



Undertaking Volumetric Analysis using Choice Based Conjoint Methods

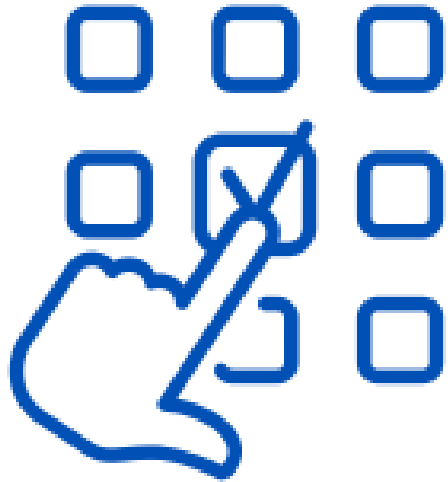
Dean Tindall



Sawtooth Software



On the face of it Volumetrics might appear simple...



What someone is likely to choose



How many of the items they want to buy

Understanding choices, is the *relatively* simply part



- ▶ Conjoint analysis has been around for over 50 years now (Luce & Tukey 1964, McFadden 1974)
- ▶ Academics sought to understand how buyers value differing components of a product, when considering all elements of the product in combination (or conjointly)
- ▶ Initially these methods worked at an aggregate level, utilising OLS or Logit regression models
- ▶ Methods have been continually improved upon over that time, especially with the introduction of Hierarchical Bayes estimation in 1993 which enabled the generation of respondent level models greatly enhancing the accuracy of predictions
- ▶ Today, Choice-Based Conjoint is relatively simple to undertake with many commercial solutions available in the market

Introduction to Choice Based Conjoint

				£0.85
				£0.90
				£0.95
				£1.00
				£1.05
				£1.10
				£1.15

We have 8 flavours of soft drinks to choose from

They can be sold at 7 potential prices

We ask respondents to undertake a series of tasks to understand the choices they make

If these were the only options available to you, which would you choose?



£0.95



£0.90



£0.85



£1.10



£0.90



£1.05



I would not
choose any
of these

- ▶ In this simple example we can understand the impact of changing the pricing of the available soft drinks
- ▶ Respondents are making choices based on the conjoined impact of both product and price
- ▶ But importantly can choose a “None Option” in each task
- ▶ By statistically varying the products and prices available across the tasks, we can understand the potential choices a respondent is likely to make

By analysing the choices made across multiple tasks, we can understand the 'utility' of each element

Level	Utility
Coke	0.6
Diet Coke	0.3
Sprite	-0.3
Dr Pepper	0.4
Pepsi	0.5
Pepsi Max	0.2
Fanta	-0.2
Lilt	-0.6

Level	Utility
£0.85	1.0
£0.90	0.8
£0.95	0.4
£1.00	0
£1.05	-0.6
£1.10	-0.8
£1.15	-1.0

Level	Utility
None	-0.5

- ▶ Despite respondents making choices based on the combination of levels, we are able to identify the value, or 'utility' they placed on each individual level
- ▶ We can see that this respondent prefers Cola, ideally Coke, and is less of a fan of Diet versions
- ▶ We can also identify that they are fairly price sensitive, as the range of utilities for price is much larger than the range of utilities for soft drinks
- ▶ We can also see the threshold at which they would decide to choose None of the products
- ▶ These utilities form the basis of our predictive models

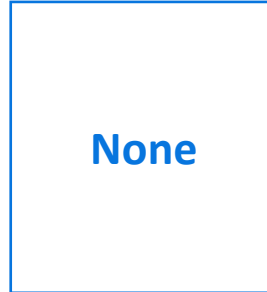
Lets assume our respondent is presented with a simple choice...



£1.00



£0.95



Product Utility	0.6	-0.6	None Utility
Price Utility	0	0.4	-0.5
Total Utility	0.6	-0.2	-0.5
Exp Utility	1.82	0.82	0.61
Share of Preference	56%	25%	19%

- ▶ In order to understand the appeal of each product we can simply sum the utilities associated with each level – this informs us as to the order of preference for the products
- ▶ But if we wish to arrive at a probabilistic measure, which we call share of preference, we need to take the exponent of the utilities for each of our products + our none option

Share of preference formula for coke =

$$\frac{\exp(\text{Coke Utility})}{\exp(\text{Coke Utility}) + \exp(\text{Lilt Utility}) + \exp(\text{None Utility})}$$

- ▶ Our respondent has a 56% likelihood of choosing Coke in this scenario.

