An Introduction to Choice-Based Conjoint

with Sawtooth Software's Lighthouse Studio

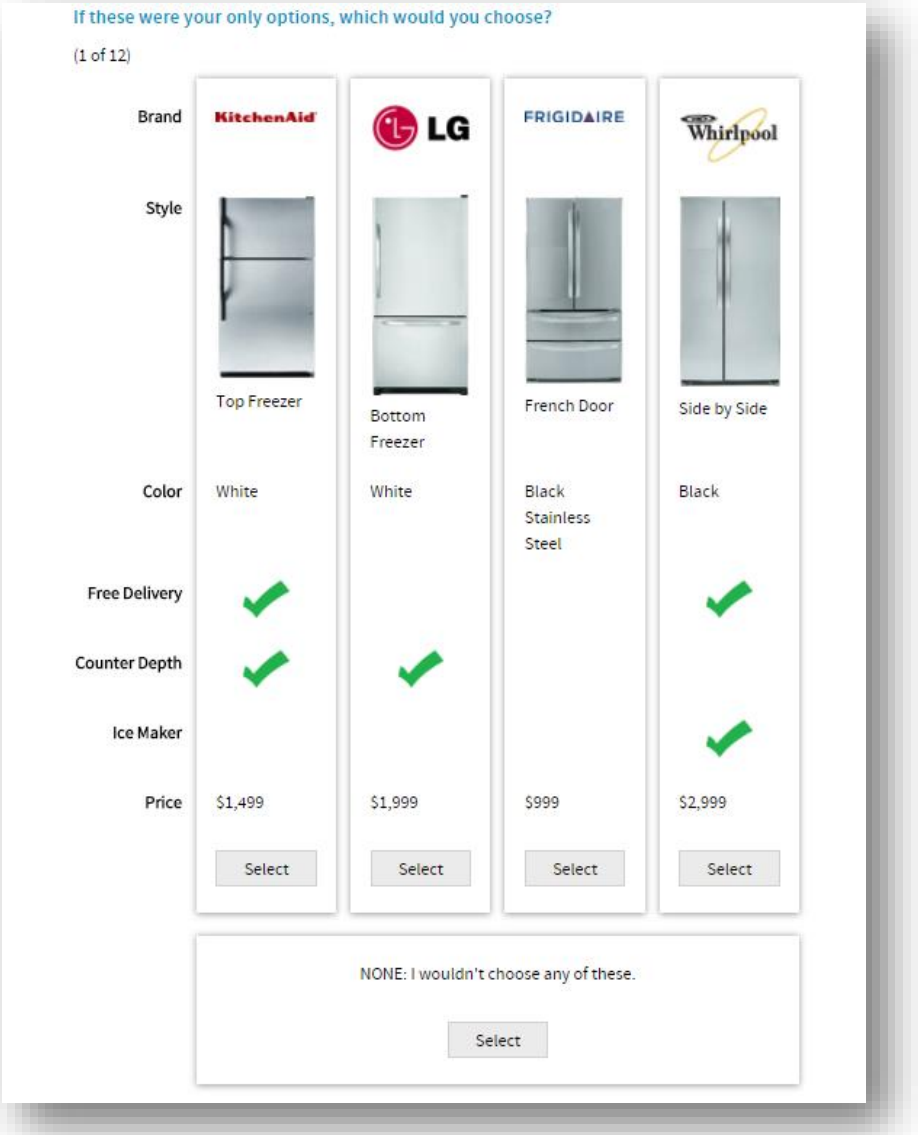
Part 2

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



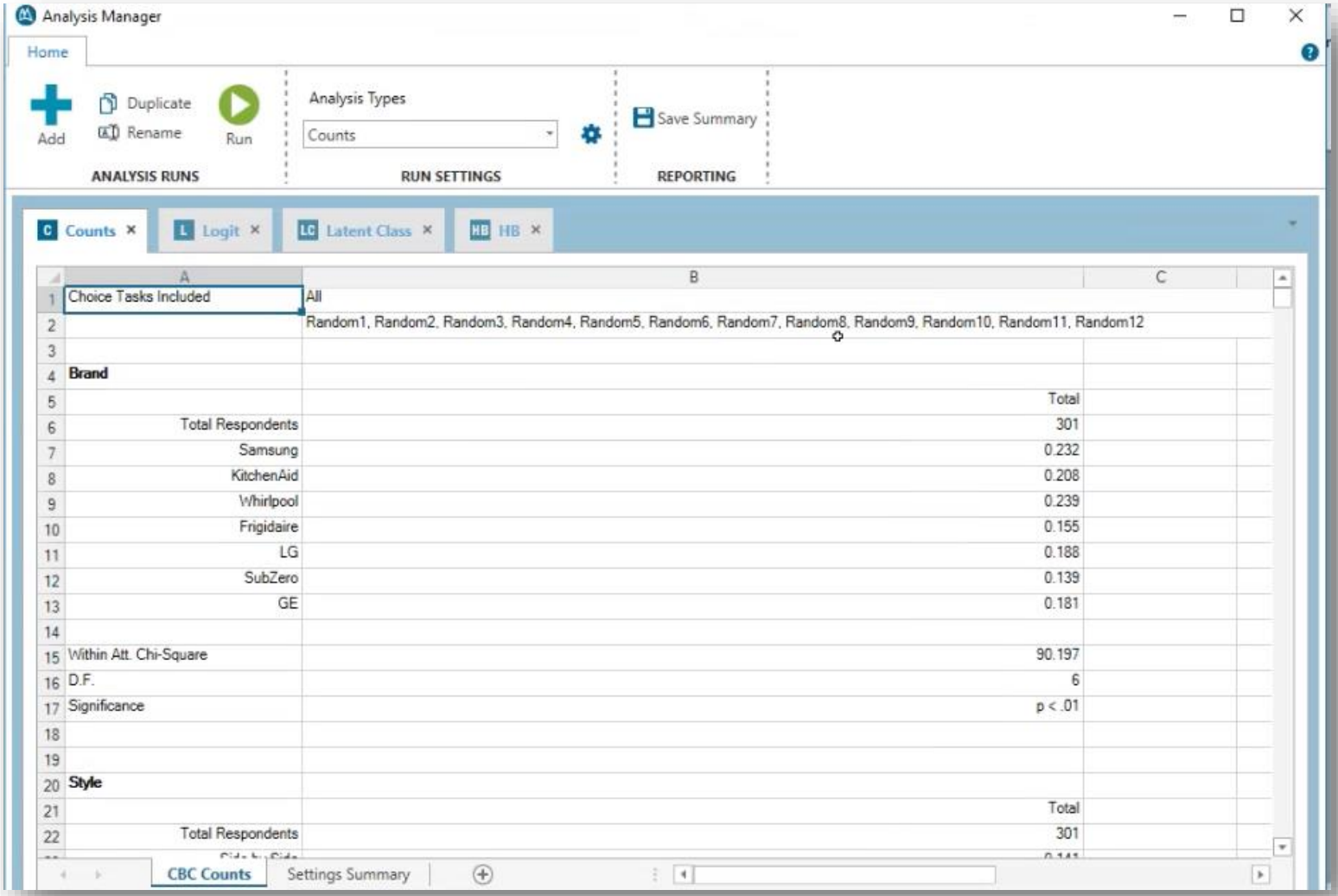
ANALYZING CBC DATA

Three Conjoint Analysis Outputs

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



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

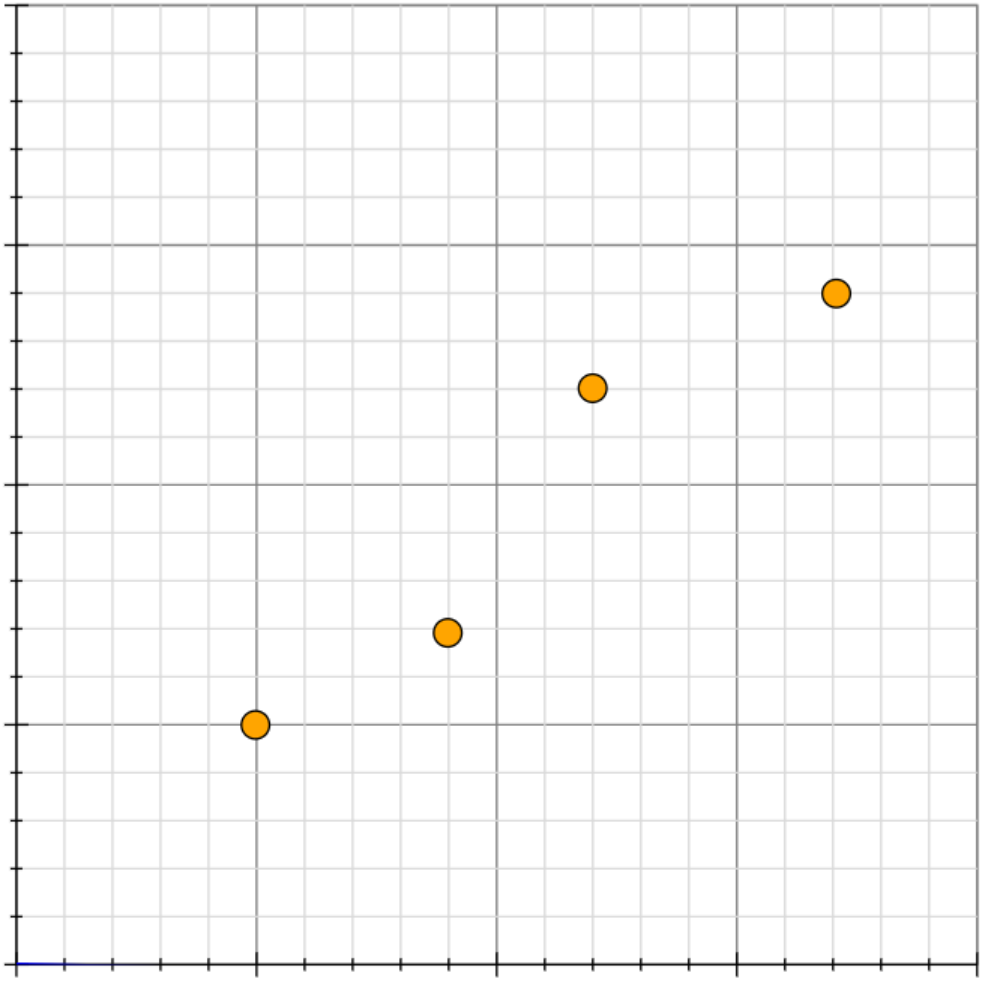
- Counts will only follow rational expectations if your design is not confounded or correlated!

