



Sawtooth Software

RESEARCH PAPER SERIES

Which Conjoint Method Should I Use?

Bryan Orme,
2003

Article originally published in Sawtooth Solutions, 1996

Which Conjoint Method Should I Use?

Bryan Orme, Sawtooth Software
Copyright Sawtooth Software, 2003

Foreword

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

It is paradoxical that many new developments in the conjoint analysis field have made the methods better than ever, but they have also made it more difficult to choose among them. 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. For a more in-depth review of the techniques, please see an excellent article presented at the 1997 Sawtooth Software Conference by Joel Huber, entitled: “What We Have Learned from 20 Years of Conjoint Research: When to Use Self-Explicated, Graded Pairs, Full Profiles or Choice Experiments” available at www.sawtoothsoftware.com.

Introduction

Conjoint analysis has become one of the most widely used quantitative tools in marketing research. When used properly, it provides reliable and useful results. There are many different 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.

Conjoint analysis comes in a variety of forms. Sawtooth Software offers a suite of conjoint software packages: Adaptive Conjoint Analysis (ACA), Traditional Full-Profile Conjoint Analysis (CVA), and Choice-based Conjoint (CBC). It makes little sense to argue which of these is the overall best approach. We have designed each package to bring unique advantages to different research situations.

Adaptive Conjoint Analysis (ACA)

The first version of ACA, released in 1985, was Sawtooth Software’s first conjoint product. ACA went on to become the most popular conjoint software tool and method in both Europe and in the US throughout the 1990s. (Shortly after the turn of the century, CBC became more widely used—more on that later.) ACA is user-friendly for the analyst and respondent alike. But ACA is not the best approach for every situation.

ACA’s main advantage is its ability to measure more attributes than is advisable with traditional full-profile conjoint. In ACA, respondents do not evaluate all attributes at the same time, which

helps solve the problem of “information overload” that plagues many full-profile studies. We believe respondents cannot effectively process more than about 6 attributes at a time in full-profile context. ACA can include up to 30 attributes, although typical ACA projects involve about 8 to 15 attributes. With six or fewer attributes, ACA’s results are similar to the full-profile approach, though there is little compelling reason to use ACA in these situations.

In terms of limitations, the foremost is that ACA must 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 exhibits another limitation with respect to pricing studies: when price is included as just one of many variables, its importance is likely to be understated, and the degree of understatement increases as the number of attributes studied increases.

ACA is a hybrid approach, combining stated evaluations about attributes and levels with conjoint pairwise comparisons. The first part of the interview approximates the self-explicated approach. Respondents rank (or rate) attribute levels, and then assign a weight (importance) to each attribute:

<p><i>Step 1, rank levels for each attribute:</i></p> <p>Rank these features from most to least preferred:</p> <ol style="list-style-type: none"> 1. Discover 2. Mastercard 3. VISA <p><i>(assume the respondent ranks VISA best and Discover worst)</i></p>	<p><i>Step 2, assign attribute importance:</i></p> <p>If two credit cards were acceptable in all other ways, how important would this difference be?</p> <p style="text-align: center;">VISA vs. Discover</p> <p>4 = Extremely Important 3 = Very Important 2 = Somewhat Important 1 = Not Important At All</p>
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

The self-explicated context puts emphasis on evaluating products in a systematic, feature-by-feature manner, rather than judging products as a whole or in a competitive context.

Using the information from the self-explicated section, ACA then presents trade-off questions. Two products are shown, and respondents indicate which is preferred, using a relative rating scale:

Which credit card would you prefer?								
Choose a number to indicate your preference								
Discover					VISA			
15% interest rate					18% interest rate			
\$5,000 credit limit					\$2,000 credit limit			
1	2	3	4	5	6	7	8	9
Strongly Prefer Left				No Preference	Strongly Prefer Right			

The product combinations are tailored to each respondent, to ensure that each is relevant and meaningfully challenging. Each of the products is displayed in partial-profile, meaning that only a subset (usually two or three) of the attributes is shown for any given question. Because of the self-explicated introductory section, the adaptive nature of the questionnaire, and the ratings-based conjoint tradeoffs, ACA is able to stabilize estimates of respondent’s preferences for more attributes using smaller sample sizes than the other conjoint methods.

Huber states that pairwise comparisons reflect the sort of purchase behavior wherein buyers compare products side-by-side. ACA does well for modeling high-involvement purchases, where respondents focus on each of a number of product attributes before making a carefully-considered decision. Purchases for low involvement product categories described on only a few attributes along with pricing research studies are probably better handled using another method.

Traditional Full-Profile Conjoint Analysis (CVA)

CVA brings full-profile conjoint to the suite of Sawtooth Software’s conjoint tools. Full-profile conjoint has been a mainstay of the conjoint community for decades now. We believe the full-profile approach is useful for measuring up to about six attributes. That number varies from project to project depending on the length of the attribute level text, the respondents’ familiarity with the category, and whether attributes are shown as prototypes or pictures. CVA is designed for paper-and-pencil studies, whereas ACA must be administered via computer. CVA can also be used for computerized interviews in a CAPI or Internet survey.

