INTRA-METROPOLITAN TRANSFERABILITY
OF MODE CHOICE MODELS

By Richard A. Galbraith and David A. Hensher*

Individual (or disaggregate) choice models have been credited with great accuracy in estimating and predicting travel choice.¹ Researchers have argued that the parameters of such models should remain stable in predicting travel behaviour not only for one area in different time periods but also for different groups of people in different areas. Hence Richards and Ben Akiva (1975, p. 14, p. 143) say:

By considering individual behaviour and seeking out . . . causal relationships, we have a much greater chance of developing a general model which can be transferred from one area to another, or from one population to another. . . . If a disaggregate model is truly a behavioural model, and if it has been estimated with data which has a high degree of variability, then it can be expected that the model can be used in different geographic locations and for populations with different economic structures without amendment to the coefficients.

If this is correct, the implication is plain—the parameters of a "well specified" model (which properly "explains" travel behaviour in one area) should be appropriate in predicting travel behaviour for other areas. The motivation for model coefficient transfers is, Atherton and Ben-Akiva (1976, p. 12) write, that:

If a model estimated for one area can be transferred to another, the costs of conducting transportation studies can be greatly reduced.

"Transferability" may be considered at various levels of generality, from that of

---

* Richard A. Galbraith is with R. Travers Morgan Pty Ltd., Sydney, N.S.W. The research for this paper was undertaken while he was in the School of Geography, University of New South Wales. David A. Hensher is at the School of Economic and Financial Studies, Macquarie University, Sydney; he was on leave at the London Business School when the paper was completed. The comments of Michael Beesley, James Crittle, David Wilson and a referee have assisted materially in the final presentation of the paper.

¹ The distinction between choice and demand must be made clear, in view of the continual misuse of the latter construct by disaggregate choice modellers. The reference to modal demand in disaggregate mode-choice models is, strictly, incorrect unless we can assume that the model is allocating the predicted demand (after a change) between modes. In most applications it is correct to talk of modal choice with choice (share) elasticities, but not of modal demand with demand elasticities. To accommodate the latter requires a modelling system that links the mode choice model with at least a trip frequency choice model and a model of expenditure on travel versus other activities (unless the expenditure elasticity with respect to the attributes of travel is unity). Hence if the term "demand" is to be used it should be qualified as conditional demand.
broad behavioural postulates through to that of model parameters and data specifications. The transferability of model parameters is the focus of this paper.

In spite of the early optimism, experience to date suggests that some adjustments must be made to model coefficients before they are transferred from one geographical area to another (see Appendix A for an overview of previous studies). Because no model is ever sufficiently specified, this is to be expected. With one possible exception (Atherton and Ben-Akiva, 1976), different models estimated on different data have generally produced inconsistent parameter estimates. How far these inconsistencies arise from behavioural differences, or from differences in other factors, is not known. However, reported tests of transferability have clearly not controlled sufficiently for “poor data” factors, which must inhibit transferability.

For example, Daly (1974) and Watson and Westin (1975) presented models using data on level of service only. Other variables which may be important (specifically socio-economic measures) were ignored. At least two of the studies reviewed used data bases containing aggregate network data. It has been argued (e.g. Horowitz, 1981) that aggregate data can produce biased estimates of model parameters, so the value of these studies must be questionable. Finally, the Charles River Associates study (1978, p. 132) concluded that the underprediction was: “clearly not because of failure for the model to transfer, . . . but rather, due to attempts to perform the application with faulty data”.

All this suggests that a precondition for model transferability is that the data used be verified and refined, to reduce the sources of error which influence the estimated values of model parameters. If this is not done, model parameters become statistical artefacts of the estimation data sets and, except by accident, are likely to predict poorly for other populations and time periods.

Further, dissimilarities in model parameters are in part due to different methods of collecting data. In none of the past studies reviewed have the data sets been designed and collected in a consistent manner. The resulting model coefficients are likely to have been influenced by different measurement procedures, different sampling procedures, variable definitions, and even differences in questionnaire wording and coding conventions. Therefore unless there can be more consistency across different data sets (Louviere, 1981), statistical discussions of transfer efficiency have little practical meaning.

This paper tests the hypothesis of intra-urban transferability of coefficients in two municipalities in Sydney, controlling for differences in data collection procedures, measurement of variables and sampling design, starting from a characteristic disaggregate model. The paper is organized as follows. First the data sets are described. Secondly, the set of transfer models are estimated and tested for parameter transferability. Thirdly, the models are tested for predictive performance, and contrasted with a “naive” model. Fourthly, updating procedures are used to adjust the model parameters in the attempt to improve transferability. The concluding section comments further on the problems of transferability.

---

2 Hansen (1981) has proposed a four-level hierarchy for transferability; level 1 (the most general), the transferability of broad behavioural postulates such as utility maximization; level 2, the transferability of a mathematical model class; level 3, the transferability of specific model form; and level 4 (the least general), the transferability of model coefficients. This paper emphasizes level 4.
THE DATA

Two spatially distinct data sets were required. To replicate a situation likely to be faced in reality, in which the transfer model would have been estimated on data collected in another area and at an earlier date, it was preferable that the data sets be collected at different time periods. The data sets selected were commuter samples collected from two suburban areas in the Sydney region: one from the Hornsby region to the Northwest of Sydney for 1971, the other from an area to the Southwest of the city comprising Fairfield, Green Valley and Macquarie Fields, for 1975 (Appendix B).

Both surveys were by home interview; the use of “reported–perceived” values is consistent over the two samples; and questions relating to similar variables were to a great extent worded consistently across the data sets, yielding consistent variable definitions. This ensures a consistency across the two data sets that other empirical studies have not achieved.

On the other hand, a potentially limiting feature of the data is that the choice of variables is restricted by data collected in both surveys. Although the two surveys sought information covering a wide range of areas, the only information of relevance to this study was that common to both data sets.

The samples

In an area where the only feasible public transport is the railway, the Northwestern (N.W.) area data set was based only on individuals who choose between car and rail as their main mode of travel for the work trip. The sample drawn from the Southwestern (S.W.) area was selected to represent this same situation. For each individual, car and rail are the chosen or alternative modes. Other journeys (for example, walk or bus as the usual or alternative mode) were excluded. This ensured that the sampling procedures were consistent. The samples drawn were random, subject to this main restriction. It suggested a number of other restrictions, imposed on both data sets: each individual had to hold a current driver’s licence, and each household had to own at least one car. People who required a car for use during work hours, or who had a company car, were excluded from the samples on the ground that no choice was possible to them.

The final samples were:

- **Northwestern (N.W.) sample**: 332 cases, of which 182 (55%) were regular car users and 150 were regular train users.
- **Southwestern (S.W.) sample**: 243 cases, of which 210 (86%) used the car as their usual mode, and 33 chose rail.

There are significant differences between the two areas in length of development, socio-economic structure, and, in particular, train service. Table 1 compares the study samples, and highlights important differences in socio-economic indicators and work-trip characteristics.

