Sawtooth Software

RESEARCH PAPER SERIES

Conducting Full-Profile Conjoint Analysis over the Internet

Bryan K. Orme
and W. Christopher King,
Sawtooth Software, Inc.
1998

Conducting Full-Profile Conjoint Analysis over the Internet

Bryan K. Orme and W. Christopher King, Sawtooth Software, Inc.

Embracing the Web

The advent of the WWW (World Wide Web) is changing the way we communicate in business. Over the past twenty years, a similar impact was felt with personal computers and software, overnight delivery services, fax machines, e-mail, and voice mail/answering machines. The WWW is building on the strengths of these advances.

The growth in Internet usage is truly astounding. According to IntelliQuest, Inc. of Austin, TX, as of Q1 1998, 32% of the U.S. population age 16 and older (or 66.5 M individuals) is online. In the period of only a year (Q4 1996 to Q4 1997), the number of Internet users in the U.S. grew by 32%. And if projections hold, 38% of the U.S. population age 16 and older (or 78.4 M individuals) will be online as of Q3 1998 (IntelliQuest 1998).

As market researchers begin to use the Internet to conduct surveys, they shouldn't feel completely disoriented. Internet surveys share much in common with traditional computerized surveys and we collectively have nearly two decades of experience in that area. The trick is to leverage what we've already learned about computer interviewing and computerized conjoint surveys and apply it to this new and exciting medium.

This article is organized in two parts. First, we'll cover general WWW survey research issues, and then we'll report on an online full-profile conjoint survey conducted over the web dealing with credit card preferences.

Computer Interviewing: Historic Perspective

Until recently, the WWW had been largely experimental in the marketing research industry. Control and access were primitive, limiting the kind of information one could collect. We see many parallels between early web research and what was felt when computerized interviewing first appeared in the '70s.

The first computerized interviewing was done using terminals connected to large computers in the mid '70s. Later, Dr. Richard M. Johnson, Chairman of Sawtooth Software, pioneered PC-based interviewing in 1979 using Apple II computers. He found that he could customize each interview not just with programmed skip patterns but using adaptive heuristics to formulate efficient preference questions for collecting conjoint data. The computer would "learn" about a respondent's preferences and customize each interview to focus on the most important attributes. In 1985, Sawtooth Software released Ci2 (Computer Interviewing) and ACA (Adaptive Conjoint Analysis) for the IBM PC to the marketing research community.

There are many anecdotal stories from those early days, including the following told by Richard Johnson at our 1992 Sawtooth Software Conference:

"I still have vivid memories of one of the first computer interviewing studies. We swallowed hard and purchased a dozen Apple II computers, which arrived a day before they had to be sent to Atlanta, where the field work was to be done. Half of the computers were dead on arrival. By combining parts of the various machines, we got nine of them to function.

"Those early machines had severe reliability problems, not the least of which was their sensitivity to static electricity. Respondents sometimes scuffed their feet on the carpet, and that caused real problems. When an interviewer would call in complaining of machine problems, our first response was to say 'Have the respondent take off her shoes.' That often seemed to work." (Johnson, 1992)

Widespread use of Disk-By-Mail (DBM) was still many years in the future when PCs became commonplace in businesses and homes. Today we face similar issues and opportunities with the Internet. Fortunately, advances in software and the booming popularity of the Internet means that WWW interviewing is rapidly becoming practical and feasible as an additional tool for the market researcher.

Using the WWW to collect market research data consists of two modalities: surveys that are e-mailed, and online surveys.

E-mail Surveys

The text-based e-mail survey is perhaps the easiest method for conducting marketing research surveys on the web. Respondents type answers into pre-specified blanks with their e-mail editor or word processor, and return the completed form to the sender.

Text-based e-mail survey pros:

- Low cost: quick and easy to put together.

Text-based e-mail survey cons:

- Lots of data cleaning.
- Respondents may delete part of the survey with their word processor.
- Questionnaires are not very attractive: no graphics, font control or colors.
- Respondent sees all questions at once: no automatic skip patterns.

The second form of e-mail survey involves a program executable (usually in a zipped file) which respondents install on their computers. The data file is e-mailed back to the sender.

E-mailed survey executable pros:

- Control of skip patterns and data entry verification.
- Attractive surveys, including graphics, font control and colors.

E-mailed survey executable cons:

- Many users fear installing software e-mailed to them.
- Installation can be time-consuming: best for computer literate respondents.
- Software compatibility across different computers   on some computers it may not work at all.

