Chapter 6
Advanced Topics on the MUSA method

6.1 Computational Issues

The computational difficulty of the MUSA method is based on the number of variables and the number of constraints in the formulated LP. The method consists of two distinct stages: in the first stage an initial LP is solved in order to obtain an optimum value for the selected error function, while in the second stage a heuristic algorithm is used (solving a number of LPs) in order to explore the multiple or near optimal solutions space.

As a rule, the computational effort (CE) of a single LP may be estimated using the following expression:

$$CE \propto N_v \cdot N_c^2 \tag{6.1}$$

where N_v and N_c are the number of variables and constraints of the LP, respectively.

For example, the initial LP in the basic, and the generalized, MUSA method has $M+2$ constraints and $2M+(\alpha-1)+\sum(\alpha_i-1)$ variables, while in the post-optimality analysis stage n LPs are solved having $M+3$ constraints (the number of variables remains the same). Without loss of generality, we may assume that $\alpha=\alpha_i\ \forall\, i$, and therefore, the computational effort for the basic MUSA method is:

$$CE \propto \left[2M+(n+1)(\alpha-1)\right]\left[(M+2)^2+n(M+3)^2\right] \tag{6.2}$$

where M is the number of customers, n is the number of criteria, and α is the number of overall (or marginal) satisfaction levels.

As shown in expression (6.2), the computational difficulty of the MUSA method is heavily affected by the number of customers (see also Figure 6.1),

E. Grigoroudis and Y. Siskos, *Customer Satisfaction Evaluation*, International Series in Operations Research & Management Science 139, DOI 10.1007/978-1-4419-1640-2_6,

which is quite reasonable, since M determines the number of cases in a regression-type model.

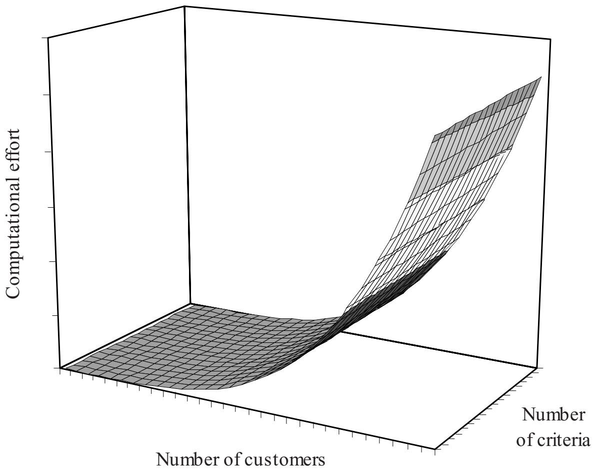

Fig. 6.1 Computational difficulty for the basic MUSA method

For this reason, the Dual Linear Program (DLP) of the MUSA formulation may be considered, in order to reduce the computational effort. In the case of the basic MUSA method, LP (4.17) may be written in the following form:

$$
\begin{array}{l}
[\min]\ \mathbf{1}'_{2M}\boldsymbol{\sigma} \\
\text{subject to} \\
\begin{bmatrix} \boldsymbol{\Theta} \\ \mathbf{1}'_{\sum_i(\alpha_i-1)} \\ \mathbf{0}'_{\sum_i(\alpha_i-1)} \end{bmatrix}\mathbf{w} + \begin{bmatrix} \boldsymbol{\Psi} \\ \mathbf{0}'_{(\alpha-1)} \\ \mathbf{1}'_{(\alpha-1)} \end{bmatrix}\mathbf{z} + \begin{bmatrix} \boldsymbol{\Lambda} \\ \mathbf{0}'_{2M} \\ \mathbf{0}'_{2M} \end{bmatrix}\boldsymbol{\sigma} = \begin{bmatrix} \mathbf{0}_M \\ 100 \\ 100 \end{bmatrix} \\
\text{with } \mathbf{w},\mathbf{z},\boldsymbol{\sigma} \geq 0
\end{array}
\tag{6.3}
$$

