A look inside the choice-modeling toolbox

| By Michael Lieberman

snapshot

An overview of five common choice models employed in marketing research.

Consumers do not usually make purchase decisions based on one single condition. More often, consumers examine a range of features or attributes and then make judgments or trade-offs to determine their final purchase choice. Choice modeling is a market research tool that outlines this decision process. Within the context of a product or brand, choice modeling incorporates attributes such as cost, prestige and environmental impact to predict purchase decisions of individuals or market segments.

There are many statistical methods that can be used to perform this analysis. The main challenge for market researchers is determining which one is most appropriate in a given situation. This article will give an overview of the five most common choice models employed in market research, describe when to use them and outline their advantages and disadvantages.

Paired-comparison analysis

Paired-comparison analysis is the most basic type of choice model. Essentially, a respondent sees two choices and then determines which one he prefers. He then sees two more and the exercise repeats.

Paired-comparison is useful when there are small numbers of products or brands to compare – five or fewer. Comparisons contain no outside attribute information, such as prices, which limits the analysis.

Paired-comparison analysis helps the researcher work out the importance of a number of options relative to one another. It also helps the researcher set priorities where there are conflicting demands on resources.

For example, a major medical foundation is choosing between several different projects that are asking for funding. To maximize impact, it only wants to contribute to a few of these and the board of the foundation has been given the following four options.

A: End-of-life closure
B: Health careers futures
C: HIV/AIDS
D: Regional health initiative

With four initiatives, looking at two at a time, the maximum number of comparisons is six. The paired-comparison chart is shown in Table 1. The letter represents the choice of which initiative the board preferred and the number represents the strength of the preference (1 = prefer, 2 = prefer strongly).

Table 1

Service Category Questions	End-of-Life Closure	Health Careers Futures	HIV/AIDS	Regional Health Initiative
End-of-Life Closure		A,2	C,1	A,1
Health Careers Futures			C,2	B,2
HIV/AIDS				C,2
Regional Health Initiative				

Finally, they add up the A, B, C and D values and convert each into a percentage of the total. These calculations yield these percent-

quirks.com/articles
ID 20130205

ages: end-of-life closure – 30 percent; health careers futures – 20 percent; HIV/AIDS – 50 percent; regional health initiative – 0 percent. The board has made its choice. It will support HIV/AIDS and end-of-life closure.

Conjoint analysis

Conjoint analysis is useful in shaping new products, determining maximum levels of product enhancement and predicting market share. It works best when assessing a product that has a maximum of six attributes. In this method, the researcher is given a list of all attributes and all levels of these attributes. From this, he creates a computer-generated design referred to as an orthogonal model.

Respondents are asked to rank various product qualities, followed by a series of product purchase interest questions. Running the data yields utility scores which allow the researcher to accurately simulate the marketplace in great detail.

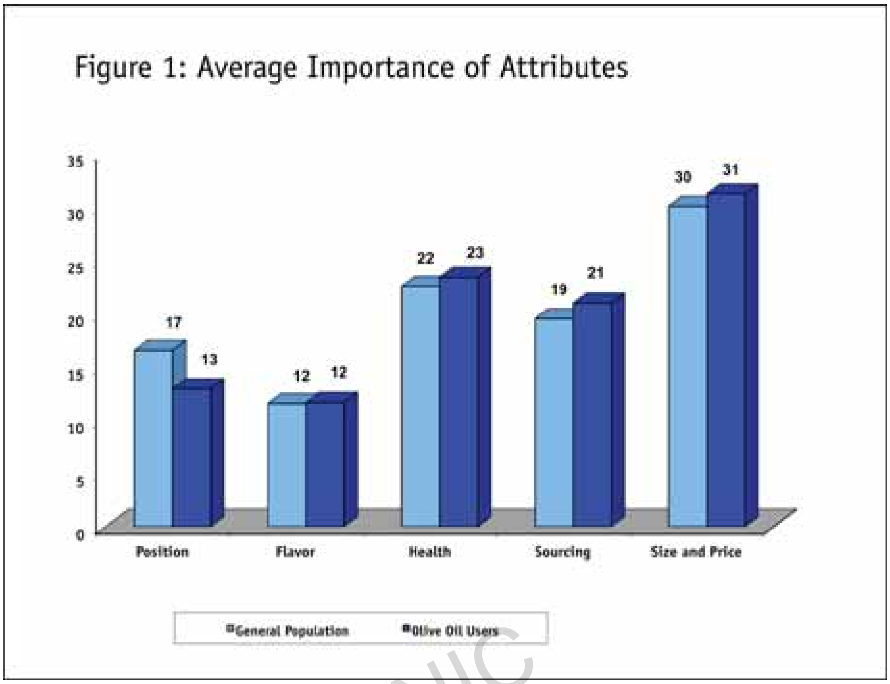
Below is an example of a conjoint scenario:

On a 1-to-5 point scale, how likely are you to purchase this olive oil with the following features?

