1. INTRODUCTION

The optimal design of Stated Preference (SP) experiments has shown the feasibility of improving the experimental designs such that better estimates can be obtained at the end of the process, when compared with models estimated with data gathered using standard experimental designs. This optimisation approach has been used to design SP experiments to value environmental aspects, to detect the presence of income effect in mode choice and to assess the subjective value of time components. In all cases Multinomial Logit models have been used to model the demand or choice process, due to its simplicity.

In this paper the attention has focused on using a different demand model, such that taste variation could be detected in waiting or travel times. A Mixed Logit model was chosen to find out the presence of this condition. Given the open form of this demand model, then a combination of optimisation and simulation had to be used when looking for the best SP experimental design. Software codes were prepared such that the optimisation procedure could be carried out, since the minimization of the coefficient variances required some numerical analysis. These codes were linked with standard simulation Mixed Logit routines, such that the optimal designs could be obtained.

This two part procedure was used to design a SP experiment dealing with the improvement of a rail service in the Great Concepcion, Chile. Cost, waiting and travel times were the explanatory variables. The choice process was presented as a binomial one, comparing the present service with an improved one. Standard experimental designs were also applied, for comparison.

Results show that better estimates can be obtained when applying the optimal designs, with respect to the standard designs. In some cases, taste variation could be detected more robustly, whereas in other cases the occurrence of taste variation could be rejected more strongly.

This article has been organised as follows. Next section deals with the use of SP experiments to model choice processes, including a description of the optimal design of experiments. In section three is explained briefly the demand model being used in this work, whilst the two-step optimisation procedure is described in section four. Experimental work and main results are shown in section five, finishing with the main conclusions and comments in last section.
2. STATED PREFERENCES

The Stated Preference (SP) technique consists of expressing preferences on hypothetical scenarios, upon alternatives that are presented to people. It has its origin in the mathematical psychology (Luce and Tukey, 1964). People's choice of one of the alternatives (or the ranking or rating of them) is utilized to derive the importance of those attributes describing the alternatives. This follows Lancaster (1966) principle regarding the fact that people choose an option due to its properties rather than for being that option per se.

The technique started being used extensively in transport in the 80s, achieving an important development over the years (see Bates (1988), Hensher (1994), Ortuzar (2000) and Louviere et al. (2000)). A later and recent application of SP has to do with the valuation of environmental assets (Bateman et al., 2000; Rizzi and Ortuzar, 2003; Wardman and Bristow, 2004; Tudela and Garcia, 2004; Tudela et al., 2006).

An important issue related to the use of SP experiments has to do with the design of the experiments. There are proved procedures for the design (Kocur et al., 1982; Louviere et al., 2000), although alternative approaches, based on optimisation, have been suggested (Toner et al., 1998; Toner et al., 1999).

Applications of the SP technique rested for many years on the classical design of experiments, in the sense of asking for orthogonal designs (Kocur et al., 1982), such that attributes should not be correlated. This condition started to be criticised for many reasons, either empirical as well as theoretical. Fowkes et al. (1993), Toner et al. (1998), Toner et al. (1999) and other authors did notice that the condition of being orthogonal on attributes was not transmitted to the estimates when using a non linear modelling of responses, being the case of the Logit model. This implied that there was no need to be concerned about preparing orthogonal SP designs. Besides, realism reasons as well as the posterior usage of estimates (for prediction or to derive economic values) implied that many times the orthogonal condition on designs should be left aside for the sake of the estimates final use (Fowkes and Wardman, 1988).

Clark and Toner (1996) proposed that better designs, than the classical ones, might be obtained using an optimisation procedure. The analyst should define what he/she wants to do with the data collected using SP experiments, such that the design could be adjusted to achieve his/her requirements. This procedure consists of defining the number of games in the experiment, and the attribute levels that will describe the alternatives. Given a preliminary design, that will be optimised, and some prior coefficients, then the likelihood function is built up. From this function is possible to derive second derivatives matrix (Hessian matrix) and the variance-covariance matrix (Ben Akiva and Lerman, 1982), whose components might be defined as the objective function to be optimised. Several objective functions can be outlined: the individual variance of estimates, a weighted sum of variances, and a non linear function of variances that considers the covariance between estimates, for example. Constraints for attribute magnitude can be added such that attribute levels are
reasonable; figure 1 shows a schema of the optimisation procedure. It can be noticed that the method rests on the existence of some previous coefficients. These estimates will come from earlier studies, which are valid for the population that will answer the SP survey, or will have to be estimated using ad-hoc surveys for small samples from the interest population.

