No More Static

By Nicholas Fisher, Alan Lee, and Ross Sparks

Value surveys have become increasingly popular in recent years as a means of monitoring how a company is performing in its chosen markets. During the last decade, marketing research into the interpretation, modeling, normative, and other aspects of customer value has ramified in many ways. In a 2001 *Journal of Consumer Marketing* article, co-authors F. Huber, A. Herrmann, and R.E. Morgan noted, "creating superior customer value is a necessary precondition for securing a niche in a competitive environment, not to mention a leadership position in the market." The way in which value is defined and modeled for quantitative purposes determines many, but not all, of the features of an attendant value survey. Some features are common to all value surveys, particularly those relating to market segmentation and survey design.

We believe there is much to be gained by exploiting possibilities afforded by continuous acquisition and analysis of survey data. Although the methods we advocate have been developed in the context of a specific approach to measuring and monitoring customer value, they would be applicable to other survey methodologies up to the point of detailed implementation.

We take as our basic approach to measuring customer value the value-tree-based methodology developed in considerable detail by Ray Kordupleski at AT&T in the late 1980s and des-

Continuously track customer satisfaction, so you can respond quickly to changing customer needs.

Executive Summary

Most monitoring of markets is still done by static survey— monthly, quarterly, or yearly. There is, however, much to be gained from acquiring and analyzing customer preference data on a continuous basis, in terms of timely response to changes in customer satisfaction.

scribed in his 2002 book with Janice Simpson, *Mastering Customer Value Management*. In this approach, value derives from the two key drivers: (1) satisfaction with quality of product or service received and (2) satisfaction with price paid. Because the whole market is surveyed, it's possible to assess the relative value added by the company, also known as "customer value added" (CVA). Further, the same approach can be used with surveys of employees and with assessing the company's performance with its other key stakeholder groups (e.g., its suppliers, strategic partners, and the community) in a more general performance measurement setting, as suggested by Stan Dransfield and his co-authors in a 1999 *International Statistical Review* article. In fact, a general value analysis approach then becomes a key component of an integrated performance measurement system.

In a value survey, the data comprise a sample of tree-structured observations, with each observation tree consisting of a set of integer scores typically in the range of one to 10. The structure of an individual observation tree is shown in Exhibit 1, for a value survey of customer experiences in acquiring and servicing a mortgage.

As described by Kordupleski and Simpson in *Mastering Customer Value Management*, the tree structure confers some important operational characteristics on the associated survey. For example:

- The tree structure recognizes a natural hierarchy of satisfaction drivers, and so facilitates statistical modeling and analysis of the data to identify the most important drivers of overall customer satisfaction.
- The tree structure connects the ultimate outcome (worth what paid for) to the core business process (the sequence of customer experiences, or key processes, of talking to the loan manager, applying for a loan, settling, and repaying) that delivers the outcome, and so leads directly to process improvement. This provides an interesting connection between two of Richard Normann's fundamental constructs: value added for customers and moments of truth.
- The resulting data need to be analyzed in such a way as to obtain results similar to those in Exhibit 2, where the weights indicate the relative importance of the various factors in determining a customer's overall satisfaction with the super-ordinate attribute, and the summary ratings indicate current competitive performance. The most important point about such summary tables is that they are easily interpreted by managers, who need to know where to focus improvement priorities in order to get the biggest impact on their enterprise. Our purpose is to show how a new approach to data acquisition and analysis can yield significant gains in terms of timeliness of management action.

CVA Analysis

In a typical CVA campaign, data are available from the survey company at regular intervals, for example every three months. For a given market segment, some 150 responses may be acquired for each company being assessed, depending on the precision required in the final estimates.

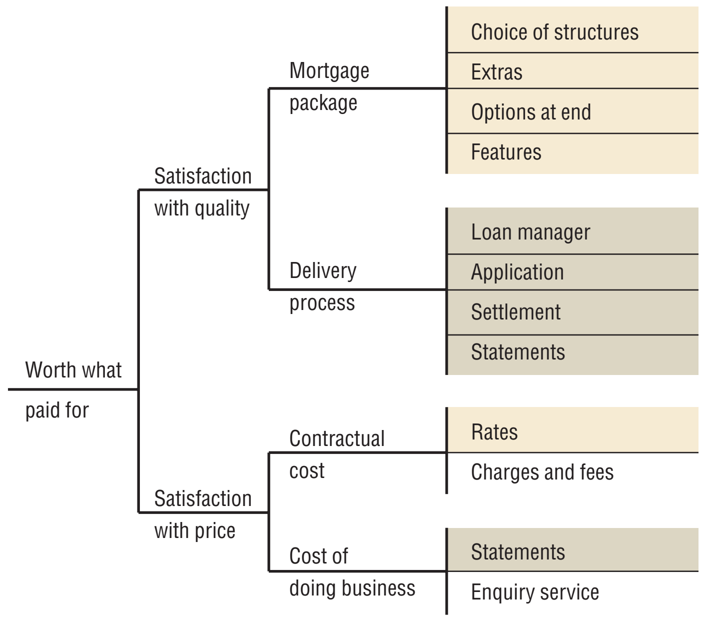
Kordupleski and Simpson recommend modeling the entire data set as a set of hierarchical linear models, leading to the weights in Exhibit 2. (The method was originally developed by Kordupleski and Rich DeNicola.) Typically, modeling and analysis are restricted to the data for the current quarter, although survey data for a 12-month period are sometimes studied in aggregate. The method of calculating the weights does not guarantee that they will be positive or add to something less than 100%. In practice, however, improvements in attributes generally lead to improvements in the higher-level responses, so the coefficients tend to be positive. (If they are not, this usually points to a problem in the survey design.) It is important to bear in mind the ultimate goal: to provide guidance about setting priorities, based on predicting the likely effects of improving performance on different attributes. This simple approach has proved very effective in practice in two critical ways:

