# Application and Test of Web-based Adaptive Polyhedral Conjoint Analysis

Ely Dahan

John R. Hauser

Duncan Simester

Olivier Toubia

January 16, 2002

Ely Dahan is an Assistant Professor of Management Science, John Hauser is the Kirin Professor of Marketing, Duncan I. Simester is an Associate Professor of Management Science, Olivier Toubia is a Ph.D. candidate all at the Sloan School of Management, Massachusetts Institute of Technology, 38 Memorial Drive, Cambridge, MA 02142-1307.

The Sloan School of Management, the Center for Innovation in Product Development at MIT and the EBusiness Center at MIT all supported this research. This paper may be downloaded from http://mitsloan.mit.edu/vc. That website also contains (1) open source code to implement the methods described in this paper, (2) demonstrations of web-based questionnaires used to collect data for the tests described in this paper, and (3) related papers and demonstrations of web-based interviewing methods. The authors are listed in alphabetical order.

We wish to thank Limor Weisberg for creating the graphics used on the web-based questionnaire and Rob Hardy, Adler Kim, and Bryant Lee who assisted in the development of the web-based questionnaire. We gratefully acknowledge the contribution of Robert M. Freund who proposed the use of the analytic center and approximating ellipsoids for the math programs required for question selection and partworth estimation. The questionnaire benefited from comments by numerous pretesters at the MIT Operations Research Center, Analysis Group Economics, Applied Marketing Science, the Center for Innovation in Product Development, and Dražen Prelec's market measurement class. Our thanks to Brennan Mulligan of Timbuk2 Designs, Inc. for supplying the laptop computer bags used in our experiments.

## Application and Test of Web-based Adaptive Polyhedral Conjoint Analysis


## Abstract

In response to the need for more rapid and iterative feedback on customer preferences, researchers are developing new web-based conjoint analysis methods that adapt the design of conjoint questions based on a respondent's answers to previous questions. Adapting within a respondent is a difficult dynamic optimization problem and until recently adaptive conjoint analysis (ACA) was the dominant method available for addressing this adaptation. In this paper we apply and test a new polyhedral method that uses "interior-point" math programming techniques. This method is benchmarked against both ACA and an efficient non-adaptive design (Fixed).

Over 300 respondents were randomly assigned to different experimental conditions and were asked to complete a web-based conjoint exercise. The conditions varied based on the design of the conjoint exercise. Respondents in one group completed a conjoint exercise designed using the ACA method, respondents in another group completed an exercise designed using the Fixed method, and the remaining respondents completed an exercise designed using the polyhedral method. Following the conjoint exercise respondents were given $100 and allowed to make a purchase from a Pareto choice set of five new-to-the-market laptop computer bags. The respondents received their chosen bag together with the difference in cash between the price of their chosen bag and the $100.

We compare the methods on both internal and external validity. Internal validity is evaluated by comparing how well the different conjoint methods predict several holdout conjoint questions. External validity is evaluated by comparing how well the conjoint methods predict the respondents' selections from the choice sets of five bags.

The results reveal a remarkable level of consistency across the two validation tasks. The polyhedral method was consistently more accurate than both the ACA and Fixed methods. However, even better performance was achieved by combining (post hoc) different components of each method to create a range of hybrid methods. Additional analyses evaluate the robustness of the predictions and explore alternative estimation methods such as Hierarchical Bayes. At the time of the test, the bags were prototypes. Based, in part, on the results of this study these bags are now commercially available.

## Recent Developments

Preference measurement has a long history in marketing. Firms routinely rely upon various forms of conjoint analysis, self-stated importances, and voice-of-the-customer methods to design new products, expand product lines, and set segmentation strategies (e.g., Cattin and Wittink 1982; Green and Srinivasan 1990; Wittink and Cattin 1989). Until recently, preference measurement could safely be called a mature technology with a rich history of applications, improvements, simulations, and empirical tests. However, three recent developments have caused researchers to refocus on this fertile area of research.

The first development is a change in the product development (PD) process. Business-to-business and complex durable goods manufacturers are adopting consumer-packaged-good methods. Because these products are typically complex with many features, PD teams require preference measurement methods than can handle more customer needs and provide prioritizations among substantially more features (Eppinger 1998; Eppinger, et. al. 1994; Ulrich and Eppinger 2000). At the same time concurrent engineering, rapid web-based product design, and "spiral" processes view preference measurement as iterative (Cusumano and Selby 1995; McGrath 1996; Smith and Reinertsen 1998; Tessler, Wada and Klein 1993). PD teams routinely screen large numbers of features, then focus, remeasure, and iterate. Accuracy need only be sufficient to separate the winning features from the losing features, shifting the tradeoff toward lower research costs per iteration. PD teams are more willing to sacrifice some precision in the measurement of preferences in order to handle more features.

The second development is the advent of the Internet as a means to collect data. Web-based panels with large numbers of respondents provide the potential for very rapid and reliable feedback (Buckman 2000; Gonier 1999; Nadilo 1999; Willkie, Adams, and Girnius 1999). A conjoint analysis study can be fielded nationally or internationally in days. Furthermore, with broadband capabilities, virtual concepts can be created with vivid representations, animation, context-sensitive help, and sound (Dahan and Hauser 2002; Dahan and Srinivasan 2000). In addition, the ability for computation between questions has renewed interest in new forms of adaptive questions – questions selected and created for each customer based on his or her prior answers. The Internet has also brought a new set of limitations. In particular, web-based interviewing techniques must address reduced respondent tolerance for answering multiple questions. In central location interviewing, respondents are remarkably tolerant of multiple questions. However, the literature and commercial experience indicate much lower tolerance levels in a web context (e.g., De Angelis 2001). In our experience, interest and reliability drop off rapidly after eight to twelve repetitive questions. This creates greater demand for conjoint methods that gather more information with fewer questions.

The third development is the introduction of projective interior-point methods in math programming. Conjoint analysis has used math programming successfully for almost thirty years (e.g., Srinivasan and Shocker 1973a, 1973b), but its application to question selection is new. For large problems, optimization for individual-respondent question selection is a complex, difficult-to-solve problem, which, if solved by traditional methods, would lead to excessive delays between questions – a situation that is unacceptable to web-based respondents. The new algorithms search the interior of a parameter space and find near-optimal solutions extremely fast (Freund 1993; Kamarkar 1984; Nesterov and Nemiroskii 1994; Sonnevend 1985a, 1985b; Vaidja 1989). This represents a change in philosophy from a primarily statistical focus, to a focus on feasible heuristics. These methods do well in simulation, but have not yet been tested on real customers.

In this paper we evaluate the interior point method recently proposed by Toubia, Simester and Hauser (2002). We use a real product – a laptop computer bag worth approximately $100. After completing the various web-based conjoint questionnaires (assigned randomly), respondents in the study were given $100 to spend on a choice set of five bags. Respondents received their chosen bag together with the difference in cash between the price of their chosen bag and the $100. This data enables us to compare the methods on both internal and external validity. Internal validity is evaluated by comparing how well the methods predict several holdout conjoint questions. External validity is evaluated by comparing how well the different conjoint methods predict which bag respondents' later chose to purchase using their $100. We then explore whether there is a question-order effect, whether the performance of the "pure-bred" methods can be improved with hybrids that combine components from the different methods, and whether alternative estimation methods (such as Hierarchical Bayes) improve internal and external validity.

We begin with a brief review of the methods and then describe the research design. The core of the paper reports on and interprets the internal and external validity tests. We close with a summary and discussion of the findings.

## Three Conjoint Analysis Methods

Toubia, Simester and Hauser's (2002) FastPace method adapts questions for each respondent using that respondent's answers to previous questions. Adapting within a respondent is a difficult dynamic optimization problem that should be distinguished from techniques that adapt across respondents. The only other approach that we found to solve this problem is Sawtooth Software's adaptive conjoint analysis (ACA). In contrast, techniques such as the Arora and Huber (2001), Huber and Zwerina (1996), and Sandor and Wedel (2001) swapping/relabeling methods adapt conjoint designs <u>across</u> respondents.

We compare the so-called FastPace method with two benchmarks; ACA and an efficient non-adaptive (Fixed) method. We also explore hybrid techniques and improvements based on these data collection models. For this first empirical test, we focus on paired-comparison questions in which respondents provide a metric rating of the extent to which they prefer one product over the other (metric-paired-comparison questions) because this is the question format that ACA uses. This is a common format, well suited to the web, standard in the ACA method, and shown to provide interval-scaled preferences (Hauser and Shugan 1980; Johnson 1991; Mahajan and Wind 1992; Wittink and Cattin 1989). We hold the question format constant across methods, but recognize that there are many forms of data collection for conjoint analysis. Most fall into one of three categories – metric paired comparisons, choice from a set of profiles, and ranking or rating of a factorial design of profiles. Polyhedral methods can be devised for all three tasks (suggesting future empirical tests).

The methods vary in at least two respects: (1) the heuristics that they use to select the product profiles that respondents compare, and (2) the manner in which partworths associated with each feature are estimated. In the discussion that follows, we review the design of the product profiles and the basic estimation procedures for the three methods. As all three methods are described elsewhere, our review is brief (readers seeking additional details should consult the references). We consider alternative estimation methods in later sections.

**Fast Polyhedral Adaptive Conjoint Estimation (FP)**

Fast polyhedral adaptive conjoint estimation (FP) uses recent advances in interior point methods to *design* questions (product profiles) and *estimate* partworths.[1] Toubia, Simester and Hauser (2002) propose two versions to address metric paired-comparison questions and choices from a set of profiles. They test both versions using simulated respondents. Their methods interpret the problem of estimating partworths for *p* different product features as searching a *p*-dimensional space.[2] A point in this *p*-dimensional space represents one set of relative importances (partworths) for the features. The goal is to identify the point in the space that best represents the true partworths for an individual respondent.

When respondents answer a metric question comparing one or more product profiles, their answers are generally consistent with only a subset of the points in the *p*-dimensional space. By eliminating points that are inconsistent with the responses, the method is designed to converge to the point that represents the true set of relative importances. Responses are more informative if they allow the researcher to eliminate a larger subset of the points, thus FP selects questions to quickly reduce the set of consistent (feasible) points. The FP method approximates the current feasible (polyhedral) region using ellipsoids. By selecting question vectors that are perpendicular to the longest axis of the ellipsoids, the feasible region shrinks rapidly (and inconsistent responses are less likely). We refer interested readers to their paper, where the authors describe the details and rationale for this approach, including methods for model-

[1] Acronyms are used throughout the paper to simplify exposition. A complete summary of the acronyms is provided in the Appendix.
[2] FP allows the initial space to be constrained by prior information or other constraints. For features with more than two levels or for features that interact, FP simply expands the dimensionality of the space accordingly.

ing respondent errors. (Their paper and open-source code are available at the website listed in the acknowledgments section of their paper.)