This procedure has been used by Tudela and Muñoz (2002), Carlsson and Martinsson (2003) and Tudela and García (2004a, 2004b), for the estimation of subjective values of time, identification of health attributes and environmental assets valuation, respectively. In general more significant estimates were achieved when using the optimal designs with respect to those obtained using the classical designs; a smaller sample size could be needed to achieve similar statistical tests.

![Optimisation scheme](image)

**Figure 1 Optimisation scheme**

### 3. MIXED LOGIT MODEL

Discrete choice models used in transport modelling rest upon the Random Utility Theory (Domencich and McFadden, 1975; Ortuzar and Willumsen, 2001) and on the fact the people choose based on the alternatives attributes (Lancaster, 1966).

Random Utility Theory states that the direct utility (total utility) can be expressed as a deterministic utility (indirect utility) plus a random term (error component). Different discrete choice models can be generated depending on the probabilistic distribution of the error term. If the error term is defined by a Gumbel IID distribution, then the resulting model is a Multinomial Logit. When the error component is given by a Normal distribution, then the obtained model is a Probit (Ben Akiva and Lerman, 1982). More complex models can be developed depending on the relaxation of the error term distribution definition (Bath, 2000; Koppelman and Sethi, 2000)
The observable fact of people choosing different even when they have similar characteristics might correspond to taste variation across the population. This situation can be tackled using Mixed Logit models. These models allow the introduction of random coefficients in the specification of the indirect utility. These random coefficients will permit the modeller to take into account taste variation and other phenomena that cannot be dealt with a Multinomial Logit (Ben Akiva and Bolduc, 1996; Alvarez and Munizaga, 2000; Hensher and Greene, 2003; Train, 2003).

The direct utility related to the i-th alternative and n-th individual, $U_{in}$, can be written as follows

$$U_{in} = V_{in} + \eta_{in} + \varepsilon_{in},$$

where $V_{in}$ is the indirect utility and $\eta_{in}$ and $\varepsilon_{in}$ are error terms. If $\varepsilon_{in}$ is Gumbel IID distributed, and $\eta_{in}$ follows a generalised distribution, defined by a density function $f(\eta / \theta^*)$, then the conditional probability of choosing the i-th option will correspond to a Multinomial Logit, i.e.,

$$P_n(i / \eta) = L_{in}(\eta) = \frac{e^{V_{in} + \eta_{in}}}{\sum_j e^{V_{jn} + \eta_{jn}}}.$$

$\theta^*$ are the distribution parameters, such as the mean and variance. More frequent probability distributions are the Normal, Triangular and Log-Normal (Hensher and Greene, 2003).

The total probability of choosing i, $P_i$, across the whole set of values for $\eta$ can be estimated integrating equation 2. The total probability expression is

$$P_i = \int L_{in}(\eta) \cdot f(\eta / \theta^*) d\eta.$$

This is an open probability model and require of simulation procedures for its evaluation. Several approaches are suggested in the literature for numerical integration of equation 3, such as Monte Carlo simulation (Train, 2003).

### 4. OPTIMISATION PROCEDURE

Since the probability given by a Mixed Logit model (MML) has an open form (equation 3), then a numerical method has to be used to assess the Hessian matrix (see figure 1), such that estimates variance can be calculated and optimised. In particular, the simulated likelihood function, given a preliminary design, coefficients vector and demand model (Mixed Logit), allows the calculation of the second derivatives matrix (Hessian matrix). Figure 2 shows the procedure followed for the optimisation process.
Since a MML model is being used to model the demand, then a part from the usual set of coefficients, it necessary to identify the attribute that might have taste variation across the population, its probability distribution, and the parameters defining such distribution: mean and variance.

