

# Polyhedral Methods for Adaptive Choice-Based Conjoint Analysis

Olivier Toubia

John R. Hauser

Duncan I. Simester

Revised, February 14, 2003

Olivier Toubia is a graduate student at the Marketing Group, Massachusetts Institute of Technology, E56-345, 38 Memorial Drive, Cambridge, MA 02142, (617) 253-7353, fax (617) 253-7597, [toubia@mit.edu](mailto:toubia@mit.edu).

John R. Hauser is the Kirin Professor of Marketing, Sloan School of Management, Massachusetts Institute of Technology, E56-314, 38 Memorial Drive, Cambridge, MA 02142, (617) 253-2929, fax (617) 253-7597, [jhauser@mit.edu](mailto:jhauser@mit.edu).

Duncan I. Simester is an Associate Professor, Sloan School of Management, Massachusetts Institute of Technology, E56-305, 38 Memorial Drive, Cambridge, MA 02142, (617) 258-0679, fax (617) 253-7597, [simester@mit.edu](mailto:simester@mit.edu).

We gratefully acknowledge the contribution of Robert M. Freund who proposed the use of the analytic center and approximating ellipsoids and gave us detailed advice on the application of these methods.

This research was supported by the Sloan School of Management and the Center for Innovation in Product Development at M.I.T. 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) open source code for the simulations described in this paper, (3) demonstrations of web-based questionnaires based on the methods in this paper, and (4) related papers on web-based interviewing methods. All authors contributed fully and synergistically to this paper. We wish to thank Ray Faith, Aleksas Hauser, Janine Sisk, Limor Weisberg, Toby Woll for the visual design, programming, and project management on the Executive Education Study. This paper has benefited from presentations at the CIPD Spring Research Review, the Epoch Foundation Workshop, the Marketing Science Conferences in Wiesbaden Germany and Alberta Canada, the MIT ILP Symposium on “Managing Corporate Innovation,” the MIT Marketing Workshop, the MIT Operations Research Seminar Series, the MSI Young Scholars Conference, the New England Marketing Conference, and Stanford Marketing Workshop, and the UCLA Marketing Seminar Series.

# **Polyhedral Methods for Adaptive Choice-Based Conjoint Analysis**

## **Abstract**

Choice-based conjoint analysis (CBC) is used widely in marketing for product design, segmentation, and marketing strategy. We propose and test a new “polyhedral” question-design method that adapts each respondent’s choice sets based on previous answers by that respondent. Individual adaptation appears promising because, as demonstrated in the aggregate customization literature, question design can be improved based on prior estimates of the respondent’s partworths – information that is revealed by respondents’ answers to prior questions. The otherwise impractical computational problems of individual CBC adaptation become feasible based on recent polyhedral “interior-point” algorithms, which provide the rapid solutions necessary for real-time computation.

To identify domains where individual adaptation is promising (and domains where it is not), we evaluate the performance of polyhedral CBC methods with Monte Carlo experiments. We vary magnitude (response accuracy), respondent heterogeneity, estimation method, and question-design method in a  $4 \times 2^3$  experiment. The estimation methods are Hierarchical-Bayes estimation (HB) and Analytic-Center estimation (AC). The latter is a new individual-level estimation procedure that is a by-product of polyhedral question design. The benchmarks for individual adaptation are random designs, orthogonal designs, and aggregate customization. The simulations suggest that polyhedral question design does well in many domains, particularly those in which heterogeneity and partworth magnitudes are relatively large. In the comparison of estimation methods, HB is strong across all domains, but AC estimation shows promise when heterogeneity is high.

We close by describing an empirical application to the design of executive education programs in which 354 web-based respondents answered stated-choice tasks with four service profiles each. The profiles varied on eight multi-level features. With the help of this study a major university is revising its executive education programs with new formats and a new focus.

## Introduction

Choice-based conjoint analysis (CBC) describes a class of techniques that are amongst the most widely adopted market research methods. In CBC tasks respondents are presented with two or more product profiles and asked to choose the profile that they prefer (see Figure 1 for an example). This contrasts with other conjoint tasks that ask respondents to provide preference ratings for product attributes or profiles. Because choosing a preferred product profile is often a natural task for respondents consistent with marketplace choice, supporters of CBC have argued that it yields more accurate responses. CBC methods have been shown to perform well when compared against estimates of marketplace demand (Louviere, Hensher, and Swait 2000).

Figure 1

Example of a CBC Task for the Redesign of Polaroid's I-Zone Camera

Question 1 2 3 4 5 6 7 8

Here is a new set of four cameras. Remember: The features of these options have changed. Click on the white square below to tell us which of the cameras, if any, that you would be willing to purchase.

Features	Camera M	Camera N	Camera P	Camera Q
Price	\$34.99	\$24.99	\$34.99	
Picture Removal	Manual	Automatic	Automatic	
Picture Taking	2 Step	1 Step	2 Step	
Styling Covers	Changeable	Permanent	Permanent	None of these cameras.
Picture Quality	Option B	Option A	Option B	
Camera Opening	Slide Open	Fixed	Fixed	
Light Selection	3 settings	Feedback	3 settings	
I would purchase: (click only one)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Click on the feature icons for a reminder.

Help ? Back Next

Important academic research investigating CBC methods has sought to improve the design of the product profiles shown to each respondent. This has led to efficiency improvements, yielding more information from fewer responses. Because an increasing amount of market research is conducted on the Internet, new opportunities for efficiency improvements have arisen. Online processing power makes it feasible to adapt questions based on prior responses. To date the research on adaptive question design has focused on adapting questions based on responses from prior respondents (“aggregate customization”). Efficient designs are customized based on parameters obtained from pretests or from managerial judgment. Examples of aggregate customization methods include Huber and Zwerina (1996), Arora and Huber (2001) and Sandor and Wedel (2001).

In this study we propose a CBC question design method that adapts questions using the previous answers from *that* respondent (“individual adaptation”). The design of each choice task presented to a

respondent varies depending on which profiles that respondent selected in prior choice tasks. The intuitive concept, drawn from the success of aggregate customization, is that the responses to questions reveal information about feasible partworths, which, in turn, enables more efficient questions to be designed. After data are collected with these adaptive questions, partworths can be estimated with standard methods (aggregate random utility or Hierarchical Bayes methods). As an alternative, we propose and test an individual-level estimation method that relies on the analytic center of a feasible set of parameters.

Our proposal differs in both format and philosophy from the only other individual-level adaptive conjoint analysis (ACA) method. We focus on stated-choice data rather than ACA’s metric paired-comparisons and we focus on analogies to efficient design rather than ACA’s utility balance subject to orthogonality goals. Polyhedral methods are also being developed for metric-paired-comparison data, but are beyond the scope of this paper. We discuss them briefly at the end of this paper.

The remainder of the paper is organized as follows. We begin by briefly reviewing existing CBC question design and estimation methods. We next propose a polyhedral approach to the design of CBC questions. We then evaluate the proposed polyhedral methods using a series of Monte Carlo simulations, where we hope to demonstrate the domains in which the proposed method shows promise (and where existing methods remain best). We compare performance against three question design benchmarks, including an aggregate customization method that uses prior data from either managerial judgment or pretest respondents. We compare the four question design methods across a range of customer heterogeneity and response error contexts, while also varying the estimation method. We next describe an empirical application of the proposed method to the design of executive education programs at a major university. The paper concludes with a review of the findings, limitations and opportunities for future research.

### Existing CBC Question Design and Estimation Methods

To date, most applications of CBC assume that each respondent answers the same set of questions or that the questions are either blocked across sets of respondents or chosen randomly. For these conditions, McFadden (1974) shows that the inverse of the covariance matrix,  $\Sigma$ , of the MLE estimates is proportional to:

$$(1) \quad \Sigma^{-1} = R \sum_{i=1}^q \sum_{j=1}^{J_i} (\bar{z}_{ij} - \sum_{k=1}^{J_i} \bar{z}_{ik} P_{ik})' P_{ij} (\bar{z}_{ij} - \sum_{k=1}^{J_i} \bar{z}_{ik} P_{ik})$$

where  $R$  is the effective number of replicates;  $J_i$  is the number of profiles in choice set  $i$ ;  $q$  is the number of choice sets;  $\bar{z}_{ij}$  is a binary vector describing the  $j^{\text{th}}$  profile in the  $i^{\text{th}}$  choice set; and  $P_{ij}$  is the probability that the respondent chooses profile  $j$  from the  $i^{\text{th}}$  choice set. (We use binary vectors in the theoretical development, without loss of generality, to simplify notation and exposition. Multi-level features are used in both the simulations and the application.) We can increase precision by decreasing a measure (norm)

of the covariance matrix, that is, by either increasing the number of replicates or increasing the terms in the summations of Equation 1. Equation 1 also demonstrates that the covariance of logit-based estimates depends on the choice probabilities, which, in turn, depend upon the partworths. In general, the experimental design that provides the most precise estimates will depend upon the parameters.

Many researchers have addressed choice set design. One common measure is D-efficiency, which seeks to reduce the geometric mean of the eigenvalues of  $\Sigma$  (Kuhfield, Tobias, and Garratt 1994).<sup>1</sup> Intuitively, if  $\bar{u}$  represents the vector of the partworths, then the confidence region for maximum likelihood estimates ( $\hat{u}$ ) is an ellipsoid defined by  $(\bar{u} - \hat{u})' \Sigma^{-1} (\bar{u} - \hat{u})$ . The length of the axes of this ellipsoid are given by the eigenvalues of the covariance matrix, so that minimizing the geometric mean of these eigenvalues shrinks the confidence region around the estimates.

Because efficiency depends on the partworths, it is common to assume *a priori* that the stated choices are equally likely. For this paper we will label such designs as “orthogonal” efficient designs. Arora and Huber (2001), Huber and Zwerina (1996), and Sandor and Wedel (2001) demonstrate that we may improve efficiency by using data from either pretests or prior managerial judgment. These researchers improve D-efficiency by “relabeling,” which permutes the levels of features across choice sets, “swapping,” which switches two feature levels among profiles in a choice set, and “cycling,” which is a combination of rotating levels of a feature and swapping them. The procedures stop when no further improvement is possible.<sup>2</sup> Simulations suggest that these procedures improve efficiency and, hence, reduce the number of respondents that are necessary. Following the literature, we label these designs as “aggregate customization”.

