What is the task?
- Show a product described by $k$ attributes.
- Attribute $q$ is shown at one of its $l_q$ levels.
- Respondents are asked which feature of the product is *best*.
- and which is *worst*.
- Repeat for several product descriptions.

Contraceptive products
Describe contraceptives by using 7 attributes
- Product
- Effect on acne
- Effect on weight
- Frequency of administration
- Contraceptive effectiveness
- Effect on bleeding
- Cost

The levels of *frequency of administration* and *contraceptive effectiveness* are nested within Product.
Below is a hypothetical contraceptive product which has the features described. Please read the description of the product and choose which in your opinion is the best feature of this product and which is the worst.

<table>
<thead>
<tr>
<th>Best</th>
<th>Worst</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra-uterine hormonal device</td>
<td>Has no effect on acne symptoms</td>
</tr>
<tr>
<td>Has no effect on acne symptoms</td>
<td>May lose up to 1kg in weight</td>
</tr>
<tr>
<td>May lose up to 1kg in weight</td>
<td>Administered once per year</td>
</tr>
<tr>
<td>Administered once per year</td>
<td>1 in 1000 women using this product get pregnant in a 12 month period</td>
</tr>
<tr>
<td>1 in 1000 women using this product get pregnant in a 12 month period</td>
<td>Cost is $7 per month</td>
</tr>
<tr>
<td>Cost is $7 per month</td>
<td>Most women using this product experience irregular bleeding</td>
</tr>
</tbody>
</table>

**Motivation for the task**

- Allows direct comparison of all levels across all attributes.
- Can average effects of all levels for an attribute and talk of the attribute impact.
- Can be more acceptable task than the usual DCE, particularly if a choice set has multiple implausible options.
- Even if all options are plausible, task may be less cognitively demanding for some respondents.

**The choice set**

Given a profile \((x_1, x_2, \ldots, x_k) = \mathbf{x}\) the implicit choice set of pairs from which a respondent is making a choice is given by:

\[ C_\mathbf{x} = \{(x_1, x_2), (x_1, x_3), \ldots, (x_1, x_k), (x_2, x_3), \ldots, (x_{k-1}, x_k), (x_2, x_1), (x_3, x_1), \ldots, (x_k, x_1), (x_3, x_2), \ldots, (x_k, x_{k-1})\} \]

**Attribute-level maxdiff model**

Marley, Flynn and Louviere (2008)

\(BW_\mathbf{x}(x_i, x_j)\) - probability that jointly (level \(x_i\) of factor \(F_i\) is chosen as best, level \(x_j\) of factor \(F_j\) is chosen as worst) from profile \(\mathbf{x}\).

\(BW_\mathbf{x} - \text{best-worst choice probability for profile } \mathbf{x}\).

Define \(BW_\mathbf{x} \forall \mathbf{x} \in P\)

Satisfies attribute-level maxdiff model iff \(\exists\) positive scale \(b\) on the attributes such that for every \(\mathbf{x} \in P\) and for any two distinct factors,

\[ BW_\mathbf{x}(x_i, x_j) = \frac{b(x_i)/b(x_j)}{\sum_{q=1}^k \sum_{r=1,r \neq q}^k (b(x_q)/b(x_r))} \]
Incorporating attributes

Let \( b(x_i) = \exp[y_i + \beta_{F_i,x_i}] \).

Then the set of best-worst choice probabilities, \( BW_x \), satisfies

- 2-invertibility; that is,
  \[ BW_x(x_i, x_j) = BW_x(x_j, x_i) = BW_x(x_i, x_q) \]

where \( 1 \leq i, j, q \leq k, i \neq j \) and \( q \neq r \);

- 3-reversibility; that is,
  \[ BW_x(x_i, x_j)BW_x(y_j, y_q) = BW_x(z_q, z_i) \]

where \( x_j = y_j, y_q = z_q \) and \( z_i = x_i \), and \( i \neq j \) and \( j \neq q \);

- 4-reversibility; that is,
  \[ BW_x(x_i, x_j)BW_y(y_j, y_q)BW_x(z_q, z_i)BW_w(w_r, w_i) = BW_w(w_i, w_r)BW_x(z_i, z_r)BW_y(y_q, y_j)BW_x(x_j, x_i) \]

where \( x_j = y_j, y_q = z_q, z_r = w_r \) and \( w_i = x_i \), and \( i \neq j, j \neq q \) and \( q \neq r \).

so this \( b \) gives an attribute-level maxdiff model.