Market Simulations

Choice Simulator - acbc

Home My WTP Scenario Settings

Simulation Method: Share of Preference
 Range Behavior: Sensitivity
 Include none option (if present)

SIMULATION SETTINGS

Apply Product Availability (Multi-store) Exponent
 Apply Product Awareness
 Apply Share Adjustment

REPORTING **SEGMENTATION** **EXTERNAL EFFECTS** **GRAPHING**

My WTP Scenario

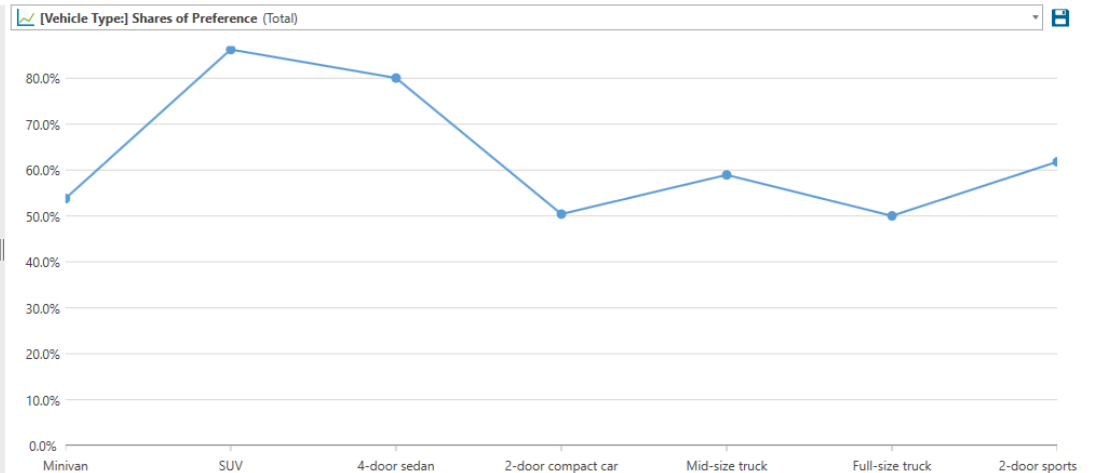
Products

Simulate Off

Label	Vehicle Type:	Manufacture Origin:	Engine Type:	Seat Covering:	Safety Rating:	Reliability Rating:	Price:
1 My Product	=RANGE(1,7)	=RANGE(1,3)	=RANGE(1,3)	=RANGE(1,2)	=RANGE(1,3)	=RANGE(1,3)	=RANGE(10500,52000)
2 Competitor	Minivan	Made in America (at least 75% of parts)	Gas engine	Cloth seats	Average safety rating (4-star)	Average reliability rating (4-star)	30000
+ <type a labi	<type or double cli	<type or double click to select>	<type or double cli	<type or double click	<type or double click to select>	<type or double click to select>	<type or double click to

Results

Label	Shares of Preference	Standard Error	Vehicle Type:	Manufacture Origin:
My Product	53.8%	2.4%	Minivan	Made in America (at least 75% of pa
Competitor	3.7%	0.6%	Minivan	Made in America (at least 75% of pa
None	42.5%	2.5%		
My Product	86.2%	1.5%	SUV	Made in America (at least 75% of pa
Competitor	1.7%	0.5%	Minivan	Made in America (at least 75% of pa
None	12.0%	1.4%		
My Product	80.0%	1.9%	4-door sedan	Made in America (at least 75% of pa
Competitor	7.2%	1.2%	Minivan	Made in America (at least 75% of pa
None	12.8%	1.5%		



Volumetric studies work on similar principles, but instead we ask consumers for volumes not choices

How many of each of the below soft drinks would you purchase in a typical week?



