Chapter 2
Basic Methodological Approaches

2.1 Types of Variables and Multivariate Analysis

In the context of quantitative techniques and tools (e.g. data analysis methods and statistical models), the problem of customer satisfaction evaluation presents the following basic characteristics (Wilkie and Pessemier, 1973; Churchill, 1991; Cooper, 1994):

- This particular subject is approached not only as a measurement problem, but also as a problem of understanding and analyzing customer satisfaction. In simple words, it is not enough for a business organization to know if its customers are satisfied or not, but it is necessary for the applied methods and techniques to identify the reasons behind customer satisfaction or dissatisfaction.
- In the majority of practical applications, it is commonly accepted that the data of the problem are based on the customers' judgments and should be directly collected from them. This justifies the necessity of conducting customer satisfaction surveys that results in the collection of a large volume of data.
- This is a multivariate evaluation problem given that customer's global satisfaction depends on a set of variables representing product/service characteristic dimensions. In addition, in several cases, it is necessary to examine and analyze customer behavior in relation to a set of competitive products.

The selection of the appropriate multivariate method depends mainly on the nature and the measurement scale of the variables used in the satisfaction evaluation model. Although extensive research on the measurement theory can be found for alternative levels of measurement, the variables used generally in market surveys may be classified to the following basic categories (Stevens, 1951):

- *Nominal variables*: These variables are only used in order to categorize various objects, and thus the containing information does not have any sense of ranking of preference. The only admissible mathematical operators in this category are

E. Grigoroudis and Y. Siskos, *Customer Satisfaction Evaluation*, International Series in Operations Research & Management Science 139, DOI 10.1007/978-1-4419-1640-2_2,

equality “=” and inequality “≠”. Thus, if nominal variables are quantified, this is purely for coding reasons (e.g. when developing a database); the numbers assigned to nominal variables carry no magnitude value (Vavra, 1997).

- *Ordinal variables*: These variables indicate the order of objects, according to a particular attribute. Along with the equality and inequality operators, the operators of “>” and “<” are also meaningful in this category. Thus, if numbers are assigned to ordinal variables, these numbers can only indicate order. For example, the central tendency of an ordinal variable may be represented by its median, but the mean cannot be defined. It should be emphasized that the ordinal scale permits the ordering of the objects, but it is unable to specify their distance. For this reason, the arbitrary quantification of an ordinal variable may lead to unexpected and erroneous results in subsequent analyses (Gerson, 1993; Vavra, 1997).
- *Interval variables*: The interval variables use a specific measurement unit and consequently they are able to order objects so that the differences between the values of the scale levels are equal. This means that the aforementioned differences are meaningful and can be compared (Vavra, 1997). A typical example of such scale is the temperature Celsius scale: 40ºC is warmer than 20ºC (ordering), an increase from 30ºC to 40ºC is the same with an increase form 40ºC to 50ºC (equal intervals), and the difference between 20ºC and 40ºC is twice the difference between 40ºC and 50ºC (comparing differences). Apart from the allowed operators of the former scales, addition (+) and subtraction (−) can also be used. However, interval variables have no meaningful zero point (usually it is arbitrarily assigned, like in the Celsius scale).
- *Ratio variables*: These variables are similar to interval variables, but with meaningful (non-arbitrary) zero point. Most of the measurement in the physical sciences and engineering is done on ratio scales, like mass, length, time, volume, etc. This scale takes its name form the fact that the measurement is the estimation of the ratio between a magnitude of a continuous quantity and a unit magnitude of the same kind. Since ratios between numbers on a ratio scale are meaningful, operators such as multiplication “*” and division “/” may be carried out directly. In fact, all available mathematical operators can be used for ratio scales.

The variables assessed on a nominal scale are also called categorical or discrete variables, while interval and ratio variables are also denoted as numerical or metric variables.

Examples of different measurement scales used in customer satisfaction surveys are presented in Figure 2.1. As shown, the most frequent use of nominal scales in these types of surveys is in collecting classification information (i.e. variables that may segment the total set of customers). On the other hand, ratio scales seldom apply to the subjective concepts measured in customer satisfaction surveys (Vavra, 1997). In fact the majority of information collected in these surveys uses ordinal variables.

(a) Nominal scale

Please indicate which product you have purchased today.

Product A [1]

Product B [2]

Product C [3]

(b) Ordinal scale

How satisfied are you with product ______ ?

Dissatisfied [1]

Somewhat dissatisfied [2]

Neither satisfied nor dissatisfied [3]

Somewhat satisfied [4]

Satisfied [5]

(c) Interval scale

Give in a 1-10 scale your satisfaction level with product ______ ?

[1] [2] [3] [4] [5] [6] [7] [8] [9] [10]

(d) Ratio scale

Which is your percentage of satisfaction with product ______ ?

Completely dissatisfied 0% ________________ 100% Completely satisfied

Fig. 2.1 Examples of different measurement scales

More specifically, the main variables considered in a satisfaction survey are directly or indirectly related to the customer satisfaction (e.g. satisfaction level, repurchase intention, loyalty level) or the performance of particular characteristics of the considered product or service. These variables are measured using the following alternatives (see also section 7.3.3):

- Using a quantitative scale (e.g a 1-10 interval) according to which the customer is asked to rate the performance or express his/her satisfaction from a product or from a product's particular characteristic. Attention should be paid to the wording of the question and the direction of the scale, so that the collected data are not biased by these factors (Naumann and Giel, 1995). It should also be noted that the size of the scale may create difficulties to respondents (Oliver, 1997).

- Using a verbal scale of an ordinal form (see for example Figure 2.1). However, as already noted, only simple descriptive statistics should be applied in these ordinal scales. For this reason, in many cases, an a priori arbitrary quantification is used (e.g. 1 for dissatisfied customers, 2 for somewhat dissatisfied customers, etc.). This particular quantification approach has been intensely criticized, because it makes the strong assumption that the "value" given by customer at each satisfaction level is known a priori. Moreover, the assumed linear relation of the satisfaction level "values" is not always compatible with the real market conditions, given that going from one satisfaction level to another neither yields the same "value" to customers nor is proportional to the effort that the organization should make. In addition, this quantification may lead to wrong conclusions, particularly when calculating averages. Finally, this approach does not take into account the demanding level of customers that may vary for different product/service characteristics.

The importance of partial satisfaction dimensions or product characteristics is another parameter included in satisfaction surveys, particularly when simple descriptive statistics are applied. The direct measurement of importance is usually accomplished with the following ways (Hauser, 1991):

- Customers are asked to assign a set of importance points (usually 100) to the defined satisfaction dimensions (this approach is also called constant sum method). Although it is widely used in several cases, its criticism concerns mainly the response difficulty that the customers face when dealing with a large number of satisfaction dimensions, and the fact that customers tend to assess the importance by using groups of 5 or 10 points, thus resulting in data that are not truly continuous.
- Customers are asked to rank the satisfaction dimensions according to their importance preference. This approach may present difficulties in case of a large number of satisfaction dimensions.
- Using an ordinal or an interval scale, similarly to the case of satisfaction judgments. This scale is either defined similarly to the satisfaction scale, or normalized in order to be combined with satisfaction data.

Other alternative techniques for measuring importance are presented by Diener et al. (1985) and Dolinsky (1994), while a large number of researches that mostly refer to employee satisfaction measurement identify the inconsistencies that this particular approach may lead to (Cohen et al., 1972; Bettman et al., 1975; Ryan and Bonfield, 1975; Locke, 1984; Rice et al., 1991; McFarlin and Rice, 1992; Taber and Allinger, 1995; McFarlin et al., 1995). These inconsistencies are caused by the so called "range of affect", or by the fact that the estimated low weight of some attributes does not necessarily imply that these are not considered important by the customers (see section 3.4.1). For this reason, many researchers suggest that the importance should not be based only on information given directly by customers, but it should be estimated using an analytical method (Mobley and Locke, 1970; Blood, 1971; Oliver, 1997).

Taking into account the aforementioned framework, it should be noted that that the selection of the appropriate multivariate method depends also on the objective of the analysis, besides the measurement scale of the considered variables. Vavra (1997) classifies the multivariate statistical techniques that may be used to analyze customer satisfaction by considering two major objectives: explore the relationships in different customer satisfaction data and determine the dependencies in these data (Figure 2.2).

2.2 Simple Quantitative Models

2.2.1 Descriptive Statistics

The simplest technique to analyze satisfaction survey data is to calculate the frequencies of customer responses to particular questions that are assumed critical. More specifically, depending on the applied scale, the percentages of satisfied and dissatisfied customers are calculated and used as a performance measure of the company. In many cases, the selection of satisfaction levels that characterize satisfied or dissatisfied customers depends on the strategy and the general philosophy of the business organization (e.g. some companies use the percentages of "very satisfied" and "satisfied" customers as their performance indicator, while others prefer to use only the percentage of "very satisfied" customers).

This approach does not violate the qualitative nature of the collected information, while in addition, if longitudinal data are available, they may be used in order to evaluate customer satisfaction trends. For example, Dutka (1995) proposes the following statistical approach:

1. Present the frequencies of customer satisfaction data in a time-series format, in order to identify which satisfaction dimensions have been improved and in which satisfaction dimensions additional effort should be put.
2. Apply a statistical hypothesis test in order to investigate potential changes in customer attitude compared to previous time periods.
3. Present the data in statistical quality control charts with predefined control limits.