The managerial problem imposed by the recent changes in product development requires estimates of partworths with relatively few questions – often fewer questions than there are features. This means that the researcher must often estimate the respondent's revealed partworths even though there remains a (hopefully small) region of feasible partworths that are consistent with the respondent's answers to the chosen questions. Based on research that justifies centrality as a criterion (Dawes and Corrigan 1974; Einhorn 1971; Huber 1975; Moore and Semenik 1988; Srinivasan and Park 1997; Srinivasan and Shocker 1973a), FP uses an estimate that characterizes the set of remaining feasible points by the center of gravity of the set. Because it is computationally difficult to identify the center of gravity of a polyhedron, they use the analytic center, which provides an excellent approximation. The analytical center can be estimated with effectively no computational delay between questions.

Tests conducted with simulated respondents suggest that the algorithm offers improved accuracy over ACA and an efficient Fixed design as long as there are many features relative to the number of questions. When the number of questions is sufficiently high all three methods have similar performance. Toubia, Simester and Hauser (2002) also consider a range of hybrid methods and show that they have the potential to improve performance.

**Adaptive Conjoint Analysis (ACA)**

Prior to FP, Sawtooth Software's Adaptive Conjoint Analysis (ACA) was the dominant approach for adapting questions within a respondent. It has become the industry standard for conjoint analysis (Green, Krieger and Agarwal 1991, p. 215).[3] Academic and commercial experience suggest that ACA is often reliable, accurate, and managerially valuable.

The ACA interview consists of four sequential tasks (Klein 1988; Sawtooth 1996). Respondents (1) eliminate unacceptable levels, (2) state the relative importances of improving a product from one feature level to another (the "self-explicated" questions), (3) respond to metric paired-comparison questions, and (4) evaluate a small set of full product profiles on a purchase-intention scale. The "adaptive" task is the third task; the first task is often skipped. The ACA algorithm uses information from a respondent's earlier responses to provide current estimates of a respondent's preferences and then designs product profiles that are closest in estimated preference, subject to constraints that ensure the overall design is nearly balanced and orthogonal. Both the intermediate estimates and the final estimates are based on a modified regression that combines the information from the self-explicated (SE) and the metric paired-comparison

[3] See also: Carroll and Green 1995; Choi and DeSarbo 1994; Green and Krieger 1995; Green, Krieger and Agarwal 1991; Huber, Wittink, Fiedler and Miller 1993; Johnson 1987; Klein 1988; Orme 1999; Wittink and Cattin 1989.

(PC) data. The intermediate regression has $p+q$ data points – one for each of $p$ SE responses and one for each of $q$ PC responses. Because the regression always has at least $p$ observations, ACA can provide intermediate estimates even after a single PC question. Once the data collection is complete, a logit regression based on the purchase intention (PI) responses (the fourth task) determines how heavily to weight the SE and PC data. For more details see Green, Krieger and Agarwal (1991) and Sawtooth (1996).

**Fixed Efficient Design**

To test whether the adaptive nature of the questions improves or degrades accuracy we included a "Fixed" benchmark, in which the set of PC questions was chosen in advance. Every respondent saw the same set of questions, although the order of the questions was randomized. For the Fixed design we used an algorithm that uses the geometric mean of the eigenvalues of the information matrix to maximize a D-efficiency criterion – a criterion that yields profile designs that are as balanced and orthogonal as feasible (Kuhfeld 1999; Kuhfeld, Tobias and Garratt 1994). Because the PC data are metric, the partworths are estimated with ordinary least-squares regression based on the $q$ questions. Regression requires more data points than estimated parameters, and so the Fixed method provides estimates only after $q$ exceeds $p$ by sufficient degrees of freedom. In later sections we investigate Hierarchical Bayes estimates that do not require that $q$ exceeds $p$.

All three methods use metric PC questions, but only ACA requires the SE (and PI) questions. Therefore, as the accuracy of the SE responses improve, we expect an improvement in the relative performance of ACA. Consistent with this, Toubia, Simester and Hauser's (2002) simulations suggest that if the SE responses are sufficiently accurate, then ACA outperforms both the Fixed and FP methods. However, in their simulations ACA does not perform as well as the other two methods when the SE responses are noisy.

## The Across-Subjects Research Design to Compare Methods

Conjoint analysis represents a product as a profile of its features and price. The preference for a product is then a function of the "partworths" assigned to the levels of each feature. The product used in the study was an innovative new laptop computer bag that includes a removable padded sleeve to hold and protect a laptop computer. At the time of our study, this product was not yet on the market so respondents had no prior experience with it. Subsequent to our research and based, in part, on data from our research, the product is now commercially available from Timbuk2 (http://www.timbuk2.com). Figure 1 offers a photograph illustrating examples of the bags.

The bag includes a range of separable product features, such as the inclusion of a mobile-phone holder, side pockets, or a logo. We focused on nine product features, each with two levels, and included price as a tenth feature. Price is restricted to two levels ($70 and $100) – the extreme prices for the bags in both the internal and external validity tests. For all three conjoint methods, we estimated the part-worths associated with prices between $70 and $100 by linearly interpolating. A more detailed description of the product features can be found on the website listed in the acknowledgements section of this paper.

**Figure 1**

**Examples of Product Category (Laptop Computer Bags)**

Based on pretests, the features appear to satisfy preferential independence, allowing us to model preference for each bag as a linear additive combination of these features (Keeney and Raiffa 1976, p. 101-105; Krantz, et. al. 1971). We recognize that in other applications features may have more than two levels and may not be separable, discrete, or additive. The separable nature of the product features may also influence the accuracy of the self-explicated (SE) questions used by ACA. In particular, we expect SE responses to be more accurate in categories where customers make purchasing decisions about features separately, perhaps by choosing from a menu of features. In contrast, we expect SE responses to be less accurate for products where the features are typically bundled together, so that customers have little experience in evaluating the importance of the individual features. Thus, our choice of separable features is likely to favor ACA relative to the FP (and Fixed) methods and, as such, provides a conservative test of the FP method (relative to ACA).

**Research Design**

We begin with an overview of the research design and then discuss each element of the design in greater detail. The overview of the research design is summarized in Table 1 and the detailed research design is summarized in Table 2. Subjects were randomly assigned to one of the three conjoint methods. After completing the respective conjoint tasks, all of the respondents were presented with the same validation exercises. The internal validation exercise involved four holdout metric paired-comparison (PC) questions, which occurred immediately after the sixteen PC questions designed by the respective conjoint methods. The external validation exercise was the selection of a laptop computer bag from a choice set of five bags. This exercise occurred in the same session as the conjoint tasks and holdout questions, but was separated from these activities by a filler task designed to cleanse memory.

**Table 1**

**Overview of the Research Design**

| FastPace | Fixed | ACA |
|---|---|---|
| FP Conjoint | Fixed Conjoint | ACA Conjoint |
| Internal Validity Task | Internal Validity Task | Internal Validity Task |
| External Validity Task | External Validity Task | External Validity Task |

**Conjoint Tasks**

Recall that ACA requires four sets of questions. Pretests confirmed that all of the features were acceptable to the target market, allowing us to skip the unacceptability task. This left three remaining tasks: self-explicated (SE) questions, metric paired-comparison (PC) questions, and purchase intention (PI) questions. ACA uses the SE questions to select the PC questions, thus the SE questions in ACA must come first, followed by the PC questions and then the PI questions. To test ACA fairly, we adopted this question order for the ACA condition.

The Fixed and FP techniques do not require SE or PI questions. Because asking the SE questions first could create a question-order effect, we asked only the PC questions (not the SE or PI questions) prior to the validation task in the Fixed condition.[4] To investigate the question-order effect we included two FP data collection procedures: one that matched that of fixed-efficient designs (FP1) and one that matched that of ACA (FP2). In FP1 the PC questions precede the validation task, while in FP2, the SE,

[4] See Alreck and Settle 1995; Green, Krieger and Agarwal 1991; Huber, et. al. 1993; Johnson 1991. For more general discussions of question-order effects, see Bickart 1993; Feldman and Lynch 1988; Nowlis and Simonson 1997; Tourangeau, Rips and Rasinski 2000

PC, and PI questions precede the validation task. This enables us to (a) explore whether the SE questions affect the responses to the PC questions and (b) evaluate hybrid techniques that combine features of the FP and ACA methods.

Our complete research design, including the question order, is summarized in Table 2. Questions associated with the conjoint tasks are highlighted in green (Rows 1, 2 and 4), while the validation tasks are highlighted in yellow (Rows 3 and 6). The filler task is highlighted in blue (Row 5). In this design, FP1 can be matched with Fixed; FP2 can be matched with ACA.

**Table 2**

**Detailed Research Design**

| Row | FP1 | Fixed | FP2 | ACA |
|---|---|---|---|---|
| 1 | | | Self-explicated | Self-explicated |
| 2 | FP paired comparison | Fixed paired comparison | FP paired comparison | ACA paired comparison |
| 3 | Internal validity task | Internal validity task | Internal validity task | Internal validity task |
| 4 | | | Purchase intentions | Purchase intentions |
| 5 | Filler task | Filler task | Filler task | Filler task |
| 6 | External validity task | External validity task | External validity task | External validity task |

**Internal Validity Task: Holdout PC Questions**

We allowed each of the conjoint methods to design sixteen metric paired-comparison (PC) questions. Following these sixteen questions, respondents were presented with four additional PC questions. These four questions served as holdout questions and were used to evaluate internal validity. This is a common test that has been used extensively in the literature.[5]

[5] The number of holdout questions varies, but is generally between two and four questions. Examples include Acito and Jain (1980), Akaah and Korgaonkar (1983), Cattin, Hermet and Pioche (1982), Elrod, Louviere and Davey (1992), Green, Goldberg and Wiley (1982), Green, Goldberg and Montemayor (1981), Green and Helsen (1989), Green, Helsen and Shandler (1988), Green and Krieger (1995), Green, Krieger and Agarwal (1991), Green, Krieger and Bansal (1988), Haaijer, Wedel, Vriens and Wansbeek (1998), Hagerty (1985), Huber (1975), Huber, Wittink, Fiedler and Miller (1993), Hauser and Koppelman (1979), Hauser and Urban (1977), Hauser, Tybout and Koppelman (1981), Jain, Acito, Malhotra and Mahajan (1979), Johnson, Meyer and Ghosh (1989), Johnson (1999), Lenk,

The same procedure was used to design the four holdout questions in each of the experimental conditions, and this design was independent of the sixteen questions that the respondents had just answered. In particular, the product profiles in these four questions were randomly selected from an independent efficient design. There was no separation between the sixteen initial questions and the four holdout questions, so that respondents were not aware that the questions were serving a different role.

**Filler Task**

The filler task was designed to separate the conjoint tasks and the external validity task. It was hoped that this separation would mitigate any memory effects that might influence how accurately the information from the conjoint tasks predicted which bags respondents chose in the external validity tasks. The filler task was the same in all four experimental conditions and comprised a series of questions asking respondents about their satisfaction with the survey questions. There was no significant difference in the responses to the filler task across the four conditions.