### Estimation

In classical logit analysis partworths are estimated with maximum likelihood techniques. Because it is rare that a respondent will be asked to make enough choices to estimate partworth values for each respondent, the data usually are merged across respondents to estimate population-level (or segment-level) partworth values. Yet managers often want estimates for each respondent. Hierarchical Bayes (HB) methods provide (posterior) estimates of partworths for individual respondents by using population-level distributions of partworths to inform individual-level estimates (Allenby and Rossi 1999; Arora, Allenby and Ginter 1998; Johnson 1999; Lenk, et. al. 1996). In particular, HB methods use data from the

---

<sup>1</sup> This is equivalent to maximizing the  $p^{\text{th}}$  root of the determinant of  $\Sigma^{-1}$ . Other norms include A-efficiency, which maximizes the trace of  $\Sigma^{-1}/p$ , and G-efficiency, which maximizes the maximum diagonal element of  $\Sigma^{-1}$ .

<sup>2</sup> The Huber-Zwerina and Arora-Huber algorithms maximize  $\det \Sigma^{-1}$  based on the mean partworths and do so by assuming the mean is known from pretest data (or managerial judgment). Sandor and Wedel (2001) include, as well, a prior covariance matrix in their calculations. They then maximize the expectation of  $\det \Sigma^{-1}$ , where the expectation is over the prior subjective beliefs.

full sample to iteratively estimate both the posterior means (and distribution) of individual-level partworths and the posterior distribution of those partworths at the population-level. The HB method is based on Gibbs sampling and the Metropolis Hastings Algorithm.

Liechty, Ramaswamy, and Cohen (2001) demonstrate the effectiveness of HB for choice menus, while Arora and Huber (2001) show that it is possible to improve the efficiency of HB estimates with choice-sets designed using Huber-Zwerina relabeling and swapping improve. In other recent research, Andrews, Ainslie, and Currim (2002, p. 479) present evidence that HB models and finite mixture models estimated from simulated scanner-panel data “recover household-level parameter estimates and predict holdout choice about equally well except when the number of purchases per household is small.”

### Polyhedral Question Design Methods

We extend the philosophy of customization by developing algorithms to adapt questions for each respondent. Stated choices by each respondent provide information about parameter values for that respondent that can be used to select the next question(s). In high dimensions (high  $p$ ) this is a difficult dynamic optimization problem. We address this problem by making use of extremely fast algorithms based on projections within the interior of polyhedra (much of this work started with Karmarkar 1984). In particular, we draw on the properties of bounding ellipsoids discovered in theorems by Sonnevend (1985a, 1985b) and applied by Freund (1993), Nesterov and Nemirovskii (1994), and Vaidya (1989).

We begin by illustrating the intuitive ideas in a two-dimensional space with two product profiles ( $J_i = 2$ ) and then generalize to a larger number of dimensions and multichotomous choice. The axes of the space represent the partworths (utilities) associated with two different product attributes,  $u_1$  and  $u_2$ . A point in this space has a value on each axis, and will be represented by a vector of these two partworths. The ultimate goal is to estimate the point in this space (or distribution of points) that best represents each respondent. The question-design goal is to focus precision toward the points that best represent each respondent. This goal is not unlike D-efficiency, which seeks to minimize the confidence region for estimated partworths. Without loss of generality we scale all partworths in the figures to be non-negative and bounded from above. Following convention, the partworth associated with the least-preferred level is set arbitrarily to zero.<sup>3</sup>

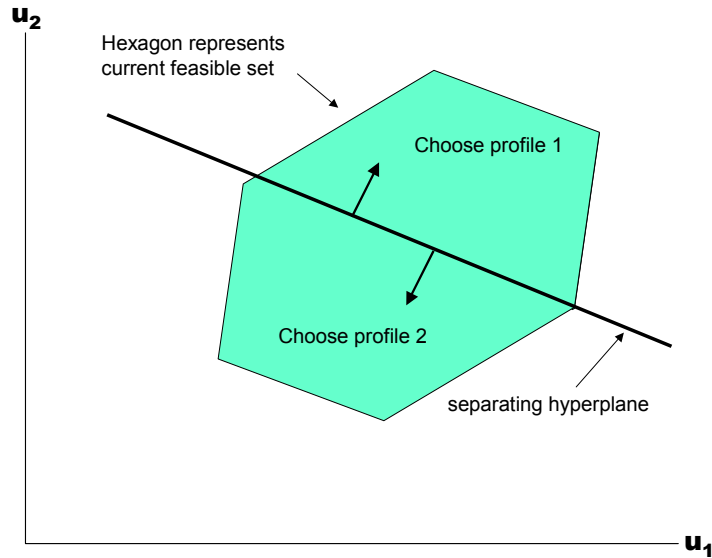
Suppose that we have already asked  $i-1$  stated-choice questions and suppose that the hexagon (polyhedron) in Figure 2 represents the partworth vectors that are consistent with the respondent’s answers. Suppose further that the  $i^{\text{th}}$  question asks the respondent to choose between two profiles with feature levels  $\bar{z}_{i1}$  and  $\bar{z}_{i2}$ . If there were no response errors, then the respondent would select Profile 1 whenever  $(z_{i11} - z_{i21})u_1 + (z_{i12} - z_{i22})u_2 \geq 0$ ; where  $z_{ijf}$  refers to the  $f^{\text{th}}$  feature of  $z_{ij}$  and  $u_f$  denotes the partworth

---

<sup>3</sup> In our application warm-up questions identify the lowest level of each feature. Such questions are common.

associated with the  $f^{\text{th}}$  feature. This inequality constraint defines a separating line or, in higher dimensions, a separating hyperplane. In the absence of response errors, if the respondent's true partworth vector is above the separating hyperplane the respondent chooses Profile 1; when it is below the respondent chooses Profile 2. Thus, the respondent's choice of profiles updates our knowledge of which partworth vectors are consistent with the respondent's preferences, shrinking the feasible polyhedron.

**Figure 2**  
**Stated Choice Responses Divide the Feasible Region**



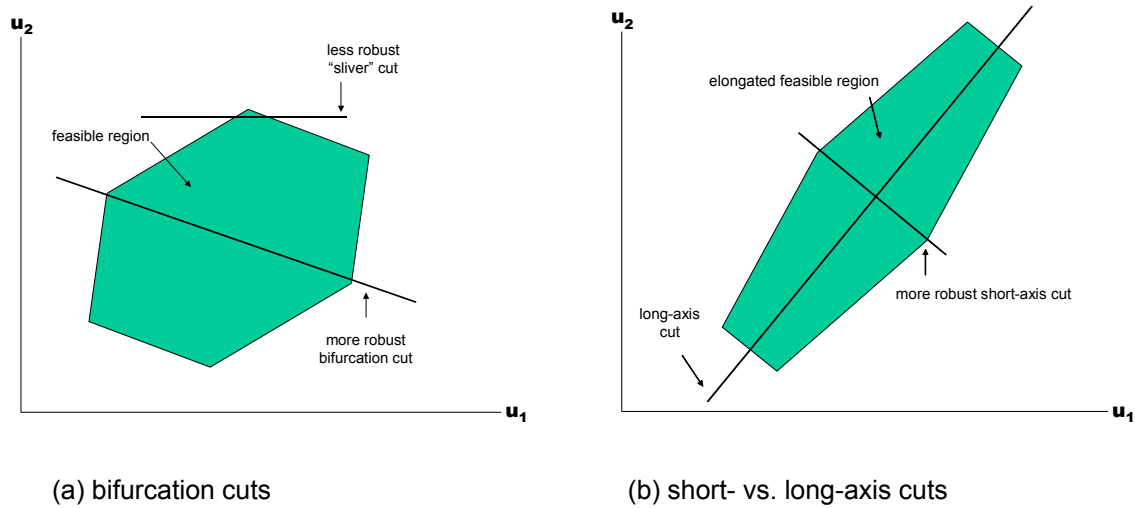
**Selecting Questions**

We seek questions (and the regions associated with each choice) so that we gain the maximum information from each question in order to reduce the feasible region rapidly. We implement this goal with four criteria. First, neither profile in the choice set should dominate the other profile for all points in the feasible region. Otherwise, we gain no information when the dominating profile is chosen and the partworth space is not reduced. Second, the separating hyperplane should intersect with and divide the feasible region as derived from the first  $i-1$  questions. Otherwise, there could be an answer to the  $i^{\text{th}}$  question that does not reduce the feasible region. A corollary of these criteria is that, for each profile, there must be a point in the feasible region for which that profile is the preferred profile.

The  $i^{\text{th}}$  question is more informative if, given the first  $i-1$  questions, the respondent is equally likely to select each of the  $J_i$  profiles. This implies that, a priori, the answer to the  $i^{\text{th}}$  question should be approximately equally likely – a criterion we call “choice balance.” Choice balance will shrink the feasible region as rapidly as feasible. For example, if the points in the feasible region are equally likely (based on  $i-1$  questions) then the predicted likelihood,  $\pi_{ij}$ , of choosing the  $j^{\text{th}}$  region is proportional to the size of

the region. The expected size of the region after the  $i^{\text{th}}$  question is then proportional to  $\sum_{j=1}^2 \pi_{ij}^2$ , which is minimized when  $\pi_{ij} = 1/2$ .<sup>4</sup> This criterion will hold approximately if we favor separating hyperplanes, which go through the center of the feasible polyhedron and cut the feasible region approximately in half. This is illustrated in Figure 3a where we favor bifurcation cuts relative to “sliver” cuts that yield unequal sized regions. Notice that if the separating hyperplane is a bifurcation cut, both the non-domination and feasibility criteria are satisfied automatically.

**Figure 3**  
**Comparing Cuts and the Resulting Feasible Regions**



However, not all bifurcation cuts are equally robust. Suppose that the current feasible region is elongated, as in Figure 3b, and we must decide between many separating hyperplanes, two of which are illustrated. One cuts along the long axis and yields long thin feasible regions while the other cuts along the short axis and yields feasible regions that are relatively more symmetric. The long-axis cut focuses precision where we already have high precision while the short-axis cut focuses precision where we now have less precision. For this reason, we prefer short-axis cuts to make the post-choice feasible regions reasonably symmetric. We can also motivate this criterion relative to D-efficiency. D-efficiency minimizes the geometric mean of the axes of the confidence ellipsoid – a criterion that tends to make the confidence ellipsoids more symmetric.

