# An Introduction to Choice-Based Conjoint

*with Sawtooth Software's Lighthouse Studio*

# Agenda

▸ What is Conjoint Analysis?

▸ What is Choice Based Conjoint Analysis?

▸ Creating a CBC Exercise in Lighthouse Studio

▸ Analyzing CBC Data in Lighthouse Studio

# Section 1

# WHAT IS CONJOINT ANALYSIS?

# What is Conjoint Analysis?

▸ Conjoint analysis is a quantitative research technique used to measure preferences for product features, to learn how changes to price affect demand for products or service, and to forecast the likely acceptance of a product if brought to market.

▸ Conjoint Analysis helps answer business questions like:

- How might our customers might react if we alter our product or change our price?

- Given multiple products, which one will consumers choose?  And what features drive those decisions?

WHAT CAN
CONJOINT ANALYSIS
DO FOR YOU?

**www.sawtoothsoftware.com/cavideo**

# Age Old Approaches

▶ Ask Direct Questions about preference:

- What brand do you prefer?

- How much would you pay for it?

- What color/flavor do you prefer?

▶ Answers often trivial and unenlightening

- Respondents prefer low prices to high prices, etc.

- Answers often have low discrimination - Everything is important!

- We don't know how the features combine

# How Does Conjoint Analysis Work?

▸ We break the product or service down into its component parts *(attributes)*, and the respective *levels* we want to test within the attributes

▸ We vary the product features (*independent variables*) to build many product concepts.

▸ We ask respondents to rate, rank or **choose** among a subset of those product concepts (*dependent variable*).

▸ Based on the respondents' evaluations of the product concepts, we figure out how much unique value (*utility*) each of the features added.

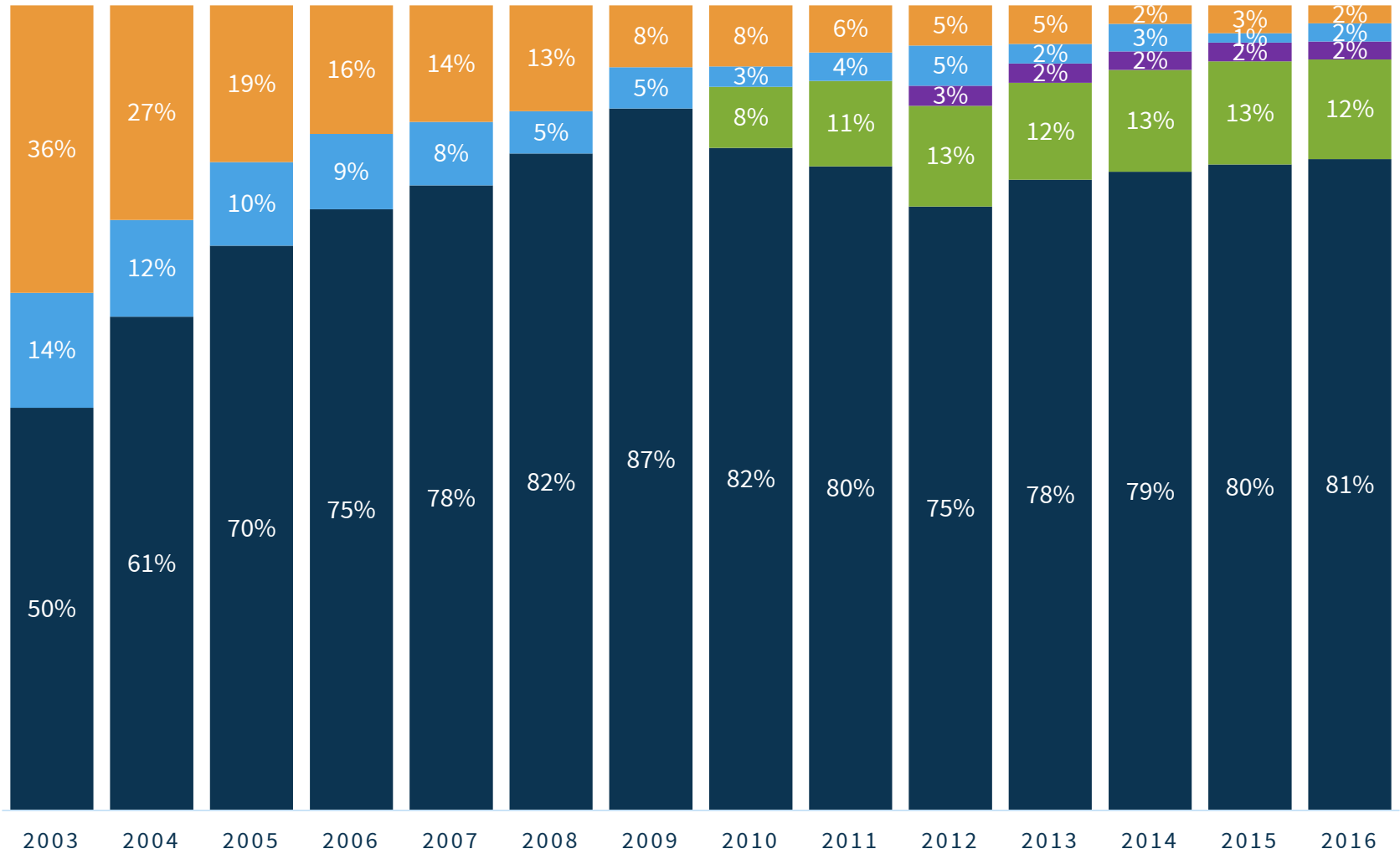When respondents are **forced to make difficult tradeoffs**, we learn what they **truly value**