1. Guiding the selection of improvement priorities, leading to improved CVA scores and improved business performance for many organizations
2. Helping managers at all levels in an organization understand and use their own customer value data

There is much to be gained from focusing more attention on actual selection of the response times.

Can more information be extracted from the data without losing these two critical benefits? In 1997, Linda Clark and co-workers at Lucent Technologies reported the results of a major study of customer value survey data, the results of which were presented at the Fourth Workshop on Case Studies in Bayesian Statistics at Carnegie Mellon University. The Lucent group built a complex statistical model for the whole data set, extending back over as many survey periods as possible. Although the approach is very comprehensive, it does not

Exhibit 1 Customer value tree for a mortgage



produce the readily actionable results consistent with the two points above. Exhibit 3 on page 18 compares the relative merits of the DeNicola-Kordupleski approach and the Lucent approach.

Continuous Customer Value

A useful starting point for an alternative approach to modeling and analyzing CVA data is to look at the basic issue of how the survey data are acquired. Typically, call center operators make telephone calls at various times of the day, throughout the week, to ensure appropriate coverage of target markets. The analyst then receives a data set every three months, comprising the consolidated responses over that period. Each response is usually tagged with the time at which the response was acquired. There is little published evidence that the additional data actually are used. However, we contend that there is much to be gained from focusing more attention on actual selection of the response times, by including a requirement for continuous acquisition of survey data as part of the survey design and using the additional information in subsequent statistical modeling and analysis.

Before developing a procedure that uses this additional information, it's helpful to list the explicit goals of the research:

- Obtain results similar to those in Exhibit 1, in which the weights convey a sense of the relative importance of the various factors and the summary ratings have associated estimates of precision, either in the form of standard errors or expressed as confidence intervals.
- Report the results in the original scale.
- Detect interesting temporal changes, either in ratings or in weights.
- Provide easily interpreted trend charts that show what's going on.

To illustrate the basic approach, we start with the simplest situation: what happens at the top of the value tree. At any instant of time t we may acquire one or more observation trees. At each node of a tree, there will be a response and a set of associated drivers. For example, at the top of the tree, the response is "worth what paid for," and the drivers are "satisfaction with quality" and "satisfaction with price."

At each node of an observation tree, and at any given time t, we assume that the response and the drivers are related by a model of the form:

$$Y_t = w_{1t}D_{1t} + \ldots + w_{kt}D_{kt} + w_tD_t$$

where Y_t is a generic response at a node, $D_{1t}, \ldots, D_{kt}$ are explanatory variables (drivers) for the response Y_t, $w_{1t}, \ldots w_{kt}$ are unknown weights attached to these drivers, and D_t is an unobserved driver representing all the other unmeasured factors that have an effect on the attribute Y_t together with its unknown weight w_t. It's assumed that the weights are non-negative and sum to one. Because data are to be acquired continuously, it makes sense to allow for the possibility that both the ratings and the weights may change with time. The times need not be equally spaced, and no complication is introduced by collecting more than one response at a particular time.

We have devised two methods to give continuously updated estimates of the weights. The first method uses the Kalman filter. The second refits a regression model each time new data are acquired, using exponentially declining regression weights. Both seem to give satisfactory results, with the latter method being easier to implement. The ratings are updated by smoothing the current and past data using a loess (or locally weighted regression) smoother.

