

An Empirical Comparison of Different Approaches for Choice Studies with a Large Number of Attributes

Jon Pinnell

Lisa Fridley

Meagan Ulrich

MarketVision Research

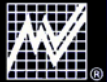
Cincinnati, Ohio

Presented at:

**Conjoint Design and
Innovations Conference**

May 2005

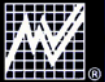
Berlin



Background

Some have expressed concern over simplifying behavior

- Two schools of thought have emerged related to choice exercises
 - Full Profile tasks with many attributes reflect reality
 - Simplification only happens in research but not purchasing



What is a large number of attributes?

- Some citations indicate the upper limit is 10
- Others say 6
- Some older references in transportation literature say 3 is the max

Common supporting citations include: Green (1984 among others)
Wright (1975)
Miller (1956)

Background

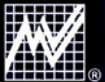
One Common Response: Partial Profile

Attributes	Full Profile Tasks	Partial Profile Task 1	Partial Profile Task 2	Partial Profile Task 3	Partial Profile Task 4
1	X	X			X
2	X		X		X
3	X	X		X	
4	X	X	X		
5	X			X	X
6	X			X	
7	X	X			
8	X		X		X
9	X		X	X	

Partial Profile tends to produce:

Utilities with larger scale (steeper)

Attribute Importances that are more similar (flatter)



Background

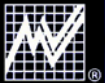
Other Recent Results Regarding Partial Profile Utilities

Pinnell (2000) showed hit rates using utilities from HB versus aggregate logit. In five of the six studies the results were positive for HB, and strongly so. The anomalous sixth study was partial profile.

Follow-up research (Pinnell and Fridley, 2001), which focused exclusively on partial profile designs, showed mixed results. Of the nine studies, HB showed a significant improvement in hit rate in three, hit rate parity in two, and a statistically significant degradation in four.

Frazier and Jones (2004) show that partial profile tasks produced MAE in share predictions nearly 50% larger than the standard full profile with a none option (when averaged across the three reported studies).

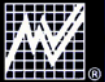
Johnson, Huber and Orme (2004) show hit rates for partial profile tasks to be 8 points lower than full profile tasks (71% vs. 63%), and MAEs in share predictions were more than 50% larger for partial profile than for full profile (7.1 vs. 4.6). The authors also showed that derived attribute importances from partial profile are flatter (more nearly equal) than those from full profile.



Experimental Design

Study Inputs – Hotel Research

- 9 attributes
- 3 alternatives per task
- 20 estimation tasks
- 6 holdout tasks (half were FP and half were PP)



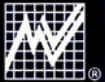
3 cells (150 business travelers per cell)

- Full Profile – Fixed Order (brand first, price last)
- Full Profile – Random Order
- Partial Profile

Findings

Statistical Efficiency and Respondent Cost

	Relative D-efficiency	Median Time (seconds / index)
FP – FO	100	503 / 100
FP – RO	100	551 / 109
PP	68	485 / 96
Index	FP – FO = 100	FP – FO = 100



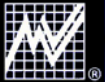
Findings

Hit Rate

	FP – FO	FP – RO	PP
Predicting FP Holdouts	69.0	60.0	59.5
Predicting PP Holdouts	66.9	68.2	69.1

MAE

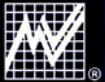
	FP – FO	FP – RO	PP
Predicting FP Holdouts	0.075	0.102	0.103
Predicting PP Holdouts	0.053	0.073	0.073



Findings

Attribute Importance

	FP – FO	FP – RO	PP
Brand	15.9	16.1	16.1
Price	19.2	18.3	17.5
Maximum	19.3	21.4	18.6
Minimum	3.7	3.2	3.9
Next Lowest	5.4	5.0	6.9
Avg. of Each Individual's Min Imp.	1.9	1.5	2.0

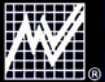


Findings

Estimate 9 logit models for each person (one for each attribute) and count the number of those models that are significant based on $-2LL$ ($\alpha = 0.05$)

Proportion of Respondents

Count	FP FO	FP RO	PP
0	0	0	0
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	1	1	0
6	3	4	2
7	18	19	16
8	45	43	29
9	33	33	53



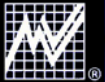
Findings

Let's look at another study...

(1220 respondents, 7 Full Profile tasks, 14 attributes)

Proportion of Respondents

Count	Proportion
0	0
1	0
2	0
3	0
4	0
5	0
6	0
7	0
8	0
9	0
10	0
11	0
12	1
13	15
14	83



Experimental Design

But still, we should consider respondent burden
 Let's look at an alternative approach: Bridging

