The Complementary Benefits of Stated Preference and Revealed Preference for Choice Modeling: Theory and Practice

Jeffrey Dumont & Stacey Falzarano

Prepared for:
The Revenue Management Workshop

5/21/2012
“Revealed preference” data
- Are actual choices made by travelers
  - Either observed directly or
  - Self-reported, such as via a survey
- Have a long history in revenue management applications in the airline industry

Benefits of revealed preference data and models
- Actual choices made in the real market environment
- Data are often available quickly, in large quantities

However, they do have their limitations
- The chosen alternative is known with certainty but often little is known about the alternatives the traveler considered and did not choose
- There is typically correlation (sometimes significant) between effects in the data
- Reflect the existing market space only
What is stated preference?

Comes in many names and varieties

- Stated Choice (SC)
- Stated Intention (SI)
- Discrete Choice Conjoint
- Direct Utility Assessment (DUA)
- Choice Based Conjoint (CBC)
- Adaptive Choice Based Conjoint Adaptive (ACBC)
- Maximum Difference Scaling (Max-Diff)
- Best-Worst Conjoint (Best-Worst)
- Adaptive Conjoint (ACA)
What is stated preference?

Stated preference data

- Are obtained when a researcher collects data
  - On what travelers, consumers, or decision makers say they *would* do
  - In a given, hypothetical choice scenario (experiment)

- Are typically collected via a primary research instrument like a survey

  **The survey asks questions like:**
  1. Which of these two flights would you choose to make your trip?
  2. Which of these vehicles would you buy?
  3. Please select the your most preferred and least preferred product from the list below.
Examples

If these options were the only options available for your weekday social or recreational trip which would you choose? Click on one of the two boxes below to select your preferred choice.

Information in green may have changed.

<table>
<thead>
<tr>
<th>Tolled route</th>
<th>Current route</th>
</tr>
</thead>
<tbody>
<tr>
<td>Travel time: 56 minutes</td>
<td>Travel time: 1 hour and 8 minutes</td>
</tr>
<tr>
<td>Toll cost: $2.45</td>
<td>Toll free</td>
</tr>
</tbody>
</table>

Next Question
Examples

Below is a set of 4 factors that you might consider when choosing among the types of vehicles that you said you were interested in considering.

On the left, please indicate which one of the factors would be **most important to you**. Then on the right, please indicate which one of the factors would be **least important to you**.

<table>
<thead>
<tr>
<th>Most Important (pick one)</th>
<th>Least Important (pick one)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Safety features (airbags, traction control, anti-roll-over)</td>
<td></td>
</tr>
<tr>
<td>Performance (acceleration/top speed)</td>
<td></td>
</tr>
<tr>
<td>Fuel economy (MPG)</td>
<td></td>
</tr>
<tr>
<td>Type of vehicle (sedan, SUV, truck)</td>
<td></td>
</tr>
</tbody>
</table>

(2 of 6)

My New Vehicle

**Type**
- Small cross-over SUV

**Brand**
1. Land Rover
2. Lincoln
3. BMW

**Price Range**
- $20,000 - $24,999

Next Question
Which of these two alternatives would you have preferred on your trip from the Manchester Regional Airport (MHT) to the Minneapolis St Paul International Airport (MSP)?

Note: Flight information may change from screen to screen.

<table>
<thead>
<tr>
<th></th>
<th>Option A</th>
<th>Option B</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Airline</strong></td>
<td>Southwest Airlines</td>
<td>JetBlue Airways</td>
</tr>
<tr>
<td><strong>Aircraft Type</strong></td>
<td>Standard Jet</td>
<td>Regional Jet</td>
</tr>
<tr>
<td><strong>Flight Departure Time</strong></td>
<td>6:55 AM Eastern Time</td>
<td>9:30 AM Eastern Time</td>
</tr>
<tr>
<td><strong>Number of Connections</strong></td>
<td>1 connection</td>
<td>Direct flight</td>
</tr>
<tr>
<td><strong>Total Travel Time</strong></td>
<td>4 hours and 55 minutes</td>
<td>4 hours and 20 minutes</td>
</tr>
<tr>
<td><strong>Flight Arrival Time</strong></td>
<td>10:50 AM Central Time</td>
<td>12:50 PM Central Time</td>
</tr>
<tr>
<td><strong>On-Time Performance</strong></td>
<td>70% of these flights are on time</td>
<td>90% of these flights are on time</td>
</tr>
<tr>
<td><strong>One-Way Fare</strong></td>
<td>$225</td>
<td>$95</td>
</tr>
</tbody>
</table>

