Sawtooth Software

RESEARCH PAPER SERIES

Three Ways to Treat Overall Price in Conjoint Analysis

Bryan Orme, Sawtooth Software

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

Three Ways to Treat Overall Price in Conjoint Analysis

[1]Bryan Orme, Sawtooth Software
May, 2007

This article discusses three ways to treat overall price in traditional ratings-based conjoint analysis or discrete choice (CBC) studies:

- Traditional Approach
- Conditional Price
- Continuous Price

The traditional approach is the easiest to manage, but the other two techniques offer benefits for more advanced applications in specialized situations. I don't know the history of the conditional price and continuous price approaches, but they probably were being used as early as the 1970s. My first experiences with them occurred in 1993. Because our recent work with Adaptive CBC uses the Continuous Price approach ("A New Approach to Adaptive CBC," Johnson and Orme 2007, forthcoming), it was important that we do some investigation into the stability of price estimates under Continuous Price. Those simulation results are reported at the end of this article.

Traditional Approach, with Price as Separate Attribute

In conjoint analysis, the typical approach to price is to include it as a separate attribute in the study design. For example, if we were studying laptop computers, we might include the following attributes:

Dell
HP
Toshiba

1 GB RAM
2 GB RAM
4 GB RAM

80 MB Hard Drive
120 MB Hard Drive
160 MB Hard Drive

2.0 GHz Processor
2.5 GHz Processor
3.0 GHz Processor

[1] Many thanks to Rich Johnson who reviewed earlier drafts of this document and provided many useful suggestions for improvement.

$700
$1,000
$1,500

With this traditional approach, we vary each attribute independently of the others (an orthogonal design). Level balance is achieved if each level within each attribute appears an equal number of times. Designs that are both level balanced and orthogonal are optimally efficient for estimating the part-worth utilities with precision (assuming respondents answer using a simple additive model). In the example above, we could estimate part-worth utility values corresponding to each of the three levels of price (a part-worth utility function), or we could estimate a single linear term to reflect the slope of price (a vector utility function). Most researchers choose the part-worth utility function, because it is more flexible and can account for non-linearities in the price function. However, it comes at the cost of increased parameters to estimate.

Despite the robust statistical qualities of orthogonal designs, some researchers and respondents have been bothered that product concepts with the best features sometimes are shown at the lowest prices (and products with the worst features are sometimes shown at the highest prices). These combinations seem illogical and often lead to obvious (dominated) choices in the questionnaire. Such questions are less informative and lead to a less realistic experience for the respondent.

Conditional Pricing

Conditional pricing is one approach for increasing the realism of the concepts shown to the respondent. With conditional pricing, incremental amounts are added for premium brands or premium features, so enhanced products are generally shown at higher prices. We still treat price as a separate attribute with just a few levels (such as three to five). But, those levels of price are described with different absolute dollar amounts, depending on product characteristics. Probably the most common use among Sawtooth Software users is to associate different brands or different brand/package size combinations with different price ranges. For example, the premium brand might be shown at $10, $15, or $20 whereas the discount brand is shown at half those prices: $5, $7.50, or $10. In the design matrix, we still treat price as a single attribute with three levels, even though a larger number of actual prices are displayed.

With conditional pricing, we use a price lookup table to determine actual prices to show in the questionnaire, based on the characteristics of each product. To create the lookup table, we first decide how many attributes (including the price attribute) will participate in the conditional pricing relationship. Our current CBC software permits price to be conditional on up to three other attributes.

Let’s assume with our previous laptop PC example that we wanted to make price conditional on RAM, Hard Drive, and Processor. We first start by choosing price premiums associated with those attribute levels. These premiums will not be explicitly shown to respondents next to each attribute level, but will be used just to determine the overall average price. Only a single total price is shown within the product concept.