The statistics on occupational status and individual income suggest a great difference between the two populations in socio-economic status. In the N.W. sample, over 80% of the workers belong to the first three (“white collar”) occupation classes,
### Table 1

**Characteristics of the Study Samples**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Northwestern Area</th>
<th>Southwestern Area</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(%)</td>
<td>(%)</td>
</tr>
<tr>
<td><strong>Occupational status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0. Professional, semi-professional</td>
<td>47.6</td>
<td>4.5</td>
</tr>
<tr>
<td>1. Army officers, proprietors, company executives</td>
<td>9.0</td>
<td>6.6</td>
</tr>
<tr>
<td>2. Store and office clerks</td>
<td>24.1</td>
<td>15.6</td>
</tr>
<tr>
<td>3. Craftsmen, foremen, skilled labourers, etc.</td>
<td>11.7</td>
<td>25.5</td>
</tr>
<tr>
<td>4. Operators, semi-skilled workers</td>
<td>3.3</td>
<td>27.2</td>
</tr>
<tr>
<td>5. Labourers and unskilled workers</td>
<td>4.2</td>
<td>20.6</td>
</tr>
<tr>
<td><strong>Annual personal income (1976)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zero or not given</td>
<td>22.3</td>
<td>25.7</td>
</tr>
<tr>
<td>Less than $4,000</td>
<td>26.0</td>
<td>22.9</td>
</tr>
<tr>
<td>$4,000--$12,000</td>
<td>41.2</td>
<td>49.2</td>
</tr>
<tr>
<td>Over $12,000</td>
<td>10.5</td>
<td>2.2</td>
</tr>
<tr>
<td><strong>Mode chosen</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car</td>
<td>54.8</td>
<td>86.4</td>
</tr>
<tr>
<td>Rail</td>
<td>45.2</td>
<td>13.6</td>
</tr>
<tr>
<td><strong>Trip destination</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBD</td>
<td>40.1</td>
<td>12.3</td>
</tr>
<tr>
<td>Other</td>
<td>59.9</td>
<td>87.7</td>
</tr>
<tr>
<td><strong>Mean number of journey stages by rail</strong></td>
<td>4.6</td>
<td>7.2</td>
</tr>
<tr>
<td><strong>Average travel times</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(mins)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car</td>
<td>45.6</td>
<td>42.5</td>
</tr>
<tr>
<td>Rail</td>
<td>66.6</td>
<td>90.2</td>
</tr>
<tr>
<td><strong>Journey distance</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0--9 miles</td>
<td>15.4</td>
<td>29.6</td>
</tr>
<tr>
<td>10--19 miles</td>
<td>75.6</td>
<td>37.9</td>
</tr>
<tr>
<td>20--29 miles</td>
<td>8.4</td>
<td>23.5</td>
</tr>
<tr>
<td>30--39 miles</td>
<td>0.6</td>
<td>8.6</td>
</tr>
<tr>
<td>40--49 miles</td>
<td>—</td>
<td>0.4</td>
</tr>
</tbody>
</table>

* Census data for 1976 were used to compare the income levels of workers in the two areas, as the figures available in the disaggregate samples were collected over a five-year gap, and thus were not strictly comparable. Income (and not the other variables) was standardized because of the great changes in income levels over the five-year period, and because information on income was readily available.

Compared with only 27% in the S.W. The income distributions appear similar at the lower end of the scale. However, most of the S.W. workers earn between $4,000 and $9,000, and only 2.2% were earning over $12,000 in 1976. Over 10% of the N.W. workers were in this high income bracket.

As the different occupational status of the samples indicates, over 40% of the N.W. workers work in the central business district (C.B.D.) of the city; while almost 90% of
the S.W. sample work locally or in other suburbs. This difference is reflected in other statistics in Table 1. On average, S.W. workers require more journey stages for their trip to work by rail. Most workers in the N.W. sample have easy access to a railway station and receive a very direct service to the C.B.D. The Sydney suburban rail network is not oriented towards the cross-city trips that the S.W. individuals demand, and their trips involve many stages. These constraints are seen in the average travel time. Work trips are almost 50% slower for the S.W. sample when rail is used, yet the average time taken to work by car is almost identical between the two samples.

ESTIMATING AND TESTING THE TRANSFER MODELS

Alternative formulations of the N.W. model were tested to identify suitable models. Models were estimated using the multinomial logit form:
\[
P_{\text{car}} = \frac{\exp(V_{\text{car}})}{\exp(V_{\text{car}}) + \exp(V_{\text{rail}})}
\]
where
\[
P_{\text{car}} \quad \text{is the probability of choosing the car mode for the journey to work,}
V_{\text{car}} \quad \text{is the (indirect) representative utility associated with the car journey, and}
V_{\text{rail}} \quad \text{is the (indirect) representative utility associated with the rail journey.}
\]

The (indirect) representative utility functions \(V_{\text{car}}\) and \(V_{\text{rail}}\) are linear in their parameters and additive in the explanatory variables. The overall measure of goodness-of-fit is rho-squared (\(\rho^2\)) with \(t\)-tests determining the significance of individual variables. The coefficients were judged by their statistical significance, and by whether their signs correspond to a priori expectations. The procedure adopted was to decide first on the form of the time and cost variables, and then to build up the models by the addition of other variables one at a time. A description of the final model variables is given in Table 2. Other variables evaluated are given in Appendix C. The final N.W. models, three in total, are given in Table 3.

Comparison of the sets of coefficients

A test of whether a model is transferable is to compare the parameter estimates obtained for the different areas. In this initial test, the specifications of three transfer models developed on the N.W. data were re-estimated on the S.W. sample.

Three pairs of models, called Models A, B and C, are presented in Table 3. For each model, the N.W. estimated coefficients are in column one (together with their \(t\)-statistics), and for the S.W. sample in column two. The \(\rho^2\) values cannot be compared across different data sets,\(^3\) but can be used to compare different models estimated on the same data set. Thus, within the S.W. data set, Models A, B and C perform equally well (\(\rho^2 = 0.486, 0.492, 0.476\) respectively) when the estimated coefficients are unrestricted.

\(^3\) It is likely, however, that the \(\rho^2\) values of the S.W. samples, while greater than their N.W. counterparts, represent a poorer goodness-of-fit; as the skewed modal split of the sample (86% car) inflates the minimum \(\rho^2\) value for the S.W. data set. See Tardiff (1976).
The "transferability test" statistically tests the differences between the two sets of coefficients, where the null hypothesis is that the two sets of coefficients are equal. The test is given by:

\[
\chi^2 = -2[L^*(T_{sw}) - L^*(S_{sw})]
\]  

where

- \(L^*(SW_{sw})\) is the log-likelihood of the S.W. coefficients on the S.W. sample, and
- \(L^*(T_{sw})\) is the log-likelihood of the S.W. sample when the coefficients are "restricted" to the values of the N.W. transfer model.

The calculated \(\chi^2\) value is compared against the critical \(\chi^2\) value (at the 5% level). In all three cases, the calculated \(\chi^2\) value exceeds the critical value; thus the null hypothesis that the two sets of coefficients are equal is rejected in each case.