The application of the approach requires the definition of the probability integration procedure, number of iterations, error level, etc.

5. APPLICATION AND MAIN RESULTS

The previous optimisation procedure was applied to a simple Stated Preference binary choice experiment. Options being considered were a current passenger train service and an improved train service. The second was actually a new transport mode being implemented in the Great Concepcion, Chile at the time of the survey. Three attributes were considered: cost, in vehicle travel time and waiting time. Experiments were applied to people that lived nearby the railway track and stations, such that the train was a real transport alternative.

5.1 Preliminary coefficients

Given that the optimisation process requires of the attributes coefficients, then a preliminary SP experiment was carried out. A classical design was developed. The boundary values approach was followed to carry out the design. Values being considered finally were as follows. Notice that a full factorial design implied 12 games, with three of them being dominated.
Current cost:  CL$ 300  Alternative cost:  CL$ 270, 340, 370  
Current waiting time: 10 minutes  Alternative waiting time: 5, 14 minutes  
Current travel time: 30 minutes  Alternative travel time: 20, 42 minutes

This design was applied to 10 people, implying a final pseudo sample of 90 observations. Responses were modelled using a binary Logit model, with a linear specification of the utility. Obtained estimates, with t-test in brackets, were

Cost:  -0.0283 (-3.9) 1/CL$
Waiting time: -0.2852 (-3.7) 1/min
Travel time:  -0.1829 (-5.0) 1/min

The subjective values of the waiting and travel time are 10.1 and 6.5 CL$/min respectively.

With these estimates it would be possible to find the best SP design if a binary Logit model were used for the modelling. Since the target of the study was to test for the use of a MML, then a MML had to be estimated using the pilot survey.

Taste variation could be assumed either for the travel time or waiting time coefficients. Preliminary estimations showed that taste variation was stronger for the travel time attribute. Using Gauss codes prepared by and reported in Train (1999), assuming a normal distribution for the coefficient, and 250 repetitions for the random simulation process based on Halton series, the following results were obtained

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Parameter</th>
<th>Estimate</th>
<th>t test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost</td>
<td>Mean</td>
<td>-0.0413</td>
<td>-0.63</td>
</tr>
<tr>
<td>Waiting time</td>
<td>Mean</td>
<td>-0.4084</td>
<td>-0.65</td>
</tr>
<tr>
<td>Travel time</td>
<td>Mean</td>
<td>-0.2608</td>
<td>-0.65</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>0.1529</td>
<td>0.32</td>
</tr>
</tbody>
</table>

The subjective values of the waiting and travel time, assessed in the coefficient mean, are 9.9 and 6.3 CL$/min respectively.

Although estimates are not very significant, these values might be considered reasonable, taking into account the sample size (90 pseudo responses) and the complexity of the demand model. Besides, ratios between coefficients are similar to those obtained for the binary Logit model. Therefore, these values were deemed as correct to be used in the optimisation process.

### 5.2 Optimal design

Coefficients stated in previous section were used to carry on the optimisation process. The preliminary design was the classic one showed previously.
The objective function being considered was the minimisation of the weighted sum of coefficient variances, giving a bigger weight to the travel time SD variance and cost mean variance. This was made because the experiment was oriented to detect taste variation in the travel time attribute.

Constraints were introduced in attributes size, such that realistic values were obtained out of the optimisation process. Train’s code and an ad-hoc code for the optimisation process were used to develop this optimisation stage. For the simulation, 250 random numbers were used, based upon Halton series.

The optimal attribute values for the alternative are shown in table 2. Attributes for the current situation are those in section 5.1.