where $\mathbf{w}$, $\mathbf{z}$, and $\boldsymbol{\sigma}$ are the vectors of the model variables, $\mathbf{1}_x$ and $\mathbf{0}_x$ are vectors of size x with ones and zeros, respectively, $\boldsymbol{\Theta}$ and $\boldsymbol{\Psi}$ are matrices of size $M\times\sum(\alpha_i-1)$ and $M\times(\alpha-1)$, respectively, where θ_{ij} and ψ_{ij} are given according to formula (4.16), and $\boldsymbol{\Lambda}$ is a matrix of size $M\times 2M$ having the following form:

$$\mathbf{\Lambda} = \begin{bmatrix} -1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & -1 & 1 & \dots & 0 & 0 \\ \vdots & & & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & -1 & 1 \end{bmatrix}$$

The dual of LP (6.3) can be written as follows:

$$\begin{aligned} &[\max]\ 100u_{M+1} + 100u_{M+2} \\ &\text{subject to} \\ &\begin{bmatrix} \mathbf{\Theta}' & \mathbf{1}_{\sum_i(\alpha_i-1)} & \mathbf{0}_{\sum_i(\alpha_i-1)} \\ \mathbf{\Psi}' & \mathbf{0}_{(\alpha-1)} & \mathbf{1}_{(\alpha-1)} \end{bmatrix} \mathbf{u} \le 0 \\ &-1 \le u_i \le 1 \quad \text{for } i = 1, 2, \dots, M \\ &u_i \ \text{ free of sign } \ \forall i \end{aligned} \tag{6.4}$$

where $\mathbf{u}$ is the vector of dual variables with size $M + 2$.

The complexity of DLP (6.4) is based only on $(\alpha-1) + \sum(\alpha_i-1)$ constraints, since $-1 \le u_i \le 1$ are just boundary constraints. Thus the computational difficulty of DLP (6.4) is significantly smaller compared to the original LP (4.17).

It should be noted that the previous discussion refers only to the basic or the generalized MUSA method. The computational difficulty changes if we consider alternative objective functions for the post-optimality analysis stage. As shown in Table 6.1, the alternative MUSA methods presented in section 5.3 have different number of constraints and variables, while a different number of LPs has to be solved during the post-optimality analysis stage. Figure 6.2 shows the computational effort for these extensions of the MUSA method for a given number of criteria n and satisfaction levels α and α_i (the computational effort has been estimated using formula (6.1)). As expected, the complexity appears smaller for the generalized MUSA and the MUSA II methods, while MUSA III variation requires the highest computational effort.

6.2 Reliability Evaluation and Error Indicators

6.2.1 Average Fitting Indices

The reliability evaluation of the results is mainly related to the level of fitting to the customer satisfaction data, and the stability of the post-optimality analysis results.

Table 6.1 Problem size of alternative post-optimality approaches

Extension	Number of LPs	Number of constraints	Number of variables
Generalized MUSA	n	$M+3$	$2M+(\alpha-1)+\sum_{i=1}^{n}(\alpha_i-1)$
MUSA I	$2n$	$M+3$	$2M+(\alpha-1)+\sum_{i=1}^{n}(\alpha_i-1)$
MUSA II	$n+1$	$M+3$	$2M+\alpha+\sum_{i=1}^{n}\alpha_i$
MUSA III	$(\alpha-1)+\sum_{i=1}^{n}(\alpha_i-1)$	$M+3$	$2M+(\alpha-1)+\sum_{i=1}^{n}(\alpha_i-1)$
MUSA IV	1	$3M+3$	$2M+\alpha+\sum_{i=1}^{n}(\alpha_i-1)$