# Section 2

# WHAT IS CBC?

Relative Use of Different Sawtooth Software Conjoint Methods

# What Does a CBC Experiment Look Like?

▸ Rather than rate options, respondents are simply asked to choose which they prefer.

### Which of the following do you prefer?

(Task 1 of 10)



**Attributes**

| Provider | MasterCard | VISA | AMERICAN EXPRESS | VISA | None – I wouldn't sign up for any of these credit cards |
|---|---|---|---|---|---|
| **Interest Rate** | 12% | 9.5% | 7.5% | 11.5% | |
| **Annual Fee** | No Fee | No Fee | $75 | $50 | |
| **Credit Limit** | $5,000 | $3,000 | $9,000 | $7,000 | |
| | ○ | ○ | ○ | ○ | ○ |

**Concept**

**"None" Option**

# Why Do Sawtooth Software Users Like It?

▶ Simpler for respondents because it is more like what they typically do in real life

- We get better data if our data-gathering technique more closely matches reality
- Less respondent fatigue caused by unfamiliar scales

▶ Probabilities In : Probabilities Out

▶ Flexible Design and Analysis

- Different design algorithms (balanced overlap, complete enumeration; alternative specific; conditional pricing)
- Different analysis techniques  (logit, latent class, hierarchical Bayes)

# Section 3

# CREATING A CBC EXERCISE

# What Goes Into a CBC Exercise?

▶ # of Attributes

▶ # of Levels per Attribute

▶ Prohibitions

▶ # of Tasks

▶ # of Concepts

▶ # of Versions

▶ Design method

# Defining Attributes

▶ Attributes are independent aspects of a product or service (Brand, Speed, Color, Price, etc.)

▶ How many attributes?

- Depends on research objectives
  - Respondents can often deal with more information than previously thought possible, especially with well-organized grids of information, graphical representation of levels, and improved computer interviewing approaches.
- Attributes should be independent, mutually exclusive
  - Brand, quality, product life expectancy may all measure the same thing.

# Defining Attribute Levels

▸ Each attribute has varying degrees, or "levels"

- Brand: Coke, Pepsi, Sprite
- MPG: 15 MPG, 20 MPG, 25 MPG
- Color: Red, Blue, Green, Black

▸ Each level is assumed to be mutually exclusive of the others (a product has one and only one level of each attribute)

▸ Levels should have concrete/unambiguous meaning

▸ If it is possible for a product to not have any level of a certain attribute, you must include a "Not Present" (NULL) level!

# Rules for Formulating Attribute Levels

▶ Don't include too many levels for any one attribute

- The usual number is about 3 to 5 levels per attribute
- The temptation (for example) is to include many, many levels of price, so we can estimate people's preferences for each
- But, you spread your precious observations across more parameters to be estimated, resulting in noisier (less precise) measurement of ALL price levels
- Better approach usually is to interpolate between fewer more precisely measured levels for "not asked about" prices
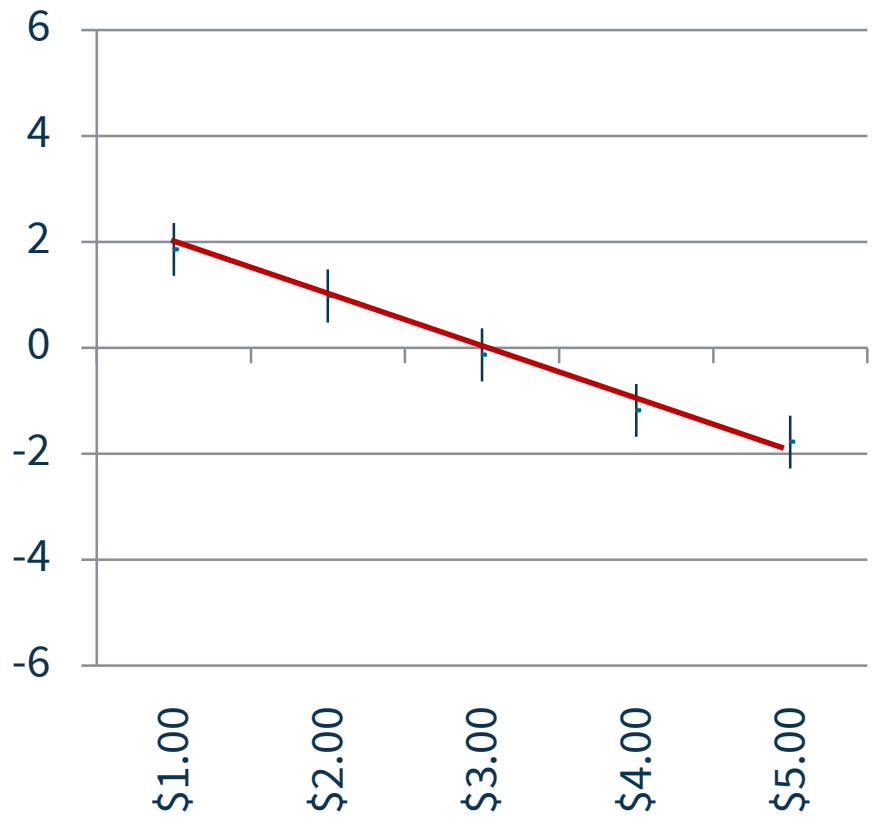
▶ HOWEVER, there are many cases where it is appropriate to have far more levels of certain attributes (Wine Brands, Shampoo SKU's, etc.)