The statistic \(\rho_{trans}^2\) indicates the goodness-of-fit of the models on the S.W. sample when the coefficients are restricted to the original N.W. estimates, and is a measure of the fit of transfer models to the S.W. data. It is defined by:

\[
\rho_{trans}^2 = 1 - \frac{|L^*(T_{sw})|/|L^*(0)|}{1}
\]  

However, the statistic rho-squared is an abstract measure; it is difficult to gauge what differences in \(\rho_{trans}^2\) actually mean in terms of model performance. Given this, it is only possible to rank the three transfer models by their goodness-of-fit to the S.W. sample.
On this criterion, Model A is the least satisfactory ($p_{trans} = 0.176$), while Model C (0.356) is marginally superior to Model B (0.334).

The differences in transfer performance can be explained by differences in the specifications of the three transfer models. Specification A, while fitting the N.W. area sample best, performed poorly on the second data set when its coefficients were restricted to the N.W. levels. Models A and B differed in the definition of the cost variables, yet Model B's coefficients were, as a set, statistically more stable across the two samples. These results indicate that the theoretically more appealing specification of travel costs in Model B, where cost is weighted by the inverse of the worker's income, is preferable to "unweighted costs" (see Train and McFadden, 1978).

The best of the three transfer models in goodness-of-fit to the S.W. area is Model C. Model C was similar to the specification of Model B, except that in-vehicle time was represented more simply as a single generic variable. This result confirms suspicions concerning the omission of a relevant variable (vis-à-vis multicollinearity in the base data). In Models A and B, in-vehicle travel time for the main mode by rail has been omitted; however, if (as expected a priori) this variable is a significant determinant of mode choice behaviour, its omission would cause bias in the parameters of the other variables. The estimated relationships in the first two models would thus only reflect statistical correlations which exist in the N.W. data. While the parameters would provide a good fit to the N.W. sample, we have little confidence in transferring them. The transfer coefficients of Model C are less subject to this error, as the total in-vehicle time variable includes all components of in-vehicle time. The weighting applied to the in-vehicle time variable in Model C is more likely to reflect the motivations of commuters in general than the correlations of the base data set.

**Comparison of pairs of coefficients**

The difference between individual coefficients can be evaluated by the $t$-statistic for the absolute difference between the transfer and unrestricted coefficients (column three of Table 3).

A $t$-statistic of less than 1.96 indicates that the null hypothesis that a pair of coefficients is equal cannot be rejected at the 5% level. Caution is required in applying this test. The $t$-statistic depends on the coefficients' standard errors as well as on their size difference. For example, in Table 3 the difference between the transfer and the unrestricted coefficient for parking cost ($PARK-C$) is substantial, yet the two coefficients are not statistically different ($t = -0.80$) because the standard error for parking cost is very large. To allow for this weakness, the test was applied only to parameters of small standard error.

The only variable to be invariant across the two samples was household car competition ($CCE/IN$). The estimated parameters were consistently similar (within the range $-1.32$ to $-1.73$) across the different model specifications and areas.

Individually, neither the "straight" cost nor the "income weighted" cost parameters were consistently similar across the two samples. Of the cost variables, the total in-vehicle car cost parameter ($TIVC-C$) was very similar, and so was rail cost divided by income ($RTOC/I$). None of the parking cost variables are similar, however. Of the time variables, only the total out-of-vehicle time ($G-TOVT$) variable parameters were
### Table 3

**The Transferability of Models A, B, C**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model A</th>
<th>Model B</th>
<th>Model C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1) N.W. Area</td>
<td>(2) S.W. Area</td>
<td>(3) Difference</td>
</tr>
<tr>
<td>ASC_Car</td>
<td>0.243</td>
<td>0.41</td>
<td>-2.382</td>
</tr>
<tr>
<td>IVTM_C</td>
<td>-0.055</td>
<td>-4.43</td>
<td>-0.024</td>
</tr>
<tr>
<td>G_IVTC</td>
<td>-0.040</td>
<td>-2.39</td>
<td>0.008</td>
</tr>
<tr>
<td>G_TIVT</td>
<td>-0.038</td>
<td>-2.06</td>
<td>-0.139</td>
</tr>
<tr>
<td>G_TOVT</td>
<td>-0.038</td>
<td>-4.01</td>
<td>-0.033</td>
</tr>
<tr>
<td>TIVC_C</td>
<td>-0.017</td>
<td>-1.70</td>
<td>0.020</td>
</tr>
<tr>
<td>TOTC_R</td>
<td>-0.009</td>
<td>-5.61</td>
<td>-0.048</td>
</tr>
<tr>
<td>CCE/IN</td>
<td>-1.321</td>
<td>-2.31</td>
<td>-1.726</td>
</tr>
</tbody>
</table>

| No. of Cases  | 332 | 243 |
| L*(0)         | -228.58 | -96.54 |
| L*(β)         | -131.30 | -49.66 |
| L*(T_0)       | -78.73 | -63.65 |
| L*(SW_con)    | -49.66 | -49.06 |
| Calculated $\chi^2$ | 58.2 | 29.2 |
| Critical $\chi^2$ | 14.1 | 14.1 |
| $r^2$         | 0.423 | 0.486 |
| $r^2_{max}$   | 0.176 | 0.334 |

$p^2_{max}$ is a goodness-of-fit measure of the transferred coefficients on the Southwestern data.
strictly comparable. These were much larger in the second area for all three models. The in-vehicle time variables, though not strictly comparable because of the high standard errors of the S.W. parameters, were also dissimilar.

Explaining the model differences

The coefficients of the three models when estimated unconstrained on the two areas are, by and large, inconsistent. In general, while the socio-economic variable CCE/IN was transferable, the transport level of service variables was not.

The income-adjusted car competition effect is an important variable in transferability of a mode-choice model, probably representing a socio-economic attitudinal dimension; that is, the higher the income level, the more care there is for the opportunity cost for the car.

Since the public transport service varies widely between the two areas (Table 1), the quality dimensions are likely to have a significant impact on coefficient estimates. We can in part account for quality differences by ascribing them to implicit psychometric weights built into the disaggregated observed variables. But unobserved quality dimensions still exist which cannot be assumed to have the same joint distribution in both areas (even though the models assume so). That is, there will be context-specific effects. To illustrate, the high coefficients for out-of-vehicle time in the unrestricted S.W. area models indicate a greater disutility associated with walk, wait and transfer time in the S.W. area. This is a reflection of the public transport service, which was inconvenient by comparison with the N.W. services.

An inconvenience proxy (relative number of journey stages NSTGSD), tested and rejected for the base sample, was highly significant when tested for the S.W. data (see Table 4). Its coefficient is positive, suggesting that, ceteris paribus, the greater the number of stages by rail relative to car, the greater the utility of car travel. Introducing this variable brings the waiting time variable closer to those of the transfer models. This suggests that inconvenience, as measured and omitted from the transfer models, is important. Moreover, inconvenience is likely to be just one of many unmeasured modal differences between the two areas. The influence of the S.W. mode-specific constant on the car utility expression suggests that other important variables in the S.W. model were absent.

In summary, the transfer model coefficients are for the most part statistically different from the “unrestricted” S.W. coefficients. In particular, the measured level of service variables (time and cost) are not transferable, the main reason being that the transfer coefficients are location specific (to the base sample) and cannot account for differences in unmeasured modal attributes that exist across the two samples.