£0.95

15



£0.90

0



£0.85

0



£1.10

0



£0.90

0



£1.05

8

- ▶ Instead of choices, respondents are asked to assign volumes across each product
- ▶ These tasks can either be 'bounded' e.g. "Thinking about your next 10 purchases". Or 'Unbounded' e.g. "How many of these products would you typically purchase within a week?"
- ▶ We capture more information, but these tasks are significantly more difficult for respondents to undertake

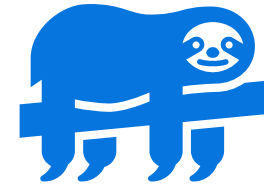
Asking for choices is easy, asking for volumes...less so...



- ▶ Respondents can often answer knowledgeably about which item they would be most likely to choose, but when we ask them how many they often struggle
- ▶ If we then ask them to share an allocation across several products, they are likely to answer with increased levels of error
- ▶ They might not know how many of a particular item they buy in set period, or even how this behaviour might change given a change in scenario
- ▶ We cannot accurately understand whether a respondent is simply stocking up in the short term vs. entering a true long term equilibrium change
- ▶ Volumetric tasks are taxing on a respondent, data quality and consistency is often an issue in these kinds of studies

Respondents can get lazy, lack knowledge or answer unrealistically

- ▶ Initially they may answer honestly, but soon get tired and resort to buying a single unit, when in earlier tasks they bought between 8-10
- ▶ Typos can result in a respondent who was previously buying 10 units suddenly purchasing 100
- ▶ Its also easy to say you will purchase 8 crates of that cheap Hoppy IPA – but how would they take them home in a real store?
- ▶ Respondents often do not have the knowledge of how many units they might ever purchase – especially in a B2B scenario



But we can help respondents with these issues...



- ▶ Before the conjoint we can ask respondents to inform us of their typical purchase behaviours and the maximum they would ever be likely to purchase – and remind them of these during the task
- ▶ We can include volume and price totals to allow respondents to understand the impact of their allocations
- ▶ We can limit the potential inputs to the volumes they enter into the task
- ▶ In order to assist with data quality we can ‘winsorize’ any volumetric data (e.g. replacing the top 5% of responses with the 95th percentile value) in order to reduce the potential impact on data quality which rogue respondents are likely to have

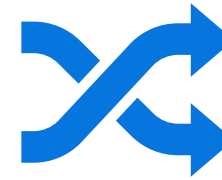
There is also as yet, no single way to undertake volumetric conjoint analysis



Bounded



Unbounded



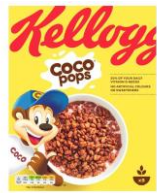
Joint Discrete /
Continuous



Menu Based

Bounded Volumetric Analysis

You told us you typically purchase 10 boxes of cereal every three months, please indicate how many of each of the below you would purchase?



£3.50

2



£2.00

2



£4.50
33% Off

0



£3.00
25% Off

4



£4.75

0



£3.50

0

- ▶ This is the simplest form of volumetric experiment, we simply ask respondents to allocate choices based on a total they already gave us
- ▶ We can identify the relative allocations and how these change throughout each task
- ▶ Any left over allocations are assigned to “None” (in this case 2 of the 10 allocations)
- ▶ Respondent shares of preferences are then weighted by the bounded volume they initially gave us

Bounded Volumetric Analysis

Unweighted Share of Preference

Weighted Bounded Volume

Respondent	Sugar Puffs	Cheerio's	Weetabix	None	X Bounded Volume	Sugar Puffs	Cheerio's	Weetabix	None
1	10%	20%	30%	50%	10	1	2	3	5
2	90%	0%	0%	10%	5	4.5	0	0	0.5
3	0%	50%	50%	0%	8	0	4	4	0
4	25%	0%	75%	0%	20	5	0	15	0
5	0%	75%	25%	0%	12	0	9	3	0

Average Units Purchased

SoP	25%	29%	36%	12%	2.1	4	5	1.1
------------	-----	-----	-----	-----	-----	---	---	-----

Bounded Volumetric Analysis

Pros

- ▶ Extremely simply to perform
- ▶ Requires limited additional data processing
- ▶ Limits the impact of poor quality data as we force respondents to answer in a realistic way
- ▶ Only requires a single model

Cons

- ▶ Respondents are fixed to a set volume, regardless of how good a scenario we present to them
- ▶ Models can only show how volume might shift between products

Unbounded Volumetric Analysis (Synthetic None)

Please indicate how many of each of the below stationery products you would purchase on a typical stock up mission?



