Sawtooth Software

RESEARCH PAPER SERIES

Within- and Across- Attribute Constraints in ACA and Full Profile Conjoint Analysis

Ivo A. van der Lans, University of Leiden
Dick R. Wittink, Cornell University
Joel Huber, Duke University
and
Marco Vriens, University of Groningen,
1992

WITHIN- AND ACROSS-ATTRIBUTE CONSTRAINTS IN ACA AND FULL PROFILE CONJOINT ANALYSIS

Ivo A. van der Lans
University of Leiden (The Netherlands)
Dick R. Wittink
Johnson Graduate School of Management, Cornell University
Joel Huber
Fuqua Graduate School of Business, Duke University
Marco Vriens
University of Groningen (The Netherlands)

Published in Sawtooth Software Conference Proceedings, 1992

INTRODUCTION

Studies by Srinivasan *et al.* (1983) and van der Lans and Heiser (1990) indicate that the predictive validity of partworth estimates in conjoint analysis can be improved by imposing constraints on the estimates. Srinivasan *et al.* imposed inequality constraints among partworth estimates within attributes. These constraints were based on *a priori* desirability orders for the levels of attributes and ensured that partworth estimates are consistent with these *a priori* orders. If *a priori* orders are valid then the imposition of constraints will bring the partworth estimates closer to their true value and this will improve the predictive validity of these estimates. For instance, Srinivasan *et al.* assumed that consumer preferences for checking accounts offered by banks can only increase with the hours of operation. Compared to unrestricted partworth estimates, they obtained better predictions with restricted partworth estimates for the full-profile method but not for the tradeoff matrix approach. (Predicative validity was defined at the level of an individual respondent in two ways: *i*) the ability to predict the rank order of preferences for a set of holdout stimuli, and *ii*) the ability to predict the most preferred stimulus in that set.)

As an alternative to *a priori* desirability orders of attribute levels that are specified by the researchers, Srinivasan *et al.* suggested that self-explicated desirability orders could also be used to impose within-attribute constraints. This suggestion was followed and extended by van der Lans and Heiser. They derived both within- and across-attribute constraints using the self-explicated utility model (Huber, 1974). The self-explicated utility model implies that, for instance, the partworth $h_{(40,\text{hours of operation})}$ for 40 hours of operation per week is the product of the desirability value $e_{(40,\text{hours of operation})}$ for that level and the importance value $v_{(\text{hours of operation})}$ for that attribute, or:

$$h_{(40,\text{hours of operation})} = v_{(\text{hours of operation})}\, e_{(40,\text{hours of operation})} \quad (1)$$

More generally, the self-explicated utility model implies that for any level *i* of attribute *j*:

$$h_{ij} = v_{ij}\, e_{ij} \quad (2)$$

Within- and across-attribute constraints (constraints consisted of inequality and/or equality constraints) upon partworth estimates were obtained by: *i*) within- and across-attribute constraints among estimates for the e_{ij}'s and *ii*) across-attribute constraints among estimates for the v_j's.

Constraints among estimates for the e_{ij}'s were based upon self-explicated attribute level desirability ratings, and constraints among estimates for the v_j's were based upon self-explicated attribute importance ratings. Again, if the self-explicated utility model and the *a priori* orders from self-explicated desirability and self-explicated importance ratings are valid, then the imposition of constraints will bring the partworth estimates closer to their true values and this will improve the predictive validity. Van der Lans and Heiser compared the predictive validities of *i*) partworth estimates under within-attribute constraints, based upon the self-explicated desirability ratings only, *ii*) estimates under within- and across-attribute constraints, and *iii)* unconstrained estimates, for the full-profile method. Like Srinivasan *et al.* (1983), they obtained better predictions for constrained than for unconstrained estimates. Also, the addition of across-attribute constraints resulted in better or equal predictive validities compared with using only within-attribute constraints. (Predictive validities were measured both at the individual level, in terms of cross-validated Pearson correlations and correct first-choice predictions for a set of holdout stimuli, and at the aggregate level, based on mean absolute errors in first-choice share predictions.)

In this paper we derive respondent-specific within- and across-attribute constraints from the self-explicated data in ACA (ACA System by Sawtooth Software). Using both ACA and full-profile data collected by Huber *et al.* (1991), three sets of partworth estimates—unconstrained, within-attribute constrained, and within- and across-attribute constrained—are compared with respect to their predictive validity. Although there has recently been a lot of interest in the comparative predictive validity of ACA and full profile (see Agarwal and Green, 1991; Finkbeiner and Platz, 1986; Green *et al.*, 1991; Green and Srinivasan, 1990; Huber *et al.*, 1991, 1993; Johnson, 1991), we are not aware of studies comparing the two under constrained estimation. In addition to comparing predictive validities, we also determine the effects of the constraints on the existence/magnitude of a number of levels effect in derived attribute importances (Wittink *et al.*, 1991). We present conclusions and end with suggestions for ways in which conjoint analysis can be improved and with suggestions for further research.

