Modern Analysis of Customer Satisfaction Surveys: *comparison of models and integrated analysis*

Ron S. Kenett[1] and Silvia Salini[2]

[1]KPA Ltd., Raanana, Israel and University of Torino, Torino, Italy
ron@kpa-group.com
[2]Department of Economics, Business and Statistics, University of Milan, Italy
silvia.salini@unimi.it

Abstract. Customer satisfaction is a key dimension driving business outcomes and performance of processes in service and product organizations. Measuring customer satisfaction is typically based on self-declared or interview based questionnaires where users or consumers are asked to express opinions on statements, or satisfaction scales, mapping out various interactions with the service provider or product supplier. The topic has gained importance in recent years with researchers proposing new models and methods for designing, implementing, and analyzing customer satisfaction surveys. The paper builds on material presented in a recent edited book entitled *Modern Analysis of Customer Satisfaction Surveys* (Kenett and Salini, 2011). The book provides a comprehensive exposition of a variety of models that have all been applied to the same data set by leading experts. These models generate a variety of management insights. Combining models opens up opportunities for further research and applications. Specifically, we suggest that an integrated analysis, aggregating several approaches to survey data analysis, may prove effective in increasing the information quality derived from of a customer satisfaction survey.

Keywords: Annual Customer Satisfaction Survey (ACSS), categorical data analysis, nominal data, ordinal data, statistical models, Information Quality (InfoQ).

1. Introduction

Self-declared or interview-based surveys are a prime research tool in many application areas such as social science research, marketing, service management, risk management and customer satisfaction management. In such surveys, customers are requested to fill in questionnaires with, typically, ten to eighty or even one hundred questions. A survey with n questions produces responses that can be considered as random variables, $X_1, \ldots, X_n$. Some of these variables, q of them, are responses to questions on overall satisfaction, recommendation or repurchasing intention, that are considered target variables. Responses to the other questions, $X_1, \ldots, X_k$, $k = n-q$, can be analyzed under the hypotheses that they are positively dependent with the target variables. The combinations (X_i, X_j), $X_i \in X_1, \ldots, X_{n-q}$, $X_j \in X_{n-q+1}, \ldots, X_n$, are either positive dependent or independent, for each pair of variable (X_i, X_j), $i \leq n-q$, $n-q < j \leq n$. In general, dependency patterns can be extracted from data by using data mining techniques and statistical models (Hand et al., 2001).

Modern analysis of customer satisfaction surveys includes methods such as Structural Equation Models, Bayesian Networks, Log Linear Models, CUB Models, Rasch models, Decision Trees, PLS Models, Non Linear PCA, Multidimensional Scaling, Multilevel models for ordinal data, Control Charts and Fuzzy Methods. Different sources of knowledge, such as subjective information from expert opinions or knowledge from literature and survey data can be integrated with Bayesian Networks and other methods. For examples see Kenett and Salini (2009, 2011) and Salini and Kenett (2009).

In this paper we present and compare several models for analyzing customer satisfaction survey data. We first introduce the models leaving details to the referenced papers and books. The models are than applied to the 2010 annual customer satisfaction survey (ACSS) data of the ABC Company. The qualitative outcomes and derived conclusions are compared using the concept of information quality (*InfoQ*) proposed in Kenett and Shmueli (2011). This comparative analysis provides insights on various models used in modern analysis of customer satisfaction survey data.

The ABC Company is a typical global supplier of integrated software, hardware, and service solutions to media and telecommunications service providers. The company develops, installs and maintains systems combining hardware, software and advanced electronics. These enabling systems support various lines of business, including video on demand, cable, and satellite TV, as well as a range of communications services, such as voice, video, data, Internet protocol, broadband, content, electronic, and mobile commerce. The company's workforce consists of more than 5,000 professionals located in 10 countries and serves customers in Europe and elsewhere. In the year 2010 the company launched its first annual customer satisfaction survey (ACSS) to provide feedback on all company touch points and interactions with customers. It covers topics such as *Equipment and Systems*, *Sales Support*, *Technical Support*, *Training*, *Supplies and Orders*, *Software Add On Solutions*, *Customer Website*, *Purchasing Support*, *Contracts and Pricing* and *System Installation*. Descriptive variables for each customer include: country, industry segment, age of the equipment from the ABC Company, profitability, customer seniority and position of respondent. The information on the ABC Company and its ACSS is derived from real applications with changes designed to ensure privacy and confidentiality, including the reference to the year of the survey as 2010.

