Sawtooth Software

RESEARCH PAPER SERIES

Which Conjoint Method Should I Use?

Bryan K. Orme
Sawtooth Software, Inc.

530 W. Fir St.
Sequim, WA 98382
(360) 681-2300
www.sawtoothsoftware.com

Which Conjoint Method Should I Use?

Bryan Orme, Sawtooth Software

Foreword

We originally published an article by this title in the Fall 1996 issue of *Sawtooth Solutions*. Interest in that paper along with a steady flow of new developments in the conjoint analysis field have led us to update this piece now four times.

It is paradoxical that new developments have made the methods better than ever, but they have also made it more challenging to determine whether a given project clearly would work better with one method or another. Many differentiating limitations which earlier caused researchers to reject one "flavor" of conjoint analysis in favor of another have been overcome, thus blurring the lines between the unique capabilities of the approaches.

Most recently, we have added an *interactive adviser* to our website that helps guide the decision of conjoint method at www.sawtoothsoftware.com/products/advisor/. This article provides greater depth for understanding the issues involved with choosing a conjoint approach.

Introduction

Conjoint analysis has become one of the most widely used quantitative tools in marketing research. According to recent Sawtooth Software customer surveys, we estimate that from 8,000 to 10,000 conjoint studies are conducted each year by our customers. When used properly, it provides reliable and useful results. There are multiple conjoint methods. Just as the golfer doesn't rely on a single club, the conjoint researcher should weigh each research situation and pick the right combination of tools.

Sawtooth Software has been producing conjoint analysis software since 1985, and over the years has offered a variety of conjoint analysis software systems. The older systems involve rating product concepts on sliding scales (such as 1 to 9) or on a 100-point scale. Our newer systems ask respondents to choose among sets of products presented in choice scenarios. Although many researchers still use the older ratings-based approaches and there is evidence they can work well when designed and executed correctly, the majority of recent academic support and practical research applications are in favor of choice-based approaches.

The Ratings-Based Systems

The first method of conjoint analysis was introduced to the market research community by Paul Green and colleagues in the early 1970s. It involved asking respondents to rate (or rank) a series of concept cards (where each card displayed a product concept consisting of multiple attributes).

Respondents typically rated between a dozen to thirty cards described on up to somewhere around six attributes. At the time, Paul Green and colleagues felt respondents couldn't deal with more than about six attributes without resorting to problematic simplification strategies. But, perhaps the greater limitation was that increasing the number of attributes meant that even more cards had to be presented to respondents to obtain good results. At some point, respondents would burn out and not give good responses to increasingly deeper decks of cards. Not surprisingly, this first conjoint technique was called "card-sort conjoint." Sawtooth Software's CVA system does this flavor of conjoint analysis, as well as an extension involving paired comparison judgments.

In 1985, Sawtooth Software released its first conjoint analysis software system called ACA (Adaptive Conjoint Analysis). ACA went on to become the most popular conjoint software tool and method in both Europe and in the US throughout the 1990s. ACA's main advantage was its ability to measure more attributes than was advisable with the earlier card-sort conjoint approach. With ACA, it was possible to study a dozen to two-dozen attributes, while still keeping the respondent engaged and providing good data. ACA accomplished this by having varying sections of the interview that adapted to respondents' previous answers. In each section, only one or a few attributes were presented at a time, so as not to overwhelm the respondent with too much information at once. The software led the respondent through a systematic investigation over all attributes, resulting in a full set of preference scores for the levels of interest (part-worth utilities) by the end of the interview.

In terms of limitations, the foremost was that ACA needed to be computer-administered. The interview adapts to respondents' previous answers, which cannot be done via paper-and-pencil. Like most traditional conjoint approaches, ACA is a main-effects model. This means that part-worth utilities for attributes are measured in an "all else equal" context, without the inclusion of attribute interactions. This can be limiting for some pricing studies where it is sometimes important to estimate price sensitivity for each brand in the study. ACA also exhibited another limitation with respect to pricing studies: when price was included as just one of many variables, its importance was likely to be understated, and the degree of understatement increased as the number of attributes studied increased.

Many researchers continue to use ACA today, but they tend to avoid pricing applications and also take care to implement the latest best practices for ACA research. For example, the self-explicated importance questions near the beginning of the interview have been problematic if not administered well. The ACA documentation and recent white papers from Sawtooth Software discuss methods to improve this potentially troublesome area.

Choice-Based Conjoint (CBC)

Choice-Based Conjoint analysis started to become popular in the early 1990s, and lately has become the most widely used conjoint technique in the world. CBC interviews closely mimic the purchase process for products in competitive contexts. Instead of rating or ranking product concepts, respondents are shown a set of products on the screen and asked to indicate which one they would purchase:

If you were shopping for a credit card, and these were your only options, which would you choose?			
VISA $40 annual fee 10% interest rate $2,000 credit limit	Mastercard $20 annual fee 18% interest rate $5,000 credit limit	Discover No annual fee 14% interest rate $1,000 credit limit	NONE: I would defer my purchase

This example shows just three product concepts and a "None." As in the real world, respondents can decline to purchase in a CBC interview by choosing "None." If the aim of conjoint research is to predict product or service choices, it seems natural to use data resulting from choices. Many CBC projects (especially packaged goods research) will show a dozen or more products on the screen, often graphically displayed as if they were on physical shelves of a store. We generally recommend, whenever it is possible and realistic, that researchers show more rather than fewer product concepts per choice task.

