Estimating Main Effects with Pareto Optimal Subsets

Damaraju Raghavarao
and
James B. Wiley

Series No. MKTG 96.064
(Revision of MKTG 95.046)
March, 1996
ESTIMATING MAIN EFFECTS WITH PARETO OPTIMAL SUBSETS

Damaraju Raghavarao
Department of Statistics
Temple University
Philadelphia, PA 19122

James B. Wiley
Department of Marketing
Hong Kong University of Science & Technology
Clear Water Bay, Kowloon, HONG KONG

Keywords and Phrases: Choice Experiment, Choice Sets, Conjoint Measurement, Trade-off Analysis.

ABSTRACT

A subset T of S is said to be a Pareto Optimal Subset of m ordered attributes (factors) if for profiles (combination of attribute levels) \((x_1, x_2, \ldots, x_m)\) and \((y_1, y_2, \ldots, y_m)\) \(\in T\), no profile "dominates" another: that is, there exist no pair such that \(x_i \leq y_i\), for \(i = 1, 2, \ldots, m\). The experimental design problem here is to select a Pareto Optimal subset, or subsets, such that we can estimate all elementary contrasts between the levels of each attribute of a main-effects design.

1. INTRODUCTION

In marketing research one likes to know how consumers choose among multiattribute alternatives that are described in terms of either benefits (attributes on which value increases with level) or costs (attributes on which value decreases with level). For example, beneficial attributes of automobiles include: roominess, riding comfort, quietness, and acceleration. Costs include: price, operating costs, and emissions. Generally, it may be taken as a priori information that buyers prefer higher levels of benefits and lower levels of costs. It is therefore not informative to ask respondents whether they would prefer to have more of a single benefit or less of it or whether they would prefer to pay a higher price for a product.
or a lower price. However, production technology, budget constraints, market forces and the like combine to create a "no-free-lunch" situation for those who want to buy products or services. That is, when faced with the offerings in a market, consumers find that they must give up an incremental level of benefit (or incur an incremental cost) in order to acquire an incremental benefit (or avoid an incremental cost) on some other attribute. In making choices in such markets, consumers reveal the value they attach to increments of benefits and costs relative to the value they attach to increments on other attributes.

Market researchers have developed "conjoint measurement" procedures (also called "trade-off analysis") to describe how consumers "trade-off" benefits and costs. In a typical study, consumers are given a set of attribute profiles (attribute level combinations) and they indicate their preference for the profiles. Two common methods for indicating preference are: a) choice, the respondent picks the most preferred profile from a set, or sets and b) ranking, profiles in a set are ranked in terms of preference. Ideally, options should be presented to respondents in choice sets in such a way that they must "give-up" something in order to "get" something. In practice, however, fractional factorial main effects designs are used to specify which subsets of all possible combinations are used in a conjoint study. These designs succeed in reducing the number of profiles presented to respondents to a manageable number. However, they do not guarantee that subsets presented to respondents will not contain profiles that dominate, or are dominated by, other profiles in a set. The objective of this paper is to provide ways of generating sets that are guaranteed not to contain dominated or dominating alternatives.

The implications of having dominated or dominating alternatives in a set can be illustrated by a simple example. Suppose a consumer is given the choice between a $9,000 car that gets 25 miles-per-gallon and a $10,000 car that gets 30 miles-per-gallon of gasoline. If the consumer
chooses the first alternative, it may be inferred that he or she values $1,000 more than 5 additional miles-per-gallon, assuming other things equal. Conversely, choice of the second alternative would indicate that the incremental gas mileage was valued more than the increment in price. If a consumer is given the choice between a $9,000 car that gets 30 miles-per-gallon and a $10,000 car that gets 25 miles-per-gallon, the first one will automatically be selected because it dominates the second alternative. Nothing is learned about the relative value the respondent attaches to cost versus gas mileage. Problems of this type arise not only in market research, but also in other areas of social research. A human resources administrator may like to know whether employees prefer a better pension plan or higher wages. Politicians may like to know whether the public likes increased taxes or a reduced budget. An instructor may like to know whether students would like more material covered or would like more problems solved in the classroom. In all these situations, respondents must be provided with choice sets where no profile of attributes dominates other profiles of attributes in that choice set.

The notion of conjoint measurement is that a fairly good idea of the relative value consumers attach to attribute levels may be inferred from their responses to profiles in a number of choice sets. This paper describes how choice sets may be constructed so that respondents must make trade-offs when indicating their preferences for the elements of choice sets.

