How to Develop a MaxDiff Typing Tool
to Assign New Cases into Meaningful Segments

Jay Magidson
Statistical Innovations

Gary Bennett
The Stats People

Statistical Innovations
Thinking outside the brackets!™

ART Forum 2016

the STATS PEOPLE

The Problem

- Segmentation studies often require production of a Typing Tool or “Golden Questions” algorithm to predict segment membership for respondents obtained in future surveys.
- Trade-offs exist between minimizing the number of predictors needed and maximizing the accuracy of prediction.
- For needs based segmentations generated from MaxDiff data there is currently no straightforward way to develop such an algorithm.
  - MaxDiff choice exercises typically use Best/Worst format on subsets of 3-6 options; it is difficult to employ even a reduced version in future surveys without substantial respondent fatigue.

Given the popularity of needs-based segmentations an improved solution is needed

Abstract

- To obtain meaningful segments, latent class analysis is often applied to MaxDiff (best-worst scaling) data – powerful way to identify segments with different needs.
- Marketers often find it valuable to include a relatively small number of "Golden Questions" in future surveys to classify respondents into the most likely of these Needs Based segments. However, the process of asking these questions with minimal respondent fatigue and still obtain accurate classifications is a challenge.
- In this presentation we describe a new approach to develop a MaxDiff typing tool and use real data to illustrate the steps of our proposed approach.
- Specifically, a relatively small number of *binary* items is included in future surveys, each item requesting a simple choice between a specified attribute pair.
- We propose an approach for selecting such items and the resulting classifications.
- Results suggest that based on these binary items, a high percentage of the original respondents could be expected to be assigned to the correct segments.
- The resulting typing tool is not only simpler but also improves over earlier approaches by allowing a customization which can provide higher accuracy for those segments judged most important to target.

- Latent class analysis and simulation is performed using Latent GOLD® Choice 5.1

Background: MaxDiff vs. Paired Comparisons

- **MaxDiff requires only 2 choices** – Best (Most preferred) and Worst (Least Preferred). Thus, MaxDiff is an efficient way to obtain preferences for 3 or more items.

Example: Choose Most and Least Preferred among A,B,C:

- If ask Best and Worst from triple (A,B,C):
- (‘A’ = Best) => A>B, A>C
- (‘B’ Worst) => C>B

2 CHOICES IMPLIES 3 PAIR PREFERENCES

- **Paired Comparisons requires 3 choices (A,B), (A,C), (B,C)**
- **The MaxDiff advantage in efficiency increases with K-tuples, K>3**

MaxDiff Case Study

Choose *highest* and *lowest* funding priorities:

Option	Improvement
1	More frequent off peak trains between major centres
2	Improved peak rail capacity
3	More frequent bus services on major routes
4	Extensions of light rail services
5	Integrated fares
6	Integrated ticketing
7	Real-time arrival information
8	New cycleways; more bike and scooter parking
9	Trains use green power

Presented as 12 sets of triples

* **Sydney transport project**: For further information on these data, see the book:
Best Worst Scaling: Theory, Methods and Applications , by Louviere, Flynn, & Marley (2015).

Choice Set #1 of 12`

Options displayed 3 at a time: Which improvement should receive:

a) Highest priority to fund?

b) Lowest priority to fund?

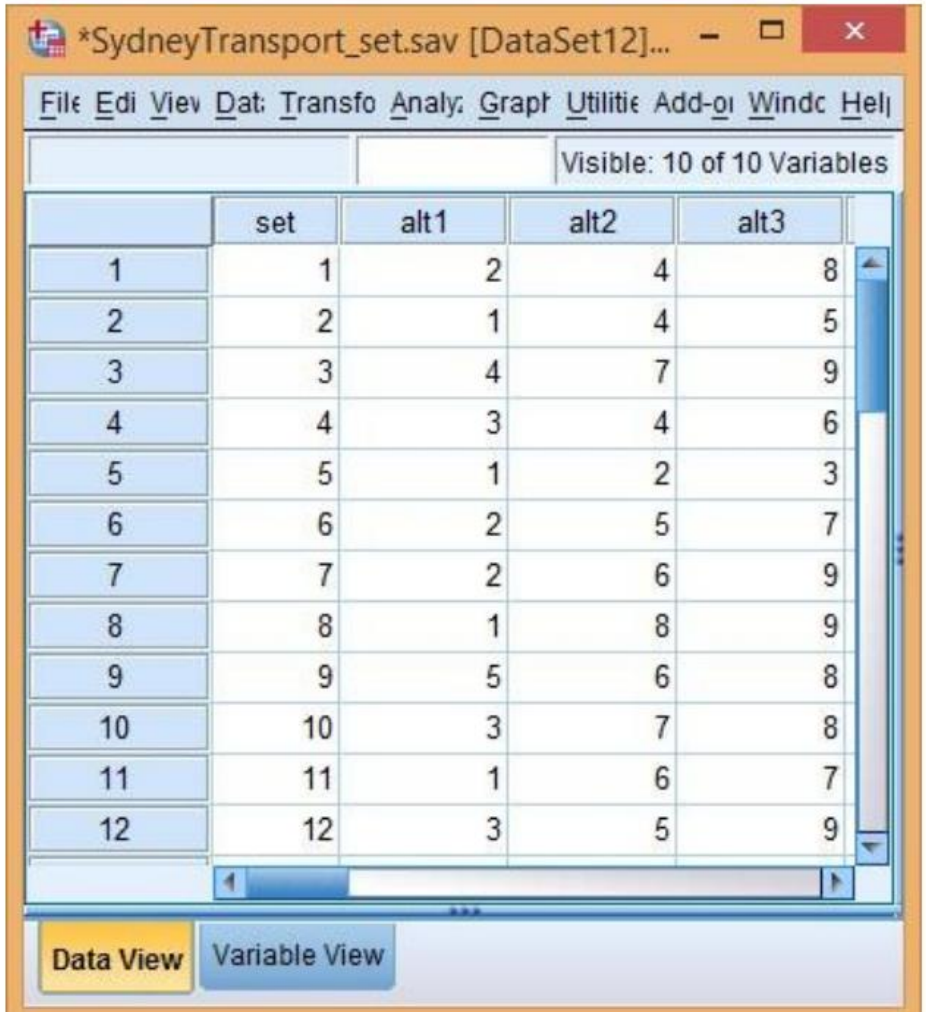
Option Number	Option	Highest Priority	Lowest Priority
2	Improved peak rail capacity		
4	Extensions of light rail services		
8	New cycleways; more bike and scooter parking		