For two profiles we satisfy non-dominance, feasibility, choice balance, and post-choice symmetry if we select profiles such that (a) the separating hyperplanes go through the center of the feasible region

<sup>4</sup> For  $J_i$  profiles, equally-sized regions also maximize entropy, defined as  $-\sum_j \pi_{ij} \log \pi_{ij}$ . Formally, maximum entropy is equal to the total information obtainable in a probabilistic model (Hauser 1978, Theorem 1, p. 411).



and (b) the separating hyperplanes are perpendicular to the longest “axis” of the feasible polyhedron as defined by the first  $i-1$  stated choices. To implement these criteria we propose the following heuristic algorithm.

**Step 1** Find the center and the longest axis of the polyhedron based on  $i-1$  questions.

**Step 2** Find the two partworth vectors that intersect with the longest axis and the boundary of the polyhedron. A separating hyperplane formed by these two vectors would be perpendicular to the longest axis. However, respondents choose among profiles; they do not choose directly among utility vectors. Thus,

**Step 3** For each of the two partworth vectors, select a profile corresponding to each partworth vector such that the separating hyperplane formed by the two profiles is approximately perpendicular to the longest axis. We select each profile such that the respondent chooses that profile when the respondent maximizes utility given those partworths.

We address the respondent’s utility maximization problem below after we discuss how this heuristic algorithm extends to  $J_i > 2$ .

### Selecting More than Two Profiles

In a choice task with more than two profiles the respondent’s choice defines more than one separating hyperplane. The hyperplanes that define the  $i^{\text{th}}$  feasible region depend upon the profile chosen by the respondent. For example, consider a choice task with three product profiles, labeled 1, 2 and 3. If the respondent selects Profile 1, then we learn that the respondent prefers Profile 1 to Profiles 2 and Profile 1 to Profile 3. This defines two separating hyperplanes – the resulting polyhedron of feasible partworths is the intersection of the associated regions and the prior feasible polyhedron. In general,  $J_i$  profiles yield  $J_i(J_i-1)/2$  possible hyperplanes. For each of the  $J_i$  choices available to the respondent,  $J_i-1$  hyperplanes contribute to the definition of the new polyhedron. The full set of hyperplanes, and their association with stated choices, define a set of  $J_i$  regions, one associated with each answer to the stated-choice question. These regions represent a collectively exhaustive and mutually exclusive division of the partworth space (except for the regions’ “indifference” borders, which have zero measure).

We extend Steps 1 to 3 as follows. Rather than finding the longest axis, we find the  $(J_i/2)$  longest axes and identify the  $J_i$  points where the  $(J_i/2)$  longest axes intersect the polyhedron. (If  $J_i$  is odd, we select randomly among the points intersecting the  $(J_i/2)^{\text{th}}$  longest axis.) We associate profiles with each of the  $J_i$  partworth vectors by solving the respondent’s maximization problem for each vector. The separating hyperplanes still divide the feasible region into  $J_i$  collectively exhaustive and mutually exclusive sub-

regions of approximately equal size. Non-dominance and feasibility remain satisfied and the resulting regions tend toward symmetry.<sup>5</sup>

### Implementation

Implementing this heuristic raises challenges. Although it is easy to visualize (and implement) the heuristic with two profiles in two dimensions, practical CBC problems require implementation with  $J_i$  profiles in large  $p$ -dimensional spaces with  $p$ -dimensional polyhedra and  $(p-1)$ -dimensional hyperplane cuts. Furthermore, the algorithm should run sufficiently fast so that there is little noticeable delay between questions.

The first challenge is finding the center of the current polyhedron and the  $J_i/2$  longest axes (Step 1). If we define the longest “axis” of a polyhedron as the longest line segment in the polyhedron, then we would need to enumerate all vertices of the polyhedron and compute the distances between the vertices. Unfortunately, for large  $p$  this problem is computationally intractable (Gritzmann and Klee 1993); solving it would lead to lengthy delays between questions for each respondent. Furthermore, this definition of the longest axes of a polyhedron may not capture the intuitive concepts that we used to motivate the algorithm.

Instead, we turn to Sonnevend’s theorems (1985a, 1985b) which state that the shape of polyhedra can be approximated with bounding ellipsoids centered at the “analytic center” of the polyhedron. The analytic center is the point that maximizes the geometric mean of the distances to the boundaries. Freund (1993) provides efficient algorithms to find the analytic centers of polyhedra. Once the analytic center is found, Sonnevend’s results provide analytic expressions for the ellipsoids. The axes of ellipsoids are well-defined and capture the intuitive concepts in the algorithm. The longest axes are found with straightforward eigenvalue computations for which there are many efficient algorithms. With well-defined axes it is simple to find the partworth vectors on the boundaries of the feasible set that intersect the axes (Step 2). We provide technical details in an Appendix.

To implement Step 3 we must define the respondent’s utility maximization problem. We do so in an analogy to economic theory. For each of the  $J_i$  utility vectors on the boundary of the polyhedron we obtain the  $j^{\text{th}}$  profile,  $\bar{z}_{ij}$ , by solving:

$$(OPT1) \quad \max \bar{z}_{ij} \bar{u}_{ij} \quad \text{subject to: } \bar{z}_{ij} \bar{c} \leq M, \quad \text{elements of } \bar{z}_{ij} \in \{0,1\}$$

---

<sup>5</sup> Because the separating hyperplanes are defined by the profiles associated with the partworth vectors (Step 3) not the partworth vectors themselves (Step 2), the hyperplanes do not always line up perfectly with the axes. For  $J_i > 2$  the stated properties remain approximately satisfied based on “wedges” formed by the  $J_i-1$  hyperplanes. It is an empirical question whether the proposed heuristic is effective (we use  $J_i = 4$  in the simulations and application).

where  $\bar{u}_{ij}$  is the utility vector chosen in Step 2,  $\bar{c}$  are “costs” of the features, and  $M$  is a “budget constraint.” We implement (approximate) choice balance by setting  $\bar{c}$  equal to the analytic center of the feasible polyhedron,  $\bar{u}_{i-1}$ , computed based on  $i-1$  questions. At optimality the constraint in OPT1 will be approximately binding, which implies that  $\bar{z}_{ij}\bar{u}_{i-1} \cong \bar{z}_{ik}\bar{u}_{i-1} \cong M$  for all  $k \neq j$ . It may be approximate due to the integrality constraints in OPT1 ( $\bar{z}_{ij} \in \{0,1\}$ ). This assures that the separating hyperplanes go (approximately) through the analytic center, which implies that the regions are approximately equal in size.

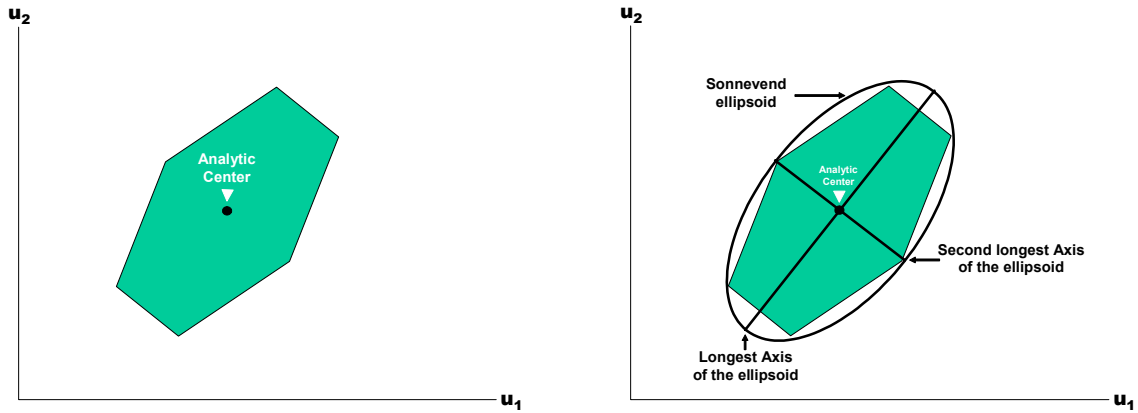
Solving OPT1 for profile selection (Step 3) is a knapsack problem which is well-studied and for which efficient algorithms exist.  $M$  is an arbitrary constant that we draw randomly from a compact set (up to  $m$  times) until all profiles in a stated-choice task are distinct. If the profiles are not distinct we use those that are distinct. If none of the profiles are distinct then we ask no further questions (in practice this is a rare occurrence in both simulation and empirical situations).

OPT1 also illustrates the relationship between choice balance and utility balance – a criterion in aggregate customization. In our algorithm, the  $J_i$  profiles are chosen to be equally likely based on data from the first  $i-1$  questions. In addition, for the partworths at the analytic center of the feasible region the utilities of all profiles are approximately equal. However, utility balance only holds at the analytic center, not throughout the feasible region.

We illustrate the algorithm for  $J_i = 4$  with the two-dimensional example in Figure 4. We begin with the current polyhedron of feasible partworth vectors (in Figure 4a). We then use Freund’s algorithm to find the analytic center of the polyhedron as illustrated by the black dot in Figure 4a. We next use Sonnevend’s formulae to find the equation of the approximating ellipsoid and obtain the  $J_i/2$  longest axes (Figure 4b), which correspond to the  $J_i/2$  smallest eigenvalues of the matrix that defines the ellipsoid. We then identify  $J_i$  target partworth vectors by finding the intersections of the  $J_i/2$  axes with the boundaries of the current polyhedron (Figure 4c). Finally, for each target utility vector we solve OPT1 to identify  $J_i$  product profiles. When the respondent chooses a profile this implies a new smaller polyhedron. We continue for  $q$  questions or until OPT1 no longer yields distinct profiles.

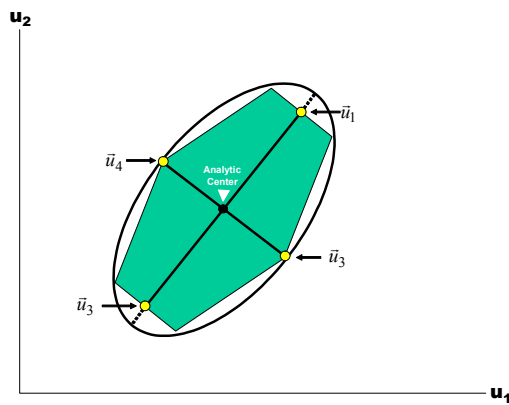
**Figure 4**

**Bounding Ellipsoids and the Analytic Center of the Polyhedra**



(a) Find the analytic center

(b) Find Sonnevend's ellipsoid and axes



(c) Find partworth values on boundary of the polyhedron

**Incorporating Managerial Constraints and other Prior Information**

Previous research suggests that prior constraints enhance estimation (Johnson 1999; Srinivasan and Shocker 1973). For example, self-explicated data might constrain the rank order of partworth values across features. Such constraints are easy to incorporate and shrink the feasible polyhedron. In addition, most conjoint analysis studies use multi-level features, some of which are ordinal scaled (e.g., picture quality). For example, if  $u_{fm}$  and  $u_{fh}$  are the medium and high levels of feature  $f$ , we can add the con-

straint,  $u_{fm} \leq u_{fn}$ , to the feasible polyhedron.<sup>6</sup> We can similarly incorporate information from managerial priors or pretest studies.

