Intro to Adaptive Choice-Based Conjoint

Agenda

- Motivations for Adaptive CBC
- The ACBC Experiment
- DEMO
- Comparing Methods

Section 1

MOTIVATIONS FOR ACBC

Choice is Dominant

- Especially over the last decade, academics and practitioners have favored choice over ratings-based methods:
  - Stronger mathematical theory
  - Stronger psychological underpinnings
  - Argued to be more accurate
- But, DCM/CBC has drawbacks, especially for small sample sizes and studying many attributes.
  - Can we do even better?

Motivation for Adaptive CBC

- Capture more information from respondent (than standard CBC), encouraging deeper rather than superficial information processing, and requiring smaller sample sizes.
- Provide more engaging interviews than standard CBC (hopefully leading to better data).
- Greater ability than standard CBC to study many attributes and levels.
- More accurate predictions than standard CBC, better segmentations.

Example Choice Task:

- If you were a respondent, how would you complete such a survey with 12+ tasks?

Assuming these were the only available options, please tell me which of the following homes you would likely purchase?

(3 of 18)

1400 Square Feet	1750 Square Feet	3000 Square Feet	4000 Square Feet	
Built green certification & superior efficiency	Standard construction & efficiency	Built green certification & superior efficiency	Standard construction & efficiency	
2 bedrooms, 1.5 bath	2 bedrooms, 1.5 bath	4 bedrooms, 3.5 baths	3 bedrooms, 2.5 baths	
Medium master bath (separate shower and tub), one walk-in closet	Small master bath (integrated shower/tub), small walk-in closet	Small master bath (integrated shower/tub), small walk-in closet	Large master bath (separate shower & deep soaking tub), two walk-in closets	
2-car garage	3-car garage	3-car garage	2-car garage	None of these homes comes very close to fitting my needs and budget.
Large lot (0.5 acres)	Small lot (0.2 acres)	Small lot (0.2 acres)	Medium lot (0.35 acres)	
Local schools ranked in bottom third of nation	Local schools ranked in top third of nation	Local schools ranked in top third of nation	Local schools ranked in middle third of nation	
Upgraded finish package (floors, countertops, millwork, fixtures)	Standard finish package (floors, countertops, millwork, fixtures)	Standard finish package (floors, countertops, millwork, fixtures)	Premium finish package (floors, countertops, millwork, fixtures)	
Home built in 2000	New home	Home built in 1990	Home built in 1980	
$236,250	$398,500	$379,750	$715,000	
○	○	○	○	○

Strong Evidence for Non-Compensatory Processing

- Most respondents answer CBC tasks applying some combination of non-compensatory behaviors: must-haves, unacceptables, elimination-by-aspects, conjunctive rules ("must be acceptable on this AND this"), etc.

- Hauser et al. have a comprehensive article on this that is well worth the read: "Non-Compensatory (and Compensatory) Models of Consideration-Set Decisions" 2009 Sawtooth Software Conference Proceedings, pp. 207-232.

Logit Rule & Efficiency

- Most all CBC research on efficient experimental designs for DCM/CBC has assumed respondents apply the logit rule (compensatory processing).
  - Level balance, minimal overlap, utility balance, d-efficiency
- But, if respondents typically don't use compensatory heuristics, then are such traditional designs really optimal?

Imagine that you required at least a 3 GHz processor (you do a lot of HB)…

Which could you choose?

Dell	Lenovo	HP
4 GB RAM	1 GB RAM	2 GB RAM
80 GB Hard Drive	160 GB Hard Drive	100 GB Hard Drive
2 GHz Processor	2.5 GHz Processor	3 GHz Processor
21-Inch Monitor	19-Inch Monitor	17-Inch Monitor
$850	$750	$1,000

If you truly require 3 GHz processor (lesser speed is "unacceptable")…

And we inform utility estimation of that fact…Wouldn't this be a more efficient question?

Dell	Lenovo	HP
2 GB RAM	1 GB RAM	4 GB RAM
80 GB Hard Drive	100 GB Hard Drive	160 GB Hard Drive
3 GHz Processor	3 GHz Processor	3 GHz Processor
21-Inch Monitor	19-Inch Monitor	17-Inch Monitor
$750	$850	$1,000

Now, if we also learn you're absolutely loyal to Dell (Dell is a "must-have")…

And we inform utility estimation of that fact…Wouldn't this be a more efficient question?

