IMPROVING PREDICTION IN
CONJOINT MEASUREMENT

by

Dov Pekelman
Tel-Aviv University

and

Subrata K. Sen
University of Rochester

REVISED
July, 1978

Working Paper No. 7811

Graduate School of Management

MGT.LIB
D837.6wo
#7811

IMPROVING PREDICTION IN
CONJOINT MEASUREMENT

by

Dov Pekelman
Tel-Aviv University

and

Subrata K. Sen
University of Rochester

REVISED
July, 1978

Working Paper No. 7811

Not to be quoted or reproduced without the authors' permission.

COMMENTS ARE WELCOME

IMPROVING PREDICTION IN CONJOINT MEASUREMENT

ABSTRACT

Conjoint measurement generally concentrates on the assessment of utilities for a few discrete attribute levels. An important practical application of such utility assessments is the prediction of the overall value of alternative new product concepts. If these product concepts are defined as combinations of attribute levels other than the original discrete levels presented to the respondents, linear interpolation is commonly used to predict utilities for the new attribute levels. However, linear interpolation may lead to poor predictions and the consequent choice of the wrong new product to introduce in the market. Therefore, we present in this paper an alternative to linear interpolation: the estimation of a utility function. We provide an analytical procedure and some numerical results to indicate the superiority of the utility function approach for the prediction of attribute utilities. The market share implications of improved utility prediction are also discussed.

IMPROVING PREDICTION IN CONJOINT MEASUREMENT

INTRODUCTION

An important application of conjoint measurement in marketing¹ is the prediction of preferences for new product concepts. These preference predictions are then used to select the product concept that is likely to lead to the highest possible market share. Since new product concepts are represented in conjoint measurement as combinations of different product attributes, predicting overall preferences for the product concepts necessitates the prediction of utilities for each possible level of the relevant attributes. Errors in predicting these utilities could result in the selection of an inferior product, leading to an opportunity loss in terms of potential market share.

Current practice in conjoint measurement produces utility measures for a few discrete levels of the relevant attributes (see, for example, Figure 1a which portrays the measured utilities for four discrete levels of a sample automobile attribute: "Top Speed"). Utility predictions for attribute levels other than the original discrete levels are then made using techniques such as linear interpolation (see the utility predictions for 110 and 120 miles per hour respectively in Figure 1b).

Such predictions could be significantly inaccurate since interpolation ignores the fact that the utility measures obtained for the original discrete attribute levels contain an error component. The error is introduced by the respondent when he translates his true underlying utilities for the input product concepts into a preference rank ordering. This rank ordering is then used by the conjoint measurement algorithm to produce the utility measures for the discrete attribute levels used to define the input product concepts.

Since the rank ordering used to obtain the utility measures is not free of error, it follows that the utility measures also contain an error component.

Since the utility measures for the discrete attribute levels are not free of error, a procedure that is preferable to linear interpolation is the estimation of the parameters of a utility function which is then used to predict utilities for the various attribute levels (see Figure 1c, for example). Since the utility function, in effect, "smooths out" the measurement errors, it produces more accurate utility predictions not only for the "new" attribute levels (such as the 110 and 120 miles per hour levels in Figure 1b) but also for the original attribute levels (e.g., the four attribute levels in Figure 1a). Greater accuracy in the prediction of attribute utilities increases the likelihood of selecting the new product concept that would result in the highest possible market share.

In this paper we develop an analytical procedure to show that under plausible assumptions, the utility function approach does lead to better utility predictions compared to linear interpolation. We then show how the utility function predictions can be used to calculate increases in expected market share. The magnitude of the market share increase is illustrated by a simple numerical example.

RATIONALE FOR THE UTILITY FUNCTION APPROACH

As typically used in marketing, conjoint measurement calculates utilities for a few discrete levels of the attributes used by a consumer in evaluating product alternatives. For example, suppose that "Top Speed" is one of the attributes used by a consumer in evaluating automobiles and that four levels of "Top Speed" are used to generate the required rank-ordered input data from the respondent. If the four levels are, say, 85, 100, 115 and 130 miles per hour, one obtains as output the consumer's utilities for these four levels of "Top Speed" (see Figure 1a).

Ideally, the marketing manager would like to ascertain the consumer's utilities for levels of "Top Speed" other than 85, 100, 115 and 130 miles per hour because he may be evaluating, for example, two other automobile brands which have top speeds of 110 and 120 miles per hour respectively. However, the manager is limited by the amount of data that a respondent can be reasonably expected to provide. Because most product concepts are characterized by several attributes, the number of concepts that a respondent has to evaluate becomes very large if each attribute is represented by a large number of levels. Since one cannot expect a respondent to provide reliable overall judgments for an unreasonable number of product concepts, researchers are forced to limit themselves to a very small number of levels of each attribute in gathering data from respondents. They then have to infer the consumer's utilities for attribute levels other than those used to generate the input data.

If the attribute is continuous (as "Top Speed" is), the method usually used to infer such utilities is linear interpolation. For example, a consumer's utility for a "Top Speed" of 110 miles per hour would be predicted by a linear interpolation between the utilities estimated for 100 and 115 miles per hour (see Figure 1b).