### Response Errors

In real questionnaires there is likely to be response error in stated choices. In the case of response error, the separating hyperplanes are approximations rather than deterministic cuts. For this and other reasons, we distinguish question selection and estimation. The algorithm we propose is a question-selection algorithm. After the data are collected we can estimate the respondents' partworths with most established methods, which address response error formally. For example, polyhedral questions can be used with classical random-utility models and HB estimation. It remains an empirical question as to whether or not response errors counteract the potential gains in question selection due to individual-level adaptation. Although we hypothesize that the criteria of choice balance and symmetry lead to robust stated-choice questions, we also hypothesize that individual-level adaptation will work better when response errors are smaller. We examine these issues in the next section.

### Analytic-Center (AC) Estimation

In addition to estimation with established methods, the analytic center of the  $i^{\text{th}}$  feasible polyhedron provides a natural summary of the information in the first  $i$  stated-choice responses. This summary measure is a good working estimate of the respondent's partworth vector. It is a natural by-product of the question-selection algorithm and is available as soon as each respondent completes the  $i^{\text{th}}$  stated-choice question. Such estimates might also be used as starting values in HB estimation, as estimates in classical Bayes updating, and as priors for aggregate customization. We hypothesize that AC estimates are more likely to be accurate when response errors are lower.

Analytic-Center estimates also give us a means to test the ability of the polyhedral algorithm to implement the feasibility and choice-balance criteria. Specifically, if we use the  $i^{\text{th}}$  AC estimate to forecast choices for  $q > i$  choice sets, it should predict 100% of the first  $i$  choices (feasibility) and  $(1/J_i)$  percent of the last  $q - i$  choices (choice balance). That is, the internal predictive percentage should approximately equal  $[i + (1/J_i)(q-i)]/q$ . We examine this statistic in the empirical application later in the paper.

Finally, we can give the AC estimate a statistical interpretation if we assume that the probability of a feasible point is proportional to its distance to the boundary of the feasible polyhedron. In this case, the analytic center maximizes the likelihood of the point (geometric mean of the distances to the boundary).

---

<sup>6</sup> In the theoretical derivation we used binary features without loss of generality for notational simplicity. An ordinal multi-level feature constraint is mathematically equivalent to a constraint linking two binary features.

### Incorporating Null Profiles

Many researchers prefer to include a null profile as an additional profile in the choice set (as in Figure 1). Polyhedral concepts generalize readily to include null profiles. If the null profile is selected from choice set  $i$  then we can add the following constraints:  $\bar{z}_{ij}\bar{u} \leq \bar{z}_k^*\bar{u} \quad \forall j, k \neq i$  where  $\bar{z}_k^*$  denotes the profile chosen from choice set  $k$  (given that the null profile was not chosen in choice set  $k$ ). Intuitively, these constraints recognize that if the null profile is selected in one choice set, then all of the alternatives in that choice set have a lower utility than the profiles selected in other choice sets (excluding other choice sets where the null was chosen). Alternatively, we can expand the parameter set to include the partworth of an “outside option” and write the appropriate constraints. After incorporating these constraints the question design heuristic (and analytic-center estimation) can proceed as described above. However, we leave practical implementation, Monte Carlo testing, and empirical applications with null profiles to future research.

### Summary

Polyhedral (ellipsoid) algorithms provide a feasible means to adapt stated-choice questions for each respondent based on that respondent’s answers to the first  $i-1$  questions. The algorithms are based on the intuitive criteria of non-dominance, feasibility, choice balance, and symmetry and represent an individual-level analogy to D-efficiency. Specifically, the polyhedral algorithm focuses questions on what is not known about the partworth vectors and does so by seeking a small feasible region. D-efficiency seeks to ask questions so that the confidence region for maximum-likelihood estimates,

$(\bar{u} - \hat{u})'\Sigma^{-1}(\bar{u} - \hat{u})$ , is as small as possible. Both criteria are based on the size and shape of the ellipsoids that define the partworth regions – feasible region or confidence region.

While both polyhedral question design and aggregate customization are compatible with most estimation methods, including AC estimation, the two methods represent a key tradeoff. Polyhedral question design adapts questions for each respondent but may be sensitive to response errors. Aggregate customization uses the same design for all respondents, but is based on prior statistical estimates that take response errors into account. This leads us to hypothesize that polyhedral methods will have their greatest advantages relative to existing methods (question design and/or estimation) when responses are more accurate and/or when respondents’ partworths are more heterogeneous.

In the next section we use Monte Carlo experiments to examine whether individually-adapted questions provide more precise estimates of the partworth vectors in some domains. We also examine whether or not AC estimation has advantages in some domains.

## Monte Carlo Experiments

We use Monte Carlo experiments to investigate whether polyhedral methods show sufficient promise to justify further development and to identify the empirical domains in which the potential is greatest. Monte Carlo experiments are widely used to evaluate conjoint analysis methods, including studies of interactions, robustness, continuity, attribute correlation, segmentation, new estimation methods, and new data-collection methods. In particular, they have proven particularly useful in the first tests of aggregate customization and in establishing domains in which aggregate customization is preferred to orthogonal designs. Monte Carlo experiments offer several advantages for an initial test of new methods. First, with any heuristic, we need to establish computational feasibility. Second, Monte Carlo experiments enable us to explore many domains and control the parameters that define those domains. Third, other researchers can readily replicate and extend Monte Carlo experiments, facilitating further exploration and development. Finally, Monte Carlo experiments enable us to control the “true” partworth values, which are unobserved in studies with actual consumers.

However, Monte Carlo experiments are but the first step in a stream of research. Assumptions must be made about characteristics that are not varied, and these assumptions represent limitations. In this paper we explore domains that vary in terms of respondent heterogeneity, response accuracy (magnitude), estimation method, and question-design method. This establishes a  $4 \times 2^3$  experimental design. We hope that subsequent researchers will extend the findings by varying other characteristics of the experiments.

### Structure of the Simulations

For consistency with prior simulations, we adopt the basic simulation structure of Arora and Huber (2001). Arora and Huber varied response accuracy, heterogeneity, and question-design method in a  $2^3$  experiment using Hierarchical Bayes (HB) estimation. Huber and Zwerina (1996) had earlier used the same structure to vary response accuracy and question-design with classical estimation, while more recently Sandor and Wedel (2001) used a similar structure to compare the impact of prior beliefs.

The Huber - Zwerina, and Arora - Huber algorithms were aggregate customization methods based on relabeling and swapping. These algorithms work best for stated-choice problems in which relabeling and swapping is well-defined. We expand the Arora and Huber design to include four levels of four features for four profiles, which ensures that complete aggregate customization and orthogonal designs are possible. Sandor and Wedel included cycling, although they note that cycling is less important in designs where the number of profiles equals the number of feature levels.<sup>7</sup>

---

<sup>7</sup>In the designs that we use, the efficiency of the Sandor and Wedel algorithm is approximately equal to the efficiency of the Huber and Zwerina algorithm.

Within a feature Arora and Huber choose partworths symmetrically with expected magnitudes of  $-\bar{\beta}$ , 0, and  $+\bar{\beta}$ . They vary response accuracy by varying  $\bar{\beta}$ . Larger  $\bar{\beta}$ 's imply higher response accuracy because the variance of the Gumbel distribution, which defines the logit model, is inversely proportional to the squared magnitude of the partworths (Ben-Akiva and Lerman 1985, p. 105, property 3). For four levels we retain the symmetric design with magnitudes of  $-\bar{\beta}$ ,  $-\frac{1}{3}\bar{\beta}$ ,  $\frac{1}{3}\bar{\beta}$ , and  $\bar{\beta}$ . Arora and Huber model heterogeneity by allowing partworths to vary among respondents according to normal distributions with variance,  $\sigma_{\beta}^2$ . They specify a coefficient of heterogeneity as the ratio of the variance to the mean. Specifically, they manipulate low response accuracy with  $\bar{\beta} = 0.5$  and high response accuracy with  $\bar{\beta} = 1.5$ . They manipulate high heterogeneity with  $\sigma_{\beta}^2 / \bar{\beta} = 2.0$  and low heterogeneity with  $\sigma_{\beta}^2 / \bar{\beta} = 0.5$ . Given these values, they draw each respondent's partworths from a normal distribution with a diagonal covariance matrix. Each respondent then answers the stated-choice questions with probabilities determined by a logit model based on that respondent's partworths. Arora and Huber compare question selection using root mean squared error (RMSE) and so we adopt the same criterion.

We select magnitudes and heterogeneity that represent the range of average partworths and heterogeneity that we might find empirically. While we could find no meta-analyses for these values, we did have available to us data from a proprietary CBC application (D-efficient design, HB estimation) in the software market. The study included data from almost 1,200 home consumers and over 600 business customers. In both data sets  $\bar{\beta}$  ranged from approximately  $-3.0$  to  $+2.4$ . We chose our high manipulation (3.0) from this study recognizing that other studies might have even higher magnitudes. For example, Louviere, Hensher and Swait (2000) report stated-choice estimates (logit analysis) that are in the range of 3.0 and higher.

After selecting  $\bar{\beta}$  for high magnitudes, we set the low magnitude  $\bar{\beta}$  to the level chosen by Arora and Huber. In the empirical data, the estimated variances ranged from 0.1 to 6.9 and the heterogeneity coefficient varied from 0.3 to 3.6.<sup>8</sup> To approximate this range and to provide symmetry with the magnitude coefficient, we manipulated high heterogeneity with a coefficient of three times the mean. Following Arora and Huber we manipulated low heterogeneity as half the mean. We feel that these values are representative of those that might be obtained in practice. Recall that, as a first test of polyhedral methods, we seek to identify domains that can occur in practice and for which polyhedral methods show promise. More importantly, these levels illustrate the directional differences among methods and, hence, provide insight for further development.