**External Validity Task: Final Bag Selection**

Respondents were told that they had $100 to spend and were asked to choose between five bags. The five bags shown to each respondent were drawn randomly from an orthogonal fractional factorial design of sixteen bags. This design was the same across all four experimental conditions, so that there was no difference, on average, in the bags shown to respondents in each condition. The five bags were also independent of responses to the earlier conjoint questions. The price of the bags varied between $70 and $100 reflecting the difference in the market price of the features included with each bag. By pricing the bags in this manner we ensured that the choice set represented a Pareto frontier, as recommended by Elrod, Louviere, and Davey (1992), Green, Helsen and Shandler (1988), and Johnson, Meyer and Ghosh (1989).

Respondents were instructed that they would receive the bag that they chose. If the bag was priced at less that $100, they were promised cash for the difference. In order to obtain a complete ranking, we told respondents that if one or more alternatives were unavailable, they might receive a lower ranked bag. The page used to solicit these rankings is presented in Figure 2. We acknowledge two tradeoffs in this design. The first is an endowment effect because we endow each respondent with $100. The second is the lack of a "no bag" option. While both are interesting research opportunities and quite relevant to market forecasting, neither should favor one of the three methods relative to the other; the endowment/forced-choice design is common to all treatments. Pragmatically, we designed the task to maxi-

---

DeSarbo, Green and Young (1996), Malhotra (1986), Moore (1980), Moore and Semenik (1988), Orme, Alpert and Christensen (1998), Orme and King (1998), Parker and Srinivasan (1976), and Tybout and Hauser (1981).

maximize the power of the statistical comparisons of the four treatments. The forced-choice also helped to reduce the (substantial) cost of this research.

At the end of the study the chosen bags were distributed to respondents together with the cash difference (if any) between the price of the selected bag and $100.

### Self-Explicated and Purchase Intention Questions

The self-explicated questions asked respondents to rate the importance of each of the ten product features. For a fair comparison to ACA, we used the wording for the questions and (four point) response scale proposed by Sawtooth (Sawtooth 1996). Further details regarding the design of the questions and the measurement scales are available in Sawtooth (1996) and by visiting the website listed in the acknowledgements section of this paper (see also Figure 3d). For the purchase intentions questions, respondents were shown six bags and were asked how likely they were to purchase each bag. We again adopted the wording, response scale, and algorithms for profile selection that are suggested by Sawtooth.[6]

**Figure 2**
**Respondents Choose and Keep a Laptop Computer Bag**

**Subjects**

The subjects (respondents) were first-year MBA students. Although the subjects had taken basic marketing, they neither knew the objectives of our experiment nor had they taken a course in which conjoint analysis was taught in detail. A total of 360 students were sent an email inviting their participation. We received 330 complete responses to the first stage of the study (there was one incomplete response) resulting in a response rate of over 91%. This high response rate reflects the perceived desirability of the laptop bag relative to the short time investment required to respond. Random assignment yielded 80 subjects for the ACA condition, 88 for the Fixed condition, and 162 for the FP conditions broken out as 88 for the standard question order (FP1) and 74 for the alternative question order (FP2).

The questionnaires were pretested on a total of 69 subjects drawn from professional market research and consulting firms, former students, graduate students in Operations Research, and second-year students in an advanced marketing course that studied conjoint analysis. The pretests were valuable for fine-tuning the question wording and the web-based interfaces. By the end of the pretest, respondents found the questions unambiguous and easy to answer. Following standard scientific procedures, the pretest data were not merged with the experimental data. However, analysis of this small sample suggests that the findings agree directionally with those reported here, albeit not at the same level of significance.

**Additional Details**

Figure 3 illustrates some of the key screens in the conjoint analysis questionnaires. In Figure 3a respondents are introduced to the price feature.[7] Figure 3b illustrates one of the dichotomous features – the closure on the sleeve. This is an animated screen which provides more detail as respondents move their pointing devices past the picture. Figure 3c illustrates one of the PC tasks. Each respondent was asked to rate his/her relative preference for two profiles that varied on three features. Sawtooth (1996) recommends that respondents can handle three features once they are familiar with the task and cites experience that little benefit is gained by using more than three features. To provide a fair comparison to ACA, we adopted this guideline. Both text and pictures were used to describe the profiles. Features that did not vary between the products were chosen to coincide with the respondent's choices in the tasks illustrated in Figure 3b. The format was identical for all four experimental treatments. Finally, Figure 3d illustrates the first three self-explicated questions. The full questionnaires for each treatment are available on the website listed in the acknowledgements to this paper.

---

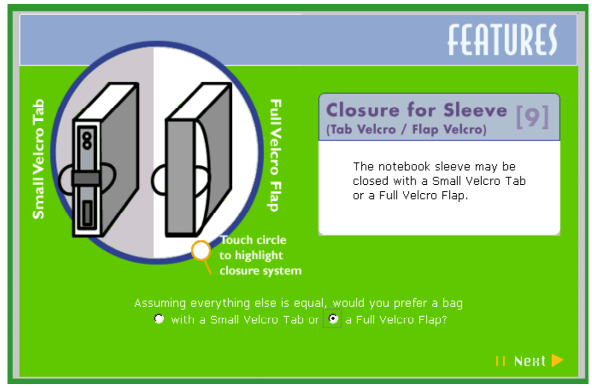
[6] The six profiles include the most attractive, the least attractive, and four "middling" profiles as determined by the respondent's preferences. The four middling profiles are chosen to differ most strongly based on partworths as determined by the SE and PC questions.

[7] Due to human-subject concerns, the details of the $100 incentive were described in an introductory screen which included other experimental protocols.

**Figure 3**

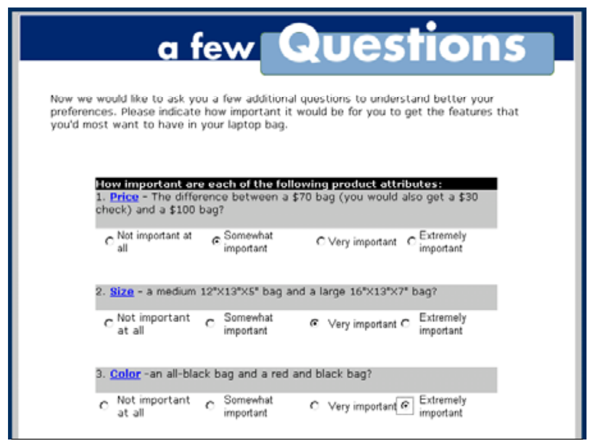
**Example Screens from Questionnaires**


(a) Price as change from $100




(b) Introduction of “sleeve” feature


(c) Metric paired-comparison (PC) question




(d) Self-explicated (SE) questions

### Test of Internal Validity

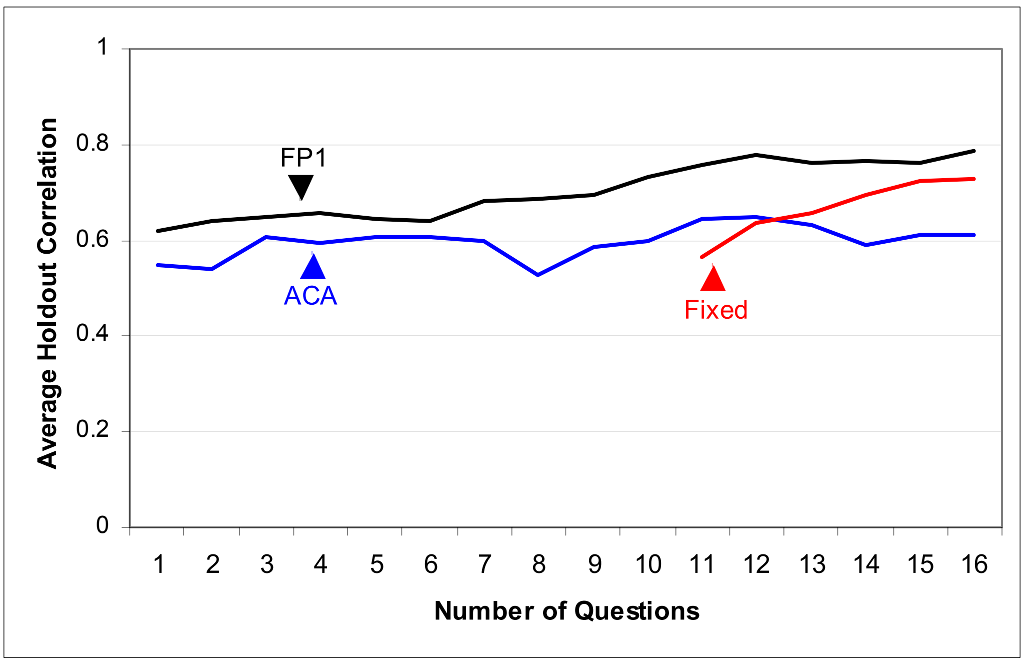
To test internal validity we compare the ability of each method to predict preferences in four holdout metric paired-comparison (PC) questions. For each respondent we calculated the correlation between the predicted and observed responses across the four questions. This correlation, averaged across respondents in each condition, is summarized in Figure 4. The correlations are recalculated after each additional PC question so that we can observe the incremental accuracy from asking additional questions.

As an alternative metric, we compared how well the methods predicted which product the respondents favored from each pair. The two metrics provide a very similar pattern of results and so, for ease of exposition, we focus on the correlation measure in Figure 4.[8] To ensure that the FP method is neither helped nor hindered by question order effects, we evaluate FP in these comparisons using FP1. In later discussion we investigate question-order effects by comparing FP1 and FP2.

The correlations are within the ranges reported in prior research. When metric data are gathered, correlations of predicted and observed ratings vary from 0.25 to 0.82 and are typically in the range of 0.60 to 0.75. For this product category and for this feature set, the adaptive methods, FP and ACA, appear to provide reasonable accuracy with far fewer questions than a non-adaptive method. We examine later whether this advantage holds for Hierarchical Bayes estimation.

**Figure 4**

**Interval Validity Test – Holdout Pairs**




For the Fixed method, OLS estimation is only possible after at least eleven PC responses (there are ten parameters to estimate). Although accuracy after eleven questions is lower than the other two methods, accuracy improves when additional PC responses are received. Toubia, Simester and Hauser (2002) report simulations suggesting a similar pattern of results using related Fixed models. They found that fixed designs do not do as well as FP and ACA when responses are available for only a small number of questions. It is only after the number of questions exceeds 150% of the number of parameters (15 questions in our experiment) that their Fixed method perform as well as either of the other methods. Were we

[8] There is an additional reason to focus on correlations in the external validity tests. First preference prediction is a dichotomous variable and has higher variance than the Spearman correlation which is based on five observations per

to ask 20 questions from an efficient design, their simulations suggest that the Fixed method would approach the accuracy of FP.

To evaluate whether the difference in the performance of the three methods is statistically significant we focus first on the final estimates using responses to all sixteen PC questions. The analysis confirms that the FP estimates are significantly ($p<0.05$) more accurate than the ACA estimates. After sixteen questions, the Fixed method is also significantly ($p<0.05$) more accurate than the ACA method, but is not significantly different from the FP method.