The use of linear interpolation for utility predictions implies that the researcher is assuming that the underlying utility function is a piecewise linear function. This is equivalent to an admission of complete ignorance about the phenomenon under study. Not only is it being assumed that the researcher knows nothing about the shape of the function, it is also being assumed that the function is not a smooth function. At the very least, the researcher should assume that the underlying utility function is a smooth

function (e.g., a polynomial) with the same number of parameters as a piecewise linear function.²

The other problem with the use of a piecewise linear function is that one is essentially trying to fit k data points perfectly by means of a $(k-1)$ order piecewise linear function which passes through all the points. The number of parameters equals the number of observations, resulting in zero degrees of freedom. A perfect fit is obtained to the generated data, but since the generated observations are not completely free of error (see the earlier discussion for the source of the error), the fitted function is unlikely to be appropriate for predictive purposes.³ Hence, it seems more sensible to estimate a smooth utility function which has several degrees of freedom (see Figure 1c) instead of merely measuring utilities for discrete attribute levels and then using piecewise functions to interpolate between the measured data points. It should be reiterated that the utility predictions provided by the utility function will be more accurate not only for the two new levels of "Top Speed" (110 and 120 miles per hour) but also for the four original discrete levels of the attribute (85, 100, 115 and 130 miles per hour respectively).

As is true in most model building, the estimation procedure would start with some prior knowledge about the phenomenon being studied. For our problem, the prior knowledge that might be used is that the underlying utility function is smooth and that it can be represented by one of several plausible functional forms which might each be analyzed separately. In this paper, we provide an analytical framework which enables us to calculate the amount by which an estimated smooth utility function (which does not contain any misspecification error) is superior to linear interpolation in

terms of the variance of the prediction error. It should be noted that in our calculations, we do not consider the misspecification bias introduced by the interpolation method. This makes the analysis comparable for the two methods. In general, little can be said about bias resulting from misspecification of the functional form when the true functional form is not known. Therefore, our analysis concentrates on those elements of the prediction error where the results can be fully generalized.

The framework presented in this paper can be used to study any functional form that is linear in its parameters (e.g., an n -degree polynomial). The specific polynomial we use in our analysis is the second-order polynomial. This is used only to demonstrate our method and also because we believe that in many applications, the utility function is single-peaked, a characteristic which can be easily represented by the second-order polynomial (this issue is discussed in more detail later in the paper).

Utility prediction is, however, only a stepping stone toward the principal objective of this paper. Correct utility predictions are managerially important only because such predictions are used by marketing managers to determine which new product concept to introduce into the marketplace. Errors in prediction may lead to the introduction of the wrong product and hence to a loss in potential market share. Therefore, the most important objective of this paper is to provide an analytical method of determining the nature and magnitude of such market share losses and thus draw the attention of both researchers and managers to the potential dangers of the unquestioning use of linear interpolation for utility prediction.

Before proceeding further, it is important to emphasize that the issue of estimating a utility function does not typically arise if all the attribute

are categorical. For example, if the product category: "restaurant menus" is represented solely by categorical attributes such as types of entrees and types of desserts [5, pp. 159-185], it is not clear how one could specify a utility function for such categorical attributes. If, on the other hand, the attributes are continuous attributes such as "percent discount," "number of stores," and "\$cost of card" (which are three attributes of retail discount cards analyzed in [5, pp. 135-158]), it is clearly preferable to use a utility function to estimate the consumer's utilities for both the original and new levels of the three attributes. In most cases, conjoint analysis problems involve both continuous and categorical attributes. For example, Green [4] describes a problem dealing with trans-Atlantic air travel which involves two categorical attributes (e.g., airline, represented by four discrete levels: TWA, BOAC, Pan Am, and Air France) and seven continuous attributes (e.g., arrival time punctuality, anticipated plane load, etc.). For such problems, Pekelman and Sen [11] have indicated a method of estimating utility functions for the continuous attributes and specific utility values for the discrete levels of the categorical attributes. Clearly, the issue of estimating a utility function remains important for the continuous attributes of such problems.

DESCRIPTION OF THE ANALYTICAL PROCEDURE

Rationale for Using Regression Analysis

Using the rank-ordered input data typically available in conjoint analysis, a smooth utility function can be estimated for continuous attributes using a technique such as linear programming (see Pekelman and Sen [11]). It was argued above that such a utility function would be theoretically superior to piecewise interpolation. But would it be a significantly more accurate predictor of consumer utilities empirically? This question cannot be answered analytically if linear programming is used to estimate the utility function. It could, however, be answered by simulation. A simulation which

experiments with various characteristics of the problem such as the number of attributes, the number of levels per attribute, the nature of the underlying utility functions, etc., is likely to be time consuming and expensive. Therefore, in order to obtain some insight into the general problem of interpolation versus the estimation of a utility function, the linear programming estimation approach is replaced by regression analysis and the performance of interpolation techniques is compared with that of regression. Regression analysis makes it possible to investigate this issue analytically. Besides, the direction of the results do not change when linear programming is used instead of regression (see Pekelman and Sen [11] for a report of the results of a small-scale simulation).

Preliminary Considerations

To compare interpolation with regression, only a single attribute will be considered and it will be assumed that a sample of n observations is available: $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$. X_i is the i th level of the attribute and Y_i is the corresponding measured utility. It is desired to predict the utilities for each of p new levels of the attribute which are denoted as $X_a, X_b, \dots, X_p$.

These predictions will be made using (1) interpolation, and (2) the regression function estimated from the sample observations: $(X_1, Y_1), \dots, (X_n, Y_n)$. Typically, the marketing decision maker (e.g., a new product manager) utilizes the utility predictions by first converting them into a rank order of the new "product concepts." The more preferred concepts are then considered as candidates for market introduction. Therefore, the predictions for both schemes (interpolation and regression) will be converted into a rank ordering of the predicted utilities: $\hat{Y}_a, \dots, \hat{Y}_p$, for the p new levels of the attribute. Assuming that the true rank order of the utilities for $X_a, X_b, \dots, X_p$ is known, a rank order correlation such as Kendall's τ (see, for example, [12, pp. 871-874

can be computed to measure the predictive accuracy of each scheme. The two schemes can then be compared in terms of these computed τ 's. The details of the procedure outlined above are described below.