Online Surveys

The other form of web-based survey is the online survey: respondents connect directly to the web site, which displays the questionnaire. Online surveys can be formatted as a single form (page). The respondent scrolls down the page from question to question, then clicks the submit button to send the information to a server.

Single-form online survey pros:

- Only a single download required at connection and a single upload when the form is completed.
- Relatively inexpensive to program and administer.
- Attractive surveys, including graphics, font control and colors.

Single-form online survey cons:

- No automatic skip logic.
- Data verification only possible at end of survey.
- Long forms can seem overwhelming and may not be completed.
- Long download time if survey is long, includes complex graphics, and/or your connection is slow.
- An entire interview might be lost if the computer, modem or net connection fails.
- Respondents cannot complete part of the form, terminate, and restart at a later time without losing all their work.

The second type of online WWW survey is the multi-form survey. Questions are presented on different pages (forms), and the data are saved when the respondent clicks the submit button at the bottom of each page.

Multi-page online survey pros:

- Permits skip logic and question-specific data verification.
- User doesn't face entire task at once.
- Attractive surveys, including graphics, font control and colors.

Multi-page online survey cons:

- Complex to program without the aid of WWW survey software.
- Delay between pages if you have a slow connection or your server has limited bandwidth.

Using Passwords to Control Access to Your Web Survey

It is usually critical with web-based surveys to limit access to your survey. Assigning passwords prevents unauthorized access to your survey and "ballot stuffing." Benefits also include control over quota cells and restarting of incomplete interviews.

Software Compatibility and Availability

Incompatibility among browsers and servers remains a major software issue. With the introduction of the Java programming language and Visual Basic (VB) scripting, additional functionality can be added to online surveys that far exceeds the restrictions of HTML. Unfortunately, Java standards are still elusive and VB is not supported by all browsers. Very little is common on the server side, and some software must be customized for each server configuration.

But, all is not hopeless. New PC-based software makes it possible to construct, administer, and host your own survey on either your own web server, your ISP s (Information Service Provider) server, or the server belonging to the manufacturer of the survey software. The advantage of hosting your own site or using an ISP is that you have control over the study. You also avoid the per-interview costs that are frequently associated with hosting on someone else's marketing research site. It also means that you can easily test your questionnaire, add questions while a study is in progress, and monitor its progress online.

Is the WWW Appropriate for Your Research?

Much has been said about the representativeness of data collected over the Internet. We trust you have studied the arguments to determine that the Internet is the right vehicle for your research study. We won't spend time addressing these arguments, but will proceed under the assumption that the Internet is appropriate for your research study.

We'll now focus our attention on conducting full-profile conjoint analysis on the Internet.

Conjoint Analysis Usage

In a 1997 survey of conjoint analysis usage in the marketing research industry, ACA (Adaptive Conjoint Analysis) was found to be the most widely used conjoint methodology in both the US and Europe (Vriens, Huber and Wittink, 1997). Traditional full-profile (FP) conjoint was also reported as a popular method. In general, we believe traditional FP conjoint is an excellent approach when the number of attributes is around six or fewer, while ACA is generally preferred for larger problems.

Paper vs. Computerized Full-Profile (FP) Conjoint

FP conjoint analysis studies can be done either as paper-based or as computerized surveys (Internet surveys, disk-by-mail, or CAPI). Because they typically involve fixed designs and, unlike ACA, are not adaptive, computerized FP surveys really offer no real benefit over the paper-based approach in terms of the reliability or validity of the results. In fact, paper-based FP may work *better* than computerized FP. With traditional paper-based card-sort, respondents can examine many cards at the same time, comparing and manipulating them into piles. This helps respondents learn the range of possibilities and settle on a reliable response strategy. With computerized approaches, respondents see only one isolated question at a time. It may take a few questions for respondents to learn about the range of possibilities and settle on a reliable response strategy. It is probably beneficial with computerized FP, therefore, to show the best and worst profiles early on in the survey.

Even though computerized FP probably offers no significant benefit over paper-based surveys in terms of reliability or validity, real benefits might be realized in survey development and data collection costs.