# Price Utility Uncertainty

▶ Many price levels =
Higher Uncertainty
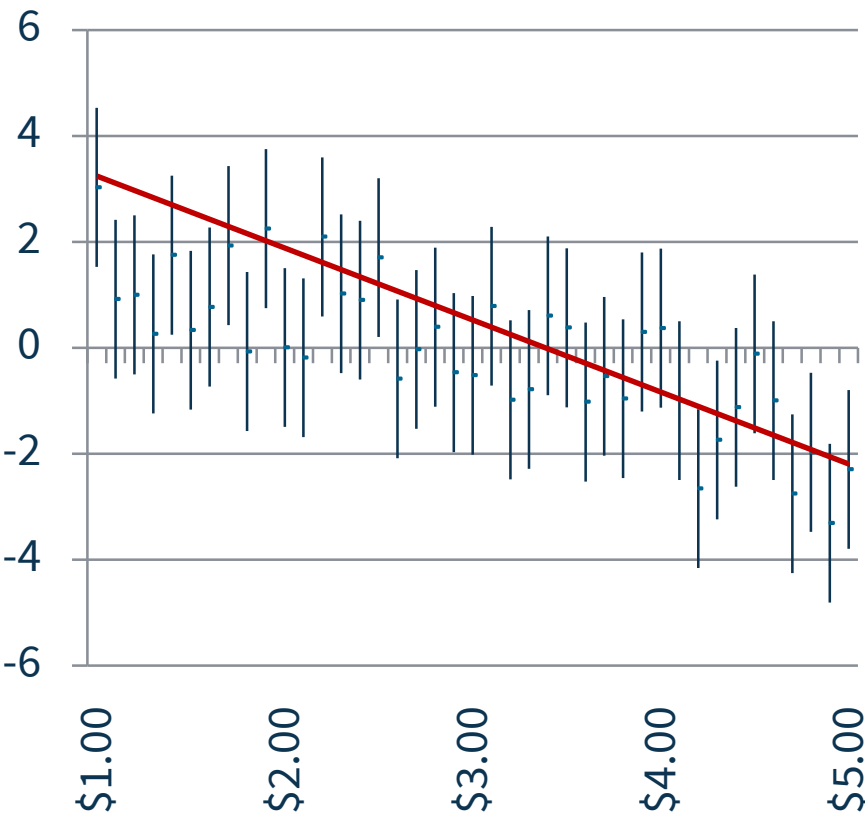
▶ Fewer price levels =
Lower Uncertainty

# Creating a CBC Exercise in Lighthouse Studio

| Brand | Style | Color | Options* | Price |
|---|---|---|---|---|
| Samsung | Side by Side | Stainless Steel | Free Delivery | $999 |
| KitchenAid | Bottom Freezer | Black Stainless Steel | Counter Depth | $1,499 |
| Whirlpool | French Door | Black | Ice Maker | $1,999 |
| Frigidaire | Top Freezer | White | | $2,499 |
| LG | | | | $2,999 |
| SubZero | | | | $3,499 |
| GE | | | | $3,999 |

▸ *Options can become 3 separate binary attributes, with levels "on" and "off", or 1 exhaustive attribute with 8 levels that states every possible combination (no items, 1 item only, 2 items, all 3 items)