---

<sup>8</sup> There was also an outlier with a mean of 0.021 and a variance of 0.188 implying a heterogeneity coefficient of 9.0. Such cases are clearly possible, but less likely to represent typical empirical situations.



## Experimental Design

In addition to manipulating magnitude (two levels) and heterogeneity (two levels) we manipulate estimation method (two levels), and question-design method (four levels). The estimation methods are Hierarchical Bayes and Analytic Center estimation. The question-design methods are random, orthogonal designs with equally-likely priors, aggregate customization (Arora and Huber), and polyhedral methods. To simulate aggregate customization, we assume that the pretest data are obtained costlessly and, based on this data, we apply the Arora - Huber algorithm. Specifically, we simulate an orthogonal “pretest” that uses the same number of respondents as in the actual study.

We set  $q = 16$  so that orthogonal designs, relabeling, and swapping are well-defined. Exploratory simulations suggest that the estimates become more accurate as we increase the number of questions, but the relative comparisons of question design and estimation for  $q = 8$  and for  $q = 24$  provide similar qualitative insights.<sup>9</sup>

## Practical Implementation Issues

In order to implement the polyhedral algorithm we made two implementation decisions: (1) We randomly drew  $M$  up to thirty times ( $m = 30$ ) for the simulations. We believe that the accuracy of the method is relatively insensitive to this decision. (2) Because, prior to the first question, the polyhedron is symmetric, we selected the first question by randomly choosing from amongst the axes.

Other decisions may yield greater (or lesser) accuracy, hence the performance of the polyhedral methods tested in this paper should be considered a lower bound on what is possible with further improvement. For example, future research might use aggregate customization to select the first question. All optimization, question selection, and estimation algorithms are described in the Appendix and implemented in Matlab code. The web-based application described later in this paper uses Perl and Html for web-page presentation. All code is available at the website listed in the acknowledgements section of this paper and is open-source.

## Comparative Results of the Monte Carlo Experiments

For comparability between estimation methods we normalize the partworths to a constant scale.<sup>10</sup> Specifically, for each respondent, we normalize both the “true” partworths and the estimated partworths

---

<sup>9</sup> The estimates at  $q = 16$  are approximately 25% more accurate than those at  $q = 8$  and the estimates at  $q = 24$  are approximately 12% more accurate than those at  $q = 16$ .

<sup>10</sup> Analytic-center estimation is unique to a positive linear transformation and thus focuses on the relative values of the partworths – as required by many managerial applications. To extend analytic-center estimates to revealed preference models and volumetric forecasts, we recommend the methods proposed by Louviere, Hensher, and Swait (2000). These methods have proven accurate, are well-documented, and do not change the relative partworths of the stated-choice models. Hence, they are not the focus of this paper.

so that their absolute values sum to the number of parameters and their values sum to zero for each feature. In this manner, the RMSEs can be interpreted as a percent of the mean partworths. Within an estimation method, subject to statistical confidence, this scaling does not change the relative comparisons among question design methods. This scaling has the added advantage of making the results roughly comparable in units for the different manipulations of magnitude (response accuracy) and heterogeneity. Table 1 reports the RMSEs for the simulation experiment using a table format similar to Arora and Huber. We use *italic bold* text to indicate the best question design method within an experimental cell and any other methods that are not statistically different from the best method.

**Table 1**  
**Comparison of Question-Design and Estimation Methods**

Magnitude	Heterogeneity	Question Design Method	RMSE of Normalized Partworths	
			Hierarchical Bayes Estimates	Analytic Center Estimates
Low	High	random	<i><b>0.898<sup>†</sup></b></i>	1.090
		orthogonal	<i><b>0.902<sup>‡</sup></b></i>	0.939
		customized	<i><b>0.876<sup>†</sup></b></i>	1.011
		polyhedral	0.943	<i><b>0.874<sup>†</sup></b></i>
Low	Low	random	1.016	1.202
		orthogonal	<i><b>0.972<sup>†</sup></b></i>	1.015
		customized	1.034	1.078
		polyhedral	1.008	1.069
High	High	random	0.615	0.811
		orthogonal	0.788	0.846
		customized	0.576	0.875
		polyhedral	0.553	<i><b>0.522<sup>†</sup></b></i>
High	Low	random	0.453	0.885
		orthogonal	0.761	0.797
		customized	0.637	1.004
		polyhedral	<i><b>0.422<sup>†</sup></b></i>	0.578

<sup>†</sup> Best or not significantly different than best at  $p < 0.05$ .

<sup>‡</sup> Best or not significantly different than best at  $p < 0.10$ .

The findings in Table 1 suggest that relative to random, orthogonal and customized designs, polyhedral question design provides more accurate estimates for high heterogeneity and/or high magnitudes. These are the domains where we hypothesized individual-level customization to have the most potential – situations where respondents provide relatively accurate data and/or where respondent partworths are relatively heterogeneous. For low magnitude and low heterogeneity, orthogonal designs are more accurate.

The comparison between the customization and orthogonal methods directionally replicate the findings reported by Arora and Huber in their simulations. Aggregate customization methods do well when magnitudes and/or heterogeneity are high. We conclude that in these contexts, aggregate customization also makes effective use of managerial priors or pretest data to improve question design.

When comparing estimation methods, HB estimation does well in all cells. However, AC estimation paired with polyhedral question design provides accurate estimates relative to existing methods for heterogeneous populations (tied for low magnitudes). Like any new method, we expect that researchers will improve AC estimation through experimentation and innovation.

### **Improving Accuracy with a Hybrid Method Based on Classical Bayes Methods**

The evidence that aggregate customization methods can improve accuracy by making effective use of managerial priors or pretest data suggests that it may be possible to improve AC estimation by making use of similar data. A variety of approaches are available to incorporate prior information. We use a classical Bayesian approach that treats the distribution of partworths across the population as a prior and the pure AC estimates as observations. The classical Bayes estimate is then the posterior mean – a convex combination of the prior mean and the pure AC estimate. The relative weights are inversely proportional to the estimated variances. This implies that the weight on the mean is greater when the population is homogeneous and less when the population is heterogeneous. The classical Bayes formulae also imply that the weight on the pure-AC estimate is higher when responses are more accurate and lower when they are less accurate.<sup>11</sup> As an illustration, Table 2 reports RMSE for this AC-hybrid when using polyhedral question selection. The results are promising. Use of the hybrid improves accuracy across all domains.

---

<sup>11</sup> To obtain the weights we use the population variance and the error variance obtained from Table 1. Other algorithms to select weights also demonstrate significant improvement relative to pure AC estimation. We could also apply classical Bayes methods to the HB estimates. However, as expected, the improvement is significantly less because the HB estimates already use population data effectively.

**Table 2**  
**Classical Bayesian AC-Hybrid Estimation**

Magnitude	Heterogeneity	RMSE of Normalized Partworths		
		Hierarchical Bayes Estimates	Analytic Center Estimates	Bayesian AC Estimates
Low	High	0.943	0.874	<b>0.847<sup>†</sup></b>
Low	Low	1.008	1.069	<b>0.858<sup>†</sup></b>
High	High	0.553	<b>0.522<sup>†</sup></b>	<b>0.497<sup>†</sup></b>
High	Low	0.422	0.578	<b>0.400<sup>†</sup></b>

<sup>†</sup> Best or not significantly different than best at  $p < 0.05$ .

### Summary of Monte Carlo Experiments

We can summarize the results of the Monte Carlo experiments as follows:

- If heterogeneity and magnitudes are both low, it may be best not to customize designs; fixed orthogonal questions appear to be the most accurate.
- Polyhedral question design shows the most promise when magnitudes and customer heterogeneity are high.
- HB estimation does well in all domains. AC estimation provides a promising alternative when heterogeneity is high.
- An AC-hybrid estimation method based on classical Bayesian methods has the potential to provide more accurate methods by combining population-level data with individual-level estimates.

Like many new technologies, we hope the methods will improve further with use, experimentation, and evolution (Christensen 1998).

### Application to the Design of Executive Education Programs

Polyhedral methods for CBC have been implemented in at least one empirical application. We briefly describe this application as evidence that it is feasible to implement the proposed methods in applications using actual respondents. We also use the data to obtain empirical estimates of magnitude and heterogeneity and to compare the consistency of the AC and HB estimates. Unfortunately the managerial environment required that we only implement a single question design method rather than compare question-design methods in this first “proof-of-concept” application.

### **Managerial Goals of the Conjoint Study**

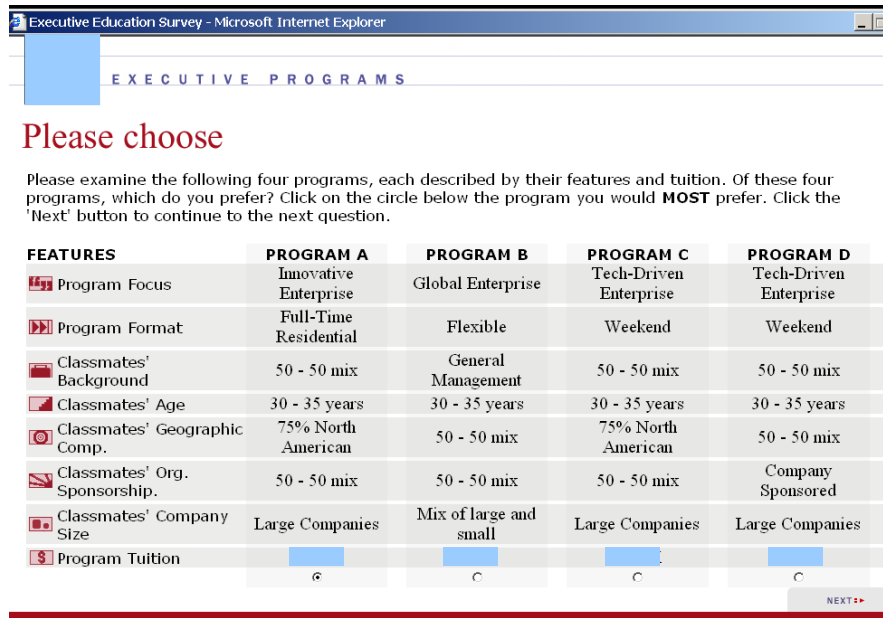
The application supported the design of new executive education programs for a business school at a major university. This school considered itself the leader in twelve-month executive advanced-degree programs. Its premier senior management program has a fifty-year track record and has produced alumni that include CEOs, prime ministers, and a secretary general of the United Nations. Its twelve-month technology-oriented program is considered the leading program of its kind providing training for both senior RD&E managers and entrepreneurs throughout the world. However, the demand for twelve-month programs is shrinking because it is increasingly difficult for executives to be away from their companies for twelve months. The senior leadership of the school was considering a radical redesign of their programs. As part of this effort, they sought input from potential students.