When respondents choose from sets, responses typically are aggregated over respondents and the dependent variable consists of the proportion of respondents selecting the profiles in each choice set. Analysis of these proportions is done within a logit or probit framework. Assume that $Y_{x_1, x_2, \ldots, x_m}$ is a consumer's response to profile $(x_1, x_2, \ldots, x_m)$ where $x_i$ is the level of the $i$th attribute, $i = 1, 2, \ldots, m$. or the transformed proportional response to that profile from that choice set. In conjoint measurement, it is typically assumed that
\[ Y_{x_1, x_2, \ldots, x_n} = \mu + \sum_{i=1}^{m} \alpha_{x_i}^i + e_{x_1, x_2, \ldots, x_n}. \] (1.1)

where \( \mu \) is the general mean, \( \alpha_{x_i}^i \) is the effect (value) of \( x_i \) level of the \( i \)th attribute \( A_i \). Equation (1.1) can be recognized as a main effects ANOVA model. Different assumptions can be made regarding the distribution of the random errors \( e_{x_1, x_2, \ldots, x_n} \). For the purpose of this paper we need only assume the expectations of the errors are zero.

If the inferences are to reveal information about trade-offs, however, the choice sets must be judiciously constructed. For example, choice or ordering of the pair \( \{ \$9.000, 25 \text{ m.p.g.} \} \) and \( \{ \$10.000, 30 \text{ m.p.g.} \} \) requires that consumers reveal information about their relative evaluation of the two attributes - price and miles-per-gallon. On the other hand, choice or ordering of the pair \( \{ \$9.000, 30 \text{ m.p.g.} \} \) and \( \{ \$10.000, 30 \text{ m.p.g.} \} \) is implicit in the assumption that the benefits and cost attributes are ordered.

Model (1.1) can be fitted with any type of choice sets. However, if information about consumers' relative evaluation of attribute levels is to be gathered, choice sets should require that trade-offs are made. Sets of this sort are called Pareto Optimal sets. Wiley (1978) provided procedures for constructing the largest Pareto Optimal set for \( s \) designs, where \( s \) is the number of attributes. Krieger and Green (1991) extended that work. The present paper provides a number of results for Pareto Optimal designs, including procedures for generating main effects designs for symmetric and asymmetric experiments.

The introduction to this point has explained the application of choice experiments in marketing research and behavioral experiments. However, the main purpose of the present paper is to show that Pareto Optimal subsets have a useful role in planning certain types of experiments and to discuss the estimability of all main effects contrasts using them. The key results established in this connection are Theorems 2.4, 3.1, and 4.1.
2. PARETO OPTIMAL SUBSETS

Consider an experiment with \( m \geq 2 \) attributes (factors) \( A_1, A_2, \ldots, A_m \). Let attribute \( A_i \) have \( s_i \) levels denoted by \( 0, 1, \ldots, s_i - 1 \); \( i = 1, 2, \ldots, m \). Assume that \( s_1 \leq s_2 \leq \ldots \leq s_m \). Let \( n = \prod_{i=1}^{m} s_i \). There are \( n \) attribute profiles denoted by \( (x_1, x_2, \ldots, x_m) \), where \( x_i = 0, 1, \ldots, s_i - 1 \); \( i = 1, 2, \ldots, m \). Let \( S \) be the set of all the \( n \) attribute profiles. Then Pareto Optimality may be defined as follows:

**DEFINITION 2.1:** A subset \( T \) of \( S \) is said to be a Pareto Optimal subset if for every two distinct profiles \( (x_1, x_2, \ldots, x_m) \), \( (y_1, y_2, \ldots, y_m) \in T \), there exist subscripts \( i \) and \( j \) (\( i \neq j \)) such that \( x_i < y_i \) and \( x_j > y_j \). We now have:

**THEOREM 2.1.** The set \( S_\alpha = \left\{ (x_1, x_2, \ldots, x_m) \mid \sum_{i=1}^{m} x_i = \alpha \right\} \) is a Pareto Optimal subset.

**PROOF:** Let \( (x_1, x_2, \ldots, x_m) \), \( (y_1, y_2, \ldots, y_m) \) be two members of \( S_\alpha \). Then

\[
\sum_{i=1}^{m} x_i = \sum_{j=1}^{m} y_j = \alpha. \tag{2.1}
\]

Now if \( x_i < y_i \) for some \( i \), then clearly there exists at least one \( j \) such that \( y_j < x_j \) in order that (2.1) is true. Thus \( S_\alpha \) is a Pareto Optimal subset.

