Getting Started with Conjoint Analysis

Strategies for Product Design and Pricing Research

Fourth Edition (Sample Chapter 11)

Bryan K. Orme

Research Publishers LLC

Getting Started with Conjoint Analysis:
Strategies for Product Design and Pricing Research
Fourth Edition
by Bryan K. Orme

Publisher
Research Publishers LLC
2711 N. Sepulveda Blvd. #320
Manhattan Beach, CA 90266-2725
USA
http://www.research-publishers.com

ISBN13 978-1-60147-115-4 (full text, paper)
ISBN13 978-1-60147-116-1 (full text, electronic, pdf)

Cover art by Paul Klee (1879–1940)
Maske Furcht, 1932, 286 (Y 6)
100 x 57 cm
Ölfarbe auf Jute auf Keilrahmen
The Museum of Modern Art, New York

Cover design by Jane Tenenbaum, Tenenbaum Design, Cambridge, Mass.
Electronic typesetting by Amy Hendrickson, TEXnology, Inc., Boston
Printed and bound by G2 Graphic Service, North Hollywood, California

Printed in the United States of America. 9 8 7 6 5 4 3 2

Contents

Chapter 11

Choice-Based Conjoint Example

This chapter presents a case study in choice-based conjoint (CBC) analysis. The data in the case are real, but the product category and attribute lists have been changed to refer to generic products of a fictional company and its competitors. The data may be downloaded from

`www.sawtoothsoftware.com/download/case4.zip`

11.1 Research Problem

Venezia-Altura manufactures a piece of equipment used in the construction industry and is in the early stages of developing a new feature to improve equipment precision. Greater precision means faster operation and less rework/waste, so it is thought that the new feature may lead to increased market share and profitability. Executives at Venezia-Altura favor a line extension strategy, offering an enhanced precision model at a higher price alongside the standard precision model at its current price. The business questions are:

- Do customers care about enhanced versus standard precision?
- The firm hopes to charge a $225 premium over the standard model to cover costs and yield a higher profit. Would consumers choose the enhanced precision model at the higher price? Who are these consumers?
- Would the line extension cannibalize Venezia-Altura's existing sales or would it draw significant share from competitors? The executives believe it would only be worth the complication and cost of maintaining two models if market share increases by at least 5 percent.

11.2 Formulating an Attribute List

The focus of this research is the value that consumers place on enhanced versus standard precision for this type of equipment. Simulating consumer choices for Venezia-Altura and its major competitors will provide an indication of whether the enhanced precision feature could lift Venezia-Altura's sales and profits. Thus, the conjoint analysis attribute list should include the top competitors in this space at their current capabilities and prices. Suppose that research shows that seven brands account for 90 percent or more of market share. Fortunately, seven brands do not pose an overly challenging measurement problem for conjoint analysis.

Product capacity and price are likely to be important attributes to the buyer decision. Although other features, such as warranty and product form factors, may be used to describe products in this category, these features do not affect buyer choice as much as as brand, capacity, and price. Accordingly, we include only four attributes and their corresponding levels in the choice study design.

Attribute 1 (Brand)
1) Bastian Brothers
2) Lordes & Co
3) Venezia-Altura
4) King Craftworks
5) Sanford Industries
6) Milroy Machines
7) Knell International

Attribute 2 (Precision)
1) Enhanced
2) Standard
3) Substandard

Attribute 3 (Capacity)
1) Mega
2) Super
3) Medium

Attribute 4 (Price)
1) $400
2) $500
3) $650
4) $850

This list of four attributes, with the most complicated attribute having seven levels, should not place excessive demands on respondents. Also, we expect to obtain precise estimates of utilities and shares of preference from a moderately sized sample. As for the price attribute, we include a wide range of values to reflect lower- and upper-end products in the construction product space. Note that it is not necessary to have equal steps between price levels. We arrange price levels so there are increasingly larger distances between adjacent prices as prices increase (a common practice).