Recommendation Evaluation

We have conducted numerous computer experiments to test the efficacy of the proposed approach in contexts that mimic

Exhibit 2 Prototype reporting format

Key purchase criteria	Importance (%)	Your company	Competition	Ratio (%)
Access	25	9.0	8.6	105
Contractual	20	8.1	8.1	100
Project work	12	8.5	8.5	100
Delivery	14	7.8	8.1	96
Payment	19	7.0	8.0	88
Follow up	10	8.4	8.2	102
Overall	100	8.1	8.3	98

Exhibit 3 Comparison of two current approaches

	DeNicola–Kordupleski approach
Advantages	**Simple** Results interpretable by all levels of management Managers can use the results themselves effectively (running "what if?" scenarios) with simple software
Disadvantages	**Approximate** May be inefficient: may not capture "all" the variation Interval estimates may not be range-preserving Uncertainty in results unclear and not reported Estimate is for middle of past measurement period
	Lucent approach
Advantages	A complete statistical analysis Appears to model the sources of variation well Provides prediction of precision of a CV survey Trellis graphics reveal/highlight subtle differences
Disadvantages	Statistical analysis requires advanced modelling and analysis skills Results not reported on original scale Presentation and interpretation of results requires an expert Results not as accessible to management for evaluation: not usable (actionable) by managers themselves No software available to work with results

practical application. The experiments were designed to mimic as closely as possible the sorts of data that arise in real consumer studies, in terms of the:

- Distributional properties of the simulated data
- Sorts of trends in ratings and relative weightings observed in practice
- Amounts of data usually acquired

We generated data to replicate the top level of the value tree, using the model $Y_t = w_{1t}D_{1t} + \ldots + w_{kt}D_{kt} + w_tD_t$ discussed earlier. Thus, in these simulations, $k = 2$, and Y_t, D_{1t}, and D_{2t} are the values of value, quality, and price at time t. The objective was to examine how well the estimated impact weights tracked the true values of the impact weights as these changed over time, which has obvious implications for timely management response to shifting market ratings or preferences. The experiment was designed similarly to a typical CVA campaign, with an initial pilot survey followed by continuous acquisition of data during a period of 12 months. The sample size of the dynamic survey was fixed at 600 respondents, assuming about 150 respondents per quarter.

Values of P (satisfaction with price) and Q (satisfaction with quality) were simulated using power binomial distribution, described in a 2002 *Applied Stochastic Models in Business & Industry* article by N.I. Fisher. In our experience with actual CVA surveys, we have found this distribution provides an acceptable model for random samples of CVA data. The means of the distributions used for the continuous survey were allowed to vary in time, with a mixture of periods of trend and constancy, and also abrupt jumps, such as the sorts of market responses observed in practice (as a result of advertising campaigns, sudden disasters, and so on). The means for the static survey were constant.

The response variable Y_t ("worth what paid for," that is, value measured at time t) was generated according to the model:

$$Y_t = a_t\, c\, P_t + c\,(1 - a_t)\, Q_t + (1 - c)\, D_t$$

In this case:

- The parameter c ($0 = c = 1$) represents the goodness of fit of the model. The larger c, the better the value score can be predicted by the scores for quality and price.
- The parameter a ($0 = a = 1$) represents the relative importance of quality and price. The larger a, the more important is price in determining the perception of value.

The parameter c was kept fixed throughout at $c = 0.7$. This value was chosen to correspond to the degree of fit obtained in practice, in typical value surveys. For the preliminary survey, the relative impact of weight at was also assumed constant. For the ongoing survey, we considered three scenarios for the relative weight a_t. The first scenario (constant) assumed no change over the year. The second (trend) assumed that a_t had a smooth linear trend over time. The third (jump) assumed a sudden level change in a_t.

Exhibit 4 shows sample time series plots of the impact weights for price, for a typical set of data for each of the three scenarios. The method can be adjusted to trade off between rapid adaptation to change, and accurate estimation in the absence of change. The trade off is controlled by the decay parameter c that defines the weights. The smaller c, the less accuracy but the more rapid the adaptation to change. Based on extensive simulations, value of c in the range 0.90 to 0.95 appears to provide an acceptable compromise.

The average ratings are as important as the impact weights. Traditionally, average ratings have been estimated by assuming that they are constant over the period of the survey, and using the overall sample mean for that period as an estimate. This approach fails to allow for change in mean rating level within a survey period and will generally result in a biased estimate if the mean level drifts. Moreover, using the standard formula for the standard error of the mean gives an optimistic impression of the error actually made because it makes no allowance for the bias.