A twelve-person committee, which included deans, senior faculty (not the authors), the Director of Executive Education, and representatives of the Boston Consulting Group (BCG), worked over a period of four months to generate the program features to be tested. The motivation for the program redesign was identified in two qualitative studies of the market, one by BCG and the second by another major consulting firm. Both studies suggested that either a weekend, flexible, or on-line-based format replace the twelve-month format (each with a 12-week residency) and that the focus of the program be changed to reflect the changing business environment.

Despite the qualitative studies, there was much debate in the committee about format, focus, and the target makeup of the classes. The committee also sought guidance on program details and tuition. The study began with an internal assessment of capability. This internal assessment, combined with the two qualitative studies, identified eight features that were to be tested with conjoint analysis. They included program focus (3 levels), format (4 levels), class composition (3 levels of interest focus, 4 age categories, 3 types of geographic focus), sponsorship strategy (3 levels), company focus (3 levels), and tuition (3 levels). This  $4^2 \times 3^6$  is relatively large for CBC applications (e.g., Huber 1997, Orme 1999), but proved feasible with professional web design. We provide an example screenshot in Figure 5 (with the university logo and tuition levels redacted). Prior to answering these stated-choice questions, respondents reviewed detailed descriptions of the levels of each feature and could access these descriptions at any time by clicking the feature's logo.

Figure 5

Example Web-Based Stated-Choice Task for Executive-Education Study



After the wording and layout was refined in pretests, potential respondents were obtained from the Graduate Management Admissions Council through their Graduate Management Admissions Search Service. Potential respondents were selected based on their age, geographic location, educational goals, and GMAT scores. Random samples were chosen from three strata – those within driving distance of the university, those within a short airplane flight, and those within a moderate airplane flight. Respondents were invited to participate via unsolicited e-mails from the Director of Executive Education. As an incentive, respondents were entered in a lottery in which they had a 1-in-10 chance of receiving a university-logo gift worth approximately \$100.

Pretests confirmed that the respondents could comfortably answer twelve stated-choice questions (recall that the respondents were experienced executives receiving minimal response incentives). Of those respondents who began the CBC section of the survey, 95% completed the section. The overall response rate was within ranges expected for both proprietary and academic web-based studies (Couper 2000; De Angelis 2001; Dillman et. al., Sheehan 2001).<sup>12</sup> No significant differences were found between

<sup>12</sup> Couper (2000, p. 384) estimates a 10% response rate for open-invitation studies and a 20-25% response rate for studies with pre-recruited respondents. De Angelis (2001) reports “click-through” rates of 3-8% from an e-mail list of 8 million records. In a study designed to optimize response rates, Dillman, et al. (2001) compare mail, telephone, and web-based surveys. They obtain a response rate of 13% for the web-based survey. Sheehan (2001) reviews all published studies referenced in Academic Search Elite, Expanded Academic Index, ArticleFirst, Lexis-Nexis, Psychlit, Sociological Abstracts, ABI-Inform, and ERIC and finds response rates dropping at 4% per year. Sheehan’s data suggest an average response rate for 2002 of 15.5%. The response rate in the Executive Education study was 16%.

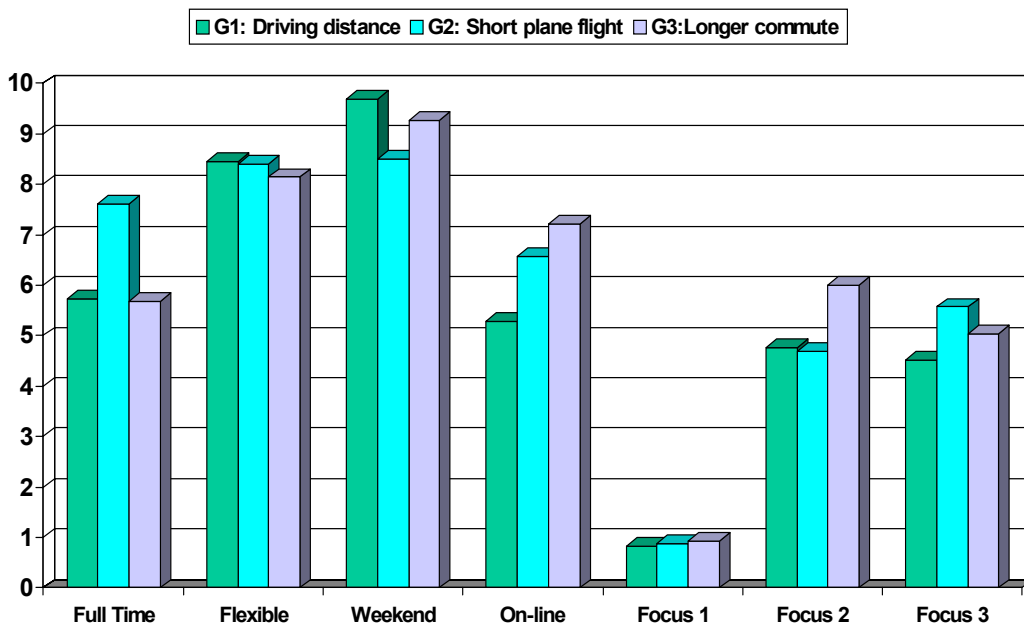
the partworths estimated for early responders and those estimated for later responders. The committee judged the results intuitive, but enlightening, and adequate for managerial decision-making.

**Managerial Implications and Actions Taken**

The committee reviewed average partworths by demographic segment, benefit segments, willingness to pay, and other standard conjoint-analysis outputs. These outputs included an automated choice simulator that enabled the committee to forecast preferences for various programs with estimates for the market as a whole and for any demographic segment. We illustrate some of the managerial insights with Figure 6, which summarizes the average partworths for focus and format. We were asked to keep the detailed format definitions confidential until the program is launched. However, we can say that Focus 1 was the current focus of one of the two twelve-month programs, while Focus 2 and Focus 3 were potential modifications to the current focus. Focus 3 matched better the internal capabilities of the school.

The message was clear; each of the two new focus proposals resonated with the market and was particularly preferred by the university’s traditional target demographics. The analyses also suggested that both flexible and weekend programs were viable new options, especially if satellite operations were established. Moreover, approximately 25% of the population prefers a full-time program and is willing to pay premium tuition for such a program. Based on detailed, segmented analyses and based on internal capabilities, the committee decided to launch new programs based on Focus 3. The new programs are likely to offer both twelve-month and more flexible options.

**Figure 6**  
**Average Partworths From Choice-Based Conjoint Analysis Study**



## Technical Results

The data provide an opportunity to examine several technical issues. First, we can estimate the magnitude and heterogeneity parameters using HB. We obtain estimates of magnitude ( $\bar{\beta}$ ) that ranged from 1.0 to 3.4, averaging 1.5. The heterogeneity coefficient ranged from 0.9 to 3.2, averaging 1.9. Thus, both the observed magnitude and the observed heterogeneity coefficients span the ranges addressed in the Monte Carlo simulations.

Hit rates are more complex. By design, polyhedral questions select the choice sets that provide maximum information about the feasible set of partworths. The Analytic-Center (AC) estimates remain feasible through the twelfth question and obtain an internal hit rate of 100% (by design). This hit rate is not guaranteed for HB estimates, which, nonetheless, do quite well with internal hit rates of 94%. As described earlier in this paper, we can examine internal consistency by comparing the hit rates based on AC estimates from the first  $i$  questions. Specifically, using the equation derived earlier, the internal hit rate should be approximately  $[i + \frac{1}{4}(12 - i)]/i$ . The observed internal hit rates, 51.1% after  $i=4$  and 74.2% after  $i=8$ , are quite close to this formula, which predicts values of 50% and 75%, respectively. This suggests that, empirically, the polyhedral question design algorithm was able to achieve approximate choice balance. These internal hit rates are not guaranteed for HB, which obtained hit rates of 57.1% and 62.4%, respectively.

Finally, Kamakura and Wedel (1995, p. 316) report a metric in which they computed how rapidly estimates converge to their final values. Following their structure, we compute the convergence rates as a function of the number of stated-choice tasks ( $i$ ) using scaled RMSE to maintain consistency with Tables 1 and 2). The scaled RMSE between the estimates based on four questions ( $i=4$ ) and those based on all twelve questions is 0.965 for AC and 0.916 for HB. However, AC appears to converge slightly faster – based on eight questions ( $i=8$ ) the scaled RMSE is 0.552 for AC and 0.806 for HB. To better compare these convergence rates we can report them as a percentage of the total change in the estimates:  $[\text{RMSE}(i=0) - \text{RMSE}(i)]/\text{RMSE}(i=0)$ . Based on this metric, AC reduces 26.3% of the RMSE after four questions and 55.6% of the RMSE after eight questions. The corresponding percentages for HB are 29.7% and 38.2%, respectively.

## Conclusions and Research Opportunities

Research on stated-choice question design suggests that careful selection of the choice sets has the potential to increase accuracy and reduce costs by requiring fewer respondents, fewer questions, or both. This is particularly true in choice-based conjoint analysis because the most efficient design depends upon the true partworth values. In this paper we explore whether the success of aggregate customization



can be extended to individual-level adaptive question design. We propose heuristics for designing profiles for each choice set. We then rely on new developments in dynamic optimization to implement these heuristics. As a first test, we seek to identify whether or not the proposed methods show promise in at least some domains. It appears that such domains exist. Like many proposed methods, we do not expect polyhedral methods to dominate in all domains and, indeed, they do not. However, we hope that by identifying promising domains we can inspire other researchers to explore hybrid methods and/or improve the heuristics.

While polyhedral methods are feasible empirically and show promise, many challenges remain. For example, we might allow fuzzy constraints for the polyhedra. Such constraints might provide greater robustness at the expense of precision. Future simulations might explore other domains including non-diagonal covariance structures, probit-based random-utility models, mixtures of distributions, and finite mixture models. Recently, Ter Hofstede, Kim, and Wedel (2002) demonstrated that self-explicated data could improve HB estimation for full-profile conjoint analysis. Polyhedral estimation handles such data readily – hybrids might be explored that incorporate both self-explicated and stated-choice data. Future developments in dynamic optimization might enable polyhedral algorithms that look more than one step ahead.