Rank Order Prediction Using Kendall's τ

The underlying utility function which generates the n observations:

$(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ is assumed to be of the following quadratic form:

$$Y_i = qX_i + rX_i^2 + \epsilon \quad (1)$$

where ϵ is the error in measuring Y . It is assumed that ϵ is normally distributed, $E(\epsilon) = 0$, and $\text{Var}(\epsilon) = \sigma^2$ for all values of X . The quadratic utility function is used only for illustrative purposes. Its advantage is that it can easily represent the two most common types of attributes: (1) attributes for which the consumer's ideal is at one extreme of the attribute, e.g., miles per gallon for an automobile, and (2) attributes for which the ideal point is somewhere in the middle of the attribute range, e.g., amount of sweetness in a cereal.⁴ It is important to note that our results do not depend upon the use of the quadratic utility function.

Now, let A, B, and C be three new product concepts (i.e., $p=3$ in this example) characterized by their predicted utility values: $\hat{Y}_a$, $\hat{Y}_b$, and $\hat{Y}_c$. $\hat{Y}_a$, $\hat{Y}_b$, and $\hat{Y}_c$ have means μ_a , μ_b , and μ_c , and variances σ_a^2 , σ_b^2 , and σ_c^2 , respectively, where the means express the true utility values of A, B, and C.

It is assumed that $\mu_a > \mu_b > \mu_c$ so that $A \succ B \succ C$. The predictions $\hat{Y}_a$, $\hat{Y}_b$, and $\hat{Y}_c$, will be normally distributed for both regression and interpolation given the assumption regarding the error term in (1). The variances: σ_a^2 , σ_b^2 , and σ_c^2 , will depend upon whether the predictions are made using linear interpolation or regression (see Appendixes A and B for the detailed derivations).

Now, A, B, and C can be rank-ordered in $3! = 6$ possible ways which can result in 0, 1, 2, or 3 violations of the true rank order of A, B, and C (see Table 1). For a particular number of violations of the true rank ordering of 3 objects, Kendall's τ takes on a specific value (again, see Table 1). Hence, by calculating the probabilities of 0, 1, 2, and 3 violations, it is possible to determine the mean value of τ for the two predictive schemes. A comparison of the mean values of τ will indicate the relative effectiveness of the two schemes in predicting the rank order of a set of p new attribute levels.

Calculation of the probabilities of m (where $m = 0, 1, 2, 3$) violations is done by combining the probabilities of the appropriate rank orders. From Table 1, the appropriate relationships are as shown below:

$$\Pr(m = 0) = \Pr[ABC]$$

$$\Pr(m = 1) = \Pr[ACB] + \Pr[BAC]$$

$$\Pr(m = 2) = \Pr[BCA] + \Pr[CAB]$$

$$\Pr(m = 3) = \Pr[CBA]$$

Therefore, it is basically necessary for the analysis to compute the probability of each possible rank ordering of the three objects.

Computing the Probability of Specific Rank Orders

The method of calculating the probabilities of the various rank orders will be indicated by taking the specific example of the rank order: [BAC]

$$\Pr[BAC] = \Pr\{(\hat{Y}_b - \hat{Y}_a) \geq 0 \text{ and } (\hat{Y}_a - \hat{Y}_c) \geq 0\}$$

$$= \Pr(s_1 \geq 0 \text{ and } s_2 \geq 0) \text{ where}$$

$$s_1 = (\hat{Y}_b - \hat{Y}_a) \text{ and } s_2 = (\hat{Y}_a - \hat{Y}_c)$$

Since $\hat{Y}_a$, $\hat{Y}_b$, and $\hat{Y}_c$ are normally distributed, s_1 and s_2 are also normally distributed. Further, the mean of s_1 is given by $(\mu_b - \mu_a)$ while the mean of s_2 is given by $(\mu_a - \mu_c)$. However, in order to compute the $\Pr(s_1 \geq 0 \text{ and } s_2 \geq 0)$, it is also necessary to determine Σ , the variance-covariance matrix for $\begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$. Σ will depend upon the specific prediction scheme employed: linear interpolation or regression. The specific expressions for the elements of Σ are developed in Appendix A for Regression and in Appendix B for Interpolation.

Knowing that s_1 and s_2 are normally distributed with variance-covariance matrix Σ and means $(\mu_b - \mu_a)$ and $(\mu_a - \mu_c)$ respectively, it is easy to calculate the $\Pr(s_1 \geq 0 \text{ and } s_2 \geq 0)$ (using numerical integration) and hence the $\Pr[\text{BAC}]$. The probabilities of the other five rank orders can be computed in a similar manner.

It is now a simple matter to compare the two prediction schemes in terms of the mean of Kendall τ for any specific set of input data:

(a) the n observations: $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$, (b) a numerical value for σ^2 in equation (1), (c) values of the parameters (i.e., q and r) of the utility function postulated in equation (1), and (d) a set of "holdout" product concepts defined by attribute levels: $X_a, X_b, \dots, X_p$.

NUMERICAL RESULTS

Input Data

Tables 2 and 3 present some numerical results of comparing the two prediction schemes using the framework developed in this paper. For both

Tables 2 and 3, the sample observations consisted of 5 data points: $X_1 = 1$, $X_2 = 2$, $X_3 = 3$, $X_4 = 4$, and $X_5 = 5$. The utility function parameters are $q = 4$ and $r = -0.5$ for Table 2 and $q = 4$ and $r = 0.0$ for Table 3. The parameter set used in Table 2 gives rise to a utility function with the ideal point within the range of the input data while the parameter set used in Table 3 results in a linear utility function. The same two sets of holdout attribute levels were used in both Tables. The first set ($X_a = 4.4$, $X_b = 3.5$, $X_c = 2.6$) consists of values in the middle of the interval defined by two sample observations, i.e., when w , as defined in equation (8) in Appendix B is close to 0.5. In contrast, the second set of holdout attribute levels ($X_a = 3.9$, $X_b = 2.8$, and $X_c = 1.9$) consists of values close to those of the sample observations, i.e., when w is close to 0 or 1.