# Adding Prohibitions
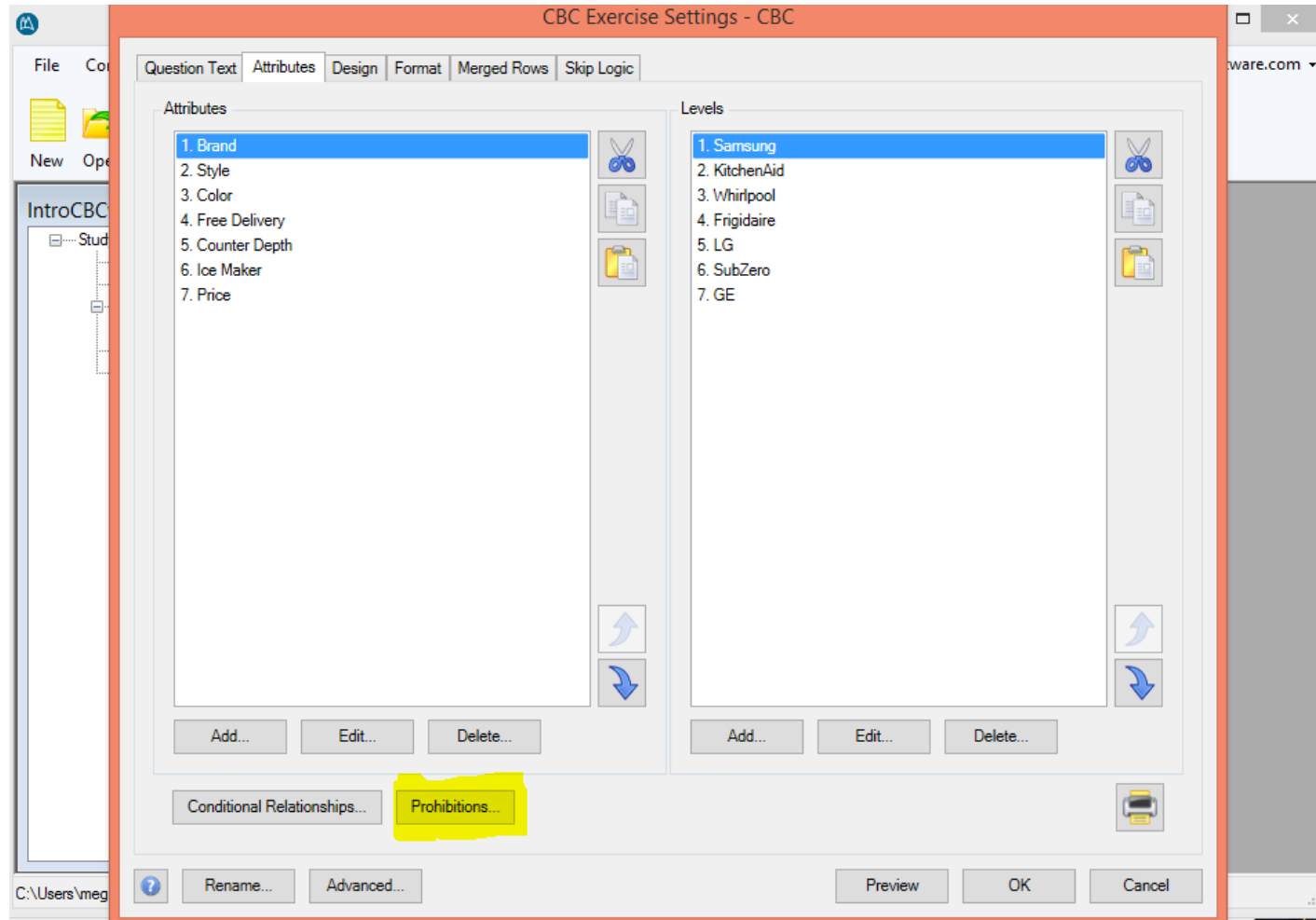
# Why are prohibitions bad?

▸ Imagine trying to create a map without visiting the entire area

▸ The CBC designer allows us to "map" the respondent's utilities

▸ Prohibitions prevent the designer from "visiting" area in the map

▸ Depending on the areas missing serious mistakes can be made

# Prohibitions

▶ The goal is to create conjoint studies without using any prohibitions, if possible.

▶ Use creative ways to eliminate prohibitions:

- Conditional pricing

- Collapsing two attributes into just one attribute

- Alternative-Specific Designs (Advanced Design Module)

# Including a "None" Option

▸ By default, a "None" concept is added to each choice task. The questionnaire author may remove the None option if desired.

▸ In LHS v9, the none option can appear inline or below the profiles (as pictured).

# Why Include a "None"?

▶ We may not want to assume that the respondent MUST choose an alternative. Therefore, we can include a "None" choice for respondents to tell us they would NOT choose any concepts in the task.

▶ Why include a none option?
- Allows respondents to avoid making uncomfortable choices
- May provide usable information about whether our concepts are acceptable, how acceptable they need to be
- Remove respondents for whom product category does not apply
  - Earlier screening also helps!
- Estimate utility for "None"

# Traditional Implementation

▸ "None" option is treated as a separate product concept in the design, is available to be selected instead of other concepts

▸ In estimation of utilities, the "None" is treated as a separate parameter

▸ In simulator, you can use "None" to estimate share of respondents that would have selected "None" in competitive context

# Problems With "None"

▸ We don't know how respondents are using it

- ▪ Fatigue / Just clicking to get through survey
- ▪ Liked option available on previous screen
- ▪ Wanting to be helpful (under-reporting None)
- ▪ Often poorly worded so definition is ambiguous

▸ It often doesn't translate well to real-world behavior, even if used appropriately

# Problems With "None"

▸ When respondents select the "None", we have less data on their relative preferences for the attributes of most interest to us

▸ Too many "None" responses effectively reduces our sample size

  ▪ Increases standard errors

# Potential Solutions to "None"

▶ Pilot test!