In case where metric variables are used in the customer satisfaction survey, it is possible to estimate an overall satisfaction index, based on the customer judgments for the performance and importance of the product/service characteristics. The customer satisfaction index *CSI* is calculated using a weighted sum formula (Hill, 1996):

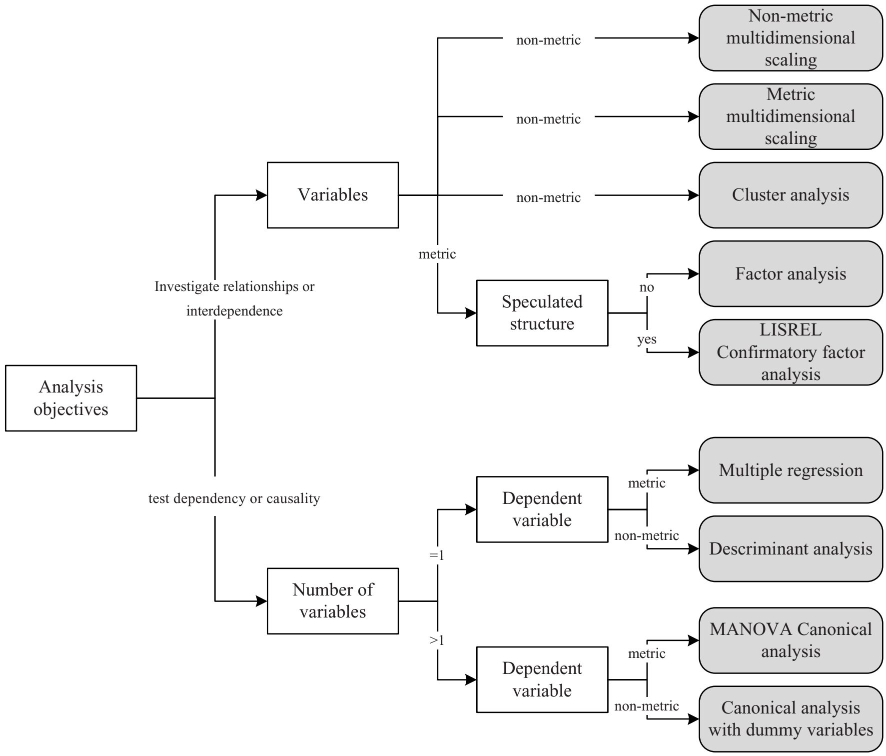


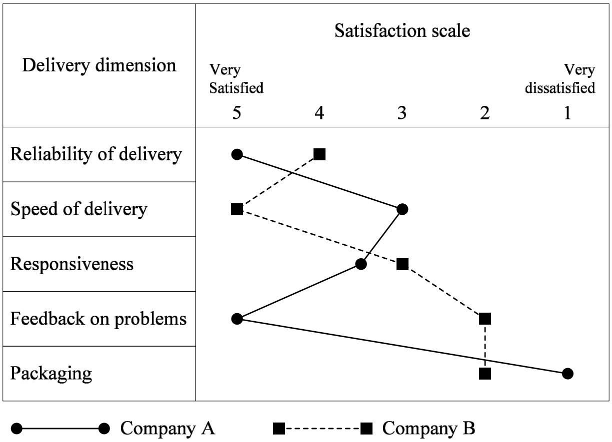
Fig. 2.2 A map of mutlivariate techniques (Vavra, 1997)

$$\begin{cases} CSI = \sum_{i=1}^{n} \overline{b}_i \overline{X}_i \\ \text{with } \overline{b}_i = \frac{1}{M}\sum_{j=1}^{M} b_{ij} \text{ and } \overline{X}_i = \frac{1}{M}\sum_{j=1}^{M} x_{ij} \end{cases} \tag{2.1}$$

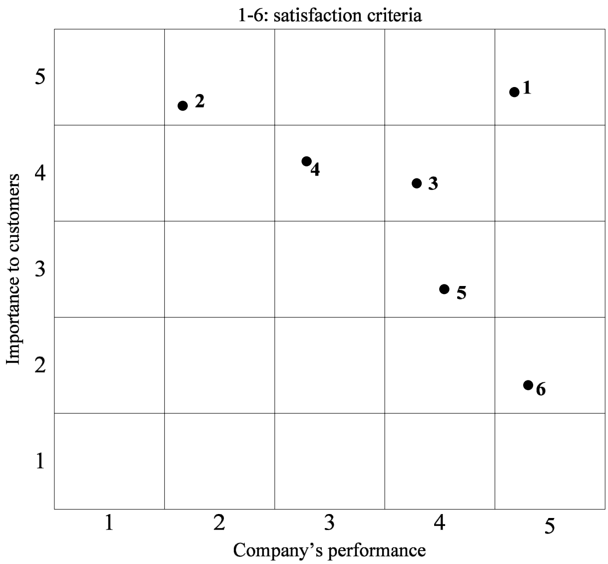
where $\overline{X}_i$ and $\overline{b}_i$ are the average scores of the satisfaction/performance and the importance of the characteristic i, respectively, x_{ij} and b_{ij} are the satisfaction/performance and the importance judgment of customer j for the characteristic i, respectively, n is the number of product/service characteristics, and M is the size of customer sample. In case that different measurement scales are used for the x and b variables, a normalization coefficient should be used in the *CSI* formula.

In several studies, the previous approach is also applied even though only ordinal satisfaction data are available. However, this requires the quantification of the ordinal scale which presents, as already mentioned, a series of problems (see section 2.1). Additional difficulties in analyzing and interpreting these types of results are also mentioned in Oliver (1997) and Vavra (1997).

Other techniques that focus on the reporting and the presentation of results are given by Dutka (1995) and Hill (1996). The most important of them refers to the performance profiles and performance matrices, examples of which are presented in Figure 2.3 (see also section 4.3.5 for a detailed presentation and discussion of performance matrices).



(a) Performance Profile



(b) Performance Matrix

Fig. 2.3 Examples of performance profiles and matrices (Hill, 1996)

Descriptive statistics methods are not able to provide an in depth analysis of customer satisfaction. Nevertheless, they can be used either during the preliminary analysis or complementary to other quantitative models.

2.2.2 Basic Statistical Approaches

Multiple regression analysis is one of the most widely used statistical methods for analyzing customer satisfaction data. The method is used to study the relation between the satisfaction/performance of the total set of product's or service's characteristics (independent variables) and the overall customer satisfaction judgment (dependent variable).

The general form of multiple regression equation is as follows:

$$Y = b_0 + b_1 X_1 + b_2 X_2 + \ldots + b_n X_n \tag{2.2}$$

where Y is the overall customer satisfaction judgment, X_i is the customer satisfaction/performance of characteristic i, b_i are the estimated regression coefficients and n is the number of product or service characteristics.

In order to apply multiple regression analysis in customer satisfaction data, the following issues should be emphasized (Grisaffe, 1993; McLauchlan, 1993; Mullet, 1994):

- All the variables in the linear model should be metric, otherwise multiple regression analysis should not be performed. Particularly in the case of ordinal variables, the arbitrary codification of the scales may lead to significant inconsistencies. In addition, if model variables are measured in different scales, a normalization procedure is necessary.
- Beside the overall customer satisfaction with a product/service, the Y variable, may also represent other related aggregated measures, such as customer loyalty level or repurchase intention level.
- The coefficients b_i indicate the contribution of the independent variables to the dependent variable Y. Thus, these coefficients may reveal the importance given by customers to each one of the product's or service's characteristics, and therefore to identify the critical satisfaction dimensions.

The major problems and the criticism of this particular approach focus on the quantification of the satisfaction data and the multicollinearity among the independent variables X_i. In addition, even when a metric scale is used, it is assumed that the model variables are continuous, which is not compatible with the type of the collected information. Moreover, the dependency among the variables X_i may affect the reliability of the results and it is possible to lead to inconsistencies. However, several approaches have been proposed in order to overcome the aforementioned problems (see for example Flury and Riedwyl, 1988).

Detailed presentation of the method is given by Draper and Smith (1967), Daniel and Wood (1980) and Flury and Riedwyl (1988), while applications of multiple regression analysis to market survey data are presented by Kerlinger and Pedhazur (1973), Cohen and Cohen (1983), Dutka (1995) and Vavra (1997).

Another statistical method widely used in analyzing customer satisfaction data is factor analysis. The aim of the method is to study the relation pattern among the product's or service's characteristics.

The main form of the factor analysis equation relates the set of variables with a minimum number of factors as follows (Harman, 1976):

$$X_i = a_{i1}F_1 + a_{i2}F_2 + \ldots + a_{im}F_m \quad \text{with } i = 1, 2, \ldots, n \tag{2.3}$$

where X_i is the customer satisfaction/performance of characteristic i, F_j is factor j, a_{ij} are the estimated coefficients, m is the number of factors, and n is the number of product/service characteristics.