Despite the benefits of choice data, they contain less information than ratings per unit of respondent effort. After evaluating a number of product concepts, the respondent tells us which one is preferred. We do not learn whether it was strongly or just barely preferred to the others; nor do we learn the relative preference among the rejected alternatives.

Our CBC system can include up to 10 attributes with 15 levels each (unless using the Advanced Design Module, where up to 30 attributes with 254 levels per attribute are permitted), though we'd never recommend you challenge these limits. CBC can be administered via CAPI or Internet surveys, or via paper-and-pencil. In contrast to either ACA or CVA (which automatically provide respondent-level part-worth preference scores), CBC results have traditionally been analyzed at the aggregate, or group, level. But with the availability of latent class and hierarchical Bayes (HB) estimation methods, group-based and individual-level analyses are now accessible and practical. In recent surveys of Sawtooth Software customers, the majority of CBC users are reporting that they use HB estimation for the final market simulation models delivered to clients.

There are a number of ways to analyze choice results, including:

> Aggregate Choice Analysis was the first and generally only form of analyzing CBC results, prior to advances in algorithms and the availability of faster computers. It was argued that aggregate analysis could permit estimation of subtle interaction effects (say, between brand and price), due to its ability to leverage a great deal of data across respondents. For most commercial applications, respondents often cannot provide enough information with even ratings- or sorting-based approaches to measure interactions at the individual level. While this advantage seems to favor aggregate analysis from choice data, academics and practitioners have argued that consumers have unique preferences and idiosyncrasies, and that aggregate-level models which assume homogeneity cannot be as accurate as individual-level models. Aggregate CBC analysis also suffers from its IIA (Independence from Irrelevant Alternatives) assumption, often referred to as the Red Bus/Blue Bus problem. Very similar products in competitive scenarios can receive too much net share. IIA models fail when there are differential

cross-effects between brands, unless steps are taken to develop sophisticated models that explicitly account for cross-effects.

Latent Class Analysis addresses respondent heterogeneity in choice data. Instead of developing a single set of part-worths to represent all respondents (aggregate analysis), Latent Class simultaneously detects homogeneous respondent segments and calculates segment-level part-worths. If the market is truly segmented, Latent Class can reveal much about market structure (including group membership for respondents) and improve the predictability over aggregate choice models. Subtle interactions also can be modeled in Latent Class, which seems to offer a compromise position, leveraging the benefits of aggregate estimation while recognizing market heterogeneity.

HB (Hierarchical Bayes Estimation) HB offers a very powerful way for "borrowing" information from every respondent in the data set to improve the accuracy and stability of each individual's part-worths. It has consistently proven successful in reducing the IIA problem and in improving the predictive validity of both individual-level models and market simulation share results. HB estimation can employ either main effects or models that additionally include interaction terms. But, researchers are finding that many (if not most) of the interaction effects that were discovered using aggregate CBC analysis were actually due to unrecognized heterogeneity. So, often main effects models with HB are sufficient to model choice. We'll explain this further.

Suppose we have individual-level part-worths in a data set, and there are two types of respondents. One group prefers Brand A and is less price sensitive; the other prefers Brand B and is more price sensitive. If we perform sensitivity simulations with no interaction terms included, we will see that the share for Brand B is more sensitive to price changes than Brand A. Brand B respondents are *more* likely to switch to Brand A due to price changes than vice-versa. Even though no interaction terms were included, a brand/price interaction was revealed due to *between*-group differences in price sensitivity operating within the context of a market simulator.

In general, we believe the benefits of individual-level part-worths make a compelling argument for HB estimation. We have consistently seen HB estimation out-perform aggregate logit for predicting shares for holdout choices and actual market shares, even when there was very little heterogeneity in the data.

Partial-Profile CBC

Many researchers that favor choice-based conjoint rather than ratings-based approaches have looked for ways to increase the number of attributes that can be measured effectively using CBC. One solution that gained some following over the last decade is partial-profile CBC. With partial-profile CBC, each choice question includes a subset of the total number of attributes being studied. These attributes are randomly rotated into the tasks, so across all tasks in the survey each respondent typically considers all attributes and levels.