Because our managerial goal was to evaluate the methods for low numbers of questions, we also evaluated significance by a method that pools the correlation measures calculated after each additional PC question. This results in a total of sixteen observations for each respondent in the FP and ACA conditions, and six observations for respondents in the Fixed condition. To control for heteroscedasticity due to the question number we estimate a separate intercept for each question number. The multivariate specification also allows us to explicitly control for heterogeneity in the respondent samples. In particular, the null model in each condition is that the ten laptop bag features are equally important. The methods each seek to improve on this null model by estimating variation in the importances of the features for each respondent. If, despite the random assignment of respondents to conditions, the responses in one condition are more consistent with the null model, then the comparisons will be biased in favor of this condition. We explicitly control for this possibility by including a measure describing how accurately the equal weights (null) model performs on the respective validity measures for each respondent.

The complete specification is described in Equation 1, where *r* indexes respondent and *q* indexes the number of PC questions used in the partworth estimates. The α's and β's are coefficients in the regression and $\varepsilon_{rq}$ is an error term. The equation is estimated using OLS.

$$(1)\qquad Correlation_{rq} = \sum_{q=1}^{15} \alpha_q Question_q + \beta_1 Fixed + \beta_2 ACA + \beta_3 EqualWeight_r + \varepsilon_{rq}$$

The independent variables are defined as follows:

| | |
|---|---|
| *Fixed* | 1 if the respondent was in the Fixed condition; 0 otherwise. |
| *ACA* | 1 if the respondent was in the ACA condition; 0 otherwise. |
| $Question_q$ | 1 if the correlation was calculated using only the first *q* PC questions; 0 otherwise. |
| $EqualWeight_r$ | Correlation obtained for respondent *r* with an equal-weights model. |

Under this specification, the intercepts ($\alpha_q$'s) describe the correlations for FP (from the FP1 condition) after each question, while the $\beta_1$ and $\beta_2$ coefficients represent the expected increase or decrease in

respondent.

this correlation in the Fixed and ACA conditions respectively. Positive (negative) values for the $\beta_1$ and $\beta_2$ coefficients indicate higher (lower) correlation for the respective method compared to FP. The $\beta_1$ and $\beta_2$ coefficients are presented in Table 3, where we also present findings from a similar analysis conducted on the external validity task. The other coefficients are omitted for ease of exposition. We also estimated a random effects model but there was almost no difference in the coefficients of interest. Moreover, the Hausman specification test favored the fixed-effects specification over the random-effects specification.

The findings from this multivariate analysis indicate that the FP method is significantly more accurate ($p<0.05$) than the ACA and Fixed methods in this internal validity task. There was no significant difference in the performance of the Fixed and ACA methods ($t=1.60$). We conclude that, based on the internal validity tests, the new polyhedral methods (FP) show promise, especially for obtaining reasonable estimates based on fewer questions. We next consider external validity by examining whether FP methods can be used to predict actual customer choice.

**Table 3**

**Multivariate Analysis of the Pooled Data**

| | **Internal Validity: Holdout PC Questions** | **External Validity: Final Bag Selection** |
|---|---|---|
| ACA ($\beta_1$) | -0.127* (0.014) | -0.142* (0.015) |
| Fixed ($\beta_2$) | -0.093* (0.021) | -0.288* (0.021) |
| Adjusted $R^2$ | 0.27 | 0.23 |
| Sample Size | 3,103 | 3,213 |

Coefficients describing the question effects and the performance of the null (*EqualWeights*) benchmark are omitted from this table.
* Significant at $p<0.05$. Standard errors are in parentheses.

## External Validity Test

While tests of internal validity are common in the conjoint-analysis literature, tests of external validity at the individual level are rare.[9] A search of the literature revealed four studies that predict choices in the context of natural experiments and one study based on a lottery choice. Wittink and Montgomery (1979), Srinivasan (1988), and Srinivasan and Park (1997) all use conjoint analysis to predict MBA job choice. Samples of 48, 45, and 96 student subjects, respectively, completed a conjoint questionnaire prior to accepting job offers. The methods were compared on their ability to predict these stu-

[9] Some researchers report aggregate predictions relative to observed market share. See Bucklin and Srinivasan (1991), Currim (1981), Davidson (1973), Green and Srinivasan (1978), Griffin and Hauser (1993), Hauser and Gaskin (1984), McFadden (2000), Page and Rosenbaum (1989), and Robinson (1980).

dents' job choices. First preference predictions ranged from 64% to 76% versus random-choice percentages of 26-36%. This corresponds to an improvement relative to random in the range of 50-56%.[10]

In another natural experiment Wright and Kriewall (1980) used conjoint analysis (Linmap) to predict college applications by 120 families. They were able to correctly predict 20% of the applications when families were prompted to think seriously about the features measured in conjoint analysis; 15% when they were not. This converts to a 16% improvement relative to their null model. Leigh, MacKay and Summers (1984) allocated 122 undergraduate business majors randomly to twelve different conjoint tasks designed to measure partworths for five features. Respondents indicated their preferences for ten calculators offered in lottery. There were no significant differences among methods with first-preference predictions in the range of 26-41% and percentage improvements of 28%. The authors also compared the performance of estimates based solely on SE responses and observed similar performance to the conjoint methods.

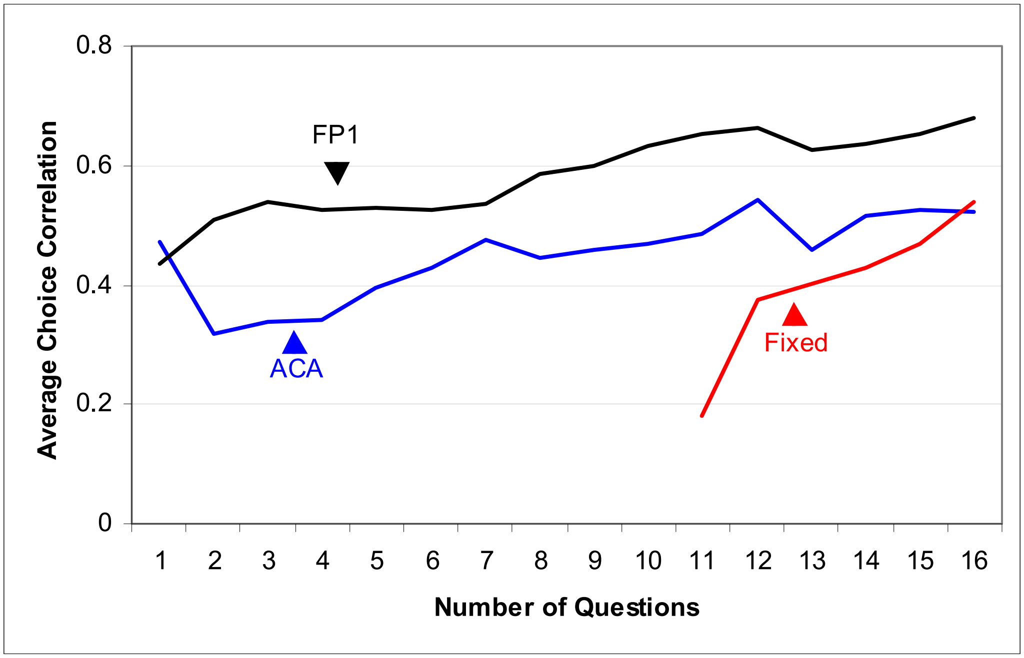
In the related choice-based conjoint literature, Louviere, Hensher, and Swait (2000) compare the convergent validity between revealed preference and "stated preference" methods to suggest that, in most cases, the two methods measure the same relative coefficients (partworths). Although this is not predictive validity per se, it does provide evidence that stated-preference questions, such as metric paired-comparison (PC) questions, provide data consistent with actual choices made by consumers. Urban and Katz (1983) provide further evidence that models based on PC questions, such as their Assessor model, predict well.

Against this background we examine the results from the external validity task in which respondents "purchased" a bag from a selection of five bags. In case their chosen bag was unavailable, the respondents ranked all five bags. To evaluate the conjoint methods we calculated the correlation between the actual and observed rankings for the five bags shown to each respondent. We then averaged these correlations across respondents in each condition. (Results based on first choice were qualitatively similar.) The results are presented in Figure 5, where we again present findings for estimates calculated after each of the PC questions.

[10] This commonly reported metric is calculated as (predicted – random)/(100% - random).

**Figure 5**

**External Validity Test – Correlation with Actual Choice**




The findings from this task are similar to the results of the internal validity test. The FP method yields more accurate predictions of the respondents' bag rankings than either the ACA or Fixed methods. When using all sixteen PC responses, the FP estimates are significantly more accurate than both the ACA estimates ($p<0.05$) and the Fixed estimates ($p<0.05$). The difference between the Fixed and ACA methods is not significant. We also compared the methods by re-estimating Equation 1 using the correlations from the external validity task. The findings from this pooled data are reported in Table 3. They reveal a similar pattern of results, with FP significantly ($p<0.05$) more accurate than ACA and Fixed. In the pooled data ACA is also significantly ($p<0.05$) more accurate than the Fixed method (over the questions for which they can be compared).

Inspection of Figure 5 reveals another qualitative finding of note. ACA's predictive ability appears to first decrease as a function of $q$ and then increase. Toubia, Simester and Hauser (2002) and Johnson (1987) report similar results when ACA's SE measures are accurate relative to the PC responses. The addition of noisy PC responses initially reduces overall accuracy because the information in these responses is outweighed by the noise that they introduce relative to the SE measures. Eventually, when there are sufficient PC responses, the information in these responses outweighs the noise.

It is also helpful to compare how frequently the different methods identified which bag each respondent most preferred. After all sixteen questions, FP correctly identified the first preference bag for approximately 59% of respondents, compared to 52% for the Fixed method and 50% for ACA.[11] A random selection would have predicted the first preference bag just 20% of the time. These correspond to percentage improvements of 49%, 40% and 38% for the FP, Fixed and ACA methods respectively. These percentage improvements are similar to the ranges reported elsewhere in the literature.

**Summary of Internal and External Validity Tests**

Table 4 reviews the performance of the three methods on the internal and external validation metrics based on all sixteen questions. For completeness we present correlations between actual and predicted choices together with the proportion of times that the methods identified the respondents' favorite bag (in each task). The findings are remarkably consistent across the two validity tasks and confirm that the estimates from the FP method are more accurate than the ACA and Fixed estimates. Recall that polyhedral methods for conjoint analysis were developed to provide reasonable estimates with fewer questions, especially for web-based questionnaires and for use in the new product development processes. Given further development and refinement, it seems likely that researchers will develop polyhedral methods that can do even better than current FP methods.

**Table 4**

**Summary of the Performance Metrics**

| | ACA | Fixed | FP1 |
|---|---|---|---|
| **Internal Validity: Holdout PC Questions** | | | |
| Correlation with holdout pairs $^{*,+}$ | 0.61 | 0.73 | 0.79 |
| First preference $^{*,\,o,\,+}$ | 0.79 | 0.87 | 0.89 |
| | | | |
| **External Validity: Final Bag Selection** | | | |
| Correlation with choice among products $^{*,\,o}$ | 0.52 | 0.54 | 0.68 |
| First preference * | 0.50 | 0.52 | 0.59 |

This summary is based predictions that use all sixteen PC questions.
* Significant difference ($p<0.05$) between FP and ACA.
$^{o}$ Significant difference ($p<0.05$) between FP and Fixed.
$^{+}$ Significant difference ($p<0.05$) between Fixed and ACA.