Beside the estimation of a_{ij} coefficients, which are able to investigate the nature and number of underlying dimensions in the survey data, factor analysis also generates data (scores) for every customer on each of the factors uncovered. These derived values for each case are called factor scores and may approximate how customers might have rated the product/service, if they were asked to give their judgments only for the discovered factors (instead of the raw variables that they originally answered). These factor scores may be also used to cluster customers (Vavra, 1997).

In general, factor analysis is used to decompose a data matrix into its bare structural essentials that can efficiently describe the original customer satisfaction data. The reduction of a large number of attributes is the most common application of factor analysis to a customer satisfaction measurement program. Usually, the application process includes the following steps (Dutka, 1995):

1. Create an exhaustive list of product/service characteristics that affect the customer satisfaction, using qualitative survey techniques, like personal interviews or customer focus groups (see section 7.1).
2. Conduct a preliminary customer satisfaction survey using a pilot questionnaire that includes the list of these characteristics.
3. Reduce the number of characteristics into the major evaluative dimensions of customers using factor analysis.
4. Implement the customer satisfaction measurement program using the defined satisfaction dimensions.

The criticism and the problems related to the application of factor analysis to market survey data do not differ from those of multiple regression analysis. In addition, Dutka (1995) notes that during the application of the method, particular attention should be paid to critical issues related to the interpretation of the results (e.g. selecting the appropriate technique to rotate the factor solution).

The mathematical development of the method is presented analytically in many textbooks on multivariate data analysis (see for example Rummel, 1970; Cooley and Lohnes, 1971; Urban and Hauser, 1980; Gorsuch, 1983), while a large number of publications refers to the application of factor analysis in market survey data (Roberts et al., 1971; Hayes, 1992; Naumann and Giel, 1995; Hill, 1996; Vavra, 1997).

2.3 Advanced Quantitative Techniques

2.3.1 Conditional Probability Models

An important category of quantitative tools that may be used in the customer satisfaction measurement problem refers to the conditional probability models. These models follow a regression-type approach, taking into account that the measurement variable has an ordinal form.

The conditional probability models, given customer evaluations for a set of product/service characteristics, estimate a satisfaction probability distribution function, i.e. the probability that a customer belongs to a particular "satisfaction group" (e.g. group of satisfied customers, group of dissatisfied customers, etc.). The main forms of these models include the linear probability model and the logit and probit models.

The linear probability model is a binary regression approach, assuming that customer's overall satisfaction (dependent variable) is a dichotomous variable taking two possible values (i.e. satisfaction or dissatisfaction). The model may be expressed by the following formula:

$$\Pr(Y = 1 | \mathbf{X}) = b_0 + b_1 X_1 + \ldots + b_n X_n \tag{2.4}$$

where Y is the dichotomous variable representing overall customer satisfaction, b_i are the regression coefficients, X_i are the customer satisfaction/performance of characteristic i, and n is the number of product/service characteristics.

It should be noted here that b_i are OLS (ordinary least square) estimates, and thus the linear probability model is used when alternative techniques based on maximum likelihood estimates are computationally difficult. Moreover, in case that Y is a multiple response variable, the model can be extended with the use of dummy variables.

There are several potential statistical problems in the application of linear probability models, although alternative techniques have been proposed in order to overcome these problems. For example, the error terms are heteroskedastic and their distribution is not normal, while without restrictions on b_i, the estimated coefficients can imply probabilities outside the unit interval [0, 1].

The logit analysis is a similar approach where the previous satisfaction probability is given by the logistic function:

$$\begin{cases} \Pr(Y = 1 | \mathbf{X}) = \dfrac{1}{1 + e^{-z}} \\ \text{with} \quad z = b_0 + b_1 X_1 + \ldots + b_n X_n \end{cases} \tag{2.5}$$

The logit analysis has numerous applications in marketing and other fields (e.g. artificial neural networks, biology, medicine, economics, mathematical psychology). The method, based on a cumulative distribution function, provides the probability of a customer to belong to one of the prescribed satisfaction classes, given his/her satisfaction/performance judgments on a set of product/service characteristics.

For the logit of the previous probability, which is the inverse of the logistic function, it can be shown that:

$$\begin{aligned} & p = \Pr(Y = 1 | \mathbf{X}) = \frac{1}{1 + e^{-z}} \Rightarrow \frac{p}{1 - p} = e^{z} \Rightarrow \\ & \Rightarrow \operatorname{logit}(p) = \ln\left(\frac{p}{1 - p}\right) = z = b_0 + b_1 X_1 + \ldots + b_n X_n \end{aligned} \tag{2.6}$$

Probit models are similar to logit analysis. The main difference is that the probability $\Pr(Y = 1|\mathbf{X})$ is given by the cumulative standard normal distribution function:

$$\begin{cases} \Pr(Y = 1 | \mathbf{X}) = \dfrac{1}{2\sqrt{\pi}} \displaystyle\int_{-\infty}^{z} e^{-\frac{1}{2}u^2} du \\ \text{with} \quad z = b_0 + b_1 X_1 + \ldots + b_n X_n \end{cases} \tag{2.7}$$

Usually, logit analysis is used as an alternative to probit analysis mainly because of the simplicity of the logistic function and the relatively lower required computational effort (Pindyck and Rubinfeld, 1985). However, these models are very similar since they assume that the probability of a customer j to be satisfied by the offered product or service is described by the relationship:

$$\Pr(\mathbf{X}_j, \mathbf{b}) = F(b_0 + \mathbf{b}^T \mathbf{X}_j) \tag{2.8}$$

where $\mathbf{X}_j$ is the satisfaction vector of customer j for the total set of product/service characteristics and $\mathbf{b}$ is the vector of estimated model parameters. The main difference is that, in order to assess $\Pr(\mathbf{X}_j, \mathbf{b})$, logit analysis uses a cumulative logistic function, while probit analysis a cumulative normal distribution function, as shown in Figure 2.4.

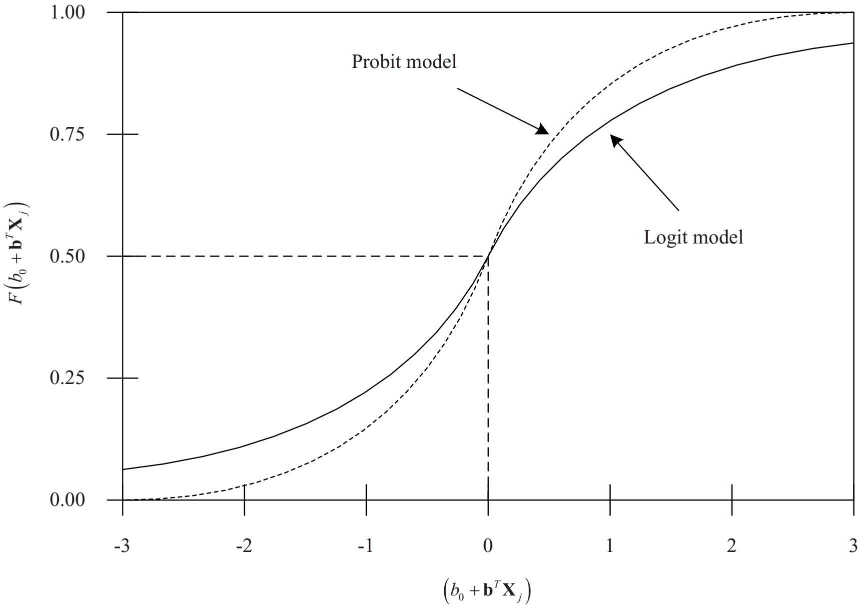


Fig. 2.4 Comparison of probit and logit analyses

Logit and Probit analyses present also similarities with the classification statistical models (e.g. discriminant analysis). However, it should be noted that their purpose is not to classify customers in prescribed satisfaction classes, but to assess the probability that a customer belongs to one of these classes.

Detailed presentation of the binary logit and probit analysis is given by Gnanadesikan (1977), Hanushek and Jackson (1977), Fienberg (1980), Andersen (1990), and Agresti (1996), while the case of multiple response models is presented in Theil (1969), McCullagh (1980), Fienberg (1980) and Agresti (1984, 1990).

The ordered conditional probability models may be considered as an extension of the previous models, taking into account that the dependent variable is ordinal. Moreover, in case of multiple responses, customer satisfaction may be modeled as follows (Agresti, 1984, 1990, 1996):

$$y_j = \begin{cases} 0 & \text{if } \ y_j^* \le \mu_0 \\ 1 & \text{if } \ \mu_0 < y_j^* \le \mu_1 \\ 2 & \text{if } \ \mu_1 < y_j^* \le \mu_2 \\ \vdots & \\ a-1 & \text{if } \ y_j^* > \mu_{a-2} \end{cases} \tag{2.9}$$

where y_j is the overall satisfaction of customer j, a is the number of satisfaction levels (ordinal scale) and μ_m are the estimated model parameters, having a role of thresholds for the dummy variable y_j^*, which is denoted by the following formula:

$$y_j^* = \sum_{i=1}^{n} b_i x_{ij} + \varepsilon_j \tag{2.10}$$

where x_{ij} is the satisfaction/performance judgment of customer j for product/service characteristic i, b_i are the estimated model coefficients, ε_j are the error terms, and n is the number of product/service characteristics.