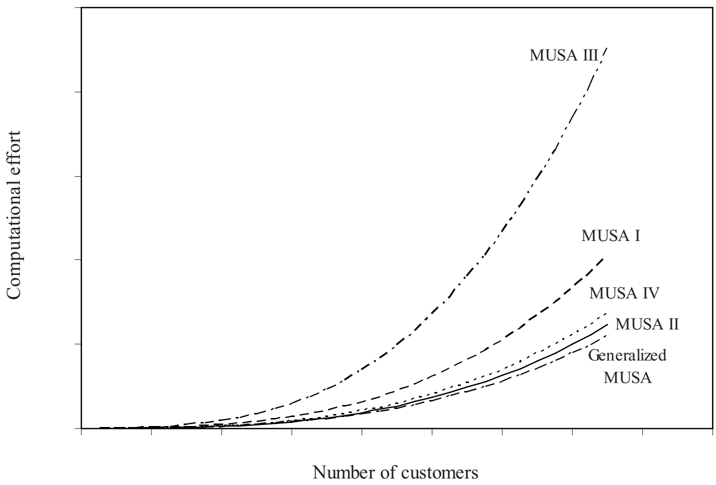

Fig. 6.2 Computational difficulty for the alternative MUSA methods

The fitting level of the MUSA method refers to the assessment of a preference collective value system (value functions, weights, etc.) for the set of customers with the minimum possible errors. For this reason, the optimal values of the error variables indicate the reliability of the value system that is evaluated.

Although several fitting measures may be assessed, all these indicators depend on the optimum error level and the number of customers. Grigoroudis and Siskos (2002) propose the following simple average fitting index AFI_1:

$$AFI_1 = 1 - \frac{F^*}{100M} \tag{6.5}$$

where F^* is the minimum sum of errors of the initial LP, and M is the number of customers.

AFI_1 is normalized in the interval [0, 1], and it is equal to 1 if $F^* = 0$, i.e. when the method is able to evaluate a preference value system with zero errors. Similarly, AFI_1 takes the value 0 only when the pairs of the error variables σ_j^+ and σ_j^- take the maximum possible values. It is easy to prove that $\sigma_j^+ \cdot \sigma_j^- = 0 \ \forall j$, i.e. the optimal solution has at least one zero error variable for each customer, given that the MUSA method is similar to goal programming modeling (Charnes and Cooper, 1961).

An alternative fitting indicator is based on the percentage of customers with zero error variables, i.e. the percentage of customers for whom the estimated preference value systems fits perfectly with their expressed satisfaction judgments. This average fitting index AFI_2 is assessed as follows:

$$AFI_2 = \frac{M_0}{M} \tag{6.6}$$

where M_0 is the number of customers for whom $\sigma^+ = \sigma^- = 0$

Although the previous fitting indicators are rather simple and can be easily calculated, they present several disadvantages. For example, AFI_1 may rarely take large values, since usually $F^* << 100M$. This is justified by the fact that it is unreasonable all the error variables in a regression-type model to have their maximum possible values, i.e. $\sigma_j^+ + \sigma_j^- = 100 \ \forall j$. For this reason, AFI_1 usually overestimates the fitting ability of the MUSA method. On the other hand, AFI_2 examines only the existence of non-zero errors, without taking into account the values of these error variables. For this reason, in several cases AFI_2 underestimates MUSA's fitting level. Additionally, the values of AFI_2 may not give a reliable indication for the overall fitting ability of the MUSA method, since a small (or high) value of AFI_2 does not imply a respective small (or high) sum of errors.

To overcome these disadvantages, a new fitting indicator may be assessed, which will be able to examine separately every level of overall satisfaction and to calculate the maximum possible error value for each one of these levels. As shown in Figure 6.3, for the estimation of y^{*m}, $0 \leq y^{*m} \leq 100$ holds and thereby, the maximum overestimation (σ^+) and underestimation (σ^-) errors are $100 - y^{*m}$ and y^{*m}, respectively. Thus, the overall maximum error for every overall satisfaction level is the maximum of the previous expressions.