4 A problem with this approach to accommodating qualitative dimensions in the theoretical derivation of the form of the representative utility expression is that the levels of goods and leisure are now “effective” levels which do not equal the actual measured levels. This makes the parameters ambiguous, especially if the value of travel time savings is a required output.
Table 4
The Influence of Number of Journey Stages on Mode Choice in S.W. Sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASC-CAR</td>
<td>-2.60</td>
<td>-2.70</td>
</tr>
<tr>
<td>G-TIVT</td>
<td>-0.017</td>
<td>-1.22</td>
</tr>
<tr>
<td>G-WAIT</td>
<td>-0.078</td>
<td>-1.51</td>
</tr>
<tr>
<td>G-WTPK</td>
<td>-0.127</td>
<td>-3.34</td>
</tr>
<tr>
<td>CIVC/I</td>
<td>-0.327</td>
<td>-3.20</td>
</tr>
<tr>
<td>RTOC/I</td>
<td>-0.275</td>
<td>-2.39</td>
</tr>
<tr>
<td>CCE/IN</td>
<td>-1.38</td>
<td>-1.94</td>
</tr>
<tr>
<td>NSTGSD</td>
<td>0.769</td>
<td>4.02</td>
</tr>
</tbody>
</table>

\[ \rho^2 = 0.567 \]

PREDICTIVE PERFORMANCE OF THE TRANSFER MODELS

So far, we have shown that, if the transferability of the N.W. models is to be judged on statistical criteria, there is little reason to claim that the coefficients of the models are transferable. However, in practice a level of transferability is required which is sufficient for specific planning needs. It would thus seem relevant to gauge the predictive ability of the transfer models in relation to the estimates which could be obtained from “unrestricted” S.W. models.

To do this, the observed modal split of the S.W. sample was compared with the modal split predicted by the transfer models, using values for the independent variables in the S.W. data set. The prediction-success tables for the three pairs of “unrestricted” and transferred models are given in Table 5.

The prediction success table is normally a goodness-of-fit measure, used on the sample which provided estimation of the model. Here it can be used to test prediction, because the transfer model coefficients are used to predict the modal split of a different population. Table 5 shows that the transfer Models B and C appear to provide good predictions of the aggregate modal split for the S.W. sample, but Model A does not. But the performance of all three models in predicting the mode-choice behaviour of rail travellers is poor. For example, of the 33 observed rail users, the best predicted level (Model B) was 16.

While the practical concern is likely to be on the aggregate modal split predictions for an area, there is interest also in knowing how well the model predicts the behaviour of each individual. From Table 5, there is very little difference between the models in the proportion of individuals successfully predicted: if anything, Model A performs better (87%) than Models B (83%) and C (82%). The index is of little value, however, as a relative measure, as it is insensitive to the relative prediction success of each mode separately; moreover, where one form of transport is revealed as a strongly

---

3 The predicted modal share is computed by calculating individual probabilities of choosing the available modes, and then summing these probabilities over the entire sample.
TABLE 5

**Prediction Success**

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Observed</th>
<th>Predicted</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car</td>
<td>Rail</td>
<td>Total</td>
<td>share</td>
</tr>
<tr>
<td>Car</td>
<td>196</td>
<td>14</td>
<td>210</td>
</tr>
<tr>
<td>Rail</td>
<td>14</td>
<td>19</td>
<td>33</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Observed</th>
<th>Predicted</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car</td>
<td>201</td>
<td>9</td>
<td>210</td>
</tr>
<tr>
<td>Rail</td>
<td>22</td>
<td>11</td>
<td>33</td>
</tr>
</tbody>
</table>

Overall proportion successfully predicted: 88%

---

**Model A—Unrestricted**

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car</td>
<td>196</td>
</tr>
<tr>
<td>Rail</td>
<td>14</td>
</tr>
<tr>
<td>Total</td>
<td>210</td>
</tr>
<tr>
<td>Observed</td>
<td>86%</td>
</tr>
</tbody>
</table>

Overall proportion successfully predicted: 88%

---

**Model A—Transferred**

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car</td>
<td>201</td>
</tr>
<tr>
<td>Rail</td>
<td>22</td>
</tr>
<tr>
<td>Total</td>
<td>223</td>
</tr>
<tr>
<td>Observed</td>
<td>92%</td>
</tr>
</tbody>
</table>

Overall proportion successfully predicted: 87%

---

**Model B—Unrestricted**

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car</td>
<td>196</td>
</tr>
<tr>
<td>Rail</td>
<td>14</td>
</tr>
<tr>
<td>Total</td>
<td>210</td>
</tr>
<tr>
<td>Observed</td>
<td>86%</td>
</tr>
</tbody>
</table>

Overall proportion successfully predicted: 88%

---

**Model B—Transferred**

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car</td>
<td>186</td>
</tr>
<tr>
<td>Rail</td>
<td>17</td>
</tr>
<tr>
<td>Total</td>
<td>203</td>
</tr>
<tr>
<td>Observed</td>
<td>84%</td>
</tr>
</tbody>
</table>

Overall proportion successfully predicted: 83%

---

**Model C—Unrestricted**

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car</td>
<td>195</td>
</tr>
<tr>
<td>Rail</td>
<td>15</td>
</tr>
<tr>
<td>Total</td>
<td>210</td>
</tr>
<tr>
<td>Observed</td>
<td>86%</td>
</tr>
</tbody>
</table>

Overall proportion successfully predicted: 88%

---

**Model C—Transferred**

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car</td>
<td>186</td>
</tr>
<tr>
<td>Rail</td>
<td>19</td>
</tr>
<tr>
<td>Total</td>
<td>205</td>
</tr>
<tr>
<td>Observed</td>
<td>84%</td>
</tr>
</tbody>
</table>

Overall proportion successfully predicted: 82%
preferred mode for an area (for example, car in the S.W. sample), the prediction success of this mode swamps the influence of the other.  

From a practical viewpoint, the ability of the model to predict the choice of mode for any one individual may not be important: for example, if accurate aggregate forecasts of model usage are all that is required. On this criterion, at least two of the three transfer models adequately predict the actual mode split for the S.W. sample. Table 6 summarises the aggregate predictions of the three models, relative to the observed modal split of the sample. Model C provides marginally better results than Model B, and both B and C perform better than Model A, which underpredicts rail usage by 40%.

In summary, on an aggregate basis, the transfer models (C and B in particular) appear to provide good predictions of the S.W. modal split. However, the models are unable to predict the choice behaviour of rail users closely.

In view of this, a logical question is—what about the predictive success of transferred "naive" models? On pragmatic grounds, where predictive ability is often the main criterion, and an understanding of the causes and consequences of change is a secondary concern, a naive model may be more suitable. Although the naive model is clearly underspecified (for example, containing only overall time and cost and a constant), it is not difficult to visualize a situation where the coefficients are statistically similar between the areas simply because the distributions of the variables are likely to be more compatible, ceteris paribus, as the coarseness of the variables increases.

To throw some light on the transferability of naive models, we estimated a model on two generic variables (overall travel time and overall travel cost) and the alternative—specific constant. Time was selected as one of variables, since it picks up much of the spatial separation of individuals from opportunities.