It should be emphasized that the values $\{0, 1, \ldots, a{-}1\}$, which the overall satisfaction variable can take, are simply a coding and do not quantify the y_j variable. In addition, the arbitrary quantification of y_j is avoided by using the dummy variable y_j^* and estimating the parameters μ_m. Usually, the thresholds μ_m are normalized by setting $\mu_m = 0$ in order to minimize the model parameters that should be estimated. Moreover, it is assumed that the error terms follow a prescribed probability distribution function (e.g. standard normal distribution, standard logistic distribution). Finally, the aforementioned modeling assumes that all possible values of the overall satisfaction y_j are present in the dataset.

Using equations (2.9) and (2.10), the probability that customer j has expressed for the m-th satisfaction level, given his/her satisfaction/performance judgments $\mathbf{X}_j = (x_{1j}, x_{2j}, \ldots, x_{nj})$ is

$$\begin{aligned} &\Pr(y_j = m) = \Pr(\mu_{m-1} < y_j^* \le \mu_m) = \\ &= \Pr\left(\varepsilon \le \mu_m - \sum_{i=1}^{n} b_i x_{ij}\right) - \Pr\left(\varepsilon \le \mu_{m-1} - \sum_{i=1}^{n} b_i x_{ij})\right) = \\ &= F\left(\mu_m - \sum_{i=1}^{n} b_i x_{ij}\right) - F\left(\mu_{m-1} - \sum_{i=1}^{n} b_i x_{ij})\right) \end{aligned} \tag{2.11}$$

or alternatively

$$\Pr(y_j \le m) = F\left(\mu_m - \sum_{i=1}^{n} b_i x_{ij}\right) \tag{2.12}$$

where F is the standard normal distribution function for the ordered probit model and the standard logistic distribution function for the ordered logit model.

The estimation of the parameters b_i and μ_m is based on the maximization of the log-likelihood function L:

$$L = \sum_{k=0}^{a-1} \log F\left(\mu_k - \sum_{i=1}^{n} b_i x_{ij}\right) \tag{2.13}$$

An analytical presentation of the ordered conditional probability models may be found in Gensch and Recker (1979), Fienberg (1980), Wickens (1989), Andersen (1990), and Agresti (1984, 1990, 1996).

The conditional probability models have been mainly applied in the marketing field (market surveys, discrete choice models), although a growing number of real-world applications in customer satisfaction surveys may be found in the literature.

Finally, it should be mentioned that logit and probit analysis may be considered as a special case of loglinear models that constitute an interesting alternative approach to the analysis of multidimensional contingency tables (Knoke and Burke, 1980; Wickens, 1989).

2.3.2 Structural Equation Modeling

Structural equation modeling (SEM) is a statistical technique for measuring relationships among latent variables. It has been around since early in the 20th century originating in the geneticist Sewall Wright's 1916 work (Bollen, 1989). SEM is a technique to specify, estimate, and evaluate models of linear relationships among a set of observed variables in terms of a generally smaller number of unobserved variables (Shah and Goldstein, 2006).

SEM, as a part of the general category of causal modeling, is focused on testing the hypothesis that the relationships among data are consistent with the assumed causal structure. These causal relationships are usually considered linear. Thus, SEM may be considered as an extension of regression models. In fact, SEM is a family of models that also include the following approaches (Raykov and Marcoulides, 2000):

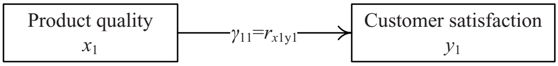
- *Path analysis*: Path analysis examines patterns of directional and non-directional relationships only among observed variables. Thus, it allows for the testing of structural relationships among observed variables, when these observed variables are of primary interest or when multiple indicators for latent variables are not available (Shah and Goldstein, 2006).

- *Confirmatory factor analysis*: Confirmatory factor analysis (CFA) models are commonly used to examine patterns of interrelationships among several constructs. CFA assumes that the observed variables are loaded on specific latent variables, which are allowed to correlate. Thus, contrary to explanatory factor analysis, CFA requires that the latent variables and their associated observed variables to be specified before analyzing data (Shah and Goldstein, 2006).

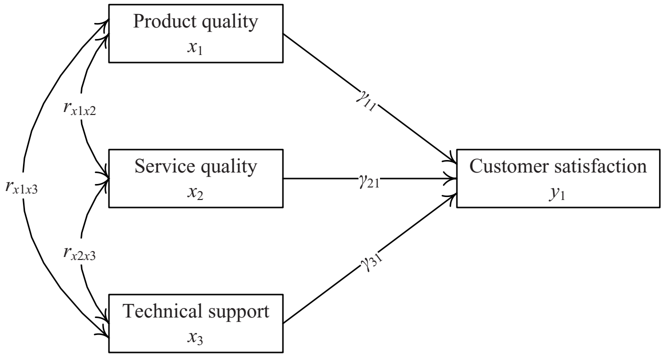
Different examples of causal modeling are presented in Figure 2.5. The simplest causal model may have only two variables: one predictor variable and one outcome variable, as shown in Figure 2.5(a). In this case, the path coefficient is equivalent to the simple correlation coefficient between these two variables. A multiple predictor causal model is depicted in Figure 2.5(b), where three predictor variables are examined, each of which has some level of covariance with the others. In this case, the path coefficients will be equivalent to correlation coefficients, only if the predictor variables are orthogonal. A more complex case is presented in Figure 2.5(c), which refers to a simple path analysis model in which five predictor variables (x_i) affect two outcome variables (y_j). In this diagram, γ_{ij} denote the path coefficients among predictor and outcome variables, while β_{kj} represent the path coefficient among outcome variables. It should be mentioned that this modeling allows predictor variables to affect both of the outcome variables, and thus both direct and indirect effects are possible (e.g. service quality influences loyalty indirectly through customer satisfaction, while loyalty is directly affected by product quality). In addition, error terms ε_j are introduced in the outcome variables and the path analysis model includes the covariance terms (CV) among all possible pairs of predictor variables.

The previous examples illustrate how causal modeling may be generalized in a path analysis context. However, the following remarks should be emphasized for the implementation of path analysis models (Allen and Rao, 2000):

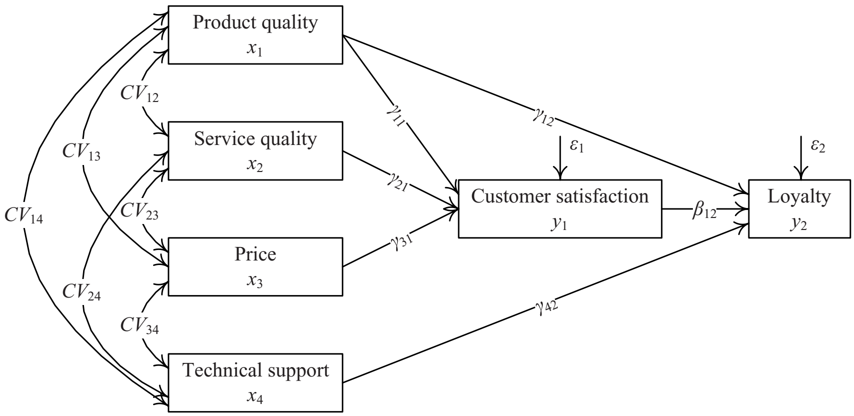
- The variance-covariance matrix is the main input data for this method. In addition, outcome variables are assumed normally distributed and measured in an interval or a ratio scale.
- Path analysis assumes that the relations between variables are linear and additive, while a sufficient number of cases are required to produce stable and robust results.
- Covariance terms among predictor variables should not be omitted, unless there are particular experiential, empirical, or theoretical reasons to do so.
- Using the hypothesized structure or the analytical model equations, it is possible to estimate direct and indirect effects of predictor variables to outcome variables.
- A saturated model (i.e. a model containing paths from each of the predictor variables to all of the dependent variables) will always fit the original data perfectly.



(a) Simple causal relationship



(b) Multiple predictor causal model



(c) Path analysis model with two dependent variables

Fig. 2.5 Examples of causal modeling (Allen and Rao, 2000)

- There are several statistical fitting indicators, but the most well-known is chi-squared, which indicates lack of fit. Moreover, in case of accepting the structural model, the errors should not be correlated.
- Since a linear equation can be written for every outcome variable, path analysis estimates a separate R^2 statistic for each of these equations (R^2 reflects the proportion of dependent variable variance accounted for by the predictor variables).

SEM refers to a general category of path analysis models having measured and latent variables, and thus it may be defined as a hypothesis of a specific pattern of relations among the aforementioned variables. In general, this category of models may be considered as a combination of path analysis and factor analysis.

The use of latent variables is the main difference from path analysis, and thus, SEM models are decomposed into their two main components:

1. The *measurement model*, which explicates the relations between measured and latent variables, and is defined as follows:

$$\begin{cases} \mathbf{y} = \mathbf{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\varepsilon} \\ \mathbf{x} = \mathbf{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta} \end{cases} \tag{2.14}$$

2. The *structural model*, which specifies relationships between latent variables through a structural equation model, and is given by:

$$\mathbf{n} = \mathbf{B}\boldsymbol{\eta} + \mathbf{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta} \tag{2.15}$$

Table 2.1 gives the necessary notation for the measurement and the structural model. The measured variables are also called manifest or observed variables, while the terms endogenous and exogenous are model specific (a latent variable is endogenous, if it is determined by variables within the model, while it is exogenous, if its causes lie outside the model).