(Question 5 of 8)
Benefits of stated preference

- Provides reliable estimates of the relative importance of each of the features
- Enables testing new products or attribute levels that do not currently exist in the market
- Minimizes confounding between estimates of effects
  - Experimental designs allow for statistically efficient recovery of effects
- Ensures full knowledge of the alternatives presented to choice makers
- Allows robust understanding of how individuals make choices
  - By observing multiple choices from one individual
- Enriches the choice model by easily matching choice behavior with socio-demographics and psychographic data
How good are SP data?

State Route 91 in California

• One of the most congested highway corridors in the U.S.

• Private initiative to build highly innovative all-electronic, dynamically-priced lanes in freeway median

• Many industry experts were skeptical of project feasibility: *Would travelers buy transponder and pay up to $5 for 10-mile trip with a competing free option?*
How good are SP data?

- Operating successfully for 10 years
- Our research indicated that drivers were willing to pay tolls equivalent to $13.40/hour of time savings: higher in PM, lower in AM
- Actual independently-measured value of time: $13.95* (within 4% of estimate)
- Need for transponder not significant barrier, but full dynamic pricing would have been problematic
- Developed revenue-optimized price structure with service targets, (reviewed monthly)
- $120 million project funded with revenue bonds
How good are SP data?

Device forecasting example

• A market forecasting model was developed to predict acquisition of electronic devices

• Model required minimal adjustments to replicate the current market place

• Adjustments that were necessary were intuitive in nature
How can RP data complement SP analyses?

- As a context/frame for stated preference experiments
- Calibration of a stated preference choice model
- Joint estimation of a RP/SP model
Using RP to Set the Choice Context

By asking detailed questions around a recent choice before showing stated preference experiments, the researcher:

• Can customize the choice experiments in a survey to a respondent’s unique choice situation resulting in more realistic trade-offs

• Present alternatives that resonate with the respondent

• More engagement from the choice maker

Typically requires an online (or other computer-based) survey instrument to implement such customization.
It is well-known that forecasts based solely on a stated preference model may not replicate existing behavior. There are good reasons for this.

In a survey choice experiment, there is increased:

- Awareness of the alternatives
- Knowledge of the benefits and costs associated with the decision
- Availability of the alternatives
When data on macro-level revealed behavior is available, adjustments to stated preference models can be made:

**Constant calibration**
- Market shares for existing products are often well-known by the researcher
- Additive constants can be introduced to the SP model to replicate these existing market shares
- Calibration constants often have a useful interpretation

**Scale Calibration**
- It is possible that changes in attribute-levels could induce a different stated behavior than what is observed in the actual market
- For example, price elasticities could be different
- A multiplicative calibration term can be introduced into the SP model to match this macro-level behavior
Joint Estimation

When revealed preference choice data is available at the individual decision maker level, estimation of a choice model on the pooled data is possible.

The researcher does not necessarily need revealed preference and stated preference data from the same set of individuals.

Typically have overlapping effects between the two data sources which can then be captured using common parameters between the two datasets.
Benefits of joint estimation:

- **Statistical efficiency** – more confidence in the parameters estimates because you are using all available data

- **Scale correction** – allows for the absolute sensitivities to be reflective of revealed behavior

- **Inclusion of new features/attributes** – can quantify the impact of new attributes or new attribute-levels
Case Study

Objective:
To help an online travel agency redefine their loyalty program to maximize profitability

Specific questions to answer:
1. How many tiers should be created in the loyalty program?
2. What specific benefits should be included in each tier?
3. What is the expected performance improvement by redefining the program in terms of profitability and retention?
4. Why are people not joining the program and what could be done to increase adoption?
Case Study

Approach

Model Existing Data

- Members
- Non-Members

Membership Choice Model

Model Survey Data

- Benefits Evaluation
  - Tier Definition
  - Program Optimization

Program Recommendations

Membership model will include examination of:
- Points Pressure
- Rewarded Behavior
- Enhanced Service
Case Study

Points Pressure

![Graph showing the percentage of customers booking as a function of accumulated points. The graph compares actual data (dashed red line) with a polynomial fit (solid black line).]
Case Study

Membership model will include examination of:
- Points Pressure
- Rewarded Behavior
- Enhanced Service