<table>
<thead>
<tr>
<th>Cost (CL$)</th>
<th>Waiting Time (min)</th>
<th>Travel Time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>240</td>
<td>3</td>
<td>48</td>
</tr>
<tr>
<td>370</td>
<td>6</td>
<td>31</td>
</tr>
<tr>
<td>360</td>
<td>9</td>
<td>16</td>
</tr>
<tr>
<td>330</td>
<td>10</td>
<td>19</td>
</tr>
<tr>
<td>350</td>
<td>4</td>
<td>51</td>
</tr>
<tr>
<td>290</td>
<td>18</td>
<td>28</td>
</tr>
<tr>
<td>260</td>
<td>16</td>
<td>35</td>
</tr>
<tr>
<td>370</td>
<td>11</td>
<td>26</td>
</tr>
<tr>
<td>330</td>
<td>17</td>
<td>15</td>
</tr>
</tbody>
</table>

It can be observed a higher dispersion of attribute values, for all three attributes, when compared with the classical design used in the initial stage.

This optimal design was applied to 22 people (198 pseudo responses). Results for the MML model, with 250 random numbers, are as follows

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Parameter</th>
<th>Estimate</th>
<th>t test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost</td>
<td>Mean</td>
<td>-0,0282</td>
<td>-4,15</td>
</tr>
<tr>
<td>Waiting time</td>
<td>Mean</td>
<td>-0,1690</td>
<td>-3,61</td>
</tr>
<tr>
<td>Travel time</td>
<td>Mean</td>
<td>-0,1262</td>
<td>-3,83</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>0,0586</td>
<td>0,91</td>
</tr>
</tbody>
</table>

The subjective values of the waiting and travel time, calculated in the coefficient mean, are 6.0 and 4.5 CL$/min respectively.
When comparing the results in table 1 and 2, clearly those estimates for the optimal design are much better than those for the classical design. Comparison might be carried out carefully due to difference in sample sizes.

5.3 Analysis of results

A fair comparison between the classical and the optimal designs requires the same size for the data base. To do so, randomly were selected 16 individuals (144 pseudo responses) from the classical and optimal SP experiments, estimating the MML models. Results are shown in table 4.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Parameter</th>
<th>Classic Design</th>
<th>Optimal Design</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Estimate</td>
<td>t test</td>
</tr>
<tr>
<td>Cost</td>
<td>Mean</td>
<td>-0.0476</td>
<td>-0.92</td>
</tr>
<tr>
<td>Waiting time</td>
<td>Mean</td>
<td>-0.3898</td>
<td>-0.92</td>
</tr>
<tr>
<td>Travel time</td>
<td>Mean</td>
<td>-0.2432</td>
<td>-0.95</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>0.1462</td>
<td>0.45</td>
</tr>
</tbody>
</table>

From these results is more evident the convenience of using an optimal SP design for the detection of taste variation, when compared with a classical SP design. Although the t test for the travel time SD is below the significance level, this might be caused by the reduced sample size.

Subjective values of time (CL$/min), when comparing both models (using the classical and optimal designs), are quite different. As in the previous figures, subjective values are relatively smaller when using optimal designs. This might have an important effect on policy making, if we consider that models based on standard SP experiments might be over estimating people’s willingness to pay for time savings. This does need further attention.

<table>
<thead>
<tr>
<th></th>
<th>Classical Design</th>
<th>Optimal Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVWT</td>
<td>8.2</td>
<td>7.0</td>
</tr>
<tr>
<td>SVTT</td>
<td>5.1</td>
<td>4.9</td>
</tr>
</tbody>
</table>

6. COMMENTS AND CONCLUSIONS

The optimal Stated Preference approach proved feasible to be applied when using a more complex demand model, such as a Mixed Multinomial Logit. Previous experiences using Probit models also showed that estimates significance can be improved using these optimal designs.

Although the idea of the optimisation procedure is clear and straightforward, the implementation procedure might be cumbersome, particularly when the demand model requires of a simulation process. Although MML codes are available at the moment, the generation of an optimisation code might be complex.
A more extensive field work should be carried out, such that bigger samples are available, and more conclusive results can be got. This is particularly important when referring to the valuation of time savings.

In this case, given the context and the experiment carried out, just the travel time attribute showed taste variation across the sample. It is expected that bigger samples might allow for the detection of this effect on other attributes, such as the waiting time.

A normal distribution was assumed for the travel time coefficient, just for simplicity. It would be interesting to find out whether the optimal design is sensible to the specification of this distribution.

Acknowledgment
This work was funded by the Universidad de Concepcion Research Office (DIUC).

Bibliography