The main assumptions of the SEM models, using the aforementioned notation, may be summarized in the following (Bollen, 1989):

1. $E(\boldsymbol{\eta}) = E(\boldsymbol{\xi}) = E(\boldsymbol{\zeta}) = E(\boldsymbol{\varepsilon}) = E(\boldsymbol{\delta}) = 0$
2. $\boldsymbol{\varepsilon}$ uncorrelated with $\boldsymbol{\eta}$, $\boldsymbol{\xi}$, and $\boldsymbol{\delta}$
3. $\boldsymbol{\delta}$ uncorrelated with $\boldsymbol{\xi}$, $\boldsymbol{\eta}$, and $\boldsymbol{\varepsilon}$
4. $\boldsymbol{\zeta}$ uncorrelated with $\boldsymbol{\xi}$
5. $(\mathbf{I}-\mathbf{B})$ nonsingular

It can be shown that the covariance matrix for the observed variables derived from raw data is a function of eight parameter matrices: $\mathbf{\Lambda}_x$, $\mathbf{\Lambda}_y$, $\mathbf{\Gamma}$, $\mathbf{B}$, $\mathbf{\Phi}$, $\mathbf{\Psi}$, $\mathbf{\Theta}_\delta$, and $\mathbf{\Theta}_\varepsilon$. Thus, given a hypothesized model in terms of fixed and free parameters of the eight-parameter matrices, and given a sample covariance matrix for the measured variables, one can solve for estimates of the free parameters of the model. The most common approach for fitting the model to data is to obtain the maximum likelihood estimates of the parameters, and an accompanying likelihood ratio chi-square test of the null hypothesis that the model holds in the population (Shah and Goldstein, 2006).

An example of a SEM model in a case of customer satisfaction measurement is presented in Figure 2.6. The model considers three latent endogenous variables (product quality, service quality, and technical support) and two latent endogenous variables (customer satisfaction and loyalty). As shown, the endogenous variables may affect loyalty directly (e.g. product quality), or indirectly through customer satisfaction (e.g. service quality). Moreover, a number of different measured variables are used in order to define all these latent variables.

Table 2.1 Notation for SEM

Type	Symbol	Dimension	Description
Variables	$\mathbf{x}$	$q \times 1$	Observed indicators of $\boldsymbol{\xi}$
	$\mathbf{y}$	$p \times 1$	Observed indicators of $\boldsymbol{\zeta}$
	$\boldsymbol{\delta}$	$q \times 1$	Measurement errors for $\mathbf{x}$
	$\boldsymbol{\varepsilon}$	$p \times 1$	Measurement errors for $\mathbf{y}$
	$\boldsymbol{\eta}$	$m \times 1$	Latent endogenous variables
	$\boldsymbol{\xi}$	$n \times 1$	Latent exogenous variables
	$\boldsymbol{\zeta}$	$m \times 1$	Latent errors in equations
Coefficients	$\boldsymbol{\Lambda}_x$	$q \times n$	Coefficient relating $\mathbf{x}$ to $\boldsymbol{\xi}$
	$\boldsymbol{\Lambda}_y$	$p \times m$	Coefficient relating $\mathbf{y}$ to $\boldsymbol{\eta}$
	$\mathbf{B}$	$m \times m$	Coefficient matrix for latent endogenous variables
	$\boldsymbol{\Gamma}$	$m \times n$	Coefficient matrix for latent exogenous variables
Covariance matrices	$\boldsymbol{\Theta}_\delta$	$q \times q$	$E(\boldsymbol{\delta\delta}')$ covariance matrix of $\boldsymbol{\delta}$
	$\boldsymbol{\Theta}_\varepsilon$	$p \times p$	$E(\boldsymbol{\varepsilon\varepsilon}')$ covariance matrix of $\boldsymbol{\varepsilon}$
	$\boldsymbol{\Phi}$	$n \times n$	$E(\boldsymbol{\xi\xi}')$ covariance matrix of $\boldsymbol{\xi}$
	$\boldsymbol{\Psi}$	$m \times m$	$E(\boldsymbol{\zeta\zeta}')$ covariance matrix of $\boldsymbol{\zeta}$

SEM has been implemented in a large number of software packages, such as LISREL (Jöreskog and Sörbom, 1993, 1996), AMOS (Arbuckle, 1997; Blunch, 2008) and EQS (Bentler, 1995).

It should be emphasized that SEM is a confirmatory rather than an exploratory approach, since its main objective is to determine whether the a priori model is valid, and not to find a suitable model (Shah and Goldstein, 2006). Thus, it is suited to theory testing rather than theory development, although, in several cases, it is used in order to explore alternative structural models.

Generally, the implementation of a SEM analysis should be based on the following main steps (Kline, 1998).

1. Specify the model (the hypotheses in the form of a structural equation model).
2. Determine whether the model is identified.
3. Select measures of the variables represented in the model and collect data.
4. Analyze the model (estimate the model parameters).
5. Evaluate model fit (determine how adequately the model accounts for the data).
6. Re-specify the model and evaluate the fit of the revised model to the same data.

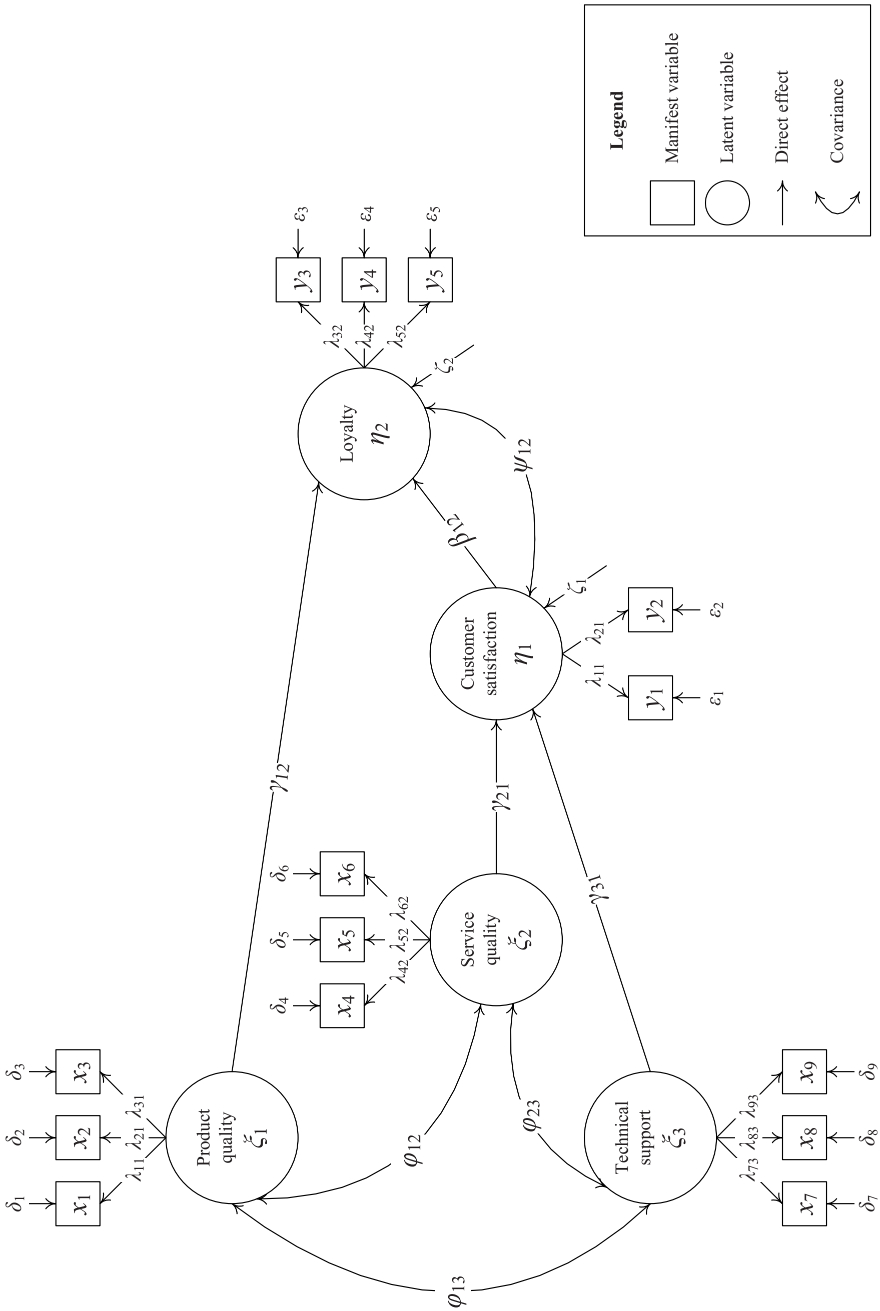


Fig. 2.6 Example of a structural equation model (Allen and Rao, 2000)

The most important strength of SEM is the ability to study latent variables. Since these variables are not directly measured, but estimated in the model from a set of measured variables, SEM models may be used to evaluate complex customer behavioral variables. For example, customer loyalty may not be measured directly, but instead, its measurement may be based on its outcomes (e.g. repurchase intention, complaints, and price elasticity). Another important advantage of SEM models, which justifies their popularity in many scientific fields of study, is that they provide a mechanism for explicitly taking into account measurement error in the observed variables (both dependent and independent) considered in the model (Raykov and Markoulides, 2000). Additionally, SEM models are able to study both direct and indirect effects of various variables included in the model. Direct effects are the effects that go directly from one variable to another, while indirect effects are the effects between two variables that are mediated by one or more intervening variable (Raykov and Markoulides, 2000).