The first part of the questionnaire consists of an assessment of overall satisfaction, with two specific variables evaluated on a five-point anchored scale; 1) a variable assessing repurchasing intentions and 2) a variable assessing likelihood of recommending the ABC Company to others; finally, a binary variable indicates if the ABC Company is considered the best supplier. In the second part of the questionnaire, there are almost fifty statements grouped according to various touch point topics representing interactions between the customer and the ABC Company. For each statement there are two types of scores: 1) the item evaluation score, based on a 1-5 scale where '1' stands for 'very dissatisfied' or 'strongly disagree' and '5' for 'very satisfied' or 'strongly agree', and 2) a measure of item importance based on a three-point anchored scale where '1' sands for 'low importance' and '3' for 'high importance'. Each topic is covered by specific items that set a context and an overall satisfaction question from the topic. Overall, the questionnaire consists of 81 questions. The 2010 ACSS questionnaire was emailed to 591 customers. The return rate was 45%, with representativeness of geographical regions determined by an M-Test (Kenett, 1991). For details on the questionnaire, the ABC Company and the models described below, see Kenett and Salini (2011).

2. The Models

This study is derived from an initiative of the Diego de Castro Statistics and Applied Mathematics Department of the University of Torino that was launched in 2005. Some of the models were originally presented in two special issues of *Quality Technology and Quantitative Management* (QTQM, 2008). The edited book *Modern Analysis of Customer Satisfaction Surveys: with applications using R* (Kenett and Salini, 2011) provides a comprehensive exposition of the models with reference to R applications, including a special R appendix. The models covered in that book are listed below, with the chapter numbers and their authors within brackets. In this paper we cover the four models presented in Chapters 11, 13, 14 and 20 that are italicized below:

- Causality Models (10, Mealli, Pacini, Rubin)
- *Bayesian Networks* (11, Kenett, Salini)
- Log Linear Models (12, Fienberg, Mandrique)
- *CUB Models* (13, Piccolo, Iannario)
- *The Rasch Model* (14, De Battisti, Nicolini, Salini)
- Tree-based Methods and Decision Trees (15, Soffritti, Galimberti)
- PLS Models (16, Boari, Cantaluppi)
- Non Linear PCA (17, Ferrari, Barbero)
- Multidimensional Scaling (18, Solaro)
- Multilevel Models for Ordinal Data (19, Rampichini, Grilli)
- *Control Charts Applications* (20, Kenett, Deldossi, Zappa)
- Fuzzy Methods (21, Zani, Morlini, Milioli).

We begin by a brief review of the four models considered here. Section 3 compares the analysis from these models using four dimensions of information quality. Section 4 presents an approach to integrate various models in order to enhance the analysis information quality and some concluding remarks. We begin with an introduction to Bayesian Networks.

2.1 Bayesian Networks

Bayesian Networks (BN) implement a graphical model structure known as a *directed acyclic graph* (DAG) that is popular in Statistics, Machine Learning and Artificial Intelligence. BN are both mathematically rigorous and intuitively understandable. They enable an effective representation and computation of the joint probability distribution (JPD) over a set of random variables (Pearl, 2000).

The structure of a DAG is defined by two sets: the set of nodes and the set of directed edges (arrows). The nodes represent random variables and are drawn as circles labeled by the variables names. The edges represent direct dependencies among the variables and are represented by arrows between nodes. In particular, an edge from node X_i to node X_j represents a statistical dependence between the corresponding variables. Thus, the arrow indicates that a value taken by variable X_j depends on the value taken by variable X_i. Node X_i is then referred to as a 'parent' of X_j and, similarly, X_j is referred to as the 'child' of X_i. An extension of these genealogical terms is often used to define the sets of 'descendants', i.e., the set of nodes from which the node can be reached on a direct path.

The structure of the acyclic graph guarantees that there is no node that can be its own ancestor or its own descendent. Such a condition is of vital importance to the factorization of the joint probability of a collection of nodes. Although the arrows represent direct causal connection between the variables, the *reasoning process* can operate on a BN by propagating information in any direction. A BN reflects a simple conditional independence statement, namely that each variable, given the state of its parents, is independent of its non-descendants in the graph. This property is used to reduce, sometimes significantly, the number of parameters that are required to characterize the JPD of the variables. This reduction provides an efficient way to compute the posterior probabilities given the evidence present in the data (Lauritzen et al, 1988, Pearl, 2000). In addition to the DAG structure, which is often considered as the "qualitative" part of the model, one needs to specify the "quantitative" parameters of the model. These parameters are described by applying the Markov property, where the conditional probability distribution at each node depends only on its parents. For discrete random variables, this conditional probability is often represented by a table, listing the local probability that a child node takes on each of the feasible values – for each combination of values of its parents. The joint distribution of a collection of variables can be determined uniquely by these local conditional probability tables.

BNs have been used to map cause and effect relationships between survey variables, see Kenett (2007), Kenett and Salini (2009) and Salini and Kenett (2009). This approach has proved very powerful. The application of BNs to the ABC Company

ACSS data is presented in section 3.1. For more details see Kenett, Perruca and Salini (2011).