Pairwise versus Single-Concept Approach

Pairwise and Single-Concept presentation are two popular approaches for FP conjoint. A Pairwise FP conjoint question administered over the Internet looks like:
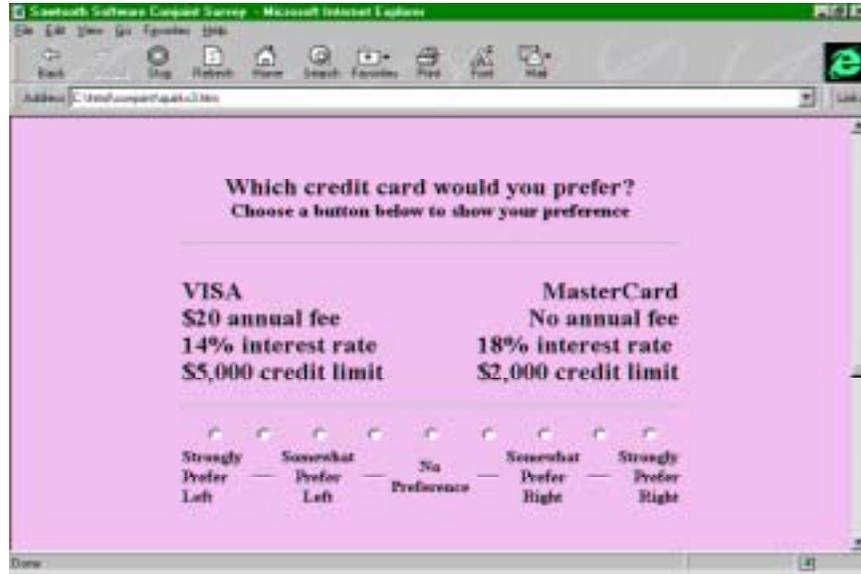
And the Single-Concept approach looks like:
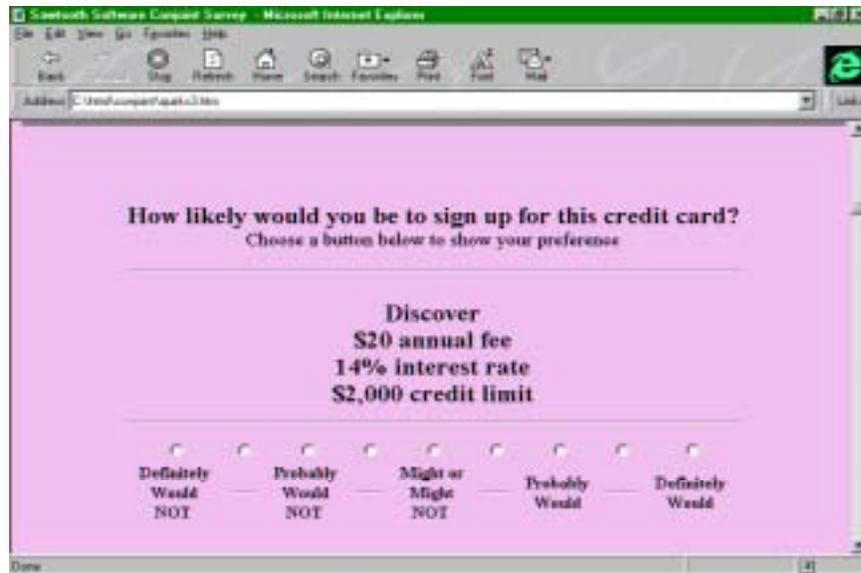
With Pairwise questions, respondents make comparative judgements regarding the relative acceptability of competing products. The Single-Concept approach probes the acceptability of a product, and de-emphasizes the competitive context. Both methods have proven to work well in practice, but we are unaware of any study other than this one that has directly compared these

two approaches.

Purchase likelihood ratings reflect the *absolute* desirability of product profiles. With Pairwise ratings, we only gain *relative* information. This potentially can be a critical distinction, depending upon the aim of the research. Consider the person who takes a Pairwise conjoint interview designed to find the optimal blend for tofu. The conjoint utilities might appear reasonable, even though he finds tofu disgusting and has absolutely no desire to ever buy it. If we use Single-Concept profiles, we can both derive utilities and learn about a respondent's overall interest in the category. Respondents who have no desire to purchase can be given less weight in simulations, or be thrown out of the data set entirely. The danger with Single-Concept ratings is that if a person gives most of the profiles the lowest (or highest) rating, there is limited variation in the dependent variable, and we may not be able to estimate very stable utilities.

One need not give up the benefit of measuring purchase likelihood when using the Pairwise approach. Both Pairwise and Single-Concept conjoint questions can be included in the same survey. Single-Concept purchase likelihood questions could be used to calibrate (scale) Pairwise utilities (as is done in ACA). We can get the benefit of the comparative emphasis of Pairwise questions while including information on purchase intent.

Details of Experiment

We designed an Internet survey to compare the Pairwise and Single-Concept approach for computerized FP conjoint analysis.