- Premium, select-quality olive oil
- Full-bodied olive oil, adding a rich layer of flavor
- Contributes to a healthful cholesterol ratio
- Select, high-quality olive oils from California
- 16 oz., $5.89

Using the same 1-to-5 scale, we would then ask the question using a different combination of attributes. Conjoint analysis allows the researcher to examine the relative importance of price against the other factors in the model. In Figure 1 we see that while price has the largest share of attribute importance, the other factors are also significant.

The next step in the conjoint analysis process is the development of a simulator to model any combination of factors. This will allow the researcher to see purchase intent for scenarios not shown in the survey – and therefore facilitate the design of the optimal product. A respondent might see only nine scenarios but with the conjoint output, we can evaluate 240 different product levels.



A limit to conjoint analysis is that it can only model combinations of levels of attributes that are included in the study. This can be a limitation if price is an attribute; conjoint is unable to model prices that are not shown.

Discrete choice modeling (also referred to as choice-based conjoint)

As the name might suggest, this choice technique best measures distinct choices – in other words, best for products or brands that already exist.

Discrete choice analysis consists of a series of questions that ask respondents to choose between three or more hypothetical products or services at different price levels. The model simulates future market states to support product and price-level decisions.

A well-constructed discrete choice model: optimizes price or brand positions within existing market realities; takes into account "non-purchase"; gives customers real-world choice by including competitive brands at different prices; and can target specific competitors with products designed to take share specifically from them.

For example, Bart's Bait Company wants to introduce a new bait into its local market. With discrete choice, it will be able to project its market share among its chief competitors. Bart specifies the competitors and a range of prices. Below are two sample scenarios:

Scenario 1
Please choose *one* of the following:
1. Bart's Skinny Chunk at the price of $2.39
2. Zoom Fat Albert Twin Tail at the price of $2.19
3. E-Bait Big Salty Chunk at the price of $2.39
4. Bracken Bait's Big Critter Craw at the price of $1.89
5. None of the above

Scenario 2
Please choose *one* of the following:
1. Bart's Skinny Chunk at the price of $2.39
2. Zoom Fat Albert Twin Tail at the price of $2.39
3. E-Bait Big Salty Chunk at the price of $1.89
4. Bracken Bait's Big Critter Craw at the price of $2.19
5. None of the above

Please notice that the only changing attribute on these two scenarios is the price. The brands remain constant.

After running the model using logistic regression, we then create a simulator, which allows Bart to plug in prices for its Skinny Chunk as well as for the three other competitors in the market.

The baseline output is shown in Figure 2: if all bait sold at a middle price of $2.19 (median market condi-

Figure 2: Discrete Choice – Bart's Bait – Skinny Chunk Market Entry
Initial Price Findings – All Brands at $2.19

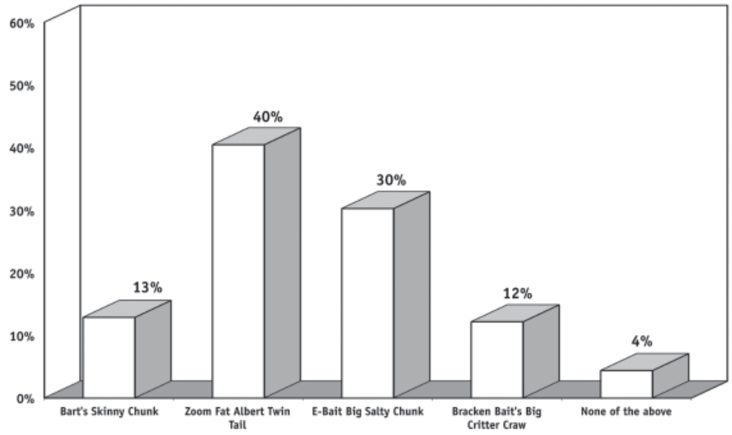
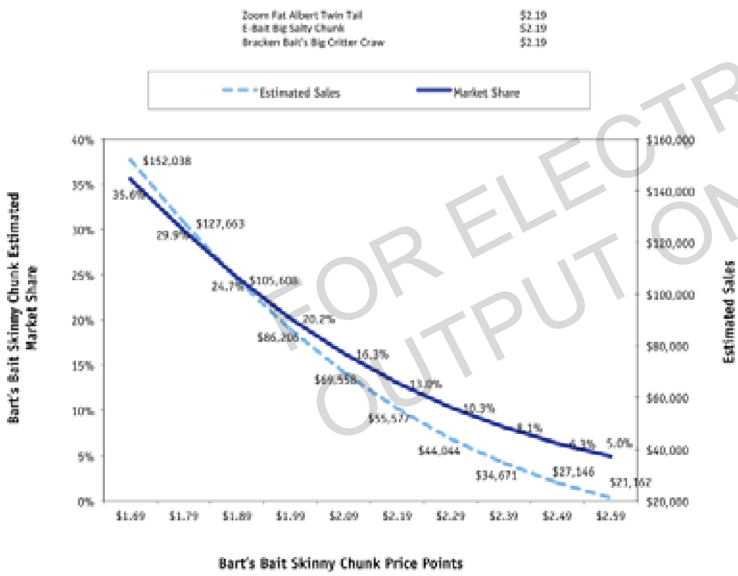