£8.00

15



£10.50

3



£4.50

1



£9.10

4



£7.90

0



£27.05

8

- ▶ In unbounded conjoint we do not force respondents to answer with a fixed volume in mind
- ▶ They are free to answer with as many units as they want across every task
- ▶ In this form of analysis, we use the maximum number of units within a single task (MEV) to determine the allocation to None in each task
- ▶ However, this makes us more open to poorer data quality

Unbounded Volumetric Analysis (Synthetic None)

Respondent x Volumetric Data

Task	Units Purchased	Units Allocated to Synthetic None
1	12	3
2	10	5
3	9	6
4	7	8
5	15	0
6	7	8
7	8	7
8	10	5

MEV = 15 Units

- ▶ In this example, the MEV (Maximum Expected Value) is determined by the maximum number of units which was purchased in across any of the respondent's task
- ▶ This is then assumed to be the greatest volume which could be purchased from this respondent
- ▶ The None allocation is then calculated as the difference between the MEV and the number of units purchased within each task
- ▶ The respondent is then weighted by the value of their MEV in the final results

Unbounded Volumetric Analysis (Synthetic None)

Unweighted Share of Preference

Weighted Unbounded Volume

Respondent	Bic Biro	Paper Reem	Notebooks	None	X MEV	Bic Biro	Paper Reem	Notebooks	None
13	5%	10%	5%	80%	35	1.75	3.5	1.75	28
14	10%	30%	0%	60%	25	2.5	7.5	0	15
15	25%	50%	0%	75%	15	3.75	7.5	0	11.25
16	5%	25%	15%	55%	90	4.5	22.5	13.5	49.5
17	0%	5%	0%	95%	100	0	5	0	95

Average Units Purchased

SoP	9%	24%	4%	73%	2.50	9.20	3.05	39.75
------------	----	-----	----	-----	------	------	------	-------

Unbounded Volumetric Analysis (Synthetic None)

Pros

- ▶ The maximum volume purchased is determined by the answers given by respondents
- ▶ An increase in the appeal of a scenario can now increase the potential volume purchased in simulations
- ▶ Only requires a single model
- ▶ Proven to work well on simpler forms of volumetric analysis (Eagle 2010)

Cons

- ▶ Respondent's MEV is based on the version of the study they saw – it could still go higher in theory - *Some have suggested using $MEV * 1.2$ as the upper limit*
- ▶ Requires more in depth data processing to calculate the MEV and Synthetic None on a respondent level - *Sawtooth has a tool for this*
- ▶ More open to poor data quality issues as this is unbounded data - *Winsorize outliers*

Joint Discrete/Continuous Approach

Please indicate how many of each of the below drinks you would purchase in a typical week?



£0.95

15



£0.90

0



£0.85

0



£1.10

0



£0.90

0



£1.05

8

- ▶ For the Joint approach we capture the data in an identical manner to the unbounded approach
- ▶ Respondents are free to answer with as many units as they want across every task
- ▶ Volumes are translated into probabilities for our HB analysis
- ▶ In tasks where no items are selected, None receives 100% of the probability of choice
- ▶ In this approach a secondary regression model is used to identify volume

Joint Discrete/Continuous Approach

Respondent x Volumetric Data

Task	Units Purchased	LNDenom
1	12	5.0245
2	10	4.0045
3	9	3.8982
4	2	-0.3817
5	15	6.0702
6	7	2.1689
7	4	1.3107
8	8	2.9912

- ▶ In order to identify the likely volume which is purchased by each respondent, we need to specify a regression model
- ▶ The model suggested by Eagle(2010) uses the number of Units Purchased in each task as the dependant variable
- ▶ With LNDenom being defined as the only independent variable – so it's a pretty simple model
- ▶ In a scenario with 3 products + None, LNDenom would be defined as;
$$LN(\exp(Ua) + \exp(Ub) + \exp(Uc) + \exp(Un))$$
- ▶ The assumption here, is that the total volume purchased in any task is directly related to the scenarios inherent appeal