STRUCTURE OF THE WITHIN- AND ACROSS-ATTRIBUTE CONSTRAINTS

In a typical ACA interview, respondents are first asked to rank order the levels of each attribute separately according to desirability. If the researcher assumes an *a priori* desirability order among the levels for all respondents, this rank order is not obtained from the respondents. Subsequently, individuals are asked to give importance ratings of the *differences* between the most desirable and least desirable levels per attribute on a four-point rating scale. From the self-explicated desirability orders, we derive inequality constraints among partworth estimates for levels within each attribute, as suggested by Srinivasan *et al.* These are the within-attribute constraints.

From the self-explicated importances, we derive inequality constraints among *differences* between partworth estimates for the most desirable and the least desirable levels per attribute. These are the across-attribute constraints. An example may serve to clarify the structure of constraints. Suppose that a respondent gave the following desirability orders (from most desirable to least desirable): ATTRIBUTE 1 level 4, level 3, level 2, level 1), ATTRIBUTE 2 (level 3, level 2, level 1), and ATTRIBUTE 3 (level 3, level 2, level 1). The within-attribute constraints on the partworth estimates are then given in Figure 1, Plot a. In Figure 1 $h_{41} \geq h_{31} \geq h_{21} \geq h_{11}$, shown in the first column, means that the partworths for attribute 1 are constrained to be consistent with this indicated pattern of inequalities. Thus, if attribute 1 is the price of a product, the constraints prevent a higher price from having a higher partworth than what occurs for a lower price.

$$
\begin{array}{ccc}
\multicolumn{3}{c}{\text{Attribute}} \\
1 & 2 & 3 \\
h_{41} & h_{32} & h_{33} \\
\rotatebox{-90}{$\geq$} & & \\
 & \rotatebox{-90}{$\geq$} & \rotatebox{-90}{$\geq$} \\
h_{31} & & \\
\rotatebox{-90}{$\geq$} & h_{22} & h_{23} \\
h_{21} & & \\
 & \rotatebox{-90}{$\geq$} & \rotatebox{-90}{$\geq$} \\
\rotatebox{-90}{$\geq$} & & \\
h_{11} & h_{12} & h_{13}
\end{array}
\qquad\qquad
\begin{array}{ccccc}
\multicolumn{5}{c}{\text{Attribute}} \\
1 & & 2 & & 3 \\
h_{41} & \geq & h_{32} & \geq & h_{33} \\
\rotatebox{-90}{$\geq$} & & & & \\
 & & \rotatebox{-90}{$\geq$} & & \rotatebox{-90}{$\geq$} \\
h_{31} & & & & \\
\rotatebox{-90}{$\geq$} & & h_{22} & & h_{23} \\
h_{21} & & & & \\
 & & \rotatebox{-90}{$\geq$} & & \rotatebox{-90}{$\geq$} \\
\rotatebox{-90}{$\geq$} & & & & \\
h_{11} & \geq & h_{12} & \geq & h_{13}
\end{array}
$$

Plot a Within-attribute constraints **Plot b** Within-and across-attribute constraints

Figure 1
Example of the structure of constraints

If in addition to these desirability orders, the respondent gave importance ratings of "1" to ATTRIBUTE 1, of "2" to ATTRIBUTE 2, and of "3" to ATTRIBUTE 3, we define the within- and across-attribute constraints depicted in Figure 1, Plot b. In plot b the row constraints prevent the partworths of the extreme (best and worst) levels from being inconsistent with the self-explicated (or *a priori* assumed) importances. If the self-explicated importances are equal for two attributes, no constraints are imposed on the conjoint-based importances.

The across-attribute constraints upon differences between partworths of the extreme levels correspond to stretching desirability values per attribute to ensure equal ranges before multiplying them in the self-explicated utility model with importance values as recommended by Srinivasan (1988), but advised against by Green *et al.* (1991). It is important to note that we are forced to impose constraints upon differences between products of desirability values and importance values, due to the incommensurability of the desirability responses across attributes. It is also important to note that the structure of within- and across-attribute constraints we propose is different from the one imposed by van der Lans and Heiser, who do not stretch desirability values per attribute before multiplying them with importance values.

Clearly, whether or not constrained estimation (under within-attribute constraints or within- and across-attribute constraints) will improve predictions depends upon the validity of the desirability orders and the order of the importance ratings.