We close by recognizing research on other optimization algorithms for conjoint analysis. Evgeniou, Boussios, and Zacharia (2002) propose “support vector machines (SVMs)” to balance complexity of interactions with fit. They are currently exploring hybrids based on SVMs and polyhedral methods. We have also examined polyhedral methods for metric paired-comparison data. Consistent with the findings reported here, these methods also show promise for domains defined by high response accuracy and high heterogeneity.

## References

- Allenby, Greg M. and Peter E. Rossi (1999), "Marketing Models of Consumer Heterogeneity," *Journal of Econometrics*, 89, (March/April), 57-78.
- Andrews, Rick L., Andrew Ainslie, and Imran S. Currim (2002), "An Empirical Comparison of Logit Choice Models with Discrete versus Continuous Representations of Heterogeneity," *Journal of Marketing Research*, 39, (November), 479-487
- 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.
- and Joel Huber (2001), "Improving Parameter Estimates and Model Prediction by Aggregate Customization in Choice Experiments," *Journal of Consumer Research*, 28, (September), 273-283.
- Ben-Akiva, Moshe and Steven R. Lerman (1985), *Discrete Choice Analysis: Theory and Application to Travel Demand*, (Cambridge, MA: MIT Press).
- Christensen, Clayton (1998), *The Innovator's Dilemma : When New Technologies Cause Great Firms to Fail* (Boston, MA: Harvard Business School Press).
- Couper, Mick P. (2000), "Web Surveys: Review of Issues and Approaches," *Public Opinion Quarterly*, 64, 464-494.
- De Angelis, Chris (2001), "Sampling for Web-based Surveys," 22<sup>nd</sup> Annual Marketing Research Conference, Atlanta, GA, (September 23-26).
- Dillman, Don A., Glenn Phelps, Robert Tortora, Karen Swift, Julie Kohrell, and Jodi Berck (2001), "Response Rate and Measurement Differences in Mixed Mode Surveys Using Mail, Telephone, Interactive Voice Response, and the Internet," (Pullman, WA: Social and Economic Sciences Research Center, Washington State University).
- Evgeniou, Theodoros, Constantinos Boussios, and Giorgos Zacharia (2002), "Generalized Robust Conjoint Estimation," Working Paper, (Fontainebleau, France: INSEAD).
- Freund, Robert (1993), "Projective Transformations for Interior-Point Algorithms, and a Superlinearly Convergent Algorithm for the W-Center Problem," *Mathematical Programming*, 58, 385-414.
- , Robert, R. Roundy, and M.J. Todd (1985), "Identifying the Set of Always-Active Constraints in a System of Linear Inequalities by a Single Linear Program," WP 1674-85, Sloan School of Management, MIT.
- Gritzmann P. and V. Klee (1993), "Computational Complexity of Inner and Outer J-Radii of Polytopes in Finite-Dimensional Normed Spaces," *Mathematical Programming*, 59(2), pp. 163-213.
- Hadley, G. (1961), *Linear Algebra*, (Reading, MA: Addison-Wesley Publishing Company, Inc.)
- Hauser, John R. (1978), "Testing the Accuracy, Usefulness and Significance of Probabilistic Models: An Information Theoretic Approach," *Operations Research*, Vol. 26, No. 3, (May-June), 406-421.

- Huber, Joel (1997), "What we Have Learned from 20 Years of Conjoint Research," Research Paper Series, (Sequim, WA: Sawtooth Software, Inc.).
- and Klaus Zwerina (1996), "The Importance of Utility Balance in Efficient Choice Designs," *Journal of Marketing Research*, 33, (August), 307-317.
- Johnson, Richard (1999), "The Joys and Sorrows of Implementing HB Methods for Conjoint Analysis," Research Paper Series, (Sequim, WA: Sawtooth Software, Inc.).
- Kamakura, Wagner and Michel Wedel (1995), "Life-Style Segmentation with Tailored Interviewing," *Journal of Marketing Research*, 32, (August), 308-317.
- Karmarkar, N. (1984), "A New Polynomial Time Algorithm for Linear Programming," *Combinatorica*, 4, 373-395.
- Kuhfeld, Warren F. , Randall D. Tobias, and Mark Garratt (1994), "Efficient Experimental Design with Marketing Research Applications," *Journal of Marketing Research*, 31, 4, (November), 545-557.
- 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).
- McFadden, Daniel (1974), "Conditional Logit Analysis of Qualitative Choice Behavior," *Frontiers in Econometrics*, P. Zarembka, ed., (New York: Academic Press), 105-142.
- Nesterov, Y. and A. Nemirovskii (1994), "Interior-Point Polynomial Algorithms in Convex Programming," SIAM, Philadelphia.
- Orme, Bryan (1999), "ACA, CBC, or Both?: Effective Strategies for Conjoint Research," Working Paper, Sawtooth Software, Sequim, WA.
- Sandor, Zsolt and Michel Wedel (2001), "Designing Conjoint Choice Experiments Using Managers' Prior Beliefs," *Journal of Marketing Research*, 38, 4, (November), 430-444.
- Sheehan, Kim (2001), "E-mail Survey Response Rates: A Review," *Journal of Computer-Mediated Communications*, 6, 2, (January).
- Sonnevend, G. (1985a), "An 'Analytic' Center for Polyhedrons and New Classes of Global Algorithms for Linear (Smooth, Convex) Programming," *Proceedings of the 12<sup>th</sup> 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, Eötvös University, Budapest, 1985.
- Srinivasan, V. and Allan D. Shocker (1973), "Linear Programming Techniques for Multidimensional Analysis of Preferences," *Psychometrika*, 38, 3, (September), 337-369.
- Ter Hofstede, Frenkel, Youngchan Kim, and Michel Wedel (2002), "Bayesian Prediction in Hybrid Conjoint Analysis," *Journal of Marketing Research*, 39, (May), 253-261.
- 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.

## Appendix: Mathematics of Polyhedral Methods for CBC Analysis

Let  $u_f$  be the  $f^{\text{th}}$  parameter of the respondent's partworth function and let  $\vec{u}$  be the  $p \times 1$  vector of parameters. Without loss of generality we (1) assume  $p$  binary features such that  $u_f$  is the high level of the  $f^{\text{th}}$  feature and (2)  $\sum_{f=1}^p u_f = 100$ . For more levels we recode the  $\vec{u}$  vector and impose constraints such as  $u_m \leq u_h$ . We handle inequality constraints by adding slack variables,  $v_{hm} \geq 0$ , such that  $u_h = u_m + v_{hm}$ . Let  $r$  be the number of externally imposed constraints, of which  $r' \leq r$  are inequality constraints. Let  $\vec{z}_{ij}$  be the  $1 \times p$  vector describing the  $j^{\text{th}}$  profile in the  $i^{\text{th}}$  choice set, and, without loss of generality, index by  $j=1$  the respondent's choice from each set. In the text we let the number of profiles in each set be  $J_i$ , however, to simplify notation, we drop the  $i$  subscript in this appendix. The extension is obvious. Let  $X$  be the  $q(J-1) \times p$  matrix of  $\vec{x}_{ij} = \vec{z}_{i1} - \vec{z}_{ij}$  for  $i = 1$  to  $q$  and  $j = 2$  to  $J$ . Then, if there were no errors, the respondent's choices imply  $X\vec{u} \geq \vec{0}$  where  $\vec{0}$  is a vector of 0's. We add slack variables and augment the  $\vec{u}$  vector such that  $X\vec{u} = \vec{0}$ . To handle additional constraints, we augment these equations such that  $\vec{u}$  and  $X$  include  $r'$  additional slack variables and  $r$  additional equations. These augmented relationships form a polyhedron,  $P = \{ \vec{u} \in \mathbb{R}^{p+q(J-1)+r'} \mid X\vec{u} = \vec{a}, \vec{u} \geq \vec{0} \}$  where  $\vec{a}$  contains non-zero elements due to the external constraints. We begin by assuming that  $P$  is non-empty, that  $X$  is full-rank, and that no  $j$  exists such that  $u_j=0$  for all  $\vec{u}$  in  $P$ . We later indicate how to handle these cases.

### Finding an Interior Point of the Polyhedron

To begin the algorithm we first find a feasible interior point of  $P$  by solving a linear program, LP1 (Freund, Roundy and Todd 1985). Let  $\vec{e}$  be a vector of 1's; the  $y_f$ 's and  $\theta$  are parameters of LP1 and  $\vec{y}$  is the  $(p+q(J-1)+r') \times 1$  vector of the  $y_f$ 's. (When clear in context, inequalities applied to vectors apply for each element.) LP1 is given by:

$$(LP1) \quad \max \sum_{f=1}^{p+q(J-1)+r'} y_f, \quad \text{subject to: } X\vec{u} = \theta\vec{a}, \quad \theta \geq 1, \quad \vec{u} \geq \vec{y} \geq \vec{0}, \quad \vec{y} \leq \vec{e}$$

If  $(\vec{u}^*, \vec{y}^*, \theta^*)$  solves LP1, then  $\theta^{*-1} \vec{u}^*$  is an interior point of  $P$  whenever  $\vec{y}^* > \vec{0}$ . If there are some  $y_f$ 's equal to 0, then there are some  $f$ 's for which  $u_f=0$  for all  $\vec{u} \in P$ . If LP1 is infeasible, then  $P$  is empty. We address these cases later in this appendix.

### Finding the Analytic Center

The analytic center is the point in  $P$  that maximizes the geometric mean of the distances from the point to the faces of  $P$ . We find the analytic center by solving OPTAC1.

$$(OPTAC1) \quad \max \sum_{f=1}^{p+q(J-1)+r'} \ln(u_f), \quad \text{subject to: } X\vec{u} = \vec{a}, \quad \vec{u} > \vec{0}$$

Freund (1993) proves with projective methods that a form of Newton's method will converge rapidly for OPTAC1. To implement Newton's method we begin with the feasible point from LP1 and improve it with a scalar,  $\alpha$ , and a direction,  $\vec{d}$ , such that  $\vec{u} + \alpha\vec{d}$  is close to the optimal solution of OPTAC1. ( $\vec{d}$  is a vector of  $d_f$ 's.) We then iterate subject to a stopping rule.