Definition of \( \Lambda \)

\[
n_C(x_1, x_2, \ldots, x_k) = n_C = \begin{cases} 1 & \text{if profile } (x_1, x_2, \ldots, x_k) \text{ in BW task,} \\ 0 & \text{if profile } (x_1, x_2, \ldots, x_k) \text{ not in BW task.} \end{cases}
\]

\( N = \sum C n_C \) is the number of profiles in the experiment.

Let \( t = (x_q, x_{q_2}), s = (x_{q_3}, x_{q_4}) \)

\[
\Lambda_{t,t} = \frac{\pi_t}{N} \sum_{C \cap t \subseteq C} \frac{n_C \sum_{w \in C \mid w \neq t} \pi_w}{\left( \sum_{w \in C} \pi_w \right)^2}
\]

and

\[
\Lambda_{t,s} = \frac{-\pi_t \pi_s}{N} \sum_{C \cap t \subseteq C \cap s \subseteq C} \frac{n_C}{\left( \sum_{w \in C} \pi_w \right)^2}.
\]

Link with the MNL model

Define

\[
\pi(F_t x_i, F_j x_j) = \frac{b(x_i)}{b(x_j)} = \exp[\beta_{F_i,x_i} - (\beta_{F_j,x_i})]
\]

\( \pi \) - vector containing the distinct \( \pi(F_t x_i, F_j x_j) \)

\( \gamma = \ln(\pi) \)

As defined \( \Lambda \)

is information matrix for \( \gamma \)

for MNL model

selection probabilities given by \( BW_x \)

for profiles in \( P \) and corresponding choice sets \( C_x \)
The entries in $\Lambda$ when $\pi = 1$

$$\Lambda_{t,s} = \frac{k(k-1)-1}{N(k(k-1))^2} \sum_{\{C|tC\subseteq C\}} n_C$$

and

$$\Lambda_{t,s} = \frac{-1}{N(k(k-1))^2} \sum_{\{C|tC, s \subseteq C\}} n_C.$$ 

The entries in $36\Lambda$ from profile 0 1 1 when $\pi = 1$

Row 0112 has 5 in column 0112.

and has -1 in columns 0113, 1213, 1201, 1301, and 1312

and has 0 in all other columns.

Rows 0113, 1213, 1201, 1301, and 1312 also have 5 on diagonal and -1 in columns corresponding to the entries in the implicit choice set.

All other rows are 0 throughout.

Example $k = 3$ attributes, each with $\ell_q = 2$ levels

Complete set of profiles is:
000, 001, 010, 011, 100, 101, 110, 111

Consider profile 0 1 1

Implied choice set is (subscript is attribute):
\{0112, 0113, 1213, 1201, 1301, 1312\} = C

Number of levels = 2+2+2 = 6

Number of pairs of levels = 2 x 4 x 3 = 24

Labels of rows and columns of $\Lambda$:

0102, 0112, 1102, 1112, 0103, 0113, 1103, 1113,
0203, 0213, 1203, 1213, 0201, 0211, 1201, 1211,
0301, 0311, 1301, 1311, 0302, 0312, 1302, 1312

The first 12 rows of $36\Lambda$:

<table>
<thead>
<tr>
<th>01</th>
<th>02</th>
<th>03</th>
<th>04</th>
<th>05</th>
<th>06</th>
<th>07</th>
<th>08</th>
<th>09</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Bringing back $\beta$

\[
\gamma(F_i, x_i, y_i) = \beta_{F_i} + \beta_{F_i x_i} - (\beta_{F_i} + \beta_{F_i x_i})
\]

Transform $\Lambda$ to the information matrix for the $\beta_{F_i}$ and the $\beta_{F_i x_i}$.