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

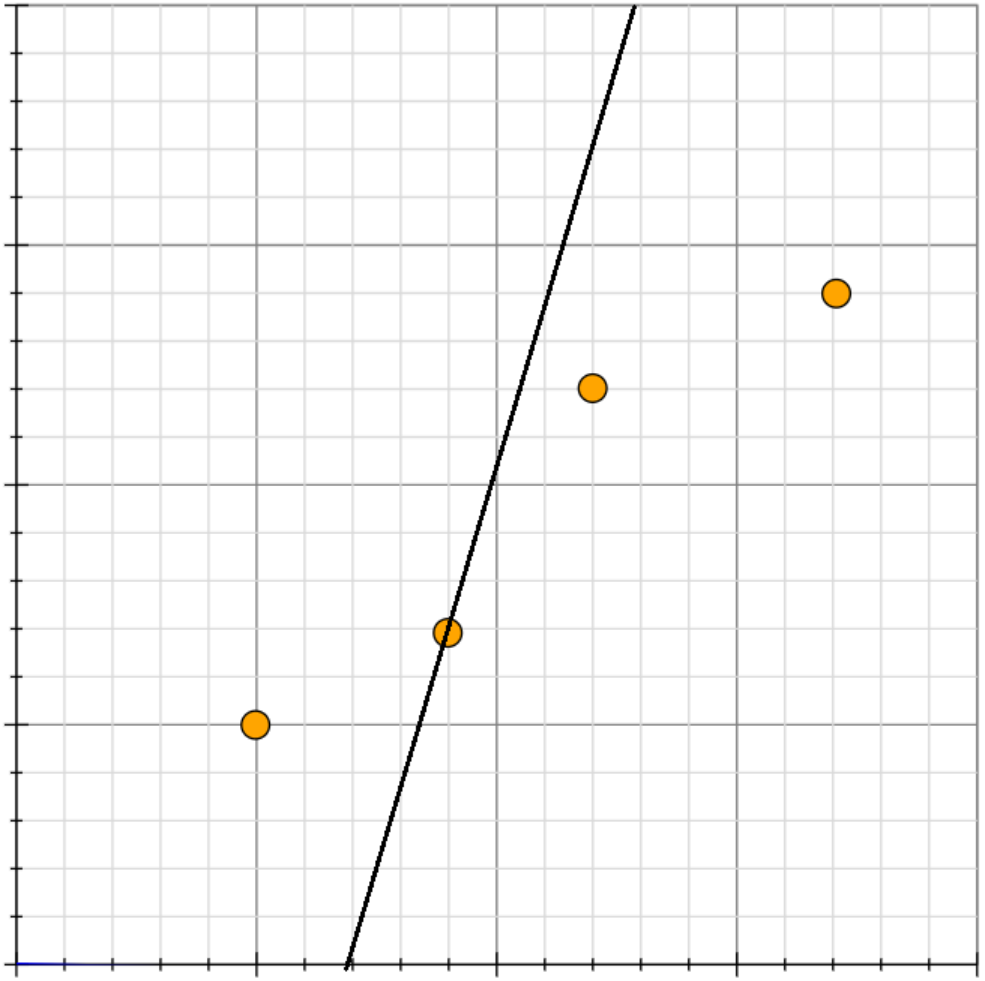
Counting Analysis

- 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

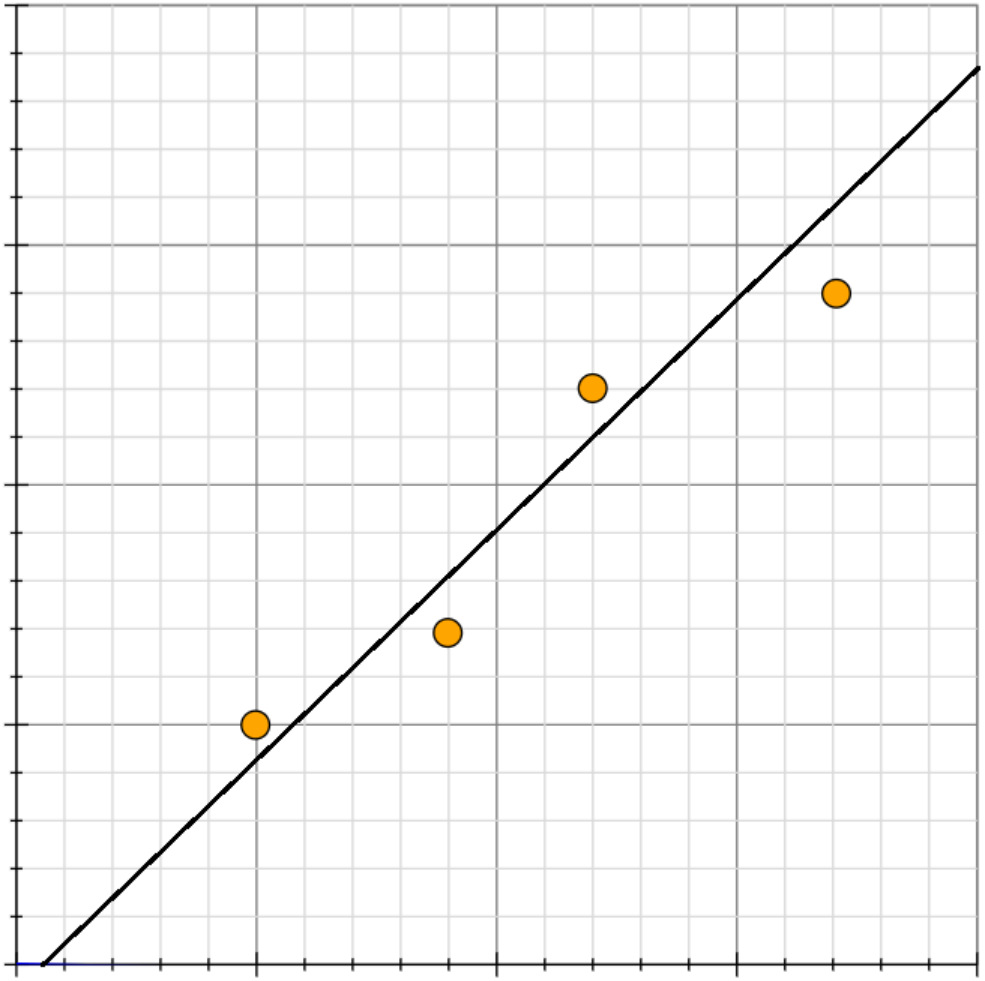
Modelling Data



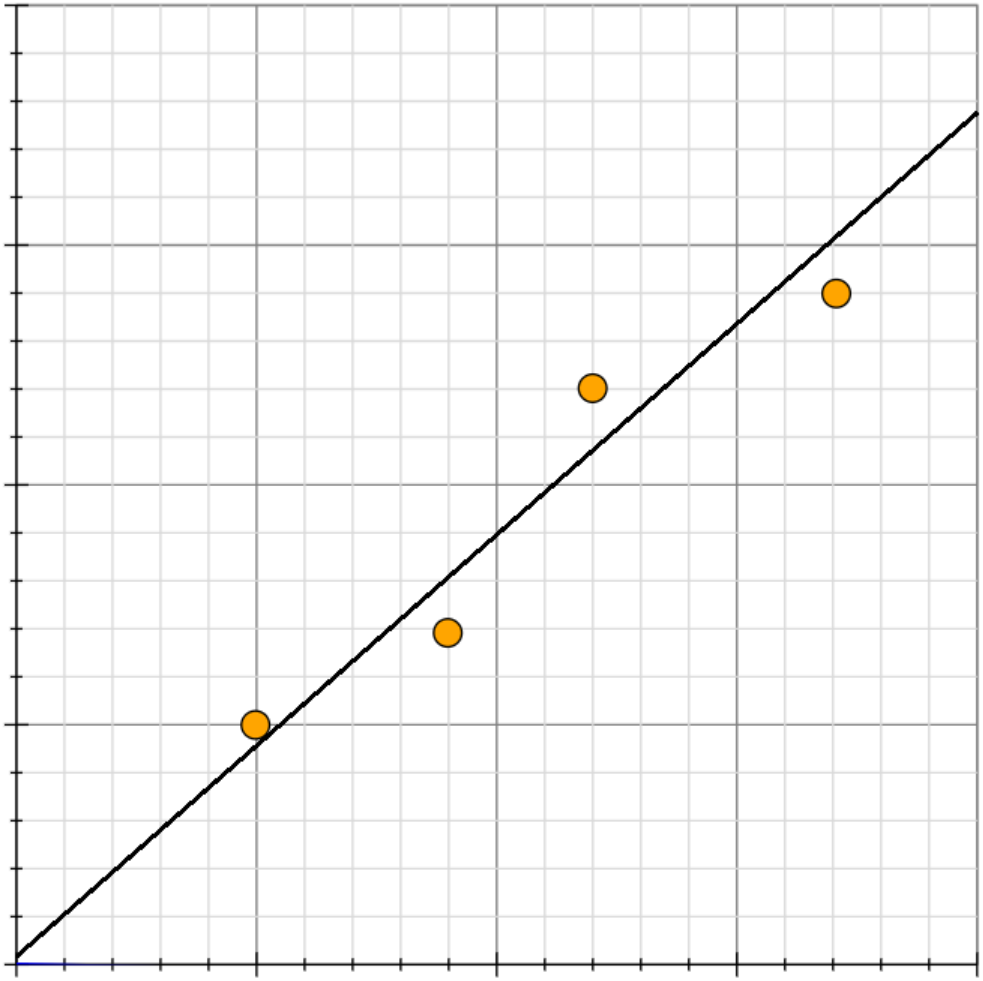