Example 2: Conditional Price

Dell
HP
Toshiba

1 GB RAM +$0
2 GB RAM +$100
4 GB RAM +$200

80 MB Hard Drive +$0
120 MB Hard Drive +$100
160 MB Hard Drive +$200

2.0 GHz Processor +$0
2.5 GHz Processor +$200
3.0 GHz Processor +$400

Low Price (-30%)
Medium Price (Average Price)
High Price (+30%)

Let's assume that the base price for the laptop is $750. We construct a look-up table to determine the prices that we should show on the screen for each possible product combination at each of the three prices. This table would have a total of 3 x 3 x 3 x 3 = 81 rows. The first five rows of the price lookup table look like:

RAM	Hard Drive	Processor	Price	Text to Display
1	1	1	1	$525
1	1	1	2	$750
1	1	1	3	$975
1	1	2	1	$665
1	1	2	2	$950
…	…	…	…	…

For example, row four of the table specifies what price should be displayed when a product with 1 GB RAM, 80 MB Hard Drive, and 2.5 GHz Processor appears at the low price. The price to show is $665. This is determined by taking the base price ($750) plus the price increments associated with the three conditional attributes, and then reduced by 30%.

The benefit of conditional pricing is that more reasonable prices are shown to respondents and in the simplest case price may be estimated using main effects for (in this example) the three levels of price in the design. And, critically, the design is still orthogonal and unencumbered by prohibitions.

There are a few challenges when working with conditional pricing:

- We no longer can interpret the main effect utilities for attributes involved in the conditional relationship independent of price. For example, we cannot interpret the levels of processor speed as the preference for each of its levels holding everything else constant. The utility of each level of processor speed is confounded with the incremental price attached to that level. The levels must therefore be interpreted as the preference for levels of processor speeds given the average prices shown for those levels. So, it's very possible to achieve a higher average utility for 2.0 GHz Processor @ +$0 than 2.5 GHz Processor @ +$200, if respondents on average did not feel that it was worth the extra $200 to have the faster processor.
- The estimation of part-worth utilities works well when the prices shown to respondents are based on a certain percentage increase or decrease from the average price. However, the resulting prices often need to be rounded to the nearest $100 (or made to end in a "9" for consumer packaged goods). Quite small relative changes in price to round to a more presentable number don't pose much problem. But, significant price changes due to rounding introduce error in the utility estimation for the price attribute.
- If the conditional pricing table is not built in a consistent, proportional manner as specified here (or if rounding resulted in significant deviations from the original formula-based values), it may become impossible to model the data correctly using the conditional price approach unless imposing interaction effects. Interaction effects may lead to overfitting.
- If respondents oversimplify by paying attention only to prices, the preference for lower levels of performance attributes will be biased upward.
- The current software limitations specify that no more than three attributes (in addition to price) may be included in the conditional relationship.

Continuous Price

Another approach not currently offered in Sawtooth Software products is continuous price. (Even though it's not a supported feature of the software, a power user can still do it, though it requires being able to reformat the text-only studyname.CHO file.) Continuous price differs from conditional price is two ways. First, it generalizes the idea of conditional pricing (beyond the software limitations of just three attributes). Second, it estimates the effect of overall price as a linear coefficient, rather than as a part-worth utility function. As with conditional pricing, we approach the problem by considering a base price for the product as well as fixed price premiums for levels of non-price attributes (plus or minus some overall independent price variation). If we consider the example from the previous section, the base price is $750 and the most expensive product option would be $1,550 (prior to varying price by some independent amount).

As with the first two pricing approaches, we also only show a single overall price within the product concept, rather than showing prices attached to each attribute level. The only difference between conditional price and continuous price is in the coding of the design matrix, where price is coded in a single column as a continuous variable. Typically, a single price coefficient is estimated based on linear price (or the natural log of price)[2].

Because values in the price column of the design matrix are determined from information in other columns of the design matrix, that column would be linearly dependent on other columns if we didn't do something to break up that dependence. We do this by adding random variation to the prices.