Interestingly, the proposed across-attribute constraints are also likely to reduce the number of levels effect in conjoint analysis (Currim *et al.,* 1981; Green and Srinivasan; Wittink *et al.*, 1982, 1989, 1991). This effect implies that when intermediate levels are added to the levels of an attribute in a conjoint design, this attribute will get a higher relative importance than when these intermediate levels would not have been added, other things being equal. Wittink et al. (1991) found a significant number of levels effect with both the full-profile method and ACA, the effect size for the full-profile method being about twice as large as for ACA. However, they did not find a significant number of levels effect for self-explicated importances. Thus, the across-attribute constraints would not be dependent upon the number of levels and any effects in the conjoint-based importances will be reduced by these constraints. ACA already incorporates the self-explicated data in the estimation of partworths. This is one reason why the number of levels effect is larger for the full-profile method than for ACA, and we may therefore expect a larger reduction in the level effect for the full-profile method, by imposing across-attribute constraints, than for ACA. For a more detailed discussion of the levels effect, see Wittink *et al.* (1991).

METHOD

To compare the predictive validity of unconstrained, within-attribute constrained, and within- and across-attribute constrained partworth estimates, for both ACA and the full-profile method the data from Huber *et al.* (1991) are used. These data consist of responses of 400 respondents to refrigerators, refrigerator attributes, and refrigerator attribute levels. The attributes and attribute levels used are given in Table 1.

Table 1 Refrigerator attributes and levels

A. Brand Name	General Electric; Sears/Kenmore; Whirlpool
B. Capacity	19; 20*; 21*; 22 (cubic feet)
C. Energy Cost	$70; $80*; $90*; $100 (annual)
D. Compressor	Extremely quiet; somewhat quiet*; somewhat noisy*; extremely noisy
E. Price	$700; $850*; $1,000*; $1,150
F. Design	Freezer on left (side by side); Freezer on top
G. Warranty	1 year; 3 years
H. Refrigerant	Soft CFC (environmentally safe); Chlorofluoro-hydrocarbon (hurts environment)
I. Dispenser	Dispenses ice and water through the door; No door dispenser for ice or water

* In full profile, half the respondents saw all four levels for attributes B (Capacity) and E (Price), and only the extreme levels for attributes C (Energy Cost) and D (Compressor). The other half saw four levels for C and D, but only the extremes for B and E. In ACA, attribute D and E were each given at four levels for half the respondents, and attributes B and C had four levels for the other half.

Respondents were interviewed at super-regional malls in 11 cities. Prior to actual interviewing they were screened for being over 18 and having refrigerators in their homes, and promised $5 each for completing the interview. Each respondent provided both full-profile ratings (on a 9-point likelihood to purchase scale) and ACA judgments (self-explicated data as before, and graded paired comparisons on a 9-point scale). Half the respondents completed the full-profile task first whereas the other half completed the ACA task first. Both tasks were administered by computer in one sitting. In the ACA task, within-attribute desirability orders for levels of capacity, energy cost, compressor, and price, were assumed to be known a *priori* and to be common across respondents. Thus, no desirability orders were asked for these attributes.

Other experimental manipulations in Huber et al. (1991) are: 1) half the respondents saw nine attributes and half the respondents saw five attributes (A,B,C,D,E), 2) the number of levels was varied for attributes B, C, D, and E, as indicated in the footnote of Table 1, and 3) the order in which attributes were listed in the full-profile method was varied. Respondents who saw nine attributes gave 16 full-profile ratings and 16 graded paired comparison ratings (ten pairs differing on two attributes and six pairs differing on three attributes). Respondents who saw five attributes gave 16 full-profile ratings and 12 graded paired comparison ratings (ten pairs differing on two attributes and two pairs differing on three attributes). The order in which attributes appeared in the self-explicated part of the ACA task was not manipulated, but was equal to one of the orders of the attributes in the full-profile task. In ACA's graded paired comparisons task, the order in which attributes were listed varied between pairs of profiles.

In addition to the full-profile task and the ACA task, respondents were asked to complete a choice task in which they indicated their most preferred alternative in two pairs and two triples of refrigerator profiles. Respondents were also asked to indicate their least preferred alternative in the two triples. The profiles were defined in terms of the five attributes and the 13 levels that were common to all respondents in the full-profile and ACA tasks. Profiles within each pairltriple were chosen such that it could be expected that no profiles were dominated on all attributes by other profiles. The choice task was administered twice, once after the first preference elicitation task and again after the second preference elicitation task.

