

An Introduction to Choice-Based Conjoint

with Sawtooth Software's Lighthouse Studio



Part 2

ANALYSIS & MARKET SIMULATIONS







Our CBC Exercise

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

<https://IntroCBCwLHS.sawtoothsoftware.com/login.html>

If these were your only options, which would you choose?

(1 of 12)

Brand	KitchenAid	 LG	FRIGIDAIRE	 Whirlpool
Style	 Top Freezer	 Bottom Freezer	 French Door	 Side by Side
Color	White	White	Black Stainless Steel	Black
Free Delivery	✓			✓
Counter Depth	✓	✓		
Ice Maker				✓
Price	\$1,499	\$1,999	\$999	\$2,999
	Select	Select	Select	Select

NONE: I wouldn't choose any of these.

Select

ANALYZING CBC DATA

Three Conjoint Analysis Outputs



Counts

Utilities

Importance
scores

Counting Analysis

- ▶ Divide the number of times a level was chosen by the number of times it was available (presented to respondents in the questionnaire)
 - Ratio scaled
- ▶ “Gut Check” of the data

Counting Analysis Demo

The screenshot displays the Analysis Manager software interface. The top navigation bar includes 'Home', 'Add', 'Duplicate', 'Rename', 'Run', 'Analysis Types' (set to 'Counts'), 'Save Summary', and a help icon. Below this are three tabs: 'ANALYSIS RUNS', 'RUN SETTINGS', and 'REPORTING'. The main workspace shows a table with columns A, B, and C. The table is divided into sections for 'Brand' and 'Style'. The 'Brand' section lists various brands and their corresponding proportions. The 'Style' section lists 'Total Respondents' and their proportion. The bottom status bar shows 'CBC Counts' and 'Settings Summary'.

	A	B	C
1	Choice Tasks Included	All	
2		Random1, Random2, Random3, Random4, Random5, Random6, Random7, Random8, Random9, Random10, Random11, Random12	
3			
4	Brand		
5			Total
6	Total Respondents		301
7	Samsung		0.232
8	KitchenAid		0.208
9	Whirlpool		0.239
10	Frigidaire		0.155
11	LG		0.188
12	SubZero		0.139
13	GE		0.181
14			
15	Within Att. Chi-Square		90.197
16	D.F.		6
17	Significance		p < .01
18			
19			
20	Style		
21			Total
22	Total Respondents		301

Counting Analysis

- ▶ “Counts” proportions only sum to 100% within an attribute if:
 - Each level for the attribute was shown exactly once in each choice task
 - No “None” option
- ▶ Counts sums exceed 100% if the number of alternatives per task is less than the number of levels in a given attribute

Brand A	Brand B	Brand C	Brand A
Red	Blue	Red	Blue
\$10	\$15	\$20	\$25

Counting Analysis

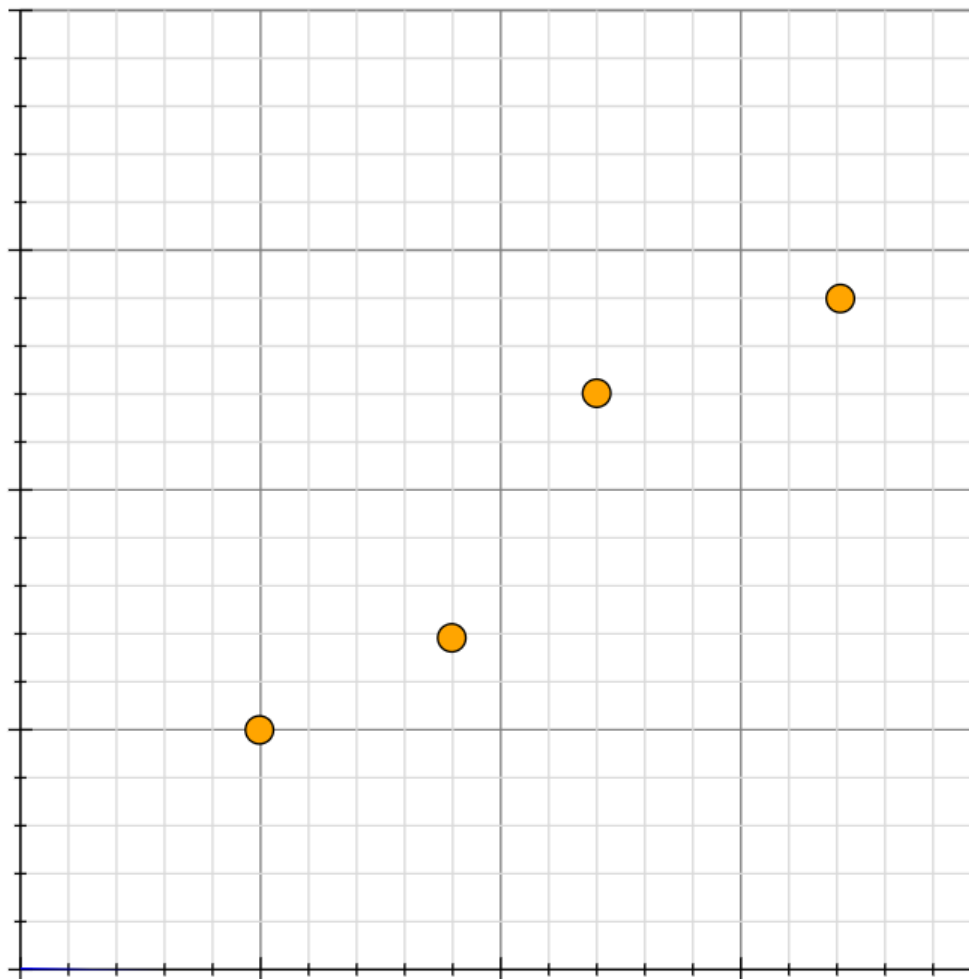
- ▶ Counts will only follow rational expectations if your design is not confounded or correlated!

Brand A	Brand B	Brand C	Brand D
Tastes Like Strawberries	Tastes Like Vanilla	Tastes Like Garbage	Tastes Like Dirt
\$25	\$20	\$5	\$10