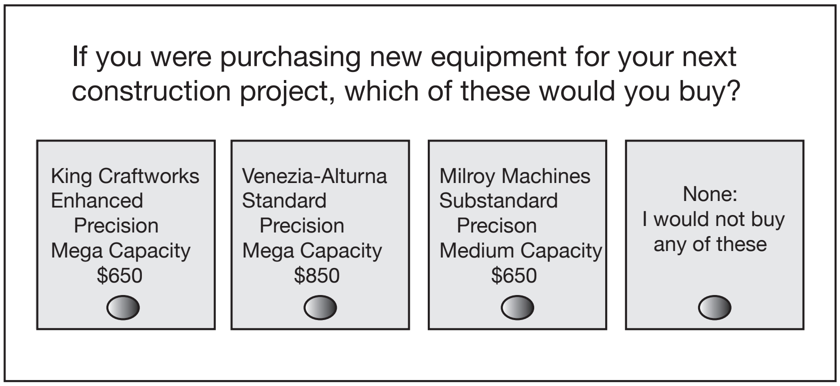

Exhibit 11.1. Sample CBC Task

11.3 Designing the CBC Questions

Practitioners have found that 8 to 12 questions or tasks work well for most online choice surveys as long as respondents are properly recruited and given appropriate incentives. We use 12 choice tasks per respondent, with each task having three product alternatives and a *none* option, as shown in exhibit 11.1. Prior to completing choice tasks, respondents would be screened and educated regarding this particular piece of equipment and its attribute levels.

Although we are primarily interested in how enhanced precision may improve Venezia-Altura's product, consumers are asked to consider any of the seven brands making this technological improvement. Allowing enhanced precision to appear freely with any of the seven brands increases measurement precision about the preference for enhanced precision on the Venezia-Altura offering.

Each of the 12 choice tasks shows distinct combinations of the four attributes. Although there are $7 \times 3 \times 3 \times 4 = 252$ unique products that could be created by taking one level from each of the four attributes, it is not necessary to show all 252 possible products to a respondent. Nor is it necessary to show all possible product combinations across respondents. What is important is that we show a sufficiently large variety of combinations in an independent (orthogonal) fashion. This ensures efficient estimation of part-worth utilities for attribute levels. Accordingly, we ensure that attribute levels (within each attribute) appear with nearly equal frequency. We also ensure that each level of each attribute appears nearly equally with each level of every other attribute. This makes for a fair and balanced experiment. In contrast, if we were to show the lowest price ($400) every time for enhanced precision products, we would not know whether a product were being chosen due to its enhanced precision or its lower price.

We design the CBC questionnaire so that each individual respondent receives his or her own version (block) of conjoint questions/tasks. This increases the variation in product concepts across respondents and also reduces the possibility of psychological order effects affecting the results.

Software for conjoint analysis can do an excellent job of selecting combinations of attribute levels for choice tasks. We favor experimental plans featuring level overlap, with attribute levels sometimes repeating within the same choice task. Note that exhibit 11.1 shows a choice set of three product concepts with mega capacity appearing twice. A modest amount of level overlap tends to improve the estimation of interaction effects (if desired), and it also creates a more informative questionnaire for a respondent who, for example, requires a mega capacity. Showing mega capacity twice within the same choice task gives us a chance to see which attribute is next most important in driving respondent choice.

11.4 Data Collection

Although we would usually favor 300 to 600 respondents for such a conjoint analysis survey, the data set for this example includes only 201 respondents prior to cleaning. Imagine that we have had difficulty recruiting respondents or have been unable to provide adequate incentives for responding.

Suppose that survey respondents are qualified based on their job function, whether they are involved in purchasing construction equipment, and whether they have influence on the purchasing decision. Prior to presenting the online conjoint survey, we educate respondents about all study attributes and levels equally, avoiding emphasis on the key issue of enhanced versus standard precision.

Suppose we collect survey data using Amazon's Mechanical Turk panel, which is an Internet crowd-sourcing marketplace for matching humans who want to earn money doing human intelligence tasks with employers willing to pay them for those tasks.[1]

Mechanical Turk respondents often complete surveys quickly. They are adept at producing data that cannot easily be flagged as "bad," so they will be paid for their work and selected for additional work. These respondents know not to straight-line, choosing the same answer repeatedly. Fortunately, conjoint analysis utility estimation yields a fit statistic (a consistency score) that helps us identify respondents who are answering near-randomly versus those who are following a logically consistent or rational choice strategy.

Top-Line View of the Raw CBC Data

Two hundred and one (201) respondents completed the questionnaire, 11 percent using a mobile phone. There were four concepts presented in each choice task, with the *none* alternative always in the fourth or right-most position. The composition of the first three concepts varied according to a fair and balanced experimental design plan.

[1] We are not claiming Amazon's Mechanical Turk is the best provider of sample; we merely refer to Mechanical Turk because this is how we actually collected the (disguised) data for this example.

The 201 respondents contributed $201 \times 12 = 2,412$ choices, which were distributed fairly evenly across choice task positions.