For each of 385 respondents (7 respondents provided incomplete data and another 8 gave equal ratings to all full profiles andlor pairs) we computed nine sets of least squares partworth estimates. Three sets were based on the full-profile ratings, three on both the ACA priors and the graded paired comparisons, and three on the graded paired comparisons only. The latter three sets were included to compare the way in which ACA uses the self-explicated data with a theoretically more justifiable way of using only the order information in the self-explicated data. We used the MORALS-algorithm (Young, De Leeuw and Takane, 1976) to compute unconstrained partworth estimates, and within- attribute constrained partworth estimates for the full-profile method. All other partworth estimates were computed by an alternating least squares procedure given by van der Lans (1992, Chapter 3; see also van der Lans, 1991). To reduce the likelihood of solutions that correspond to a local minimum, we computed solutions from multiple starts and retained the best ones. The implementation of constrained estimation requires additional considerations. Two or more attributes could obtain the same importance rating. These ties were treated by the so-called primary approach to ties (Kruskal, 1964). That is, differences between partworths of extreme levels were not constrained to be equal in the final conjoint solution for attributes that obtained equal self-explicated importances. The primary approach to ties was chosen because the coarse-grained importance rating scale could be expected to result in a large number of ties, and van der Lans and Heiser found better predictive validity for the primary approach to ties even with a finer-grained importance scale.

The different sets of partworth estimates were used: 1) to predict individual choices in the (replicated) choice task, 2) to predict aggregate choice shares via the first choice rule and the multinomial logit rule, and 3) to investigate the number of attribute levels effect.

RESULTS

Individual-Level Choice Prediction

In Table 2 we show proportions of correctly predicted individual choices. For computing the entries in Table 2, each triple was converted to three pairs (except for the last subtable). Thus, with two pairs and two triples each evaluated twice we have sixteen pairs per respondent. The successive columns show the proportions of correctly predicted individual choices (hit rates) for unconstrained, within-attribute constrained, and within- and across- constrained partworth estimates. The last column gives the highest proportion across the three sets of estimates. The highest proportion (based on three digits precision) within each row is underlined.

Part A of Table 2 gives the overall proportions. In the rightmost column, we see that the best set of estimates for each method yields approximately equal proportions correctly predicted choices. However, without constraints ACA provides the highest predictive validity. Its validity does not improve when constraints are imposed. On the other hand, Full Profile and Paired Comparisons Only do gain from constrained estimation under both within-attribute constraints and within- and across-attribute constraints. The low proportion for unconstrained partworth estimates from Paired Comparisons Only is due to the nonorthogonality of the paired comparisons design which results in a large variance of the partworth estimates. This variance is reduced by the imposition of constraints.

In Part B of Table 2 hit rates are decomposed according to the number of attributes that the respondents saw. Looking at the highest hit rate within each row, we see that, somewhat surprisingly, ACA and Paired Comparisons Only do better than Full Profile with five attributes, but not with nine. This seems to contradict (at least up to about ten attributes) the suggestion by Huber *et al.* (1991) that ACA should become increasingly attractive as the number of attributes to be included in the study increases. However, they did not consider the benefit of imposing constraints. Without constraints, ACA is superior for both five and nine attributes. A tentative explanation for the result would be that ACA's graded paired comparisons design matrices were more efficient for estimating the tradeoff between the attributes involved in the choice sets with five attributes than with nine attributes. This might also explain why with five attributes the predictive validities of ACA and Paired Comparisons Only become slightly worse under within- and across-attribute constraints compared to no constraints and within-attribute constraints. Unlike the case with nine attributes, for five attributes the tradeoff between attributes may already be estimated precisely, and the constraints (which contain some error) only worsen the predictive validity. On the other hand, for Full Profile with nine attributes, compared to with five attributes, the effect of *i*) this increased task complexity, and *ii*) the lower ratio of observations to parameters, on the predictive validity, seems to be counteracted by the imposition of constraints. Furthermore, with nine attributes within- and across-attribute constraints do better than within-attribute constraints.

Table 2 Individual-Level Choice Predictions (Proportion Predicted Correctly)