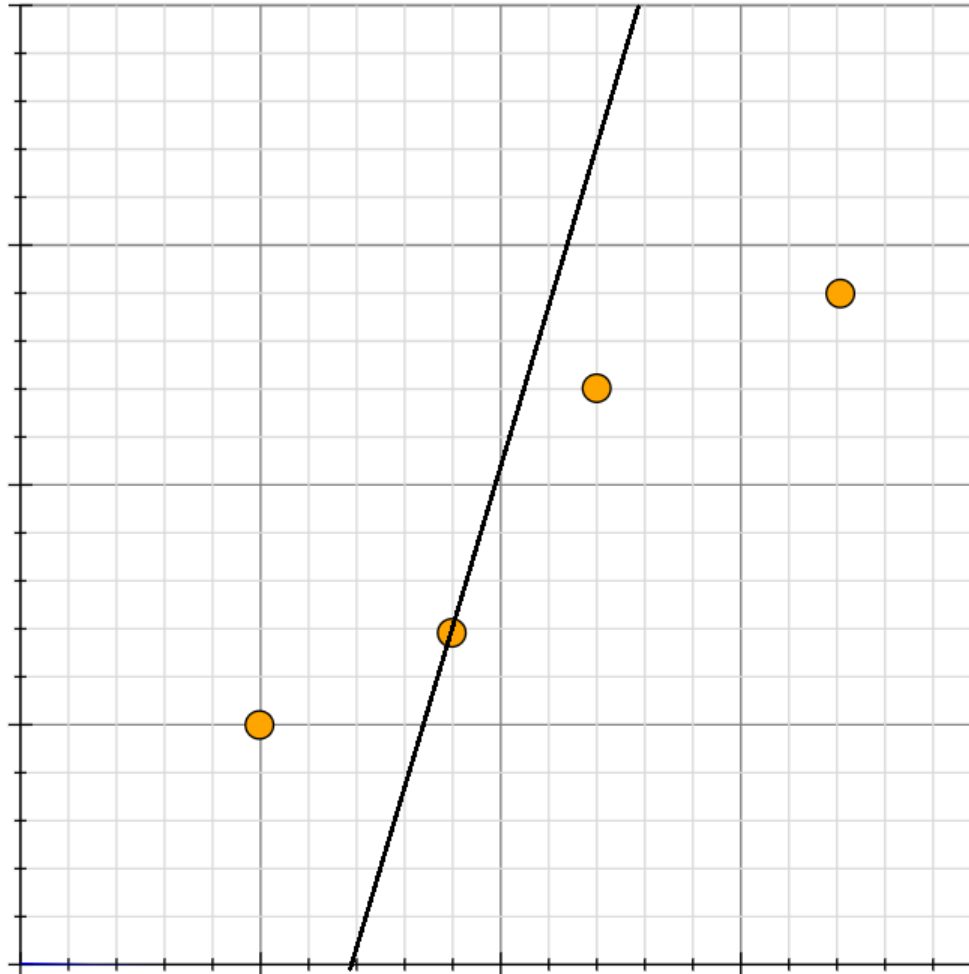
Counting Analysis

- ▶ Counts proportions show on average which levels were preferred for a sample of respondents
- ▶ Better to build a generalizable model of preference
 - How do you feel about Brand A versus Brand B
 - How important are Brand and Price in your decision
- ▶ Generalizable models allows us to understand the preferences of combinations of levels