We first approximate the objective function with a quadratic expansion in the neighborhood of  $\vec{u}$ .

$$(A1) \quad \sum_{f=1}^{p+q(J-1)+r'} \ln(u_f + d_f) \approx \sum_{f=1}^{p+q(J-1)+r'} \ln(u_f) + \sum_{f=1}^{p+q(J-1)+r'} \left( \frac{d_f}{u_f} - \frac{d_f^2}{2u_f^2} \right)$$

If we define  $U$  as a diagonal matrix of the  $u_f$ 's, then the optimal direction solves OPT2:

$$(OPT2) \quad \max \bar{e}^T U^{-1} \bar{d} - (\gamma/2) \bar{d}^T U^{-2} \bar{d}, \quad \text{subject to: } X\bar{d} = \bar{0}$$

Newton's method solves OPT1 quickly by exploiting an analytic solution to OPT2. To see this, consider first the Karush-Kuhn-Tucker (KKT) conditions for OPT2. If  $\bar{z}$  is a vector parameter of the KKT conditions that is unconstrained in sign then the KKT conditions are written as:

$$(A2) \quad U^{-2} \bar{d} - U^{-1} \bar{e} = X^T \bar{z}$$

$$(A3) \quad X\bar{d} = \bar{0}$$

Multiplying A2 on the left by  $XU^2$ , gives  $X\bar{d} - XU\bar{e} = XU^2 X^T \bar{z}$ . Applying A3 to this equation gives:  $-XU\bar{e} = XU^2 X^T \bar{z}$ . Since  $U\bar{e} = \bar{u}$  and since  $X\bar{u} = \bar{a}$ , we have  $-\bar{a} = XU^2 X^T \bar{z}$ . Because  $X$  is full rank and  $U$  is positive, we invert  $XU^2 X^T$  to obtain  $\bar{z} = -(XU^2 X^T)^{-1} \bar{a}$ . Now replace  $\bar{z}$  in A2 by this expression and multiply by  $U^2$  to obtain  $\bar{d} = \bar{u} - U^2 X^T (XU^2 X^T)^{-1} \bar{a}$ .

According to Newton's method, the new estimate of the analytic center,  $\bar{u}'$ , is given by  $\bar{u}' = \bar{u} + \alpha \bar{d} = U(\bar{e} + \alpha U^{-1} \bar{d})$ . There are two cases for  $\alpha$ . If  $\|U^{-1} \bar{d}\| < \gamma/4$ , then we use  $\alpha = 1$  because  $\bar{u}$  is already close to optimal and  $\bar{e} + \alpha U^{-1} \bar{d} > \bar{0}$ . Otherwise, we compute  $\alpha$  with a line search.

### Special Cases

If  $X$  is not full rank,  $XU^2 X^T$  might not invert. We can either select questions such that  $X$  is full rank or we can make it so by removing redundant rows. Suppose that  $\bar{x}_k$  is a row of  $X$  such that  $\bar{x}_k^T = \sum_{\ell=1, \ell \neq k}^{q(J-1)+r} \beta_\ell \bar{x}_\ell^T$ . Then if  $a_k = \sum_{\ell=1, \ell \neq k}^{q(J-1)+r} \beta_\ell a_\ell$ , we remove  $\bar{x}_k$ . If  $a_k \neq \sum_{\ell=1, \ell \neq k}^{q(J-1)+r} \beta_\ell a_\ell$ , then  $P$  is empty and we employ OPT4 described later in this appendix.

If in LP1 we detect cases where some  $y_f$ 's = 0, then there are some  $f$ 's for which  $u_f = 0$  for all  $\bar{u} \in P$ . In the later case, we can still find the analytic center of the remaining polyhedron by removing those  $f$ 's and setting  $u_f = 0$  for those indices.

### Finding the Ellipsoid and its Longest Axis

If  $\bar{u}$  is the analytic center and  $\bar{U}$  is the corresponding diagonal matrix, then Sonnevend (1985a, 1985b) demonstrates that  $E \subseteq P \subseteq E_{p+q(J-1)+r}$  where,  $E = \{ \bar{u} \mid X\bar{u} = \bar{a}, \sqrt{(\bar{u} - \bar{u})^T \bar{U}^{-2} (\bar{u} - \bar{u})} \leq 1 \}$  and  $E_{p+q(J-1)+r}$  is constructed proportional to  $E$  by replacing 1 with  $(p+q(J-1)+r)$ . Because we are interested only in the direction of the longest axis of the ellipsoids we can work with the simpler of the proportional ellipsoids,  $E$ . Let  $\bar{g} = \bar{u} - \bar{u}$ , then the longest axis will be a solution to OPT3.

$$(OPT3) \quad \max \bar{g}^T \bar{g} \quad \text{subject to: } \bar{g}^T \bar{U}^{-2} \bar{g} \leq 1, \quad X\bar{g} = \bar{0}$$

OPT3 has an easy-to-compute solution based on the eigenstructure of a matrix. To see this we begin with the KKT conditions (where  $\phi$  and  $\gamma$  are parameters of the conditions).

$$(A4) \quad \bar{g} = \phi \bar{U}^{-2} \bar{g} + X^T \bar{\gamma}$$

$$(A5) \quad \phi(\bar{g}^T \bar{U}^{-2} \bar{g} - 1) = 0$$

$$(A6) \quad \bar{g}^T \bar{U}^{-2} \bar{g} \leq 1, \quad X\bar{g} = \bar{0}, \quad \phi \geq 0$$

It is clear that  $\bar{g}^T \bar{U}^{-2} \bar{g} = 1$  at optimal, else we could multiply  $\bar{g}$  by a scalar greater than 1 and still have  $\bar{g}$  feasible. It is likewise clear that  $\phi$  is strictly positive, else we obtain a contradiction by left-multiplying A4 by  $\bar{g}^T$  and using  $X\bar{g} = \bar{0}$  to obtain  $\bar{g}^T \bar{g} = 0$  which contradicts  $\bar{g}^T \bar{U}^{-2} \bar{g} = 1$ . Thus, the solution to OPT3 must satisfy  $\bar{g} = \phi \bar{U}^{-2} \bar{g} + X^T \bar{\gamma}$ ,  $\bar{g}^T \bar{U}^{-2} \bar{g} = 1$ ,  $X\bar{g} = \bar{0}$ , and  $\phi > 0$ . We rewrite A4-A6 by letting  $I$  be the identify matrix and defining  $\eta = 1/\phi$  and  $\bar{w} = -\bar{\gamma}/\phi$ .

$$(A7) \quad (\bar{U}^{-2} - \eta I) \bar{g} = X^T \bar{w}$$

$$(A8) \quad \bar{g}^T \bar{U}^{-2} \bar{g} = 1$$

$$(A9) \quad X\bar{g} = \bar{0}, \quad \phi > 0$$

We left-multiply A7 by  $X$  and use A9 to obtain  $X\bar{U}^{-2} \bar{g} = XX^T \bar{w}$ . Since  $X$  is full rank,  $XX^T$  is invertible and we obtain  $\bar{w} = (XX^T)^{-1} X\bar{U}^{-2} \bar{g}$  which we substitute into A7 to obtain

$(\bar{U}^{-2} - X^T (XX^T)^{-1} X\bar{U}^{-2}) \bar{g} = \eta \bar{g}$ . Thus, the solution to OPT3 must be an eigenvector of the matrix,  $M \equiv (\bar{U}^{-2} - X^T (XX^T)^{-1} X\bar{U}^{-2})$ . To find out which eigenvector, we left-multiply A7 by  $\bar{g}^T$  and use A8 and A9 to obtain  $\eta \bar{g}^T \bar{g} = 1$ , or  $\bar{g}^T \bar{g} = 1/\eta$  where  $\eta > 0$ . Thus, to solve OPT3 we maximize  $1/\eta$  by selecting the smallest positive eigenvector of  $M$ . The direction of the longest, next longest, etc. axes are then given by the associated eigenvectors of  $M$ .

We need only establish that the eigenvalues of  $M$  are real. To do this we recognize that  $M = P\bar{U}^{-2}$  where  $P = (I - X^T (XX^T)^{-1} X)$  is symmetric, i.e.,  $P = P^T$ . Then if  $\eta$  is an eigenvalue of  $M$ ,  $\det(P\bar{U}^{-2} - \eta I) = 0$ , which implies that  $\det[\bar{U}(\bar{U}^{-1} P\bar{U}^{-1} - \eta I)\bar{U}^{-1}] = 0$ . This implies that  $\eta$  is an eigenvalue of  $\bar{U}^{-1} P\bar{U}^{-1}$ , which is symmetric. Thus,  $\eta$  is real (Hadley 1961, 240).

### Selecting Profiles for Target Partworth Values

We then select the values of the  $\bar{u}_{ij}$  for the next question ( $i = q+1$ ) based on the longest axes. Each axis provides two target values. For odd  $J$  we randomly select from target values derived from the  $[(J+1)/2]^{\text{th}}$  eigenvector. To find the extreme estimates of the parameters,  $\bar{u}_{ij}$ , we solve for the points where  $\bar{u}_{ij} = \bar{u}' + \alpha_1 \bar{g}_1$ ,  $\bar{u}_{ij} = \bar{u}' - \alpha_2 \bar{g}_1$ ,  $\bar{u}_{ij} = \bar{u}' + \alpha_3 \bar{g}_2$ , and  $\bar{u}_{ij} = \bar{u}' - \alpha_4 \bar{g}_2$  intersect  $P$ . For each  $\alpha$  we do this by increasing  $\alpha$  until the first constraint in  $P$  is violated. To find the profiles in the choice set we select, as researcher determined parameters, feature costs,  $\bar{c}$ , and a budget,  $M$ . Without such constraints, the best profile is trivially the profile with all features set to their high levels. Subject to this budget constraint, we solve the following knapsack problem with dynamic programming.

$$(OPT4) \quad \max \bar{z}_{ij} \bar{u}_{ij} \quad \text{subject to: } \bar{z}_{ij} \bar{c} \leq M, \quad \text{elements of } \bar{z}_{ij} \in \{0,1\}$$

For multi-level features we impose constraints on OPT4 that only one level of each feature is chosen. In the algorithms we have implemented to date, we set  $\bar{c} = \bar{u}'$  and draw  $M$  from a uniform distribution on  $[0, 50]$ , redrawing  $M$  (up to thirty times) until all four profiles are distinct. If distinct profiles cannot be identified, then it is likely that  $P$  has shrunk sufficiently for the managerial problem. For null profiles, extend the constraints accordingly, as described in the text.