---

6 To illustrate, if every individual in the S.W. sample was predicted to go by car the overall prediction success would be 86%. McFadden (1979) suggests developing other indices, such as the modal success index and the overall success index; but these indices are of little value, as it is still not clear for what range of the indices we should regard the model as satisfactory.
TRANSFERABILITY OF MODE CHOICE MODELS  R. A. Galbraith and D. A. Hensher

The calculated $\chi^2$ value (Table 7) exceeds the critical value (at the 5% level). Thus the null hypothesis that the two sets of coefficients are equal is rejected, a result in accordance with Models A to C. The predictive ability of the naive model (Table 8) on all criteria applied to Models A to C is identical to that for Model C. Naïve models such as this may have predictive capability, but they are likely to be too general for assessing particular policy changes (e.g. increase in walk times due to route reassignment); however, this possible disadvantage may be compensated by the usefulness of such a model in broad (or system) policy assessment (e.g. global fare increases of 40%). Judgement of the suitability of naïve models must be left to the domain of empiricism.\(^3\)

EMPIRICAL EVALUATION OF UPDATING PROCEDURES

In this section, it is assumed that a sample of observations of behaviour of individual travellers from the S.W. area is available for use in updating each of the three transfer models. An empirical test is undertaken of the scaling and the “Bayesian update” methods, both developed by Atherton and Ben Akiva (1976).

Scaling update

The scaling procedure uses the disaggregate sample to re-estimate the mode-specific constant and to estimate a “scalar”, which is used to scale the other coefficients, so as to keep the ratios between them unchanged. The rationale is that the average effects of unmeasured factors are likely to vary between the two suburban areas, and these effects can be reflected in the adjustment of the constant. The rationale for adjusting the scale of the other coefficients is that travellers in different areas may differ in the level of importance they attach to variables in the mode choice equation, but it is also likely that the relative weights attached to these variables will be the same for both areas.\(^8\)

The update constant and scalar were calculated by estimating a model with one independent variable $V$ (Charles River Associates, 1978). This variable is formed by combining the vector of the original transfer model coefficients ($a$) for all variables except the constant, and the vector of values of the independent variables $x_q$ for the $q$th individual, so that:

$$V_q = \alpha x_q, \quad q = 1, \ldots, Q$$

(4)

The constant estimated from this single variable model is the constant appropriate for

\(^3\) Another “naïve” model was estimated, with only two variables—the alternative specific constant and the car competition effect (CCE). $\chi^2$ was 0.240; the overall percentage successfully predicted was 77%; and the percentage of observed car (train) predicted to use car (train) was 87% (15%).

\(^8\) Maintaining the scale of the coefficients assumes constant trade-offs between attributes; for example, a constant “value of travel time savings” is preserved.
### Table 7

*A Naive Model and Relative Power of Transferability*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Northwestern Area</th>
<th>Southwestern Area</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unrestricted</td>
<td>Transferred</td>
</tr>
<tr>
<td></td>
<td>$\beta$</td>
<td>$t$</td>
</tr>
<tr>
<td><strong>ASC-Car</strong></td>
<td>0.636</td>
<td>2.51</td>
</tr>
<tr>
<td><strong>G-TIME</strong></td>
<td>-0.025</td>
<td>-3.09</td>
</tr>
<tr>
<td><strong>G-COST</strong></td>
<td>-0.050</td>
<td>-7.11</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>No. of cases</th>
<th>$L^*(0)$</th>
<th>$L^*(\beta)$</th>
<th>$L^*(<em>{T</em>{ex}})$</th>
<th>$L^*(SW_{ex})$</th>
<th>Calculated $\chi^2$</th>
<th>Critical $\chi^2$</th>
<th>$\rho^2$</th>
<th>$\rho^2_{trans}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>332</td>
<td>-228.58</td>
<td>-155.36</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.320</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>243</td>
<td>-96.54</td>
<td>-65.35</td>
<td>—</td>
<td>-65.35</td>
<td>13.22</td>
<td>5.99</td>
<td>—</td>
<td>0.255</td>
</tr>
</tbody>
</table>

### Table 8

*Prediction Success—Naive Models*

<table>
<thead>
<tr>
<th></th>
<th>Unrestricted Model</th>
<th>Transferred Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Predicted Car</td>
<td>Observed Rail</td>
</tr>
<tr>
<td><strong>Observed:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car</td>
<td>190</td>
<td>20</td>
</tr>
<tr>
<td>Rail</td>
<td>20</td>
<td>13</td>
</tr>
<tr>
<td><strong>Predicted total</strong></td>
<td>210</td>
<td>33</td>
</tr>
<tr>
<td><strong>Predicted share</strong></td>
<td>86%</td>
<td>14%</td>
</tr>
<tr>
<td><strong>Overall proportion successfully predicted</strong></td>
<td>83.5%</td>
<td></td>
</tr>
</tbody>
</table>
the updated model, and the estimated coefficient of \( V \) is the scalar, which is then multiplied by the original coefficients to obtain the updated coefficients.

**Bayesian update**

The Bayesian procedure combines the information contained in the base (N.W.) and new disaggregate (S.W.) samples, by computing the updated coefficient on any one variable as the weighted average of the coefficient of that variable as estimated in the base model, and the coefficient as estimated with the sample from the new area.\(^9\) The weights used are the inverses of the variances of the two coefficient estimates. Thus:

\[
\beta_{\text{upd}} = \left( \frac{\beta_{nw}}{\sigma_{nw}^2} + \frac{\beta_{sw}}{\sigma_{sw}^2} \right) / \left( \frac{1}{\sigma_{nw}^2} + \frac{1}{\sigma_{sw}^2} \right)
\]

(5)

where \( \beta_{nw} \) and \( \beta_{sw} \) are the original and new sample coefficients, \( \sigma_{nw} \) and \( \sigma_{sw} \) are standard deviations of the original and sample coefficients, and \( \beta_{\text{upd}} \) is the new updated coefficient for the variable.

**Updating results**

The results obtained from applying the updating procedures to each of the transfer models are presented in Table 9. For each of the three models, column 1 lists the unrestricted S.W. model coefficients. Column 2 gives the updated coefficients using the scaling procedure, and column 3 lists the Bayesian updated parameters. \( \rho^2 \) is again used as a relative measure of the goodness-of-fit of each model to the S.W. data.

For each transfer model, the Bayesian updating procedure resulted in a model which, predictably, performed better than the “no update” model (\( \rho^2_{\text{no update}} \)), but not as well as the unrestricted model. In contrast, the scaling update procedure improved the goodness-of-fit of transfer Model A only. In Models B and C, the \( \rho^2 \) decreased when the scaling update was applied. The scaling approach assumes that if the weightings applied to variables by people differ in different areas, only the absolute levels vary; the marginal rates of substitution between the variables remain invariant. While this may or may not be a realistic hypothesis, the problem in practice is that the pairs of coefficients do not all differ in the same direction, so that, when a single scaling factor is applied, some coefficients in the model benefit (that is, become closer to the unrestricted value), while others do not. Thus it is possible for the procedure to decrease the fit of a model, as in Models B and C.

Following the approach used in a previous section, the statistical difference between the updated models and the unrestricted models was tested using the chi-square test.