2.2 CUB Models

Responses to customer satisfaction surveys are governed by specific experience and psychological considerations. When faced with discrete alternatives, people make choices by pairwise comparison of the items or by sequential removals. Such choices are affected by both uncertainty in the choice and pure randomness. Modeling the distribution of responses is far more precise than considering single summary statistics. Such considerations lead to the development of the CUB (Combination of uniform and shifted binomial random variables) model, originally proposed in Piccolo (2003). The CUB model is applied to the study of sampling surveys where subjects express a definite opinion selected from an ordered list of categories with m alternatives. The model differentiates between satisfaction level from an item and randomness of the final choice. These unobservable components are defined as *feeling* and *uncertainty*, respectively.

Feeling is the result of several factors related to the respondent such as country of origin, position in the company and years of experience. This is represented by a sum of random variables which converges to a unimodal continuous distribution. To model this, CUB models feeling by a shifted Binomial random variable, characterized by a parameter ξ and a mass b_r for response r where:

$$b_r(\xi) = \binom{m-1}{r-1} \xi^{m-r} (1-\xi)^{r-1} \text{ , } r = 1, 2, \ldots, m.$$

Uncertainty is a result of variables such as the time to answer, the degree of personal involvement of the responder with the topic being surveyed, the availability of information, fatigue, partial understanding of the item, lack of self-confidence, laziness, apathy, boredom etc… A basic model for these effects is a discrete Uniform random variable:

$U_r(m)=1/m$, $r = 1, 2, \ldots, m$.

The integrated CUB discrete choice model is:

$$\Pr(R = r) = \pi b_r(\xi) + (1-\pi) U_r(m) \text{ , } r = 1, 2, \ldots, m$$

$0 \leq \pi \leq 1$,

and

$$E(R) = \frac{(m+1)}{2} + \pi(m-1)\left(\frac{1}{2} - \xi\right)$$

In the ABC Company 2010 ACSS, *m*=5 in most questions. For more details see Iannario and Piccolo (2011).

2.3 The Rasch Model

The Rasch Model (RM) was first proposed in the 1960s to evaluate ability tests (Rasch, 1960). These tests are based on a set of the items and the assessment of the ability of a subject depends on two factors: his relative *ability* and the item's intrinsic *difficulty*. Subsequently the RM has been used to evaluate behaviors or attitudes. In this case, the two factors become the subject's *property* and the item's *intensity*, respectively. In recent years the model has been employed in the evaluation of services (De Battisti, Nicolini and Salini, 2005, 2010, 2011). In this context, the two factors become an individual customer *satisfaction* and the item (question) intrinsic level of *quality*. These two factors are measured by the parameters θ_i, referring to the satisfaction of person (customer) i, and β_j, referring to the quality of item (question) j. It is then possible to compare these parameters. Their interaction is expressed by the difference: $\theta_i - \beta_j$. A positive difference means that the customer' satisfaction is superior to the item's quality level. The difference $\theta_i - \beta_j$, determines the probability of a specific response to question *j*. In particular, in the dichotomous case where the question's response is '0' for 'not satisfied' and '1' for 'satisfied', the probability of a response $x_{ij} = 1$ by customer i with satisfaction level θ_i, when answering question j of quality β_j, is:

$$P\left\{x_{ij} = 1 \mid \theta_i, \beta_j\right\} = \exp\left(\theta_i - \beta_j\right) \Big/ \left(1 + \exp\left(\theta_i - \beta_j\right)\right) = p_{ij}$$

In the dichotomous model, data is collected in a *raw score matrix*, with *n* rows (one for each customer) and *J* columns (one for each question), whose values are equal to 0 or 1. The sum of each row $r_i = \sum_{j=1}^{J} x_{ij}$ represents the total score of customer i for all the items, while the sum of each column $s_j = \sum_{i=1}^{n} x_{ij}$ represents the score given by all the customers to the question j. The RM possesses several important properties. The first property is that the items measure only one latent feature (*one-dimensionality*). This is a limitation in the applications to customer satisfaction surveys where there are usually several independent dimensions. Another important characteristic of RM is that the answers to an item are independent of answers to other items (*local independence*). In the customer satisfaction survey context, this is an advantage. For parameters where no assumptions are made, by applying the logit transformation, $\log \frac{p_{ij}}{1 - p_{ij}}$, θ_i and β_j can be expressed on the same scale (*parameters linearity*); the estimations of θ_i and β_j are test and sample free (*parameters separability*); and the row and column totals on the raw score matrix are sufficient statistics for the estimation of θ_i and β_j (*sufficient statistics*). For more on these properties see Andrich (1988),

There are three main approaches to estimate parameters of a RM: *joint maximum likelihood* (JML), *conditional maximum likelihood* (CML) and *marginal maximum likelihood* (MML). In JML and CML person specific parameters are considered fixed effects, whereas in MML they are assumed random and independent variables drawn from a density distribution that describes the population. For an overview of parameter estimation techniques in the logistic models with one, two and three parameters see Baker (1987). For the examination of theoretical features linked to the existence and uniqueness of the Maximum Likelihood Estimates for the Rasch model see the papers by Bertoli-Barsotti (2003, 2005).