Concept Choices by Position
683 first concept choices (28.3%)
725 second concept choices (30.1%)
697 third concept choices (28.9%)
307 *none* choices (12.7%)

This fairly even distribution across choice task positions is an expected result, given that attribute levels are distributed evenly across the three positions. The 12.7 choice percentage for the *none* alternative falls within a typical *none* usage range, around 5 to 20 percent. We should interpret *none* choice frequencies with caution, noting that clients may mistakenly think of *none* frequencies as inversely related to real-world purchasing frequencies.

Examining choices for each of the 201 respondents, we find no straight-liners (no person answered the same way across all 12 choice tasks). Furthermore, the maximum number of *none* choices for any one individual is 9 out of 12 tasks. If a person had answered *none* to all 12 tasks, we would discard that person's data as someone who was not in the market for the product under study.

The median time to complete the 12 choice tasks is 136 seconds, or 11.3 seconds per task (minimum time 21 seconds, maximum time 623 seconds). Although this is a relatively simple CBC study with limited text per choice task, a respondent completing 12 tasks in 21 seconds (an average of 1.75 seconds per task) would probably not have been conscientiously considering attribute tradeoffs.

Data Cleaning (Identifying "Bad" Respondents)

It is typical in market research to collect more respondents than needed, allowing for the possibility of discarding 10 to 30 percent of the sample as "bad." None of the Mechanical Turk respondents straight-lined, but there are other metrics to consider when data cleaning. Suppose we examine respondent inconsistency (randomness) and speeding.

How do we identify random or near-random responders? We use hierarchical Bayes (HB) estimation to obtain a fit/consistency score for each respondent. It is possible that random-responding actors can get lucky and appear to be consistent, just as it is possible to throw a long series of heads in a row when repeatedly flipping a coin. So, we consider the range of outcomes that we would expect across hundreds of random responders. We generate a few hundred random responders and estimate their utilities using HB. We sort the fit/consistency scores from lowest to highest and note the 95^{th} percentile (best) fit. If a respondent is truly random, he or she will have a fit score lower than this 95^{th} percentile cutoff 95 percent of the time. On the other hand, a truly random responder still has a 5 percent chance of getting lucky enough to have a fit score higher than the 95^{th} percentile cutoff, so this method of detecting random respondents is not foolproof.

Suppose we employ a consistency check, requiring each respondent to have an HB fit score above the 95$^{\text{th}}$ percentile cutoff of random responders. We examine HB fit scores for respondents in the CBC study, deleting 19 of the 201 respondents due to their failing this consistency check.

Speeding relates to respondents who rush through surveys, not conscientiously considering attribute tradeoffs. Respondents who complete 12 choice tasks with an average speed of 5 seconds or less per task finish the survey in 60 seconds (one minute) or less. These respondents are speeding through the survey. We identify 21 of the 201 respondents as speeders and delete them from further analysis.

We could stop data cleaning at this point or go on to apply an additional rule to ensure data quality. Suppose we delete respondents who are moderate speeders (top one-third in speed) and also have moderately low consistency (bottom one-third in fit/consistency scores). Applying this third data cleaning rule leads to the deletion of 7 additional respondents.

Data cleaning rules involve subjective judgment and need to be adjusted to avoid deleting too many respondents (throwing away too much good with the bad). Applying the three data cleaning rules or quality checks we have described here involves deleting 19 + 21 + 7 = 47 respondents or 47/201 = 23 percent of survey respondents. We are left with 201 - 47 = 154 good respondents for analysis. This is a lower sample size than we would prefer for a four-attribute CBC study, but we will proceed with this case study illustration nonetheless.

11.5 Utility Estimation

We fit an HB multinomial logistic model to the case study data, estimating utility scores (part-worth utilities) for each attribute level plus a utility for the *none* concept.[2] We begin answering Venezia-Altura's management questions by referring to estimated utilities.

Does the Market Care about Enhanced Precision?

We can answer the first business question posed at the beginning of this chapter by estimating utilities for attribute levels. We conduct a matched sample *t*-test on enhanced versus standard precision utilities (where each respondent's utilities are normalized to be on the same scale). The difference in utility between having enhanced versus standard precision is positive and statistically significant, with a confidence level greater than 0.9999. So, we are extremely confident that the market prefers enhanced over standard precision for this piece of machinery.

It is possible to report average utilities and importance scores for all attributes, as we have for other examples in this book. But a more telling and perhaps more intuitive rendering of study results comes from market simulations.

[2] The details of HB estimation in choice-based conjoint (CBC) studies are described in *Becoming an Expert in Conjoint Analysis* (Orme and Chrzan 2017).