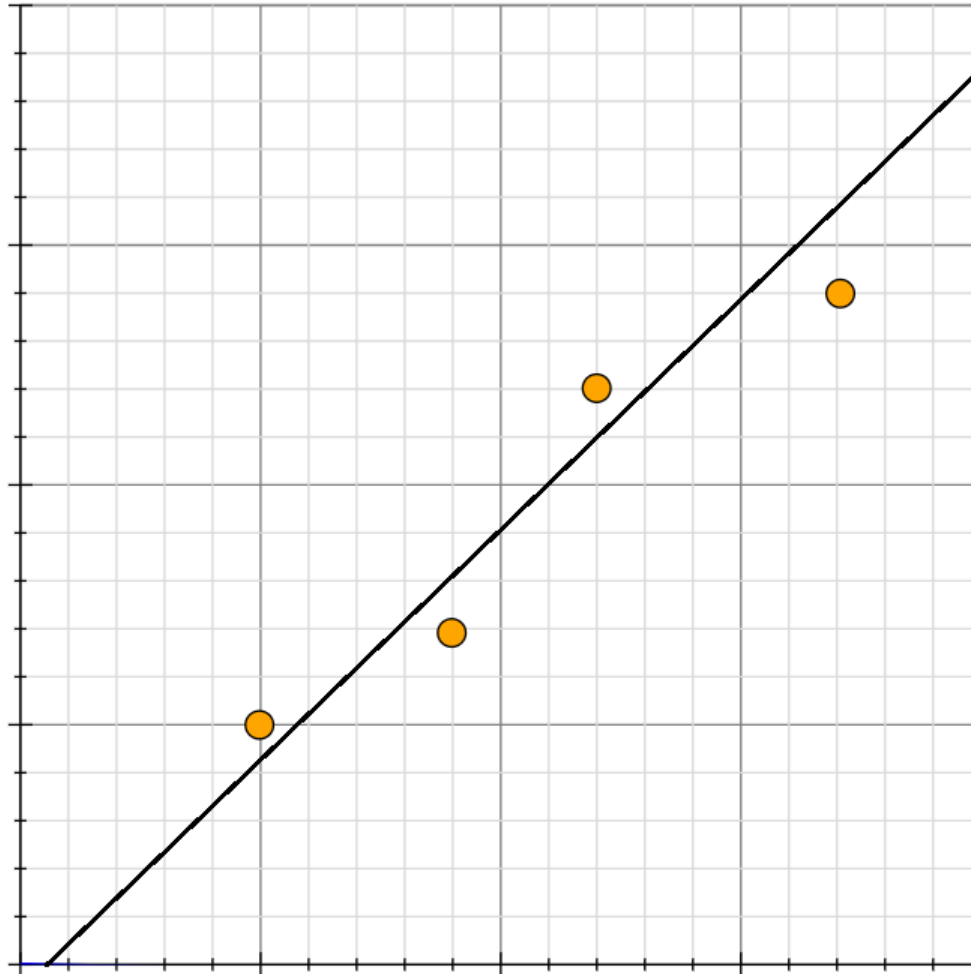
Modelling Data



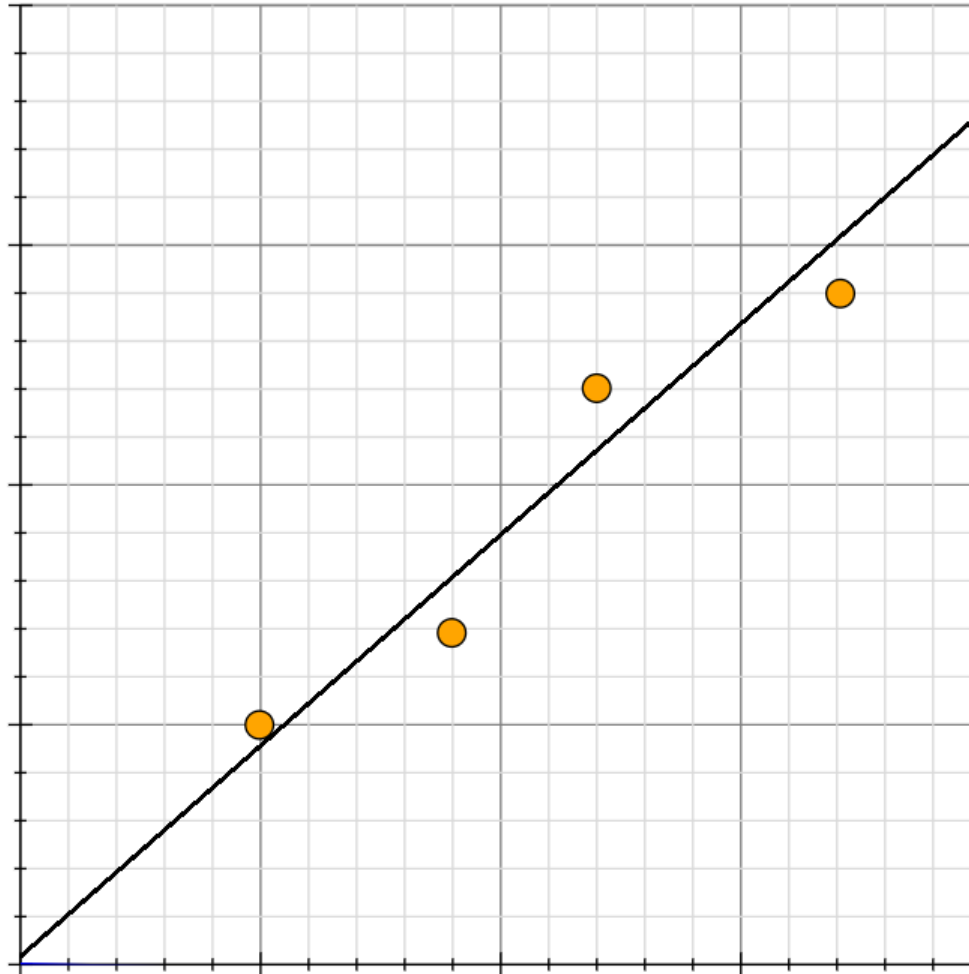
Modelling Data – Linear Regression



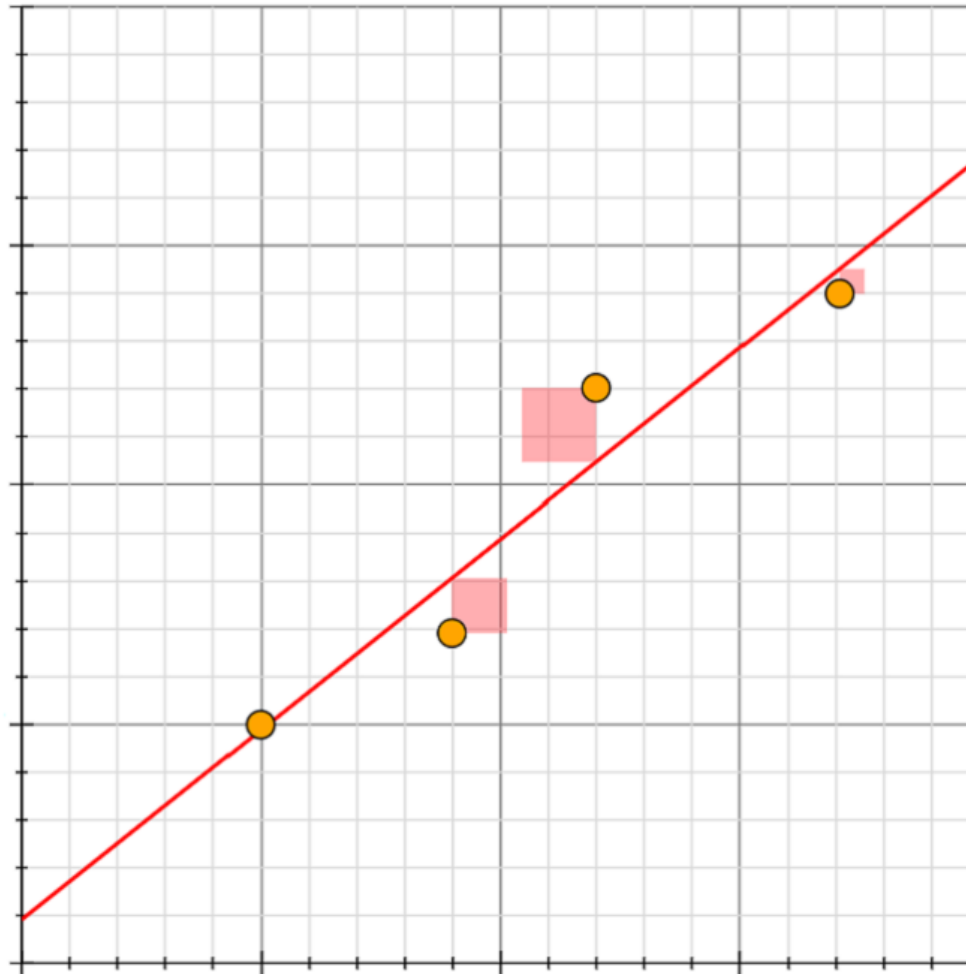
Modelling Data – Linear Regression



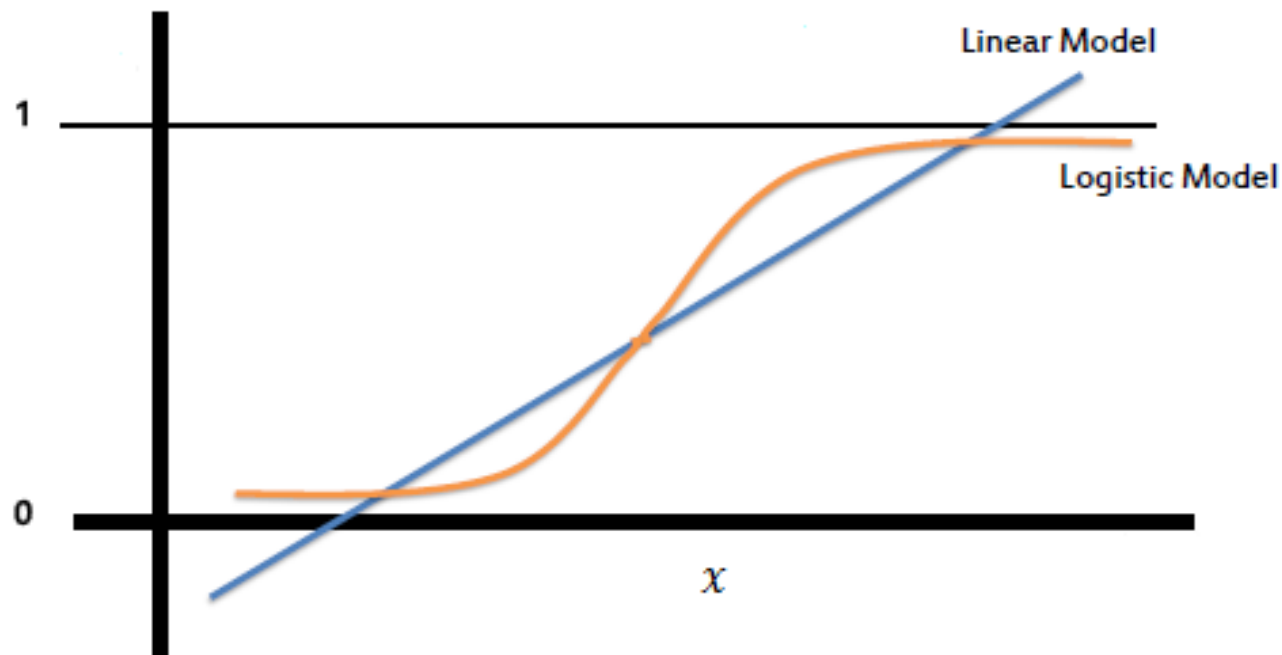
Modelling Data – Linear Regression



Modelling Data – Least Squares Regression



Modelling Data - Logistic Regression



Logistic Regression

- ▶ “Utility values” are developed for each level such that the following model produces “maximum likelihood fit” to the actual choices:

Consider three alternatives (A, B, C) in a choice set:

$$P_A = \exp(U_A) / [\exp(U_A) + \exp(U_B) + \exp(U_C)]$$

where:

P_A = “Probability of choosing alternative A”

exp means “exponentiate or antilog” or raising e to the power of the total utility

U_A = Total utility for alternative A, etc.

Logistic Regression

Brand X	Brand Y
Red	Blue
Feature B	Feature A
Yes	Yes
Automated	Automated
\$150	\$100



X	
Y	
Red	
Blue	

Logistic Regression

Brand X	Brand Y
Red	Blue
Feature B	Feature A
Yes	Yes
Automated	Automated
\$150	\$100



X	.2
Y	-.2
Red	.2
Blue	-.2

$$P(A) = \frac{\exp(A)}{(\exp(A) + \exp(B))}$$

$$P(A) = \frac{\exp(.4)}{(\exp(.4) + \exp(-.4))} = 69\%$$

$$P(B) = \frac{\exp(-.4)}{(\exp(.4) + \exp(-.4))} = 31\%$$

Logistic Regression

Brand X	Brand Y
Red	Blue
Feature B	Feature A
Yes	Yes
Automated	Automated
\$150	\$100



X	.2
Y	-.2
Red	.2
Blue	-.2

$$P(A) = \frac{\exp(A)}{(\exp(A) + \exp(B))}$$

$$P(A) = 69\%$$

$$P(B) = 31\%$$

Logistic Regression

Brand X	Brand Y
Red	Blue
Brand X	Brand Y
Blue	Red
Feature B	Feature A
Yes	Yes
Automated	Automated
\$150	\$100