Using this approach, an alternative formulation of AFI_1 may be developed. The new average fitting index AFI_3 takes into account the maximum values of the error variables for every global satisfaction level, as well as the number of customers that belongs to this level:

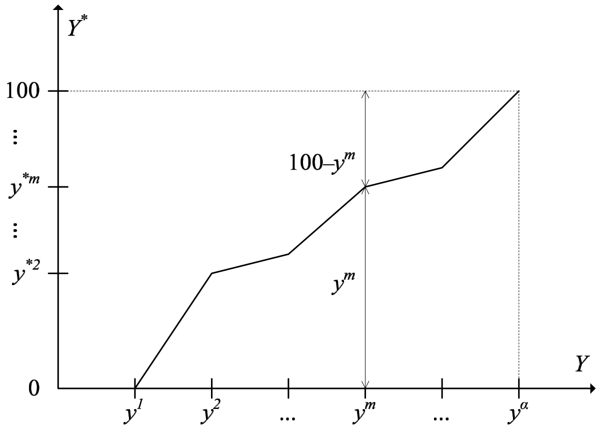

Fig. 6.3 Maximum error values for the m-th overall satisfaction level

$$AFI_3 = 1 - \frac{F^*}{M \sum_{m=1}^{\alpha} p^m \max\left\{ y^{*m}, 100 - y^{*m} \right\}} \tag{6.7}$$

where p^m is the frequency of customers belonging to the y^m satisfaction level.

Consequently, AFI_3 may be considered as a variation of AFI_1, for which $AFI_3 \leq AFI_1$ can be proved to hold. Although AFI_3 appears more reliable, all of the aforementioned average fitting indicators are highly affected by potential inconsistencies in customer satisfaction judgments. Therefore, the examination of all these indices may give a more complete view for the fitting ability of the MUSA method.

6.2.2 Other Fitting Indicators

One of the most useful tools, which may serve as an alternative fitting indicator of the MUSA method, is the variance diagram of the added value curve. This variance diagram (Figure 6.4) shows the value range that the customers' set gives for each level of the ordinal satisfaction scale. Therefore, it can be considered as a confidence interval for the estimated added value function.

This diagram depends upon the estimated satisfaction values and the optimal values of the error variables as well. The development process of this diagram consists of the following steps (Grigoroudis and Siskos, 2002):

Step 1:
For each customer j, the evaluated satisfaction value $\tilde{y}_j^{*m}$ is calculated according to the formula:

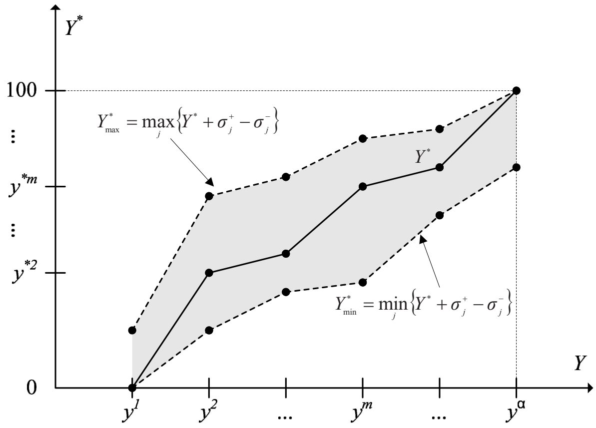

Fig. 6.4 Variance diagram of the added value curve

$$\tilde{y}_j^{*m} = y^{*m} + \sigma_j^{+} - \sigma_j^{-} \tag{6.8}$$

where y^{*m} is the satisfaction value of level *m*, and $\sigma_j^{+}, \sigma_j^{-}$ are the corresponding error variables for customer *j*.

Step 2:
The maximum and minimum satisfaction curves $y_{\max}^{*m}$ and $y_{\min}^{*m}$ accordingly, are calculated for each level *m* of the ordinal satisfaction scale, using the following formula:

$$\begin{cases} y_{\max}^{*m} = \max\limits_j \left\{ \tilde{y}_j^{*m} \right\} \\ y_{\min}^{*m} = \min\limits_j \left\{ \tilde{y}_j^{*m} \right\} \end{cases} \quad \text{for } m = 1, 2, \ldots, \alpha \tag{6.9}$$

Another fitting indicator is the prediction table of global satisfaction, which is developed in a similar way according to the following steps (Grigoroudis and Siskos, 2002):

Step 1:
For each customer *j*, the evaluated satisfaction value $\tilde{y}_j^{*m}$ is calculated according to (6.8).