The Rasch dichotomous model has been extended to the case of more than two ordered categories such as a 1-5 Likert scale. This approach assumes that, between each category and the next, there is a threshold that qualifies the item's position as a function of the quality level presented by every answer category. A threshold is where two adjacent categories have the same probability to be chosen so that, for example, the probability to choose the first category is the probability not to exceed the first threshold. Thus, the answer to every threshold *h* of an item *j* depends on a value $\beta_j + \tau_h$, where β_j characterizes responses to item *j*. The second term represents the *h*-th threshold of β_j referring to the item *j*. The thresholds are ordered ($\tau_{h-1} < \tau_h$), because they reflect the category order. For more details see De Battisti, Nicolini and Salini (2011). This extension will allow us to model responses on a 1-5 scale.

2.4 Control Chart Applications

Perceived quality, satisfaction levels and customer complaints can be effectively controlled with control charts used in the context of statistical process control (SPC). SPC methods were originally developed in the 1920s to improve the quality of products. Control charts are generally classified into two groups. If the quality characteristic is measured on a continuous scale, we have a *control chart for variables*. When the quality characteristic is classified as conforming or not conforming on the basis of whether or not it possesses certain attributes, then *control charts for attributes* are used. For an introduction to basic and advanced control charts see Kenett and Zacks (1998). In analyzing customer satisfaction survey data, we can use control charts to identify a shift from previous surveys or investigate the achievement of pre-set targets. In general, we test the hypothesis: $\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta \neq \theta_0 \end{cases}$ where θ can be the mean, the standard error, or a proportion, depending on the particular kind and scope of the control chart (i.e., for variables or for attributes). All the above details also hold when we are interested in testing a specific shift of the parameter such as $\theta > \theta_0$ or $\theta < \theta_0$. In these cases, only one control limit, either upper control limit (UCL) or lower control limit (LCL), is reported on the control chart. Specifically, the p chart with control limits = $\overline{p} \pm k\sqrt{\overline{p}\left(1-\overline{p}\right)/n}$ is used to monitor the percentage of respondents who answered "5" (Very High) to a question on overall satisfaction, where n is the number of respondents and k is a constant multiplier of the binomial standard deviation used to set up the control limits. The value k=2 is often applied in applications of control charts to the analysis of customer satisfaction data. For more details see Kenett, Deldossi and Zappa (2011).

3. The ABC Company Survey Data Analysis: A Comparison of the Methods

As mentioned, the four models and techniques described above have been applied to the same problem, an analysis of the ABC Company 2010 ACSS data. In order to compare the models we apply the concept of information quality (*InfoQ*) defined as the potential of a dataset to achieve a specific (scientific or practical) goal using a given empirical analysis method (Kenett and Shmueli, 2011). In assessing *InfoQ* one first needs to describe a specific research study with four components: i) a specific analysis goal (g), ii) the available dataset (X), iii) the method or model that was used (f) and iv) a utility measure (U). As a generic concept, we define $InfoQ(f,X,g) = U\left(f\left(X \mid g\right)\right)$, i.e. the derived utility from an application of a model to a certain data set, given the research goals. In our case the data set X, the goals, g, and the utility U are assumed identical. The only difference is in the model used (f). This definition describes *what* is done by a specific analysis. In order to assess *how* it is carried out, Kenett and Shmueli (2011) propose 8 dimensions. These are: (1) Data resolution, (2) Data structure, (3) Data integration, (4) Temporal relevance, (5) Generalizability, (6) Chronology of data and goal, (7) Construct operationalization and (8) Communication. Since we apply the models to the same data set with the same objectives, in evaluating the four models, we focus only on dimensions (3), (5), (7) and (8).

In this section we compare what has been achieved with the four models described above in analyzing the ABC Company 2010 ACSS data. We build on the original work presented in the chapters in Kenett and Salini (2011) and present here only their main findings. For more details the reader is referred to the respective chapters in the book edited by Kenett and Salini (2011). Our uniform ruler for comparing the models applications will be the following questions:

A. *Data Integration*: How well is the analysis integrating various data sources and types to increase information quality derived from application of the model?

B. *Generalizability*: Is the analysis capable of generalizing the findings to a wide population frame or to other populations?

C. *Construct Operationalization*: Are the analysis outputs provided from the model useful to decision makers and can they trigger effective follow up actions?

D. *Communication*: Effective communication of analysis is obviously key to information quality. Can the model outcomes be effectively communicated?

We proceed to revisit the four models described in Section 2 and assess their outcomes with these four dimensions (A-D).

3.1 Bayesian Networks

The Bayesian Network analysis provides a visual causality map linking the various survey variables and target variables such as overall satisfaction; recommendation and repurchasing intentions. Figures 1 and 2, adapted from Kenett and Salini (2009), represent the BN of variables representing overall satisfaction from the various questionnaire topics, the country of the respondent and responses to Overall Satisfaction, Recommendation and Repurchasing Intention. Figure 1 presents the DAG with the variable names as nodes, Figure 2 shows the distribution of responses on a 1-5 scale. The BN was prepared with the GeNIe V 2.0 software (GeNIe, 2006).

