CONJOINT ANALYSIS:

AN INTRODUCTION

Introduction

Conjoint analysis has received a great deal of attention from both practitioners and academics. Because of this attention, conjoint analysis has grown from a single concept into a family of related techniques—many of which are referred to by several names. All of these conjoint methods, though, share the basic tenet of decomposing products[1] into their component parts to *analyze* how decisions are made and then *predict* how decisions will be made in the future. That is, conjoint analysis is used to understand the importance of different product components or product features, as well as to determine how decisions are likely to be influenced by the inclusion, exclusion, or degree of that feature.

Conjoint analysis is sometimes referred to as "trade-off" analysis because respondents in a conjoint study are forced to make trade-offs between product features. In this sense, conjoint analysis is able to infer the "true" value structures that influence consumer decision making; something that other research methods typically cannot. Traditionally, researchers used direct methods of questioning, such as scalar importance questions, for product design research. A researcher who wanted to understand the importance of product features when designing a VCR might ask respondents the following questions:

> ***Using a scale from 1 to 9 where 1 means NOT AT ALL IMPORTANT and 9 means EXTREMELY IMPORTANT, how important is....***
>
> ***....the number of channels on the VCR?***
> ***....the playback picture quality of the VCR?***
> ***....the price of the VCR?***
> ***....the brand of the VCR?***

The researcher would likely learn that all of the features are very important and the typical consumer wants all of the features at a low price. Results such as these are not actionable and therefore largely unusable. Consumers do not have the option of having more of every product characteristic that is desirable and less of every product characteristic that is undesirable. Rather, when purchasing products, buyers must trade-off some of one characteristic to get more of another. This technique, then, is based on the premise that purchase decisions are not made based on a single factor; but on several factors *considered jointly*.

[1] While we will refer to products throughout this document, this discussion is equally appropriate for products or services.

Key Terms

Before discussing the design of conjoint studies, it is important to be familiar with the key terms. Several of these are discussed below.

Using the previous VCR example, the product features such as number of channels and playback quality are referred to as **attributes** in conjoint analysis. These represent the dimensions on which a product can be defined and on which consumers make choices between competitive products. Unfortunately, in the previous VCR example, if the researcher learned that the number of channels was very important, he would have only slight information to guide product design. Is 50 a large number of channels, or is 500? Conjoint analysis, therefore, includes **levels** that represent specific amounts of particular attributes. For instance, there might be four appropriate levels of the attribute *number of channels* as well as four levels for *price*.

Number of Channels[2]	Price
25	$100
75	$200
150	$300
250	$400

After conducting a conjoint study, the researcher will develop quantitative measures of preference for both the attributes (Number of Channels and Price) and their levels (75 Channels and $300). The measures focusing on levels are referred to as **utilities**[3] and the measures focusing on attributes are referred to as **importances**.

Utilities are simply numerical representations that express the value consumers place on each level.[4] For example, the following might represent one respondent's utilities for the two attributes shown above.

Number of Channels		Price	
Level	***Utility***	***Level***	***Utility***
25	0	$100	95
75	20	$200	75
150	35	$300	45
250	65	$400	0

The interpretation of utilities will be discussed in detail later.

Importances are based on utilities and express the range between the most preferred and least preferred level of each attribute. In this way, attribute importances measure the total impact that a particular attribute can have on total preference. To calculate importances, examine the range of the utilities for an attribute. The maximum and minimum utility for each of four attributes is shown on the next page.

[2] Note that the levels for number of channels are not equally spaced (25 to 75 channels represents a 50 channel interval while 75 to 150 represents a 75 channel interval) while those for price are presented with equally spaced increments of $100. Conjoint does not require equally spaced levels, but care must be taken with certain approaches to conjoint analysis when the levels are not equally spaced. Also note that the levels for both attributes are metric, i.e., defined quantitatively. Levels can also be non-metric, as they might be for picture quality or brand name.

[3] Utilities are also sometimes referred to as part-worth utilities or just part-worths.

[4] The unit of measurement for utilities is the "utile".