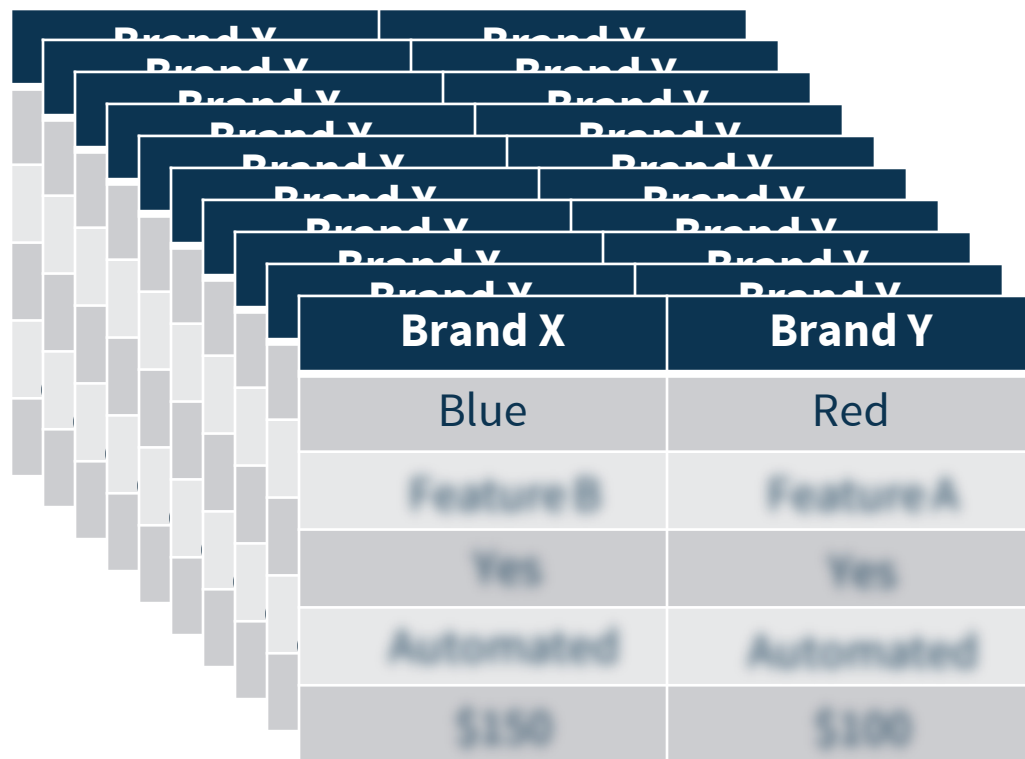
X	1
Y	-1
Red	0
Blue	0

$$P(A) = \frac{\exp(A)}{(\exp(A) + \exp(B))}$$

$$P(A1) = 88.1\% \quad P(A2) = 88.1\%$$

$$P(B2) = 11.9\% \quad P(B2) = 11.9\%$$

Logistic Regression



Brand X	Brand Y
Blue	Red
Feature B	Feature A
Yes	Yes
Automated	Automated
\$150	\$100

X	?
Y	?
Red	?
Blue	?

What utilities can we come up with to maximize the probability of the items the respondent actually chose?

Aggregate Logit

- ▶ Aggregate multinomial logit module
- ▶ Combines all respondents' choice tasks together
- ▶ Estimates a single set of utilities to fit the total sample





Aggregate Logit

- ▶ Pooling respondents provides statistical power to estimate utilities for the sample with potentially excellent precision
- ▶ But, only having a single set of utilities to reflect all respondents assumes everybody has equal preference (homogeneity)
- ▶ This is a problematic assumption for market simulations

Aggregate Logit Demo



Analysis Manager


Home

 Add
  Duplicate
  Rename
  Run

Analysis Types

Logit

 Utility Report
  Export Utilities

 Save Summary

ANALYSIS RUNS RUN SETTINGS REPORTING

C Counts x L Logit x LG Latent Class x HB HB x

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
16	Alkaike Info Criterion	9461.31166																						
17	Consistent Alkaike Info Criterion	9619.51166																						
18	Bayesian Information Criterion	9597.51166																						
19	Adjusted Bayesian Info Criterion	9527.60667																						
20	Chi-Square	2196.39231																						
21	Relative Chi-Square	99.83601																						
22																								
23																								
24	Variable	Effect	Std Error	t Ratio																				
25	Samsung	0.10216	0.05227	1.95455																				
26	KitchenAid	0.03093	0.05467	0.56569																				
27	Whirlpool	0.23007	0.05245	4.38662																				
28	Frigidaire	-0.27375	0.05989	-4.57112																				
29	LG	-0.05862	0.05580	-1.05056																				
30	SubZero	-0.24496	0.06380	-3.83802																				
31	GE	0.21406	0.05830	3.67155																				
32																								
33	Side by Side	-0.30205	0.04153	-7.27362																				
34	Bottom Freezer	0.37966	0.03632	10.45263																				
35	French Door	0.53827	0.03545	15.18377																				
36	Top Freezer	-0.61589	0.04461	-13.80492																				
37																								
38	Stainless Steel	0.32361	0.03605	8.97730																				
39	Black Stainless Steel	0.48116	0.03556	13.53027																				
40	Black	-0.21908	0.04064	-5.39096																				
41	White	-0.58569	0.04443	-13.18300																				
42																								
43	Yes	0.13826	0.02191	6.31018																				
44	No	-0.13826	0.02191	-6.31018																				
45																								
46	Yes	0.10455	0.02192	4.76958																				
47	No	-0.10455	0.02192	-4.76958																				
48																								
49	Yes	0.22317	0.02198	10.15452																				
50	No	-0.22317	0.02198	-10.15452																				
51																								
52	\$999	1.06359	0.04967	21.41512																				
53	\$1,499	0.80825	0.04869	16.59854																				
54	\$1,999	0.74158	0.04996	14.84254																				
55	\$2,499	-0.10168	0.06041	-1.68312																				
56	\$2,999	-0.13329	0.06029	-2.21074																				
57	\$3,499	-1.08563	0.08114	-13.37913																				
58	\$3,999	-1.29282	0.08609	-15.01786																				
59																								
60	NONE	0.00000	0.00000	0.00000																				

Latent Class

Latent

adjective

la·tent

\ 'lā-tənt \

existing but not yet developed or manifest; hidden; concealed

Latent Class

- ▶ Intuitively, it is like cluster analysis and logit analysis combined:
 - We find different groups of respondents (latent classes) that have similar preferences
 - Within each group (class) of respondents, we develop a logit utility solution to reflect the part worth utilities for that class
- ▶ But, respondents are not fully assigned membership in just a single group but instead have a probability of membership in each class.

Latent Class

- ▶ A great option to look for segmentation strategies
- ▶ Has better statistics (based on likelihood fit) than cluster analysis for determining an appropriate number of groups.
- ▶ Relaxes the assumption of population homogeneity—but still we don't have true individual-level data!
- ▶ “Friendly” algorithm that will always produce an N-group solution if we ask it to

Latent Class Demo

The screenshot displays the Analysis Manager software interface. The main window shows the 'Latent Class' analysis results, including a table of 'Summary of best replications' and 'Summary of all replications'. A dialog box titled 'CBC Latent Class Settings' is open, showing various configuration options for the analysis.

Analysis Manager Interface:

- Home** tab: Add, Duplicate, Rename, Run, Analysis Types (Latent Class), Utility Report, Export Utilities, Save Summary.
- ANALYSIS RUNS** tab: Counts, Logit, Latent Class, HB.
- Latent Class Estimation (11/15/2017 8:50:06 AM)**

Summary of best replications:

Groups	Replication	Log-lik
2	4	

Summary of all replications:

Groups	Replication	Log-lik
2	1	
2	2	
2	3	
2	4	
2	5	

CBC Latent Class Settings:

- Respondent Id**: Internal Interview Numbers
- Respondent Filter**: Include All Respondents
- Weights**: Equal
- Tasks**: All
- Attribute Coding**: Part-Worth, Main-Effects
- Constraints**: Default
- Estimation Settings**: Modified
- Advanced**: Default
- Advanced Output Options**: Default

Solutions:

- Minimum number of groups: 2
- Maximum number of groups: 2
- Number of replications for each solution: 5
- Maximum number of iterations: 100
- Convergence limit for log likelihood: 0.01

Estimation will run solutions between 2 and 2 groups, replicating each solution 5 times, for a total of 5 estimation runs.