CVA calculates a set of part worths for each individual, using traditional full-profile card-sort (either ratings or ranked) or pairwise ratings. Up to 30 attributes with 15 levels can be measured, though we’d never recommend you approach these limits with a real study.

Through the use of compound attributes, in a limited way CVA can measure interactions between attributes such as brand and price. Compound attributes are created by including all combinations of levels from two or more attributes. For example, two attributes each with two levels can be combined into a single four-level attribute. However, interactions can only be

measured in a limited sense with this approach. Interactions between attributes with more than 2 or 3 levels each are probably better measured using CBC.

CVA can design pairwise conjoint questionnaires (like the ACA example above), or single-concept (card-sort) designs. Showing one product at a time encourages respondents to evaluate products individually, rather than in direct comparison with a competitive set of products. It focuses more on probing the acceptability of an offering rather than the differences between competitive products. If the comparative task is desired, CVA’s pairwise approach may be used. Another alternative is to conduct a card-sort exercise. Though respondents view one product per card, in the process of evaluating the deck they usually compare them side-by-side and in sets.

Because respondents see the products in full-profile (all attributes at once), respondents tend to use simplification strategies if faced with too much information to process. Respondents may key on two or three salient attributes and largely ignore the others. Huber points out that buyers in the real world may also simplify when facing complex decisions for certain categories, so simplification isn’t by definition always a bad thing.

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 (in full-profiles) 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

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.

Huber argues that choice tasks are more immediate and concrete than abstract rating or ranking tasks. They seem to ask respondents how they would choose now, given a set of potential offerings. Choice tasks show sets of products, and therefore mimic buying behavior in competitive contexts. Because choice-based questions show sets of products in full-profile, they encourage even more respondent simplification than traditional full-profile questions. Attributes that are important will get even greater emphasis (importance), and less important factors will receive less emphasis relative to CVA or ACA.

CBC can measure up to 30 attributes with 15 levels each, 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. There are a number of ways to analyze choice results:

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.

ICE (Individual Choice Estimation) is largely an interesting historical side note now, though some researchers still enthusiastically use this technique for individual-level estimation from CBC data. Before computers became fast enough to permit estimation of moderate to large CBC datasets under hierarchical Bayes (HB) within a reasonable amount of time, ICE offered a compelling speed advantage. Computers are now fast enough that this is usually not an issue, and in general, the industry has embraced HB.

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.

If interactions occur principally *within* individual preference structures (person *i*'s disutility for spending money depends on the brand), then explicitly modeling interaction terms may be necessary for accurate share predictions. Which approach is appropriate for your situation may be difficult to tell. 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 effectively be measured using CBC. One promising solution that is gaining momentum over the last few years 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 (relative to ACA), 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 have 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. Partial-profile CBC results tend to reflect greater discrimination between most and least important attributes relative to ACA, though it is not a given that this means improved accuracy in predicting real world choices.

One obvious question that hasn't completely been resolved is whether partial-profile CBC is subject to the same price bias as ACA. We suspect that some of the price bias in ACA is due to the partial-profile nature of the task. But, a few split-sample studies comparing partial- and full-profile CBC suggest the price bias is not a problem for partial-profile CBC. Some researchers approach the problem by always including price in each task, and randomly rotating in the

remaining attributes. It is not known whether this results in more *accurate* price importance (elasticity) estimates, though it certainly increases the *precision* of the price estimates.

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. Is the product a high-involvement category for which respondents deliberate carefully on all of the features, or should the conjoint task encourage simplification?

If you need to study many attributes, ACA or possibly partial-profile CBC should be considered. If you need to include attribute interactions in your models, you should probably use CBC. In many cases, survey populations don't have access to PCs, and it may be too expensive to bring PCs to them, or vice-versa. If your study must be administered paper-and-pencil, consider using CVA or CBC with its paper-and-pencil module.

If you are dealing with relatively small sample sizes, you should be cautious about using CBC, unless respondents are able to answer more than the usual number of choice tasks. ACA and CVA are able to stabilize estimates using relatively smaller samples than CBC.

Many researchers include more than one conjoint method in their surveys. For example, some studies need to measure a dozen or more attributes, and also require brand-specific demand curves. ACA followed by CBC can solve this problem within a single questionnaire. ACA would include all the attributes, while brand, price, and a few key performance variables would be studied using CBC. ACA provides the product design and feature importance model, while CBC provides price sensitivity estimates for each brand and a powerful pricing simulator. Another approach is to include holdout choice questions involving price, and adjusting the ACA price estimates to best fit the holdout choices. (Please see an article entitled "Calibrating Price in ACA: The ACA Price Effect and How to Manage It" in the technical papers library at www.sawtoothsoftware.com.)

For some projects, it may be difficult to decide on which method to use. With the introduction of HB and partial-profile CBC, the lines which have defined the distinct capabilities of conjoint methods have become blurred. If this ambiguity still vexes you, it is comforting to recognize that the methods, though different in their approach, tend to give similar overall results.