In each of the three scaling update models, the calculated \( \chi^2 \) values were large, so that the null hypothesis of statistical transferability was rejected in each case. Applying the Bayesian update to Model A produced a similar result. However, for Models B and C, the Bayesian procedure produced calculated \( \chi^2 \) values (16.2 and 15.0 respectively) which were only marginally greater than the respective critical values (14.1 and 12.6).

---

\(^9\) Thus, this procedure approximates the “pooling” of the base sample with the new sample, but obviates the need to go back to the original sample data.
### Table 9
The Updated Models

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model A</th>
<th></th>
<th></th>
<th>Model B</th>
<th></th>
<th></th>
<th>Model C</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1) Unrestricted</td>
<td>(2) Scaling</td>
<td>(3) Bayesian</td>
<td>(1) Unrestricted</td>
<td>(2) Scaling</td>
<td>(3) Bayesian</td>
<td>(1) Unrestricted</td>
<td>(2) Scaling</td>
<td>(3) Bayesian</td>
</tr>
<tr>
<td></td>
<td>$\beta$</td>
<td>$t$</td>
<td>$\beta$</td>
<td>$\beta$</td>
<td>$t$</td>
<td>$\beta$</td>
<td>$\beta$</td>
<td>$t$</td>
<td>$\beta$</td>
</tr>
<tr>
<td>ASC-CAR</td>
<td>-2.382</td>
<td>-2.57</td>
<td>0.505</td>
<td>-0.523</td>
<td>-2.244</td>
<td>-2.53</td>
<td>1.868</td>
<td>-0.0215</td>
<td>-2.731</td>
</tr>
<tr>
<td>IVTM-C</td>
<td>-0.0243</td>
<td>-1.59</td>
<td>-0.0302</td>
<td>-0.0428</td>
<td>-0.0254</td>
<td>-1.87</td>
<td>-0.0481</td>
<td>-0.0405</td>
<td></td>
</tr>
<tr>
<td>G-IVTC</td>
<td>0.0077</td>
<td>0.20</td>
<td>-0.0217</td>
<td>-0.0316</td>
<td>0.0021</td>
<td>0.06</td>
<td>-0.0575</td>
<td>-0.0457</td>
<td></td>
</tr>
<tr>
<td>G-TIVT</td>
<td>-0.1388</td>
<td>-4.55</td>
<td>-0.0174</td>
<td>-0.0535</td>
<td>-0.1393</td>
<td>-4.57</td>
<td>-0.0329</td>
<td>-0.0524</td>
<td>-0.136</td>
</tr>
<tr>
<td>TIVC-C</td>
<td>-0.0326</td>
<td>-2.44</td>
<td>-0.0210</td>
<td>-0.0365</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PARK-C</td>
<td>0.0196</td>
<td>0.44</td>
<td>-0.0941</td>
<td>-0.0173</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTC-R</td>
<td>-0.0472</td>
<td>-3.04</td>
<td>-0.0486</td>
<td>-0.0676</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CIVC/I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.2496</td>
<td>-2.45</td>
<td>-0.0675</td>
<td>-0.0752</td>
<td>-0.323</td>
</tr>
<tr>
<td>PARK/I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.1499</td>
<td>0.37</td>
<td>-0.1014</td>
<td>-0.0991</td>
<td>0.1178</td>
</tr>
<tr>
<td>RTOC/I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.3537</td>
<td>-3.36</td>
<td>-0.1326</td>
<td>-0.1573</td>
<td>-0.334</td>
</tr>
<tr>
<td>CCE/IN</td>
<td>-1.726</td>
<td>-2.81</td>
<td>-0.723</td>
<td>-1.446</td>
<td>-1.697</td>
<td>-2.80</td>
<td>-1.460</td>
<td>-1.543</td>
<td>-1.73</td>
</tr>
<tr>
<td>$L^*(0)$</td>
<td>-96.54</td>
<td>-96.54</td>
<td>-96.54</td>
<td>-96.54</td>
<td>-96.54</td>
<td>-96.54</td>
<td>-96.54</td>
<td>-96.54</td>
<td>-96.54</td>
</tr>
<tr>
<td>$L^*(\beta)$</td>
<td>-49.66</td>
<td>-68.03</td>
<td>-60.97</td>
<td>-49.06</td>
<td>-68.84</td>
<td>-57.16</td>
<td>-50.56</td>
<td>-67.46</td>
<td>-58.10</td>
</tr>
<tr>
<td>$L^*(\text{IVTC})$</td>
<td>-68.03</td>
<td>-60.97</td>
<td>-68.84</td>
<td>-57.16</td>
<td>-67.46</td>
<td>-58.10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$L^*(\text{SW}))$</td>
<td>-49.66</td>
<td>-49.66</td>
<td>-49.06</td>
<td>-49.06</td>
<td>-50.56</td>
<td>-50.56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calculated $\chi^2$</td>
<td>36.6</td>
<td>22.6</td>
<td>39.4</td>
<td>16.2</td>
<td>33.8</td>
<td>15.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Critical $\chi^2$</td>
<td>14.1</td>
<td>14.1</td>
<td>14.1</td>
<td>14.1</td>
<td>12.6</td>
<td>12.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho^2$</td>
<td>0.486</td>
<td>0.288</td>
<td>0.362</td>
<td>0.492</td>
<td>0.279</td>
<td>0.402</td>
<td>0.476</td>
<td>0.294</td>
<td>0.391</td>
</tr>
<tr>
<td>$\rho^2_{\text{no update}}$</td>
<td>0.175</td>
<td>0.334</td>
<td>0.356</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* $\rho^2_{\text{no update}}$ refers to the goodness-of-fit of the original "no update" Northwestern coefficients on the Southwestern sample. It is included here to allow comparison with the $\rho^2$ values of the updated models.*
TRANSFERABILITY OF MODE Choice MODELS R. A. Galbraith and D. A. Hensher

In fact, at the 99% confidence level, the null hypothesis that the models are not significantly different could not have been rejected. This provides some measure of the statistical similarity of the Bayesian update transfer model and the unrestricted S.W. model for specifications B and C.10

In summary, the Bayesian updating approach performs quite well if judged on statistical criteria. This is not surprising, since, by definition, the Bayesian approach offers a direct compromise between the two models. In practice, however, the benefits resulting from a Bayesian update are likely to be less pronounced unless the estimated coefficients of the small disaggregate sample have small standard errors. If the standard errors are not small, the updated coefficients will be based primarily on the original transfer model coefficients, and there will be little benefit from updating.

CONCLUSION

This article has sought to evaluate the hypothesis of transferability of disaggregate travel-choice model coefficients. Although transferability is frequently linked to validation of the disaggregate modelling approach, model validity is a necessary but not sufficient condition for transferability. Theoretically the argument for transferability may be realistic. However, the potential decreases as we move from theory to practical estimation, and as additional assumptions and imperfect data reduce the range and validity of the model.

The empirical study was designed to reduce the data errors implicit in previous studies, by analysing two compatible data sets. In general, different models estimated on the two data sets produced coefficients which were somewhat dissimilar. Short of updating the transfer models to improve their goodness-of-fit, the models were not on the whole statistically or observably transferable.