Buttons: Restore Defaults, OK, Cancel.

Hierarchical Bayes

- ▶ Our recommended approach for estimating models is HB, which leverages Bayesian statistics to provide individual respondent-level models
 - Use Counts, Aggregate Logit for initial guidance and “gut check”
 - Use Latent Class to investigate segments
 - Use HB for your actual model

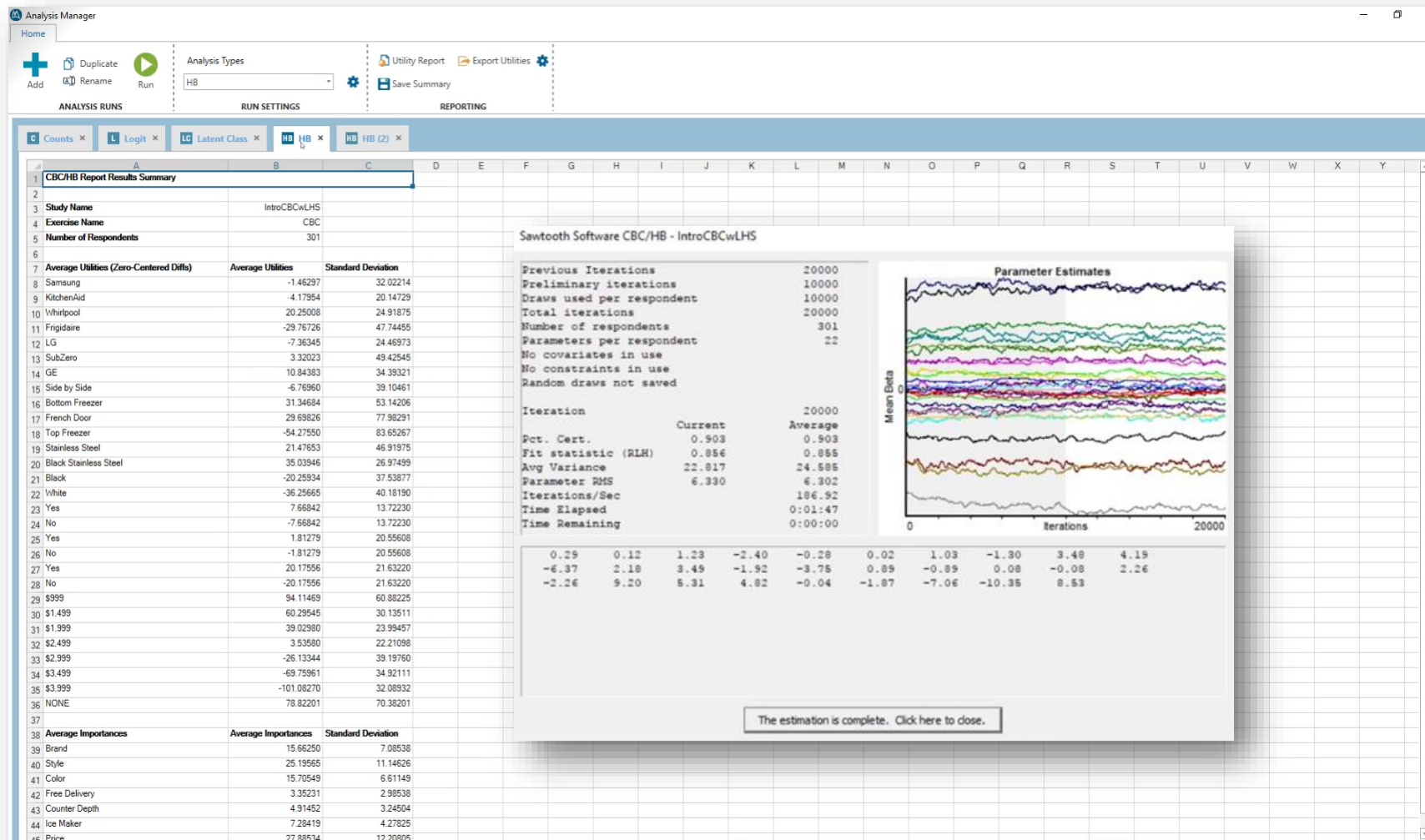
Strengthening Individual Estimates

- ▶ At the **upper level**, we assume individuals are distributed in some specified way, such as multivariate normal, with means and covariances to be estimated.
- ▶ At the **lower level**, we assume each individual's answers conform to a separate model, such as logit or OLS regression.
- ▶ Hierarchical Bayes determines the optimal degree to which the upper-level model and the lower-level model influence the parameters for each individual.
- ▶ Lower Level model dominates if...
 - There's a lot of information per respondent
 - Respondent is consistent

Hierarchical Bayes: Benefits

- ▶ Individual level utility models
 - Dramatically improves modeling, improved understanding of the market (heterogeneity)
 - Much greater flexibility in analysis (post-process by filters, cross-tabs, etc. without having to re-run the model each time for segments)
 - Need for explicitly modeling higher-order effects (interactions, substitutions, irrelevant alternatives, etc.)

Hierarchical Bayes Demo



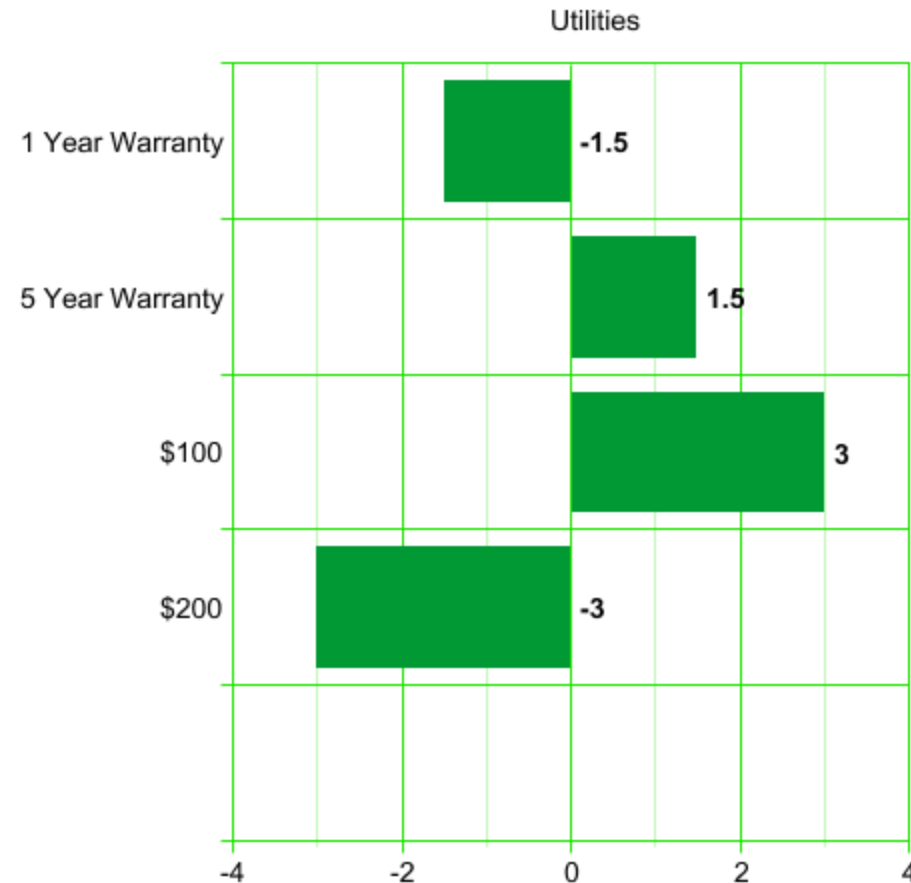
Special Features

- ▶ Constraints force parameters to conform to external rule, i.e. lower prices should always be preferred over higher prices
 - May mask important findings such as price being an indicator of quality (bicycle locks, laser eye surgery, etc.)