Figure 3: Tested Price Points for Bart's Bait Skinny Chunk



tions), Bart could expect about a 13 percent entry market share.

This model works as a market simulator. Bart can project its Skinny Chunk market share and sales for different price points. Figure 3 shows a graph of Skinny Chunk sales and market share by price-point if competitors are all priced at $2.19.

Clearly, the lower the price Bart charges, the greater the market share it will gain. However, the lower prices may not be realistic price points and the higher ones may give Bart too low a sales figure. Moreover, Bart's competitors might raise or lower prices around the median of $2.19, which would render this analysis obsolete.

One of the great advantages of the discrete choice model is its flexibility. The graph (Figure 3) can be reproduced in literally infinite iterations if market conditions change. For example, if E-Bait priced its bait at $2.29 and Zoom priced its at $1.99, a new chart can easily be created. Or if Bracken's suddenly exits the market, the model can be reset for only three brands.

So Bart's Bait can determine the best price for its product based on changes in the competition.

Large numbers of attributes

Users of the three methods seen so far would probably agree that their most serious problem is dealing with large numbers of attributes. Corporate clients often seek to include every possible scenario, which may lead to long, grueling interviews and reduced incidence – not to mention respondent fatigue and inferior-quality data. There is a definite trade-off between including "the kitchen sink" and performing an actionable choice model.

There are, however, choice models that deal with large numbers of attributes. The final two examples are useful when these come up.

Maximum-difference (max-diff)

Whereas a paired-comparison question asks a respondent to make a binary choice, maximum-difference has the respondent specify "best" and "worst" choices from sets of three or more objects.

Max-diff can test a large number of attributes without requiring respondents to see them all. Moreover, max-diff is flexible – data from a max-diff analysis can be used to create a choice simulator like those of conjoint and discrete choice.

A maximum-difference choice model is easily administered, has multiple levels of analysis and is a very effective tool in establishing the relative priority of such items as: potential message for a new product; features or benefits of a service; which extras to include in a loyalty program; fundamental customer interests and activities; and unmet/future needs.

Maximum-difference eliminates the awkwardness of a large set of customer choices, moves rapidly through the survey and eases respondent fatigue. Moreover, max-diff not only reveals the descriptive results that companies are looking for but can also be applied to predict future customer behavior.

As an example of a max-diff study, let's say that a hotel chain, Malone Gardens, wants to know which benefits members of its loyalty program prefer. Moreover, it would like to know if it can expect a bump in expected visits if a given benefit is included.

Initially Malone Gardens provided a long list benefits it wanted to test. For the sake of simplicity, we tested only 12:

- free hotel nights
- experience getaways
- dream vacations
- premium merchandise
- airline miles
- bonus points
- complimentary health club privileges
- hotel room upgrades
- reward planner services
- spa or golf packages
- partner car rental privileges
- shopping and dining

Figure 4: Maximum Difference Point Mean Allocations

			Total Sample
11	Free Hotel Nights	Level 1	75.0
7	Hotel Room Upgrades		66.5
8	Shopping and Dining	Level 2	27.7
13	Dream Vacations		14.9
10	Experience Getaways	Level 3	1.4
6	Complimentary Health Club Privileges		0.9
1	Premium Merchandise		-2.7
4	Airline Miles		-3.7
9	Spa Or Golf Packages		-6.2
2	Partner Car Rental Privileges	Level 4	-25.8
12	Bonus Points		-30.8
5	Reward Planner Services		-64.7

Level 1: WINNERS - highly desirable items that have high/low scores near 100
Level 2: DESIRABLE - items with a high/low over 10
Level 3: NEUTRAL - items with high/low near 0, indicating indifference
Level 4: NOT DESIRABLE - items with high/low negative scores over 20

Each of the 12 benefits can be shown randomly (thus no need to do an orthogonal design). However, they must be shown in groups of four and each benefit must be shown the same number of times. Two example sets are shown below.