Two additional research cells

Standard Full Profile

Brand
X
X
X
X
X
X
X
X
Price

Neutral Bridge

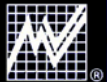
Brand
X
X
X
X

X
X
X
Price

Brand & Price Bridge

Brand	Brand
X	
X	
X	
X	
	X
	X
	X
Price	Price

Estimate two models for each cell and combine the results



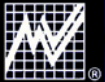
Findings

MAE

Predicting Full Profile Holdouts

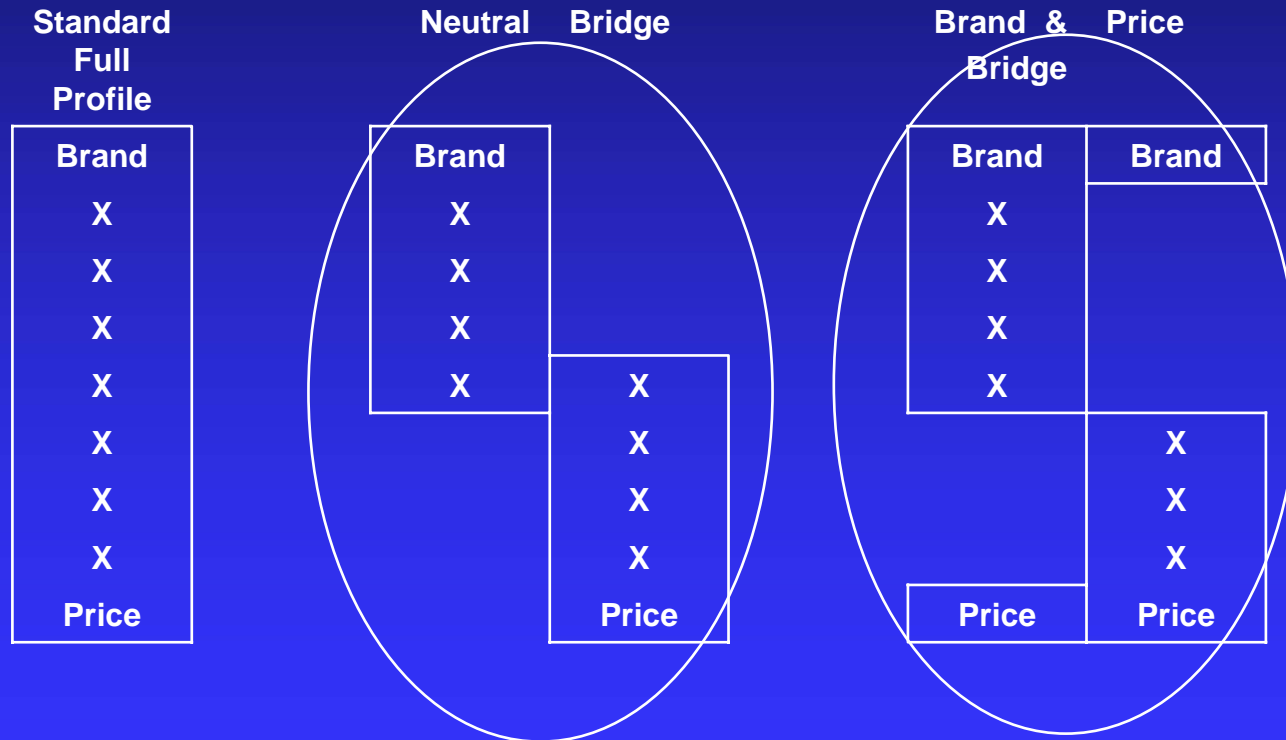
	No Scaling	Individual Level Scaling	Aggregate Level Scaling
Brand and Price	0.105	0.120	0.104
Neutral	0.070	0.069	0.069

(FP-FO = 0.075)



Experimental Design

Yet another approach to estimating utilities from these same tasks: Blocking



Only estimate one model per cell

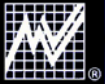
Findings

MAE

Predicting Full Profile Holdouts

	No Scaling	Individual Level Scaling	Aggregate Level Scaling	Blocking No Bridge
Brand and Price	0.105	0.120	0.104	0.087
Neutral	0.070	0.069	0.069	0.057

(FP-FO = 0.075)



Findings

Attribute Importance

	FP-FO	FP-RO	PP	PP-NB	PP-BP
Brand	15.9	16.1	16.1	14.7	14.7
Room Service	8.9	7.2	9.7	10.2	9.1
Fitness Center	7.9	7.5	6.1	6.9	8.7
Airport Shuttle	5.4	5.0	6.9	7.4	8.6
Breakfast Buffet	11.7	10.0	11.8	11.6	10.5
Internet Access	19.3	21.4	18.6	18.8	16.4
Mini Bar	3.7	3.2	3.9	3.0	4.4
Location	8.0	11.2	9.3	8.8	8.1
Price	19.2	18.3	17.5	18.4	19.4



Findings

Attribute Importance

	FP-FO	FP-RO	PP	PP-NB	PP-PB
Max	19.3	21.4	18.6	18.8	19.4
Min	3.7	3.2	3.9	3.0	4.4
Avg. of Each Individual's Minimum Importance	1.9	1.5	2.0	1.8	2.1



Summary

Respondents can and do process many attributes

The traditional full profile exercise doesn't seem to promote simplification

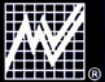
Randomly ordering attributes might appear to create frustration which can lead to simplification

Even if a respondent ignores an attribute it doesn't necessarily indicate a problem with the choice design or simplification of a complex task

It might just truly be unimportant

Simplification is only a problem if the simplification in choice exercises is different than simplification in reality

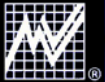
Always letting price be the last attribute impacts the importance of price



Summary

Traditional PP Designs have proved problematic

- Lower statistical efficiency
- Poorer predictive validity
- Overfitting with HB



Recommendations

If you are concerned about respondent burden...

Employ a blocked partial profile

This test included only two, but that is not a limit, for example, this has four blocks