Step 2:
Based on the previous value, the evaluated satisfaction level $\tilde{y}_j^{m}$ is calculated for each customer *j*, according to the formula:

$$\tilde{y}_j^m = \begin{cases} y_j^1 & \text{if } \tilde{y}_j^{*m} \le \dfrac{y^{*2}}{2} \\ y_j^2 & \text{if } \dfrac{y^{*2}}{2} < \tilde{y}_j^{*m} \le \dfrac{y^{*3}+y^{*2}}{2} \\ \vdots & \\ y_j^\alpha & \text{if } \tilde{y}_j^{*m} > \dfrac{100+y^{*\alpha-1}}{2} \end{cases} \tag{6.10}$$

Step 3:
Using the actual (as expressed by the customers) and the estimated level of global satisfaction, y_j^m and $\tilde{y}_j^m$ accordingly, the number of customers belonging to each of these levels is calculated.

The general form of a prediction table is presented in Figure 6.5, and includes the following results for each actual and evaluated satisfaction level:

- N_{ij}: the number of customers that have declared to belong to global satisfaction level i, while the model classifies them to level j.
- R_{ij}: the percentage of customers of actual global satisfaction level i that the model classifies to level j.
- C_{ij}: the percentage of customers of estimated global satisfaction level j that have declared to belong to level j.

R_{ij} and C_{ij} are calculated according to the formulas:

$$R_{ij} = \frac{N_{ij}}{\sum_{i=1}^{\alpha} N_{ij}}, \quad C_{ij} = \frac{N_{ij}}{\sum_{j=1}^{\alpha} N_{ij}} \quad \forall i, j \tag{6.11}$$

while the overall prediction level (OPL) is based on the sum of the main diagonal cells of the prediction table, and it represents the percentage of correctly classified customers:

$$OPL = \frac{\sum_{i=1}^{\alpha} N_{ii}}{\sum_{i=1}^{\alpha}\sum_{j=1}^{\alpha} N_{ij}} \tag{6.12}$$

In general, it should be mentioned that the fitness of the MUSA method is not satisfactory when a high percentage of customers appears away from the main diagonal of the prediction table, i.e. a significant number of customers having declared to be very satisfied is predicted to have a low satisfaction level and vice versa.

Actual Global Satisfaction Level \ Predicted Global Satisfaction Level	$\tilde{y}^1$	$\tilde{y}^2$		$\tilde{y}^j$		$\tilde{y}^\alpha$
y^1	N_{11} R_{11} C_{11}	N_{12} R_{12} C_{12}	…	N_{1j} R_{1j} C_{1j}	…	$N_{1\alpha}$ $R_{1\alpha}$ $C_{1\alpha}$
y^2	N_{21} R_{21} C_{21}	N_{22} R_{22} C_{22}	…	N_{2j} R_{2j} C_{2j}	…	$N_{2\alpha}$ $R_{2\alpha}$ $C_{2\alpha}$
	⋮					⋮
y^i	N_{i1} R_{i1} C_{i1}	N_{i2} R_{i2} C_{i2}	…	N_{ij} R_{ij} C_{ij}	…	$N_{i\alpha}$ $R_{i\alpha}$ $C_{i\alpha}$
	⋮					⋮
y^α	N_{a1} R_{a1} C_{a1}	N_{a2} R_{a2} C_{a2}	…	N_{aj} R_{aj} C_{aj}	…	$N_{\alpha\alpha}$ $R_{\alpha\alpha}$ $C_{\alpha\alpha}$