Dell	Dell	Dell
1 GB RAM	4 GB RAM	2 GB RAM
160 GB Hard Drive	80 GB Hard Drive	160 GB Hard Drive
3 GHz Processor	3 GHz Processor	3 GHz Processor
17-Inch Monitor	19-Inch Monitor	21-Inch Monitor
$750	$1000	$850

New Approach

- Following Hauser *et al.* presentation on modeling non-compensatory effects in conjoint (Sawtooth Software conference 2006), we considered that we might be "barking up the wrong tree" with our previous adaptive CBC attempts.
- Other academics were also showing that respondents often used non-compensatory processing of conjoint tasks (e.g. Gilbride, Allenby).
  - If respondents regularly did not adhere to logit-rule assumptions, attempts to increase d-efficiency via adaptive methods might be futile.
- We decided what was needed was a new kind of Adaptive CBC interview that capitalized on non-compensatory heuristics and produced an overall more relevant and information-filled questionnaire.

Tests Show Promise

- Three comparison tests vs. CBC (prior to releasing ACBC v1) gave us a lot of hope that ACBC may be a good improvement.
  - Respondents preferred it to CBC
    - More realistic (3 of 3 tests)
    - Less monotonous (2 of 3 tests)
  - It produced more accurate hit rates than CBC
  - It produced better share predictions than CBC
- Main drawback was time for interview is longer.

Subsequent Findings

- After v1's release, additional research has shown that ACBC performs well:
  - Chapman (Microsoft) found ACBC to predict actual purchases *slightly* better than standard CBC. Information for segmentation and optimization routines was superior from ACBC. (2009 Sawtooth Software Conference)
  - Goodwin (Lifetime Products) found ACBC to work well for a 16-attribute study with 45 total levels. (2009 Sawtooth Software Conference)
  - Binner (bms) found that even non-technical trades people could do in-person ACBC interviews on 14 attributes. They found the interview engaging and enjoyable. (SKIM/Sawtooth Software Conference 2009)
  - Neggers and Hoogerbrugge compared ACBC to CBC for a mobile telephony study on 11 attributes, finding ACBC to predict holdout shares more accurately than CBC, with much lower sample size. (SKIM/Sawtooth Software Conference 2009)

Section 2

THE ACBC EXPERIMENT

ACBC Questionnaire Flow

- Build-Your-Own (BYO) Section
- Screening Section
- Choice Tasks Tournament
- Calibration Concepts

BYO: Drop-Down

Please assume you were in the market to buy a home. If these were the available options, what home would you be most likely to buy?

As you select different options, the total price of the home will change. When you have finished choosing the home that suits you best, click the Next button to continue.

Feature	Select Option	Option Price
Size of Home	3000 Square Feet $506,250	$ 506,250
Eco-Friendly Features	Standard construction & efficiency	$ 0
Bedroom/Bathroom Configuration	4 bedrooms, 3.5 baths (+ $12,500)	$ 12,500
Master Bath	Medium master bath (separate shower and tub), one walk-in closet	$ 0
Garage	3-car garage (+ $12,500)	$ 12,500
Lot	Select Option	$
Schools		$
Finish		$
Age of Home	Select Option	$
	Total:	$ 531,250

Select Option
Small lot (0.2 acres) (subtract $25,000)
Medium lot (0.35 acres)
Large lot (0.5 acres) (+ $25,000)

BYO: Radio-Button Grid

Bitte wählen Sie diejenigen Eigenschaften aus, die Sie beim Neukauf eines Fahrzeugs, am wahrscheinlichsten wählen würden. Bitte entscheiden Sie für jede der Eigenschaften ob Sie sie gerne hätten oder nicht.

Eigenschaften:	Bitte wählen Sie eine der Eigenschaften:	Aufpreis für diese Eigenschaft:
Design und Typenbezeichnung:	○ Reguläres Karosserie- und Interieurdesign mit Serientypenbezeichnung (+ €23.000) ○ Reguläres Karosserie- und Interieurdesign mit Öko-Typenbezeichnung (+ €23.000) ○ Besonderes, auf die Umweltfreundlichkeit des Fahrzeugs hinweisendes Karosserie und Interieur-Design (+ €23.000)	€
Anzahl Türen:	○ 2 ○ 4 (+ €150) ○ 5 (+ €250)	€
Karosserieform:	○ Stufenheck ○ Fliesheck ○ Kombi (+ €1.500)	€
Kraftstoffsorte:	○ Benzin ○ Diesel (+ €1.000) ○ Hybrid (+ €1.500)	€
Hubraum:	○ 1,1 ltr. ○ 1,2 ltr. (+ €200) ○ 1,3 ltr. (+ €400)	€