Although FP performed significantly better than ACA and the non-adaptive Fixed design, neither ACA nor Fixed should be rejected. Both do well and both have proven valuable in many managerial applications. ACA and Fixed have strengths that might allow them to outshine FP in other empirical con-

[11] FP is significantly better than ACA ($p<0.05$). The difference between FP and Fixed is only marginally significant ($p<0.07$).

texts. Moreover, components of these methods might be combined with components of the FP method to produce an even better method. In a later section we evaluate opportunities to improve predictions through hybrid techniques that draw components from each of the three methods.

## Question-Order Effects – FP1 vs. FP2

In the comparisons reported in the previous sections there is an important difference between the conditions. In the ACA condition, self-explicated (SE) questions precede the metric paired-comparison (PC) questions. This was not the case in the FP1 and Fixed conditions and it is possible that this difference affected the results (recall that we used FP1 in these comparisons). For example, the SE questions may tire respondents causing them to pay less attention to the PC questions (e.g., Alreck and Settle 1995). If this were true, then question order could explain the improved performance of FP1 (or Fixed) relative to ACA. Alternatively, the SE questions may improve the accuracy of the PC questions by acting as training task. Some researchers have observed a learning or priming phenomenon, in which the SE questions help respondents clarify their values, increasing the accuracy of the PC questions (Green, Krieger and Agarwal 1991; Huber, et. al. 1993; Johnson 1991). The literature advances several theories that support this training hypothesis, including task learning, self-preference learning, memory accessibility, and context effects (Bickart 1993; Feldman and Lynch 1988; Nowlis and Simonson 1997; Tourangeau, Rips and Rasinski 2000; Simmons, Bickart and Lynch 1993). Because any "learning" effect might counter the wear out effect, the directional impact of question order is an empirical issue.[12] By comparing FP1 and FP2 we can examine whether or not there is a significant question-order effect. If the effect favors FP1, then we can use FP2 to parse the improved performance of FP1 relative to ACA. If the effect favors FP2, then our results are strengthened and FP2 becomes the preferred version of FP. In either case, future experiments can explore the source of any question-order effect.

When using all sixteen questions, the predictive accuracy of the FP1 and FP2 conditions are not statistically different. This suggests that there was no measurable wearout due to the SE questions, and that by the sixteenth question any warm-up/learning advantage had disappeared. However, there might still be an effect for low numbers of questions. When we plot FP1 and FP2 for questions 1 through 16 we observe that FP2 performs slightly better than FP1 after asmall number of questions (this plot is not reported). However, when we estimate the performance using a version of Equation 1, the effect is not significant for the choice correlations ($t = 0.72$), the choice hit rate ($t = 0.34$), and the holdout hit rate ($t = 0.72$). FP2 is significantly better on the holdout correlations ($t = 2.48$).

[12] We might also observe a question-order effect if the PC questions preceded the SE questions or if the choice task preceded the SE questions. For example, Green, Krieger and Agarwal (1991) observe an effect that is consistent with PC-affects-SE hypothesis, but Huber, et. al. (1993) do not. In addition, theories in Allen 1982; Allen and Dil-

In summary, the evidence is mixed. FP2 performs better than FP1 on all four metrics, but the difference is only significant for one of these metrics. The evidence that FP2 performs better than FP1 does confirm that the favorable performance of FP1 compared to ACA cannot be explained by the question order, and is instead due to differences in the question design and/or estimation methods.

## Improving Predictive Performance with Hybrid Methods

The term heterosis is used in biology to describe the principle under which genetically-diverse offspring of two pure species often have traits superior to either of their purebred parents (Campbell 1996, p. 428). Analogously, there is evidence that combinations of conjoint methods often yield more accurate or more efficient predictions than either of the parent methods.[13] Successful hybrid conjoint models have combined the power of compositional methods, such as self-explicated (SE) questions, and decompositional methods, such as paired-comparison (PC) questions to produce new estimates. Empirical tests suggest that both SE and PC questions add incremental information (Green, Goldberg, and Montemayor 1981; Huber, et. al. 1993, Johnson 1999; Leigh, MacKay, and Summers 1984). In this context, ACA itself can be considered a hybrid. Although there are instances in which purebred methods either outperform or provide equivalent accuracy to hybrid methods, there are many situations and product categories in which hybrid methods outperform their purebred progenitors (Green 1984; Wittink and Bergestuen 2001).

### Alternative Estimation Procedures Using FP Question Selection

In this section we hold the FP question-selection method constant and investigate the accuracy of alternative estimation methods. In particular, we consider two different approaches for incorporating self-explicated (SE) responses into the FP estimates. To do so we focus on responses from the FP2 condition in which SE responses were collected.

In the first approach, which we label *Balanced Hybrid*, we combine the SE and PC responses by minimizing a least squares norm in which both the SE responses and the PC responses are treated as data. For $p$ SE responses and $q$ PC responses, the regression has $p+q$ observations.[14] This procedure is similar to the approach suggested by Sawtooth (Sawtooth 1999) as a modification of the ACA method and is

lon 1982; Bem 1972; Folkes and Kiesler 1991; and Tybout and Yalch 1980 suggest that SEs that follow a choice task might be stated to justify choices.

[13] See for example: Akaah and Korgaonkar 1983; Cattin, Hermet, and Pioche 1982; Carroll and Green 1995, Green 1984; Green, Goldberg and Montemayor 1981; Green, Goldberg and Wiley 1982; Green and Krieger 1995; Huber 1974; Huber, et. al. 1993; Moore and Semenik 1988; Wind, et. al. 1989.

[14] In particular, for the Balanced Hybrid we minimize the following norm: $\left\| \begin{bmatrix} I \\ X \end{bmatrix} \vec{u} - \begin{bmatrix} \vec{s} \\ \vec{a} \end{bmatrix} \right\|_2$ where $\vec{s}$ is the vector of prior self-explicated partworth estimates, $X$ is the design matrix for the paired comparisons (chosen adaptively), $\vec{a}$ is the vector of respondent answers to the paired comparisons, $I$ is the identity matrix, and $\vec{u}$ are the estimates.

similar to the OLS model that ACA uses to obtain intermediate partworth estimates. In the second benchmark, which we label *FP with SE-based Constraints (FPSE)*, we introduce the SE responses as constraints on the feasible polyhedron that forms the basis of the FP estimation procedure. For example, if Feature 1 is given a higher SE importance weight than Feature 2, we add a constraint to the FP polyhedron such that the partworth of Feature 1 is larger than the partworth of Feature 2.

**Figure 6**

**Hybrid Methods Based on FP Question-Selection**

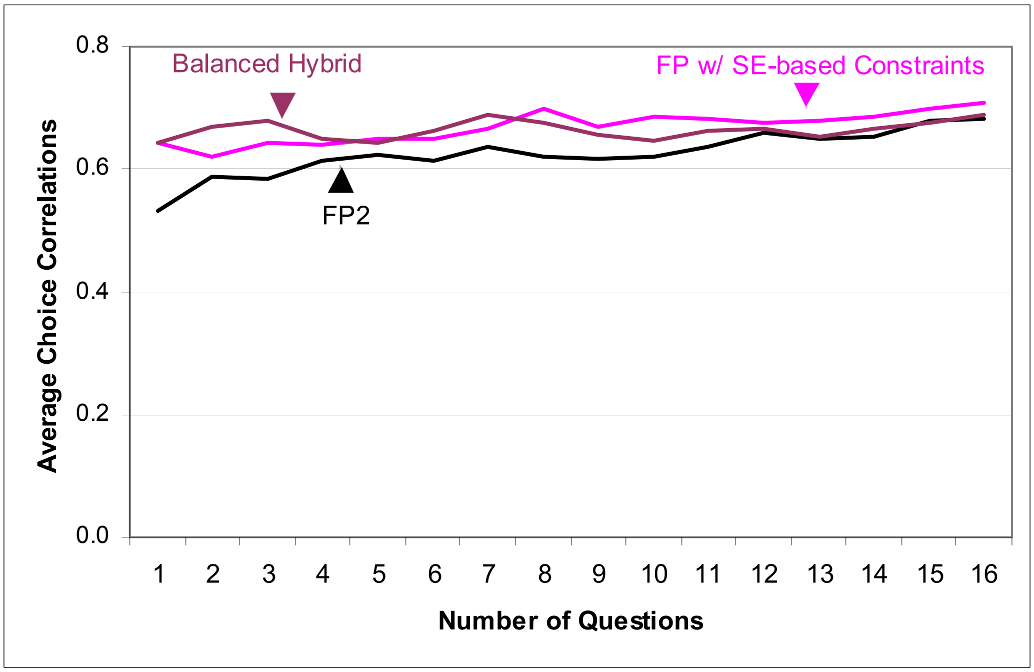



Figure 6 compares the two hybrid methods to a purebred FP. Note that all of these methods use data from the FP2 condition, so that the only difference between the estimates is the estimation method. Because we obtain similar results for both the holdout pairs and the choice of bags, we report only the latter correlations in Figure 6. Based on Equation 1, both hybrids are significantly better ($p<0.05$) than a purebred FP. Of the two hybrids, FPSE appears to do slightly better than the Balanced Hybrid. However, this difference is not significant. Thus, it appears that both the PC and the SE questions add incremental information and that hybrid methods have the potential to predict significantly better than either FP or SE alone – at least when the PC questions are chosen by FP.[15] These findings are consistent with Toubia, Simester and Hauser's (2002) simulations which suggest that SE-PC hybrids have the potential to improve FP. These findings also suggest that further research on hybrid estimation could improve the meth-

[15] For q=16 on the FP2 sample, both hybrids and FP perform better than SE alone. Average correlations are 0.71, 0.69, 0.68, and 0.64 for FPSE, Balanced Hybrid, FP, and SE, respectively.

ods further. However, we must caution the reader that these estimation procedures were tested after the data were collected.