Fig. 6.5 Prediction table of global satisfaction

6.2.3 Average Stability Index

The stability of the results provided by the post-optimality analysis is not related to the degree of fitness of the MUSA method. More specifically, during the post-optimality stage, n LPs are formulated and solved, which maximize repeatedly the weight of each criterion. The mean value of the weights of these LPs is taken as the final solution, and the observed variance in the post-optimality matrix indicates the degree of instability of the results. Thus, an average stability index *ASI* may be assessed as the mean value of the normalized standard deviation of the estimated weights:

$$ASI = 1 - \frac{1}{n}\sum_{i=1}^{n} \frac{\sqrt{n\sum_{j=1}^{n}\left(b_i^j\right)^2 - \left(\sum_{j=1}^{n} b_i^j\right)^2}}{100\sqrt{n-1}} \tag{6.13}$$

where b_i^j is the estimated weight of the i-th criterion in the j-th post-optimality analysis LP.

ASI is normalized in the interval [0, 1], and it should be noted that when this index takes its maximum value, then:

$$ASI = 1 \Leftrightarrow b_i^j = b_i \quad \forall i, j \tag{6.14}$$

where b_i is the final estimated weight for criterion i.

On the other hand, if *ASI* takes its minimum value, then:

$$ASI = 0 \Leftrightarrow b_i^j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad \forall i, j \tag{6.15}$$

It should be emphasized that the aforementioned stability index refers to the basic or the generalized MUSA method. In case of alternative objective functions during the post-optimality analysis stage, formula (6.13) should be modified taking into account the number of LPs solved during this stage (see Table 6.1).

Generally, apart *ASI*, the variance of the weights during post-optimality analysis (see section 9.5.4) is also able to provide valuable information for the stability analysis of the results provided by the MUSA method. This diagram can give a confidence interval for the estimated weights, and can identify possible competitiveness in the criteria set, i.e. the existence of certain customer groups with different importance levels for the satisfaction criteria.

6.3 Selection of Parameters and Thresholds

6.3.1 Preference Thresholds

The problem of selecting appropriate model parameters is focused on the preference values γ, γ_i, and the tradeoff threshold ε during the post-optimality analysis.

In this section, it is examined how different values of these parameters may affect the fitting and stability level of the MUSA results. For this reason, a large number of indicative customer satisfaction data sets have been used. These data sets present different characteristic properties (e.g. number of criteria, number of satisfaction levels, consistency of judgments and stability level, etc.). One of the most important results of this analysis is that the selection of preference thresholds γ and γ_i depends mainly on the stability of the results.

In particular, in case of stable results, the average fitting index AFI_1, as well the average stability index *ASI*, have high values (~100%) for $\gamma = \gamma_i = 0$. The increase of γ and γ_i will cause a relatively small reduction of the fitting and stability level of the results, as shown in Figure 6.6(a). This finding may be justified by the fact that the preference thresholds provide a lower bound for the model variables z_m and w_{ik} (see formula (5.3)). For example, by increasing γ_i, the MUSA method is forced to assign a minimum weight of $\gamma_i\,(\alpha_i-1)$ to each criterion. Thereby, the initially achieved fitting and stability level of the results is decreased. Consequently, in case of stable results, it is preferred to set $\gamma = \gamma_i = 0$ (or at least very small values for the preference thresholds).

In case of unstable results, *ASI* may take rather small values (e.g. <50%) for $\gamma = \gamma_i = 0$, while AFI_1 may retain a relatively high level (e.g. >80%). Figure 6.6(b) reveals a competitive relation between *ASI* and AFI_1 in this case: the increase of preference thresholds γ and γ_i may improve the stability of the results, but it will decrease the fitting level of the model. As previously noted, this is justified by the fact that the preference thresholds determine the minimum value of the criteria weights. Thus, in case of instability, the increase of γ and γ_i will decreas the variability observed in the post-optimality table, and therefore, it will increase the average stability index.