	No Constraints	Within-Attribute Constraints	Within- and Across Attribute	Highest
<u>A: Overall</u>				
Full Profile	.68	.71	.<u>72</u>	.72
ACA	.<u>73</u>	.<u>73</u>	.72	.73
Paired Comparisons Only	.66	.<u>73</u>	.72	.73
<u>B: Number of Attributes</u>				
		Nine Attributes		
Full Profile	.65	.69	.<u>72</u>	.72
ACA	.71	.<u>71</u>	.70	.71
Paired Comparisons Only	.61	.70	.<u>71</u>	.71
		Five Attributes		
Full Profile	.71	.<u>72</u>	.<u>72</u>	.72
ACA	.<u>76</u>	.76	.74	.76
Paired Comparisons Only	.70	.<u>75</u>	.73	.75
<u>C: Consistency</u>				
		≤4		
Full Profile	.<u>57</u>	.56	.56	.57
ACA	.59	.<u>60</u>	.59	.60
Paired Comparisons Only	.57	.59	.<u>60</u>	.60
		5 or 6		
Full Profile	.59	.64	.<u>66</u>	.66
ACA	.<u>69</u>	.69	.68	.69
Paired Comparisons Only	.62	.<u>69</u>	.68	.69
		7		
Full Profile	.69	.71	.<u>72</u>	.72
ACA	.<u>74</u>	.73	.72	.74
Paired Comparisons Only	.67	.72	.<u>72</u>	.72
		8		
Full Profile	.78	.81	.<u>83</u>	.83
ACA	.82	.<u>82</u>	.80	.82
Paired Comparisons Only	.71	.<u>82</u>	.79	.82

Part C of Table 2 decomposes the hit rates according to the respondents' consistency in the replicated choice task. Each respondent chose from two pairs and two triples twice. By decomposing the choices with each triple into three pairs, this yielded eight replicated choices. Both ACA and Paired Comparisons Only do better than Full Profile when the number of consistent choices is less than 4, or 5 or 6. With eight consistent choices, the difference between Full Profile, ACA, and Paired Comparisons Only completely disappears. Apparently, respondents who are not perfectly consistent in their choices gain most from the simplicity of the ACA task. Constrained estimates do not improve over unconstrained estimates for Full Profile when respondents are very inconsistent ($\leq$4). With more than 4 consistent choices constrained estimates do improve for Full Profile. In that case, within- and across-attribute constraints also improve over within-attribute constraints.

Hit rates were also decomposed on the combination of task order (ACA first or Full Profile first) and number of attributes (see Table 2, Part B). Apart from the main effect of the number of attributes, two interesting results were found. First, within-constrained partworth estimates always improve upon unconstrained partworth estimates (difference of about five percent correctly predicted choices) for Full Profile except when ACA comes first and respondents see only five attributes. Apparently, ACA serves as a warm-up by which respondents become able to judge full profiles based on five attributes in a manner that is consistent with the desirability orders on the attribute levels. Secondly, within- and across-attribute constraints improve upon within-attribute constraints for Full Profile only when Full Profile comes first and the number of attributes is nine (difference of four percent correctly predicted choices). Apparently, only when no warm-up task precedes the full-profile judgments and the number of attributes is relatively high do the respondents seem to be less accurate in making the tradeoff between attributes in their full-profile judgments. This latter result differs from the result in van der Lans and Heiser, where within- and across-constraints improve upon within-constraints for Full Profile, given six attributes and given that full-profile judgments were preceded by self-explicated desirability and importance ratings.

In Part A of Table 3 we distinguish between the proportions of correctly predicted choices for pairs and triples. As suggested by Huber *et al.* (1991), two reasons for examining the triples separately are *i*) the fact that real world choices are usually not restricted to pairs, and *ii*) ACA's paired comparisons intensity ratings may be more similar to the pairs in the choice tasks than the triples. We see that the large difference in hit rates between Full Profile and ACA for the pairs from triples found by Huber *et al.* (1991) (with unconstrained partworth estimates), becomes very small when partworth estimates for Full Profile are constrained. And, except for when there are no constraints, hit rates for Paired Comparisons Only are virtually equal to the hit rates for ACA.

Table 3 Individual-Level Choice Predictions Split by Type of Choice (Proportion Predicted Correctly)

	No Constraints	**Within-Attribute Constraints**	**Within- and Across-Attribute Constraints**	**Highest**
A: Choice Task				
		Pairs from Triples		
Full Profile	0.67	0.71	0.72	0.72
ACA 0.74	0.74	0.72	0.74	
Paired Comparisons Only	0.65	0.73	0.72	0.73
		Pairs from Pairs		
Full Profile	0.70	0.70	0.71	0.71
ACA 0.72	0.72	0.71	0.72	
Paired Comparisons Only	0.66	0.72	0.71	0.72
B: Triples				
		Most Likely		
Full Profile	0.52	0.59	0.61	0.61
ACA	0.64	0.65	0.62	0.65
Paired Comparisons Only	0.53	0.63	0.61	0.63
		Least Likely		
Full Profile	0.57	0.60	0.61	0.61
ACA	0.63	0.63	0.61	0.63
Paired Comparisons Only	0.52	0.62	0.61	0.62

Part B of Table 3 shows the results for the choices from triples separately for choosing the "most likely" and the "least likely" to-be-purchased profile from the triple. Note that the reduced percentages for most and least likely choices (relative to the pairs) simply result from the greater difficulty in predicting choices from triples than from pairs. Differences between the three methods are larger for most likely choices than for least likely choices. Of course, the most likely choices are the ones for which predictions are most critical. Constrained partworth estimates for Full Profile give the largest improvement compared to unconstrained partworth estimates for most likely choices.

