctly predicted (maximum utility classification)

b) For Validation Data Set.

FPR(%) = First preference recoveries expressed as a percentage a/
LRT b/ = Likelihood ratio test, performed directly on the validation data

---

a/ It is actually equivalent to % Right.
b/ When comparing models estimated with the same data base (see Gunn and Bates, 1982).

TABLE 4: STATISTICAL MEASURES AND TESTS USED IN MODEL COMPARISONS

variable CBD, included in the Leeds case, was irrelevant for Santiago as all the trips considered had their destination in the CBD area. The results obtained are shown in the second colum of Table 6, under the heading MNL-1. It is quite obvious that the Leeds specification does not behave well at all when calibrated on the Santiago data. Most coefficients shown a wrong sign and/or are not significant at the 95% level. The only coefficient which is significant and of the proper sign is that of the variable WAIT, however its value is much higher than in the English case (see Tables 5 and 6).

Model MNL-2 is the result of literally dozens of alternative specification trials (subject to using only information as was available in the English case, ie. not using the full richness of the Chilean data). It was found that to obtain a reasonable coefficient for the variable number of cars (CAROWN), this should not enter the utilities of modes 2 and 7 (car passenger and K & R) as was the case in England. Also contrary to what was found in Leeds, we did not find a non-linear influence of the number of cars in Santiago, ie. CO_1 and CO_2 came with almost the same value, and the specification including them was clearly worse than model MNL-2 in general statistical terms. The coefficients of the variables COST and WALK came always with incorrect (positive) sign and were not significant at the 95% level; the variable IVT also came always with incorrect sign, however when separating it into three components, time in Metro (MTIM), time in other public transport modes (PTTIM) and time in car, it was found that the first two showed correct signs and were significant but the latter was also significant and of incorrect sign so it was dropped from the model. Finally the variable WAIT remained significant and with a correct sign throughout but also always with a higher coefficient than the one obtained in England.

VARIABLES (alternatives entered a/)	Constants only	MNL
CBD (4-7)	-	0.915 (3.01)
CO_1 (1, 2, 6, 7)	-	2.469 (6.12)
CO_2 (1, 2, 6, 7)	-	2.968 (5.86)
COST (1, 2, 4-7)	-	-0.052 (-3.60)
IVT (1, 2, 4-7)	-	-0.072 (-2.37)
WAIT (4-7)	-	-0.243 (-5.02)
WALK (1, 2, 4-7)	-	-0.149 (-5.98)
CAR (1)	2.047 (13.42)	2.615 (9.01)
TRAIN (4)	0.492 (1.85) b/	6.686 (7.32)
BUS (5)	0.967 (8.23)	5.732 (8.97)
P & R (6)	-0.527 (-1.53)	1.469 (2.15)
K & R (7)	-1.191 (-4.92)	0.955 (1.68) b/
LR(0)	408.27	571.30
(degrees of freedom)	(5)	(12)
% Right	64.8	66.2
ρ^2	0.269	0.374
$\bar{\rho}^2$	Not applicable	0.146
Estimation sample size	600	600
FPR(%)	-	64.5
LRT to model MNL c/	0.0165	1
Validation sample size	124	124

Notes: a/ Alternatives are: 1:Car; 2:Car passenger;4:Train;5:Bus;6: P&R; 7: K&R b/ Not significant at 95% level c/ This means that the observed data in the validation sample is something like 61 (ie. $(0.0165)^{-1}$ times as likely under the best model than under the constants only model. We would judge the former to be better although both models are consistent with the validation sample (see Gunn and Bates, 1982).

TABLE 5: BEST MNL MODEL FOR THE GARFORTH CORRIDOR

VARIABLES (alternatives entered)	COEFFICIENTS (t-ratios) Constants-only	MNL-1	MNL-2
CAROWN (1,6)	-		0.529 (3.71)
CO1 (1,2,6,7)	-	-0.904 (-2.95) c/	-
CO2 (1,2,6,7)	-	-0.001 (-0.41) c/ b/	-
COST (1-9)	-	0.023 (0.90) c/ b/	-
IVT (1-9)	-	0.004 (1.98) c/	-
MTIM (4,6-9)	-	-	-0.085 (-1.96)
PTTIM (3,5,8,9)	-	-	-0.065 (-2.94)
WAIT (3-9)	-	-0.372 (-2.89)	-0.310 (-2.38)
WALK (1-9)	-	0.023 (0.95) c/ b/	-
CAR(1)	1.557 (7.56)	1.400 (6.02)	0.706 (2.25)
SH. TAXI(3)	-0.162 (-0.63) b/	-0.453 (-1.18) b/	1.125 (2.68)
METRO (4)	2.917 (11.79)	2.980 (6.91)	4.286 (8.90)
BUS (5)	0.573 (2.52)	0.650 (1.24) b/	2.917 (4.74)
P & R (6)	1.628 (7.57)	2.082 (6.87)	1.960 (3.64)
K & R (7)	-0.059 (-0.21) b/	0.465 (1.35) b/	1.123 (2.21)
SH.TAXI-M(8)	0.349 (1.30) b/	0.840 (1.54) b/	2.373 (3.72)
BUS-METRO (9)	1.169 (4.61)	2.065 (3.17)	3.679 (5.08)
LR(O) (degrees of freedom)	341.23 (8)	368.55 (14)	376.05 (12)
% Right	43.9	-	47.6
ρ^2	0.199	-	0.219
$\bar{\rho}^2$	Not applicable	-	0.025
Estimation sample size	517	517	517
FPR (%)	41.4	-	43.4
LRT d/	0.0008	-	1
Validation sample size	99	-	99