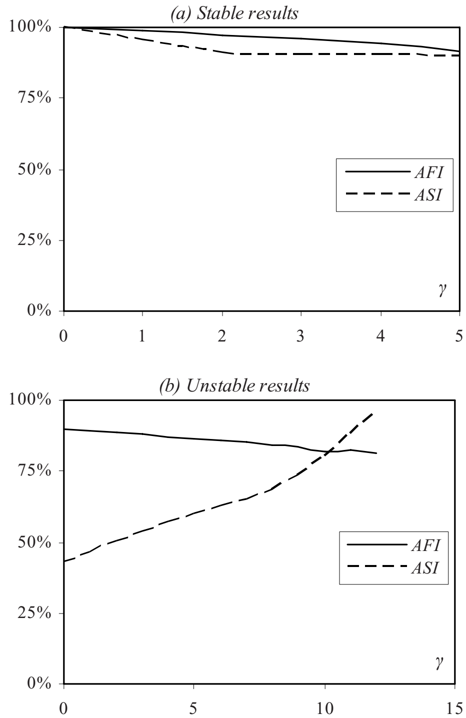

Fig. 6.6 Modification of AFI_1 and *ASI* for different values of γ

Generally, the process proposed in Figure 6.7 should be considered when selecting appropriate values for the preference thresholds γ and γ_i. This process is based on the work of Jacquet-Lagrèze and Siskos (1982) in the area of ordinal regression modeling. Moreover, it should be emphasized that special attention should be given when modifying the preference thresholds, because of the following main reasons:

- An arbitrarily large increase of the preference thresholds may falsify the customer satisfaction data set; large values of γ and γ_i require stronger assumptions for the preference conditions (5.1).
- Based on the assessed values of γ_i the minimum weight of criterion i is $\gamma_i\,(\alpha_i - 1)$. This assumption should be verified by the decision-maker.

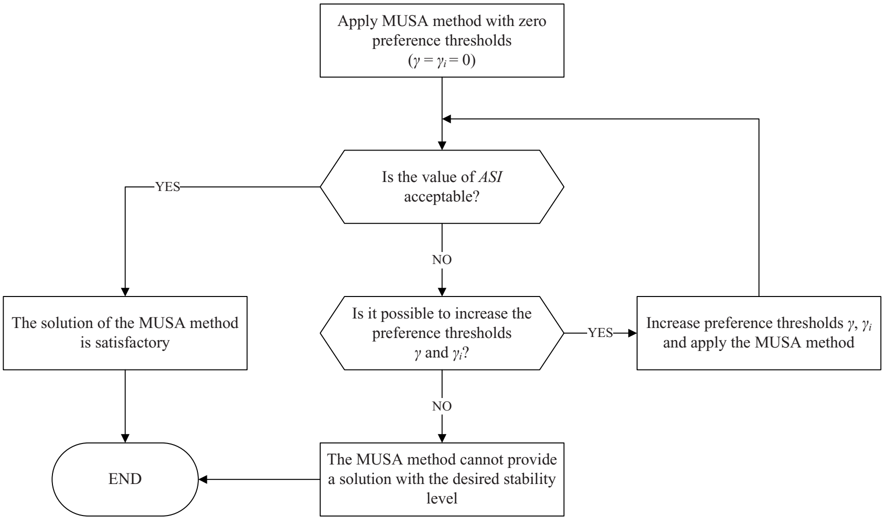

Fig. 6.7 A process for selecting preference thresholds

6.3.2 Post-optimality Thresholds

The post-optimality threshold ε does not affect the fitting ability of the model, since all the alternative fitting indices do not depend on the post-optimality results. Moreover, it should be noted that usually, in real world applications, $F^* > 0$, and thus ε may be assessed as a small percentage of the optimal value of the objective function F.

Similarly to the previous analyses, a large number of customer satisfaction data sets have been used, in order to examine the effect of post-optimality threshold on the stability level of the MUSA results. These experiments show that the increase of ε causes a decrease of the average stability index *ASI*, regardless of the stability level of results. This is rather expected, since an increase of ε implies an increase of the near optimal solutions space (see Figure 4.10).