Note: The current approach is applicable to options displayed 3 (or more) at a time, and to larger Max-Diff problems (e.g. 30+ options). Thus, can also be applied to large designs with several versions and 5+ alternatives per set.

MaxDiff Design Assesses Preferences for All Possible Option Pairs

12 Sets of triples:

Contains each of 36 option pairs <u>exactly</u> once (permits comparison with our proposal to use simulated data)



	set	alt1	alt2	alt3
1	1	2	4	8
2	2	1	4	5
3	3	4	7	9
4	4	3	4	6
5	5	1	2	3
6	6	2	5	7
7	7	2	6	9
8	8	1	8	9
9	9	5	6	8
10	10	3	7	8
11	11	1	6	7
12	12	3	5	9

Compared to $\binom{9}{2}$ = 36 total paired comparisons

Accurate Typing Tool Attainable with 14 Pairs? 12 Pairs? Less?

Binary items are much simpler to implement than MaxDiff in future surveys.

After developing segmentation with MaxDiff data, can one select a small number of binary items to use to assign a new case to most appropriate segment?

Item	A_OPTION	B_OPTION
1	6	9
2	7	8
3	4	9
4	3	6
5	7	9
6	1	8
7	2	6
8	5	9
9	4	7
10	3	9
11	1	6
12	6	8
13	4	5
14	3	7

Overview of Methodology

Step 1: Segmentation: Use LC to Segment MaxDiff data to obtain segments.

Step 2: Simulation: Use LC model parameter estimates to **simulate** population data consisting of segment membership (dependent variable) and responses to all possible pairs (predictors).

Step 3: Variable reduction to **obtain the reduced number of pairs** for inclusion in future surveys.

Step 4: Obtain scoring equation based on reduced pairs and implement in **typing tool to classify new cases into the most likely segment**.

Methodology

Implement MaxDiff: N=200

1. Segment the MaxDiff data using the LC Sequential Best-Worst Model and BIC to determine # segments. Obtain the expected classification error for the key segments of interest. (See Technical slides)
2. Use the model parameter estimates from Step 1 (including the class-specific utilities) to simulate N=8,000 'best' choices from all possible $\binom{9}{2}$ = 36 pairs, and use Bayes rule to get class labels for these cases.
3. Use a variable reduction method, such as stepwise multinomial logistic regression or a regularized regression approach, such as Correlated Component Regression to analyze the data obtained in Step 2. The dependent variable is the class label and the 36 binary responses are candidate predictors. Use class-specific classification errors as criteria for determining the number of items to include.
4. Use the selected pairs in future surveys and apply the prediction scoring equations obtained in Step 3 to get the class predictions for the new cases.

Methodology

Implement MaxDiff: N=200

1. **Segment the MaxDiff responses using the LC Sequential Best-Worst Model and BIC to determine # segments. Obtain the expected classification error for the key segments of interest. (See Technical slides)**
2. Use the model parameter estimates from Step 1 (including the class-specific utilities) to simulate N=8,000 'best' choices from all possible $\binom{9}{2}$ = 36 pairs, and use Bayes rule to get class labels for these cases.
3. Use a variable reduction method, such as stepwise multinomial logistic regression or a regularized regression approach, such as Correlated Component Regression to analyze the data obtained in Step 2. The dependent variable is the class label and the 36 binary responses are candidate predictors. Use class-specific classification errors as criteria for determining the number of items to include.
4. Use the selected pairs in future surveys and apply the prediction scoring equations obtained in Step 3 to get the class predictions for the new cases.