Let \( k = \sum_{i=1}^{m} s_i - m \). Then \( S_0, S_1, \ldots, S_k \) are \( k+1 \) Pareto Optimal subsets covering \( S \). The attribute profiles \( (0, 0, 0, \ldots, 0), (1, 0, 0, \ldots, 0), (1, 1, 0, \ldots, 0), (1, 1, 1, \ldots, 0), \ldots, (s_1 - 1, s_2 - 1, \ldots, s_m - 1) \) are \( k+1 \) in number and no two of them can occur in the same Pareto Optimal subset. Thus there are at least \( k+1 \) Pareto Optimal disjoint subsets in \( S \) and we establish the following:

**THEOREM 2.2.** There are at least \( k+1 \) Pareto Optimal disjoint subsets in \( S \) and \( S_0, S_1, \ldots, S_k \) is a minimal covering of \( S \).
From Sperner's problem (1928), it is known that the largest Pareto Optimal subset is $S_{k+1}$, if $k$ is even; and $S_{k+1}$, or $S_{k-1}$, if $k$ is odd (see Krieger and Green (1991)). By standard counting methods (see Feller (1957)), we now establish:

**THEOREM 2.3.** The number of elements in $S_\alpha$ denoted by $|S_\alpha|$ is

$$|S_\alpha| = \binom{m + \alpha - 1}{m - 1} - \sum_{i=1}^{n} \binom{m + (\alpha - s_i) - 1}{m - 1} + \sum_{i<j}^{m} \binom{m + (\alpha - s_i - s_j) - 1}{m - 1} - \ldots \quad (2.2)$$

**PROOF:** Consider $\alpha$ stars and $m+1$ bars to create the attribute profiles using 1 bar at left and 1 bar at the right end. Ignoring the levels of the attributes, the number of possible configurations are

$$\binom{m + \alpha - 1}{m - 1}. \quad (2.3)$$

The number of these configurations that exceed $s_i - 1$ levels for attribute $A_i$ are those in which $s_i$ levels are given to $A_i$ and the remaining $\alpha - s_i$ stars are assigned to the $m$ attributes. The number of such configurations are

$$\binom{m + (\alpha - s_i) - 1}{m - 1}. \quad (2.4)$$

The number of configurations enumerated in (2.3) that exceed $s_i - 1$ levels for attribute $A_i$ and $s_j - 1$ levels for attribute $A_j$ are obtained by assigning $s_i$ levels to $A_i$ and $s_j$ levels to $A_j$ and randomly assigning the remaining $\alpha - s_i - s_j$ to the $m$ attributes and this number is

$$\binom{m + (\alpha - s_i - s_j) - 1}{m - 1}. \quad (2.5)$$

Continuing this argument and counting by the inclusion and exclusion principle, we get (2.2).
Borrowing the notation from factorial experiments, we call the experiment \( s_1 \times s_2 \times \ldots \times s_m \) for the \( m \) attribute case with \( s_i \) levels for \( A_i \). If \( s_1 = s_2 = \ldots = s_m = s \), we call it an \( s^m \) experiment.

The following is a corollary for Theorem 2.3.

**COROLLARY 2.3.1:** In a \( s^m \) experiment

\[
|S_\alpha| = \binom{m + \alpha - 1}{m - 1} - \binom{m - (\alpha - s) - 1}{m - 1} \\
+ \binom{m}{2} \binom{m + \alpha - 2s - 1}{m - 1} - \ldots \\
+ (-1)^q \binom{m}{q} \binom{m + \alpha - qs - 1}{m - 1},
\]

(2.4)

where \( \alpha = qs + r, \ 0 \leq r < s \).

As noted in the introduction, the experimental setting in a conjoint measurement study is such that responses will be recorded for each attribute profile in a Pareto Optimal choice set. Let \( Y_{x_1, x_2, \ldots, x_m} \) be a response (or a transformed response) to the profile \((x_1, x_2, \ldots, x_m)\). We assume that the consumers' response to the profile is determined as a function of the attribute levels as specified by the linear model (1.1). It may be noted that the statistical properties of Non-Pareto Optimal and Pareto Optimal designs are difficult to distinguish. The consequence of Non-Pareto Optimal choice sets will be manifested in the values assumed by the dependent variable and not in the statistical properties of the designs used to generate the observations.