**Alternative Estimation Procedures Using ACA Question Selection**

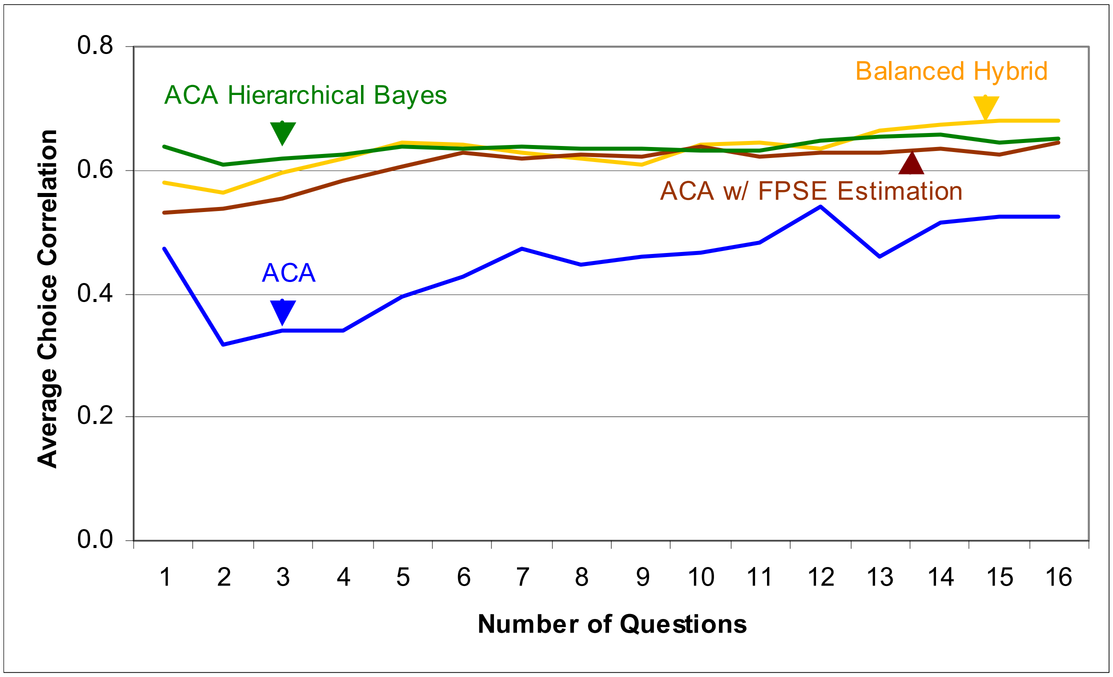
For the ACA questions we consider three alternative approaches for combining self-explicated responses with the ACA metric paired-comparison responses. The first two approaches are the same as those used in the FP analysis above. In particular, we apply FPSE estimation and Balanced Hybrid estimation using the self-explicated responses and paired-comparison responses from the ACA condition. In addition, we attempt to improve predictions with Hierarchical Bayes (HB) estimation. HB is based on recent developments in the estimation of conjoint-analysis data suggesting that predictions at the level of the individual respondent can be improved using data from the population of respondents. These methods use the data from the full sample to iteratively estimate individual-level partworths and do so by constraining the distribution of those partworths. The estimation is based on Gibbs sampling and the Metropolis Hastings Algorithm (Sawtooth 1999).[16] Sawtooth's Hierarchical Bayes algorithm provides four alternatives for incorporating the SE data. Preliminary analysis revealed that the most accurate of these alternatives included the SE data as constraints but did not include them in the objective function. We used this alternative, which is recommended by Sawtooth, in our comparisons.

The predictions based on these hybrid methods are shown in Figure 7. Based on the multiple-question statistical tests (Equation 1), all three hybrid methods are significantly better ($p<0.05$) than the purebred ACA. Thus, for ACA-based questions, it appears that heterosis again improves predictions. The findings for ACA HB are particularly interesting in Figure 7. By using population-level PC data, ACA HB performs well after only a single PC question. However, increasing the number of paired comparison responses did not yield an improvement in the accuracy of this technique. This lack of improvement contrasts with the other hybrid techniques. As the number of PC questions increases, the Balanced Hybrid method, which relies only on data from each respondent, eventually outperforms ACA HB. We caution that this difference is not significant.

[16] See also Allenby and Rossi (1999); Arora, Allenby and Ginter (1998); Lenk, et. al. (1996); and Liechty, Ramaswamy and Cohen (2001).

**Figure 7**

**Hybrid Methods Based on ACA Question-Selection**




**Alternative Estimation Procedures Using Fixed-Efficient-Design Question Selection**

In the Fixed condition there is no suitable self-explicated data available to combine with metric paired-comparison responses. However, even without self-explicated data, Hierarchical Bayes (HB) may improve the performance of the Fixed method by using data from the population of respondents to moderate respondent heterogeneity. HB offers the added advantage that it can estimate individual partworths even if the individual has provided fewer PC responses than the number of parameters. It uses information from the population of respondents to make up for the lack of individual responses.

In the ACA HB analysis we used the self-explicated responses to provide starting values. Given the absence of self-explicated responses in the Fixed condition, we used equal weights as starting values for the importance of the nine product features and randomly drew the starting values for price from a uniform distribution.[17] Further investigation revealed that the final parameter estimates (although not the performance metrics) were sensitive to this random draw. For this reason, we repeated the analysis three times, with different random draws in each case, and then averaged the findings across the three repetitions.

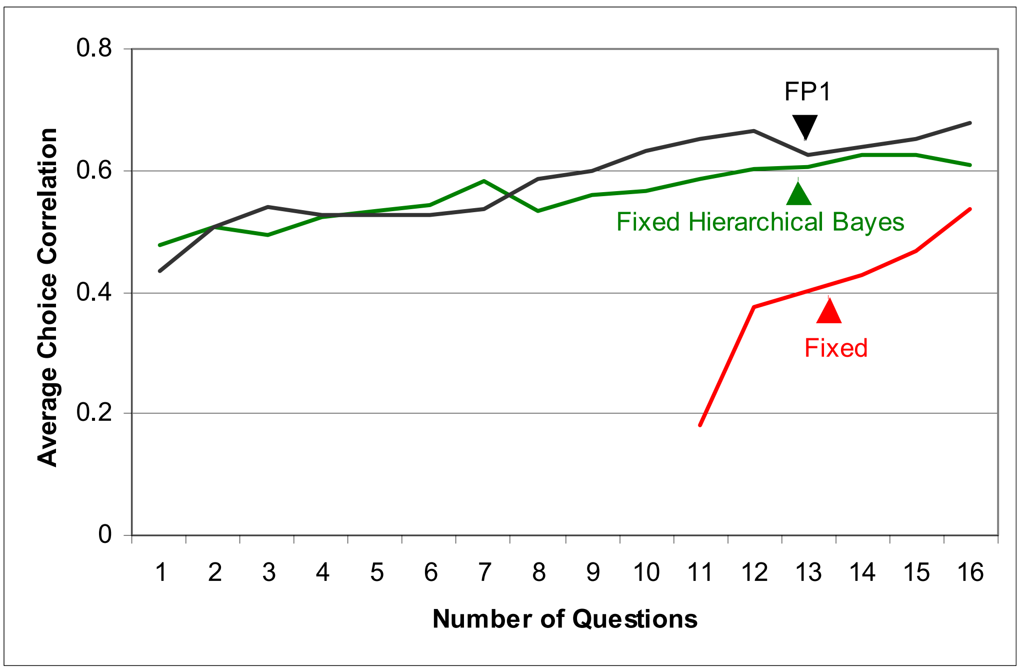
---

[17] If price also had the same starting value the initial covariance matrix would be singular.

Figure 8 reports the results obtained for the external validity data when estimating partworths using Hierarchical Bayes in the Fixed condition. For comparison, we include the standard Fixed results from Figure 5, where the partworths are estimated using OLS. In the absence of an alternative Fixed benchmark, we also include the FP1 results from Figure 5. We used Equation 1 to evaluate the significance of the difference between the Hierarchical Bayes estimates and these two benchmarks. The Hierarchical Bayes estimates are significantly more accurate than the standard OLS estimates ($p<0.05$).[18] However, this improvement is still not sufficient to improve upon the FP estimates, which remain significantly more accurate ($p<0.05$) than even the Hierarchical Bayes version of the Fixed method.

**Figure 8**

**Hybrid Methods Based on Fixed Question Selection**




**Summary of Hybrid Analyses**

While we are encouraged by these findings, we must be cautious in our interpretations. The tested hybrid methods do well and all add significantly to a product development team's knowledge of customer preferences. The SE-based hybrids appear to improve predictions, but are based on post analysis of the data in a category where SE questions do well. It should not surprise us that the addition of accurate SE responses can improve predictive accuracy. However, before embracing SE-based hybrids for all categories, further testing is required in product categories for which SE responses are less accurate.

[18] Similar results can be found elsewhere in the literature (cf. Judge, et. al., 1985, Chapter 3).

This may include categories where the product is less separable and more holistic and categories where the direct scaling of partworths by respondents is difficult. We expect that such categories exist. For example, in a review of forty-two SE applications, Wilkie and Pessemier (1973) identify 45% with favorable results, 33% questioning the basic model, and the remainder mixed.

HB estimation also appears to perform well and provides an alternative method to obtain individual-respondent-level estimates with fewer questions than there are parameters.

## Discussion

Polyhedral methods do well in simulation, but, prior to this paper, they were unproven in empirical situations. This paper provides initial tests of polyhedral methods in an actual choice involving a new product. We report tests of internal validity and tests of external validity. The internal validity tests comprise four holdout metric paired-comparison questions, while the external validity tests represent an actual choice from a set of five laptop computer bags. The internal and external tests are consistent with one another and consistent with prior simulations. In this first head-to-head comparison, the FastPace (FP) method appears to improve predictions relative to the benchmarks of ACA and an efficient Fixed (non-adaptive) design.

We also tested several hybrid methods that combine separate components from each technique. These hybrids perform well and confirm that both the self-explicated and metric paired-comparison questions contribute incremental information. Although we recognize that self-explicated questions may be less useful in other applications, these findings suggest that hybrid improvements are worth considering. The greatest improvement for the FP method appears to be the use of self-explicated responses to impose constraints on the feasible polyhedron – a method we have labeled FPSE. The greatest improvement for ACA appears to be a hybrid estimation method (from a previous version of ACA) in which data from the SE and PC questions are balanced. Hierarchical Bayes (HB) estimation also improves ACA with slightly better prediction than the balanced method for low numbers of PC questions and slightly lower prediction for high numbers of PC questions.

We also investigated question-order effects. The findings suggest that asking self-explicated questions before metric paired-comparison questions can improve the accuracy of the paired-comparison questions. However, the effect is not very strong.

In the end we conclude that all of the methods tested appear to perform well on these validity tests. This is good news for the many academics and product-development professionals that rely on conjoint analysis. We conclude further that the polyhedral methods appear to have passed their first empirical test. The performance of FP and the FP hybrids suggest that further developments based on polyhedral methods have the potential to advance conjoint analysis theory and practice.

## References


Acito, Franklin and Arun K. Jain (1980), "Evaluation of Conjoint Analysis Results: A Comparison of Methods," *Journal of Marketing Research*, 17, (February), 106-112.

Akaah, Ishmael P. and Pradeep K. Korgaonkar (1983), "An Empirical Comparison of the Predictive Validity of Self-explicated, Huber-hybrid, Traditional Conjoint, and Hybrid Conjoint Models," *Journal of Marketing Research*, 20, (May), 187-197.

Allen, Chris T. (1982), "Self-Perception Based Strategies for Stimulating Energy Conservation," *Journal of Consumer Research*, 8 (March), 381-390.

_____, and William R. Dillon (1983), "Self-Perception Development and Consumer Choice Criteria: Is There A Linkage?" in *Advances in Consumer Research*, 10, Eds. Richard P. Bagozzi and Alice M. Tybout. Ann Arbor, MI: Association for Consumer Research, 45-50.