The benefits of summed price relative to conditional pricing include:

- In contrast to conditional pricing, the utility of each feature level is estimated independently of any price premium associated with the level[3]. Thus, we would expect the utilities for levels of processor speed to look just like they would when using the standard conjoint approach with no conditional pricing[4]. See Appendix A for more in-depth explanation.
- Since we are estimating price as a continuous function, there is no worry about whether rounding prices to the nearest "9" or the nearest $100 will lead to errors in fitting the data.

But, these benefits come with a serious potential drawback: the price attribute is positively correlated with any attributes that involve incremental prices in the study, leading to less precise estimates of all effects, but most especially the price coefficient. The amount of correlation among attributes depends on the magnitude of the random variation in overall price as well as the size of the base fixed component of price relative to the incremental prices associated with each feature level. In the worst case, with no random variation, continuous price is simply the sum of the prices associated with the attribute levels. In that case, price would be perfectly predicted by a linear combination of the attributes and the design would be deficient. But, if we additionally vary the overall price by a large enough random amount (see guidelines further below), we can obtain sufficient precision of the estimates for overall price sensitivity as well as the other features in the study.

Continuous price is not an option in the current implementations of Sawtooth Software's CBC or CVA products. However, a power user could implement it for either type of study and estimate the results properly using Latent Class, CBC/HB or, in the case of a CVA study, HB-Reg.

The final section of this paper includes a simulation study to investigate what variation should be specified in the overall price attribute to lead to reasonable estimates with continuous price.

---

[2] More complex curve fitting might be considered, as well as piecewise coding. These approaches may provide better fit to the data, but risk overfitting and also introduce some correlation in the independent variable matrix.

[3] Because the non-price attribute levels are estimated independently of both overall price and any incremental pricing attached to the product features, it shouldn't matter what incremental price levels are attached to each level (within reason, of course). Preliminary evidence from a recent split sample CBC study we conducted supports this.

[4] For evidence of this, see the part-worth utility results for a split-sample study comparing traditional pricing and continuous price in Johnson and Orme, 2007.

Simulation Study

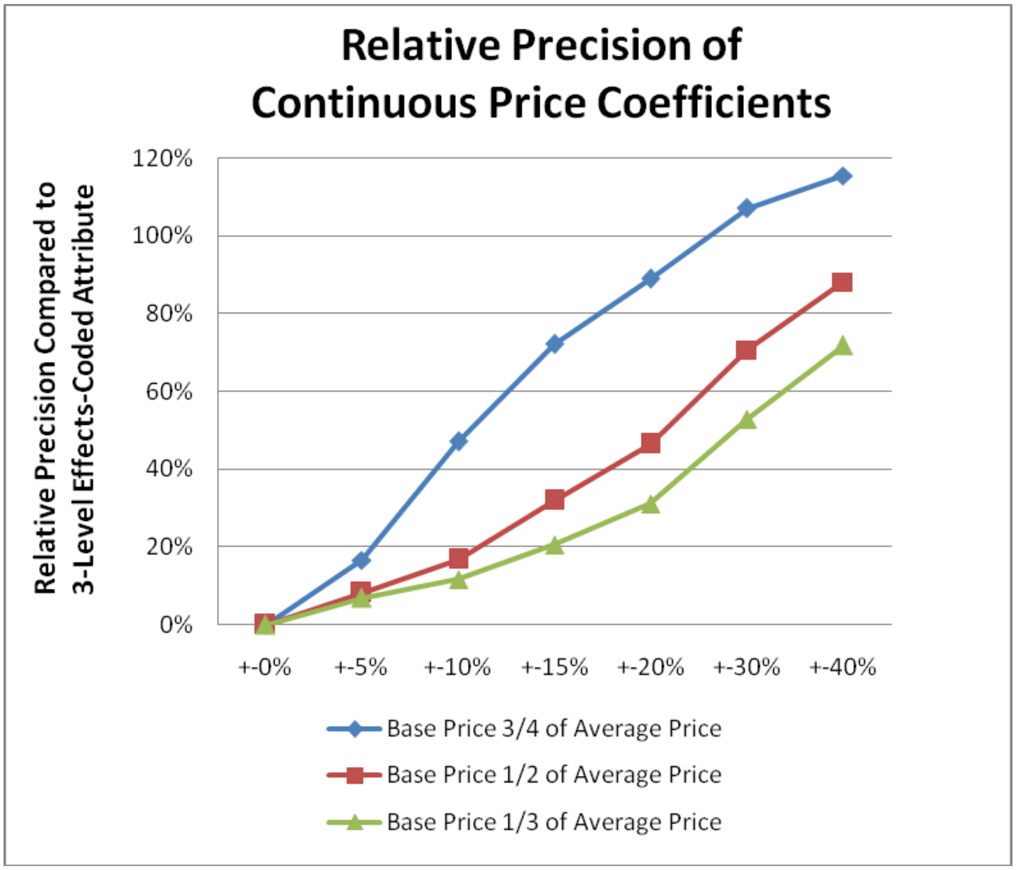
As mentioned earlier, the amount of random variation given to continuous price has a direct impact on the efficiency of the estimates. To provide guidelines regarding how much random variation in price we should include in continuous price designs, we conducted a synthetic study with 300 simulated respondents. Each (computer generated) respondent received 10 choice tasks that were answered randomly. There were five attributes, each with three levels, along with overall price. The overall price was found using the base price plus incremental prices associated with each of the other five attributes. Then, if we were varying the overall price as much as +/-10%, we used five distinct price disturbances within that range of -10%, -5%, +0%, +5%, and +10%[5].