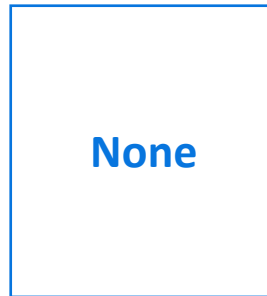
Joint Discrete/Continuous Approach



£1.00



£0.95



Total Utility	0.6	-0.2	-0.5
Exp Utility	1.82	0.82	0.61

- ▶ LNDenom is defined by as the natural log of the denominator in the MNL share of preference model
- ▶ Using our simple scenario we saw earlier this would be

$$LN(\text{Denom}) = LN(\text{Sum}(\text{Exp}(\text{Coke Utility}) , \text{Exp}(\text{Lilt Utility}), \text{Exp}(\text{None Utility})))$$

$$LN(1.82 + 0.82 + 0.61) = 1.178$$

- ▶ We are therefore using the aggregate appeal of the scenario to predict the number of units which are likely to be purchased
- ▶ Respondents are then weighted by their individual regression predictions

Joint Discrete/Continuous Approach

Unweighted Share of Preference

Weighted Joint Volume

Respondent	Coke	Lilt	None	LNDenom	Predicted Units	Coke	Lilt	None
55	55%	0%	45%	1.18	5	2.75	0	2.25
56	10%	56%	34%	2.03	8	0.8	4.48	2.72
57	0%	67%	33%	0.67	2	0	1.34	0.66
58	80%	5%	15%	3.34	12	9.6	0.6	1.8
59	70%	5%	25%	1.89	7	4.9	0.35	1.75

Average Units Purchased

SoP	Coke	Lilt	None
	43%	27%	30%

3.61	1.35	1.84
------	------	------

Joint Discrete/Continuous Approach

Pros

- ▶ The maximum volume purchased is determined by the regression model, based on respondents inputs
- ▶ An increase in the appeal of a scenario will lead to an increase the potential volume purchased in simulations
- ▶ Proven to work well on more complex forms of volumetric analysis (Eagle 2010)

Cons

- ▶ Requires two models
- ▶ Volume prediction models can become overly complex
- ▶ More open to poor data quality issues as this is unbounded data - *Winsorize outliers*

In Summary...

- ▶ One of the key barriers to undertaking CBC volumetric analysis is the quality of responses we are likely to receive from respondents
- ▶ Single choices are easy, but varying allocations can be very difficult to answer
- ▶ Can or should we expect respondents to be able to answer on a volumetric level with any real level of precision?
- ▶ We can offset some of this impact by running a bounded exercise, but we then enforce a huge assumption that the market volume is fixed
- ▶ Unbounded approaches give us the flexibility to model changes in market size, but leave us at the mercy of the quality of responses we receive
- ▶ Data Quality is key to a robust Volumetric CBC

Additional References

1. “Volumetric modeling,” paper presented at Sawtooth Software’s Turbo CBC: A Power User’s Forum on Discrete Choice Modeling, October 8, 2009, Anaheim, CA.
2. “Practical Approaches to Modeling Demand,” with Mark Garratt, paper presented at the 2010 Advanced Research Techniques Forum, American Marketing Association, June 6-9, 2010, San Francisco, CA.
3. “Modeling Demand Using Simple Methods: Joint Discrete/Continuous Modeling,” paper presented at the 2010 Sawtooth Software Conference, October 6-8, 2010, Newport Beach, CA.
4. “A Comparison of Volumetric Models,” with Jordan Louviere and Towhidul Islam, paper presented at the Sawtooth Software Conference, March 7-9, 2018, Orlando, FL.
5. “Volumetric Choice Experiments (VCEs),” with Richard Carson, Towhidul Islam, and Jordan Louviere, *Journal of Choice Modelling*, vol 42 (2022).

THANK YOU

Questions?



Sawtooth Software

UK : +44 1617 685 267

US : +1 801 477 4700

sawtoothsoftware.com

dean@sawtoothsoftware.com