Let

\[
\beta' = (\beta_{F_1}, \beta_{F_2}, \ldots, \beta_{F_k}, \beta_{F_1 l_1}, \ldots, \beta_{F_k l_k-1}, \ldots, \beta_{F_k l_k-1}),
\]

Define $X$ by $\gamma = X \beta$.

---

Example $k = 3$ attributes, each with $\ell_q = 2$ levels (cont)

\[
\beta' = (\beta_{F_1}, \beta_{F_2}, \beta_{F_3}, \beta_{F_1 l_1}, \beta_{F_1 l_2}, \beta_{F_2 l_1}, \beta_{F_2 l_2}, \beta_{F_3 l_1}),
\]

with corresponding row of $X$:

\[
1 - 0 1 0 - 0 0
\]
The design question

What set of profiles should we show to get as much information as possible about the scale of the attribute levels?

If we are only fitting main effects, can we get as much information about the attribute levels by showing (the right) subset of all possible combinations of attribute levels, or do we have to show all possible combinations of attribute levels?

Example  $k = 3$ attributes, each with $\ell_q = 2$ levels

Complete set of profiles is:
000, 001, 010, 011, 100, 101, 110, 111

Each profile can be included or not in a BW task.

So there are $2^8 - 1 = 255$ different possible designs to compare.

For each of these calculate $\det(F)$.

Comparing Designs

To compare designs for their ability to estimate $\beta$s, we need $F$, information matrix for $\beta$s. Get this by calculating
- information matrix for $\gamma$, $\Lambda$,
- the reduced matrix $R$
- Evaluate $F = R'\Lambda R$.

Use the generalised variance, that is, the determinant of the variance-covariance matrix, to compare designs.

The $D$-optimal design has the smallest determinant of the variance-covariance matrix.

Equivalently, it has the largest determinant for the information matrix.
The best designs are:

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0 0</td>
<td>0 1</td>
<td>0 0</td>
<td>0 0</td>
<td>0 1</td>
<td>0 0</td>
<td>0 1</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td>0 1</td>
<td>0 1</td>
<td>0 0</td>
<td>0 1</td>
<td>0 0</td>
<td>0 1</td>
<td>0 0</td>
<td>0 1</td>
<td>0 0</td>
</tr>
<tr>
<td>1 0</td>
<td>1 0</td>
<td>0 1</td>
<td>1 0</td>
<td>0 1</td>
<td>1 0</td>
<td>0 1</td>
<td>1 0</td>
<td>0 1</td>
</tr>
<tr>
<td>1 1</td>
<td>1 1</td>
<td>0 1</td>
<td>1 1</td>
<td>0 1</td>
<td>1 1</td>
<td>0 1</td>
<td>1 1</td>
<td>0 1</td>
</tr>
</tbody>
</table>

Orthogonal array - \( \text{OA}[N; \ell_1, \ell_2, \ldots, \ell_k; t] \)

\( N \times k \) array with elements from a set of \( \ell_q \) symbols in column \( q \) such that any \( N \times t \) subarray has each \( t \)-tuple as a row equally often. \( t \) is the strength of the OA.

\( \text{OA}[N = 8; \ell_1 = 2, \ell_2 = 2, \ell_3 = 2, \ell_4 = 2, \ell_5 = 4; t = 2] \)

\[
\begin{array}{cccc}
0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 \\
0 & 1 & 0 & 1 \\
0 & 1 & 1 & 0 \\
1 & 0 & 0 & 1 \\
1 & 0 & 1 & 0 \\
1 & 1 & 0 & 0 \\
1 & 1 & 1 & 0 \\
\end{array}
\]