Figure 1: Bayesian Network of the ABC 20110 ACSS data (with names of variables)

Figure 2: Bayesian Network of the ABC 20110 ACSS data (with variable distributions)

By studying the network one can see that an intervention to improve satisfaction levels from Technical Support or Equipment and Systems will increase *Overall Satisfaction* and eventually *Recommendation* and *Repurchasing Intentions*. This result was derived by applying various algorithms for constructing the DAG. An application of BN that provides diagnostic and predictive capabilities is presented in Kenett and Salini (2009) and Salini and Kenett (2009). As an example, consider the BN with and without conditioning on the highest recommendation level. Without conditioning, the highest level of satisfaction from Technical Support (percentage of “5”) is 26%. When conditioning the network on the response “5” to recommendation, 26% increases to 37%. The implication is that if the organization increases the percentage of customers with top level satisfaction from Technical Support from 26% to 37%, recommendation levels will reach their maximum. Management can use this analysis to justify a target of 37% for the percentage of customers rating “5” their overall satisfaction from Technical Support. We summarize now the characteristics of a Bayesian Network analysis of the ACSS data.

A. *Data Integration*: Bayesian Networks are particularly effective in integrating qualitative and quantitative data.

B. *Generalizability*: The diagnostic and predictive capabilities of Bayesian Networks provide generalizability to population subsets. The causality relationship provides further generalizability to other contexts such as organizational processes or specific job functions.

C. *Construct Operationalization*: The use of a model with conditioning capabilities provides an effective tool to set up improvement goals and diagnose pockets of dissatisfaction.
D. *Communication*: The visual display of a Bayesian Network makes it particularly appealing to decision makers who feel uneasy with mathematical models.

3.2 CUB Models
In applying the CUB model to the ABC Company 2010 ACSS data, Iannario and Piccolo (2011) observe that customers express a judgment with no uncertainty parameters with regard to *Equipment and Systems*, *Supplies and Orders* and *Contracts and Pricing* and with a limited uncertainty in the other items. They also note that satisfaction is higher for questions on *Technical Support* and *Equipment and Systems*. Thus, in this case study, customers are relatively satisfied with the equipment supplied by the ABC Company and, to a lesser extent, with *Sales Support*. Specifically the authors report the estimates presented in Table 1.

Table 1: CUB model estimates and their standard deviations

Summarizing the main characteristics of the CUB model provides the following remarks:
A. *Data Integration*: CUB models integrate the intensity of *feeling* towards a certain item with the response *uncertainty*. These two components can be also explained by using appropriate covariates.
B. *Generalizability*: The model is not generalizable per se. Its components offer however interesting cognitive and psychological interpretations.
C. *Construct Operationalization*: The model is mostly focused on explaining the outcomes of a survey. Insights on uncertainty and feelings can lead to interesting diverse initiatives.
D. *Communication*: The model estimates can be visually presented with bar plots or otherwise.

3.3 The Rasch Model
De Battisti, Nicolini and Salini, (2011) apply the Rasch model to the overall satisfaction levels of the 6 dimensions: *Equipment and Systems, Sales Support, Technical Support, Supplies and Orders, Purchasing Support, Contracts and Pricing*. Using the Partial Credit model implemented in the ***eRm*** R application (Mair and Hatzinger, 2007), the location and thresholds of individual items (questions) is computed. By sorting items by location, parameter items are ranked from best quality rating to poorest quality rating (see Table 2). The item with the highest quality rating is *Technical Support* and the item with the lowest quality rating is *Contracts and Pricing*. The Andersen Likelihood Ratio statistic, that tests the assumptions that the estimates of the difficulty parameters are equal whatever the level of the latent trait, is equal to 33 with 38 *df* and the Chi-Square p-value is 0.70. We therefore have evidence supporting the Rasch model fits the data well. Except for the reversal in *Technical*

Support, all topics show an expected monotone increasing set of thresholds marking the switch between rating values.

==

Table 2: Item location and Item thresholds of the Overall Model

==

The Rasch model provides many diagnostic tools such as item characteristic curves, goodness-of-fit plots, person-item maps, pathway maps and a wide range of statistical tests (for more details see Chapter 14 in Kenett and Salini, 2011).

Summarizing the main characteristics of the Rasch model:
A. *Data Integration*: The Rasch model integrates item and individual specific characteristics. These two components can be also explained by using appropriate covariates.
B. *Generalizability*: The model is highly generalizable, as originally conceived by Georg Rasch under the concept of specific objectivity.
C. *Construct Operationalization*: The model provides a clear distinction between individual tendencies and item specific satisfaction levels.
D. *Communication*: The model estimates can be presented visually with bar plots or otherwise. Its various diagnostic plots provide effective data analysis tools.