Modelling Data – Linear Regression



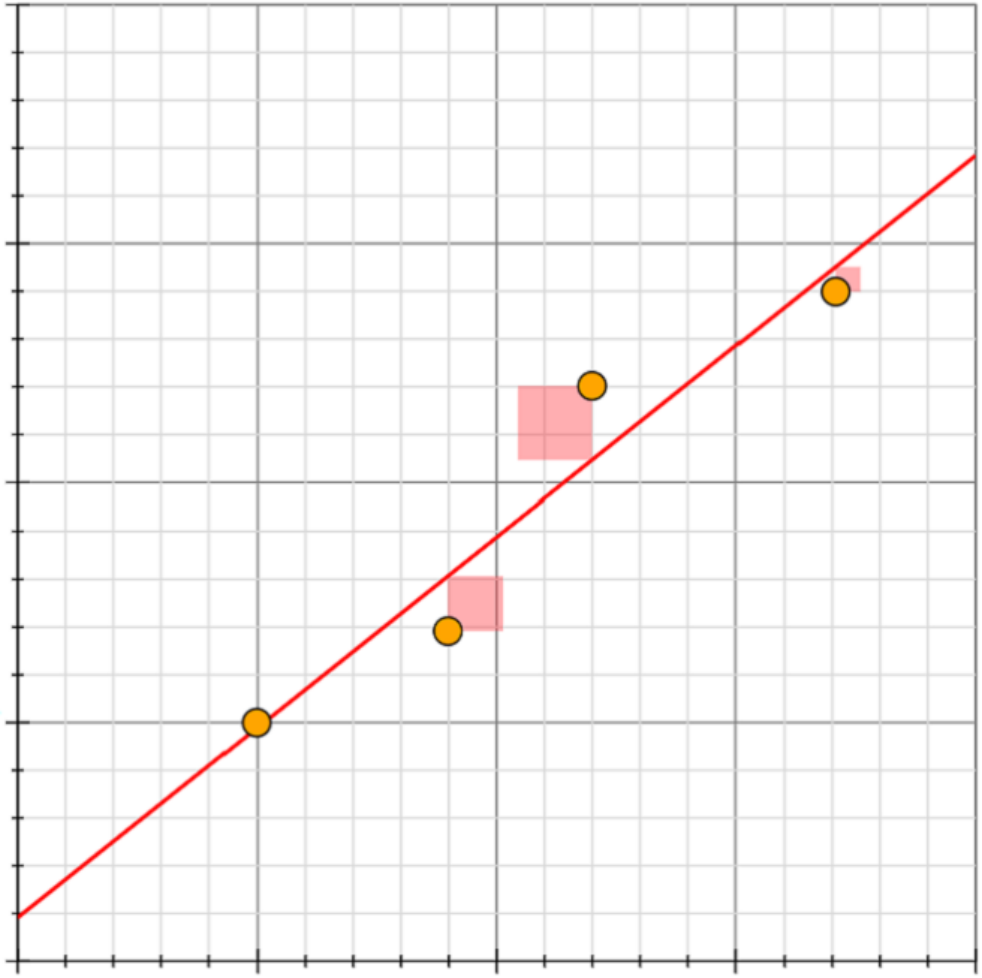
Modelling Data – Linear Regression



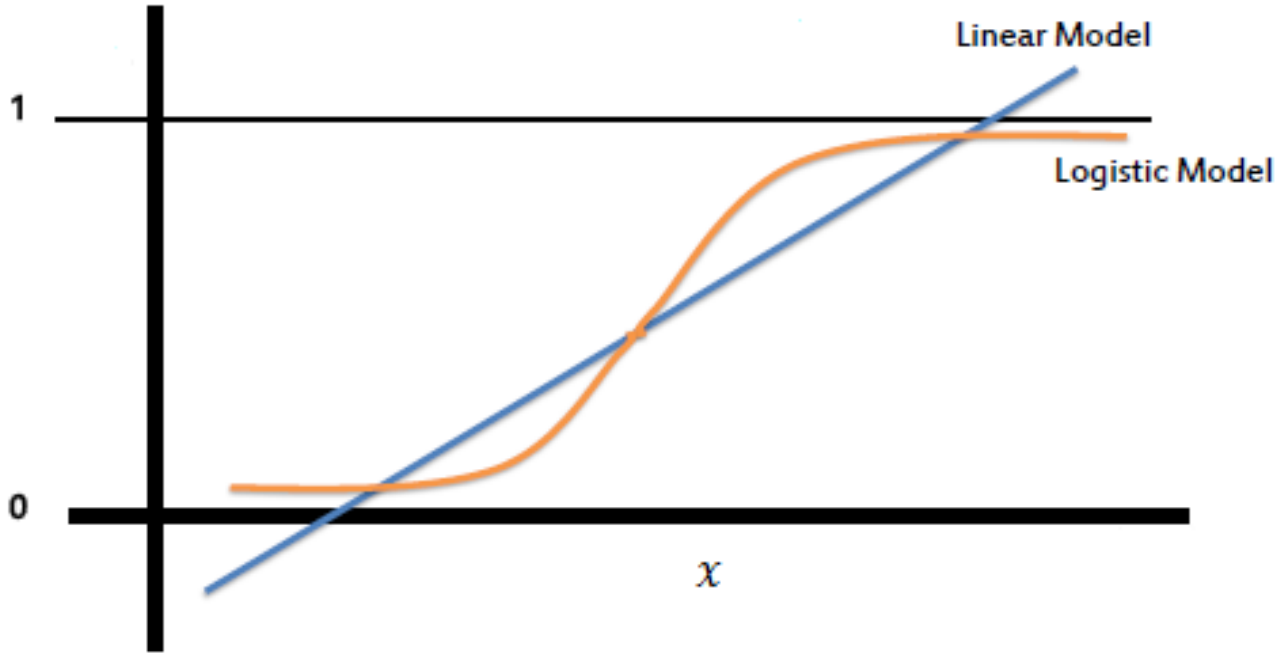
Modelling Data – Linear Regression



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