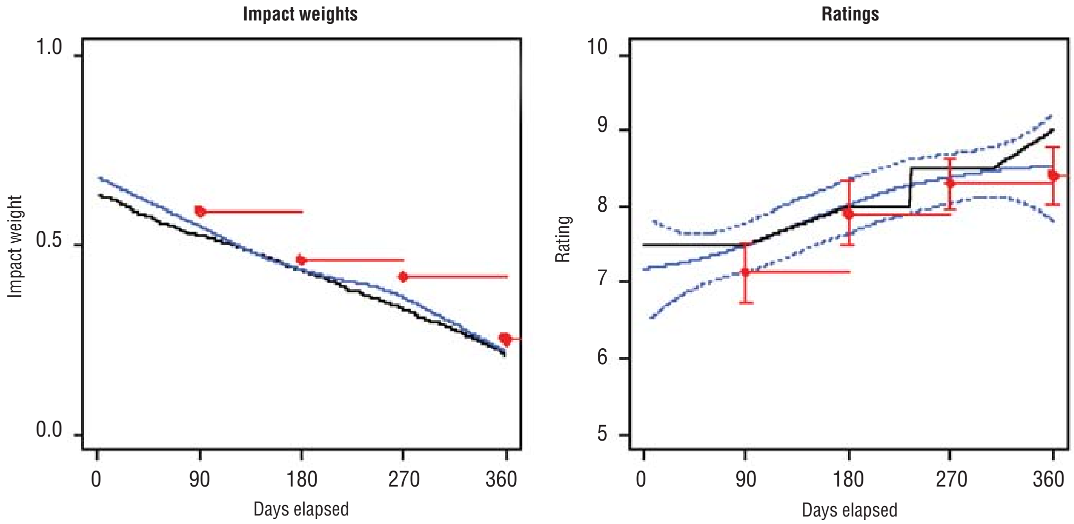
In contrast, our approach regards mean ratings as functions of time that we need to estimate. A modern statistical smoothing method (such as the loess method described by William S. Cleveland and Susan J. Devlin in a 1988 *Journal of the American Statistical Association* article) provides an approximately unbiased estimate of the mean rating in the presence of change, and a realistic estimate of its error. And of course, there is the obvious benefit of detecting any trend in mean rating over time.

Improvements Made

It is interesting to contrast our results with what would be obtained using a traditional approach, where data are analyzed quarterly, with no other use being made of the dates of their acquisition. Exhibit 4 shows the results from a synthetic example, in which the true relative importance (impact weight a_t) of quality relative to price is declining steadily while the rating is increasing at the same time.

At the end of each quarter, 150 respondents are surveyed, either continuously or en bloc. The black line represents the estimate of the trend curve for a_t. Because continuously acquired data are available for analysis at any time, the estimate of a_t can be updated whenever desired and timely actions are taken. On the other hand, if one relies simply on quarterly surveys, the estimates of impact weight can be significantly out of date within a relatively short period of time. Similar remarks attach to the estimated ratings.

Exhibit 4 Impact weights and mean ratings



Impact weights and mean ratings estimated by the continuous method (blue line) and quarterly method (red line), for a typical customer survey sampling protocol. The black line in the impact weights graph on the left shows a steady decline in the relative importance of quality, compared with price. But the rating of quality is increasing in a somewhat erratic fashion, as shown by the black line in the right hand graph. Continuous surveying and analysis allows for changes in impact weights or rating to be detected and managed in a timely fashion.

Overall, the results of the simulations are very encouraging indicators of the new approach's performance. We believe that the new approach:

- Meets all the requirements demanded of it
- Leads to more informative reporting of results
- Provides an improved basis for timely management action, since important trends both in ratings of factors, and in their relative importance, can be detected and managed in real time
- Results in significant savings in survey costs, because the daily survey load is easily piggy-backed onto other surveys

This approach has already produced excellent results in the context of staff satisfaction surveys and should do so for customer markets, particularly those in which consumer preferences move rapidly. ●

Additional Reading

Desarbo W.S., K. Jedidi, and I. Sinha (2001), "Customer Value Analysis in Heterogeneous Market," *Strategic Management Journal*, 22, 845-857.

Dransfield S.B., N.I. Fisher, and N.J. Vogel (1999), "Using Statistics and Statistical Thinking to Improve Organizational Performance. *International Statistical Review*, 67, 99-150.

Eggert A. and W. Ulaga (2002), "Customer Perceived Value: A Substitute for Satisfaction in Business Markets?" *The Journal of Business and Industrial Marketing*, 17, 107-118.

Kass., R., ed. (1997), "Proceedings of the 4th Workshop on Case Studies in Bayesian Statistics," held at Carnegie Mellon University, Pittsburgh, Penn. (September 26-27).

Pihlens D. and L. Ricci (2002), "Understanding the Drivers of Customer Value: The Fusion of Macro and Micromodeling," *Journal of Targeting, Measurement and Analysis for Marketing*, 10, 269-281.

Authors' Note: The authors are most grateful to Ray Kordupleski and Rich DeNicola for helpful discussions about their methodologies.

Nicholas Fisher is principal of ValueMetrics Australia and visiting professor of statistics at the University of Sydney. He may be reached at nif@valuemetrics.com.au. **Alan Lee** is professor of statistics at the University of Auckland, New Zealand. He may be reached at lee@stats.auckland.ac.nz. **Ross Sparks** is a research scientist in CSIRO Mathematical and Information Sciences in Australia. He may be reached at Ross.Sparks@csiro.au.