Notes: a/ Alternatives are: 1: Car; 2: Car passenger; 3: Shared taxi; 4: Metro; 5: Bus; 6: P&R; 7: K&R; 8: Shared taxi-Metro; 9: Bus-Metro. b/ Not significant at 95% level; c/ Wrong sign d/ This means that the observed data in the validation sample is something like 1250 (ie. $0.0008)^{-1})$ times as likely under model MNL-2 than under the constants only model. We would judge the former to be better although both models are consistent with the validation sample.

TABLE 6: MNL MODELS CALIBRATED WITH THE SANTIAGO DATA

5. COMPARISON OF RESULTS AND CONCLUSIONS.

First of all it is very important to remark how difficult it is to examine the issue of model transferability. If one uses the notion that transferable models are those that when estimated with different data sets (ie. different areas, points in time or indeed cultural environments) are stable or invariant with respect to their estimated parameters, then it is rather obvious that no model can be wholly transferable. However if one uses the softer definition proposed by Lerman (1981) of 'close enough' models, in the sense of being substitutable for a given specific purpose, one would go a long way to resolve some reported contradictory findings (see for example Talvitie and Kirshner, 1978 ; and Ben-Akiva, 1981). Unfortunately our findings do not support the transferability of the MNL even in this latter sense.

In Table 2 a comparison is made of the frequency of choice of each alternative, and also of the number of people that had each alternative available, in both samples. As it can be seen several alternatives vary significantly their importance in both cases. For example, BUS changes from being the most popular alternative in the English case, with 43.8% of the people choosing it, to being chosen by only 10% of the Chilean sample. The reverse occurs with TRAIN wich jumps from 3.5% in the English case to METRO being the second most popular option (21.1%) in the Chilean case. CAR PASSENGER also decreases sigficantly (16.7% in England to only 6.4% in Chile), and P & R shows a spectacular increase (from 1.7% in England to 18.2% in Chile). These changes should help to explain some of the findings we present below but clearly cannot be held resposible for the most important differences observed.

As it is well known the mode specific constants of a MNL model represent those characteristics of the alternative not considered by the socio economic and level of service variables included in the specification. Thus, one can say that they are affected by the modal perceptions of the users with respect to comfort, privacy, environment, safety, reliability. etc. They are also affected by the way in which the availability of each mode is defined for the individuals included in the data set. Therefore, the differences observed between the values obtained for the mode specific constants in the Leeds and Santiago cases should be mainly due to diferences in those factors. However as we already noted, choice set availability was defined in both cases in the same 'objective' manner (the only difference was that in Santiago we added the empirically found constraint that no individual with a high income had BUS available). It is interesting to note that in both cases the prefered mode (ceteris paribus) is TRAIN/METRO and the less prefered one, CAR PASSENGER. In this sense also BUS comes second in England and third in Chile (after BUS-METRO). The rest of the ordering however is not so similar. We cannot be certain whether this is just a coincidence or if there is some similarity in the unexplained factors that determine the size of the mode-specific constants in both cases. Sadly though, this is one of the strongests signs of transferability -of any sort - that we were able to detect!

The importance and behaviour of the car ownership variable was very different in both cases; further, it was found that the Chilean specification needed a slight change to produce reasonable results (ie. to drop CAROWN from the utilities of modes 2 and 7). Besides, the coefficients of the variables COST and WALK were not significant and of incorrect sign in Chile 1/.

The most encouraging findings were related with the variables WAIT, which although of much higher importance in Chile than in England was at least always correctly specified, and the disaggregation of IVT into MTIM and PTTIM in Chile.

1/ It is interesting to mention that COST has been found not significant in most previous studies in Chile (see for example, Fernández et al 1981).

Although in England it was found that such disaggregation was not warranted (Ortúzar, 1980), the coefficients of MTIM and PTTIM are rather close to that of IVT in Leeds (see Table 5 and 6). However, it is difficult to explain why the coefficient of the variable "time in car" came positive and very significant.

In general then, we can conclude that the results obtained do not grant the transferability of MNL models betwen different cultural settings. There are several reasons why this was something to be expected. The main ones are related with the kind of model used and the variables included in its specification. Firstly, it is well known that the use of the MNL is bounded to the limitations imposed by the independence of irrelevant alternatives property of the model. Both in the Leeds and Santiago cases the set of alternatives used were far away from being totally independent, therefore violating a basic assumption of the model[1/]; this should have important consecuences for the correct calibration in each case and therefore could reduce transferability. Secondly the specification used in the Leeds case, on which the comparison is based, was (as usual in practice) limited by the information available and does not necessarily correspond to the ideal correct specification of the true model behind the situation studied. In particular this is true for the socio-economic variables used and the specification of the COST variable. It is difficult to compare cost figures without reference to income.