Aggregate-Level Choice Share Prediction

In Table 4 mean absolute errors in choice share predictions are given for each method under each type of constraint. Predicted and actual choice shares were computed for each choice (from two pairs and two triples, that is, for a total of ten profiles) across respondents and across the replication of the choice task. Predicted choices were determined via the first choice rule and the multinomial logit rule. For the multinomial logit rule, slope parameters were computed first by maximizing the log likelihood of the multinomial logit model across respondents, choice sets, and the replication. In the multinomial logit model, the chance $p_{(k)is}$ that respondent *k* will choose profile *i* from choice set *s* is given by:

$$(3) \qquad p_{(k)is} = \frac{e^{bV_{(k)i}}}{\sum_{jes} e^{bV_{(k)j'}}}$$

in which $V_{(k)i}$ gives the predicted overall utility that profile *i* has for respondent *k*. One benefit of the multinomial logit rule is that it adds random error to the predicted overall utilities which may improve choice share predictions when the noise in the choices is greater than the noise in the predicted conjoint-based overall utilities for the profiles (Elrod and Kumar, 1989).

Looking at Table 4 we see that Full Profile within-attribute constrained partworth estimates with the first choice rule give by far the lowest mean absolute error in predicted choice share. Interestingly, the absolute error in predicting choice shares increases when random error is added for Full Profile, whereas for ACA it decreases.

This result can be explained by using the following argument: It seems that overall utilities predicted from Full Profile contain more noise than the choices. Adding random error to the predicted utilities can only increase the mean absolute error in predicted choice shares. Furthermore, it seems that for Full Profile, within attribute-constraints reduce the noise in the predicted utilities to the level of the noise in the choices, and within- and across-attribute constraints reduce the noise in the predicted overall utilities even further (below the level of noise in the choices). On the other hand, the overall utilities predicted from ACA seem to contain less noise than the choices. The fact that adding random error does not lower the mean error further than 5.6 may be due to systematic differences in ACA judgments and holdout choices. Clearly, systematic differences cannot be eliminated by adding random error. Adding constraints in ACA reduces the noise even further below the noise present in the choice data.

Note that this explanation, and the somewhat counterintuitive finding that the method that does better at individual-level choice prediction (ACA) does worse at aggregate choice share predictions, is consistent with Hagerty's (1986) results. His results imply that for individual-level predictions noise in partworth estimates is quite important, whereas for aggregate-level predictions the noise tends to cancel itself out across respondents. Results for Paired Comparisons Only seem somewhat erratic. However, they can be explained if unconstrained partworth estimates for Paired Comparisons contain less noise than unconstrained partworth estimates for Full Profile.

Table 4 Mean Absolute Error in Aggregate-Level Choice Share Prediction (%)

	No Constraints	**Within-Attribute Constraints**	**Within- and Across-Attribute Constraints**	**Lowest**
First Choice Rule				
Full Profile	5.3	<u>1.7</u>	3.8	**1.7**
ACA	<u>10.4</u>	11.5	12.2	**10.4**
Paired Comparisons Only	<u>4.1</u>	10.0	11.3	**4.1**
Multinomial Logit Model				
Full Profile	10.0	6.7	<u>6.4</u>	**6.4**
ACA	<u>5.6</u>	5.9	6.2	**5.6**
Paired Comparisons Only	12.5	<u>6.9</u>	14.4	**6.9**

<u>Number of Attribute Levels Effect</u>

The plots in Figure 2 show the number of attribute levels effect for the attributes whose numbers of levels were manipulated. As in Wittink *et al.* (1991), importances were first normalized per respondent such that the importances of the five common attributes sum to 100. We see that the different sets of partworth estimates yield different effect sizes. Unconstrained partworth estimates for Full Profile and Paired Comparisons Only display the largest number of level effects throughout. ACA shows much smaller effects.