2.3.3 Other Statistical and Data Analysis Models

Satisfaction dimensions related to a customer segment may diversify compared to another segment. Thus, several quantitative methods and techniques aim to identify product or service attributes that best discriminate customer segments, which are assessed according to the expressed satisfaction level (i.e. satisfied vs. dissatisfied customers), or a particular customer characteristic (e.g. frequency of use).

Discriminant analysis is one among the most widely used multivariate methods that, given a customer's satisfaction judgments on set of the product/service characteristics, estimates whether this customer belongs to one of the prescribed satisfaction classes.

Discriminant analysis estimates a z-score for each customer i, based on the following formula:

$$z_j = a_1 x_{1j} + a_2 x_{2j} + \ldots + a_n x_{nj} \tag{2.16}$$

where x_{ij} is the satisfaction judgment of customer j for product/service characteristic i, a_i are the estimated model coefficients, and n is the number of product/service characteristics.

The classification of customers is achieved using these z_j values and the calculation of appropriate cutoff scores. Detailed presentation of the method is given by Cooley and Lohnes (1971) and Klecka (1980).

For applying discriminant analysis in customer satisfaction surveys, the following should be taken into account (Vavra, 1997):

- The assessment of the classification groups constitutes one of the most difficult and important decisions when applying this particular method, given that it refers to the selection of the classification variable (e.g. overall satisfaction, repurchase intention), as well as the determination of the variable levels that discriminate the particular customer classes.

- The potential problems referring to the application of the method do not differ from these that were mentioned in the case of multiple regression analysis due to the relative similarity of the two methods (e.g. the parameters a_i may be interpreted in the same way with the regression coefficients).
- Usually, the set of customers is divided in two subsets, the first of which is used for the estimation of the model parameters (training set) and the second for testing the reliability the results (test set).
- Stepwise discriminant analysis is a different version of this particular method, which may be used when the set of satisfaction dimensions that classifies customers is not known and defined.

Characteristic examples of discriminant analysis applications to customer satisfaction problems are presented by Dutka (1995) and Vavra (1997).

Another important objective of satisfaction data analyses is the identification of priorities and the development of improvement strategies for the business organization. In this context, conjoint analysis is used to assess the effects of the trade-offs made by customers, when they purchase or express satisfaction evaluations for a particular product or service. According to this method, customers evaluate a series of product or service profiles having different performance levels on a set of defined attributes. This trade-off analysis is able to reveal the relative importance of these component attributes.

Conjoint analysis may be considered as a reasonable extension of customer satisfaction surveys, given that the most important trade-off decisions made by customers include the critical performance dimensions of a product or service that have been identified during the satisfaction survey process. The implementation of conjoint analysis includes the following main steps (Dutka, 1995):

1. Identification of the trade-off choices among the critical performance attributes.
2. Development of an experimental design to measure trade-offs.
3. Conduction of consumer surveys to implement the experimental design.
4. Computation of utility functions that measure the importance of the various trade-offs.
5. Analysis of the impact of changes in the product or service.

A large number of publications refer to the presentation of this particular approach (Green and Rao, 1971; Green and Wind, 1973; Johnson, 1974; Green and Sprinivasan, 1978; Green et al., 1983; Green, 1984), while a detailed review of alternative versions of conjoint analysis is given by Louviere (1988). The applications of the method not only refer to cases of customer satisfaction surveys, but also to general market surveys (Gattin and Wittink, 1982; Joseph et al., 1989; Anderson and Bettencourt, 1993).

Another important data analysis technique refers to correspondence analysis, which is one of the most popular mathematical tools for developing perceptual maps in the marketing field. Customer satisfaction research is an ideal application for perceptual mapping, since the relationship among questionnaire variables (e.g. satisfaction or performance judgments for particular product/service attributes, demographics, competitors' performance) may be investigated (Dutka, 1995).

The most important characteristics of the method, in relation to other statistical models are (Dutka, 1995):

- Correspondence analysis is mainly a descriptive technique providing qualitative information of an explanatory nature, in contrast to discriminant and regression analysis, which are quantitative methods allowing the evaluation of overall customer satisfaction on the basis of a specific mathematical formula.
- The method uses cross-tabulations as input data, thus it can analyze simultaneously row and column variables of this table (e.g. performance attributes in relation to customer demographic characteristics) However, a significant portion of the information from the raw satisfaction survey data is lost.
- Physical interpretations of the axes presented in the perceptual maps are not necessary, in contrast to factor analysis where this particular task is rather difficult. This may be justified by the fact that correspondence analysis relies on point-to-point distances rather than distances from axes.

The detailed development of the method is presented by Hoffman and Franke (1986) and Weller and Romney (1990), while conclusively, it should be noted that conjoint analysis is not able to evaluate and analyze customer satisfaction, but it is usually applied either during the preliminary stage of the data analysis process, or complementary to other methods and techniques.

Other statistical models and quantitative tools, applied for analyzing customer satisfaction, include (Wilk and Gnanadesikan, 1968; Aldenderfer and Blashfield, 1984; Denby et al., 1990; Douglas, 1995; Vavra, 1997; Löthgren and Tambour, 1999, Allen and Rao, 2000):

- Data Envelopment Analysis (DEA)
- Multidimensional scaling
- Confirmatory factor analysis
- Kruskal's relative importance approach
- Cluster analysis
- Canonical correlation analysis
- Dominance analysis
- Probability plotting methods

Finally, recent research efforts in the problem of measuring and analyzing customer satisfaction include approaches from the field of dominance-based rough sets, support vector machines, fuzzy logic, and neural networks.

2.4 Consumer Behavioral Models

2.4.1 Consumer Psychology and Satisfaction

Customer ratings on a set of product/service attributes do not explain why a particular attribute is considered important (or unimportant) and why its performance level is considered excellent (or poor). Thus, this performance approach is not able to reveal the psychological intricacies that customer brings to the firm's product or service. This important shortcoming of customer satisfaction performance analysis is emphasized by several researchers who argue that levels of performance exist only as external stimuli to consumers (Oliver, 1977).

The approach of psychology and consumer behavioral analysis is based on the assumption that satisfaction is a mental condition of the customer. The performance evaluation of a provided product or service (or some of their characteristics) is quite subjective and for this reason it should be linked with some comparison standards.

A generic model of consumer behavioral analysis considers the working on a customer's mind as a "black box", implying that consumer's psychology mediated the impact of performance observations on satisfaction judgments (Figure 2.7). Alternative behavioral models try to describe and explain what exactly happens in this "black box" in order to unravel the processing of future performance (Oliver, 1977, 1997).

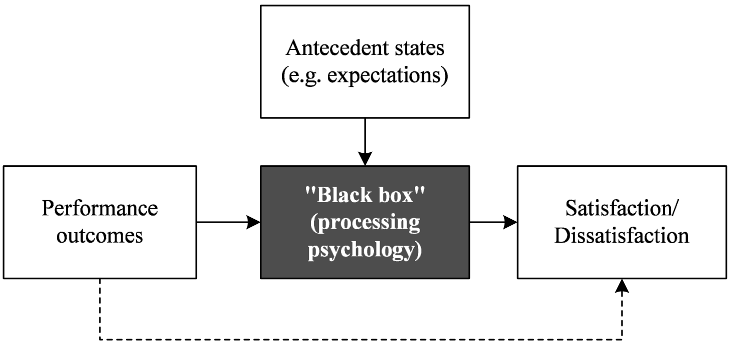


Fig. 2.7 The mediated performance model of satisfaction (Oliver, 1997)

The nature of comparison standards used in this customer satisfaction judgment process received increasing attention during the last years. A typical definition of satisfaction is focused on customer expectations as the main comparison standard (see section 2.1). However, as Woodruff and Gardial (1996) note, there are several comparison standards used by customers, which may vary across stages in a consumption process (e.g. pre-purchase, purchase, use, and disposal). These different comparison standards may lead to completely different satisfaction judgments, and they include the following (Woodruff and Gardial, 1996):

- *Expectations*: they represent how the customer believes the product/service will perform.
- *Ideals*: they represent how the customer wishes the product/service would perform.
- *Competitors*: the performance of competitors in the same product/service category may be adopted by customers as a standard for comparison.
- *Other product categories*: products or services in completely different categories may also provide comparison standards for customers.
- *Marketer promises*: they refer to promises that were made by the salesperson, the product/service advertisement, the company spokesperson, or some other form of corporate communication.
- *Industry norms*: they are related to a "model" or average performance level developed by customers with considerable experience in a product category (across companies and brands) or access to industry standards.