Segmentation Result: 8-Class Model

		Class1	Class2	Class3	Class4	Class5	Class6	Class7	Class8
	class size	0.2083	0.177	0.136	0.123	0.1183	0.1123	0.0849	0.0401
		Class-Specific Utility Parameters							
	Alternatives	Class1	Class2	Class3	Class4	Class5	Class6	Class7	Class8
1	More frequent off peak trains bet	1.2	1.0	0.4	-0.1	-0.3	-0.7	-0.1	-0.4
2	Improved peak rail capacity	1.3	3.1	0.8	3.6	0.7	2.2	1.1	1.3
3	More frequent bus services on ma	1.4	1.8	0.2	2.6	0.6	2.2	0.8	0.6
4	Extensions of light rail services	0.5	-1.9	-1.2	-1.8	-0.3	-1.4	-2.4	1.3
5	Integrated fares	-0.5	0.4	0.3	-1.0	-0.3	-0.3	2.0	2.0
6	Integrated ticketing	-0.7	0.4	0.1	-1.2	-0.4	-0.4	2.7	2.6
7	Real-time arrival information	-0.8	-0.1	0.4	-1.9	-2.0	-1.1	-1.3	0.4
8	New cycleways; more bike and sc	-1.3	-2.6	-1.6	-2.0	0.4	0.9	-1.9	-3.6
9	Trains use green power	-1.0	-2.2	0.6	1.7	1.6	-1.4	-0.8	-4.1

Expected Class-specific error = 10.3% 6.2% 10.3% 8.9% 7.3% 7.5% 5.0% 2.3%

Expected correct classification = 92% based on MaxDiff responses to 12 sets

These are the parameters used to simulate population data

Methodology

Implement MaxDiff: N=200

1. Segment the MaxDiff data using the LC Sequential Best-Worst Model and BIC to determine # segments. Obtain the expected classification error for the key segments of interest. (See Technical slides)
2. **Use the model parameter estimates from Step 1 (including the class-specific utilities) to simulate N=8,000 'best' choices from all possible $\binom{9}{2}$ = 36 pairs, and use Bayes rule to get class labels for these cases.**
3. Use a variable reduction method, such as stepwise multinomial logistic regression or a regularized regression approach, such as Correlated Component Regression to analyze the data obtained in Step 2. The dependent variable is the class label and the 36 binary responses are candidate predictors. Use class-specific classification errors as criteria for determining the number of items to include.
4. Use the selected pairs in future surveys and apply the prediction scoring equations obtained in Step 3 to get the class predictions for the new cases.

Step 2. Get Simulated Data (N=8,000)

ID	Choice.12	Choice.13	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	C h	Choice.89	Class#
1	2	2	1	1	1	1	1	1	2	1	1	2	1	1	1	1	1	2	2	1	1	2	2	1	2	2	2	1	2	1	1	1	1	1	1	2	6
2	2	1	1	1	2	1	1	2	2	1	2	2	1	1	2	1	2	2	1	1	1	2	2	1	2	2	1	1	1	1	1	1	1	2	2	2	4
3	2	2	2	1	2	1	2	2	2	2	1	1	1	1	2	1	1	2	2	2	2	2	2	1	2	2	1	1	2	2	1	2	2	2	2	2	2
4	1	2	1	1	1	1	2	2	1	1	2	1	2	1	1	1	2	2	1	1	1	2	2	1	2	2	2	2	1	2	2	1	2	1	2	2	5
5	2	2	1	2	1	1	1	2	2	1	1	1	1	1	1	1	2	1	1	2	1	1	2	2	1	2	2	1	1	2	2	1	2	2	2	1	8
6	2	2	1	2	1	1	2	2	2	1	1	1	1	2	1	1	2	1	1	2	2	1	2	1	1	2	1	1	2	2	1	2	2	2	2	2	2
7	2	2	2	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	2	2	1	2	2	1	1	2	1	2	2	2	2	2	8
8	2	1	1	2	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	1	1	2	2	1	1	1	1	1	1	1	2	6
9	2	2	1	2	2	1	1	2	2	1	2	2	1	1	1	1	2	2	2	1	1	2	2	2	2	2	2	1	1	2	1	1	1	1	1	1	4
10	2	2	1	1	2	1	1	1	1	1	2	1	1	1	1	1	2	2	1	1	1	2	2	2	1	2	1	1	1	1	1	1	1	1	1	2	4
11	2	2	2	2	2	1	1	1	1	1	2	2	1	1	1	1	2	1	1	1	1	2	2	2	2	2	2	1	1	2	1	1	1	2	1	2	4
12	1	1	1	1	1	1	1	1	1	1	2	1	2	1	1	1	1	1	1	2	1	2	1	1	1	1	2	2	1	2	1	2	1	2	1	1	1
13	2	2	1	2	2	1	1	1	2	1	2	2	1	1	1	1	2	2	1	1	1	2	2	2	1	2	2	1	1	1	1	1	1	2	2	2	4
14	2	2	1	1	2	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	1	2	1	2	2	1	2	1	2	2	8
15	2	1	2	2	2	2	1	2	2	1	1	1	2	1	2	1	2	1	1	1	2	2	2	2	1	2	2	1	2	1	2	1	2	1	2	1	5
16	2	1	2	2	2	2	1	1	1	1	1	1	1	1	2	2	2	2	2	1	2	2	2	1	2	2	1	1	1	1	2	2	2	1	1	2	5
17	2	2	2	1	2	1	2	2	1	2	1	2	1	2	2	1	1	2	1	2	2	2	2	1	1	2	1	1	2	2	1	2	2	2	2	1	2
18	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	2	1	1	2	1	2	2	1	2	2	1
19	2	2	1	1	2	2	1	1	2	1	2	2	2	1	2	1	1	1	2	1	1	2	2	2	1	2	1	1	1	2	2	1	1	1	2	2	5
20	2	1	1	2	2	1	1	2	1	1	2	2	1	1	1	1	2	2	1	1	2	2	2	2	2	2	2	1	1	2	1	1	1	1	2	2	4

Step 3. Analyze these data to get the best prediction pairs.