Allenby, Greg M. and Peter E. Rossi (1999), "Marketing Models of Consumer Heterogeneity," *Journal of Econometrics*, 89, (March/April), 57-78.

Alreck, Pamela L. and Robert B. Settle (1995), The Survey Research Handbook: Guidelines and Strategies for Conducting a Survey, 2E, (New York, NY: McGraw Hill), p. 102.

Arora, Neeraj, Greg M. Allenby and James L. Ginter (1998), "A Hierarchical Bayes Model of Primary and Secondary Demand," *Marketing Science*, 17, 1, 29-44.

Ben-Akiva, Moshe and S. Gershenfeld (1998), "Multi-featured Products and Services: Analyzing Pricing and Bundling Strategies," *Journal of Forecasting*, 17, 175-196.

Bem, Daryl (1972), "Self Perception Theory," in *Advances in Experimental Social Psychology*, 6, Ed. Leonard Berkowitz, New York: Academic Press.

Bickart, Barbara A. (1993), "Carryover and Backfire Effects in Marketing Research," *Journal of Marketing Research*, 30, 1, (February), 52-62.

Bucklin, Randolph E. and V. Srinivasan (1991), "Determining Interbrand Substitutability Through Survey Measurement of Consumer Preference Structures," *Journal of Marketing Research*, 28, (February), 58-71.

Campbell, Neil A. (1996), *Biology, 4E*, (Reading, MA: The Benjamin/Cummings Publishing Company, Inc.).

Carroll, J. Douglas and Paul E. Green (1995), "Psychometric Methods in Marketing Research: Part I, Conjoint Analysis," *Journal of Marketing Research*, 32, (November), 385-391.

Cattin, Philippe and Dick R. Wittink (1982), "Commercial Use of Conjoint Analysis: A Survey," *Journal of Marketing*, 46, (Summer), 44-53.

______, Garard Hermet and Alain Pioche (1982), "Alternative Hybrid Models for Conjoint Analysis: Some Empirical Results," in *Analytic Approaches to Product and Marketing Planning: The Second Conference*, R. K. Srinivasan and Allen D. Shocker, eds., (Cambridge, MA: Marketing Science Institute), 142-152.

Choi, S. Chan and Wayne S. DeSarbo (1994), "A Conjoint-based Product Designing Procedure Incorporating Price Competition," *Journal of Product Innovation Management*, 11, 451-459.

Currim, Imran S. (1981), "Using Segmentation Approaches for Better Prediction and Understanding from Consumer Mode Choice Models," *Journal of Marketing Research,*18, (August), 301-309.

Cusumano, Michael A. and Richard W. Selby (1995), *Microsoft Secrets*, (New York, NY: The Free Press).

Dahan, Ely and John R. Hauser (2002), "The Virtual Customer,", forthcoming, *Journal of Product Innovation Management*.

______ and V. Srinivasan (2000), "The Predictive Power of Internet-Based Product Concept Testing Using Visual Depiction and Animation," *Journal of Product Innovation Management*, 17, 99-109.

Davidson, J. D. (1973), "Forecasting Traffic on STOL," *Operations Research Quarterly*, 22, 561-569.

Dawes, Robin M. and Bernard Corrigan (1974), "Linear Models in Decision Making," *Psychological Bulletin*, 81, (March), 95-106.

De Angelis, Chris (2001), "Sampling for Web-Based Surveys," paper presented at 22nd Annual Marketing Research Conference, Atlanta, GA, September.

Einhorn, Hillel J. (1971), "Use of Nonlinear, Noncompensatory, Models as a Function of Task and Amount of Information," *Organizational Behavior and Human Performance*, 6, 1-27.

Elrod, Terry, Jordan Louviere, and Krishnakumar S. Davey (1992), "An Empirical Comparison of Ratings-Based and Choice-based Conjoint Models," *Journal of Marketing Research* 29, 3, (August), 368-377.

Eppinger, Steven D. (1998), "Information-Based Product Development," Presentation to the Research Advisory Committee of the Center for Innovation in Product Development, MIT, Cambridge, MA (January 12).

_______, Daniel E. Whitney, Robert P. Smith, and David A. Gebala (1994), "A Model-Based Method for Organizing Tasks in Product Development", *Research in Engineering Design*. 6, 1, 1-13.

Feldman, Jack M. and John G. Lynch, Jr. (1988), "Self-generated Validity: Effects of Measurement on Belief, Attitude, Intention, and Behavior," *Journal of Applied Psychology*, 73, (August), 421-435.

Folkes, Valerie S., and Tina Kiesler, (1991), "Social Cognition: Consumers' Inferences About the Self and Others," in Thomas S. Robertson and Harold H. Kassarjian, eds., *Handbook of Consumer Behavior*, 281-315.

Freund, Robert (1993), "Projective Transformations for Interior-Point Algorithms, and a Superlinearly Convergent Algorithm for the W-Center Problem," *Mathematical Programming*, 58, 385-414.

Gonier, Dennis E. (1999), "The Emperor Gets New Clothes," Paper presented at the Advertising Research Foundation's On-line Research Day and available at www.dmsdallas.com. (January).

Green, Paul E., (1984), "Hybrid Models for Conjoint Analysis: An Expository Review," *Journal of Marketing Research*, pp. 155-169.

Green, Paul E. and Kristiaan Helsen (1989), "Cross-Validation Assessment of Alternatives to Individual-Level Conjoint Analysis: A Case Study," *Journal of Marketing Research*, pp. 346-350.

______, ______, and Bruce Shandler (1988), "Conjoint Internal Validity Under Alternative Profile Presentations," *Journal of Consumer Research*, 15, (December), 392-397.

______, Stephen M. Goldberg, and Mila Montemayor (1981), "A Hybrid Utility Estimation Model for Conjoint Analysis," *Journal of Marketing*, pp. 33-41.

______, ______, and James B. Wiley (1982), "A Cross Validation Test of Hybrid Conjoint Models," in *Advances in Consumer Research*, 10, Richard P. Bagozzi and Alice M. Tybout, eds., (Ann Arbor, MI: Association for Consumer Research), 147-150.

______ and Abba Krieger (1995), "Attribute Importance Weights Modification in Assessing a Brand's Competitive Potential," *Marketing Science*, 14, 3, Part 1 of 2, 253-270.

______, ______, and Manoj K. Agarwal (1991), "Adaptive Conjoint Analysis: Some Caveats and Suggestions," *Journal of Marketing Research*, pp. 215-222.

______, ______, and Pradeep Bansal (1988), "Completely Unacceptable Levels in Conjoint Analysis: A Cautionary Note," *Journal of Marketing Research*, 25, (August), 293-300.

______ and V. Srinivasan (1990), "Conjoint Analysis in Marketing: New Developments With Implications for Research and Practice," *Journal of Marketing*, pp. 3-19.

______ and ______ (1978), "Conjoint Analysis in Consumer Research: Issues and Outlook," *Journal of Consumer Research*, 5, 2, (September), 103-123.

Griffin, Abbie and John R. Hauser (1993), "The Voice of the Customer," *Marketing Science*, vol. 12, No. 1, (Winter), 1-27.

Haaijer, Rinus, Michel Wedel, Marco Vriens, and Tom Wansbeek (1998), "Utility Covariances and Context Effects in Conjoint MNP Models," *Marketing Science*, 17, 3, 236-252.

Hagerty, Michael R. (1985), "Improving the Predictive Power of Conjoint Analysis" The Use of Factor Analysis and Cluster Analysis," *Journal of Marketing Research*, 22, (May), 168-184.

Hauser, John R., and Steven P. Gaskin (1984), "Application of the `Defender' Consumer Model," *Marketing Science*, Vol. 3, No. 4, (Fall), 327-351.

______ and Frank S. Koppelman (1979), "Alternative Perceptual Mapping Techniques: Relative Accuracy and Usefulness, *Journal of Marketing Research*, Vol. 16, No. 4, (November), 495-506.

______ and Steven M. Shugan (1980), "Intensity Measures of Consumer Preference," *Operations Research*, 28, 2, (March-April), 278-320.

______, Alice M. Tybout, and Frank Koppelman (1981), "Consumer-Oriented Transportation Service Planning: "Consumer Analysis and Strategies," *Applications of Management Science*, Vol. 1, 91-138.

______ and Glen L. Urban (1977), "A Normative Methodology for Modeling Consumer Response to Innovation," *Operations Research,* 25, 5, (July-August), 579-619.

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

Huber, Joel (1975), "Predicting Preferences on Experimental bundles of Attributes: A Comparison of Models," *Journal of Marketing Research*, 12, (August), 290-297.

______, Dick R. Wittink, John A. Fiedler, and Richard Miller (1993), "The Effectiveness of Alternative Preference Elicitation Procedures in Predicting Choice," *Journal of Marketing Research*, pp. 105-114.

Jain, Arun K., Franklin Acito, Naresh K. Malhotra, and Vijay Mahajan (1979), "A Comparison of the Internal Validity of Alternative Parameter Estimation Methods in Decompositional Multiattribute Preference Models," *Journal of Marketing Research*, 16, (August), 313-322.

Johnson, Eric J., Robert J. Meyer, and Sanjoy Ghose (1989), "When Choice Models Fail: Compensatory Models in Negatively Correlated Environments," *Journal of Marketing Research*, pp. 255-270.

Johnson, Richard (1987), "Accuracy of Utility Estimation in ACA," Working Paper, Sawtooth Software, Sequim, WA, (April).

______ (1991), "Comment on `Adaptive Conjoint Analysis: Some Caveats and Suggestions," *Journal of Marketing Research*, 28, (May), 223-225.

______ (1999), "The Joys and Sorrows of Implementing HB Methods for Conjoint Analysis," Working Paper, Sawtooth Software, Sequim, WA, (November).

Judge, G. G., W. E. Griffiths, R. C. Hill, H. Lutkepohl, and T. C. Lee (1985), *The Theory and Practice of Econometrics*, (New York, NY: John Wiley and Sons).

Karmarkar, N. (1984), "A New Polynomial Time Algorithm for Linear Programming," *Combinatorica*, 4, 373-395.

Keeney, Ralph and Howard Raiffa (1976), *Decisions with Multiple Consequences: Preferences and Value Trade-offs*, (New York, NY: John Wiley & Sons).

Klein, Noreen M., "Assessing Unacceptable Attribute Levels in Conjoint Analysis," *Advances in Consumer Research* vol. XIV, pp. 154-158.

Krantz, David H., R. Duncan Luce, Patrick Suppes, and Amos Tversky (1971), *Foundations of Measurement*, (New York, NY: Academic Press).

Kuhfeld, Warren F. (1999), "Efficient Experimental Designs Using Computerized Searches," Working Paper, SAS Institute, Inc.

______, Randall D. Tobias, and Mark Garratt (1994), "Efficient Experimental Design with Marketing Research Applications," *Journal of Marketing Research*, (November), 545-557.

Leigh, Thomas W., David B. MacKay, and John O. Summers (1984), "Reliability and Validity of Conjoint Analysis and Self-Explicated Weights: A Comparison," *Journal of Marketing Research*, pp. 456-462.