3.4 Control Chart Applications

A *p* chart of the percentage of respondents which rated their satisfaction level as "5" in the Equipment and Systems and the Sales Support questions is presented in Figure 3. We call these percentages 'TOP5'. The chart shows an average TOP5 proportion for Equipment and Systems questions of 14.4%. Question 9 on "uptime" is showing up with a TOP5 proportion significantly higher than the average, indicating that "uptime" is an area that stands out as an area of excellence from the customer's point of view. The Sales Support average TOP5 proportion is 18.2% with question 14, on *satisfaction from response time by sales personnel*, significantly high. Because of the small number of questions, the UCL and LCL are positioned at 2 standard deviations above and below the average central line (CL). Figure 2 is drawn with MINITAB™ version 16.0 using the questionnaire topic as a 'stage' (MINITAB, 2011).

Figure 3: *p* chart of proportion of "5" in questions on Equipment and Systems and Sales Support

Specifically, the questions displayed in Figure 3, and, in brackets, the number of response "5" out of 262 responses, are:

Equipment and Systems

q6 The equipment's features and capabilities meet your needs (32).
q7 Improvements and upgrades provide value (40).
q8 Output quality meets or exceeds expectations (30).
q9 Uptime is acceptable (49).

Sales Support

q12 Verbal promises have been honored (39).
q13 Sales personnel communicate frequently enough with you (50).
q14 Sales personnel respond promptly to requests (60).
q15 Sales personnel are knowledgeable about equipment (43).
q16 Sales personnel are knowledgeable about market opportunities (45).

Summarizing the control chart analysis of the ABC Company 2010 ACSS:
A. *Data Integration*: Control charts can be split by covariate values. Basic univariate control charts do not provide an effective data integration approach.
B. *Generalizability*: The analysis provides insights relevant to the data at hand without generalizable theory.
C. *Construct Operationalization*: The findings clearly distinguish significant from random effects, thereby helping decision makers to effectively focus their improvement efforts.
D. *Communication*: The visual display of a control chart makes it very appealing for communication and visualization of the analysis.

4. Integration of Models for Customer Survey Data Analysis

The previous section shows that the application of more than one data analysis technique to a given data set increases knowledge and information quality (*InfoQ*). In other words, combining models increases the derived utility from the application of such models to a certain data set, given the research goals. In order to integrate the techniques, specific analysis goals are described and the four models described above are considered. Our goal is to describe conceptually the relationship between *InfoQ*, the number of goals and the number of models used.

If one considers the models listed in Section 2, there are several possible outputs produced with such models that provide complementary information. For example, the *Bayesian Network* analysis provides a visual causality map linking the various survey variables and target variables. *Log-linear models* provide an important and powerful approach for examining the dependence structure among categorical random variables. *CUB models* differentiate between satisfaction levels from an item and randomness of the response choice. *The Rasch model* gives a quantitative measure of satisfaction (person parameters) and a quantitative measure of item quality (item parameters). Both are reported on the same scale with a measure of *misfit* of the item that can be used to calibrate the questionnaire, *Decision Trees* provide conceptually simple ways of understanding and summarizing the main features of the data; in particular, they exploit tree-graphs to provide visual representations of the rules underlying a given data set. *PLS models*, using latent variables, determine a dependency model between the variables and satisfaction and loyalty. *Non Linear PCA* provides a measure of satisfaction for each customer (scores), a measure of importance for the items (factor loadings) and the optimal quantifications of the ordinal categories of the variables. *Muldimensional Scaling* has, as a main objective to set up satisfaction-related-dimensions that can adequately reproduce differences among customers. *Control Charts* can be used to combine importance of an item with the mapping of items of excellence and problem area in order to identify strengths and weaknesses. In analyzing data over time, control charts can be used in a more traditional way, for identifying trends and pointing out improvement or deterioration areas. *Fuzzy Methods* allow us to obtain rankings of the subjects, from very dissatisfied to completely satisfied. For each of the different subsets of the questions of the survey, the distribution of the respondents, according to the value of their fuzzy composite indicator, can be obtained. The results for different subsets of variables can be correlated with the latent concept of customer satisfaction and fuzzy composite indicators for each categories of the overall satisfaction can be computed.

So, considering the many results achieved with different models and techniques, it is possible to identify different goals in which one could be interested. These goals should be explicitly elaborated and decided when a customer satisfaction survey is conducted. Examples of such goals include:

goal 1. Decide where to launch improvement initiatives
goal 2. Identify the drivers of overall satisfaction
goal 3. Detect positive or negative trends
goal 4. Highlight best practices by comparing products or marketing channels
goal 5. Improve the questionnaire
goal 6. Assess the reliability of the questionnaire
goal 7. Set up improvement goals
goal 8. Design a balanced scorecard using customer inputs
goal 9. Determine the meaning of the rating scales
goal 10. Effectively communicate the results using graphics or otherwise

Some of the goals can be reached using only one technique, a combination of goals usually requires applying several models. In general, the *Rasch model* can help achieve goal 1, goal 5, and goal 9. *Decision Tree, PLS, Bayesian Network* can help meet goal 2 in different ways and starting from different assumptions. In any case, even if only one technique could be enough, usually *infoQ* increases if one technique confirms the results obtained with another technique and/or shows the results in different way. An important issue that affects the analysis of survey data is the handling of missing data (Fuchs and Kenett, 2007). Harel (2009) and Harel and Schafer (2009) propose an approach extending the Rubin Multiple Imputation approach to ignorable and non-ignorable missingness which account for an underlying model.