Benefits of BYO

- Is quick and easy for respondents
- Allows respondents to tell us exactly what they want in a product, subject to their budget threshold
- May be analyzed with MNL (together with other data)
- Contains less random error than CBC tasks
- Provides excellent training task to educate respondents about attributes and levels

(Source: Johnson, Orme, Pinnell 2006)

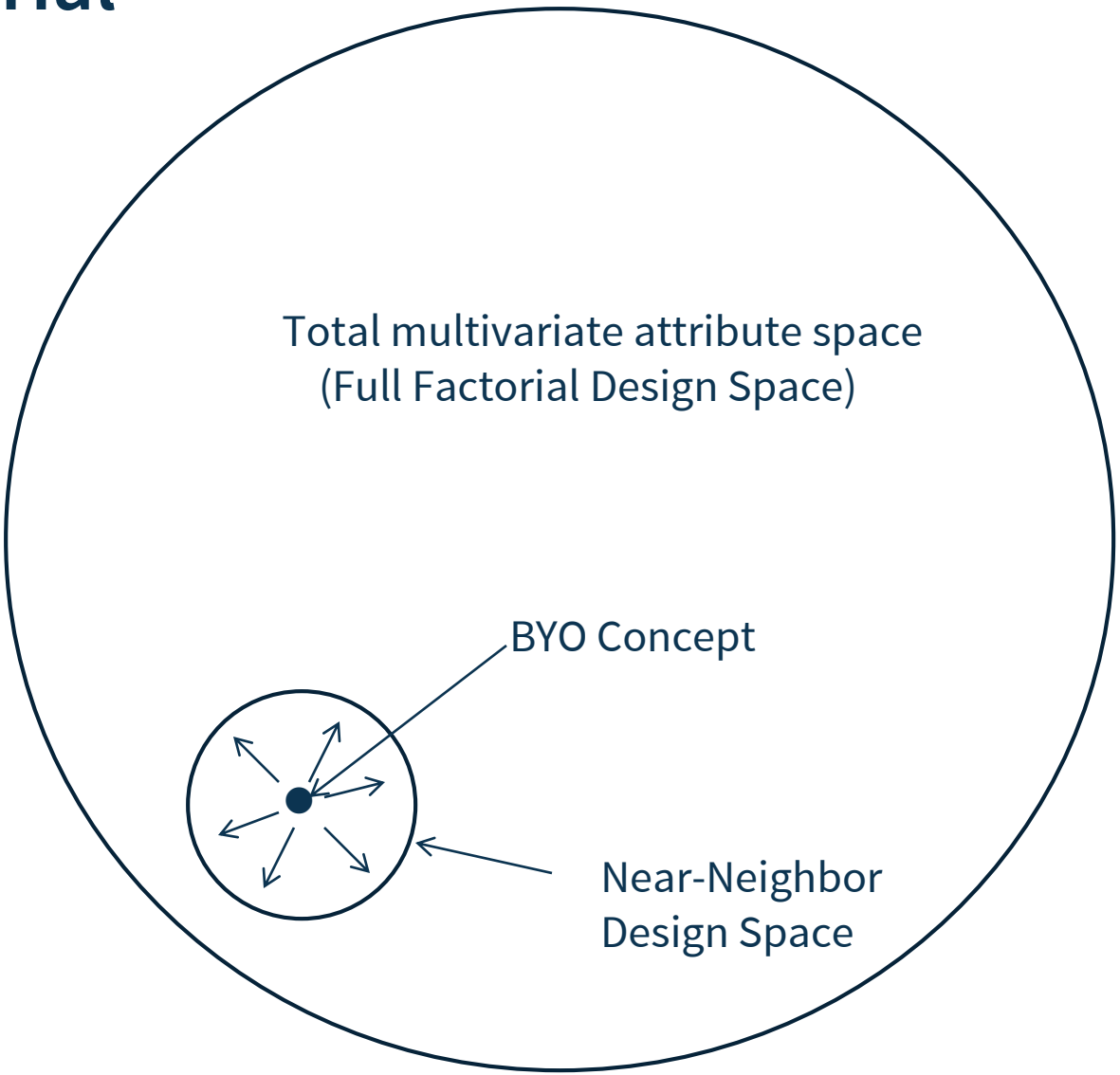
Also, clients typically like BYO tasks!

BYO in the ACBC Scheme

- It informs utility estimation regarding the most preferred level within each attribute.