The construction of max-diff gives it an advantage in that if a client wants to test 30 attributes he or she can. Each respondent would see only 10 or so scenarios, with the flexibility that each respondent may likely see all different sets of benefits.

One drawback for max-diff is that attributes are unable to have different levels (i.e., how many bonus points should be given). Secondly, if a client wants to test a large number of attributes, he or she will most likely need a large sample to fulfill data requirements.

First-level max-diff scores may be calculated as such. Among scenarios with each reward present:

- a reward scores +100 if "Most Appealing"
- a reward scores zero if "Not Chosen"
- a reward scores -100 if "Least Appealing"

Figure 4 shows the first-level descriptive scores for the Malone Gardens project. Not surprisingly, guests want free nights.

Another advantage of max-diff is that it can be formed to compute a choice model, similar to a conjoint. In our example, guests were asked "How likely would you be to join Malone Gardens Priority Club?" The average before running through the max-diff exercise was 5.6. Figure 5 shows that if the hotel adds three of the more popular benefits, the likely-to-join average rises from 5.6 to 7.3. This kind of agility makes max-diff a leading choice model choice for clients who want to test a large number of attributes.

Least Important	Reward	Most Important
	Free Hotel Nights	X
X	Experience Getaways	
	Dream Vacations	
	Premium Merchandise	

Least Important	Reward	Most Important
	Bonus Points	
	Complimentary Health Club Privileges	
	Hotel Room Upgrades	X
X	Reward Planner Services	

Adaptive conjoint analysis

Like max-diff, adaptive conjoint analysis (ACA) was developed specifically for situations where there are many attributes. For example, given the large number of variables that go into car design, the automobile industry employs ACA. Generally, ACA is recommended for scenarios with more than six attributes when pricing research isn't

Figure 5: Maximum-Difference Difference Simulator	
Malone Gardens Hotel	
	1=Included in The Items
Loyalty Benefits	
Free Hotel Nights	1
Hotel Room Upgrades	1
Shopping and Dining	0
Dream Vacations	0
Experience Getaways	1
Complimentary Health Club Privileges	0
Premium Merchandise	0
Airline Miles	0
Spa Or Golf Packages	0
Partner Car Rental Privileges	0
Bonus Points	0
Reward Planner Services	0
Likelihood to Join Malone Gardens Priority Club Member	7.3

the goal of the study.

In ACA, the researcher designs a computer-interactive interview. The questions change based on a respondent's last answer. ACA allows the researcher simulate respondent preferences for new or modified products, giving the analysis the ability to test what-if scenarios such as product formation or marketing activities. Respondent utilities are used to estimate strengths of preference, purchase likelihoods, and, with cumulated respondents data, to provide a simulated market share.

The downside of ACA is that the data collection vendor must have the ACA module. This makes the technique expensive. In addition, employing a large number of attributes and levels often confuses the results. Respondents may have difficulty keeping in mind that all other attributes not involved in the current question.

For a new-car study with 12 attributes, for example, ACA might cut the question down to something like this:

Please rank order the following:
Power windows/locks
Automatic transmission
No antilock brakes
Five-year loan
Six airbags

To be followed by:

Which of these two are more important?
Automatic transmission
OR
Six airbags

ACA interviews generally take a minimum of 45 minutes. In the past decade, the use of ACA is declining. Firms have come out with hybrid methods that are simpler, such as max-diff. Researchers have also shifted to discrete choice, as choices are viewed as more realistic than concept ratings.

Many ways to apply

There are many ways to apply choice modeling to business solutions. (Table 2 summarizes my recommendations. ACA is not included, as I generally do not recommend this methodology.) Each choice methodology can yield extraordinarily accurate, real-life results if applied correctly. That is why it is important to zero-in on the specific problem the client wants to solve. Deliverables must be made clear, for they determine which method we choose.

Whenever we begin a kickoff meeting, I always ask, "At the end of this project, what would you like to be holding?" From there we work backwards to form a game plan to best answer the questions. And then the fun part begins.

Table 2	
Service Category Questions	Choice Model
Small Number Of Existing Attributes	Paired-Comparison
Product Design	Traditional Conjoint
Market Share, Pricing	Discrete Choice Model
Large Number Of Attributes	Maximum-Difference

Michael Lieberman is founder and president of Multivariate Solutions, a New York research consulting firm. He can be reached at 646-257-3794 or at michael@mvsolution.com.