Discussion of the Results

The mean predicted τ for both prediction schemes depends upon two main factors:

1. Measurement Noise: As measurement noise increases (i.e., σ^2 becomes larger), the mean τ should decrease.
2. Mean Predicted Utilities: As the mean predicted utilities of the new attribute levels (i.e., μ_a , μ_b , μ_c) become sufficiently different from one another, the predicted rank ordering of these attribute levels is more likely to match the actual rank ordering (ceteris paribus). This should lead to higher Kendall τ s.

The results reported in Tables 2 and 3 are uniformly consistent with the above. First, the mean τ falls for both prediction schemes for both Tables as σ^2 increases. Next, the τ s in Table 3 are generally higher than those in Table 2 since μ_a , μ_b , and μ_c differ more from one another in Table 3. In

Table 2 itself, the three μ 's differ more from one another for the second set of holdout attribute levels, leading to higher mean τ 's for both prediction schemes.

While the above patterns attest to the general plausibility of the results, it is of greater interest to study the comparative performance of Regression and Interpolation. An examination of the two Tables reveals that in all cases, the mean τ for Regression is greater than or equal to the mean τ for Interpolation. Note also that the advantage for Regression is particularly pronounced for higher values of σ^2 (see the "Improvement" column in Tables 2 and 3).

It is interesting to note the effect of the magnitude of w (defined in equation 8 in Appendix B) in determining the performance of Interpolation. As is indicated in the discussion of equation (14) in Appendix B, Interpolation is expected to perform less well when w is close to 0 or 1 (as is the case for the second set of holdout attribute levels) than when w is close to 0.5 (as is the case for the first set of holdout attribute levels). When the differences between the μ 's is roughly the same for the two holdout sets (as is the case in Table 3), Interpolation does perform better for the first holdout set. However, this effect is swamped in Table 2 where the three μ 's are very close to one another for the first holdout set.

In summary, the numerical results indicate the predictive superiority of Regression over Interpolation for a variety of cases. The results presented above have primarily been for three "holdout" product concepts. Theoretically, the results can be easily generalized to p (where $p > 3$) "holdout" concepts. Practically, it would be difficult to deal with cases where $p > 4$ since integration of multivariate normal distributions with

more than three dimensions is not an easy task. However, for most marketing applications, a practical limitation of four new concepts does not appear to be overly restrictive.

PERFORMANCE OF INTERPOLATION AND REGRESSION IN TERMS OF MARKET SHARE

Although utility prediction is considerably improved by utilizing Regression instead of Interpolation, it is of greater managerial interest to investigate the impact of improved prediction on the decision for which these predictions are typically used, i.e., the selection of a new product concept to introduce in the marketplace. Presented below is a formulation of this decision problem in which the computed Kendall τ values for Regression and Interpolation are used to assess the relative desirability of the two prediction schemes in terms of their impact on the market share likely to be attained for the new product that is introduced. The analysis presented below is for a homogeneous segment of the market. Extension to several market segments is straightforward.

Development of the Computational Formulae

Consider the situation in which three new product concepts are being considered for market introduction, and management is interested in determining the most preferred concept. Assume that the true ranking is $A \succ B \succ C$ and that the potential market shares of the three products: MS_a , MS_b , and MS_c are proportional to their true preference values: μ_a , μ_b , and μ_c . In other words, $\frac{MS_a}{MS_b} = \frac{\mu_a}{\mu_b}$, $\frac{MS_a}{MS_c} = \frac{\mu_a}{\mu_c}$, $\frac{MS_b}{MS_c} = \frac{\mu_b}{\mu_c}$. It is shown below how the two prediction schemes can be compared in terms of expected market share.

If the ranks [ABC] or [ACB] are predicted, product concept A will be chosen, resulting in a market share of MS_a . Similarly, predicting [BAC] or

[BCA] will result in a market share of MS_b while a market share of MS_c will be obtained if [CAB] or [CBA] is predicted. Let P_a , P_b , and P_c denote the respective probabilities of predicting that product concepts A, B, and C will be the most preferred. P_a , P_b , and P_c can be calculated from the probabilities of occurrence of each of the six possible rank orders. For example, $P_a = \Pr[ABC] + \Pr[ACB]$. Since it is known how to compute the probabilities of the various rank orders, it is easy to compute P_a , P_b , and P_c . Knowing P_a , P_b , and P_c , $\overline{MS}$ and $\text{Var}(MS)$, the mean and the variance of the market share can be computed for both estimation schemes.

$$\begin{aligned}
 \overline{MS} &= P_a MS_a + P_b MS_b + P_c MS_c \\
 &= MS_a \left\{ P_a + P_b \frac{\mu_b}{\mu_a} + P_c \frac{\mu_c}{\mu_a} \right\} \\
 &= \frac{MS_a}{\mu_a} (P_a \mu_a + P_b \mu_b + P_c \mu_c) \\
 &= k MS_a \\
 &\text{where } k = (P_a \mu_a + P_b \mu_b + P_c \mu_c) / \mu_a
 \end{aligned} \tag{2}$$

$$\begin{aligned}
 \text{Var}(MS) &= P_a (MS_a - \overline{MS})^2 + P_b (MS_b - \overline{MS})^2 + P_c (MS_c - \overline{MS})^2 \\
 &= P_a (MS_a - \overline{MS})^2 + P_b \left\{ \frac{\mu_b}{\mu_a} MS_a - \overline{MS} \right\}^2 + P_c \left\{ \frac{\mu_c}{\mu_a} MS_a - \overline{MS} \right\}^2
 \end{aligned} \tag{3}$$