The subject for our study was credit cards, with the following attribute levels:

Brand	Annual Fee	Interest Rate	Credit Limit
VISA	No annual fee	10% interest rate	$5,000 credit limit
MasterCard	$20 annual fee	14% interest rate	$2,000 credit limit
Discover	$40 annual fee	18% interest rate	$1,000 credit limit

Respondents completed both Pairwise and Single-Concept conjoint questions (in rotated order). Enough conjoint questions (nine) were included to estimate utilities (12 part-worths) for both the Pairwise and Single-Concept designs at the individual level. These designs had only one degree of freedom. In general, we would not recommend conjoint designs with so few observations relative to estimated parameters. For the purposes of our methodological study (respondents were required to complete both designs in the same interview) these saturated designs seemed satisfactory. Additionally, holdout choice sets were administered both before and after the traditional conjoint questions.

A total of 280 respondents completed the survey. Respondents self-selected themselves for the survey, which was launched from a hyperlink on Sawtooth Software's home page. This sampling strategy is admittedly poor had we been interested in collecting a representative

sample. But the purpose of our study was not to achieve outwardly projectable results, but rather to compare the within-respondent reliability of alternative approaches to asking FP computerized conjoint.

We took three steps to help ensure the quality of our data: 1) we required respondents to give their name and telephone number for follow-up verification, 2) we included repeated holdout choice tasks for measuring reliability and flagging "suspect" respondents, and 3) we examined the data for obvious patterned responses.

Measuring the Reliability of Conjoint Methods

Reliability and validity are two terms often used to characterize response scales or measurement methods. *Reliability* refers to getting a consistent result in repeated trials. *Validity* refers to achieving an accurate prediction. Our study focuses only on issues of reliability.

Holdout conjoint (or choice) tasks are a common way to measure reliability in conjoint studies. We call them "holdout" tasks because we don't use them for estimating utilities. We use holdouts to check how well conjoint utilities can predict answers to observations not used in utility estimation. If we ask some of the holdout tasks twice (at different points in the interview), we also gain a measure of test-retest reliability.

We included a total of three different holdout choice questions in our Internet survey, which looked like:
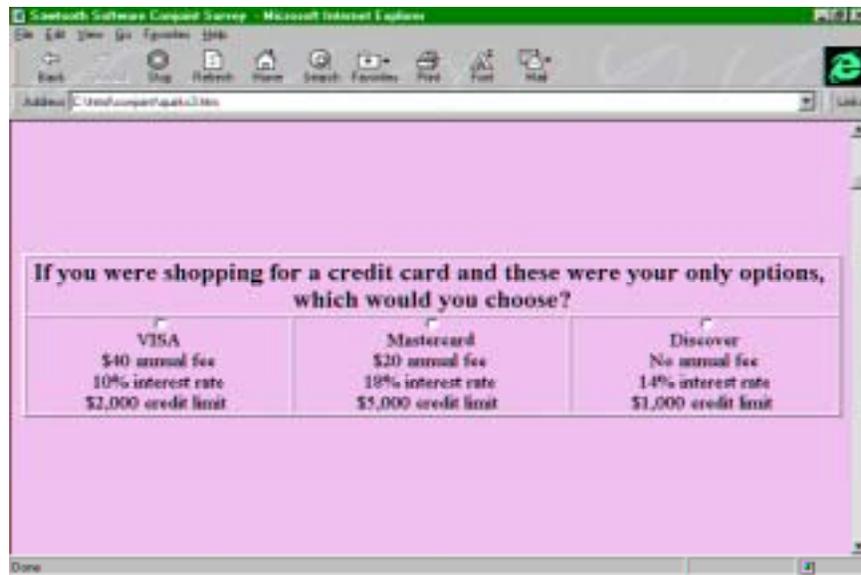
These questions came at the beginning of the interview, and then the same ones (after rotating the product concepts within set) were repeated at the end of the survey. Respondents on average answered these holdouts the same way 83.0% of the time. This test-retest reliability is in line with those reported for other methodological studies we've seen that were not collected over the Internet. But one can argue that our respondents (marketing and market research professionals) were a well-educated and careful group. We cannot conclude from our study that Internet interviewing is as reliable as other methods of data collection.

We use the holdout choice tasks to test the reliability of our conjoint utilities. We would hope that the conjoint utilities can accurately predict answers to the holdout questions. We call the percent of correct predictions the *holdout hit rate*. Some have referred to hit rates as a validity measurement, but prediction of holdout concepts asked in the same conjoint interview probably say more about reliability than validity.