Reason for Simulation

- Each segment (population k) is defined by its associated MaxDiff model parameters. With these parameters, N[k]=1,000 cases were simulated using the MNL formula to generate responses to all 36 paired comparisons.
- The resulting simulated data (N=8,000 for the total 8 segments) were then analyzed to determine the subset of pairs that best predict segment membership and parameter estimates for these predictors to be used in a typing tool.
- **WHY SIMULATION?**
  - **The large sample size is needed to prevent the analysis from being overly affected by sampling fluctuation.**
  - **Allows us to work directly with TRUE segments, rather than MODAL assignments**
  - **Generates a consistent set of observations generalizable to projects where there is >1 version/rotation of design**

Methodology

Implement MaxDiff: N=200

1. Segment the MaxDiff data using the LC Sequential Best-Worst Model and BIC to determine # segments. Obtain the expected classification error for the key segments of interest. (See Technical slides)
2. Use the model parameter estimates from Step 1 (including the class-specific utilities) to simulate N=8,000 'best' choices from all possible $\binom{9}{2} = 36$ pairs, and use Bayes rule to get class labels for these cases.
3. **Use a variable reduction method, such as stepwise multinomial logistic regression or a regularized regression approach, such as Correlated Component Regression to analyze the data obtained in Step 2. The dependent variable is the class label and the 36 binary responses are candidate predictors. Use class-specific classification errors as criteria for determining the number of items to include.**
4. Use the selected pairs in future surveys and apply the prediction scoring equations obtained in Step 3 to get the class predictions for the new cases.

Generate Golden Questions Algorithm from Simulated data

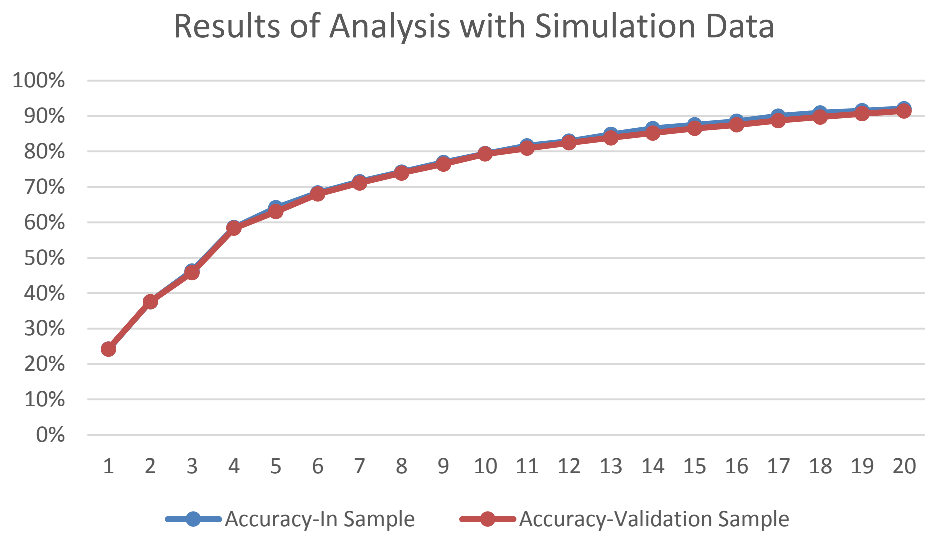
- **Obtain simulated data –**
  - Simulate each segment separately – 1,000 cases for each
  - Responses to all 36 binary pairs generated using segment parameters

	Alternatives	Class1
1	More frequent off peak trains bet	1.2
2	Improved peak rail capacity	1.3
3	More frequent bus services on ma	1.4
4	Extensions of light rail services	0.5
5	Integrated fares	-0.5
6	Integrated ticketing	-0.7
7	Real-time arrival information	-0.8
8	New cycleways; more bike and sc	-1.3
9	Trains use green power	-1.0

- **Combine data from all 8 simulations and predict segments using binary responses as predictors**
- **Use statistical modeling to reduce # predictors from 36 to 14 or less using:**
  - stepwise multinomial logistic regression (OK in this application),
  - discriminant analysis (not appropriate with dichotomous predictors),
  - or regularized regression (preferred)

Stepwise Regression Analysis of Simulated Data

Number of Pairs	Accuracy-In Sample	Accuracy-Validation Sample
1	24.2%	24.2%
2	37.6%	37.6%
3	46.3%	45.8%
4	58.5%	58.3%
5	64.1%	63.1%
6	68.3%	68.0%
7	71.5%	71.1%
8	74.2%	73.9%
9	76.9%	76.5%
10	79.4%	79.3%
11	81.6%	80.9%
12	82.9%	82.5%
13	84.8%	83.9%
14	86.5%	85.2%
15	87.5%	86.5%
16	88.5%	87.5%
17	90.0%	88.7%
18	90.9%	89.7%
19	91.5%	90.7%
20	92.1%	91.5%