▶ Ask "Why did you choose None?"

▶ Make sure your definition of "None" has real meaning

- "None. I wouldn't choose any of these." vs. "None. I'd go without medical treatment." or, "I'd rather have my power cut off than pay for any of these services." or, "I'd stick with my current provider."

# Dual Response: An Alternative

▶ Rather than including "None" as a concept, split the question in two:

- First, ask respondents to choose among alternatives, not including the "None" option

- Second, ask whether they would really buy the "best" alternative

# Benefits of Dual Response None

▶ **More information, more efficient**

■ Learn about preferences (even if all alternatives are poor) with greater precision

■ It increases the utility of the None (None threshold is higher), resulting in larger share of preference for None in the market simulator

■ Many researchers believe that the higher None threshold is actually more realistic than the traditional None weight

# Benefits of Dual Response None

▶ Respondents don't feel obligated to choose a product just to be helpful

  ▪ Dual response leads to "None" choices 3-4 times higher than traditional CBC task that includes a None option. (Brazell et al. 2003)

  ▪ The prevailing hypothesis is that respondents want to help us as researchers, and feel that answering None in traditional CBC tasks is not helpful.

▶ Reduces incentive to "click through" just to get done

▶ Better information leads to better individual level utilities

# Are There Drawbacks?

▶ It does take longer – twice as many questions to answer

▶ Still doesn't necessarily lead to accurate "take rates"

- Some researchers have proposed using a 5-point purchase likelihood scale as an alternative

▶ Can't be used with chip allocation

# How Many Concepts per Task?

▶ Generally, 3 or more concepts are used (sometimes as high as 50 or more)

▶ Showing just 2 concepts is typically considered sub-optimal
   ▪ It doesn't take respondents twice as much time to respond to four concepts per task than two.

▶ Attribute text length, graphical representation affect the decision
   ▪ Notable exception: highly emotional decisions such as cancer treatment, where patients are unfamiliar with attributes and must make extremely difficult choices.

# How Many Tasks per Survey?

▸ Respondents are expensive to recruit. It makes sense to ask respondents multiple choice tasks.

▸ For AGGREGATE ANALYSIS (assuming an adequate N to begin with), doubling the number of choice tasks is nearly as valuable as doubling the sample size.

  ▪ Less true for HB analysis, where we need good individual level data.

  ▪ In typical consumer goods studies, respondents take about 5-7 minutes on average to answer 20 tasks (~20 seconds per task)

▸ CBC is very flexible in terms of how many tasks to include. Minimum is just one task! (but you'll need huge sample size, and will face limitations in analysis).

▸ Typical is **10 to 14 random** choice tasks (default is 12).

# Additional Recommendations

▸ As a rule of thumb, for individual-level estimation, we recommend that each level occur approximately 6x or more for each respondent.

$$\frac{(\#\ of\ Tasks\ \times\ \#\ of\ Concepts)}{(\#\ of\ levels\ in\ largest\ attribute)} \geq 6$$

# Fixed ("Holdout") Tasks

▶ Fixed (attribute levels not varying) CBC questions that we recommend you add to every CBC study

- Tasks (observations) are "held out" (not included) in the utility estimation
- Holdouts used to test specific scenarios the client is interested in
- Holdouts used to test the internal validity of different models (main effects/interactions) and different estimation methods (logit, Latent Class, HB)
- Holdout can be used to identify inconsistent respondents (should repeat holdouts, and compute test-retest reliability)

# Number of Versions

▸ With computerized interviewing, it is easy to administer hundreds of versions to respondents

  ▪ Avoids multicollinearity

  ▪ Control for context effects, order effects

▸ Our recommendations are:
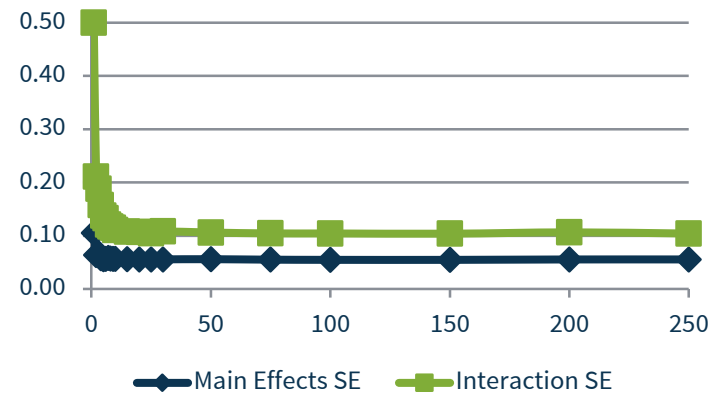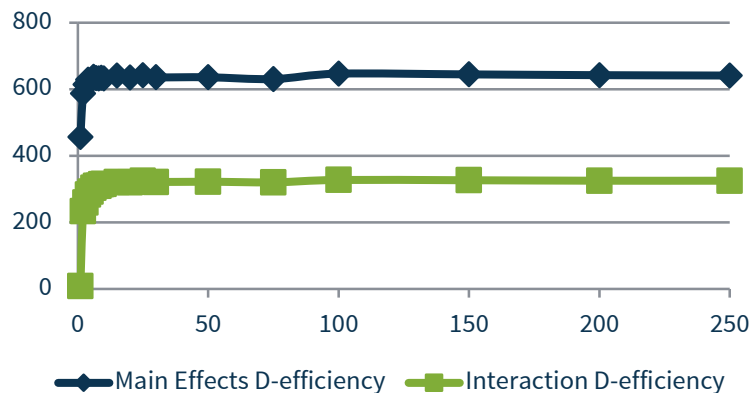
# Generating Designs