\( \text{OA}[N = 18; \ell_1 = 3, \ell_2 = 3, \ell_3 = 3, \ell_4 = 3; t = 2] \)

subdivided into \( \text{OA}[N = 9; \ell_1 = 3, \ell_2 = 3, \ell_3 = 3, \ell_4 = 3; t = 2] \)

\[
\begin{array}{cccc}
0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 \\
0 & 2 & 2 & 2 \\
0 & 2 & 2 & 0 \\
1 & 0 & 2 & 1 \\
1 & 0 & 2 & 2 \\
1 & 1 & 0 & 2 \\
1 & 1 & 0 & 0 \\
1 & 2 & 1 & 0 \\
1 & 2 & 1 & 1 \\
2 & 0 & 1 & 2 \\
2 & 0 & 1 & 0 \\
2 & 1 & 2 & 0 \\
2 & 1 & 2 & 1 \\
2 & 2 & 0 & 1 \\
2 & 2 & 0 & 2 \\
\end{array}
\]
What is the best-worst choice task?
The model
The design question
Contraceptive Choices Example

General result when $\pi = 1$
If respondents see all the treatment combinations in the complete factorial, or just those in an orthogonal array with $t = 2$, then the information matrix for the attribute-level maxdiff model is

$$
\begin{bmatrix}
\frac{2}{k-1} l_{k-1} & \frac{2}{k(k-1)} J_{k-1} & \cdots & 0_k - 1, \ell_{k-1} & \cdots & 0_k - 1, \ell_{k-1} \\
0_{\ell_{k-1}, k-1} & \frac{1}{\ell_{k-1}(a-\ell_{k-1})} l_{\ell_{k-1}} & \cdots & 0_{\ell_{k-1}, \ell_{k-1}} & \cdots & 0_{\ell_{k-1}, \ell_{k-1}} \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
0_{\ell_{k-1}, k-1} & 0_{\ell_{k-1}, \ell_{k-1}} & \cdots & \frac{1}{\ell_{k-1}(a-\ell_{k-1})} l_{\ell_{k-1}} & \cdots & 0_{\ell_{k-1}, \ell_{k-1}}
\end{bmatrix}
$$

where $a = \sum q \ell_q$.

Performance when $(0_1, 0_2, 0_3, 1_1, 1_2, 1_3) = (1, 1, 1, 1, 1, 1)$

Performance under alternative hypotheses
$(0_1, 0_2, 0_3, 1_1, 1_2, 1_3) = (1.2, 1.1, 1.1, 0.8)$

Performance under alternative hypotheses
$(0_1, 0_2, 0_3, 1_1, 1_2, 1_3) = (1.2, 1.2, 1.2, 0.8, 0.8, 0.8)$
What is the best-worst choice task?

The model

The design question

Contraceptive Choices Example

Performance under alternative hypotheses

$$(0_1, 0_2, 0_3, 1_1, 1_2, 1_3) = (1.4, 1.2, 1.0, 9, 0.7, 0.5)$$

How much is lost by using a non-optimal design?

$$\begin{array}{cccccc}
(0_1, 0_2, 0_3, 1_1, 1_2, 1_3) & 105 & 131 & 219 & 255 \\
(1,1,1,1,1,1) & 1 & 0.830 & 0.985 & 1 \\
(1.6,1.5,1.4,0.6,0.5,0.4) & 0.997 & 0.948 & 1 & 0.997 \\
(2.0,1.9,1.8,0.4,0.3,0.2) & 0.940 & 1 & 0.962 & 0.941 \\
\end{array}$$

Design

000 001 010 011 100 101 110 111

105 0 1 1 0 1 0 0 1
150 1 0 0 1 0 1 1 0
131 1 0 0 0 0 0 1 1
137 1 0 0 0 0 0 1 1
219 1 1 0 1 1 0 1 1
255 1 1 1 1 1 1 1 1
Can all level orderings be recovered?

Let \( k = 3 \) and \( \ell_1 = \ell_2 = \ell_3 = 2 \)

Suppose that \( 0_1 > 0_2 > 0_3 > 1_1 > 1_2 > 1_3 \).