- **Very little drop-off in validation data**
- **14 Pairs yields 85.2% overall accuracy**

Results from MNL Stepwise Regression 14 Selected Items

Pair	p-value
(6,9)	1.9E-198
(7,8)	4.4E-217
(4,9)	2.0E-280
(3,6)	2.3E-167
(7,9)	9.2E-204
(1,8)	6.9E-222
(2,6)	1.4E-183
(5,9)	4.0E-170
(4,7)	3.0E-158
(3,9)	7.5E-171
(1,6)	2.4E-156
(6,8)	2.7E-151
(4,5)	2.4E-142
(3,7)	9.3E-134

Notes:
- Pairs listed in order of model entry
- p-value assesses the effect of excluding predictor

14 Pairs In Sample – Accuracy

(Simulated Pairs)

Observed | Predicted

	1	2	3	4	5	6	7	8	TOTAL
1	823	54	29	31	17	34	5	7	1000
2	59	839	33	2	0	7	38	22	1000
3	30	45	816	41	14	6	46	2	1000
4	17	1	25	859	82	15	1	0	1000
5	9	0	12	94	854	31	0	0	1000
6	39	16	4	14	30	888	9	0	1000
7	3	25	26	1	1	9	898	37	1000
8	3	36	1	0	0	0	19	941	1000
TOTAL	983	1016	946	1042	998	990	1016	1009	86.5%

Observed | Predicted

Row %	1	2	3	4	5	6	7	8
1	82%	5%	3%	3%	2%	3%	1%	1%
2	6%	84%	3%	0%	0%	1%	4%	2%
3	3%	5%	82%	4%	1%	1%	5%	0%
4	2%	0%	3%	86%	8%	2%	0%	0%
5	1%	0%	1%	9%	85%	3%	0%	0%
6	4%	2%	0%	1%	3%	89%	1%	0%
7	0%	3%	3%	0%	0%	1%	90%	4%
8	0%	4%	0%	0%	0%	0%	2%	94%

14 Pairs Validation – Accuracy

(Simulated Pairs)

Observed / Predicted

	1	2	3	4	5	6	7	8	TOTAL
1	7165	542	234	388	183	426	19	43	9000
2	497	7411	313	43	0	94	428	214	9000
3	260	338	7260	509	175	90	336	32	9000
4	175	10	254	7680	766	114	1	0	9000
5	129	0	141	865	7566	292	7	0	9000
6	406	179	50	151	305	7846	56	7	9000
7	23	229	256	9	24	86	8057	316	9000
8	67	378	9	0	0	7	183	8356	9000
TOTAL	8722	9087	8517	9645	9019	8955	9087	8968	85.2%

Observed / Predicted

Row	1	2	3	4	5	6	7	8
1	80%	6%	3%	4%	2%	5%	0%	0%
2	6%	82%	3%	0%	0%	1%	5%	2%
3	3%	4%	81%	6%	2%	1%	4%	0%
4	2%	0%	3%	85%	9%	1%	0%	0%
5	1%	0%	2%	10%	84%	3%	0%	0%
6	5%	2%	1%	2%	3%	87%	1%	0%
7	0%	3%	3%	0%	0%	1%	90%	4%
8	1%	4%	0%	0%	0%	0%	2%	93%

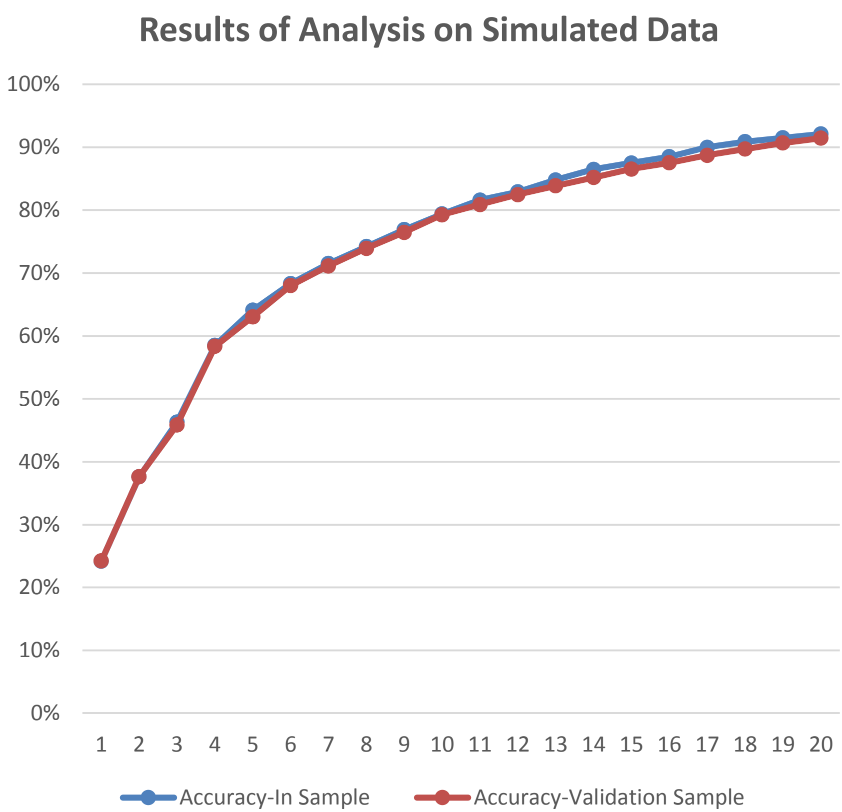
Alternative Results from MNL Stepwise Model Built on Original Sample <u>No Simulations</u>