	Min	Max	Range
Number of channels	0	65	65
Price	0	95	95
Picture quality	0	80	80
Brand	0	60	60
SUM OF RANGES			***300***

This range represents the maximum impact that an attribute can contribute to a product.[5] Frequently, these ranges are presented in terms of **relative attribute importance**, which is calculated by percentaging each range against the sum of the ranges. For example, the relative attribute importances are calculated below for the four attributes in our example.

	Range		Relative Attribute Importance
Number of channels	65	65/300	21.66%
Price	95	95/300	31.66%
Picture quality	80	80/300	26.66%
Brand	60	60/300	20.00%
			100.00%

Background and History

Conjoint analysis is a relatively recent creation of marketing researchers. Even though citations appear back into the 1950s, conjoint analysis, as we know it today, came about in the early 1970s. Conjoint has its roots in decision making and information processing from the field of psychometrics. The development of conjoint was originally a very theoretical search to identify interval level preference data. While many of the highly theoretical questions from the psychometricians remain unanswered,[6] the practical aspects of interval level preference data have found a great home in conjoint analysis. Four specific approaches to conjoint analysis are discussed below.

- Trade-off matrices
- Full profile (ratings-based) card sort
- Hybrid (ratings-based) conjoint
- Discrete choice modeling (choice-based)

Each will be discussed in turn.

[5] It is important to point out that the researcher controls the importance of the attribute by determining the range of levels studied. For example, price, which is currently the most important attribute, would become the least important attribute had the $400 price level been excluded from the design. It is therefore always critical to discuss the levels studied when reporting relative attribute importances.

[6] Consider the seminal 1971 publication by Krantz, Luce, Suppes, and Tversky, Foundations of Measurement, Volume 1. Even though discussed in the first volume, Volume II wasn't published until 1989.

Trade-off Matrices

Some of the earliest practical applications of conjoint analysis were conducted using trade-off matrices. An example of a trade-off matrix is shown below:

		Number of Channels			
		25	75	150	250
Price	$100				
	$200				
	$300				
	$400				

This matrix presents all 16 combinations of the levels of two attributes. Respondents completing this task would fill in their rank order of preferences for all of the 16 cells, keeping an "all other things equal" mindset.[7] For attributes in which there was a clear *a priori* order of preference, the ranks of two of the cells were always known. That is, the combination that offers the most channels at the lowest price is the most preferred combination, and the fewest channels at the highest price is the least preferred combination. The matrix, when completed by a respondent, would look like the following:

		Number of Channels			
		25	75	150	250
Price	$100	6	3	2	1
	$200	9	7	5	4
	$300	12	11	10	8
	$400	16	15	14	13

Since matrices such as this could only handle two attributes at a time, the respondent burden was large in studies with many attributes.

Full Profile (ratings-based) Card Sort

While trade-off matrices were being developed, other approaches to conjoint were gaining support among other researchers. Full profile card sort conjoint is what most researchers would think of as traditional conjoint. While trade-off matrices dealt with only two attributes at a time, full profile conjoint required respondents to evaluate several product concepts, one at a time, defined on all attributes simultaneously. Full profile conjoint, therefore, did away with the direct attribute trade-offs.

These concepts were frequently printed on separate sheets of paper, referred to as "cards." Each card has one level of each attribute and respondents are asked to either rate or rank each concept. The process of sorting these profiles into stacks caused this approach to be referred to as "card sort." An example of a full profile card is shown below.

[7] Note that the resulting utilities data were interval scaled, but that the input was of only ordinal scale. While this issue of how to deal with rank ordered data has largely gone away today, early conjoint utilities were derived using MONANOVA (Monotonic ANOVA), non-metric regression, or other approaches like LINMAP. Today, most conjoint analysis is conducted by either eliciting metric dependent variables and is analyzed with OLS regression or by having respondents make choices (such as pick one of k) and is analyzed with multinomial logit or similar models, such as probit.

Product Concept 12

Brand B
75 Channels
Somewhat fuzzy picture
$300

Rating: ______

Rate this concept on a 0 to 100 scale where 0 means "I would never purchase this product" and 100 means "I would definitely purchase this product."