Lenk, Peter J., Wayne S. DeSarbo, Paul E. Green, and Martin R. Young (1996), "Hierarchical Bayes Conjoint Analysis: Recovery of Partworth Heterogeneity from Reduced Experimental Designs," *Marketing Science*, 15, 2, 173-191.

Liechty, John, Venkatram Ramaswamy, Steven Cohen (2001), "Choice-Menus for Mass Customization: An Experimental Approach for Analyzing Customer Demand With an Application to a Web-based Information Service," *Journal of Marketing Research*, 38, 2, (May).

Louviere, Jordan J., David A. Hensher, and Joffre D. Swait (2000), *Stated Choice Methods: Analysis and Application*, (New York, NY: Cambridge University Press), pp. 354-381.

Mahajan, Vijay and Jerry Wind (1992), "New Product Models: Practice, Shortcomings and Desired Improvements," *Journal of Product Innovation Management*, 9, 128-139.

______ (1986), "An Approach to the Measurement of Consumer Preferences Using Limited Information," *Journal of Marketing Research*, 23, (February), 33-40.

McGrath, Michael E. (1996), *Setting the Pace in Product Development: A guide to Product and Cycle-Time Excellence*, (Boston, MA: Butterworth-Heinemann).

Montgomery, David B. and Dick R. Wittink (1980), "The Predictive Validity of Conjoint Analysis for Alternative Aggregation Schemes," in *Market Measurement and Analysis: Proceedings of the 1979 ORSA/TIMS Conference on Marketing*, David B. Montgomery and Dick R. Wittink, eds, Cambridge, MA: Marketing Science Institute, 298-309.

Moore, William L. (1980), "Levels of Aggregation in Conjoint Analysis: An Empirical Comparison," *Journal of Marketing Research*, pp. 516-523.

______ and Richard J. Semenik (1988), "Measuring Preferences with Hybrid Conjoint Analysis: The Impact of a Different Number of Attributes in the Master Design," *Journal of Business Research*, pp. 261-274.

Nadilo, Rudy (1999), "On-line Research: The Methodology for the Next Millennium," *Advertising Research Foundation Journal*, (Spring). Available at www.greenfield.com.

Nesterov, Y. and A. Nemirovskii (1994), "Interior-Point Polynomial Algorithms in Convex Programming," SIAM, Philadelphia.

Nowlis, Stephen M. and Itamar Simonson (1997), "Attribute-Task Compatibility as a Determinant of Consumer Preference Reversals," *Journal of Marketing Research*, 36, (May), 205-218.

Orme, Bryan (1999), "ACA, CBC, of Both?: Effective Strategies for Conjoint Research," Working Paper, Sawtooth Software, Sequim, WA.

______, M. I. Alpert, and E. Christensen (1997), "Assessing the Validity of Conjoint Analysis – Continued," *Sawtooth Software Conference Proceedings*, 209-225.

Page, Albert L. and Harold F. Rosenbaum (1987), "Redesigning Product Lines with Conjoint Analysis: How Sunbeam Does It," *Journal of Product Innovation Management*, 4, 120-137.

Park, C. Whan, Sung Youl Jun, and Deborah Macinnis (2000), "Choosing What I Want versus Rejecting What I Do Not Want: An Application of Decision Framing to Product Option Choice Decisions," *Journal of Marketing Research*, 37, 2, (May), 187-202.

Parker B. R. and V. Srinivasan (1976), "A Consumer Preference Approach to the Planning of Rural Primary Health-Care Facilities," *Operations Research*, 24, 5, (September-October), 991-1025.

Robinson, P. J. (1980), “Applications of Conjoint Analysis to Pricing Problems,” in *Market Measurement and Analysis: Proceedings of the 1979 ORSA/TIMS Conference on Marketing*, David B. Montgomery and Dick Wittink, eds., (Cambridge, MA: Marketing Science Institute), 183-205.

Sandor, Zsolt and Michel Wedel (2001), “Designing Conjoint Choice Experiments Using Managers’ Prior Beliefs,” *Journal of Marketing Research*, 38, 4, (November), 430-444.

Sawtooth Software, Inc. (1999), “CVA: Full-Profile Conjoint System from Sawtooth Software,” *CVA Manual*, (Sequim, WA: Sawtooth Software, Inc.)

______ (1996), “ACA System: Adaptive Conjoint Analysis,” *ACA Manual*, (Sequim, WA: Sawtooth Software, Inc.)

______ (1999), “The ACA/HB Module for Hierarchical Bayes Estimation,” *ACA/HB Manual*, (Sequim, WA: Sawtooth Software, Inc.)

Sawyer, Alan G. (1975), “Demand Artifacts in Laboratory Experiments in Consumer Research,” *Journal of Consumer Research*, 1. (March), 20-30.

Sheth, Jagdish N. and Wayne W. Talarzyk (1972), “Perceived Instrumentality and Value Importance as Determinants of Attitudes,” *Journal of Marketing Research*, 9, 6-9.

Shimp, Terry, Eva Hyatt, and David Snyder (1991), “A critical Appraisal of Demand Artifacts in Laboratory Experiments in Consumer Research,” *Journal of Consumer Research,* 18, (December), 273-283.

Simmons, Carolyn J., Barbara A. Bickart, and John G. Lynch, Jr. (1993), “Capturing and Creating Public Opinion in Survey Research,” *Journal of Consumer Research*, 30, (September), 316-329.

Smith, Preston G. and Donald G. Reinertsen (1998), *Developing Products in Half the Time, 2E*, (New York, NY: John Wiley & Sons, Inc.).

Sonnevend, G. (1985a), “An ‘Analytic’ Center for Polyhedrons and New Classes of Global Algorithms for Linear (Smooth, Convex) Programming,” *Proceedings of the 12th IFIP Conference on System Modeling and Optimization*, Budapest.

______ (1985b), “A New Method for Solving a Set of Linear (Convex) Inequalities and its Applications for Identification and Optimization,” Preprint, Department of Numerical Analysis, Institute of Mathematics, Eőtvős University, Budapest, 1985.

Srinivasan, V. (1988), “A Conjunctive-Compensatory Approach to The Self-Explication of Multiattributed Preferences,” *Decision Sciences*, pp. 295-305.

______ and Chan Su Park (1997), “Surprising Robustness of the Self-Explicated Approach to Customer Preference Structure Measurement,” *Journal of Marketing Research*, 34, (May), 286-291.

______ and Allen D. Shocker(1973a), “Estimating the Weights for Multiple Attributes in a Composite Criterion Using Pairwise Judgments,” *Psychometrika*, 38, 4, (December), 473-493.

______ and ______ (1973b), “Linear Programming Techniques for Multidimensional Analysis of Preferences,” *Psychometrika*, 38, 3, (September), 337-369.

Tessler, Amy, Norm Wada, and Robert L. Klein (1993), “QFD at PG&E,” *Transactions from The Fifth Symposium on Quality Function Deployment,* (June).

Toubia, Olivier, Duncan Simester, John R. Hauser (2002), “Fast Polyhedral Adaptive Conjoint Estimation,” working paper, Cambridge, MA: Center for Innovation in Product Development, MIT, (January).

Tourangeau, Roger, Lance J. Rips, and Kenneth A. Rasinski (2000), *The Psychology of Survey Response*, (New York, NY: Cambridge University Press).

Tybout, Alice M. and John R. Hauser (1981), "A Marketing Audit Using a Conceptual Model of Consumer Behavior: Application and Evaluation," *Journal of Marketing*, Vol. 45, No. 3, (Summer), 81-101.

Ulrich, Karl T. and Steven D. Eppinger (2000), *Product Design and Development*, (New York, NY: McGraw-Hill, Inc.).

Urban, Glen L. and Gerald M. Katz, "Pre-Test Market Models: Validation and Managerial Implications," *Journal of Marketing Research*, Vol. 20 (August 1983), 221-34.

Vaidja, P. (1989), "A Locally Well-Behaved Potential Function and a Simple Newton-Type Method for Finding the Center of a Polytope," in: N. Megiddo, ed., *Progress in Mathematical Programming: Interior Points and Related Methods*, Springer: New York, 79-90.

Wilkie, William L. and Edgar A. Pessemier (1973), "Issues in Marketing's Use of Multi-attribute Attitude Models," *Journal of Marketing Research*, 10, (November), 428-441.

Willkie, Joseph., Adams, Christine and Azra Girnius (1999), "First Systematic Comparison Between Mall-Intercept and Internet Interviewing," European Society for Opinion and Marketing Research, Worldwide Internet Conference, London UK, (February 21).

Wind, Jerry, Paul E. Green, Douglas Shifflet, and Marsha Scarbrough (1989), "Courtyard by Marriott: Designing a Hotel Facility with Consumer-Based Marketing Models," *Interfaces*, pp. 25-47.

Wittink, Dick R. and Trond Bergestuen (2001), "Forecasting with Conjoint Analysis," *Principles of Forecasting: A Handbook for Researchers and Practitioners*, J. Scott Armstrong, ed., (Norwell, MA: Kluwer Academic Publishers),

______ and Philippe Cattin (1989), "Commercial Use of Conjoint Analysis: An Update," *Journal of Marketing*, pp. 91-96.

______ and David B. Montgomery (1979), "Predictive Validity of Trade-off Analysis for Alternative Segmentation Schemes," *1979 AMA Educators' Conference Proceedings,* Neil Beckwith ed., Chicago, IL: American Marketing Association.

Wright, Peter and Mary Ann Kriewall (1980), "State-of-Mind Effects on Accuracy with which Utility Functions Predict Marketplace Utility," *Journal of Marketing Research*, 17, (August), 277-293.

## Appendix

## Summary of Acronyms Used to Simplify Exposition

| | |
|---|---|
| ACA | Refers to both Adaptive Conjoint Analysis and an experimental condition (see Table 2). |
| ACA HB | Hierarchical Bayes estimation using data from the ACA condition. |
| Fixed | Refers to both an experimental condition (see Table 2) and describes a question-selection method in which the metric paired-comparison questions are chosen by a non-adaptive D-efficient experimental design. |
| FP | Fast Polyhedral Adaptive Conjoint Estimation (FastPace). |
| FP1 | Refers to both an experimental condition (see Table 2) and FastPace estimated using data from that condition. In the FP1 (and Fixed) conditions, the SE and PI questions are not collected prior to the external validity task. |
| FP2 | Refers to both an experimental condition (see Table 2) and FastPace estimated using data from that condition. In FP2 (and ACA) conditions, the SE and PI questions are collected prior to the external validity task. |
| FPSE | FastPace with Self Explicated constraints. |
| HB | Hierarchical Bayes. |
| PC | Metric Paired-Comparison Questions. The respondent provides a rating to indicate his or her strength of preference between two alternatives. All tested methods use PCs. |
| PD | Product Development. |
| PI | Purchase Intention Questions. These questions are used in standard ACA applications. They are only collected in the ACA and FP2 conditions prior to the external validity task. |
| SE | Self-Explicated Questions. The respondent states the importance of a feature directly. They are only collected in the ACA and FP2 conditions prior to the external validity task. |