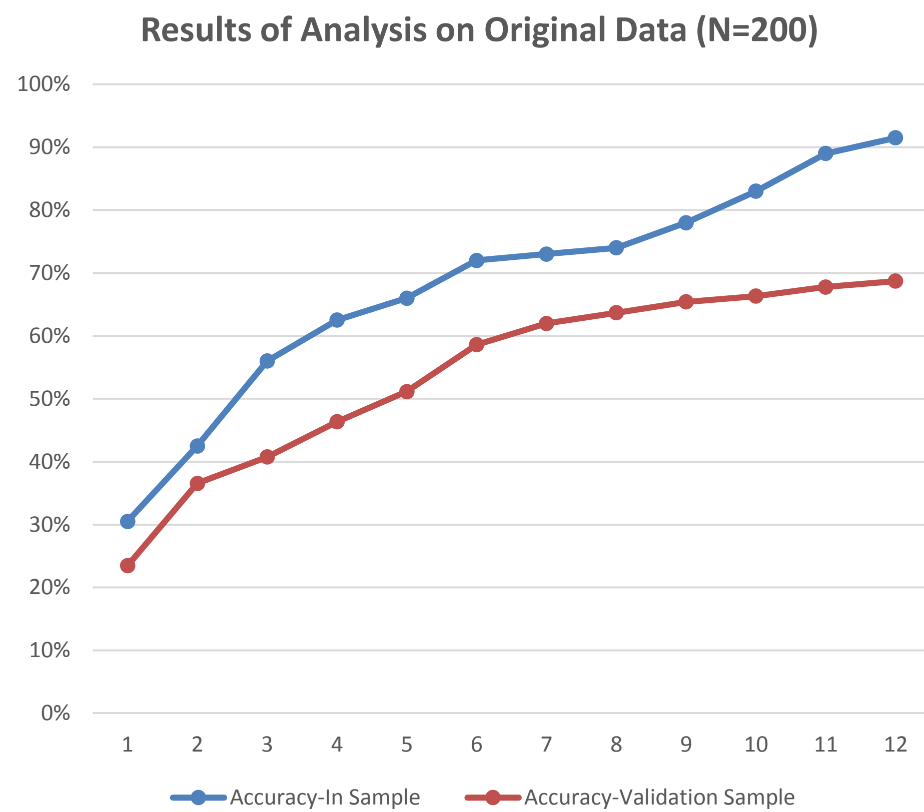
Number of Pairs	Accuracy-In Sample	Accuracy-Validation Sample
1	30.5%	23.4%
2	42.5%	36.5%
3	56.0%	40.7%
4	62.5%	46.3%
5	66.0%	51.1%
6	72.0%	58.6%
7	73.0%	62.0%
8	74.0%	63.7%
9	78.0%	65.4%
10	83.0%	66.3%
11	89.0%	67.8%
12	91.5%	68.7%

* Note 12th step added one predictor but removed another

- Experimental design for this project assessed each of the 36 pairs exactly once & was a one version design
- Allows test of MNL stepwise selection ON N=200 OBSERVATIONS DIRECTLY (DV = Modal Segment)
- RESULT: Best Model OVERFITS (91.5% in-sample drops off to 69%) and PERFORMS MUCH WORSE than model built on simulated data: Validation Accuracy 69% vs 85%.