- ▶ Interaction effects allow for adjustments in utility for specific combinations of levels
 - Price attribute is perceived differently depending on brand, high resolution is more valued on bigger TV sizes

Utilities

- ▶ “Utilities” are developed for each level to indicate relative preference
 - Utilities are zero-centered (positive does not indicate “good” and negative does not indicate “bad”)
 - Utilities are interval scaled (zero point is arbitrary, no ratio operations)
 - You *CANNOT* compare one level from one attribute with one level from another attribute.



Attribute Importance

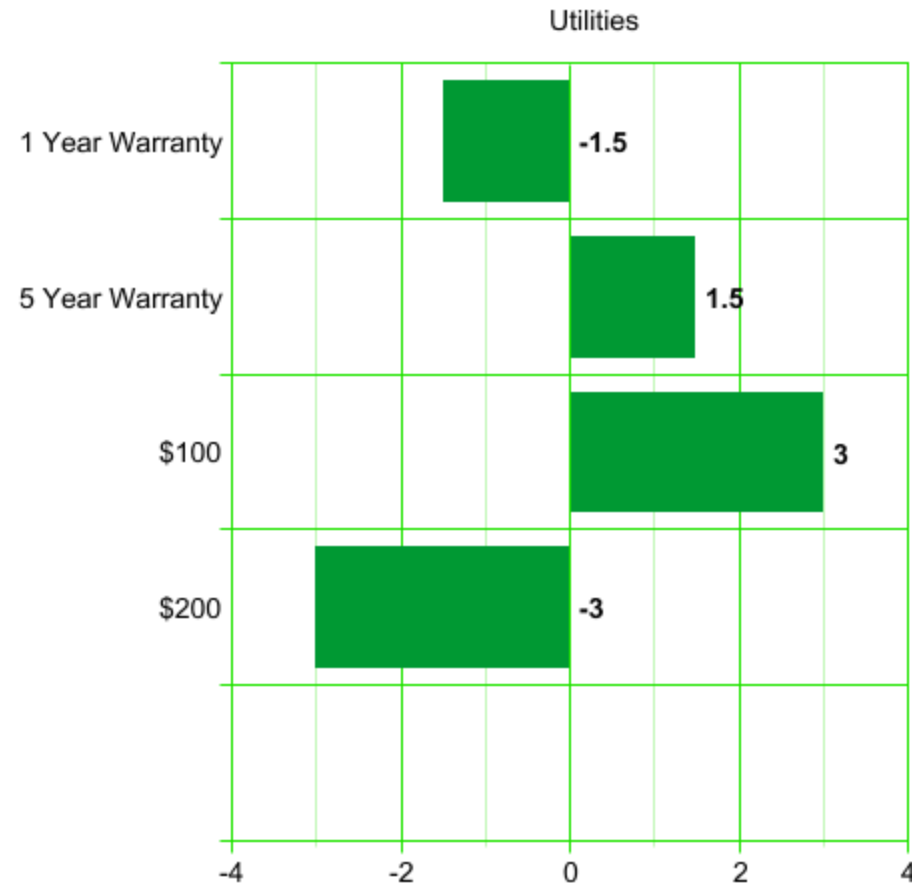
- Quantifies how much potential influence each attribute has in the model

- Highest level – lowest level divided by total range

Warranty: $1.5 - (-1.5) = 3 / 9 = 33.3\%$

Price: $3 - (-3) = 6 / 9 = 66.7\%$

Total Range: 9



MARKET SIMULATIONS

Why Conduct Market Simulations?

- ▶ Utilities and importance scores only go so far (and have baggage!)
- ▶ Average utilities can be misleading
- ▶ A “What if?” laboratory for testing multitude of real-world possibilities
- ▶ Results expressed in terms that make sense to management and are actionable

Market Simulator Demo

Choice Simulator - IntroCBCwLHS2

Home Scenario 1 Settings

Utility Sets: HB (CBC)

Attribute Info Conditional Pricing (display) Utility Report Create Client Simulator Package Precision

Prohibitions Revenues & Costs Create Excel Report Create PowerPoint Report Create Excel Simulator Create Excel Simulator Currency

SCENARIOS PROJECT INFORMATION REPORTING CLIENT SIMULATIONS OPTIONS

Scenario 1 x Sensitivity x Search x Multi search x

Products Simulate

Label	Brand	Style	Color	Free Delivery	Counter Depth	Ice Maker	Price
Samsung	Samsung	Side by Side	Stainless Steel	No	No	No	1499
LG	LG	French Door	Stainless Steel	Yes	Yes	Yes	\$2,499
SubZero	SubZero	Bottom Freezer	Black Stainless Steel	Yes	No	Yes	\$3,499
GE	GE	French Door	Black Stainless Steel	Yes	Yes	Yes	\$3,999

Results

Simulation Results

Label	Shares of Preference	Revenue
Samsung	11.4 %	\$ 8,526,752.6
LG	52.1 %	\$ 65,129,790.7
SubZero	10.2 %	\$ 17,928,528.6
GE	1.0 %	\$ 2,013,176.5
None	25.2 %	N/A

Standard Errors

Label	Shares of Preference	Revenue
Samsung	1.5 %	\$ 1,127,774.3
LG	2.6 %	\$ 3,305,920.8
SubZero	1.6 %	\$ 2,743,009.5
GE	0.4 %	\$ 773,310.2
None	2.2 %	N/A

Product Specifications

Label	Brand	Style	Color	Free Delivery	Counter Depth	Ice Maker	Price
Samsung	Samsung	Side by Side	Stainless No	No	No	No	1499
LG	LG	French Door	Stainless Yes	Yes	Yes	Yes	\$2,499
SubZero	SubZero	Bottom Freezer	Black St Yes	No	Yes	Yes	\$3,499
GE	GE	French Door	Black St Yes	Yes	Yes	Yes	\$3,999

Shares of Preference

Simulation Results

Standard Errors

Product Specifications

Market Simulation

	Resp 1	Resp 2	...	Resp 500
Samsung	2.5	-1.0		3.7
KitchenAid	1.8	1.0		0.5
Side by Side	5.3	1.2		1.0
Bottom Freezer	3.2	0.7		0.8
KitchenAid Side by Side	7.1	2.2		1.5
Samsung Bottom Freezer	5.7	-0.3		4.5
Winner	KitchenAid	KitchenAid		Samsung

“First Choice” Market Simulations

- ▶ For each respondent, assume respondent chooses the product with the highest utility
- ▶ Accumulate (aggregate) respondent choices and call them “Shares of Choice” or “Shares of Preference” (some refer to them as “Market Shares”)

First Choice: How Realistic?

- ▶ First choice model is simple to do and easy to understand, but is usually an overly simple, extreme model of consumer behavior
 - Assumes a product barely preferred over another is chosen 100% of the time
- ▶ Less efficient use of data: we learn about which product is preferred, but don't capture anything about relative preferences of not preferred options

Randomized First Choice (RFC)

- ▶ Respondents go on “shopping trips” many times
- ▶ Special error is added to the utilities each shopping trip
 - Options close in utility will “win” relatively evenly
 - Options with much more utility will “win” much more often
- ▶ Results of all the shopping trips are aggregated to produce overall shares of preference

Simulator Options

- ▶ **Sawtooth Software offers two off-the-shelf options:**
 - Lighthouse Studio: Integrated into Sawtooth Software's base product
 - Online Simulator: web-based simulator

- ▶ **Build-Your-Own in Excel**
 - Export a started excel simulator from Lighthouse Studio!