- It tells us about the most relevant product for this respondent (establishes the centroid for our design space).
  - Is used to create maximally relevant concepts *for later sections of the interview*.

Design Strategy: Near-Neighbors instead of Full Factorial



Generating a Near-Neighbor Concept

	BYO Concept	Near-Neighbor Concept #1
	Made in USA	Made in USA
	4-door Sedan	4-door Sedan
Change this attribute →	Red exterior	Black exterior
	Black interior	Black interior
Change this attribute →	All-Wheel Drive	Front-Wheel Drive
	35 MPG	35 MPG
	0-60 in 9 seconds	0-60 in 9 seconds
	Premium trim package	Premium trim package
Change this attribute →	3-Year Extended Warranty	No Extended Warranty

In this example, we have randomly selected 3 of the 9 attributes to change from the BYO concept. For each of those 3 attributes, we randomly pick a new level. This process is repeated to create typically 24 to 36 near-neighbor concepts for the respondent to evaluate.

Some Recommendations

- The number of attributes to vary from the BYO selections depends on the total number of attributes (not including "summed price") in the study:

4 attributes	6 attributes	10 attributes
(vary 1 -2)	(vary 2-3)	(vary 2-4)

Note: need more empirical research to see if these guidelines hold!

Orthogonality? D-Efficiency?

- ACBC's designs are near-orthogonal (especially with no prohibitions).
- If you were to compare them to regular CBC designs using the standard ideas about design efficiency, they would appear to be inferior:
  - e.g. Lower D-Efficiency
- But, we would argue that standard efficiency measures make inappropriate assumptions about how respondents answer CBC interviews (logit rule), and don't reveal the true value of a design in terms of reducing uncertainty about the part-worths for attribute levels, given human information processing tendencies, limitations, and heuristics.

Screening Section

Here are a few homes you might like. Do any of these look like they are possibilities? It's helpful if you can keep about half of them for further consideration. But, it's up to you.

(1 of 6)

1750 Square Feet	3000 Square Feet	3000 Square Feet	4000 Square Feet
Standard construction & efficiency	Standard construction & efficiency	Built green certification & superior efficiency	Standard construction & efficiency
4 bedrooms, 3.5 baths	4 bedrooms, 3.5 baths	4 bedrooms, 3.5 baths	3 bedrooms, 2.5 baths
Large master bath (separate shower & deep soaking tub), two walk-in closets	Medium master bath (separate shower and tub), one walk-in closet	Medium master bath (separate shower and tub), one walk-in closet	Medium master bath (separate shower and tub), one walk-in closet
2-car garage	3-car garage	3-car garage	3-car garage
Medium lot (0.35 acres)	Large lot (0.5 acres)	Medium lot (0.35 acres)	Medium lot (0.35 acres)
Local schools ranked in middle third of nation	Local schools ranked in bottom third of nation	Local schools ranked in top third of nation	Local schools ranked in middle third of nation
Upgraded finish package (floors, countertops, millwork, fixtures)	Upgraded finish package (floors, countertops, millwork, fixtures)	Standard finish package (floors, countertops, millwork, fixtures)	Standard finish package (floors, countertops, millwork, fixtures)
Home built in 1990	Home built in 2000	Home built in 2000	Home built in 1990
$380,750	$472,750	$636,500	$623,500
○ A possibility ○ Won't work for me	○ A possibility ○ Won't work for me	○ A possibility ○ Won't work for me	○ A possibility ○ Won't work for me

(Respondent answers 6 screens like this)

Screening Section: Background Theory

- Many academics and researchers have argued that buyers use a Consider, then Choose heuristic to navigate complex marketplaces with many/complex product choices (see Gaskin, Evgeniou, Bailiff, Hauser 2007 for a literature review).

- Non-compensatory (cutoff) rules are commonly employed in the Consider stage to develop a manageable consideration set.

- More careful consideration (perhaps more compensatory processing) is given to products within the consideration set before Choosing the final product to buy.

“Unacceptable” Questions

After a few screens (e.g. around 12 concepts), we ask the following after each subsequent Screener page:

Would any home that had the features below be totally unacceptable? If so, mark the one feature that is most unacceptable to you, and I won't bother showing you any more homes with that feature.

I've noticed that you've never said a home was a possibility if it had:

- 1400 Square Feet
- Small lot (0.2 acres)
- Home built in 1980

- None of the above are completely unacceptable.

Observe, then Confirm

- Past research with ACA suggests that respondents are too aggressive in marking levels as unacceptable if a list of levels is presented and they are asked to mark all unacceptable levels.