Continuous price usually consists of a fixed base price plus additional upgraded feature costs. If the base price is relatively large compared to the incremental prices attached to upgraded features, then (after disturbing overall price by the price variation) the resulting price attribute will be relatively uncorrelated with the linear combination of the other attributes. However, if the base price is relatively small compared to the incremental prices for the other levels in the study, then the resulting overall price will be more strongly correlated with a linear combination of the other attributes. Therefore, we needed to consider how different relative sizes of the fixed base price relative to incremental prices would affect the results.

Study Procedure and Results

We simulated different amounts of independent price variation on continuous price, from as low as +/-5% to as much as +/-40%. We estimated price as a single coefficient, to be applied to the natural log of total price. Prior to estimating (with aggregate logit) for each price condition, we normalized the log price variable to have a variance of unity, so that the standard errors would be comparable across the different simulation runs that featured substantial differences in the amount of absolute price variation. For each study condition, we recorded the standard error for log price. The precision of the estimated price coefficients relative to a standard three-level attribute in a parallel study without continuous price is shown in the following chart.

[5] One could choose any number of discrete price variations within the +/-10% range; since the attribute is presented as a continuous variable with usually hundreds or thousands of unique summed prices, there is no additional Number Of Levels bias if using more rather than fewer discrete price variations.



Recommendations
According to our simulations, the precision of the utility estimate for log price depends strongly on both the amount of added random price variation and the relative size of the constant base price compared to the feature-based prices. When the fixed base price of the product is 3/4 the total average price[6], as little as +/-10% price variation on continuous price will achieve precision of estimates that are nearly 50% as efficient as a standard 3-level attribute coded as a part-worth function. In terms of absolute magnitude, the standard error for price was 0.038 compared to 0.026 for levels from the standard three-level attribute. Based on additional simulations, we found that a standard five-level attribute (if placed within the same study instead of continuous price) would also achieve standard errors of estimates for its levels of about 0.038. Generally, in practice, we'd be comfortable with such precision.