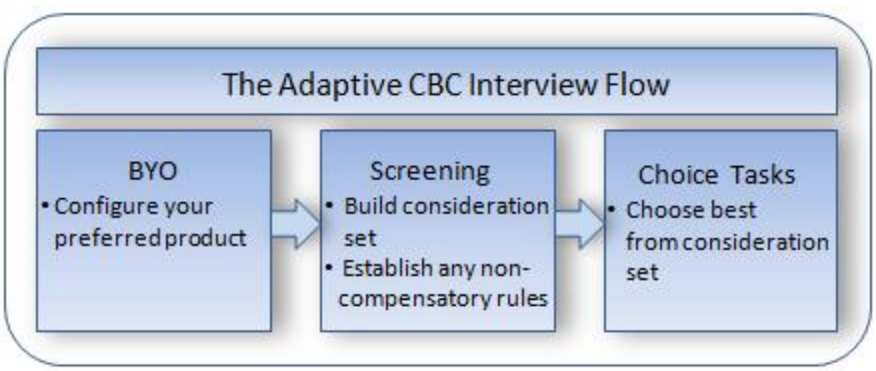
The problem with partial-profile CBC is that the data are spread quite thin, because each task has many attribute omissions, and the response is still the less informative (though more natural) 0/1 choice. As a result, partial-profile CBC requires larger sample sizes to stabilize results, and individual-level estimation under HB doesn't always produce stable individual-level part-worths. Despite these shortcomings, some researchers who used to use ACA for studying many attributes shifted to partial-profile choice. The individual-level parameters have less stability than with ACA, but if the main goal is achieving accurate market simulations (and large enough samples are used), some researchers are willing to give up the individual-level stability.

Lately, we've come to realize that partial-profile CBC studies may be subject to a similar price bias as ACA (though not as pronounced). Recent split-sample studies presented at the Sawtooth Software conferences have shown that price tends to carry less weight, relative to the other attributes, when estimated under partial-profile CBC rather than full-profile. Furthermore, partial-profile methods assume that respondents can ignore omitted attributes and base their choice solely on the partial information presented in each task. If respondents cannot, then this biases the final part-worth utilities. For this and other reasons, most researchers and academics favor full-profile conjoint techniques that display all attributes being studied within each choice task.

Adaptive CBC (ACBC)

Choice-based rather than ratings-based conjoint methods have become dominant in our industry. Yet, standard CBC questionnaires can seem tedious to respondents, repetitive, and to lack relevance. The same-looking choice tasks repeat and repeat. Products shown to respondents seem all "over the board" and not very often near what the respondent actually wants.

Recently, Sawtooth Software developed a new approach called Adaptive CBC (ACBC), leveraging aspects of adaptive conjoint analysis (ACA) and CBC. It first asks respondents to identify the product closest to their ideal using a configurator (Build Your Own—BYO) question. The BYO task also serves as an excellent training exercise, to acquaint respondents with the attributes and levels being studied. Next, we build typically a couple dozen product concepts for the respondent to consider, all quite similar (near neighbors) to the BYO product. Respondents indicate which of those they would consider. Considered products are taken forward to a choice tournament to identify the overall best concept, where the choice tasks look very much like standard CBC tasks.



Recent evidence suggests that respondents find the ACBC interview more engaging and realistic, even though the interview generally takes longer than CBC to complete. But, sample size requirements are smaller than standard CBC, because more information is captured from each individual. More information at the individual level also leads to better segmentation work. Early evidence also suggests validity (accuracy of predicting actual sales) on par or slightly better than CBC. Furthermore, ACBC interviews directly capture what percent of respondents find each attribute level to be "must have" or "unacceptable."

ACBC is a relatively new technique, and there is much more to learn regarding its range of usefulness and effectiveness. For example, the BYO + Consideration + Choice Tournament approach doesn't seem to be as useful for packaged goods research involving only a few attributes. Standard CBC and the store shelf display are considered best practice for such projects. ACBC seems more appropriate for problems involving more complex products and services with about five attributes or more.

So Which Should I Use?

You should choose a method that adequately reflects how buyers make decisions in the actual marketplace. This includes not only the competitive context, but the way in which products are described (text), displayed (multi-media or physical prototypes), and considered.

Key decision areas and how they affect choice of conjoint method are as follows:

Number of Attributes:
If you need to study many attributes (especially eight or more), ACA historically was considered a solid approach. More recently, ACBC seems more effective—especially for projects involving price as an attribute. Three or fewer attributes would favor CBC.

Mode of Interviewing:
In many cases, survey populations don't have access to computers. If your study must be administered paper-and-pencil, first consider using CBC, with CVA also being a option under conditions of very small sample size (see below).

Sample Size:
If you are dealing with relatively small sample sizes (especially less than 100), you should be cautious about using CBC, unless respondents are able to answer more than the usual number of choice tasks. ACBC and the older ratings-based approaches (such as ACA and CVA) are able to stabilize estimates using relatively smaller samples than CBC. If interviewing must be done on paper, and very small sample sizes are the norm, consider CVA.

Interview Time:
If you only have a few minutes to use in conjoint questions, CBC is a good alternative, though you may need to compensate for the limited information from each individual by sharply increasing the sample size. With about eight or more minutes available, ACBC is feasible.

Pricing Research:
If studying price, CBC and ACBC are generally preferred.