In general, most goals and objectives of a customer satisfaction survey require the application of more than one technique. Take for example goal 1. The item parameter of the Rasch model combined with the factor loading of Non Linear PCA (Ferrari and Barbiero, 2011, Figure 17.7) show that an intervention on *Contracts and Pricing* (low quality but high importance) could increase *Overall Satisfaction*. The satisfaction level, in this example is a latent trait. If one considers the observed *Overall Satisfaction*, like in the Bayesian network application, or in the Decision Tree example (Galimberti and Soffitti, 2011, Figure 15.3), it seems that *Contracts and Pricing* is not so relevant for Overall Satisfaction. On the other hand, an intervention in the *Technical Support* areas could increase *Overall Satisfaction* and *Repurchase* and *Recommendation*. However, considering the CUB Model outputs and Rasch Model results, Table 1 and Table 2 respectively, the satisfaction/perceived quality for *Technical Support* is just reported as high. Another relevant aspect is the coherence of the single questions in each dimension and the overall satisfaction variable (g7). Looking at the Control Chart of TOP5 values for specific questions (Figure 2), it appears that *Equipment and Systems* has lower perceived quality than *Sales Support*. This is not derived from the Rasch Model and CUB Model analysis that considers only the overall satisfaction questions.

Note that the latent trait, in a specific question, is not necessarily consistent with the *Overall Satisfaction* declared for that dimension. Using partial least square models (*PLS),* Boari and Cantaluppi (2011, Figure 16.5) consider the relation between specific questions, in each dimension, and overall satisfaction. Their results are consistent with Non Linear PCA and the Rasch Model. The impact-performance matrix (Boari and Cantaluppi, 2011, Figure 16.6) suggest that an intervention on *Contracts and Pricing* is necessary.

To generalize these examples we consider a company that has several goals and is seeking to achieve, by analysis of a customer satisfaction survey, a high level of *infoQ*. To attain this, different models can be applied according to their ability to reach individual goals. This is represented conceptually in Table 3, where N_g represents the number of goals that a technique f_i is able to meet, and N_f is number of models needed to be applied in order to reach goal g_j

Table 3. Relation between *InfoQ* and N_g and N_f

In this example, goal 1 (g_1) is achieved by applying models or techniques f_1, f_2, f_3 and f_4. On the other hand, model f_4 will help us reach goals g_1, g_2 and g_3.

For achieving high *InfoQ,* we suggest to first clearly list the goals of a survey and that further considerations are given to various models for the data analysis. Summarizing these steps, as in Table 3, provides the data analyst with an integrated view of what models to apply and why.

Figure 4 gives a conceptual graphical representation of the relationship between *InfoQ,* N_g and N_f. If only one goal is present, one technique could be enough to reach a satisfactory level of *InfoQ*, even if using more than one technique can slightly increase *InfoQ*. When the number of goals (N_g) increases, but the number of techniques (N_f) does not increase, *InfoQ* will typically become very low. More goals usually require applying more data analysis techniques. The chapters in the book edited by Kenett and Salini (2011) provide examples of such techniques. An effective methodology for integrating various models requires more research and analytical developments.

Figure 4. Relation between *InfoQ* (the size of the bubble), N_g and N_f

In conclusions, the higher the number of goals from a customer satisfaction survey, the more models and more efforts are needed for generating information of adequate quality. *InfoQ* provides an approach for assessing the level of information provided by the survey data analysis. The models used in this paper are a sample of a large collection of relevant models presented in the book edited by Kenett and Salini (2011).

References

Andrich D. (2004), Controversy and the Rasch model: a characteristic of incompatible paradigms?, *Medical Care*, 42, pp. 1-16.

Baker F.B. (1987), Methodology review: item parameter estimation under the one, two and three-parameter logistic models, *Applied Psychological Measurement*, XI, 2, pp. 111-141.

Bertoli Barsotti L. (2003), An order-preserving property of the maximum likelihood estimates for the Rasch model, *Statistics and Probability Letters*, 61, pp. 91-96.

Bertoli Barsotti L. (2005), On the existence and uniqueness of JLM estimates for the Partial Credit Model, *Psychometrika*, 70, 3, pp.517-531.

Boari G. and Cantaluppi G. (2011). PLS Models, in *Modern Analysis of Customer Satisfaction Surveys*, Kenett R.S. and Salini S. Editors, John Wiley and Sons, Chichester: UK.

