Sawtooth Software

RESEARCH PAPER SERIES

A Perspective on Adaptive CBC (What Can We Expect from Respondents?)

Rich Johnson,
Sawtooth Software, Inc.

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

A Perspective on Adaptive CBC

(What Can We Expect from Respondents?)

Rich Johnson
Sawtooth Software, Inc.
July, 2008

Since the earliest days of Conjoint Analysis, there has been tension between our hopes of what we can get from sophisticated models and the reality of what respondents are willing or able to tell us in questionnaires.

Before conjoint analysis, the demands on respondents were usually modest. Most market research consisted of simple product tests or attitude measurement with rating scales. Neither of those kinds of questions really taxed respondents' abilities to provide meaningful answers. Of course, questionnaires could be long, and respondents could certainly have problems with fatigue and motivation; but individual questions were not hard to answer.

The Demands of Conjoint Analysis

With the invention of conjoint analysis, the experience of the respondent changed for the worse. Questionnaires became longer and more repetitive, and questions became more difficult. One might expect researchers to have given more attention to the limitations of conjoint respondents, but we have often just assumed respondents would be able and willing to do as we ask. Here are some brief examples.

> In the early '70s some of us used "Tradeoff Matrices." Indeed, for respondents who could think abstractly and who never got tired this could have been a wonderful method of obtaining conjoint information. Unfortunately, most respondents realized that it was much easier to decide whether the row attribute or the column attribute was more important, and having decided that, they responded trivially, filling in the matrix with a pattern of row or column dominance. Trade Off Analysis seemed to work despite this limitation, but it's clear that *respondents were giving us less information than we had expected.*
>
> One of the loveliest creations of the past several decades was Full Profile Conjoint Analysis, as proposed and first used by Paul Green and his colleagues. A collection of a few dozen product concepts was designed so that by merely rank-ordering them for preference, a respondent would reveal the value he placed on each level of each attribute. Everyone should be able to produce a preference order! Yet, results from respondents suggested that when there were more than about six attributes, the task became so complex that many of them found ways to simplify their approach, perhaps paying attention to only a few most important attributes. Again, *respondents gave us less than we expected*. Recognizing how difficult it can be for respondents to consider many attributes simultaneously, full profile analysis is now usually avoided for all but the smallest problems.

In these early conjoint applications, researchers were too optimistic about how respondents would actually perform in interviews. We gave respondents the capability of informing us in great detail about their needs and desires, but they often chose ways of responding that were less informative, though much easier.

Compensatory Models

Throughout the "Conjoint Era" we have probably been getting less information than we have believed from respondents. Nearly all conjoint work uses a model that assumes respondents behave compensatorily. That is, respondents prefer products with greatest utility, and utility is obtained by adding up values associated with components of each product. This is a very general model, and subsumes other more specific ones:

> If a respondent is only paying attention to one attribute, the corresponding compensatory model simply has partworths of zero for other attributes.
>
> If a respondent behaves lexicographically, the compensatory model has more extreme partworths for each attribute than for those with less impact.

Because the compensatory model fits so many types of behavior, it is not very informative about what kind of behavior is occurring. The fact that it fits the data does not imply that respondents are in fact behaving compensatorily!

Consider the differences between a conjoint questionnaire and an actual purchase environment. The conjoint respondent is surely less motivated than the product purchaser to give careful consideration to all aspects of the decision. Respondents may bring a certain level of mental energy that they are willing to spend within a certain time they are willing to give. When the demands of the questionnaire exceed those limits, respondents (wanting to please the researcher by providing the requested completed record, or wanting to receive the reward for completing the questions) may employ shortcuts that they would not employ if facing the real-world decision, with real-world consequences for their choices.

In recent years there has been a shift from ratings-based to choice-based conjoint analysis. There are many reasons for this shift. One important reason is that choice tasks seem similar to what a buyer does when deciding which product to purchase. Another reason is that choice tasks are easy for respondents, and virtually everyone should be able to make meaningful choices. But perhaps those choices are too easy! As reported in several recent Sawtooth Software conference presentations, respondents answer choice questions so rapidly that they can't conceivably be weighing all the attributes and levels. Also, Gilbride and Allenby, as well as Hauser, Dahan, Yee, and Orlin, showed that choice results for most respondents they studied could be accounted for by non-compensatory processes in which they are assumed to pay attention to at most a few attribute levels.

If you have any doubts about limitations on how respondents process information in choice tasks, I invite you to introspect. As a researcher who has fielded choice-based conjoint studies, you have probably pre-tested many questionnaires. Think about how you do that. Do you use a compensatory model? Or do you make decisions based on a few critical attributes and levels?

It seems almost certain that most respondents do use simplification strategies to make their tasks easier in questionnaires, but that they would be less likely to use such short cuts in real purchase decisions where the stakes were higher for them. Thus, although choice tasks may be similar to purchase situations in some ways, they might be quite different in terms of the way people approach the corresponding decisions.

The Lure of Minimal Overlap

"Minimal Overlap" is a term coined by Huber and Zwerina to describe a characteristic of efficient choice designs. The logit model can estimate main effects most efficiently if the concepts in a choice set differ as much as possible. For example, suppose there are three concepts in a choice set, and also that three brands are being investigated. In the most efficient design, every choice set will contain all three brands. One can generalize this characteristic by saying that in the most efficient designs, attribute levels within a choice set will not be duplicated more than necessary.