- In ACBC, we observe a series of previous choices, see what screening rules might be in play, and then ask respondents to confirm any rules that indeed are firm cutoffs being employed.
  - Only allow one rule specification per “Unacceptable” screen

“Must Have” Questions

- Similar in function to “Unacceptables”

Based on your choices so far, I've noticed you might have some minimum standards. Are any of these absolute requirements?

Check the one most important requirement, and I'll only show you homes that meet that standard.

- ○ At least: 2250 Square Feet
- ○ At least: 3 bedrooms, 3 baths
- ○ At least: Medium lot (0.35 acres)
- ○ At least: Home built in 2000

- ○ None of the above are absolute requirements.

Choice Task Tournament

- The final (required) section of ACBC looks like a standard CBC questionnaire (with no “None”).
- But, rather than design each concept experimentally for level balance, minimal overlap, and orthogonality, we just show product concepts from the consideration pool.
- We (typically) show them in triples, with the goal to identify the best overall product.

Choice Tasks Example question:

Among these three, which is the best option?
(We've grayed out any features that are the same, so you can just focus on the differences.)

(4 of 8)

2250 Square Feet Standard construction & efficiency	2250 Square Feet Standard construction & efficiency	2250 Square Feet Standard construction & efficiency
3 bedrooms, 3 baths Small master bath (integrated shower/tub), small walk-in closet 2-car garage	5 bedrooms, 3.5 baths Large master bath (separate shower & deep soaking tub), two walk-in closets 3-car garage	3 bedrooms, 3 baths Medium master bath (separate shower and tub), one walk-in closet 2-car garage
Medium lot (0.35 acres)	Medium lot (0.35 acres)	Medium lot (0.35 acres)
Local schools ranked in middle third of nation Upgraded finish package (floors, countertops, millwork, fixtures) New home $372,000	Local schools ranked in middle third of nation Premium finish package (floors, countertops, millwork, fixtures) Home built in 1990 $539,000	Local schools ranked in bottom third of nation Standard finish package (floors, countertops, millwork, fixtures) Home built in 1990 $374,000
○	○	◉

Calibration Concepts (Likert Scale Questions)

- This is an optional section, and not necessary unless:
  - You want to establish a new "None" threshold utility (rather than the one automatically coming from the binary screening tasks) based on a fixed threshold point on a purchase likelihood scale.

- Procedure:
  - Show "BYO" concept
  - Show a "not a possibility" concept
  - Show a "winner" from the Tournament
  - Show a "loser" from the Tournament
  - (Repeat pattern of last three, if needed)
  - Show "winning" concept from the Tournament

Calibration Concepts

How likely would you be to purchase this laptop?

*** This is the original laptop you configured ***

Size	15 inch screen, 6 lbs.
Brand	Dell
Processor	Intel Core 2 Duo T7400 (2.16GHz)
Operating System	Vista Home Premium
Memory	1 GB
Hard Drive	100 GB
Video Card	128 MB Video card, adequate for most use
Battery	3 hours
Productivity Software	Microsoft Office Small Business (Basic + PowerPoint, Publisher)
Price	$1,550

Definitely Would Not	Probably Would Not	Might or Might Not	Probably Would	Definitely Would
○	○	○	○	○

Next

Section 3

DEMO

Programming Exercise

- Simplified “Beach” Example
  - No price attribute

Specifying Attributes in ACBC

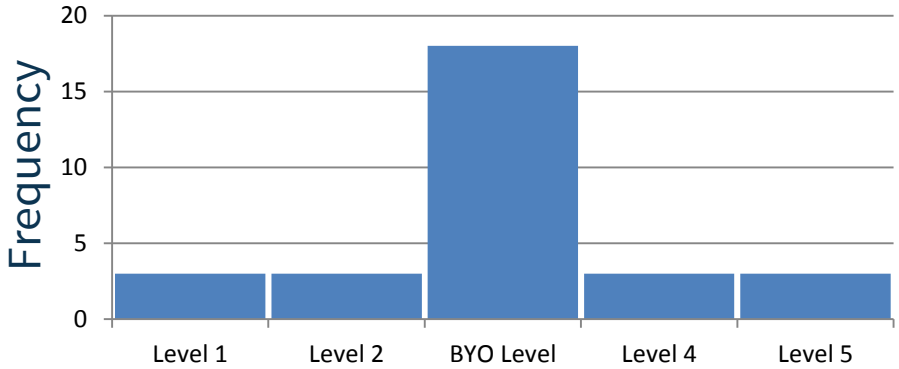
- All attributes and levels come from Lists in SSI Web
  - Either pre-defined or dynamically "constructed"
- Attribute list contains one list element per attribute: the attribute label.
- Level lists (one per attribute) contain one list element per level label within that attribute.
- Thus, an 8-attribute ACBC study will have 9 total lists for use in ACBC:
  - 1 attribute list (with 8 list elements)
  - 8 level lists (one for each attribute)