De Battisti F., Nicolini G. and Salini S. (2005). The Rasch Model to Measure Service Quality, *The ICFAI Journal of Services Marketing*, Vol. III, No. 3, pp. 58-80.

De Battisti F., Nicolini G. and Salini S. (2010). The Rasch Model in Customer Satisfaction Survey Data, *Quality Technology & Quantitative Management*, 7, 1, pp. 15-34.

De Battisti, F., Nicolini G. and Salini, S. (2011). The Rasch Model, in *Modern Analysis of Customer Satisfaction Surveys*, Kenett R.S. and Salini S. Editors, John Wiley and Sons, Chichester: UK.

Ferrari P. and Barbiero A. (2011). Nonlinear Principal Component Analysis, in *Modern Analysis of Customer Satisfaction Surveys*, Kenett R.S. and Salini S. Editors, John Wiley and Sons, Chichester: UK.

Fuchs C. and Kenett R.S. (2007). Missing Data and Imputation in *Encyclopedia of Statistics in Quality and Reliability*, Ruggeri F., Kenett R. S. and Faltin F. (editors in chief), Wiley. UK.

Galimberti G. and Soffitti G. (2011). Tree Based methods and Decision Trees, *Modern Analysis of Customer Satisfaction Surveys*, Kenett R.S. and Salini S. Editors, John Wiley and Sons, Chichester: UK.

GeNIe (2006). Decision Systems Laboratory, University of Pittsburgh, USA, http://genie.sis.pitt.edu.

Hand D.J., Mannila H., and Smyth P. (2001). *Principles of data mining*, MIT Press.

Harel O. (2009). *Strategies for Data Analysis with Two Types of Missing Values: From Theory to Application*, Lambert Academic Publishing.

Harel O. and Schafer J.L. (2009). Partial and latent ignorability in missing-data problems, *Biometrika*, 96, 37-50

Iannario M. and Piccolo D. (2011). CUB Models: Statistical Methods and Empirical Evidence, in *Modern Analysis of Customer Satisfaction Surveys*, Kenett R.S. and Salini S. Editors, John Wiley and Sons, Chichester: UK.

Kenett R.S. and Zacks S. (1998). *Modern Industrial Statistics: Design and Control of Quality and Reliability*, Duxbury Press, San Francisco, 1998, Spanish edition 2000, 2nd paperback edition 2002, Chinese edition, 2004.

Kenett R.S. (1991). Two Methods for Comparing Pareto Charts, *Journal of Quality Technology*, 23, pp. 27-31.

Kenett R.S. (2007). Cause and Effect Diagrams in *Encyclopaedia of Statistics in Quality and Reliability*, Ruggeri F., Kenett R. S. and Faltin F. (editors in chief), Wiley, UK.

Kenett R.S. and Salini S. (2009). New Frontiers: Bayesian networks give insight into survey-data analysis, *Quality Progress*, pp. 31-36, August.

Kenett R.S. and Salini S. (2011). *Modern Analysis of Customer Satisfaction Surveys: with applications using R*, John Wiley and Sons, Chichester: UK.

Kenett R.S. and Shmueli G., (2011). *On Information Quality*, submitted, http://ssrn.com/abstract=1464444.

Kenett R.S., Deldossi L. and Zappa D. (2011). Quality standards and control charts applied to customer surveys, in *Modern Analysis of Customer Satisfaction Surveys*, Kenett R.S. and Salini S. Editors, John Wiley and Sons, Chichester: UK.

Kenett R.S., Perruca G. and Salini S. (2011). Bayesian Networks, in *Modern Analysis of Customer Satisfaction Surveys*, Kenett R.S. and Salini S. Editors, John Wiley and Sons, Chichester: UK.

Lauritzen S. L., Spiegelhalter D. J. (1988). Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society, Series B (Methodological*), 50, 2, pp.157-224.

Mair P. and Hatzinger R. (2007). Extended Rasch Model: The eRm package for the Application of IRT models in R, *Journal of Statistical Software*, 20, 9. Pp. 1-20.

MININTAB (2011), Software for Quality Improvement, www.minitab.com.

Pearl J. (2000). *Causality: Models, Reasoning, and Inference*, Cambridge University Press, UK.

Piccolo D. (2003). On the moments of a mixture of uniform and shifted binomial random variables, *Quaderni di Statistica*, 5, pp. 85–104.

Quality Technology and Quantitative Management (2008). *Special Issue on Non-Standard Analysis of Customer Survey Data*, Corradetti, R., Furlan R. and Salini, S. (guest editors), http://www.cc.nctu.edu.tw/~qtqm

Rasch G. (1960). Probabilistic models for some intelligence and attainment tests, Copenhagen, Danish Institute for Educational Research, expanded edition (1980) with foreword and afterword by B.D. Wright. The University of Chicago Press, Chicago.

Salini S. and Kenett R.S. (2009). Bayesian Networks of Customer Satisfaction Survey Data, *Journal of Applied Statistics*, 36, 11, pp. 1177-1189.