2.4.2 Expectancy Disconfirmation

The most important theory for customer satisfaction analysis, in the context of consumer behavior, concerns Oliver's approach (Oliver, 1977, 1980, 1997; Churchill and Suprenant, 1982; Vavra, 1997). According to this particular methodological approach, satisfaction may be defined as a pleasant past-purchasing experience from a product or service given the ante-purchasing expectancy of the customer. The performance judgment process made by customers is presented in Figure 2.8, where the following should be noted:

- Customer perceptions play the most important role in the satisfaction creation process. Perceived performance is not necessarily the same with actual performance, as already emphasized in section 1.1.
- Perceived performance is compared with a standard that may refer to customer expectations (Oliver, 1997), or other comparison standards, as already mentioned (Woodruff and Gardial, 1996).
- The previous comparison results in disconfirmation, i.e. the difference between what was expected and what was received.
- Satisfaction is the evaluation or feeling that results from the disconfirmation process. As Woodruff and Gardial (1996) urge, it is not the comparison itself (i.e. the disconfirmation process), but it is the customer's response to the comparison, given the emotional component of satisfaction.
- Finally, satisfaction feeling leads to various attitude and behavioral outcomes, such as repeat purchase intentions, word of mouth, brand loyalty, etc.

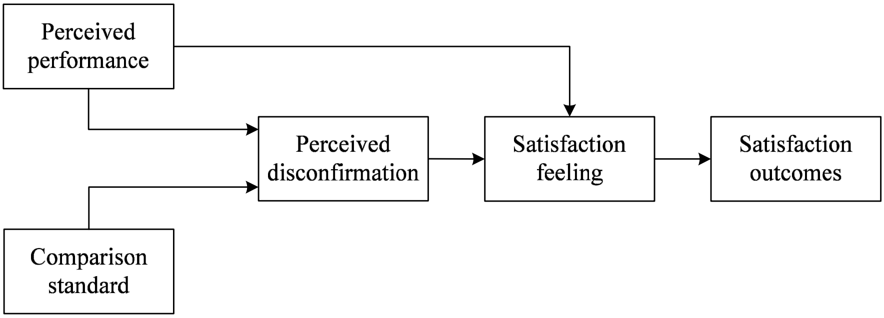


Fig. 2.8 Expectancy disconfirmation model (Woodruff and Gardial, 1996)

The aforementioned comparison process of the customer given his/her expectations is the key concept of this particular methodology. For this reason, Oliver's approach is also called as expectancy disconfirmation model.

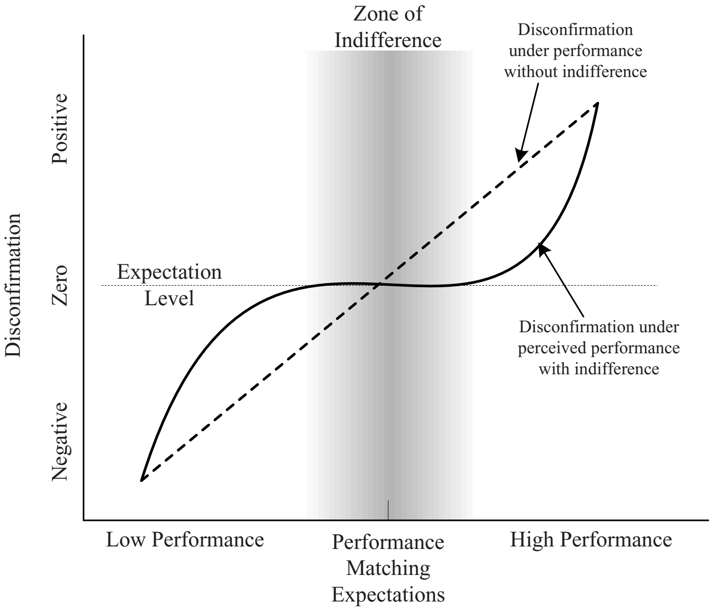
The existence of an indifference zone is an important aspect of the expectance disconfirmation process, since it suggests that disconfirmation and performance level are not proportionally related. This zone, which is also called attitude of acceptance in the assimilation-contrast theory, indicates that, from the consumer's perspective, there may be some latitude within which product performance may vary but it still fulfills the consumer's needs (Figure 2.9).

An analytical review of the expectancy disconfirmation model, which is one of the dominant theories of customer satisfaction influencing several research efforts, may be found in Churchill and Suprenant (1982), Yi (1991), and Erevelles and Leavitt (1992).

2.4.3 Fornell's model

Fornell's satisfaction model (Johnson and Fornell, 1991; Anderson and Fornell, 1991; Anderson and Sullivan, 1991; Anderson, 1994; Fornell, 1995) constitutes the basic measurement and analysis tool that is used in both the American Customer Satisfaction Index (ACSI) and the Swedish Customer Satisfaction Barometer (SCSB), as analytically presented in section 7.6.

This particular approach is based on an economic structural model that links different customer satisfaction measures (e.g. expectations, loyalty, complaints, etc.) with specific and pre-defined formulas. Given these defined relations between included variables, the model produces a system of cause and effect relationships.



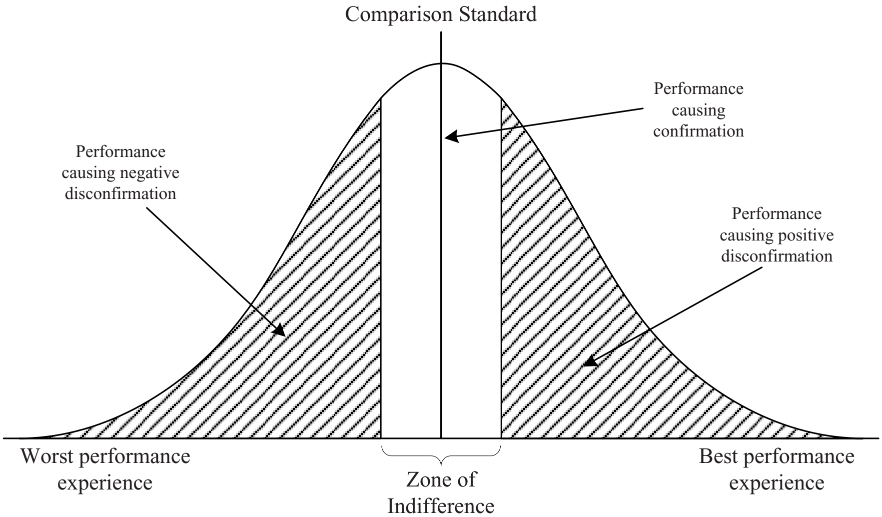


Fig. 2.9 Indifference zone in expectancy disconfirmation (Oliver, 1997; Woodruff and Gardial, 1996)

Generally, as presented in Figure 2.10, the model variables are analyzed in the following main categories:

1. *Satisfaction causes*: One of the most important assumptions of the model is that customer satisfaction has three antecedents: perceived quality, perceived value,

and customer expectations. The positive relation between customer satisfaction and perceived quality is consistent with several studies from marketing and consumer behavioral analysis (Churchill and Suprenant, 1982; Westbrook and Reilly, 1983; Tse and Wilton, 1988; Yi, 1991; Fornell, 1992). According to Deming (1981) and Juran and Gryna (1988), the evaluation of perceived quality should take into account the customization of the product or service to customer needs, as well as the product/service reliability. On the other hand, the quality/price ratio may be considered as the main estimate of perceived value, since it is used by customers for comparing similar products and services (Johnson, 1984). Another determinant of satisfaction refers to customer expectations (Oliver, 1980; Van Raaij, 1989). While perceived quality and value are based on recent customer experiences, customer expectations refer to all previous product/service purchase and usage experiences.

2. *Satisfaction*: Customer satisfaction is evaluated using a set of additional parameters, like disconfirmation of expectations and distance from the ideal product/service. These parameters are weighted in order to provide final estimates, while it should be noted that the model assumes that the previous three antecedents may be positively related (Howard, 1977; Johnson et al., 1995).
3. *Satisfaction results*: Following Hirschman's (1970) exit-voice, the consequences of customer satisfaction are focused on customer complaints and loyalty (Fornell and Wernefelt, 1987, 1988). Loyalty is the main dependent variable in the model because of its value as a proxy for profitability.

In this approach, customer satisfaction is based on multiple indicators and it is measured as a latent variable using Partial Least Squares (PLS). PLS is able to estimate this causal model and it is preferred because it is an iterative procedure that does not impose distributional assumptions on the data. PLS estimates weights for the variable measures that maximize their ability to explain customer loyalty as the ultimate endogenous or dependent variable (Fornell et al., 1996).

Furthermore, confirmatory factor analysis and linear equation modeling have been conducted to validate the relationships depicted in the model and the overall framework (Vavra, 1997).

2.4.4 Other Behavioral Models

There are several approaches from social psychology and consumer behavioral analysis that have been used in the customer satisfaction analysis problem. These approaches attempt to give a clearer understanding on how and why satisfaction is created, rather than to provide a quantitative measurement framework.