A chief impediment to transferability appeared to be the range and types of explanatory variables applied in a different context. The principal implication is that if disaggregate choice models are to realize the potential for transferability, the models must include quantitative variables better able to account for socio-economic factors, “unmeasured” level of service attributes, and situational or contextual factors which explain individual travel behaviour. The majority of current data sets fail in this respect. Unfortunately, despite the intuitive plausibility of this conclusion, and the wealth of studies with similar comments in the broader context of choice modelling, we continue to witness a classical (1960s) approach to data selection and collection.11 Alas, it is feared that these conclusions will continue to be ignored.

One kind of variable deserves particular attention. This can be termed “context”, and involves a number of separate measures related to the spatial structure, the

---

10 Note that the null hypothesis can only be accepted for Models B and C above if $\alpha$ is decreased from 5% to 1%.

11 There are exceptions, such as the Oxford University Transport Studies Unit Activity data set. However, major planning agencies maintain traditional data configurations, including those which have recently applied complete systems of individual travel-choice models. The Oxford group are currently developing travel-relevant stage-in-the-lifecycle segments so that socio-economic effects can be accommodated.
physical environment and the cultural environment of an area. Because most of the past modelling efforts have focused on single areas, there has been little reason to consider the spatial character of urban systems, their physical/environmental make-up or their cultural differences. They are falsely assumed to be constant within the urban area.

Thus, if we require a model which is sufficiently well specified to account for differences in tastes and preferences in different areas, we need to consider contextual variables which account for place variation. There would be a need to develop variables concerning the nature of transport networks and supplies, and other factors such as natural features and barriers, and differences in cultural experiences in living and travelling (Louviere, 1981). This, however, describes a whole new direction for modelling research.

In conclusion, there is a fairly limited scope for achieving greatly improved model specifications above and beyond variables such as those used in this study, within the current “state of the art” in disaggregate mode-choice modelling. These considerations suggest that past assurances that disaggregate models are transferable have been premature. True, the statistical transfer performance of the models tested in this study was weakened by the inadequate size and variability of the samples—a defect which applies to most disaggregate data sets, and which could be avoided. However, many problems cannot be resolved simply by increasing sample sizes, particularly if a variable just is not important in the new context. In future studies where transferability is a prime aim, we should consider estimating a model in the base area on criteria which define the new area.

APPENDIX A

Past empirical studies of transferability of model coefficients

Studies dealing with transferability have sought to answer the questions: Can disaggregate travel-choice models be transferred from one area to another without modification of the coefficients? If not, are modifications short of complete re-estimation of the models feasible? This review is structured broadly according to the “transfer hierarchy”, beginning with studies where no modification has taken place, and going on the studies which have used various procedures to update the coefficients, using information from the area of interest.

Tests of the “no update” procedure

Four studies have attempted to test the transferability of model parameters, estimated for one area, and applied without modification to a second area (Atherton and Ben Akiva, 1976; Watson and Westin, 1975; Talvitie and Kirshner, 1978; Stopher and Wilmot, 1979).

---

12 This emphasis tends to counter the argument that disaggregate choice models can be estimated on smaller samples than aggregate models.
Atherton and Ben Akiva (1976) explained travel behaviour in Los Angeles and New Bedford, Massachusetts, from a work trip mode choice model estimated on Washington, D.C., data. Their model predicted the probability of choosing to drive alone, share a ride, or use public transit. The independent variables included network-derived travel times and costs, income, car availability, a dummy variable indicating whether the trip maker was household head, and mode-specific constants. The test was to use the variables for the original (Washington) model to estimate new models with Los Angeles and New Bedford data, and then to compare the coefficients of the new models with those of the old model. The comparison consisted of statistical tests of the null hypothesis that the coefficients of the new models are equal to those of the old. For both models the coefficients of the car availability variables were the only ones significantly different from their Washington counterparts. The authors concluded that the evidence was “encouraging”; but, as no model is ever completely specified and hence perfectly transferable, they suggested that updating procedures were required.

Watson and Westin (1975) studied the transferability of logit mode-choice models among different subareas within a single urban area. Their data were for the Edinburgh–Glasgow area of Scotland, and their model was a binary choice model containing level of service variables and a mode-specific constant, but no socio-economic variables. The data were grouped into six categories according to whether the trip origins and destinations were in the central city, the suburbs, or the area peripheral to the urban area. Each of the six models was then used to predict the mode splits of the other five samples. They also conducted tests for significant differences of all coefficients of the six models. In the three categories that contained at least one trip end in the central city, they found that within the group the models predicted well. The remaining three categories performed poorly within themselves. As between the non-central groups, the coefficients were significantly different. Watson and Westin concluded that the predictive ability of the model of the central city was fairly favourable to transferability, but that the results for the other group indicated a need to refine the models for locational differences.

Talvitie and Kirshner’s results (1978) were substantially less optimistic. They used four data sets, comprising standard home interview survey data appended with network-supplied information on modal cost and time. The model coefficients were highly sensitive to model specification, having a significant impact on explanatory (and hence transferability) power. Moreover, judging from statistical criteria, they found “little ground to claim that the coefficients of the work mode-choice models are transferable” (p. 132); that is, the model coefficients did not appear to be transferable within a region, between regions, or over time.

As part of a study of individual choice models of work-trip mode choice in Johannesburg, Stopher and Wilmot (1979) attempted to draw comparisons between South African models and several recent United States models of work-trip mode choice. The South African data concerned mode choices for the work-trip mode by high income, high car-owning residents of a large metropolitan area with reasonable public transport service. The population segment was thus similar to all the U.S.A. populations. The results suggested, for the comparisons that could be made, that the coefficients of the South African model fell within the range of the coefficients
determined in the U.S., providing some evidence of geographic transferability, at least within a socio-economic stratum.

Studies concerned with model updating

Updating procedures use information from the area to which a model is to be transferred to improve the model's predictive ability. They have been discussed by Atherton and Ben-Akiva (1976) as well as by Daly (1974) and Charles River Associates (1978).

Atherton and Ben-Akiva (1976) tested four updating procedures. The first assumes the existence of aggregate (e.g. census) data. The others require a small disaggregate sample from the area of interest.

1. The first procedure uses data on the aggregate population shares of the various modes, along with population averages of the independent variables, to adjust the mode-specific constant to predict these modal shares more accurately. The rationale is that, since the constant captures the mean effects of unobserved factors, these constants are likely to vary between areas.

2. The second procedure estimates a new model from a small disaggregate data set, using the specification of the old model. The drawback is that for small samples the coefficient estimates are likely to be unreliable.

3. The third procedure uses the disaggregate sample to estimate the mode-specific constant, and to estimate a "scalar" used to scale all the other coefficients so as to keep the ratios between them unchanged.

4. The fourth procedure uses Bayesian techniques to combine the coefficients obtained from the transfer model and the small disaggregate sample model.

The authors found that the fourth procedure performed best, and was a slight improvement over the "no update" model (which, as was stated in the previous section, performed quite well). The third procedure performed poorly, but better than the second procedure.