The levels that appeared on the cards were developed using the principles of experimental design by creating either a full factorial[8] or a fractional factorial design. The number of respondent ratings required with most of these approaches is unworkable for most practical applications. Even with fractional designs, finding a design that represents both reasonable respondent tasks and provides enough degrees of freedom to estimate reasonable parameters can be very difficult.[9]

A second concern with full profile designs centers around respondents' ability to process many attributes. For example, a full profile card with 18 attributes would be very difficult for most respondents to process. In this situation, respondents are likely to ignore certain attributes in the task. This respondent *simplification* would be acceptable if it mirrored the way actual purchases are made, but it is not clear that this is always the case.

If the number of attributes is small (six or fewer) and the number of levels per attribute is small (four or fewer), fractional factorial full profile is generally the method of choice. When there are more than six attributes, however, alternate methods typically offer a superior approach.[10]

Hybrid (ratings-based) Conjoint

The hybrid methods, which are better at handling six or more attributes, include respondents' self-explicated utilities. That is, respondents are asked directly to indicate their preference structure for specific levels of attributes, and this information is included in part-worth estimates. The most widely used hybrid conjoint method is Adaptive Conjoint Analysis (ACA) from Sawtooth Software. ACA combines hybrid self-explicated data with paired comparisons constructed with partial profiles to estimate part-worth utilities. Respondents are first asked to indicate rank order of preference for levels within each attribute and then the importance of the attribute. Respondents then evaluate a series of paired-comparison questions. In the paired comparisons, respondents are presented with two product concepts and asked to indicate their preference using a rating scale, with the middle point indicating both concepts are equally liked.

[8] In a full factorial design, the cards consist of all possible combinations of attribute levels. This approach becomes extremely burdensome even with few attributes. For example, 3 attributes with 4 levels each would require 64 cards (4 X 4 X 4). An additional 4-level attribute would require a total of 256 cards. In a fractional factorial design, a small subset of the possible combinations is used that meets certain requirements for orthogonality and level balance.

[9] To illustrate, consider a 3^8 design (eight 3-level attributes). There are over 6500 possible combinations from which to select the fractional factorial design. Even if 24 could be selected so that the resulting profiles comprised a good design, there are only 8 degrees of freedom remaining. (With 24 dummy-coded levels, there would be 16 parameters to estimate.)

[10] An alternate approach to handling many attributes is through bridging designs. In bridging designs, several small conjoint studies are conducted with the same respondent. In each study the attributes are different, except for one, which is common to all of the studies. This common attribute acts as a bridge allowing the studies to be combined mathematically.

An example of a paired-comparison question is shown below.

Which would you purchase

Brand B 75 Channels $200	Brand A 250 Channels $300
STRONGLY PREFER PRODUCT ON LEFT	***STRONGLY PREFER PRODUCT ON RIGHT***

1 2 3 4 5 6 7 8 9

<u>Discrete Choice Modeling (choice-based)</u>

Finally, the fourth class of conjoint methods is choice-based conjoint. Choice-based conjoint, also referred to as discrete choice modeling, does not ask people to rate their preference for concepts. Rather, choice modeling presents multiple concepts to respondents and asks which one they would choose. This "pick one" task tends to be far easier for respondents.

Choice modeling tends to use full profiles, so the number of attributes is generally limited. Also, since the measurement contains far less information compared to ratings-based conjoint, the analysis has historically been limited to aggregate or subgroup level. However, recent advances in Bayesian techniques[11] are allowing individual estimation from choice studies.

An example of a discrete choice question is shown below.

Which would you purchase?

Brand A	***Brand C***	***Brand B***	***None***
75 Channels	250 Channels	150 Channels	If these were my only alternatives I would not purchase anything
Extremely clear picture quality	Clear picture quality	Somewhat fuzzy picture quality	
$300	$200	$100	

Designing and Executing a Conjoint Study

In designing and executing a conjoint study, the researcher is faced with three steps that are unique to conjoint research:

- Selecting the appropriate type of conjoint
- Selecting the attributes and levels
- Developing and interpreting utilities

Each is discussed below.