Substituting $\overline{MS} = k MS_a$ in (3):

$$\text{Var}(MS) = MS_a^2 \left\{ P_a (1-k)^2 + P_b \left(\frac{\mu_b}{\mu_a} - k \right)^2 + P_c \left(\frac{\mu_c}{\mu_a} - k \right)^2 \right\} \tag{4}$$

Numerical Illustration

The above approach is illustrated by considering a case where Regression

is considerably superior to Interpolation: the $\sigma^2 = 49$ case in Table 3 for the second set of holdout attribute levels. For this case, $\mu_a = 15.6$, $\mu_b = 11.2$, and $\mu_c = 7.6$. P_a , P_b , and P_c are calculated to be:

	P_a	P_b	P_c
Regression	0.940	0.059	0.000
Linear Interpolation	0.643	0.241	0.115

Using (2) and (4):

$$\overline{MS} = 0.982 MS_a, \quad \text{Var}(MS) = 0.0044 MS_a^2 \quad \text{for Regression}$$

$$\overline{MS} = 0.872 MS_a, \quad \text{Var}(MS) = 0.0333 MS_a^2 \quad \text{for Interpolation}$$

For both Regression and Interpolation, the mean market share, $\overline{MS}$, and the variance in market share, $\text{Var}(MS)$, have been calculated as a function of MS_a , the potential market share of product A. Even though MS_a is not known, the results presented above allow us to make a comparative statement about the relative performance of Regression and Interpolation. We note that the Regression prediction provides both a higher mean market share ($0.982 MS_a$ vs. $0.872 MS_a$) and a lower variance of the market share ($0.0044 MS_a^2$ vs. $0.0333 MS_a^2$). In percentage terms, these figures translate into a 12.6 percent increase in expected market share if we use Regression instead of Interpolation, while the variance in market share is reduced by 87 percent. Hence, it appears that in this example Regression would be preferred to Interpolation in terms of both the mean and the variance of market share.

The analytical framework presented in this paper allows us to obtain a quantitative measure of the negative impact of using an inferior utility prediction procedure such as linear interpolation. Use of this framework makes it possible to gauge the potential loss in market share that might result

due to the introduction of the wrong new product as a consequence of using linear interpolation.

SUMMARY AND CONCLUSIONS

An alternative to the common practice of calculating utilities for several discrete attribute levels is the estimation of utility functions for continuous attributes. Compared to the discrete utility measures obtained by conventional conjoint measurement procedures, the utility function approach provides more accurate estimates of utilities for the original attribute levels used to generate the input data. For attribute levels not used to generate the input data, the utility function approach also appears to be superior to the conventional approach of interpolating between discrete attribute levels. A procedure to numerically determine the usefulness of estimating a utility function by comparing the predictive performances of regression analysis and linear interpolation was presented in this paper. Regression analysis was used instead of a technique such as linear programming because the comparison with interpolation could be made analytically. In other words, the use of regression made it possible to obtain an estimate of the magnitude of improvement that could be expected from the use of utility functions without having to resort to a costly and time-consuming simulation.

In terms of mean Kendall τ for a set of holdout attribute levels, the regression utility function was superior to interpolation for predictive purposes in the cases analyzed in this paper. Of greater interest to marketing managers is the translation of the Kendall τ figures into improvements in the market share of a new product. For the case analyzed in the paper, there was a 12.6 per cent increase in expected market share if Regression is used instead of Interpolation, while the variance in market share was reduced by 87 per cent.

APPENDIX A
VARIANCE - COVARIANCE MATRIX OF RANK ORDER
PREDICTIONS FOR REGRESSION ANALYSIS

The predicted utilities $\hat{Y}_a$, $\hat{Y}_b$, and $\hat{Y}_c$ are determined from equation (1) using the regression estimates: $\hat{q}$ and $\hat{r}$. The method of determining the elements of Σ will be illustrated for the rank order [BAC] where

$$s_1 = (\hat{Y}_b - \hat{Y}_a) \quad \text{and} \quad s_2 = (\hat{Y}_a - \hat{Y}_c).$$

$$\begin{aligned} \text{Let } x_{ba} &= (x_b - x_a) & , & & \bar{x}_{ba} &= (x_b^2 - x_a^2) \\ x_{ac} &= (x_a - x_c) & , & & \bar{x}_{ac} &= (x_a^2 - x_c^2) \end{aligned}$$

Using the above notation:

$$\begin{aligned} \text{Var } s_1 &= \text{Var } [x_{ba} \hat{q} + \bar{x}_{ba} \hat{r}] \\ &= x_{ba}^2 \text{Var } \hat{q} + \bar{x}_{ba}^2 \text{Var } \hat{r} + 2x_{ba} \bar{x}_{ba} \text{Cov } \hat{q} \hat{r} \end{aligned}$$

Similarly:

$$\text{Var } s_2 = x_{ac}^2 \text{Var } \hat{q} + \bar{x}_{ac}^2 \text{Var } \hat{r} + 2x_{ac} \bar{x}_{ac} \text{Cov } \hat{q} \hat{r}$$

$$\text{Now, Cov } s_1 s_2 = E(s_1 - \bar{s}_1)(s_2 - \bar{s}_2)$$

where $\bar{s}_1$ and $\bar{s}_2$ are the expected values of s_1 and s_2 respectively.