However, if the base price only accounts for 1/3 of the total average price of the product (most of the price is explained by the incremental feature prices), then we'd need to vary continuous price +/-30% to achieve similar precision. Based on this simulation study, we can make general recommendations for continuous price:

[6] It may be useful at this point to illustrate how to distinguish fixed base price from incremental prices for the purposes of applying our results to future studies. Consider an attribute such as brand with 3 levels and incremental prices of Dell +$700, HP +$800, and Toshiba +$900. The fixed component (irrespective of brand) is $700, and the incremental prices are Dell +$0, HP +$100, and Toshiba +$200. Thus, when decomposing price into fixed versus incremental prices, you should ensure that there is no fixed component "hiding" within what appear to be incremental prices. The lowest incremental price for each attribute should start with +$0.

Recommended Minimum Independent Price Variation for Continuous Price

If base price is 3/4 of total average price: +/-10%
If base price is 1/2 of total average price: +/-20%
If base price is 1/3 of total average price: +/-30%

Let's apply the recommendations above to the laptop computer example we introduced earlier, the base price was $750, and the most expensive product (prior to introducing any independent random price variation) was $1,550. The average price falls about half-way between that interval, or at $1,150. Therefore, the fixed price component is 65% (750/1150) of the total average price. Conservatively, we would recommend varying continuous price by at least +/-20% in this situation, though one could interpolate between the functions in the relative precision chart above to justify variation of at a minimum +/-15%.

Of course, choosing the price variation also depends on the client's needs and the market simulations to be run. You should avoid extrapolating beyond the total range of price included in the questionnaire. Increasing the random price variation will improve your ability to simulate extreme priced products, at the risk of making the questionnaire present products that seem to have unreasonable prices, given their features.

Appendix A

In this section, I'll describe how coding the overall price attribute as a single column in the design matrix (estimating a single coefficient for price) makes it possible to interpret the utilities of features that have incremental prices associated with them independently of those incremental prices. Consider the simplest of conditional pricing studies, where price is conditional on just one other attribute (brand). Imagine a study with three brands, with prices to be shown as follows:

	Low Price (-30%)	**Middle Price (+0%)**	**High Price (+30%)**
Brand1	$7.00	$10.00	$13.00
Brand2	$10.50	$15.00	$19.50
Brand3	$14.00	$20.00	$26.00

With the conditional pricing approach, we treat the price attribute as a categorical attribute with three levels, -30%, +0%, and +30%. So, if we presented two product alternatives (each at their lowest prices):

Brand1 @ $7.00
Brand2 @$10.50

We would code the alternatives as follows (level 3 is the reference level, and is therefore omitted from the design to avoid linear dependency):

Brand1	Brand2	Price1	Price2
1	0	1	0
0	1	1	0

With this coding scheme, the effect of price is captured as dummy (or effects) coded parameters. Level one is the utility associated with a 30% reduction in price for the given brand (irrespective of the absolute value of price). So, even though the two prices shown to respondents are different ($7.00 for Brand1 vs. $10.50 for Brand2), the coding of price in the design matrix is identical. To account for the fact that there is a price difference in $3.50 between those two product concepts, the disutility for the extra $3.50 in price must be accounted for in the effect for Brand1 relative to Brand2. Thus, the price dummies account for the slope of the curve, and the brand dummies account for inherent preferences for the brands plus the average amount of price difference between prices shown for the brands. Thus, we cannot interpret the brand effects separate from the price premiums reflected by those brands.

However, consider what happens when we code price as a single linear coefficient, as with the "Continuous Price" approach. The two product alternatives are coded as:

Brand1	Brand2	Price
1	0	\$7.50
0	1	\$10.50

Thus, the information contained in the Price column accounts not only for the relative differences in price, but also that one brand is consistently shown at different prices than the other. The Price coefficient partials out all the effect of price (both the slope and the shift component), and the effects for Brand are captured independently of price. The utilities for brand may be interpreted as their inherent desirability, in the traditional manner of interpreting main effects.