▸ With computerized interviewing, it is easy to administer hundreds of versions to respondents

- Avoids multicollinearity
- Control for context effects, order effects
- Exposure to more combinations of features across respondents

▸ With paper questionnaires, it is typical to use just 4 or 5 different versions of the design

- Avoid headache of managing lots of questionnaire versions
- Reduce data input errors

# Four Design Methods in CBC

| Criteria | Complete Enumeration | Shortcut | Balanced Overlap | Random |
|---|---|---|---|---|
| One Way Level Balance | Strong | Strong | Strong | Weak |
| Two Way Level Balance | Strong | Weak | Strong | Weak |
| Overlap | Minimal | Minimal | Encouraged | Significant |
| Main Effects Efficiency | Great | Very Good | Good | Okay |
| Interaction Efficiency | Okay | Okay | Great | Best |
| Computation Time | Slow with large # of attributes, design space | Speedy | Slow with large # of attributes, design space | Speedy |
| Notes | Minimal overlap can encourage trivial behavior | Minimal overlap can encourage trivial behavior | Software default – good balance between all criteria. | Can show concepts that are nearly identical;  best at handling very extreme prohibitions |

# Balanced Overlap Strategy (Default)

▶ Features:
- One-way Level Balance (each level occurs an equal number of times)
- Two-way Level Balance (each pair of levels occurs an equal number of times)
- Some Level Overlap (an advantage for precision of interactions)

▶ Strengths:
- More efficient than Shortcut and Complete Enumeration for estimating interaction effects
- Encourages respondents to use "deeper" choice strategies than minimal overlap approaches

▶ Weaknesses:
- As slow as Complete Enumeration
- Sacrifices some efficiency in estimating main effects

# Complete Enumeration Strategy

▶ Features:

- One-way Level Balance (each level occurs an equal number of times)
- Two-way Level Balance (each pair of levels occurs an equal number of times)
- Minimal Overlap (ideal for precision of main effects)

▶ Strengths:

- Usually slightly more efficient than Shortcut (except in extreme prohibitions)

▶ Weaknesses:

- Can be slow to compute with more than about 8 attributes.  Often not good for interactions.  Minimal overlap can encourage trivial choice behavior in respondents (e.g. always picking the one "must have" brand and not learning anything else about the other attributes).

# Shortcut Strategy

▶ Features:

- One-Way Level Balance (each level occurs an equal number of times)
- Loose (not strict) orthogonality
- Minimal Overlap (ideal for precision of main effects)

▶ Strengths:

- Fast computation
- Handles some extreme cases of prohibitions better than other techniques

▶ Weaknesses:

- Two-way level balance (between attributes) not controlled
- Therefore, slightly less efficient than "Complete Enumeration"
- Poor choice for interactions.  Can encourage trivial choice behavior.

# Random (Truly) Strategy

▸ Features:

- Only loose one-way Level Balance (each level occurs approximately an equal number of times)

- Only loose two-way Level Balance (each pair of levels occurs approximately an equal number of times)

- Significant Level Overlap (an advantage for precision of interactions)

▸ Strengths:

- Very fast computation

- Most efficient of four methods at interaction effects and cross-effects

- Best at handling very extreme prohibitions

▸ Weaknesses:

- Sacrifices a great deal of efficiency in estimating main effects

# Test CBC Design

▶ ALWAYS test your design!

- CBC automatically runs a quick frequency test when it generates the design file, warning you if it sees potential problems

▶ Test different sample sizes. Test with and without a None. Test interactions.