Level Settings

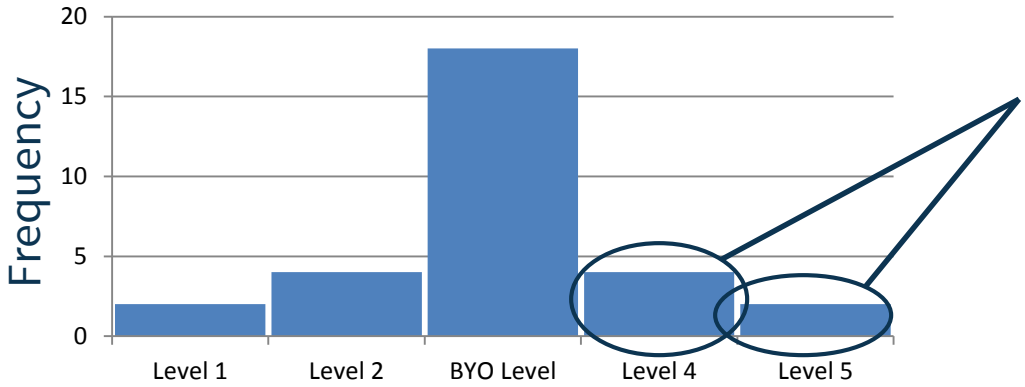
- For each attribute, you need to specify:
  - Whether it has sequential order or not
    - 10 PPM, 15 PPM, 20 PPM, etc.
  - Whether it has preference order or not
  - If preference order, the direction of preference
  - Whether to include in BYO or not

Level-Sampling Strategy Depends on Type of Attribute

- Unordered Attribute (e.g. Color)



- Ordered Attribute (e.g. MPG—fuel efficiency)



Adjacent levels 2x as likely to be selected as non-adjacent levels

Drop Attribute from BYO?

- Some attributes may seem to not make sense to ask in BYO question
  - "Obvious" answer (all respondents expected to pick same best level)
  - No tradeoff versus component price possible
- Results of dropping from BYO:
  - Lose information regarding which level is preferred that would have been provided by the BYO section
  - Even frequency balance for all levels within that attribute in near-neighbor design

Testing the Design

- ACBC includes a Test Design capability that is useful to:
  - Check if each level of each attribute appears about 2 or 3 times per respondent (rule of thumb for traditional ACBC studies)
  - Check the effect of prohibitions on design efficiency

How Many Times Each Level Appears per Respondent (Level Counts)

- Let's use the "Beach" study we just programmed
- Generate 5 test respondents (Design + Test Design)

beach Test Design

This report tabulates how many times each level appeared within each test respondent's core set of near-neighbor cards (prior to any additional replaceme
BYO-selected levels are oversampled. In general, we recommend each respondent see each level a minimum of 2 times and preferably 3 times (dependi
Test respondents are sorted from those having the lowest minimum counts on any level to those who have the highest minimum counts. Frequencies leve

		Sand softness:				Water temperature:				Public Facilities		
Respondent Number	Minimum Times Any Level Shown	Sand and rocks	Coarse sand	Medium sand	Fine sand	60°F/15°C average (wetsuit required)	70°F/21°C average	80°F/27°C average	90°F/32°C average (bathwater)	restrooms, showers or changing facilities	but no showers or changing facilities	sh ch fa av
1	2	3	2	17	2	2	2	3	17	4	17	
2	2	15	3	3	3	3	2	17	2	3	17	
3	2	2	3	17	2	3	16	3	2	18	3	
4	2	3	16	2	3	2	2	3	17	18	3	
5	2	2	17	3	2	17	3	2	2	3	17	

From the Help Manual

- "Any levels displayed to a respondent fewer than 2 times are coded in red to warn you that the design is sparse. Any levels displayed 2 times are coded in yellow, as a moderate caution. If you are interviewing 100s of respondents and are willing to sacrifice some precision at the individual level in favor of shorter questionnaires, you may decide that fielding a sparse design is perfectly suitable for your situation."