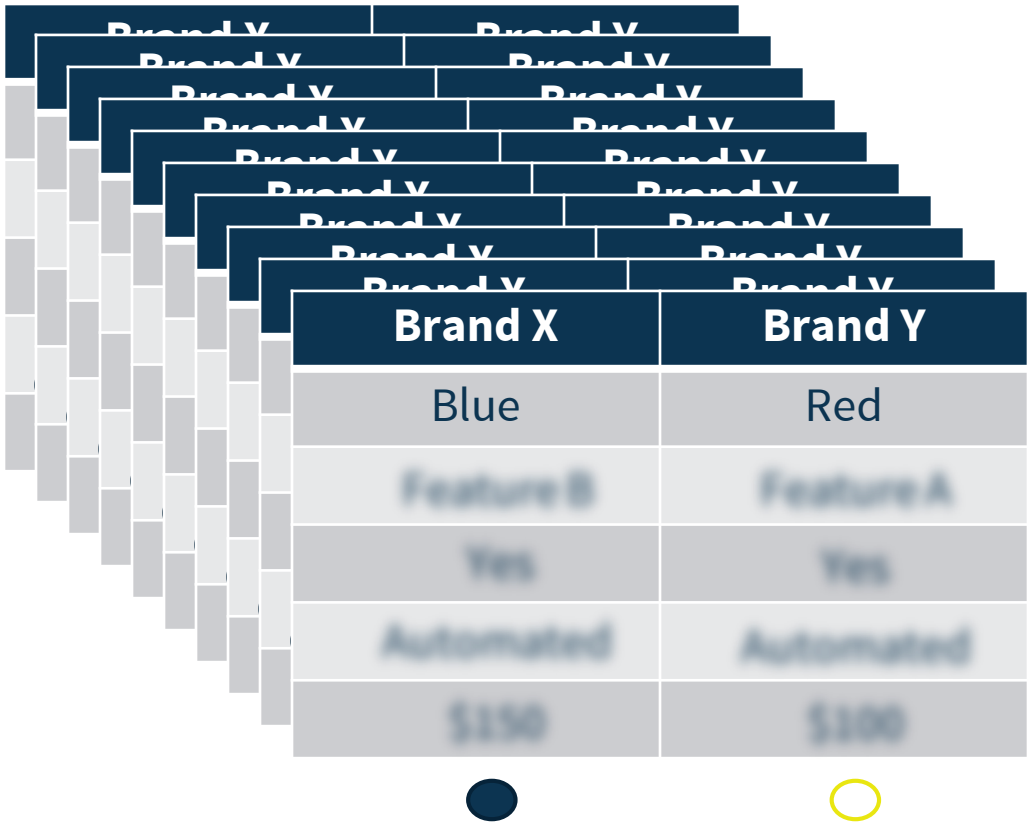
X	1
Y	-1
Red	0
Blue	0

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

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

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

Logistic Regression



X	?
Y	?
Red	?
Blue	?