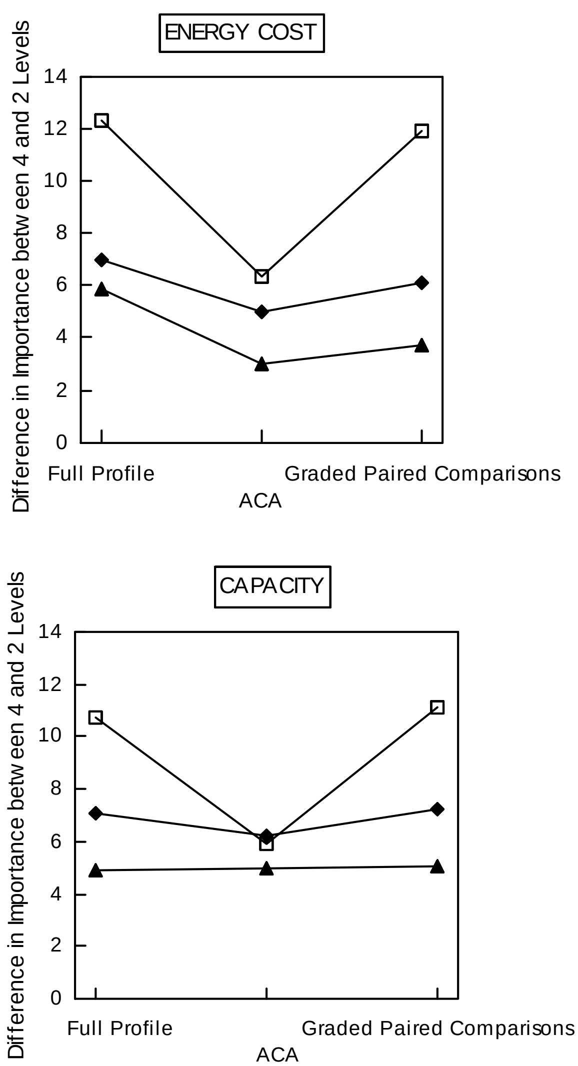
Multivariate analyses of variance were run to test the significance of the two-way interactions between number of levels and type of constraints, and between number of levels and method, as well as the three-way interaction, upon average importance. The number of levels x type of constraints interaction has a significant effect for all attributes ($p<.01$). The number of levels x method interaction has a significant effect for energy ($p<.05$) and price ($p<.05$). The three-way interaction is significant for energy ($p<.05$), capacity ($p<.01$), and price ($p<.01$), but not for compressor noise.

The large effects for unconstrained partworth estimates in Full Profile and Paired Comparisons Only are reduced considerably by imposing within-attribute constraints, except for compressor noise and price with Full Profile. Thus it seems that the orders of compressor noise levels and price levels are less often violated than the orders of energy cost levels and capacity levels.

Apparently, price is not confounded with quality as Srinivasan *et al.* (1983) warned. Adding across-attribute constraints reduces the number of attribute levels effect about equally for all three methods.

Figure 2
Number of attribute level effects for different attributes, methods and constraints

■ No Constraints
◆ Within-Attribute Constraints
∆ Within- and Between-Attribute Constraints



CONCLUSIONS

The differences in predictive validity for unconstrained ACA and Full Profile conjoint analysis are substantially reduced when constraints are imposed on the partworths. For individual-level choice predictions, Full Profile, ACA, and Paired Comparisons Only are about equally valid when within-and across-attribute constrained partworth estimates are used for Full Profile; unconstrained or within-constrained partworth estimates are used for ACA; and within-constrained partworth estimates are used for Paired Comparisons Only. ACA and Paired Comparisons Only outperform Full Profile when *i*) the Full Profile task comes first and the number of attributes is five, and *ii*) when respondents have low consistency in their choices. The almost equal predictive validities of ACA and Paired Comparisons Only seem to indicate that the way in which ACA combines the self-explicated priors with the paired comparisons data accomplishes the same result as the theoretically more justifiable idea of imposing constraints (at least for individual-level predictions).

For aggregate choice share predictions we have used both the first choice rule and the multinomial logit model. Although the differences between the two choice rules, the different conjoint analysis methods, and the different constraints can be explained by results from Elrod and Kumar and Hagerty, we think that our results should be interpreted carefully until more evidence favoring either of the two choice rules is found. Nevertheless, the much higher mean absolute error in choice share prediction for Full Profile compared to ACA when using unconstrained partworth estimates (see, Huber *et al.*, 1993) is considerably reduced when imposing constraints upon the partworth estimates for Full Profile. The difference between the effect of constraints in Full Profile and Paired Comparisons Only might be due to the fact that the design matrix for the paired comparisons is nonorthogonal. As such, imposing constraints upon some partworth estimates may alter other partworth estimates for levels that are not involved in the constraints. It is conceivable that this causes a deterioration in the predictive validity.

As expected, the imposition of constraints reduces the number of levels effect considerably for Full Profile and Paired Comparisons Only, and only slightly for ACA.

Further research could consider the use of finer-grained importance scales and commensurable desirability scales across attributes as suggested by Green, Krieger, and Agarwal. Finer-grained importance scales seem especially desirable because of the finding of van der Lans and Heiser that the primary approach to ties does better than the secondary approach, even with very fine-grained importance scales.