Simulation Approach Works Much Better

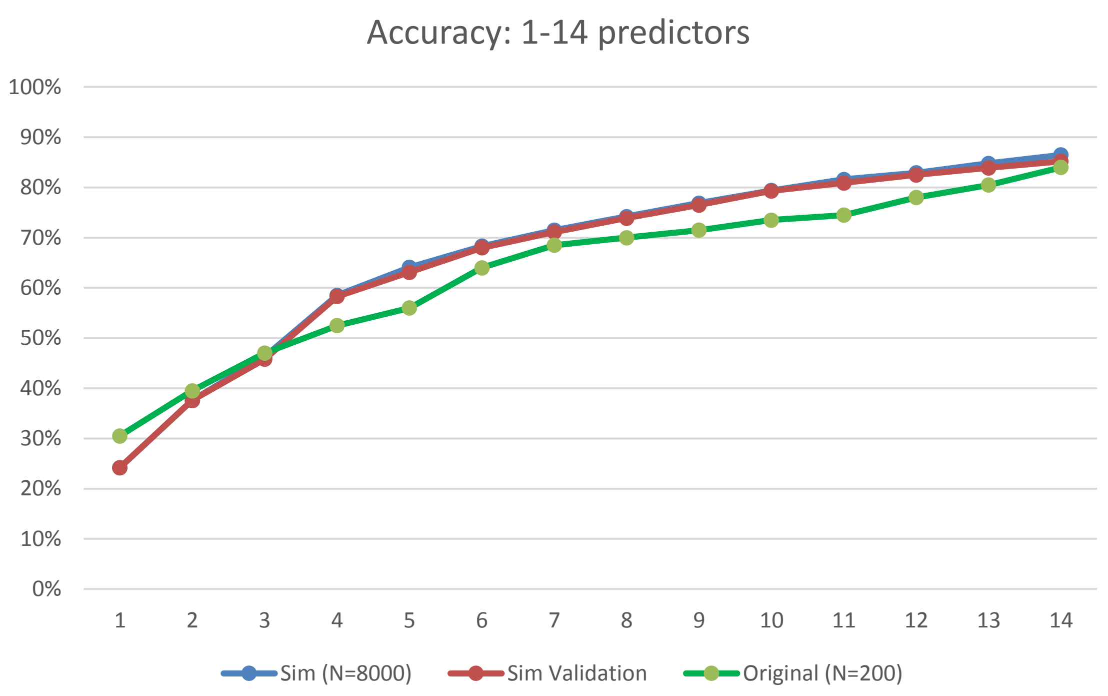




Responses to binary items (pairs) predict segments much better when simulated (N=8,000) than when obtained from observed MaxDiff responses in original sample (N=200)

14-item and other Simulated Data Models also Predict Well on Original Data

Number of Pairs	Accuracy-In Sample	Accuracy-Validation Sample	Accuracy-Original Sample
1	24.2%	24.2%	30.5%
2	37.6%	37.6%	39.5%
3	46.3%	45.8%	47.0%
4	58.5%	58.3%	52.5%
5	64.1%	63.1%	56.0%
6	68.3%	68.0%	64.0%
7	71.5%	71.1%	68.5%
8	74.2%	73.9%	70.0%
9	76.9%	76.5%	71.5%
10	79.4%	79.3%	73.5%
11	81.6%	80.9%	74.5%
12	82.9%	82.5%	78.0%
13	84.8%	83.9%	80.5%
14	86.5%	85.2%	84.0%



Simulated data models also validate when applied to small original sample (N=200)

Regularized Regression

Regularized regression is especially important to use when the sample size is relatively small and the number of correlated predictors is large, with possible high correlations. In this example, sample size of simulated data is large and number of predictors (pairs) is relatively small (36) so stepwise MNL appears to be OK – relatively small drop-off in validation sample.

Regularized regression (e.g., lasso, ridge regression, CCR) is also important to use with BIG DATA (large sample size and thousands of predictors).

Example of Pair not Selected in Model

Pairs may not contribute because they do not discriminate between the classes or because they are 'unbalanced'.

Pair (1,2) is an example of the latter where option 2 is favored over option 1 for each class:

Class-specific Utilities (Parameters)

	Alternatives	class 1	2	3	4	5	6	7	8
1	More frequent off peak trains between major centres	1.2	1.0	0.4	-0.1	-0.3	-0.7	-0.1	-0.4
2	Improved peak rail capacity	1.3	3.1	0.8	3.6	0.7	2.2	1.1	1.3

Restructured Parameters

Pair 12	1	2	3	4	5	6	7	8	All
Choice 1	0.48	0.11	0.40	0.02	0.27	0.05	0.22	0.16	0.24
2	0.52	0.89	0.60	0.98	0.73	0.95	0.78	0.84	0.76

Methodology

Implement MaxDiff: N=200

1. Segment the MaxDiff data using the LC Sequential Best-Worst Model and BIC to determine # segments. Obtain the expected classification error for the key segments of interest. (See Technical slides)
2. Use the model parameter estimates from Step 1 (including the class-specific utilities) to simulate N=8,000 'best' choices from all possible $\binom{9}{2}$ = 36 pairs, and use Bayes rule to get class labels for these cases.
3. Use a variable reduction method, such as stepwise multinomial logistic regression or a regularized regression approach, such as Correlated Component Regression to analyze the data obtained in Step 2. The dependent variable is the class label and the 36 binary responses are candidate predictors. Use class-specific classification errors as criteria for determining the number of items to include.
4. **Use the selected pairs in future surveys and apply the prediction scoring equations obtained in Step 3 to get the class predictions for the new cases.**

Future Survey Implementation Options

Options	Correct Classification	Fatigue
All 36 binary pairs	<96%	High
20 selected pairs	92%	Medium
Only 14 selected pairs	85%	Low
For comparison		
MaxDiff with 12 sets of triples	<92%	Very High
Best response only w original 12 sets	83%	

Benefits of this Method

Provides a typing tool which matches the researchers needs:

- Yields a set of data, generated consistent with the way the segments were obtained that can be used to select the most important predictors.
- Since all candidate predictors measure attribute preference, they are important based on theory.
- Scalable to large MaxDiff problems and complex designs – not dependent on having observed trade-offs for every possible pair for every individual
- Selection of pairs possible using readily available software (e.g. Stepwise Multinomial Logistic Regression and regularized regression)
- Delivers good trade-off between simplicity (few pairs) and accuracy

Easily Implemented in an Excel Typing Tool or programmable on a survey script

Possible Criticisms

- **Simulations are under the assumption the Latent Class MNL population model is correct:**

  Typing will always be conditional on a model. If desired, this LC model can be extended to allow *different scale factors* for 'best' and 'worst', and/or for different respondents*).