It is usually the case that several choice sets are given to each respondent and that the union of these sets is not Pareto optimal. The fact that the union of the sets is not Pareto optimal will not influence choice so long as the individual sets from which choices are made are Pareto Optimal. The key issue is whether the respondent must make trade-offs between attributes when choosing from a given choice set.
The design $D$ consisting of attribute profiles from a single Pareto Optimal subset, or several Pareto Optimal subsets (using each subset as a choice set), is said to be a connected main effects plan, if for model (1.1) we can estimate all the elementary contrasts of the form:

$$\alpha_u - \alpha_{u'}; \ u, u' = 0, 1, ..., s_i - 1; \ u \neq u'; \ i = 1, 2, ..., m.$$  

We now prove the following theorem of interest.

**THEOREM 2.4.** The design $D$ of attribute profiles resulting from a single Pareto Optimal subset $S_a$ is not a connected main effects plan.

**PROOF:** Let $X$ be the design matrix of order $|S_a|$ by $(k+m+1)$ for the design $D$ of attribute profiles from a single Pareto Optimal subset $S_a$. Then

$$X = \{1X_{10} \ X_{11} \ \ldots \ X_{1(s_1-1)} \ \ X_{20} \ X_{21} \ \ldots \ X_{2(s_2-1)} \ \ldots \ X_{m0} \ X_{m1} \ \ldots \ X_{m(s_m-1)}\},$$

where $1$ is a column vector of ones of order $|S_a|$ and the $i$th component of the column, $X_{ju}$ has $1$ or $0$ according as the $u$th level of $A_j$ is in the $u$th attribute profile of $S_a$, or not, for $u = 0, 1, ..., s_j - 1; \ j = 1, 2, ..., m$. Clearly

$$1 = \sum_{u=0}^{s_1-1} X_{1u} = \sum_{u=0}^{s_2-1} X_{2u} = \ldots = \sum_{u=0}^{s_m-1} X_{mu},$$

$$\sum_{i=1}^{m} \sum_{u=0}^{s_j-1} u \ X_{ju} = \alpha 1. \quad (2.5)$$

Note that for a linear model $E(Y) = X\beta$, the linear parametric functions corresponding to the rows of $L\beta$ are estimable if and only if

$$\text{Rank (}X) = \text{Rank} \left( \frac{X}{L} \right). \quad (2.6)$$

The elementary contrasts of the effects between any two levels of each attribute are not estimable because the columns of the augmented right hand side matrix of (2.6) do not satisfy the last condition of (2.5), and the theorem is proved.
Since Theorem 2.4 rules out the existence of connected main effects plans based on a single Pareto Optimal subset, we develop connected main effects plans for symmetric and asymmetric cases in the subsequent sections, using two or more Pareto Optimal subsets.

3. CONNECTED MAIN EFFECTS PLAN – SYMMETRIC CASE

For a symmetric experiment, two Pareto Optimal subsets provide a connected main effects plan as given in the following theorem.

**THEOREM 3.1.** For a $s^m$ experiment, the design $D$ based on Pareto Optimal subsets $S_{k/2}$ and $S_{k-1}$. If $k$ is even, and $S_{(k-1)/2}$, and $S_{(k-3)/2}$ if $k$ is odd, is a connected main effects plan.

**PROOF:** Note that $k/2 \geq s - 1$. In design $D$, the following attribute profiles occur for even $k$, where $x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_m$ are the same and $x_i$ takes the indicated values in each block:

<table>
<thead>
<tr>
<th>$x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_m$</th>
<th>$\sum_{j=1 \atop j \neq i}^m x_j = \frac{k}{2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$k/2$</td>
</tr>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$k/2 - 1$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>$s - 2$</td>
<td>$k/2 - (s - 2)$</td>
</tr>
<tr>
<td>$s - 1$</td>
<td></td>
</tr>
</tbody>
</table>

For odd $k$, the following attribute profiles occur, where $x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_m$ are the same and $x_i$ takes the indicated value in each block.
\[
\begin{array}{c|c}
\sum_{j=1}^{m} x_j & = \\
0 & (k+1)/2 \\
1 & (k+1)/2 - 1 \\
\vdots & \\
s - 2 & (k+1)/2 - (s-2) \\
s - 1 & \\
\end{array}
\]

From the indicated attribute profiles, when \( k \) is odd or even, we can easily construct unbiased estimators for \( \alpha_{u}^i - \alpha_{u'}^i \) for \( u \neq u' \); \( u, u' = 0, 1, \ldots, s-1 \), using the differences of responses in each block. These results hold for every \( i \) and hence the design \( D \) is a connected main effects plan.

In Theorem 3.1, we only indicated a pair of Pareto Optimal subsets that provide a connected main effects plan. However, other pairs or several Pareto Optimal subsets may also give connected main effects plans.