REFERENCES

Agarwal, Manoj K. and Paul E. Green (1991). "Adaptive Conjoint Analysis Versus Self-Explicated Models: Some Empirical Results." *International Journal of Research in Marketing,* 8, 141-146.

Currim, Imran S., Charles B. Weinberg and Dick R. Wittink (1981). "The Design of Subscription Programs for a Performing Arts Series." *Journal of Consumer Research,* 8 (June), 67-75.

De Leeuw, Jan, Forest W. Young and Yoshio Takane (1976). "Additive Structure in Qualitative Data: An Alternating Least Squares Method with Optimal Scaling Features." *Psychometrika*, 41 (December), 505-29.

Elrod, Terry and S. Krishna Kumar (1989). "Bias in the First Choice Rule for Predicting Share." *Sawtooth Software Conference Proceedings*, 259-271.

Finkbeiner, Carl T. and Patricia J. Platz (1986). "Computerized Versus Paper and Pencil Methods: A Comparison Study." Paper presented at the Association for Consumer Research Conference, Toronto, Canada, October.

Green, Paul E., Abba M. Krieger and Manoj K. Agarwal (1991). "Adaptive Conjoint Analysis: Some Caveats and Suggestions." *Journal of Marketing Research,* 28 (May), 215-22.

Green, Paul E. and Catherine M. Schaffer (1991). "Importance Weight Effects on Self-Explicated Preference Models: Some Empirical Findings." *Advances in Consumer Research,* 18, Association for Consumer Research.

Green, Paul E. and V. Srinivasan (1990). "Conjoint Analysis in Marketing Research: New Developments and Directions." *Journal of Marketing,* 54 (October), 3-19.

Hagerty, Michael R. (1986). "The Cost of Simplifying Preference Models." *Marketing Science*, 5, 298 - 319.

Huber, George P. (1974). "Multi-Attribute Utility Models: A Review of Field and Field-like Studies." *Management Science,* 20,1393-1402.

Huber, Joel C., Dick R. Wittink, John A. Fiedler, and Richard L. Miller (1991). "An Empirical Comparison of ACA and Full Profile Judgments." *Sawtooth Software Conference Proceedings,* 189-202.

Huber, Joel C., Dick R. Wittink, John A. Fiedler, and Richard L. Miller (1993). "The Effectiveness of Alternative Preference Elicitation Procedures in Predicting Choice." *Journal of Marketing Research,* forthcoming.

Johnson, Richard M. (1991). "Comment on "Adaptive Conjoint Analysis: Some Caveats and Suggestions." *Journal of Marketing Research,* 28 (May), 223-5.

Kruskal, Joseph B. (1964). "Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis." *Psychometrika,* 29 (March),1-27.

Srinivasan, V. (1988). "A Conjunctive-Compensatory Approach to the Self-Explication of Multiattributed Preferences." *Decision Sciences,* 19, 295-305.

Srinivasan, V., Arun K. Jain and Naresh K. Malhotra (1983). "Improving Predictive Power of Conjoint Analysis by Constrained Parameter Estimation." *Journal of Marketing Research, 20* (November), 433-8.

van der Lans, Ivo A. (1991). "Optimal Scaling under Order Restrictions upon Parameters and (Differences between) Products of Parameters." paper presented at the Third Conference of the International Federation of Classification Societies, Edinburgh, Scotland, August.

van der Lans, Ivo A. (1992, in preparation). "Nonlinear Multivariate Analysis for Multiattribute Preference Data." Dissertation, University of Leiden, The Netherlands.

van der Lans, lvo A. and Willem J. Heiser (1990). "Constrained Part-Worth Estimation in Conjoint Analysis using the Self-Explicated Utility Model." Research Report RR 90-02, Department of Data Theory, University of Leiden, The Netherlands.

Wittink, Dick R., Joel C. Huber, John A. Fiedler and Richard L. Miller (1991). "The Magnitude of and an Explanation/Solution for the Number of Levels Effect in Conjoint Analysis." Working Paper, Cornell University at Ithaca, NY, August.

Wittink, Dick R., Lakshman Krishnamurthi and Julia B. Nutter (1982). "Comparing Derived Importance Weights across Attributes." *Journal of Consumer Research*, 8 (March), 471-4.

Wittink, Dick R., Lakshman Krishnamurthi and David J. Reibstein (1989**).** The Effects of Differences in the Number of Attribute Levels on Conjoint Results." *Marketing Letters,* 1, 113-23.

Young, Forest W., Jan De Leeuw and Yoshio Takane (1976). "Regression with Qualitative and Quantitative Variables: An Alternating Least Squares Method with Optimal Scaling Features." *Psychometrika,* 41 (December), 505-29.