▶ *Also, make sure you actually **test your survey personally**. Sometimes a statistically efficient design will have quirks in what it shows to respondents, question wording does not make sense, etc., requiring adjustments to the design.*

# Test Design Capability

▸ Simulated Respondents, Logit Rule, D-Efficiency

  ▪ Reports absolute precision of levels, given your sample size and expected None% rate

  ▪ Reports overall precision (strength) of design in terms of D-Efficiency

▸ Uses aggregate logit utility estimation to estimate the precision of the utilities (standard errors)

▸ Our objective is to minimize the standard errors, which means we are estimating the utilities with good precision

# Test Design Report

▶ Rules of Thumb

- Standard Error for Main Effects less than 0.05
- Standard Error for Interactions/alt-specific effects less than 0.10
- No warning about "ridge adjustment" or *********'s

▶ Standard Error can be decreased by increasing sample size, concepts per task or number of tasks, removing prohibitions, increasing number of versions

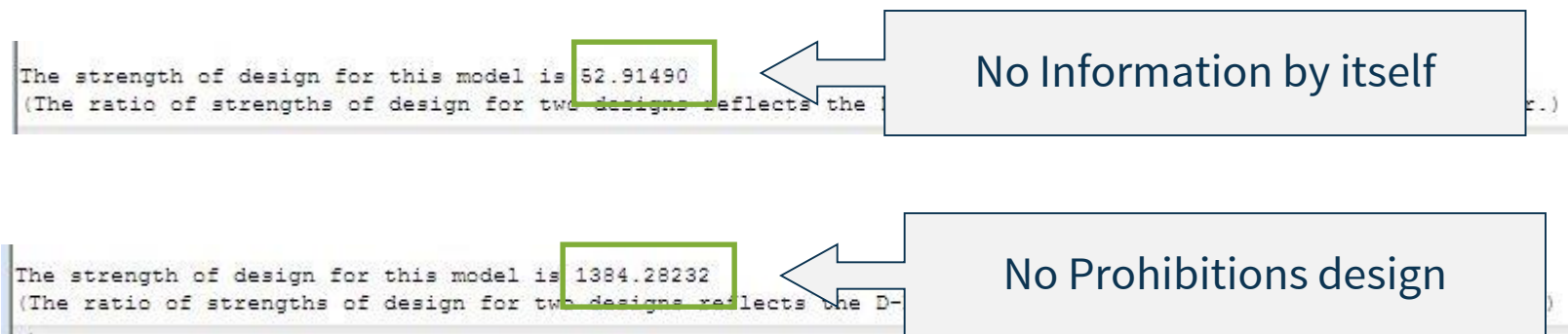# Do we NEED low standard errors?

▶ The rules of thumb we give are generally safe for most studies where the purpose is to establish stable simulation results

▶ Sometimes, though, (if we're careful!) we can get away with much higher standard errors if we are just interested in directionality

- Should we implement Feature X or Feature Y?
  - If simulations show large enough preference for Feature X over Feature Y, even after accounting for large standard errors, you may be able to get by with much smaller sample size

# D-Efficiency

▶ Measures the relative amount of information the design can collect

▶ The larger the number the better

▶ Dependent on the number of attributes and levels in the design.  (Number of parameters)

- Cannot be compared across designs with different number of attributes and levels  or interactions

# D-Efficiency

▸ Relative D-Efficiency (relative strength of design) lets you evaluate the overall (average) precision of one design relative to another (considering all parameters of interest).

▸ For example, if design A has a strength of design of 52.91 and design B has strength of 1384.28, design A is 52.91/ 1384.28= 3.8% as efficient as design B!

The strength of design for this model is 52.91490
(The ratio of strengths of design for two designs reflects the ... r.)

No Information by itself

The strength of design for this model is 1384.28232
(The ratio of strengths of design for two designs reflects the D-

No Prohibitions design

# Manual Testing

▶ **Survey Instrument Testing**

- How long does the survey take?

- Did you accidentally duplicate a level? Forget a level?

- Do all the labels/concepts make sense? Is instruction text or "ToolTip" help needed?

- Smartphone vs. laptop, scrolling issues, layout issues

▶ **Process Testing**

- How will you combine multi-language survey data together?

- Generate fake data and run through the analysis. Does the simulator give you what you need?

- (Clean out test data before going live)

# Exporting the Design

▸ CBC gives you an easy way to share the design (combinations of levels included in each choice task) with another person

▸ Go to CBC Settings | Design Tab | Import/Export Design button
  ▪ Saves the results in .csv format

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | Version | Task | Concept | Att 1 - Brands: | Att 2 - Colors: | Att 3 - Size: | Att 4 - Price: |
| 2 | 1 | 1 | 1 | 2 | 3 | 3 | 3 |
| 3 | 1 | 1 | 2 | 1 | 1 | 1 | 1 |
| 4 | 1 | 1 | 3 | 3 | 2 | 2 | 2 |
| 5 | 1 | 2 | 1 | 2 | 2 | 1 | 1 |
| 6 | 1 | 2 | 2 | 3 | 1 | 2 | 3 |
| 7 | 1 | 2 | 3 | 1 | 3 | 3 | 2 |
| 8 | 1 | 3 | 1 | 3 | 3 | 1 | 2 |
| 9 | 1 | 3 | 2 | 2 | 1 | 3 | 1 |
| 10 | 1 | 3 | 3 | 1 | 2 | 2 | 3 |

# Importing Designs

▸ You can import designs from the very same .csv format!