Research in the U.K. similar to that of Atherton and Ben-Akiva investigated the applicability of a binary-logit choice model (containing only a mode-specific constant and level of service variables) to mode choice situations in several English towns (Daly, 1974). The models for different towns were sufficiently different to prevent a model for one town from being successfully used in another. However, a satisfactory "global model" (a single model estimated on data pooled across all towns) could be constructed by estimating common coefficients for the level of service variables and then, for each town, adjusting the mode-specific constant to capture factors peculiar to that town (as in the first procedure above).

Charles River Associates (1978) applied Atherton and Ben-Akiva's scaling procedure. Instead of a small disaggregate data set, the procedure used aggregate data to select appropriate market segments. The approach attempts to minimise error in aggregation by constructing relatively homogeneous "cells" of individuals similar in observed attributes (such as levels of car availability, income, socio-economic class) that presumably determine travel choice behaviour. The principle argument for this procedure is that it provides a better fit to the data on market segments, while
TRANSFERABILITY OF MODE CHOICE MODELS  R. A. Galbraith and D. A. Hensher

preserving the ratio of the coefficients. The test was to predict the ridership of a new Park'n Ride service in Baltimore. The model was found to under-predict Park'n Ride use by 50%.

APPENDIX B

The Study Areas

Northwestern sample

The first data set was collected in April 1971 (Hensher, 1972). The Hornsby area was chosen because it contains a good cross-section of residents both of relatively newly developed areas and of the older, well established areas, and is well served by public transport.

The Hornsby area is a predominantly middle class outer-residential area on the North Shore, 25 kilometers from the CBD. Some of the older residences date from the 1920s, and most of the area was developed by the early 1960s. The study area follows the Main North railway for about four kilometers in a north-south direction, while the North Shore line, which joins the Main North line at Hornsby station, provides a direct service to the CBD. All the residences in the study area are less than 1.5 kilometers from the railway, while connecting buses to railway interchanges have good scheduling during peak hours. Commuters living in the Hornsby area are thus faced with an effective choice between public transport and car as their mode of transport.

Southwestern sample

The second data set has been drawn from data collected in the Southwestern region of Sydney in March 1975, for use in a study by what was then the Commonwealth Bureau of Roads (1976). The areas surveyed included the suburbs of Fairfield, Green Valley and Macquarie Fields. These locations had been identified as areas in which persons were considered to be relatively transport disadvantaged in the context of the journey to and from work; reflected in relatively high work journey costs, low average incomes and poor provision of public transport.

The Southwestern study area is approximately 35 kilometers from the central city, and consists of mainly “working class” suburbs, subdivided extensively for residential purposes. The northern extremity of the area, Fairfield, was developed in the early 1960s. Green Valley, to the southwest, has slightly newer houses, in crescents and culs-de-sac. Further south still is Macquarie Fields, an area at the very fringe of urban development in 1975. Fairfield and Green Valley are located away from the railway network. The Main Southern Line passes through the southern region of the area, but feeder bus services were generally not well established in 1975.

---

13 In this “market segmentation” updating approach, forecasting is accomplished by applying the model to each cell separately to predict behaviour; the relative frequencies of the cells for the market segments are then used to weight the cells and predict aggregate modal choice behaviour. Hence, the market segmentation approach is computationally simple and requires only aggregate census data (to segment the population); but a very large data set is required (to enable all cells to be satisfactorily represented).
# APPENDIX C

## The Variables

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Unit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASC-Car</td>
<td></td>
<td>Alternative specific constant—car</td>
</tr>
<tr>
<td>IVTC-C</td>
<td>minutes</td>
<td>In-vehicle time—connecting modes</td>
</tr>
<tr>
<td>WAIT-C</td>
<td>minutes</td>
<td>Waiting time</td>
</tr>
<tr>
<td>WKTP-C</td>
<td>minutes</td>
<td>Walking, transfer and parking time</td>
</tr>
<tr>
<td>TOVT-C</td>
<td>minutes</td>
<td>Total out-of-vehicle time</td>
</tr>
<tr>
<td>TOTT-C</td>
<td>minutes</td>
<td>Total time</td>
</tr>
<tr>
<td>TIVC-C</td>
<td>cents</td>
<td>Total in-vehicle cost</td>
</tr>
<tr>
<td>TOTC-C</td>
<td>cents</td>
<td>Total cost</td>
</tr>
<tr>
<td>DIST-C</td>
<td>miles</td>
<td>Distance travelled by car</td>
</tr>
<tr>
<td>INCOME</td>
<td>$000</td>
<td>Gross annual individual income</td>
</tr>
<tr>
<td>IC-C/I</td>
<td>cents per minute</td>
<td>Car cost divided by income</td>
</tr>
<tr>
<td>NSTGSD</td>
<td></td>
<td>Number of journey stages by car less rail</td>
</tr>
<tr>
<td>NSTGSR</td>
<td></td>
<td>Number of journey stages by car divided by rail</td>
</tr>
<tr>
<td>CCE</td>
<td></td>
<td>1 if number of licenced drivers &gt; number of cars; 0 otherwise</td>
</tr>
<tr>
<td>IVTC*1</td>
<td>minutes*cents</td>
<td>In-vehicle car time multiplied by income</td>
</tr>
<tr>
<td>OVTIC*1</td>
<td>minutes*cents</td>
<td>Out-of-vehicle car time multiplied by income</td>
</tr>
<tr>
<td>IVTM-R</td>
<td>minutes</td>
<td>In-vehicle time—main mode</td>
</tr>
<tr>
<td>IVTC-R</td>
<td>minutes</td>
<td>In-vehicle time—connecting modes</td>
</tr>
<tr>
<td>TIVT-R</td>
<td>minutes</td>
<td>Total in-vehicle time</td>
</tr>
<tr>
<td>WAIT-R</td>
<td>minutes</td>
<td>Waiting time</td>
</tr>
<tr>
<td>WKTP-R</td>
<td>minutes</td>
<td>Walking, transfer and parking time</td>
</tr>
<tr>
<td>TOVT-R</td>
<td>minutes</td>
<td>Total out-of-vehicle time</td>
</tr>
<tr>
<td>TOTT-R</td>
<td>minutes</td>
<td>Total time</td>
</tr>
<tr>
<td>IVTR*1</td>
<td>minutes*cents</td>
<td>In-vehicle rail time multiplied by income</td>
</tr>
<tr>
<td>OVRTIC*1</td>
<td>minutes*cents</td>
<td>Out-of-vehicle rail time multiplied by income</td>
</tr>
<tr>
<td>G-IVTM</td>
<td>minutes</td>
<td>In-vehicle time—main mode</td>
</tr>
<tr>
<td>G-WAIT</td>
<td>minutes</td>
<td>Waiting time</td>
</tr>
<tr>
<td>G-WKTP</td>
<td>minutes</td>
<td>Walking, transfer and parking time</td>
</tr>
<tr>
<td>G-TOTT</td>
<td>minutes</td>
<td>Total travel time</td>
</tr>
<tr>
<td>G-TIVC</td>
<td>cents</td>
<td>Total in-vehicle cost</td>
</tr>
<tr>
<td>G-TOTC</td>
<td>cents</td>
<td>Total travel cost</td>
</tr>
</tbody>
</table>

## REFERENCES


28
TRANSFERABILITY OF MODE CHOICE MODELS  R. A. Galbraith and D. A. Hensher