Comparing different conjoint methods using holdouts will usually favor the conjoint method that most resembles the holdouts. The comparative nature of the Pairwise approach seems to more closely resemble the choice tasks (showing three concepts at a time) than does Single-Concept presentation.

Holdout predictions are not the only way to measure reliability. We can also examine whether part-worth utilities conform to *a priori* expectations. Three of the attributes (annual fee, interest rate, and credit limit) were ordered attributes (i.e. low interest rates are preferred to high interest rates). When part-worth utilities violate known relationships, we refer to these as *reversals*.

Reliability of Pairwise Versus Single-Concept Approach

The holdout hit rates for the Pairwise and Single-Concept approach were 79.3% and 79.7%, respectively. This is a virtual tie; the difference is not statistically significant. These findings suggest that both methods perform equally well in predicting holdout choice sets.

The average number of reversals per respondent was 1.5 and 1.3 for Pairwise and Single-Concept designs, respectively. The difference was significant at the 90% confidence level. These findings suggest that utilities from Pairs questions may contain a bit more noise than Singles. The difference was small, however, and we caution drawing general conclusions without more corroborating evidence.

Qualitative Evidence

In addition to completing conjoint tasks, we asked for qualitative evaluations of the Pairwise versus the Single-Concept approach. Respondents perceived that the Pairwise questions took only 13% longer than the Singles. We asked a battery of questions such as whether respondents felt the conjoint questions were enjoyable, easy, frustrating, or whether the questions asked about too many features at once. We found no significant differences between any of the qualitative

9

dimensions for Pairwise vs. Single-Concept presentation.

Conjoint Importances and Utilities

We calculated attribute importances in the standard way, by percentaging the differences between the best and worst levels for each attribute. Conjoint importances describe how much impact each attribute has on the purchase decision, given the range of levels we specified for the attributes.

We constrained the utilities to conform to *a priori* order for annual fee, interest rate and credit limit. Further, we scaled the conjoint utilities (at the individual level) so that the worst level was equal to zero, and the sum of the utility points across all attributes was equal to 400 (the number of attributes times 100). Importances were computed at the individual-level, then aggregated.

Importances and utilities for Pairs vs. Single-Concept presentation were as follows:

Conjoint Importances

	Pairs	Single-Concept
Brand	18%	19%
Annual fee	37%	37%
Interest rate	21%	20%
Credit limit	24%	24%

Conjoint Utilities

	Pairs	Single-Concept
VISA	36	38
MasterCard	27	31
Discover	13	12
No annual fee	104	104
$20 annual fee *	44	34
$40 annual fee	0	0
10% interest rate	55	55
14% interest rate	30	30
18% interest rate	0	0
$5,000 credit limit	64	67
$2,000 credit limit	27	29
$1,000 credit limit	0	0

* statistically significant difference at 99% confidence level

The only significant difference for either conjoint importances or utilities between the two full-profile methods occurred in the utility for the middle level of annual fee ($20). In a presentation at our 1997 Sawtooth Software Conference, Joel Huber of Duke University argued that respondents may adopt different response strategies for sets of products versus Single-Concept presentation. He argued that when faced with comparisons, respondents may simplify the task by avoiding products with particularly bad levels of attributes. Annual fee was the most important attribute. The larger gap between the worst and middle level (44-0) for Pairs versus Single-Concept (34-0) is statistically significant at the 99% confidence level (t=3.93) and supports Huber's "undesirable levels avoidance" hypothesis.

Pairwise Versus Single-Concept FP Conjoint: Conclusions and Suggestions

Our data tell a comforting story, suggesting that both computerized Pairwise and Single-Concept FP ratings-based conjoint are equally reliable and result in the same importances and roughly the same utilities. Computerized FP conjoint seems to have worked well for a small design such as our credit card study. Given that the researcher has determined that the Internet is an appropriate vehicle for interviewing a given population, our findings suggest that FP conjoint can be successfully implemented via the Internet for a small study including four attributes.

References

Huber, Joel (1997), "What We Have Learned from 20 Years of Conjoint Research: When to Use Self-Explicated, Graded Pairs, Full Profiles, or Choice Experiments" *Sawtooth Software Conference Proceedings*, 243-56.

IntelliQuest Information Group, Inc. (1998), Worldwide Internet/Online Tracking Study (WWITS).

Johnson, Richard M. (1992), "Ci3: Evolution & Introduction" *Sawtooth Software Conference Proceedings*, 91-102.

Vriens, Marco, Joel Huber, and Dick R. Wittink (1997), "The Commercial Use of Conjoint in North America and Europe: Preferences, Choices, and Self-Explicated Data," working paper in preparation.