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

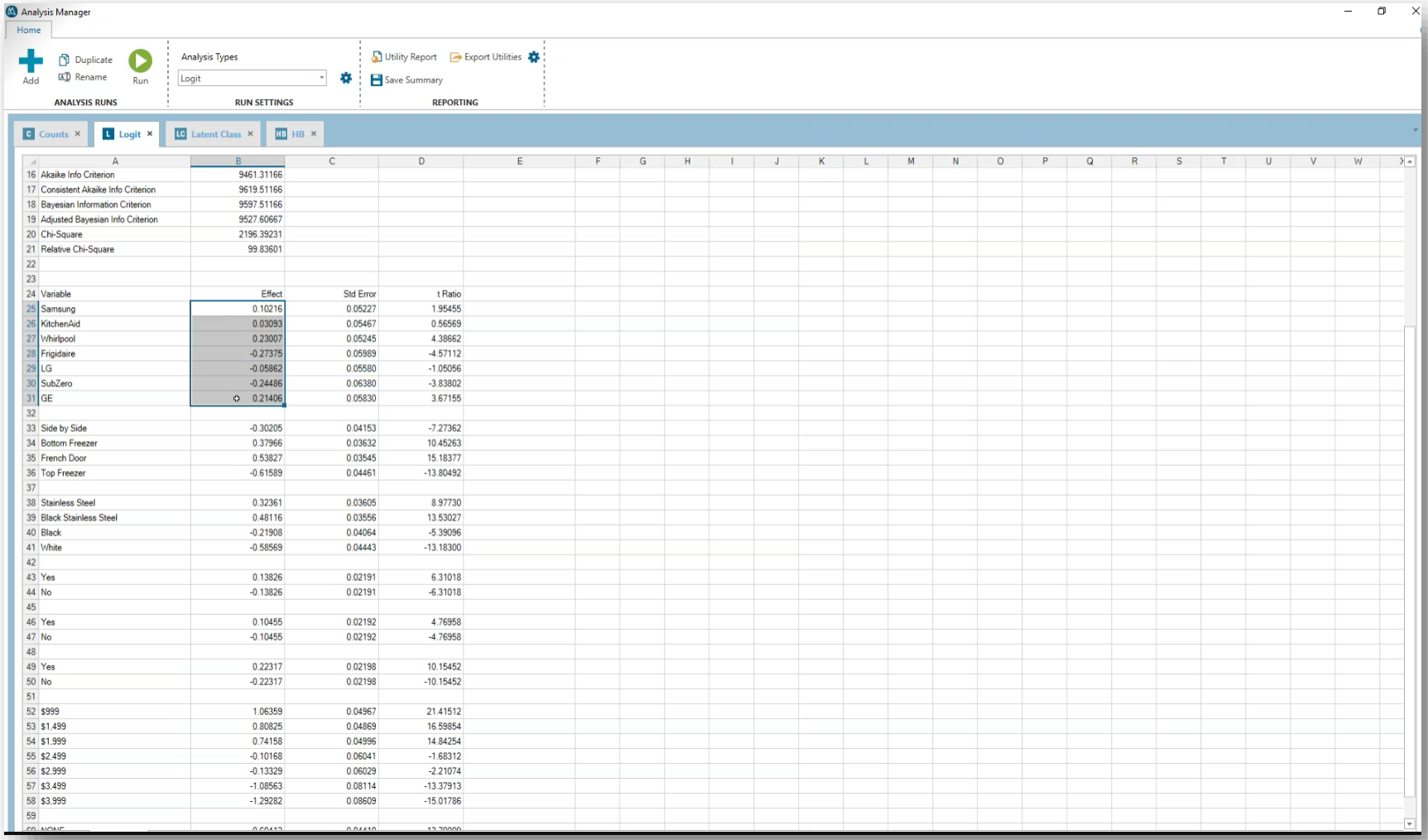
Aggregate Logit

- 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



	A	B	C	D
16	Akaike Info Criterion	9461.31166		
17	Consistent Akaike 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	-0.24486	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				