[11] The interested reader is referred to Pinnell (2000) or Pinnell and Fridley (2001) for empirical findings.

Selecting The Appropriate Type of Conjoint

In practice, trade-offs matrices are rarely used, narrowing the choice to either ratings-based or choice-based methods. While researchers are divided on this topic, we typically recommend methods that allow respondents to make comparative judgments, such as paired-comparisons with ratings-based conjoint and choice-based conjoint.

We believe the choice between these two approaches depends on the point in the product development cycle. The earlier in the cycle, the more appropriate ratings-based methods are since they more easily accommodate a large number of attributes. Later in the cycle, choice-based methods are more appropriate because many of the development priorities have been solved and the focus is more on the final product configuration, price and brand.

Selecting Attributes and Levels

The single most important component of executing a conjoint study is selecting conjoint attributes and levels. In general, attributes describe product features, such as number of channels and picture quality in the previous example. Conjoint analysis also frequently includes the attributes of price and brand. Many other attributes are possible though, including distribution channel, service or warranty options, product promotions, or positioning statements. The actual attributes used should also follow these guidelines:

1. The attributes must all influence real decisions. That is, the attributes must be determinant.[12]
2. The attributes must be independent. For example, if studying televisions, it would be inappropriate to include one attribute with the level *large screen television* and another attribute that provides screen measurements. While this example seems obvious, many others are not as clear, such as the potential overlap between reliability and quality.
3. The attributes should measure only one dimension. For example, consider the following VCR attribute:

Number of Channels
25
75
150 Cable ready
250 Cable ready

In general, this attribute should be broken into two attributes[13] as follows:

Number of Channels	Cable Ready
25	Yes
75	No
150	
250	

4. Levels must be chosen so that each product can be defined by only one of the levels. For example, consider the following attribute and levels:

Format
Large Screen (48”) Television
29” Television
Includes Picture in Picture

[12] For a thorough discussion of attribute selection, see MacLachlan, Mulhern, and Shocker (1988)

[13] In some instances, leaving the attribute as a compound attribute will better allow the researcher to measure interactions, but care should be taken with this approach.

The Picture in Picture feature should not be included with the format attribute, as either size of television could include the Picture in Picture feature.

5. The levels should include a wide enough range to allow the current and future markets to be simulated. In general, extrapolation of utilities to levels not included should be avoided. If, after including a complete range of levels, the researcher finds many unrealistic combinations of levels, the category definition needs to be revised or respondents need to be given customized conjoint studies. That is, if someone is in the market for a large screen television, asking trade-offs that include small screens will offer little information. Rather, the design should focus on trade-offs specific to large screen televisions.

6. The researcher should work to include a nearly equal number of levels for each attribute. Recent research has indicated the presence of an artificial "number of levels effect" that inflates the relative importance of attributes, which have larger numbers of levels.[14]

Interpreting Results

Interpreting Utilities

One of the most important things about understanding conjoint analysis is understanding how to interpret utilities. Misinterpreting utilities is one of the most common mistakes made by inexperienced conjoint researchers. Consider the following hypothetical VCR utilities:

Number of Channels		Price		Picture Quality		Brand	
Level	***Utility***	***Level***	***Utility***	***Level***	***Utility***	***Level***	***Utility***
25	0	$100	95	Extremely clear	80	Brand A	40
75	20	$200	75	Clear	70	Brand B	60
150	35	$300	45	Somewhat fuzzy	50	Brand C	0
250	65	$400	0	Very fuzzy	0		

To begin, it is important to note that the scaling of the utilities is arbitrary.[15] Because of this arbitrary scaling, the utility of 65 (corresponding to 250 channels) has no meaning by itself. There are, however, two comparisons that can be made to facilitate interpretation of these utilities. This first comparison is a within-attribute comparison. That is, these utilities can be interpreted as interval level data on an *intra*-attribute basis. To illustrate, the following statements can be made regarding this respondent's preference for the number of channels:

1. Prefers 250 channels to all other levels.
2. Least prefers 25 channels.
3. The 20 utile improvement that results from increasing the number of channels from 25 to 75 is larger than the 15 utile improvement that results from increasing the number of channels from 75 to 150.