Hence:

$$\begin{aligned}
 \text{Cov } s_1 s_2 &= E[X_{ba}(\hat{q} - q) + \bar{X}_{ba}(\hat{r} - r)][X_{ac}(\hat{q} - q) + \bar{X}_{ac}(\hat{r} - r)] \\
 &= E[X_{ba}X_{ac}(\hat{q} - q)^2 + \bar{X}_{ba}\bar{X}_{ac}(\hat{r} - r)^2 + \\
 &\quad (X_{ba}\bar{X}_{ac} + \bar{X}_{ba}X_{ac})(\hat{q} - q)(\hat{r} - r)] \\
 &= X_{ba}X_{ac} \text{Var } \hat{q} + \bar{X}_{ba}\bar{X}_{ac} \text{Var } \hat{r} + (X_{ba}\bar{X}_{ac} + \bar{X}_{ba}X_{ac}) \text{Cov } \hat{q} \hat{r}
 \end{aligned}$$

X_{ba} , $\bar{X}_{ba}$, X_{ac} , and $\bar{X}_{ac}$ are known. However, to compute $\text{Var } s_1$, $\text{Var } s_2$, and $\text{Cov } s_1 s_2$, it is necessary to know $\text{Var } \hat{q}$, $\text{Var } \hat{r}$, and $\text{Cov } \hat{q} \hat{r}$, the elements of $\text{COV } \hat{q} \hat{r}$, the variance-covariance matrix for $\begin{bmatrix} \hat{q} \\ \hat{r} \end{bmatrix}$. $\text{COV } \hat{q} \hat{r}$ can be calculated from

the standard regression formula (see, for example, [7, p. 357]):

$$\text{COV } \hat{q} \hat{r} = \sigma^2 (X'X)^{-1}$$

Given the utility function (1), the X matrix is:

$$X = \begin{bmatrix} X_1 & X_1^2 \\ \vdots & \vdots \\ X_n & X_n^2 \end{bmatrix}$$

where X_i = attribute level of the i^{th} observation, and

n = number of available observations.

Therefore:

$$\text{COV } \hat{q} \hat{r} = \sigma^2 \left\{ \left[\frac{1}{\sum X_i^2 \sum X_i^4 - (\sum X_i^3)^2} \right] \begin{bmatrix} \sum X_i^4 & -\sum X_i^3 \\ -\sum X_i^3 & \sum X_i^2 \end{bmatrix} \right\}$$

The elements of Σ can be computed similarly for the other five rank orders. The variance-covariance matrix, $\text{COV } \hat{q} \hat{r}$, will be the same for all six rank orders for a given data set. However, $\text{Var } s_1$, $\text{Var } s_2$, and $\text{Cov } s_1 s_2$ will be different since s_1 and s_2 are defined differently for each rank order.

APPENDIX B
VARIANCE - COVARIANCE MATRIX OF RANK
ORDER PREDICTIONS FOR INTERPOLATION

Consider, again, the rank order [BAC] where $s_1 = (\hat{Y}_b - \hat{Y}_a)$ and

$$s_2 = (\hat{Y}_a - \hat{Y}_c). \quad \text{Now:}$$

$$\text{Var } s_1 = \text{Var } \hat{Y}_b + \text{Var } \hat{Y}_a - 2 \text{Cov } \hat{Y}_a \hat{Y}_b \quad (5)$$

$$\text{Var } s_2 = \text{Var } \hat{Y}_a + \text{Var } \hat{Y}_c - 2 \text{Cov } \hat{Y}_a \hat{Y}_c \quad (6)$$

Therefore, it is necessary to determine the quantities $\text{Var } \hat{Y}_b$, $\text{Var } \hat{Y}_a$, $\text{Var } \hat{Y}_c$, $\text{Cov } \hat{Y}_a \hat{Y}_b$, and $\text{Cov } \hat{Y}_a \hat{Y}_c$. Start by determining $\text{Var } \hat{Y}_b$, $\text{Var } \hat{Y}_a$, and $\text{Var } \hat{Y}_c$. For this purpose, consider again the utility function of equation

(1):

$$Y_i = qX_i + rX_i^2 + \epsilon$$

where ϵ is normally distributed, $E(\epsilon) = 0$, and $\text{Var}(\epsilon) = \sigma^2$ for all values of X .

It is desired to predict the value of Y_a corresponding to attribute level X_a such that $X_i \leq X_a \leq X_{i+1}$ where X_i and X_{i+1} are the attribute levels of two of the n available sample observations. The linear interpolation scheme yields the following prediction of Y_a :

$$\hat{Y}_a = Y_i + w(Y_{i+1} - Y_i) \quad (7)$$

where
$$w = \frac{X_a - X_i}{X_{i+1} - X_i} \quad (8)$$

This is illustrated in Figure 2. Equation (7) can be rewritten as:

$$\hat{Y}_a = wY_{i+1} + (1-w)Y_i \quad (9)$$

FIGURE 2

Prediction Using Linear Interpolation

Substituting (1) in (9):

$$\begin{aligned} \hat{Y}_a &= w(qX_{i+1} + rX_{i+1}^2 + \epsilon_{i+1}) + \\ &\quad (1 - w)(qX_i + rX_i^2 + \epsilon_i) \end{aligned} \quad (10)$$

Substituting the value of w from (8) in (10), and simplifying:

$$\hat{Y}_a = Z_a + \epsilon_k \quad \text{where} \quad (11)$$

$$Z_a = qX_a + r \{X_{i+1}(X_a - X_i) + X_a X_i\}$$

$$\epsilon_k = w\epsilon_{i+1} + (1 - w)\epsilon_i \quad (12)$$

The variance of the prediction is given by:

$$\sigma_I^2 = E\{(Z_a + \epsilon_k) - E(Z_a + \epsilon_k)\}^2 = E\{Z_a + \epsilon_k - Z_a\}^2 = E(\epsilon_k)^2 \quad (13)$$

Hence, $\sigma_I^2 = \sigma_{\epsilon_k}^2$. From (12):

$$\sigma_I^2 = w^2 \sigma^2 + (1 - w)^2 \sigma^2 \quad \text{or}$$

$$\sigma_I^2 = \sigma^2(2w^2 - 2w + 1) \quad (14)$$

Note that for $w = 0.5$, σ_I^2 achieves its minimum value of $(0.5\sigma^2)$ while for $w = 0$ or $w = 1$, σ_I^2 takes on its maximum value of σ^2 . In other words, the variance of the prediction for the linear interpolation scheme is greatest when X_a is identical to a sample value, X_i , while the variance of the prediction is least when X_a is exactly between two sample values: X_i and X_{i+1} .