Option #1: Lighthouse Studio Simulator

- ▶ Fully integrated into Lighthouse Studio
- ▶ All data flow directly from fielding through utility estimation into simulator
- ▶ Standard simulations, but also store-based simulations, sensitivity analysis, auto-tuning shares, etc.
- ▶ Can be distributed to your clients
 - They need to install software, though!
- ▶ Can also export an Excel simulator with the RFC plug-in
 - RFC plug-in only available on Windows machines, use SOP if your client has a Mac!

Excel Simulator Example

Create an Excel Simulator
from LHS Choice Simulator!

The screenshot displays the Choice Simulator software interface. A purple arrow points to the 'Create Excel Simulator' button in the top right corner. The interface includes a 'Multi search Settings' panel on the left, a 'Products' table, a 'Results' table, and a 'Simulation Results' summary table. An Excel spreadsheet is also visible in the background.

Products Table:

Label	Brand	Style	Color
Samsung	Samsung	Side by Side	Stainless Steel
LG	LG	French Door	Stainless Steel
SubZero	SubZero	Bottom Freezer	Black Stainless Steel
GE	GE	French Door	Black Stainless Steel
Samsung New	Samsung	1-4	1-4

Simulation Results Summary Table:

Domination Rank	Result Index	Share of Preference
1	281	74.4
2	464	58.8
3	451	47.3
4	295	70.8
5	450	55.9
6	465	44.1
7	463	69.2
8	492	54.3
9	239	67.5
10	520	52.7
11	507	42.4
12	225	67.4
13	506	52.4
14	449	66.6
15	493	41.0
16	323	66.4
17	282	48.4
18	283	40.9
19	519	65.8
20	310	48.1

Excel Spreadsheet Data:

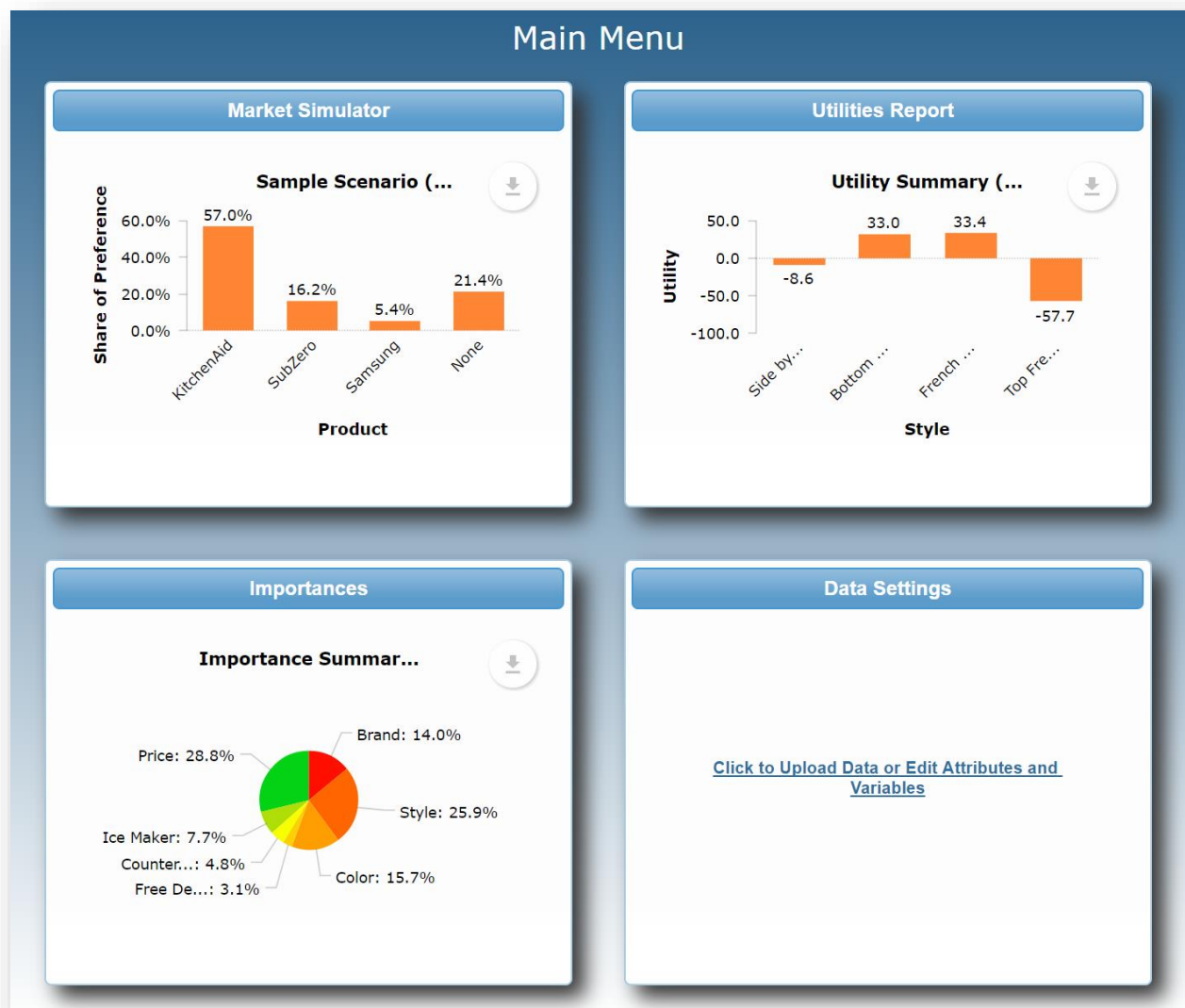
	Product 1	Product 2	Product 3	Product 4	Product 5	Product 6	None
1	1	1	1	1	1	1	1
2	1.00	2.00	3.00	4.00	5.00	6.00	
3	2.00	3.00	3.00	4.00	1.00	2.00	
4	2.00	1.00	3.00	4.00	2.00	2.00	
5	1.00	2.00	1.00	1.00	2.00	1.00	
6	1.00	2.00	2.00	2.00	1.00	2.00	
7	2.00	2.00	1.00	1.00	1.00	1.00	
8	999.00	2,499.00	1,499.00	1,499.00	999.00	2,999.00	
9	27.59 %	5.87 %	22.47 %	2.51 %	25.40 %	11.51 %	4.64 %
10	2.45 %	1.16 %	2.24 %	0.81 %	2.25 %	1.57 %	1.03 %
11	1.00						