The second comparison that can be made is across attributes. When comparing attribute utilities on an inter-attribute basis, **only differences between levels can be compared**. For example, we can say that the difference between 25 channels and 75 channels (20 utiles) is less important than the difference between $200 and $300 (30 utiles). Similarly, the difference between 75 channels and 250 channels (45 utiles) is the same as the difference between $300 and $400 (45 utiles). Put another way, this consumer will be indifferent between the following two VCRs:

---

[14] It is still not clear whether the effect is algorithmic or psychological. The effect, though, has been shown to inflate derived attribute importance by 150%. Interestingly, the effect is generally half as large in ACA as in traditional full profile methods. The interested reader is referred to Wittink, Krishnamurthi, and Reibstein (1989) or Wittink, Huber, Fiedler, and Miller (1991).

[15] These utilities have been scaled so that the least preferred level of each attribute is set to zero. Any constant can be added to the utility of all levels of an attribute and not impact the interpretation.

	Product A		Product B	
Number of Channels	75	Utility=20	250	Utility=65
Price	$300	Utility=45	$400	Utility= 0
TOTAL UTILITY		**65**		**65**

Therefore, on an intra-attribute basis, the utilities can be interpreted independently. However, on an inter-attribute basis, only differences between utilities can be compared. Frequently, one wants to comment that $200 (Utility = 75) is more preferred than 250 Channels (Utility = 65). This is wrong! Recall that the scaling of the utilities is entirely arbitrary. In this instance, the utilities were scaled such that the lowest value was set to zero. It would be absolutely "legal" to have scaled the utilities such that the *highest* value was zero. In that case we would have the following utilities:

Number of Channels		Price	
Level	***Utility***	***Level***	***Utility***
25	**-65**	**$100**	**0**
75	**-45**	**$200**	**-20**
150	**-30**	**$300**	**-50**
250	**0**	**$400**	**-95**

To illustrate the similarity of these utilities to the previous set, consider the following example with the rescaled utilities. The difference between 25 channels and 75 channels (20 utiles) is still less important than the difference between $200 and $300 (30 utiles). One might be tempted to say that 250 channels (utility = 0) is more preferred than $200 (utility = -20), the opposite of our previous example. However, this is still wrong! We can only compare differences in utilities across attributes.

Every statement made above is also true with the rescaled utilities except for the statement that $200 (utility = -20) is more preferred than 250 channels (utility = 0).

Conducting Preference Simulations

Conjoint utilities are most frequently used in market simulators that are used to answer "what if" scenarios. After conducting a conjoint study and modeling the current market, a researcher might be interested in the effect of a possible product design change. These scenarios can be investigated in a market simulator. Simulations produce *shares of preference* that resemble—but are not the same as—market share. The researcher must make several decisions in conducting preference simulations. The first decision is which choice model to use.

There are basically two choice models in common use today: the first choice model and the probabilistic model. Each will be discussed in detail below. The discussion of these models is centered on individual level data.

First Choice Model

The first choice model is the more straightforward of the two models discussed. In the first choice model, the researcher sums the utility of each product configuration being simulated and, for each respondent, assumes that the respondent will buy the product with the highest utility. The share of preference estimates, then, become the proportion of respondents for which each product had the maximum utility. While this initially seems reasonable, it might be too simplistic. Very minor differences in summed utilities can have a huge impact on predicted shares of preference.[16]

[16] This model is deterministic. That is, it says that a respondent will prefer one and only one product, no matter how small the difference between the most and second most preferred product. This can make the results too extreme and unstable, as documented in Wiley and Low (1983) and discussed in Elrod and Krishna Kumar (1989).

Probabilistic Model

Instead of assigning 100% of a respondent's purchase to the product with the highest utility, the probabilistic model assigns a probability of purchase to all products, ranging from 0 to 100%. This model is derived from Bradley-Terry-Luce (BTL) and says that the higher the utility for a specific product configuration, the higher the probability of purchase, but that sometimes consumers will buy a product which is not the best on the attributes measured. While the first choice model tends to overstate popular products, the probabilistic model frequently tends to make all products too similar.