The difficulties mentioned above obviously affect transferability and are difficult to overcome in practice. The next stage in this research is to compare more general structures, as the nested logit model (Williams, 1977), which was already found to be preferable to the simple MNL model in the English case (Ortúzar, 1983) and which appears likely to be also preferable in Chile.

1/ We suspect the violation is much more severe in Chile from preliminary findings obtained when calibrating nested logit models.

6. REFERENCES

BEN-AKIVA, M.E. (1981) Issues in transfering and updating travel-behaviour models. In P.R. Stopher, A.H. Meyburg and W. Brog (eds.), New Horizons in Travel-Behaviour Research, D.C. Heath & Co., Lexington.

FERNANDEZ, J.E., COEYMANS, J.E., DIAZ, J. and ORTUZAR, J. de D. (1981) Evaluación de alternativas de ampliación de la Red de Transporte Colectivo Independiente en el Gran Santiago. Final Report to the Ministry of Transport, Department of Transport Engineering, Pontificia Universidad Católica de Chile, Santiago (in Spanish).

GUNN, H.F. and BATES, J.J. (1982) Some statistical considerations in fitting travel demand models. Transportation Research 16A (5/6), 371-382.

LERMAN, S.R. (1981) A comment on interspatial, intraspatial and temporal transferability. In P.R. Stopher, A.H. Meyburg and W. Brog (eds.), op cit.

MC.FADDEN, D. (1974) Conditional logit analysis of qualitative choice behaviour. In P. Zarembka (ed.), Frontiers in Econometrics, Academic Press, New York.

MC.FADDEN, D. (1976) The theory and practice of disaggregate demand forecasting for various modes or urban transportation. Working Paper 7623, Institute of Transportation Studies, University of California, Berkeley.

ORTUZAR, J. de D. (1980) Modal choice modelling with correlated alternatives: Applications to mixed-mode travel demand forecasting. Ph.D. Thesis, Departament of Civil Engineering, University of Leeds.

ORTUZAR, J. de D. (1983) Nested logit models for mixed-mode travel in urban corridors. Transportation Research 17A(4), in press.

ORTUZAR, J. de D. and DONOSO, P.C.F. (1983) Survey design, implementation, data coding and evaluation for the estimation of disaggregate choice models in Santiago, Chile. 2nd International Conference on New Survey Methods in Transport, 12-16 September 1983, Hungerford Hill, Australia.

ORTUZAR, J. de D., DONOSO, P.C.F. and HUTT, G.A. (1983) The effects of measurement techniques, variable definition and model specification on disaggregate demand model functions. PTRC Summer Annual Meeting, University of Sussex, England.

TALVITIE, A.P. and KIRSHNER, D. (1978) Specification, transferability and the effects of data outliers in modelling the choice of mode in urban travel. Transportation 7(3), 311-332

TARDIFF, T.J. (1976) A note on goodness-of-fit statistics for probit and logit models. Transportation 5(4), 377-388.

WILLIAMS, H.C.W.L. (1977) On the formation of travel demand models and economic evaluation measures of user benefit. Environment and Planning 9A, 285-344.

ACKNOWLEDGEMENTS

The first part of this research was done while one of the authors was in England. Hugh Gunn, Huw Williams and Dirck Van Vliet were instrumental in helping the task being satisfactorily completed. The U.K. Science Research Council also deserves thanks for having provided the grant that financed the work.

The second part was conducted in Chile and we wish to credit the important financial support given by DIUC, the Universidad Católica's Directorate of Research. We would also like to thank the help of Gloria Hutt and Patricio Donoso in the data collection effort and computing aspects of the work.

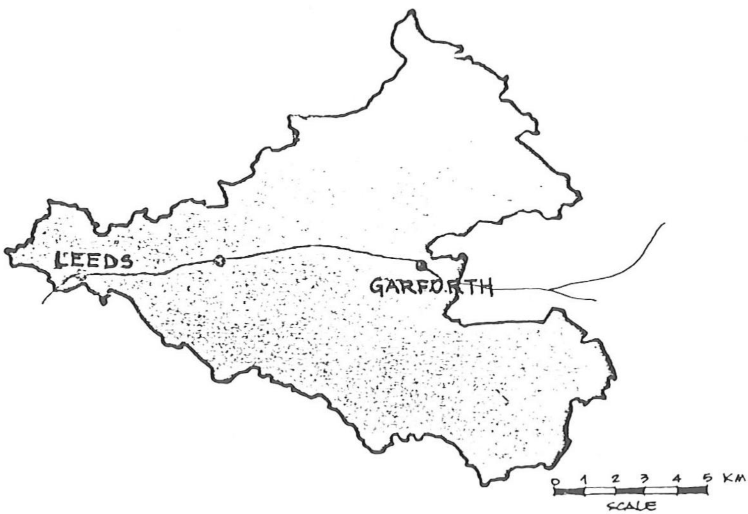

figure 1: GARFORTH - LEEDS CORRIDOR

figure 2: LAS CONDES - CBD CORRIDOR