Option #2: Online Simulator

► Web-based “Software As A Service” Tool

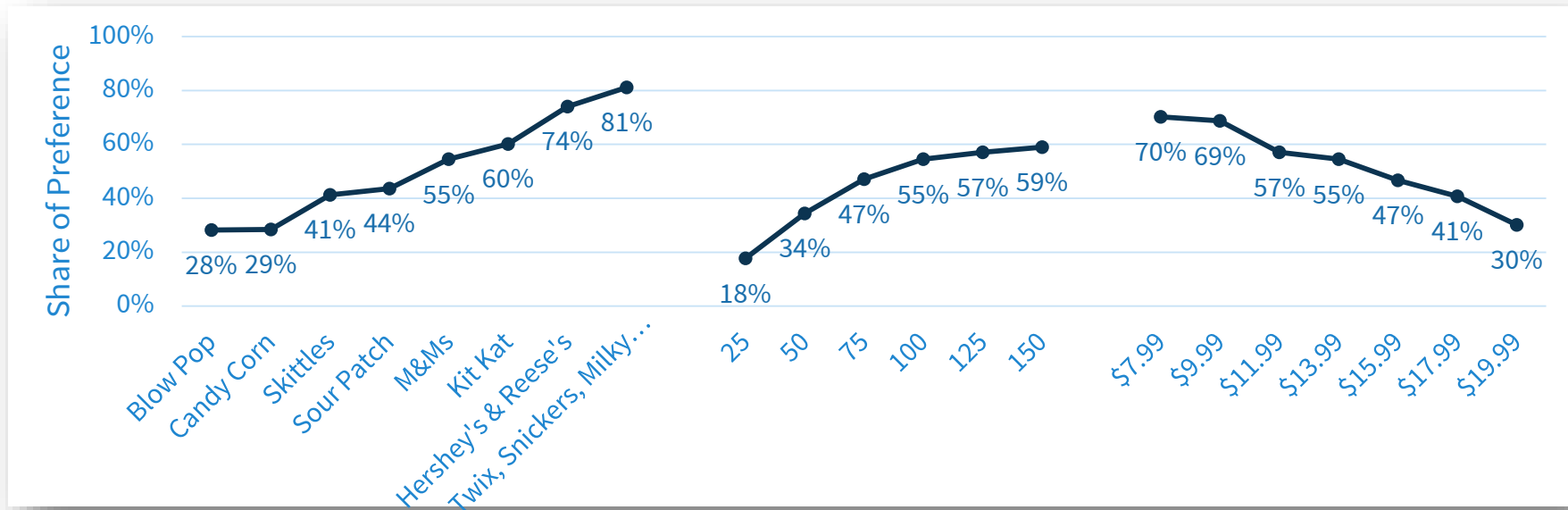
- Designed for simplicity, speed, ease of use
 - Point-and-click interface designed to make your client’s experience as simple as possible
 - Manage all simulations/analysis in one place, while allowing you to “publish” specific projects to individual colleagues/clients
 - Automatic charts and graphs of simulation results
 - Toggle products in & out of the simulation
 - Control which features each account/user may access, and lock down settings like simulation method, exponent, etc.
- But...
 - Not as full-featured as other simulators (no Advanced Simulations, no sensitivity analysis)
 - Requires internet connection

Online Simulator Example



Sensitivity Analysis

- ▶ Sensitivity analysis allows us to observe the change in share of preference due to changing product specifications.
 - Here we compare the base scenario - M&Ms, 100 count, at \$13.99 - to the “None” option. Then we vary brand, holding count and price constant. Then we vary count, holding brand and price constant. And then we vary price, holding brand and count constant.



Price Sensitivity Example

- ▶ Holding all other brands at 100 count and \$13.99, we can vary the price of M&Ms and see how sensitive consumers are.



- ▶ M&Ms has a significant loss in share when increasing price above \$9.99. There is less sensitivity above \$13.99, but very low share.

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
<u>CBC (Choice-Based Conjoint)</u>	10 attributes, 15 levels per attribute	\$5,500
<u>CBC ADM (CBC + Advanced Design Module)</u>	250 attributes	\$7,000
<u>Choice Simulator</u> Free simulator for CBC, ACBC, ACA, or CVA Lighthouse Studio subscriptions. Also creates free client simulators.		Free

Suite Subscriptions

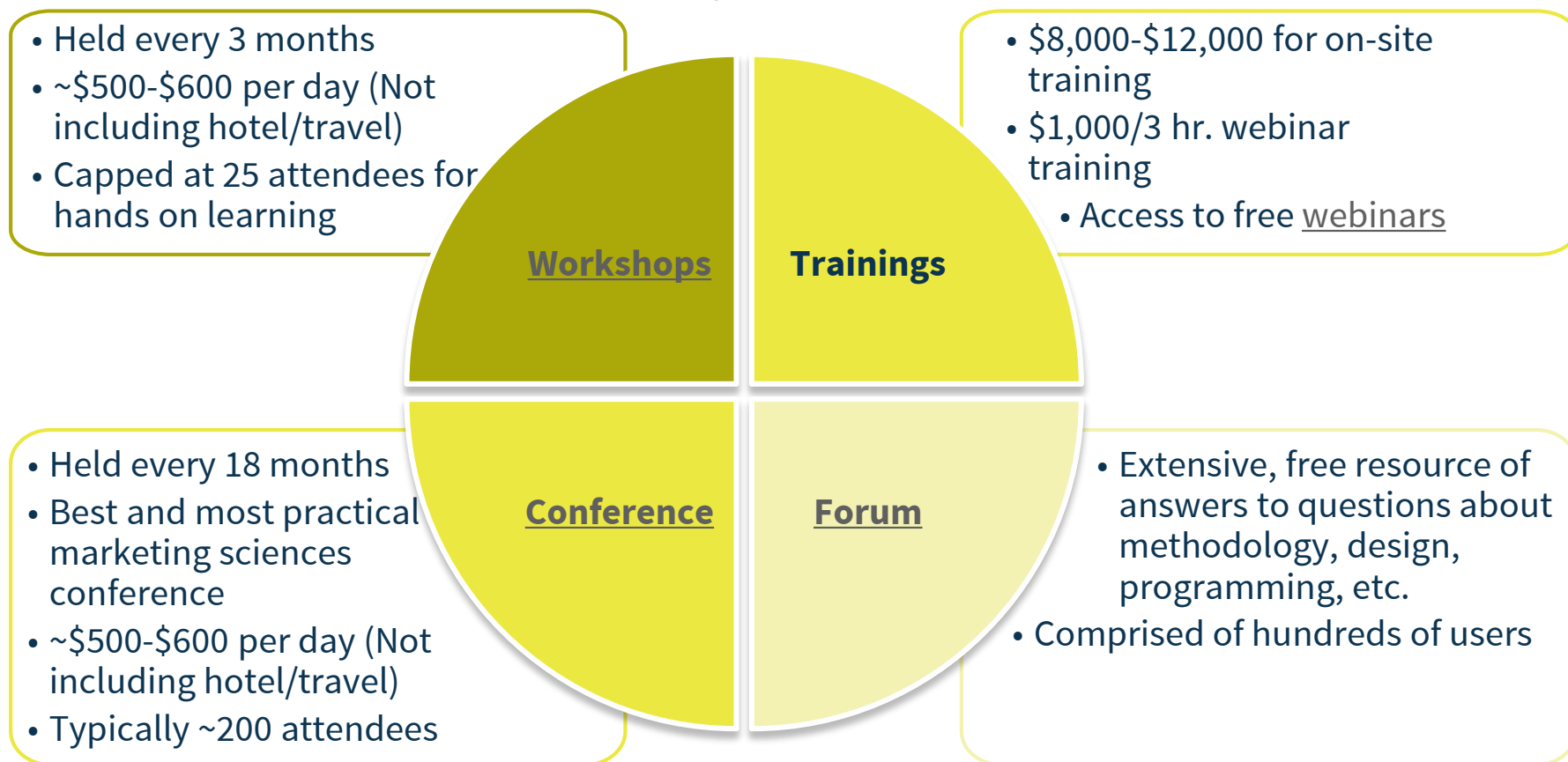
- Should you intend to execute more than one methodology, the bundled Suite Subscription offers incredible savings.

Suite Licenses	1 Person	5 Person	10 Person
<u>Premium Suite Plus</u> Includes Premium Suite, five hours of free consulting, and 15% off workshops and conferences	\$15,500	\$25,100	\$31,900
<u>Premium Suite</u> Includes Standard Suite, MBC, and unlimited CCS desktop client conjoint simulators	\$13,900	\$23,500	\$29,900
<u>Standard Suite</u> 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.



QUESTIONS OR FREE 30 DAY TRIAL



Megan Peitz
Ingenuity Ambassador
megan@sawtoothsoftware.com



Brian McEwan
Director, Client Services
brian@sawtoothsoftware.com

www.sawtoothsoftware.com

+1 801 477 4700



@sawtoothsoft