- **Stepwise MNL selection for the "pairs" is not ideal:**

  In this example with only 36 pairs stepwise MNL worked fine (very little drop-off in validation). Regularized regression may be better when dealing with very large designs (hundreds or thousands of pairs).

*Note: Scale-adjusted LC (SALC) MaxDiff models fit these data somewhat better than standard LC MaxDiff models – see Magidson and Vermunt (2015) and Magidson (2016).

Approach Dependent upon Latent Class Modeling?

- The approach relies on the ability to reproduce the segmentation on new cases via simulation.
- With LC, this can be done using the model parameters.
- With 2-step (tandem) approaches such as HB to get individual utilities followed by clustering, simulation is difficult or impossible.

Magidson, Dumont and Vermunt (ART Forum 2015) suggest that the widely used HB/cluster tandem approach is similar to:

1) assuming segments don't exist (MVN assumption),
2) then trying to find them (clustering), and
3) possibly being surprised when they are found to be difficult to interpret or reproduce.

References

Louviere (1993) “The Best-Worst or Maximum Difference Measurement Model: Applications to Behavioral Research in Marketing,” The American Marketing Association's 1993 Behavioral Research Conference, Phoenix.

Magidson, Dumont and Vermunt, 2015 A new modeling tool for identifying meaningful segments and their willingness to pay: Improving Validity by Reducing the Confound between Scale and Preference Heterogeneity, 2015 ART Forum.

Magidson, Jay and J.K. Vermunt (2016 forthcoming September), breakout session at 2016 Sawtooth Software Conference.

Magidson (2013) Correlated Component Regression: Re-thinking Regression in the Presence of Near Collinearity. Chapter in New perspectives in Partial Least Squares and Related Methods, Springer Verlag.

Magidson and Vermunt (2015) Latent GOLD Choice tutorial 8A: Analyzing MaxDiff Data with Scale Factors http://www.statisticalinnovations.com/wp-content/uploads/LGChoice_tutorial_8A.pdf

Marley, A.A.J. and J.J. Louviere (2005). Some probabilistic models of best, worst, and best– worst choices. Journal of Mathematical Psychology 49, pp. 464-480.

Vermunt and Magidson (2015) LG-Syntax User’s Guide: Manual for Latent GOLD 5 Syntax Module, http://www.statisticalinnovations.com/wp-content/upl

Questions & Discussion

Technical Appendix

- Simulation: How was simulation performed?
- Variable selection: Which regression model should be used to select the reduced set of binary choices?
- Segmentation: Which LC MaxDiff model should be used?

Simulation Performed with Latent GOLD Syntax*

	Alternatives	Class1
1	More frequent off peak trains bet	1.2
2	Improved peak rail capacity	1.3
3	More frequent bus services on ma	1.4
4	Extensions of light rail services	0.5
5	Integrated fares	-0.5
6	Integrated ticketing	-0.7
7	Real-time arrival information	-0.8
8	New cycleways; more bike and sc	-1.3
9	Trains use green power	-1.0

* Example to simulate cases from Segment 1

- outfile 'sim10000.sav' **simulation=**10000; ← '**simulation**' is a keyword in Latent GOLD syntax
- choice = 3 ← 3-file format (responses, alternatives, and choice sets)
- alternatives 'SydneyTransport_alt.sav'
- id=AltID
- choicesets 'SydneyTransport_set_inac0.sav'
- id=set;
- variables
- caseid ID;
- choicesetid Set ;
- dependent choice CHOICE; ← simulating first (Best) choices only
- attribute Alt nominal;
- latent
- Class nominal 1; ← each segment simulated separately (these can be omitted)
- equations
- choice <- Alt ; ← MNL model specification
- {
- 1.212271
- 1.307404
- 1.434362 ← (9-1 = 8 parameters for each segment – here, segment 1)
- …}

Variable selection: Which regression model should be used to select best binary choices?

- Multinomial Logistic Regression (stepwise approach used here, is not optimal with very large number of correlated predictors)
- Stepwise discriminant analysis is not appropriate here because predictors are dichotomous – not continuous
- A regularized multinomial logit regression approach has theoretical advantages with high dimensional or big data. We are currently working on the extension of Correlated Component Regression (Magidson, 2013) to work with the multinomial case.

Sequential Best-Worst Model Recommended

Model	Consistent Margins	BIC suggest # classes	Parameters Unbiased
MaxDiff	NO	YES	YES
MaxDiff-Ind	YES	NO -- too many	NO
Best-Worst	**YES**	**YES**	**YES**

Notes: There are 3 primary MaxDiff (Best-Worst) modeling approaches:

- We used the Sequential logit method ("Best-Worst" above) as implemented in the Latent GOLD program.
- The "MaxDiff" model is described by Marley and Louviere (2005), and
- The "MaxDiff-Ind" model was proposed earlier by Louviere (1993) and implemented by Sawtooth Software.

- Our "Best-Worst" model with 8 classes contains 71 parameters.
- Additional parameters can be included in the models:
  - Scale parameters – account for additional within-class variability
- Different scale factors for Best and Worst choices (generalizes the +1, -1 coding) appears to almost always be supported with real data.