Latent Class

Latent

adjective la·tent \ ˈlā-tᵊnt \

existing but not yet developed or manifest; hidden; concealed

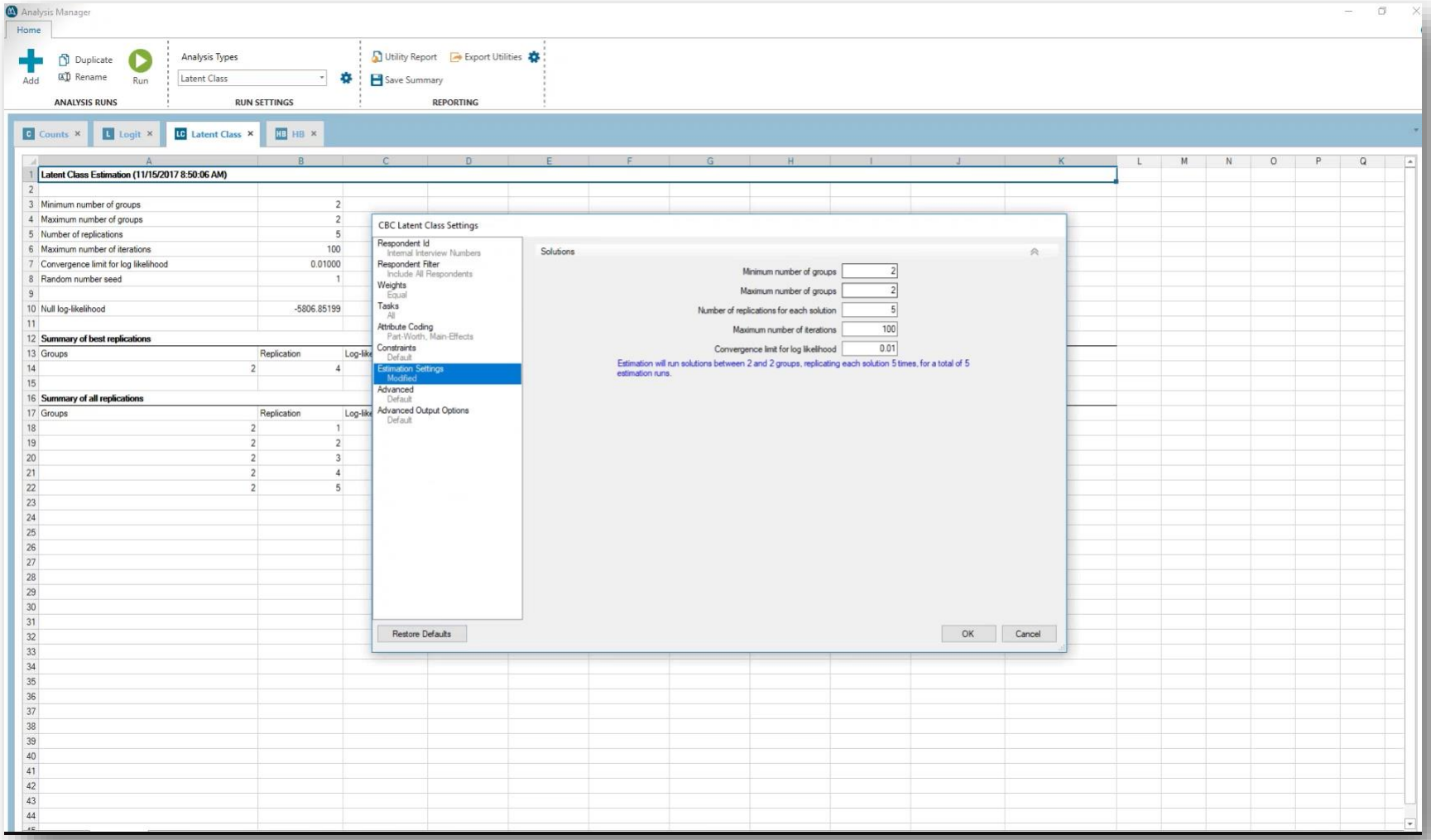
Latent Class

- 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



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

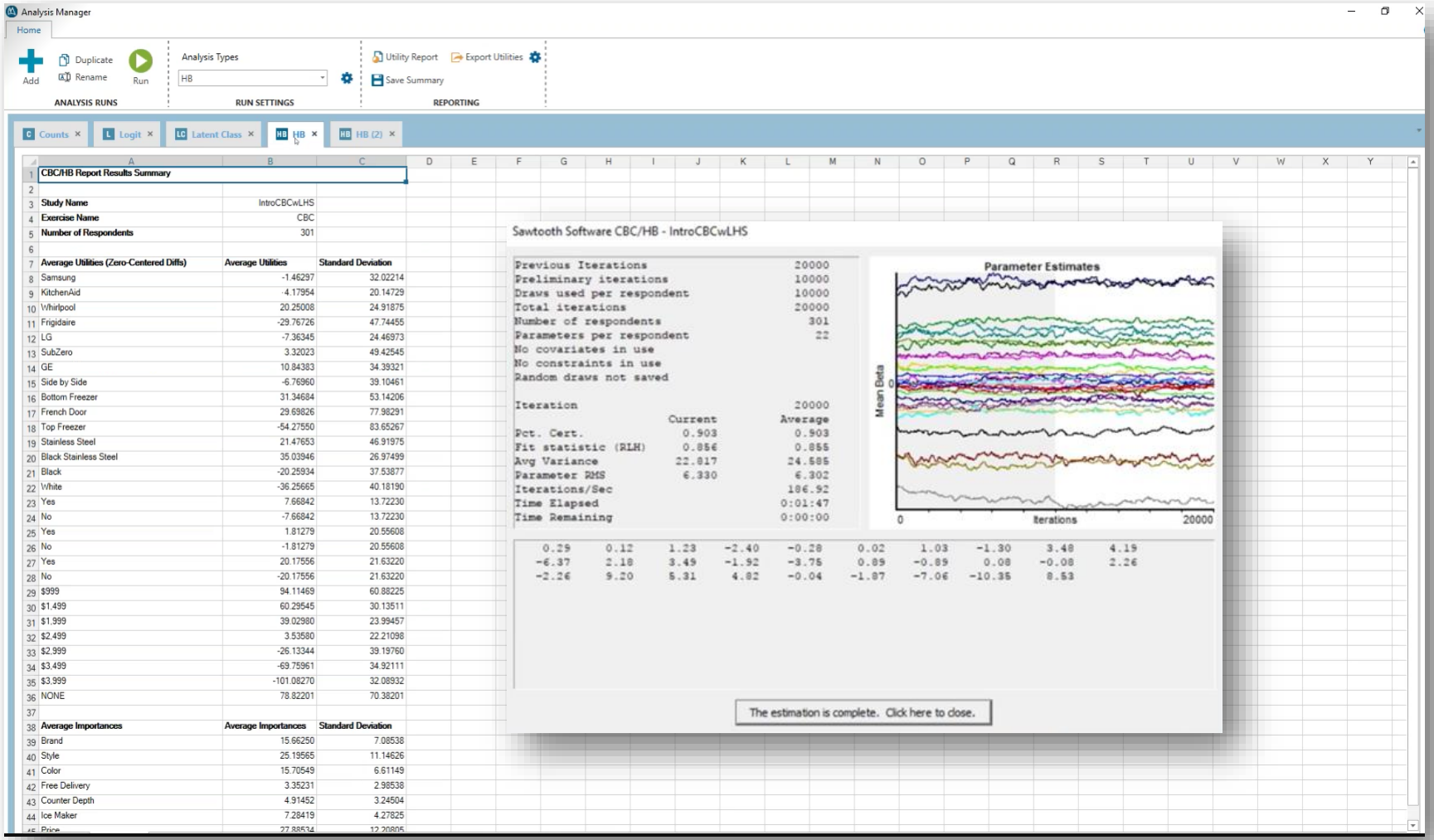
Strengthening Individual Estimates

- 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

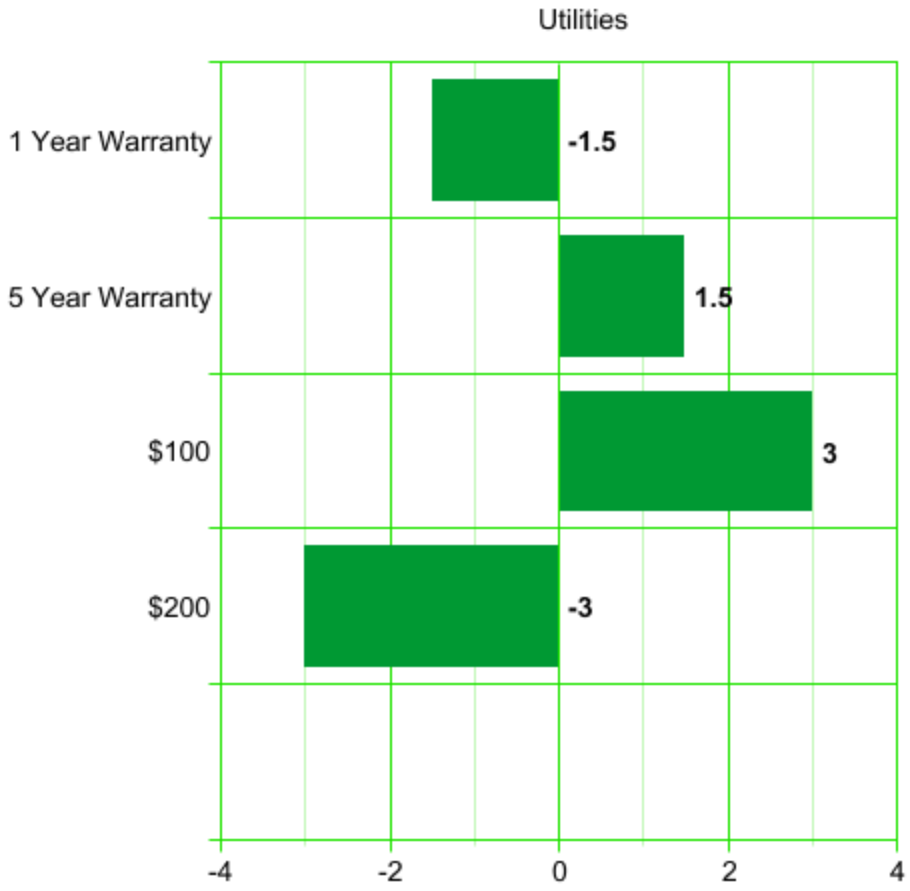


Special Features

- 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.