Thus, using equation (14):

$$\text{Var } \hat{Y}_b = \sigma^2(2w_b^2 - 2w_b + 1)$$

and similarly for $\text{Var } \hat{Y}_a$ and $\text{Var } \hat{Y}_c$. Note that w in equation (8) is equivalent to w_a .

Returning to equations (5) and (6), it can be seen that $\text{Cov } \hat{Y}_a \hat{Y}_b$ and $\text{Cov } \hat{Y}_a \hat{Y}_c$ have still to be determined. To show how this can be done, consider Figure 3 where a set of sample locations of the holdout concepts: X_a , X_b , and X_c , are indicated relative to four of the original observations: $X_i, \dots, X_{i+3}$. From equations (11) and (12):

$$\hat{Y}_a = Z_a + w_a \epsilon_{i+3} + (1 - w_a) \epsilon_{i+2}$$

$$\hat{Y}_b = Z_b + w_b \epsilon_{i+2} + (1 - w_b) \epsilon_{i+1}$$

FIGURE 3

Sample Locations of Holdout Concepts

$$\text{Now, } \text{Cov}(\hat{Y}_a \hat{Y}_b) = E[(\hat{Y}_a - \mu_a)(\hat{Y}_b - \mu_b)]$$

where μ_a and μ_b are the expected values of $\hat{Y}_a$ and $\hat{Y}_b$ respectively.

$$\begin{aligned} \text{Hence, } \text{Cov}(\hat{Y}_a \hat{Y}_b) &= E[(w_a \epsilon_{i+3} + (1 - w_a) \epsilon_{i+2})(w_b \epsilon_{i+2} + (1 - w_b) \epsilon_{i+1})] \\ &= E[w_b(1 - w_a) \epsilon_{i+2}^2] = w_b(1 - w_a) \sigma^2 \end{aligned}$$

$$\text{In general, } \text{Cov}(\hat{Y}_a \hat{Y}_b) = \begin{cases} w_b(1 - w_a) \sigma^2 & \text{if } X_a \text{ and } X_b \text{ are adjacent} \\ 0 & \text{if } X_a \text{ and } X_b \text{ are not adjacent} \end{cases}$$

" X_a and X_b are adjacent" means that a common data point (X_{i+2} in this case) is used to determine $\hat{Y}_a$ and $\hat{Y}_b$. Hence, in terms of the locations of X_a , X_b , and X_c indicated in Figure 3:

$$\text{Cov}(\hat{Y}_b \hat{Y}_c) = w_c(1 - w_b) \sigma^2$$

$$\text{Cov}(\hat{Y}_a \hat{Y}_c) = 0$$

Thus, for Interpolation, one can calculate $\text{Var } s_1$ and $\text{Var } s_2$ for all six rank orders. What remains to be done is the calculation of $\text{Cov } s_1 s_2$ for each of the six rank orders. Again, the procedure for computing

$\text{Cov } s_1 s_2$ will be illustrated for the rank order [BAC] where $s_1 = (\hat{Y}_b - \hat{Y}_a)$

and $s_2 = (\hat{Y}_a - \hat{Y}_c)$. From equations (11) and (12) and Figure 3:

$$s_1 = (Z_b - Z_a) + w_b \epsilon_{i+2} + (1 - w_b) \epsilon_{i+1} - w_a \epsilon_{i+3} - (1 - w_a) \epsilon_{i+2}$$

$$s_2 = (Z_a - Z_c) + w_a \epsilon_{i+3} + (1 - w_a) \epsilon_{i+2} - w_c \epsilon_{i+1} - (1 - w_c) \epsilon_i$$

$$\text{Cov } s_1 s_2 = E[(s_1 - \bar{s}_1)(s_2 - \bar{s}_2)]$$

where $\bar{s}_1$ and $\bar{s}_2$ are the expected values of s_1 and s_2 respectively.

Hence:

$$\begin{aligned}
 \text{Cov } s_1 s_2 &= E[\{\epsilon_{i+1}(1 - w_b) + \epsilon_{i+2}(w_b - 1 + w_a) - \epsilon_{i+3}w_a\}\{-\epsilon_i(1 - w_c) \\
 &\quad - \epsilon_{i+1}w_c + \epsilon_{i+2}(1 - w_a) + \epsilon_{i+3}w_a\}] \\
 &= E[\epsilon_{i+2}^2(w_b - 1 + w_a)(1 - w_a) - \epsilon_{i+1}^2(1 - w_b)w_c - \epsilon_{i+3}^2w_a^2] \\
 &= \sigma^2[(w_b - 1 + w_a)(1 - w_a) - (1 - w_b)w_c - w_a^2]
 \end{aligned}$$

It is easy to see that the above expression also holds for $\text{Cov } s_i s_2$ for the rank order [CAB]. Similarly, it can be shown that for rank orders [ABC] and [CBA]:

$$\text{Cov } s_1 s_2 = \sigma^2[(1 - w_b)(w_b + w_c - 1) + w_b(1 - w_a - w_b)]$$

Also, for rank orders [ACB] and [BCA]:

$$\text{Cov } s_1 s_2 = \sigma^2[w_b(w_a - 1) + w_c(1 - w_b - w_c) - (1 - w_c)^2]$$

Thus, the elements of the variance-covariance matrix, Σ , can be computed for all six rank orders if interpolation is used to make the utility predictions.