If we want an orthogonal main effects plan for \( D \), we have the following result:

**Theorem 3.2.** Each of the following \( s \) plans is an orthogonal main effects plan:

Plan \( i \): \( S_{\alpha}, \quad \alpha \equiv i \pmod{s}, \quad 0 \leq \alpha \leq k, \)

\[ i = 0, 1, \ldots, s - 1. \]
PROOF: The proof easily follows from the fact that the respective $i$ are a $1/s$ fraction of an $s^n$ experiment.

To keep a good mix of the attribute levels, it is desirable to choose plan $i$ such that $k/2 \equiv i \pmod{s}$, if $k$ is even, and $(k+1)/2 \equiv i \pmod{s}$, if $k$ is odd.

4. CONNECTED MAIN EFFECTS PLAN – ASYMMETRIC CASE

We cannot get a connected main effects plan using only two Pareto Optimal subsets as in Theorem 3.1 in any asymmetric case. In a $2 \times 10$ experiment, the largest Pareto Optimal subset $S_3$ has only 2 elements and obviously we cannot get a connected main effects plan in 2 Pareto Optimal subsets. However, the following result can be proved on similar lines as Theorem 3.1.

THEOREM 4.1. In a $s_1 \times s_2 \times \ldots \times s_m$ experiment ($s_1 \leq s_2 \leq \ldots \leq s_m$), the main effects plan $D$ based on $S_\alpha$, for $\alpha = s_1 - 1$, $s_1$, ..., $s_m - 1$, is connected.

PROOF: Let $s_i - 1 = q_i(s_m - s_1) + r_i$, where $0 \leq r_i < s_m - s_1$ for $i = 1, 2, \ldots, m$. In $D$ the following attribute profiles occur, where $x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_m$ are the same and $x_i$ takes the indicated values in a block. (Note that $s_i - s_1 \leq s_m - s_1$):
<table>
<thead>
<tr>
<th>$\ell_1, \ell_2, \ldots, \ell_{i-1}, \ell_i, \ell_{i+1}, \ldots, \ell_m$</th>
<th>$\sum_{j=1}^{m} \ell_j = \ell_i - 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$\ell_i - 1$</td>
</tr>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$\ell_m - \ell_i$</td>
<td></td>
</tr>
<tr>
<td>$\ell_m - \ell_i - 1$</td>
<td>$\ell_i - 1 - (\ell_m - \ell_i)$</td>
</tr>
<tr>
<td>$\ell_i - 1 - q \ell_i$</td>
<td></td>
</tr>
<tr>
<td>$s_i - 1$</td>
<td>$\ell_i - 1 - q \ell_i$</td>
</tr>
<tr>
<td>$\ell_i - 1$</td>
<td></td>
</tr>
<tr>
<td>$s_i$</td>
<td></td>
</tr>
<tr>
<td>$s_i - 1$</td>
<td>0</td>
</tr>
</tbody>
</table>

Unbiased estimators for $\alpha_u^i - \alpha_{u'}^i$ for $u \neq u'$; $u, u' = 0, 1, \ldots, s_i - 1$. can now easily be constructed, by considering differences in responses in each block. The above is true for every $i$ and hence the design $D$ is a connected main effects plan.

It is possible to construct connected main effects plans using other Pareto Optimal subsets and the statistical optimality issues remain an open problem for further investigation.
5. CONCLUDING REMARKS

A survey was done by the authors and Ms. Julie Hargreaves on job selection based on job attributes using both Pareto Optimal and non-Pareto optimal choice sets and the results were presented at the Marketing Science Meetings, Sydney, Australia in June, 1995. The results do show that when a dominating alternative is placed into a set almost everyone picks it and the reverse for dominated alternatives. With Pareto optimal choice sets, it is not the case. The use of Pareto optimal choice sets is to account for the behavioral issues and is difficult to statistically distinguish Pareto optimal and non-Pareto optimal choice sets. The consequences of non-Pareto optimality will be manifested in the response variable and not in the design. The main focus of the present paper is to consider the estimability problems while using Pareto optimal choice sets.

ACKNOWLEDGEMENT

The authors wish to thank Professor Milton Parnes for a helpful comment leading to the proof of Theorem 2.2.

BIBLIOGRAPHY


Please forward your requests for working papers to the following address:

Executive Officer
Department of Marketing
School of Business and Management
Hong Kong University of Science & Technology
Clear Water Bay
Kowloon, Hong Kong

www: http://www.bm.ust.hk/~mark/
Enquiries: Tel (852) 2358 7700 Fax (852) 2358 2429