One of the most important categories of these approaches refers to motivation theories. As already noted, satisfaction is related to the fulfillment of customer needs. Thus, motivation theories may be used in order to indentify needs and study human motivation.

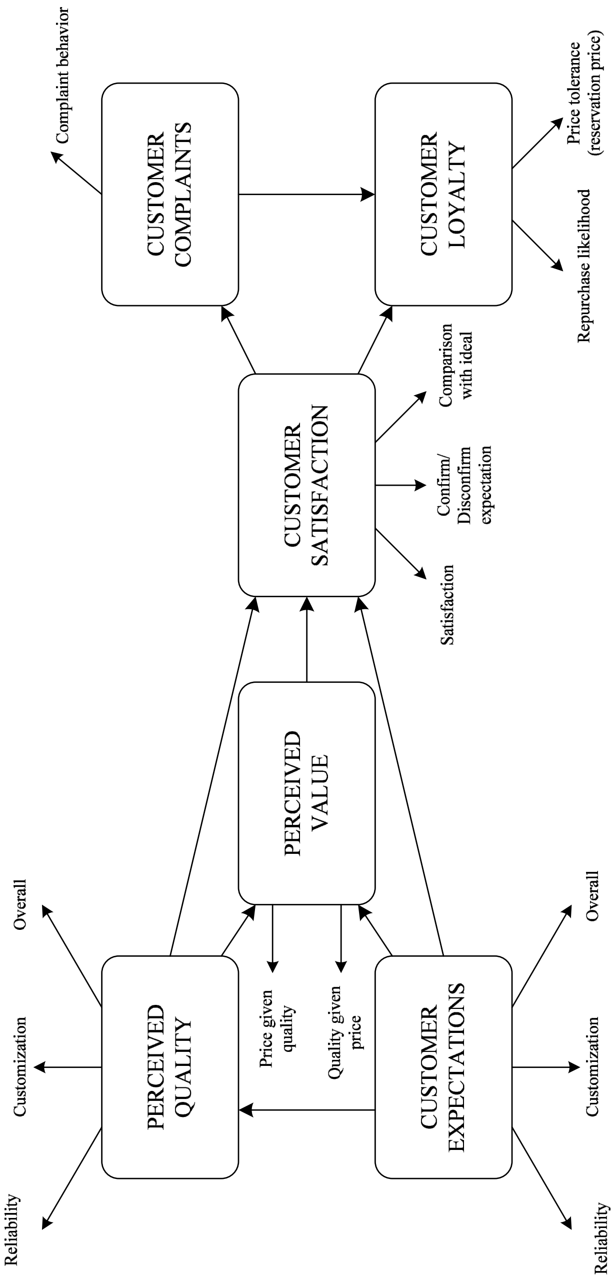


Fig. 2.10 Fornell's satisfaction model (Vavra, 1997)

In this context, early research efforts were focused not only on the determination, but also on the hierarchy of human needs (Murray, 1938; McClelland, 1961; Kassarjian, 1974; Horton, 1974). Maslow's need hierarchy is one the most popular approaches on human motivation. This hierarchy is often presented as a pyramid, and consists of the following stages (Maslow, 1943):

1. *Physiological needs*: biological needs necessary for human survival, like food, water, sleep, etc.
2. *Safety needs*: needs for safety and security, which include personal and financial security, health and well-being, etc.
3. *Love needs*: needs for love, affection and belongingness; they are also referred as needs for affiliation.
4. *Esteem needs*: needs for both the self-esteem and the esteem a person gets from others.
5. *Need for self-actualization*: need for self-fulfillment; self-actualization is described as a person's need to be and do that for which the person was "born to do".

The previous stages are presented in order of importance: the higher needs in this hierarchy only come into focus when the lower needs are met. It should be noted that later Maslow (1970) added a sixth stage: need for self-transcendence (i.e. the need to integrate with the human community rather than to remain as an individualist pursuing self-goals).

Alternative categorizations and hierarchies of human needs have also been proposed in the works of Herzberg et al. (1959), McClelland (1961), Alderfer (1972), and Alderfer et al. (1974).

The contribution of motivation theories to the customer satisfaction analysis problem is focused on the determination of "critical" satisfaction dimensions (Swan and Combs, 1976; Maddox, 1981). Product or service attributes may be classified to satisfiers and dissatisfiers, i.e. attributes that may cause satisfaction and dissatisfaction, respectively, according to their performance. Moreover, it should be mentioned that motivation theories have been focused on job satisfaction studies (see for example Herzberg et al., 1959; Herzberg, 1966, 1968). Oliver (1997) notes that these approaches are not widely adopted in consumer behavior, because they are not capable of generating an exhaustive set of satisfaction drivers, or even of choice criteria.

Another alternative behavioral approach refers to the equity theory, where equity is also referred as fairness, rightness, or deservingness to other entities, whether real or imaginary, individual or collective, person or non-person (Oliver, 1997). The "rule of justice", as proposed by Homans (1961) is the main concept of the equity theory: *"…A person's reward in exchange with other should be proportional to his/her investment…"*

Homan's approach suggests an outcome/input ratio, while reward and investment are used in a rather generic way. For example, in the customer satisfaction problem, customer reward may refer to the satisfaction caused by the usage of a product/service, or by the performance of its attributes. Similarly, investment may refer to the effort, time, or money paid by the customer in order to purchase or use a particular product/service.

According to the equity theory, satisfaction may be seen as the outcome of comparing rewards to investments, taking into account:

- the expectations (or predictions) of the customer,
- the rewards and investments of the company or the seller, and
- the rewards and investments of other customers.

A large number of studies referring to the application of the equity theory in the customer behavioral analysis problem may be found in the literature (Huppertz et al., 1978; Huppertz, 1979; Fisk and Coney, 1982; Mowen and Grove, 1983; Brockner and Adsit, 1986; Goodwin and Ross, 1990; Martins and Monroe, 1994; Lapidus and Pinkerton, 1995), while in several cases the approach is combined with the expectancy disconfirmation theory (Fisk and Young, 1985; Oliver and DeSarbo, 1988).

Finally, a relatively new approach in the context of social psychology that may be used in this particular problem is the regret theory. Since in many cases satisfaction is considered as a comparison outcome, the regret theory suggests that this outcome includes those that might have happened or those that did happen to another consumer who made a different choice of product/service (Bell, 1980; Loomes and Sugden, 1982). For example, a consumer may regret about his/her purchasing decision, thinking that he/she might have purchase an alternative product/service, or even take no purchasing decision at all.

The formulation of these comparison standards, i.e. the way a consumer thinks what might have happened, is mainly based on the following (Oliver, 1997):

- proactive observation (personal intentional direct observations),
- vicarious experience (observing the outcome of others who have made alternative choices), and
- simulation (imagine what might have happened in a hypothetical situation).

The effects of these comparison results on customer satisfaction are analytically presented in Figure 2.11.

The regret theory is one of the most recent research directions of consumer behavioral analysis that studies the customer satisfaction problem, while in several cases its applications are combined with marketing choice models (Harrison and March, 1984; Roese and Olson, 1993; Boninger et al., 1994; Roese, 1994).

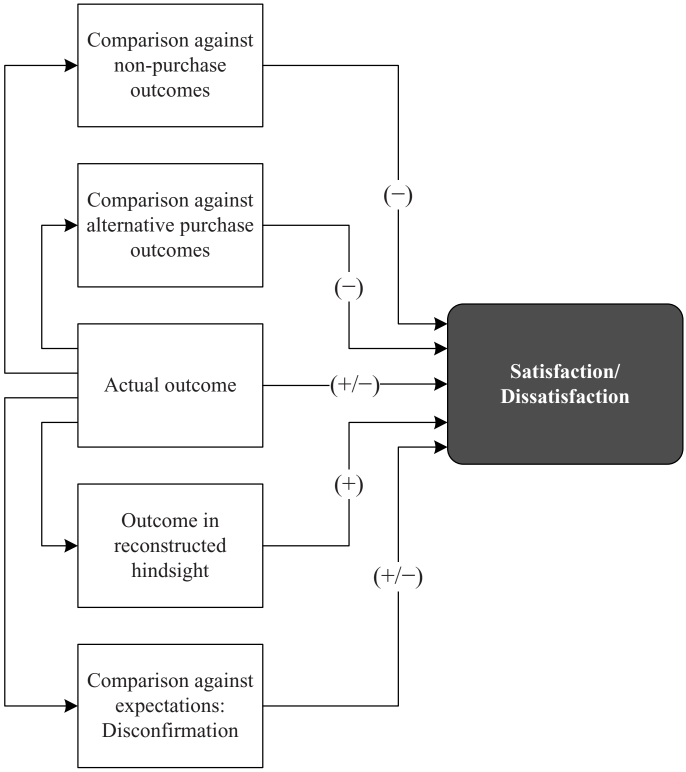


Fig. 2.11 Regret and hindsight effects on satisfaction (Oliver, 1997)

Springer
http://www.springer.com/978-1-4419-1639-6

Customer Satisfaction Evaluation
Methods for Measuring and Implementing Service
Quality
Grigoroudis, E.; Siskos, Y.
2010, X, 308 p. 129 illus., Hardcover
ISBN: 978-1-4419-1639-6