FOOTNOTES

1. See [5, pp. 54-59, 114-126] and [6] for descriptions of the technique and [2], [3], and [5, pp. 133-215] for some marketing applications.
2. When the true utility function is indeed smooth, the assumption of a piecewise linear function has a major drawback. Its derivatives are not continuous and, therefore, information is not transferred from one linear segment to another. This reduces its efficiency as an estimation tool.
3. If, for some reason, piecewise linear functions must be used, researchers should at least estimate the slopes and starting points of each linear segment (see, for example, [1] and [10]). However, estimation of these parameters is not a simple matter because it is necessary to use computationally-complex techniques such as dynamic programming [1].
4. The quadratic function is a unimodal function and is a reasonable specification of an individual consumer's utility function for a given attribute. We can, of course, obtain a multi-modal function for an attribute if the utility function is estimated empirically by aggregating the responses of several individuals (see, for example, the preference distribution for sudsiness in [9, Exhibit 2]).

FIGURE 1

UTILITY FUNCTIONS FOR 'TOP SPEED'

TABLE 1

KENDALL τ FOR THE SIX POSSIBLE RANK
ORDERINGS OF THREE OBJECTS

Possible Rank Orders	[ABC]*	[ACB]	[BAC]	[BCA]	[CAB]	[CBA]
Number of Violations	0	1	1	2	2	3
Kendall τ	1.00	0.33	0.33	-0.33	-0.33	-1.00

*The rank order [ABC] means that A is most preferred, B is preferred next, and C is least preferred.

TABLE 2

COMPARISON OF REGRESSION AND INTERPOLATION
 FOR THE QUADRATIC UTILITY FUNCTION: $Y = 4X - 0.5X^2$

Holdout Attribute Levels and Mean Predicted Utilities	Variance, ^a σ^2	Mean of Kendall τ		Improvement ^b	
		Regression	Interpolation		
$X_a = 4.4, \mu_a = 7.920$ $X_b = 3.5, \mu_b = 7.875$ $X_c = 2.6, \mu_c = 7.020$	1	0.619	0.496	24.8	Hol and
	4	0.465	0.286	62.6	X_a
	16	0.281	0.149	89.7	X_t
	49	0.169	0.086	96.5	X_c
$X_a = 3.9, \mu_a = 7.995$ $X_b = 2.8, \mu_b = 7.280$ $X_c = 1.9, \mu_c = 5.795$	1	0.974	0.737	32.2	
	4	0.863	0.454	90.2	X_c
	16	0.681	0.242	181.0	X_t
	49	0.507	0.140	262.0	X_c

^a

Variance of the error term in the utility function (see equation 1 in the text)

^bPercentage improvement of Regression over Interpolation in terms of mean Kendall τ

a

b

TABLE 3

COMPARISON OF REGRESSION AND INTERPOLATION

FOR THE LINEAR UTILITY FUNCTION: $Y = 4X$

Holdout Attribute Levels and Mean Predicted Utilities	Variance ^a , σ^2	Mean of Kendall τ		Improvement ^b
		Regression	Interpolation	
$X_a = 4.4, \mu_a = 17.6$ $X_b = 3.5, \mu_b = 14.0$ $X_c = 2.6, \mu_c = 10.4$	1	1.000	1.000	0.0
	4	1.000	0.996	0.4
	16	0.978	0.864	13.2
	49	0.841	0.610	37.8
$X_a = 3.9, \mu_a = 15.6$ $X_b = 2.8, \mu_b = 11.2$ $X_c = 1.9, \mu_c = 7.6$	1	1.000	1.000	0.0
	4	1.000	0.949	5.4
	16	0.998	0.713	40.0
	49	0.951	0.469	103.0

^aVariance of the error term in the utility function (see equation 1 in the text)^bPercentage improvement of Regression over Interpolation in terms of mean Kendall τ

REFERENCES

1. Bellman, Richard and Robert Roth. "Curve Fitting by Segmented Straight Lines," Journal of the American Statistical Association, 64 (September 1969), 1079-84.
2. Davidson, J. D. "Forecasting Traffic on STOL", Operational Research Quarterly, 24 (December 1973), 561-69.
3. Fiedler, John A. "Condominium Design and Pricing", in M. Venkatesan, ed., Proceedings of the Third Annual Conference (Association for Consumer Research, 1972), 279-93.
4. Green, Paul E. "On the Design of Choice Experiments Involving Multifactor Alternatives", Journal of Consumer Research, 1 (September 1974), 61-68.
5. _____, and Yoram Wind. Multiattribute Decisions in Marketing. Hinsdale, Illinois: The Dryden Press, 1973.
6. Johnson, Richard M. "Trade-Off Analysis of Consumer Values", Journal of Marketing Research, 11 (May 1974), 121-27.
7. Kmenta, Jan. Elements of Econometrics. New York, New York: Macmillan, 1971.
8. Kruskal, Joseph B. "Analysis of Factorial Experiments by Estimating Monotone Transformations of the Data", Journal of the Royal Statistical Society, Series B, 27 (1965), 251-63.
9. Kuehn, Alfred A. and Ralph L. Day. "Strategy of Product Quality," Harvard Business Review, 40 (November-December 1962), 100-10.
10. McGee, Victor E. and Willard T. Carleton. "Piecewise Regression," Journal of the American Statistical Association, 65 (September 1970), 1109-24.
11. Pekelman, Dov and Subrata K. Sen. "Measurement and Estimation of Conjoint Utility Functions," Working Paper, Graduate School of Management, University of Rochester (July 1978)
12. Winkler, Robert L. and William L. Hays. Statistics: Probability, Inference, and Decision. (2nd ed.) New York: Holt, Rinehart, and Winston, 1975.