<table>
<thead>
<tr>
<th>Choice set</th>
<th>Choice</th>
<th>Ordering</th>
</tr>
</thead>
<tbody>
<tr>
<td>010203</td>
<td>0103</td>
<td>01 &gt; 02 &gt; 03</td>
</tr>
<tr>
<td>011213</td>
<td>0113</td>
<td>01 &gt; 12 &gt; 13</td>
</tr>
<tr>
<td>110213</td>
<td>0213</td>
<td>02 &gt; 11 &gt; 13</td>
</tr>
<tr>
<td>111203</td>
<td>0312</td>
<td>03 &gt; 11 &gt; 12</td>
</tr>
</tbody>
</table>

from which we can recover the original ordering.

Practical design advice

Find several candidate designs that do well under the null hypothesis, since then it appears likely that you will be able to estimate everything for any values.

Check by doing a simulation study that a range of values for the levels can be recovered.

Recovering Orderings (cntd)

Let \( k = 3 \) and \( \ell_1 = \ell_2 = \ell_3 = 2 \)

Suppose that \( 0_1 > 0_2 > 0_3 > 1_1 > 1_2 > 1_3 \).

<table>
<thead>
<tr>
<th>Choice set</th>
<th>Choice</th>
<th>Ordering</th>
</tr>
</thead>
<tbody>
<tr>
<td>010203</td>
<td>0103</td>
<td>01 &gt; 02 &gt; 03</td>
</tr>
<tr>
<td>011213</td>
<td>0113</td>
<td>01 &gt; 12 &gt; 13</td>
</tr>
<tr>
<td>110213</td>
<td>0213</td>
<td>02 &gt; 11 &gt; 13</td>
</tr>
<tr>
<td>111203</td>
<td>0312</td>
<td>03 &gt; 11 &gt; 12</td>
</tr>
</tbody>
</table>

from which we can recover

\( 0_1 > 0_2 > 0_3 > 1_1 > 1_2 > 1_3 \) and \( 0_1 > 0_2 > 0_3 > 1_1 > 1_2 \)

but we do not know the relative rankings of \( 0_2 \) and \( 1_2 \).

Having the other 4 profiles from the complete factorial will not help sort this out since \( 0_2 \) and \( 1_2 \) are never in the same profile.

Could be addressed by using other designs that did not just consider profiles.

Contraceptive Choices BW task - The attributes

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Number of levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product</td>
<td>8</td>
</tr>
<tr>
<td>Acne</td>
<td>3</td>
</tr>
<tr>
<td>Weight</td>
<td>4</td>
</tr>
<tr>
<td>Frequency of administration</td>
<td>3</td>
</tr>
<tr>
<td>Contraceptive effectiveness</td>
<td>3</td>
</tr>
<tr>
<td>Effect on bleeding</td>
<td>8</td>
</tr>
<tr>
<td>Cost</td>
<td>4</td>
</tr>
</tbody>
</table>

Levels of frequency and effectiveness are nested in product.
What is the best-worst choice task?

The design

- Find an orthogonal array for all attributes except product.
- This has 32 level descriptors.
- Each of these 32 is linked with each of the 8 products to get $8 \times 32 = 256$ descriptions in total.
- These are then subdivided into 16 versions of 16 descriptions, with two descriptions for each product in each version and 16 different level descriptors.
- Each woman completes one version.
- Ability to recapture “known” prior values confirmed by a simulation study.

Simulation study

- Choose values for $\beta_q$ and $\beta_q x_q$ (or for $\beta_q + \beta_q x_q$).
- Calculate the value of $\gamma(x_q, x_q)$ for each pair in the implied choice set.
- Add a random error term of extreme value type 1.
- Choose the pair with the largest value.
- Analyse the results using the MNL.
- Repeat 100 times.
- Compare estimated and assumed terms.
What is the best-worst choice task? The model The design question Contraceptive Choices Example

Open Questions

Make a systematic general comparison of the performance of designs under alternative hypotheses.

We have been talking about the max-diff model. What are good designs for the weighted max-diff model?
References


