Estimation of Logit and Probit models using best, worst and best-worst choices

Paolo Delle Site*

*Niccolò Cusano University Rome

Dottorato in Infrastrutture e Trasporti, Sapienza University Rome
Best choices in stated preference surveys are relevant to market

Beggs, Cardell and Hausman (1981): Rank-Ordered Logit - ROL (aka exploded logit)

Potential advantages of rank-ordered data:
  - more data per respondent from each choice task
  - greater precision for model estimates

Best-Worst (BW) scaling: data collection approach including W choice in addition to B choice
BW scaling literature

- Finn and Louviere (1992): max-difference model
- Anderson and de Palma (1999): reverse Logit model
- Marley and Louviere (2005): axiomatic approach to probabilistic models of B, W and B&W choices
- de Palma, Kilani and Laffond (2017): relations between B, W and B&W choices
Contributions

- Simultaneous B&W case
- Review of identities for B, W and B&W probabilities: Logit and Probit
- Estimation of Logit and Probit
- Experiment with modal responses to congestion charging in urban area
- Comparisons B, W and B&W: estimates, precision, predictive accuracy, Value of Time (VOT)
General identities for Random Utility Models - RUM

- **Probability that \( x \) is chosen as \( B \)**

\[
B_X(x) = \mathbb{P}\left\{ U_x = \max_{z \in X} U_z \right\}, \ x \in X
\]

- **Probability that \( x \) is chosen as \( W \)**

\[
W_X(x) = \mathbb{P}\left\{ U_x = \min_{z \in X} U_z \right\}, \ x \in X
\]

- **Probability that \( x \) is chosen as \( B \) and \( y \) as \( W \)**

\[
BW_X(x, y) = \mathbb{P}\left\{ U_x \geq U_z \geq U_y, z \in X \setminus \{x, y\} \right\}, \ x, y \in X, x \neq y
\]
Identities for additive RUMs

- Total utilities

\[ U_z = v_z + \varepsilon_z, \ z \in X \]

- Probability that \( x \) is chosen as \( W \) (de Palma, Kilani and Laffond, 2017, theorem 5)

\[ W_X (x) = \sum_{\{x\} \subseteq Y \subseteq X} (-1)^{|Y|-1} B_Y (x), \ x \in X \]
Identities for additive and reverse (subtractive) RUMs

- **Total utilities in reverse RUMs**
  \[ U^R_z = v_z - \varepsilon_z, \quad z \in X \]

- **Probability that \( x \) is chosen as \( B \) in reverse RUMs**
  \[ B^R_X (x, v) = W_X (x, -v), \quad x \in X \]

- **Probability that \( x \) is chosen as \( W \) in reverse RUMs**
  \[ W^R_X (x, v) = B_X (x, -v), \quad x \in X \]

- **Probability that \( x \) is chosen as \( B \) and \( y \) as \( W \) in reverse RUMs**
  \[ BW^R_X (x, y, v) = BW_X (y, x, -v), \quad x, y \in X, \quad x \neq y \]
Identities for additive Logit

- Random terms i.i.d. according to
  \[ P(\varepsilon_z \leq t) = \exp(-e^{-t}), \ -\infty < t < \infty, \ z \in X \]

- Probability that \( x \) is chosen as \( B \)
  \[ B_X(x) = \frac{e^{v_x}}{\sum_{z \in X} e^{v_z}}, \ x \in X \]

- Probability that \( x \) is chosen as \( W \)
  \[ W_X(x) = \sum_{\{x\} \subseteq Y \subseteq X} (-1)^{|Y|-1} \frac{e^{v_x}}{\sum_{z \in Y} e^{v_z}}, \ x \in X \]
Identities for additive Logit (cont.)