▸ Thus, if you have statisticians in house who prefer to generate their own designs rather than use Sawtooth Software's designs, you can say "No problem, send it to me and we'll use it instead!"

▸ Plus, you can use the Test Design options to test the quality of the imported design (for sanity check).

# New Mobile Features in LHS v9

# Link to Another Survey



▸ LHS surveys may interact with other survey and panel providers!

▸ Our surveys have the ability to "pass-in" and "pass-out" data through standard URL query strings. Not only can you link to a Sawtooth survey from a survey created by a different survey company, but you can also redirect the respondent into a different survey.

▸ Our software is also designed to accommodate the requirements of most third-party panel providers.

▸ Check out the 5-minute tutorial here
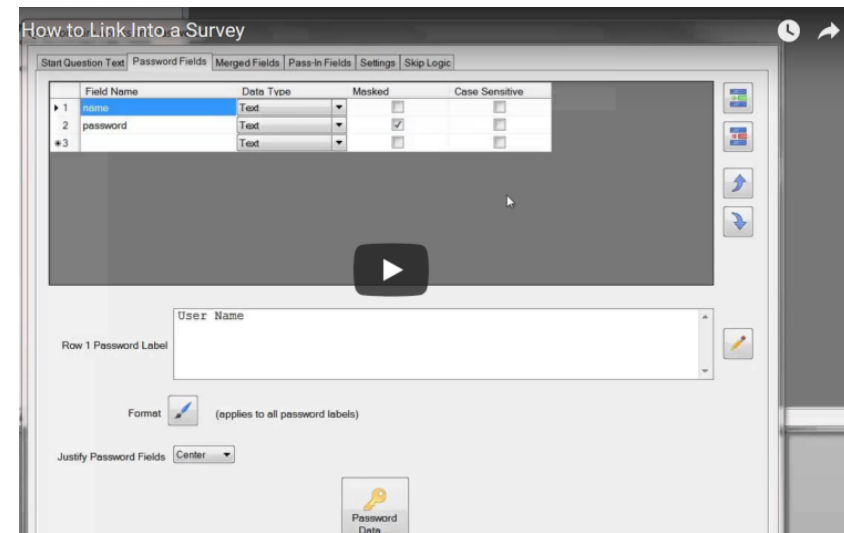
# So, What Does the Final CBC Exercise Look Like?

▶ https://IntroCBCwLHS.sawtoothsoftware.com/login.html

# START USING LHS TODAY!

# Lighthouse Studio Component Subscriptions

▸ Lighthouse Studio is sold as a 12-month subscription. Free technical support, maintenance, online hosting for data collection, and upgrades are included.

▸ Each component license comes with General Interviewing 50 license and the conjoint components come with the Choice Simulator.

| Component Subscriptions | Size | 1 Person |
|---|---|---|
| **CBC (Choice-Based Conjoint)** | 10 attributes, 15 levels per attribute | $5,500 |
| **CBC ADM (CBC + Advanced Design Module)** | 250 attributes | $7,000 |

# Suite Subscriptions

▶ Should you intend to execute more than one methodology, the bundled Suite Subscription offers incredible savings.

| Suite Licenses | 1 Person | 5 Person | 10 Person |
|---|---|---|---|
| **Premium Suite Plus**<br>Includes Premium Suite, five hours of free consulting, and 15% off workshops and conferences | $15,500 | $25,100 | $31,900 |
| **Premium Suite**<br>Includes Standard Suite, MBC, and unlimited CCS desktop client conjoint simulators | $13,900 | $23,500 | $29,900 |
| **Standard Suite**<br>Includes all components except MBC and client simulators | $9,900 | $19,500 | $25,600 |

▶ In addition to free technical support, maintenance, online hosting for data collection, and free upgrades, the Standard Suite offers the following:

Discover • CBC ADM • ACBC • ACA 30 • CVA • General Interviewing Unlimited • MaxDiff 500 • MaxDiffAnalyzer.com • CBC/HB • CBC/Latent Class • ACA/HB • CVA/HB • CCEA • HB-Reg • Choice Simulator • Advanced Simulation Module • Online Market Simulator

# Learn Sawtooth Software through our Workshops, Trainings, Forum and Conference

▶ In addition to free tech support and extensive manuals, our users will have access to many additional resources.

- Held every 3 months
- ~$500-$600 per day (Not including hotel/travel)
- Capped at 25 attendees for hands on learning

- $8,000-$12,000 for on-site training
- $1,000/3 hr. webinar training
  - Access to free webinars

**Workshops**

**Trainings**

**Conference**

**Forum**

- Held every 18 months
- Best and most practical marketing sciences conference
- ~$500-$600 per day (Not including hotel/travel)
- Typically ~200 attendees

- Extensive, free resource of answers to questions about methodology, design, programming, etc.
- Comprised of hundreds of users

# QUESTIONS?

Megan Peitz
Ingenuity Ambassador
megan@sawtoothsoftware.com

Brian McEwan
VP, Client Services
brian@sawtoothsoftware.com

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