A second limitation of the BTL model is an assumption referred to as *independence from irrelevant alternatives* (IIA). This is also referred to as the constant-odds property. IIA says that the probability of choice for two alternatives will always maintain the same ratio, independent of other products introduced into the competitive landscape. This is frequently illustrated with a common example referred to as Red Bus/Blue Bus.

Assume that each morning an individual can drive to work or take the Blue Bus that stops at a street corner one block from his house. Assume that out of every 10 days, he drives nine days and takes the bus one day. The probability of taking the bus is 0.10 and of driving 0.90. Now assume that a Red bus is entered into service following the exact same route and timing as the Blue Bus. The IIA assumption mandates that the probabilities of taking Blue Bus and of driving will always have the same nine-to-one relationship. Separately, the two buses should have an equal share since they offer equal services. With the IIA assumption, the resulting probabilities would be: Drive = 0.8182, Red bus = 0.0909, Blue bus = 0.0909. In actuality, the shares are more likely to be: Drive 0.90, Red bus = 0.05, Blue bus = 0.05.

Adjustments have been offered to account for IIA. Other model specifications, such as hierarchical logit, are designed to circumvent the limitation brought about by IIA.

Assumptions Underlying Choice Simulations

A number of assumptions are made in calculating shares of preference. If these assumptions are all met, shares of preference will resemble market shares. The general assumptions of choice simulators are the following:

- All brands and all products are equal in terms of distribution, promotion, awareness, advertising, and that all buyers have perfect information concerning each product's true specifications and are aware of all the options available to them.
- Each respondent is actually in the market and each person will purchase the same number of units.
- The researcher has not violated other assumptions of conjoint analysis related to design, such as selecting and fully specifying orthogonal and determinant attributes.

External Effects and Re-weighting

If these assumptions are not met, the researcher can often make adjustments to the data to better match market shares. These adjustments generally fall under the umbrella term of "external effects."

For example, a very low awareness or low distribution brand will likely garner a share of preference larger than its actual market share. Similarly a high awareness brand's predicted shares will likely be too low. Generally, this results from a violation of the equal awareness assumption outlined above. As researchers, we frequently make respondents more aware of product alternatives and make these alternative products far more available than they are in the market. One solution is to include posterior weights to the resulting shares of preference that reflect differential brand positions (such as unaided brand awareness). These external effects typically transform shares of preference to more closely resemble actual market positions.

References

Elrod, Terry and S. Krishna Kumar (1989), “Bias in the First Choice Rule for Predicting Share,” *Proceedings of the Sawtooth Software Conference.*

MacLachlan, Douglas, Michael Mulhern, and Allan Shocker (1988), “Attribute Selection and Representation in Conjoint Analysis: Reliability and Validity Issues,” *Proceedings of the Sawtooth Software Conference.*

Pinnell, Jon (1994), “Multistage Conjoint Methods to Measure Price Sensitivity,*" Proceedings of the AMA's Advanced Research Techniques Forum.*

Pinnell, Jon (2000), Customized Choice Designs: "Incorporating Prior Knowledge and Utility Balance in Discrete Choice Experiments," *Proceedings of Sawtooth Software Conference.*

Pinnell, Jon and Lisa Fridley (2001), "Effect of Disaggregation with Partial Profile Choice Experiments," *Proceedings of Sawtooth Software Conference*.

Wiley, James and James Low (1983), “Monte Carlo Simulation Study of Two Approaches for Aggregating Conjoint Data,” *Journal of Marketing Research*, 20, 405-417.

Wittink, Dick, Joel Huber, John Fiedler, and Richard Miller (1991), “An Empirical Comparison of ACA and Full Profile Judgments,” Proceedings *of the Sawtooth Software Conference*.

Wittink, Dick, Lakshman Krishnamurthi, and David Reibstein (1989), “The Effect of Differences in the Number of Attribute Levels on Conjoint Results,” Marketing *Letters,* 2, 113-123.

For more information please contact MarketVision Research at:

info@mv-research.com

513.791.3100

•

or visit our website at:

www.mv-research.com