- Probability that $x$ is chosen as $B$ and $y$ as $W$ (Marley and Louviere, 2005, proposition 9)
  \[ BW_X(x,y) = B_X(x) W_{X\setminus\{x\}}(y), \quad x, y \in X, \quad x \neq y \]

- which yields
  \[ BW_X(x,y) = \frac{e^{v_x}}{\sum_{z \in X} e^{v_z}} \sum_{\{y\} \subseteq Y \subseteq X \setminus \{x\}} (-1)^{|Y|-1} \frac{e^{v_y}}{\sum_{z \in Y} e^{v_z}}, \quad x, y \in X, \quad x \neq y \]
Identities for reverse Logit

- Probability that \( x \) is chosen as \( B \)

\[
B_X(x) = \sum_{\{x\} \subseteq Y \subseteq X} (-1)^{|Y|-1} \frac{e^{-v_x}}{\sum_{z \in Y} e^{-v_z}}, x \in X
\]

- Probability that \( x \) is chosen as \( W \)

\[
W_X(x) = \frac{e^{-v_x}}{\sum_{z \in X} e^{-v_z}}, x \in X
\]

- Probability that \( x \) is chosen as \( B \) and \( y \) as \( W \)

\[
BW_X(x,y) = \frac{e^{-v_y}}{\sum_{z \in X} e^{-v_z}} \sum_{\{x\} \subseteq Y \subseteq X \setminus \{y\}} (-1)^{|Y|-1} \frac{e^{-v_x}}{\sum_{z \in Y} e^{-v_z}}, x, y \in X, x \neq y
\]
Estimation of additive Logit

- Systematic utilities

\[ v_z = \xi_z \theta, \quad z \in X \]

- \( \mathcal{M} \) be the matrix obtained by concatenating the rows of the \( N \) observation-specific \( [(|X| - 1) \times K] \) matrices whose rows are \( \xi_y - \xi_x, \quad y \in X \setminus \{x\} \), where \( x \in X \) is any arbitrarily chosen alternative in \( X \)

**Theorem 1**

In the linear-in-the-coefficients additive Logit, both the total log-likelihood functions, of the best and worst choices \( LL^B \) and \( LL^W \), are strictly concave in \( \theta \) iff the matrix \( \mathcal{M} \) has full rank.
The attribute matrix

- Observation-specific sub-matrix

\[
\begin{bmatrix}
\xi_{y,1} - \xi_{x,1} & \cdots & \xi_{y,K} - \xi_{x,K} \\
\cdots & \cdots & \cdots \\
\xi_{y|X|-1,1} - \xi_{x,1} & \cdots & \xi_{y|X|-1,K} - \xi_{x,K}
\end{bmatrix}
\]

- Matrix $\mathcal{M}$ has full rank $\iff$ all columns are linearly independent $\iff$ no multi-collinearity in attributes
Identities for Probit

- Random terms distributed according to multivariate normal

**Proposition 1**
For additive Probit, the following identity between B choice and W choice probability holds
\[ W_X(x, v) = B_X(x, -v), \; x \in X \]

**Proposition 2**
For Probit, the following identities relating to B choice and W choice probabilities between additive and reverse model hold:
\[ B^R_X(x, v) = B_X(x, v), \; W^R_X(x, v) = W_X(x, v), \; x \in X \]
Estimation of Probit

- Systematic utilities

\[ v_z = \xi_z \theta, \ z \in X \]

- \( \mathcal{M} \) be the matrix obtained by concatenating the rows of the \( N \) observation-specific \( [(|X| - 1) \times K] \) matrices whose rows are \( \xi_y - \xi_x, \ y \in X \setminus \{x\} \), where \( x \in X \) is any arbitrarily chosen alternative in \( X \)

**Theorem 2**

In the linear-in-the-coefficients Probit, both the total log-likelihood functions, of the best and worst choices \( LL^B \) and \( LL^W \), are strictly concave in \( \theta \) iff the matrix \( \mathcal{M} \) has full rank.
Numerical evidence

Case study:

- Rome Urban Traffic Master Plan (PGTU - Piano Generale del Traffico Urbano), 2015
- provisions for congestion charging (daily charge) in the Railway Ring (Anello Ferroviario) area (34 km²)

Survey:

- 146 car users interviewed in Spring 2015
- computer-assisted face-to-face interviews
- questionnaire: revealed preference + stated preference + personal characteristics
Rome: planned charged area
Attributes and levels in the choice experiment