Minimal overlap is intuitively appealing, and its use has led to efficient questionnaires, so long as we assume that responses are made according to a logit rule. However, minimal overlap can be a bad idea if respondents answer by always looking for the best level of a particular attribute. For example, suppose a respondent decides he will always select a particular brand. If there are as many brands as alternatives, and if the questionnaire design has minimal overlap, then even with such trivial processing, the respondent could be completely consistent while expending almost no effort, and his partworths would perfectly predict his answers to holdout concepts.

Questionnaire design with minimal overlap is a good idea if respondents behave according to a logit rule. But it is not such a good idea if they use simplification strategies, such as always choosing a preferred brand, or the lowest price, or the most extreme level of some other attribute. If respondents are not behaving according to the logit rule, then minimal overlap exposes us to the risk of overstating the predictive ability of our conjoint data.

Following Hauser, Dahan,Yee, & Orlin's presentation at our 2006 conference, we reexamined recent data sets. For a typical data set we found that 80 percent of respondents had answers that could be accounted for by the presence or absence of only three or fewer attribute levels! This finding, coupled with other information reviewed above, led to the belief that we needed a different way of collecting conjoint data which would encourage, and perhaps even require, respondents to provide more information in conjoint interviews.

Why Had Our Previous Efforts at ACBC Failed?

For the past several years we at Sawtooth Software had been attempting to produce more efficient CBC designs, by using information obtained early in the interview to choose subsequent choice sets so as to be most efficient. Although there were some early indications of success, our general conclusion is that our efforts failed. We were indeed able to produce designs that were more statistically efficient according to the logit model, but they were not better at predicting respondent behavior in holdout choices.

That work ***assumed*** that respondents behaved compensatorily, and framed questions which would provide efficient estimation under the logit model. It seems likely that we failed because that assumption was false. It now seems likely that most respondents ***do not*** use logit models in answering choice questions! If they are not using logit models, then statistical efficiency under the logit model may be irrelevant in attempts to understand their choices.

So, What Next?

When we saw the Hauser,Yee, & Orlin results, we first considered providing a similar kind of analysis as an alternative for CBC data. However, we concluded that what was needed was not a further kind of analysis of CBC data, but rather a new kind of data. It is interesting to compare CBC with classic Full-Profile based on sorting several dozen product concepts. Full Profile Conjoint contains much more information that could be used to determine whether respondents are behaving in non-compensatory ways. With "regular" CBC we obtain so little data from each respondent that we have very little power to distinguish between compensatory and non-compensatory behavior.

One of the supposed benefits of choice-based conjoint is that choice tasks are similar in some ways to actual real-world choice situations. Although respondents do seem to use simplification strategies when filling out questionnaires, they probably work harder when making important real-life choices. So simplification in answering questionnaires is not a good thing. We learn less than we might if we pushed respondents to use deeper processing. We need to find ways to do that.

One way respondents may approach complex choice tasks is first to eliminate alternatives with unacceptable attribute levels, and then to choose among the remaining alternatives using a more refined method, such as a logit rule. Our early work with ACA did something like this, first asking each respondent to specify all attribute levels that would be "totally unacceptable," and then constructing a conjoint questionnaire to assess the remaining levels. However, we soon found this approach to be flawed. Most respondents were too quick to specify attribute levels as unacceptable. Too many levels were eliminated, and sometimes it was impossible to configure realistic products not containing unacceptable levels. Since partworths for unacceptable levels were coded as large negative constants, in some conjoint simulations the winning product was just the one with fewest unacceptable levels.

Unlike our previous attempts at Adaptive CBC, our current approach does not focus solely on producing designs that are efficient under a logit model. Rather, it acknowledges that respondents may first screen concepts on critical features, eliminate those that would be unacceptable, and then use a more refined method, such as a logit rule, to choose among those that remain.

As formerly with ACA, we identify unacceptable levels and remove them from future questions. However, we do not ask respondents to list the levels that would be unacceptable. Rather, we observe respondent behavior in assessing the first few concepts they are offered, and if we find that they behave as though some levels are unacceptable, we give them the opportunity to confirm that.

Thus, toward the end of the questionnaire, the respondent is asked to choose among concepts which should all be acceptable. In this way we may obtain deeper information about the respondent's values, which should be helpful in predicting real-life choices where respondents probably work harder to process information than they would in a "regular" CBC questionnaire.

We have compared Adaptive CBC with "regular" CBC in several studies, observing similar conclusions each time. In each study we gave respondents several holdout tasks, and then a final task composed of the alternatives that had been chosen in the previous tasks. Assuming that respondents had used simplification strategies to answer the first few holdout tasks, those strategies would not work in the final task, in which all alternatives would pass any simplified screening criteria. Adaptive CBC was dramatically more successful than "regular" CBC in predicting choices in the final holdout tasks. Presumably this is because ACBC requires respondents to use deeper processing than is required by "regular" CBC.

It is too early to tell whether this approach produces better predictions in the real world, outside of the conjoint interview. But all the signs are positive. Respondents report that they are better able to express their preferences. We are eager to put ACBC into the hands of researchers, and expect soon to have additional information about external validity.

References

Gilbride, Timothy and Greg M. Allenby (2004), "A Choice Model with Conjunctive, Disjunctive, and Compensatory Screening Rules," Marketing Science, 23, 3, (Summer) 391-406.

Hauser, John R., Ely Dahan, Michael Yee, and James Orlin (2006) "'Must Have' Aspects vs. Tradeoff Aspects in Models of Customer Decisions," Sawtooth Software Conference Proceedings, 169-181.

Huber, Joel and Klaus Zwerina (1996), "The Importance of Utility Balance in Efficient Choice Designs," Journal of Marketing Research, 33 (August) 307-317.