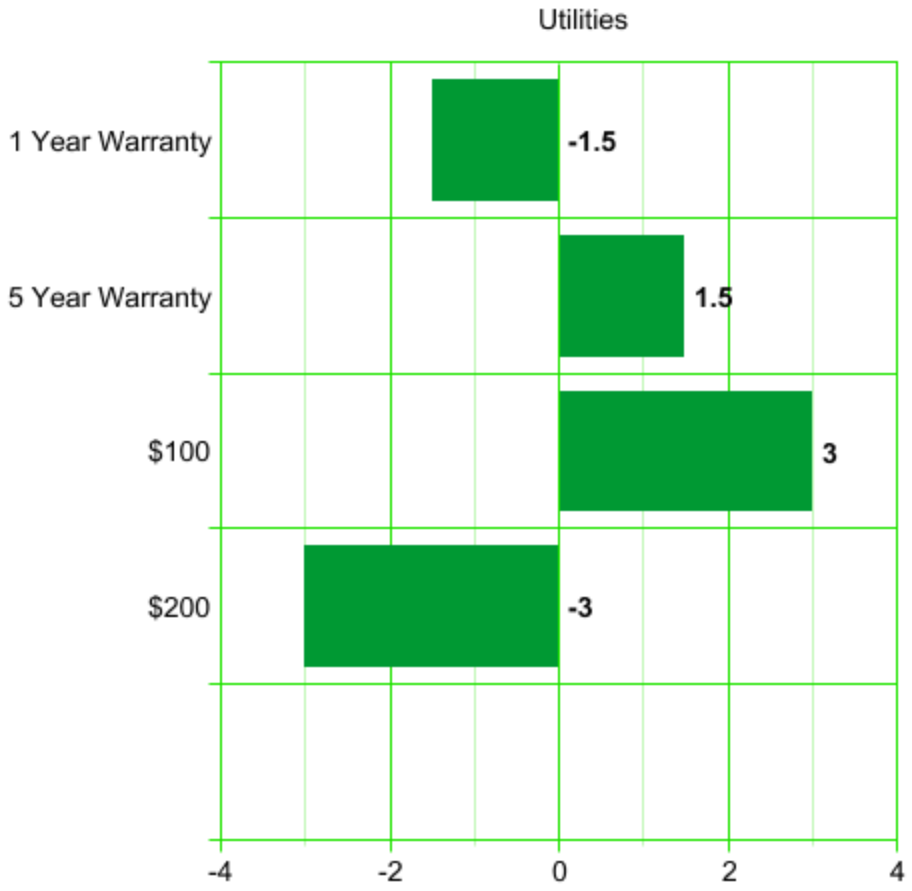
Attribute Importance

- 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

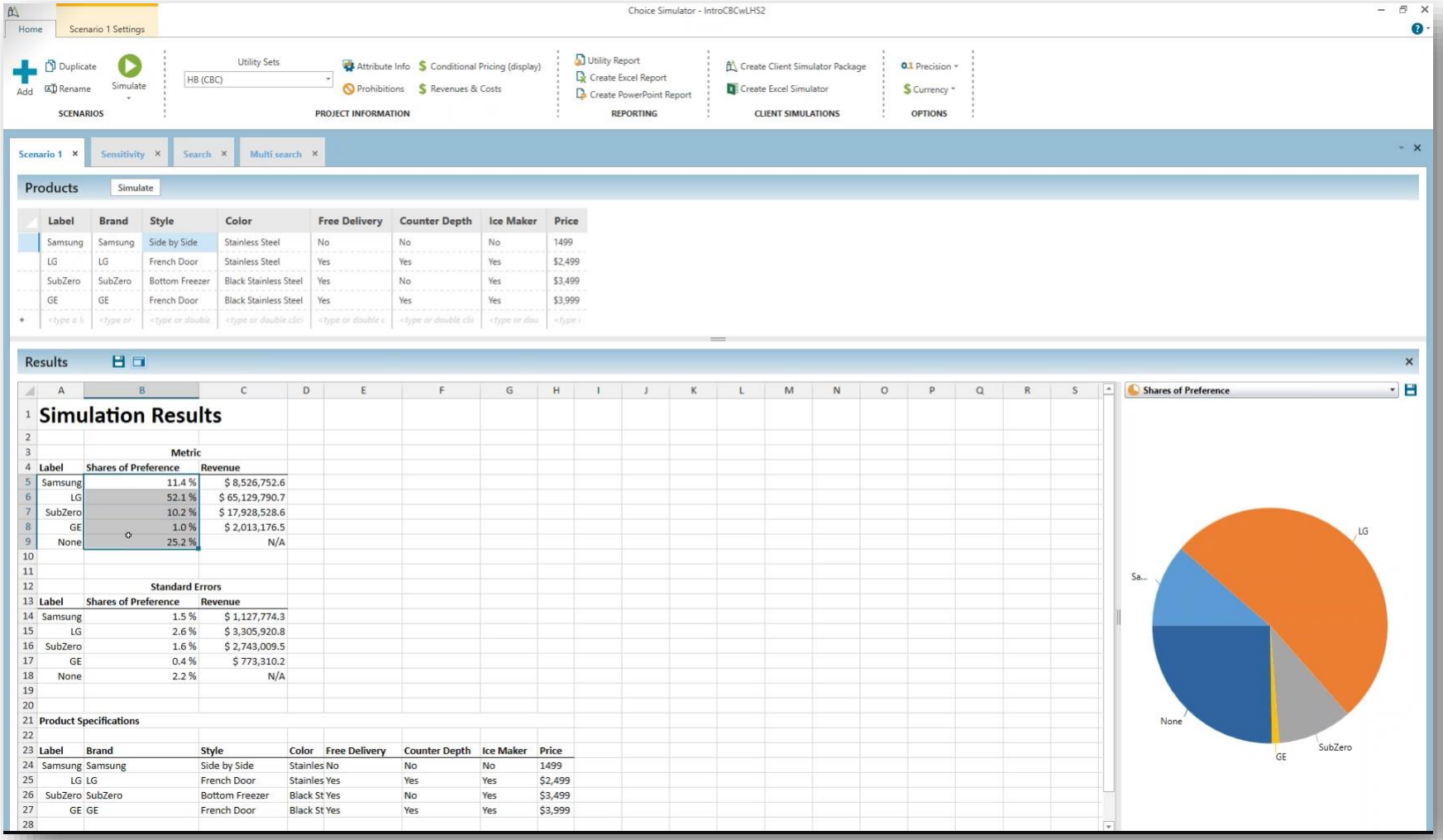


MARKET SIMULATIONS

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



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

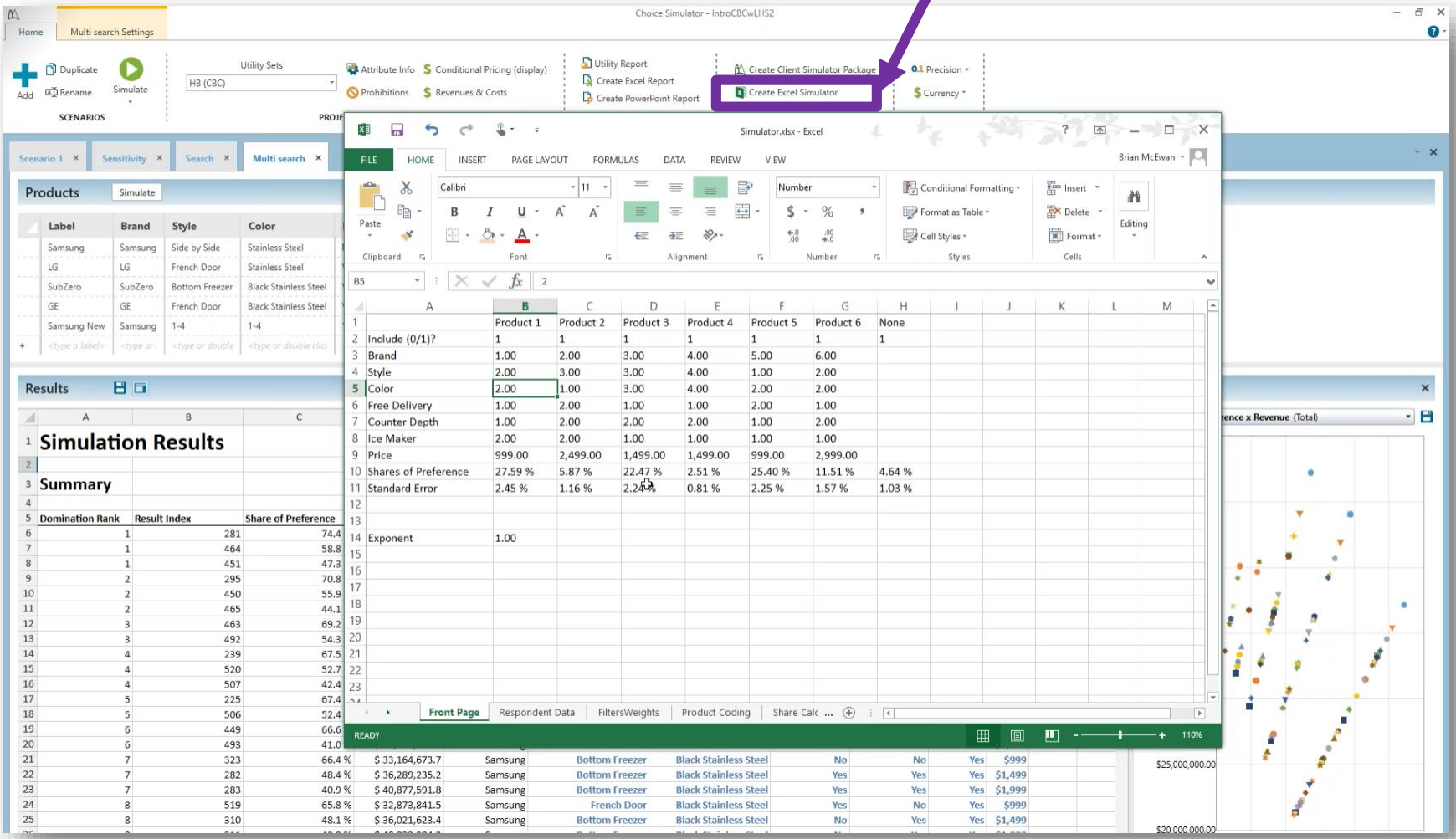
- Sawtooth Software offers two off-the-shelf options:
  - Lighthouse Studio: Integrated into Sawtooth Software's base product
  - Online Simulator: web-based simulator
- Build-Your-Own in Excel
  - Export a started excel simulator from Lighthouse Studio!

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!



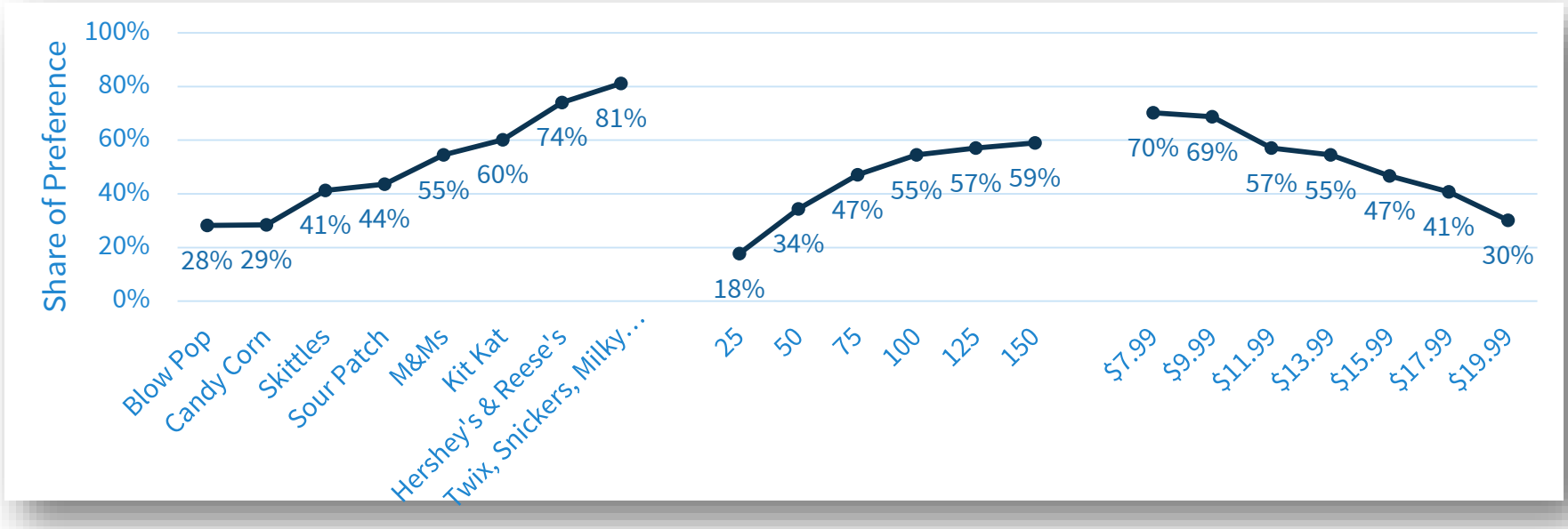
Option #2: Online Simulator

- 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
CBC (Choice-Based Conjoint)	10 attributes, 15 levels per attribute	$5,500
CBC ADM (CBC + Advanced Design Module)	250 attributes	$7,000
Choice Simulator 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
Premium Suite Plus Includes Premium Suite, five hours of free consulting, and 15% off workshops and conferences	$15,500	$25,100	$31,900
Premium Suite Includes Standard Suite, MBC, and unlimited CCS desktop client conjoint simulators	$13,900	$23,500	$29,900
Standard Suite 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

- 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 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