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Alt.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>car (EUR/day)</td>
</tr>
<tr>
<td>access charge</td>
<td>0, 2, 4, 6</td>
</tr>
<tr>
<td>travel time (minutes)</td>
<td>RP</td>
</tr>
</tbody>
</table>

RP = revealed preference
Example of choice task

<table>
<thead>
<tr>
<th>attributes/alternatives</th>
<th>car</th>
<th>PTW</th>
<th>public transportation</th>
</tr>
</thead>
<tbody>
<tr>
<td>access charge (EUR/day)</td>
<td>4</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>travel time (minutes)</td>
<td>30</td>
<td>20</td>
<td>45</td>
</tr>
<tr>
<td>best alternative</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>worst alternative</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>
## Estimation results - additive Logit

<table>
<thead>
<tr>
<th></th>
<th>Alt.</th>
<th>Coefficient (t-stat)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>B</td>
</tr>
<tr>
<td>charge (EUR/day)</td>
<td>1,2</td>
<td>-.947***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-7.85)</td>
</tr>
<tr>
<td>travel time (minutes)</td>
<td>1,2,3</td>
<td>-.095***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-9.82)</td>
</tr>
<tr>
<td>alt. specific constant</td>
<td>2</td>
<td>-2.18***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-11.06)</td>
</tr>
<tr>
<td>alt. specific constant</td>
<td>3</td>
<td>-.453***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-2.69)</td>
</tr>
<tr>
<td>travel time × gender (female = 1)</td>
<td>1,2,3</td>
<td>.043***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4.81)</td>
</tr>
<tr>
<td>charge × income (six income classes: 1,...,6)</td>
<td>1,2</td>
<td>.121***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4.20)</td>
</tr>
</tbody>
</table>

583 (pseudo)observations

Key: ***,**, * = statistical significance at 1%, 5%, 10% level
W choice coefficients plotted against B choice coefficients - additive Logit
<table>
<thead>
<tr>
<th>Net income (EUR/month)</th>
<th>B</th>
<th>W</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>(t-stat.)</td>
</tr>
<tr>
<td><strong>Male</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤500</td>
<td>6.9</td>
<td>(6.74)</td>
</tr>
<tr>
<td>501-1000</td>
<td>8.1</td>
<td>(7.46)</td>
</tr>
<tr>
<td>1001-1500</td>
<td>9.8</td>
<td>(8.30)</td>
</tr>
<tr>
<td>1501-2500</td>
<td>12.4</td>
<td>(8.35)</td>
</tr>
<tr>
<td>2501-5000</td>
<td>16.9</td>
<td>(5.91)</td>
</tr>
<tr>
<td>&gt;5000</td>
<td>26.5</td>
<td>(2.89)</td>
</tr>
<tr>
<td><strong>Female</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤500</td>
<td>3.7</td>
<td>(7.02)</td>
</tr>
<tr>
<td>501-1000</td>
<td>4.4</td>
<td>(7.56)</td>
</tr>
<tr>
<td>1001-1500</td>
<td>5.3</td>
<td>(7.96)</td>
</tr>
<tr>
<td>1501-2500</td>
<td>6.7</td>
<td>(7.39)</td>
</tr>
<tr>
<td>2501-5000</td>
<td>9.2</td>
<td>(5.18)</td>
</tr>
<tr>
<td>&gt;5000</td>
<td>14.4</td>
<td>(2.69)</td>
</tr>
</tbody>
</table>
Coefficient attenuation: an empirical regularity

- Ben-Akiva, Morikawa and Shiroishi (1992): route alternatives in public transport network
  - differences in variances are statistically significant, increasing variance when moving to W choice
  - coefficients are statistically different even accounting for different scale parameters

- Giergiczny, Dekker, Hess and Chintakayala (2017): route alternatives in public transport network
  - willingness to pay for delay reduction higher for B choices
Conclusions

Key results:

- identities for Probit, conditions for strict logconcavity of B and W Logit and Probit likelihood
- increase of precision in B&W estimates not found
- coefficient attenuation from B to W confirmed in additive and reverse Logit and in Probit
- VOT larger for W choices and with larger confidence intervals in additive and reverse Logit and in Probit