Standard Errors (n=5)

- Aggregate logit report showing standard errors without prohibitions and standard errors with prohibitions
- Due to the adaptive nature of the designs (oversampling of BYO-selected levels), standard errors for robotic respondents will not stabilize until you have about 500 or more robotic respondents.

beach Test Design

Attribute Level	Standard Errors: Design as specified	Standard Errors: Without prohibitions	Relative Efficiency: Design as specified vs. Without Prohibitions
Sand and rocks	0.2793	0.2545	83.0182 %
Coarse sand	0.3020	0.3535	136.9870 %
Medium sand	0.2769	0.2867	107.2488 %
Fine sand	0.3543	0.2670	56.8136 %
60°F/15°C average (wetsuit requir	0.3392	0.2696	63.1544 %
70°F/21°C average	0.2898	0.3769	169.1505 %
80°F/27°C average	0.2996	0.3011	101.0360 %
90°F/32°C average (bathwater)	0.2757	0.2614	89.8748 %
No public restrooms, showers or c	0.2504	0.2948	138.5896 %

Part-Worth Utility Estimation

- Johnson's Monotone Regression
- Generic HB
- Task-Specific Scale Factors HB (Otter's Method)

Johnson's Monotone Regression

- Based on "A Simple Method of Pairwise Monotone Regression", Psychometrika, 1975, pp 163-168.
- No data borrowing, so it is appropriate for very small sample sizes (n<30).
  - Can even be used for n=1.
- Utility constraints may be (and probably should be?) imposed.
- Longer questionnaires should be used (larger pool of near-neighbor concepts: each level shown at least 4 times?).

Generic HB

- HB (like that implemented in our CBC/HB software) may be used to estimate part-worths.
  - It's built in to the ACBC functionality in the SSI Web system
- Works well with as little as n=10.
- Problem: respondents have different error levels when answering BYO, Screeners, and Choice Tasks.
  - Fortunately, this problem hasn't proven very detrimental in practice

Otter's Scale-Specific HB Model

- Thomas Otter has demonstrated that part-worths can be developed with HB using a model that (in addition to the part-worth parameters) estimates an aggregate scale factor for each of the three sections.
  - BYO has largest scale (smallest error)
  - Choice tasks have smallest scale (largest error)

- His model is also included as an option within the HB estimation dialog for ACBC.

Section 4

COMPARING METHODS

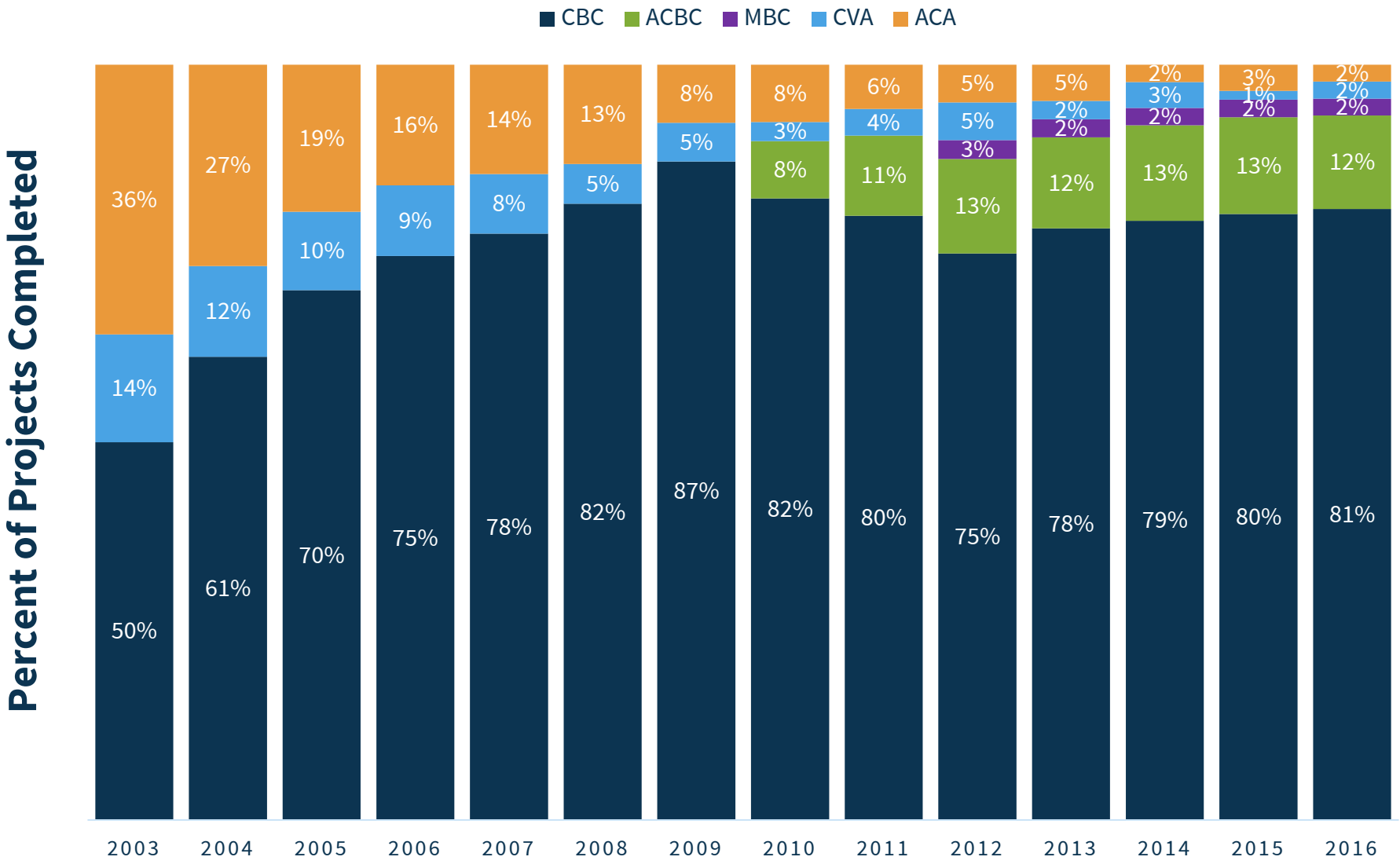
Compared to standard CBC, our work suggests that ACBC…