Brand	Precision	Capacity	Price	Share of Preference
Bastian Brothers	Standard	Super	$500	30.60%
Lordes & Co	Substandard	Mega	$475	6.50%
Venezia-Altura	Standard	Super	$575	28.10%
King Craftworks	Substandard	Super	$450	1.10%
Sanford Industries	Standard	Medium	$550	5.20%
Milroy Machines	Standard	Super	$600	20.00%
Knell International	Substandard	Super	$575	5.90%
None				2.50%
			Total:	100.00%

Exhibit 11.2. Base-Case Shares of Preference

11.6 Market Simulation

Conducting a market simulation is like creating a voting machine. We place multiple product offerings in competition with one another, creating a base-case set of products or market scenario. Using the 154 respondents with clean/good data, we estimate part-worth utilities and predict the likelihood that each respondent will pick each product. Then we estimate shares of preference for products in the market scenario. Simulation results for an initial set of products or base-case scenario are shown in exhibit 11.2,

Would consumers choose an enhanced precision model at a higher price?

Recall that Venezia-Altura hopes to charge a $225 premium over the standard precision product to cover development and manufacturing costs and to yield a profit. How many consumers would choose the enhanced precision product at its premium price?

To answer this question, we add a product alternative to the base-case scenario or choice set—a Venezia-Altura product with enhanced precision at a $225 price premium over the firm's standard product. We assume other product attributes and prices do not change from the base-case scenario, and we run an additional market simulation, estimating shares of preference under the expanded scenario.

Simulation results for the expanded scenario are shown in exhibit 11.3. An enhanced precision Venezia-Altura model can be expected to capture a 14.2 percent share of preference if delivered at a $225 premium over the standard model. Furthermore, moving from a seven-product base-case scenario to an eight-product expanded scenario increases Venezia-Altura's total share of preference.

Would the line extension cannibalize Venezia-Altura's existing sales?

Comparing base-case and enhanced product scenarios, we note that Venezia-Altura's share of preference grows from 28.10 percent to 23.00 + 14.20 = 37.20 percent. This increase of 9.1 percent in share of preference is due mostly to new

Brand	Precision	Capacity	Price	Share of Preference
Bastian Brothers	Standard	Super	$500	27.20%
Lordes & Co	Substandard	Mega	$475	5.60%
Venezia-Altura	Standard	Super	$575	23.00%
Venezia-Altura	Enhanced	Super	$800	14.20%
King Craftworks	Substandard	Super	$450	1.00%
Sanford Industries	Standard	Medium	$550	4.70%
Milroy Machines	Standard	Super	$600	16.40%
Knell International	Substandard	Super	$575	5.60%
None				2.30%
			Total:	100.00%

Exhibit 11.3. Shares of Preference after New Product Introduction

Venezia-Altura consumers choosing the enhanced precision model, rather than to current Venezia-Altura customers moving from the standard precision model to the enhanced precision model (cannibalization). If we were to run a matched sample *t*-test for these data, we would see that this 9.1 difference in share of preference is statistically significant at a high confidence level. The executives believe it would only be worth the complication and cost of maintaining two models if market share were to increase by at least 5 absolute percentage points. Indications are that it does.

Who are the respondents choosing the enhanced precision product?

To describe those respondents who choose particular products, we rely on information about the demographics or business characteristics of those respondents. We filter or sort respondents into groups and note predicted product choices across those groups.

For any group of respondents, we can use shares of preference to estimate the average probability of selecting one product among a set of products. Suppose we compare shares of preference for standard versus enhanced precision products for commercial versus residential builders. For the 154 respondents in the CBC study, we see that, compared with residential builders, commercial builders have a much higher share of preference for the enhanced precision product.

Enhanced Precision Product Share of Preference by Buyer Group
Residential builders (n = 72): 8.9%
Commercial builders (n = 82): 18.8%

Market Simulation Summary

There are caveats with interpreting shares of preference. As noted in chapters 3 and 10, shares of preference often do not translate directly to market shares in the real world. Nonetheless, market simulation results suggest a very good opportunity for Venezia-Altura. The net share of preference for two products (standard and enhanced precision products) is significantly larger than for just one product (standard precision product alone). And many customers are expected to prefer the enhanced precision product compared to the standard product, even when that enhanced precision product is sold at a $225 price premium. In sum, offering an enhanced precision product could increase the firm's total market share and profits.

No market is static. If Venezia-Altura were to offer both standard and enhanced precision products, its competitors may follow suit, introducing enhanced precision products of their own and/or changing product prices. What happens, for example, if Bastian Brothers introduces an enhanced precision product selling at a $225 premium compared to its own standard precision product? We would run additional market simulations to answer this and many other questions.