- Generates more accurate individual-level predictions and market simulators, especially if respondents employ non-compensatory processes.
- Requires somewhat smaller sample sizes to estimate population parameters or shares of preference.
- May even be used with n=1 for understanding a single customer, whereas standard CBC generally wouldn't be considered.
- Provides additional information for clients regarding what levels respondents screen on (must-haves and unacceptables).
- Probably better for market segmentation (latent class or cluster analysis).
- Makes respondents and clients happier.
- Can deal with greater number of attributes and levels more effectively.

But, ACBC does have its weaknesses.

- Survey is often 2-3 times longer than a comparable CBC
- Currently no support for some CBC “goodies” (chip allocation, traditional none, etc.)
- More complex to program, analyze
- Must be administered on computerized device (desktop, laptop, mobile phone, tablet but can be done “offline”)
- May be overkill for small-attribute studies (4 or fewer attributes)

Relative Use of Different Sawtooth Software Conjoint Methods



Which Conjoint Method?

Method	Minimum Sample Size	Attributes	Levels	Pricing?	Complexity (do atts freely combine?)	Fielding	Typical Use
CVA	Small	Up to 6-7	Up to 4-5	Yes, but limited	No	Paper, computer	Small attribute studies, situations where objective is to measure purchase likelihood or other direct scale elicitation, small sample size studies; may be used to generate generalized designs; situations where small, fixed design is required.
ACA	Small	Up to 30	Up to 15, but usually no more than 5	Not recommended	No	Computer only	Large attribute studies; situations where objective is to measure purchase likelihood.
CBC	Large	Standard: up to 6-7 Advanced: up to 250 (30 pre-Version 8)	Standard: up to 15, but usually no more than 5 Advanced: up to 250	Yes	Yes	Paper, computer	Competitive scenarios where choice is among multiple alternatives; pricing studies; alternative-specific studies; chip allocation studies; shelf-facing studies; fixed alternatives/competitors; many other…
ACBC	Small	Any	Any	Yes	Some	Computer	Pricing studies; large number of attributes; focus is on finding best product; allow respondents or situation to determine which attributes/levels are shown.
MaxDiff	Medium	No Attributes	Usually up to 30-40	N/A	No	Paper, computer	Lists of brands, positioning statements, specific product concepts, flavors, etc.
MBC	Very Large	Any	Any	Yes	Yes	Paper, computer	Multi-part decisions; complex models; bundling; mixed designs (CVA &CBC together).

QUESTIONS?

Aaron Hill
VP of Client Services
aaron@sawtoothsoftware.com

Megan Peitz
Ingenuity Ambassador
megan@sawtoothsoftware.com

www.sawtoothsoftware.com
+1 801 477 4700
@sawtoothsoft