As shown in Figure 6.8, the decrease of *ASI* is larger in case of unstable results because F^* is larger and, thus, the overall tradeoff value $(1+\varepsilon)F^*$ is larger in the post-optimality analysis. For this reason, the results presented in Figure 6.8(a)-(b) are not straightforward comparable (i.e. for the same value of ε, the tradeoff value $1+\varepsilon)F^*$ is larger for unstable results than for stable results).

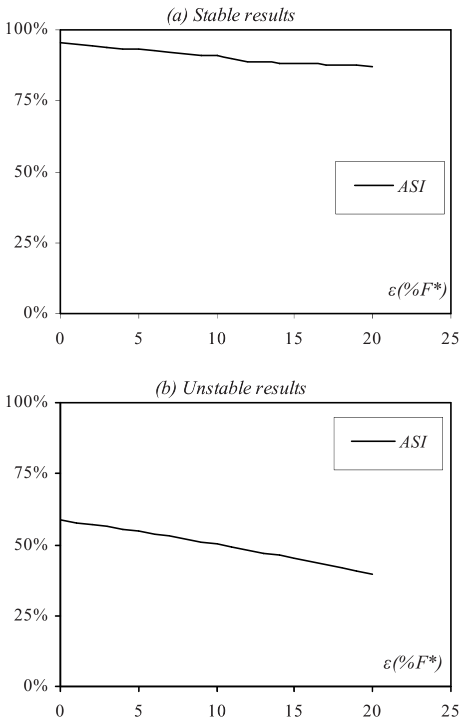

Fig. 6.8 Modification of *ASI* for different values of ε

Consequently, ε is a near optimal solutions threshold that should be always selected as a small percentage of F^*. The modification of ε should take into account the following:

- A very large value of ε will falsify the information provided by the post-optimality analysis, and decrease the stability ability of the model.
- A very low value of ε will not give the ability to explore the near optimal solutions space during post-optimality analysis.

6.4 Experimental Comparison Analysis

6.4.1 Design of the simulation process

The experimental research is the most important approach for comparing alternative methodologies. The main aim of this section is to present an experimental comparison analysis for different customer satisfaction evaluation models. The comparison concerns the MUSA method and the ordered conditional probability models (ordered Logit-Probit analysis), since all these models require the same type of input data, while they respect the qualitative form of the examined variables.

The first stage of the experimental comparison analysis refers to the design of the simulation process and aims at generating customer satisfaction data sets with different predefined characteristics. In particular, the data generation procedure is based on the principal that customer behavior (satisfaction judgments) may be explained through an explicitly defined set of value functions for a set of satisfaction criteria.

As Figure 6.9 shows, the data generation algorithm for the presented experiment consists of the following main steps:

Step 1:
In this initial step, the main parameters of the data sets are defined. These parameters include:

1. The number of satisfaction criteria n.
2. The number of the overall satisfaction levels α, as well as the number of the satisfaction levels of each satisfaction criterion α_i ($i = 1, 2, \ldots, n$).
3. The deviation level D_e (with $D_e \in [0, 1]$).
4. The desirable size of the data set M.

In addition, a set of value functions for the overall satisfaction Y^* and the marginal satisfaction $b_i X_i^*$ ($i = 1, 2, \ldots, n$) is selected in this step. For these value functions, the following monotonicity and normalization constraints must hold:

$$\begin{cases} y^{*1} = 0 \text{ and } y^{*a} = 100 \\ y^{*m} \le y^{*m+1} \quad m = 1,2,\ldots a-1 \end{cases}$$
$$\begin{cases} b_i x_i^{*1} = 0 \text{ and } \sum_{i=1}^{n} b_i x_i^{*a_i} = 100 \quad i = 1,2,\ldots,n \\ b_i x_i^{*k} \le b_i x_i^{*k+1} \quad k = 1,2,\ldots a_i - 1 \text{ and } i = 1,2,\ldots,n \end{cases} \tag{6.16}$$

where it should be noted that the marginal value functions are written in a non-normalized form, in order to reduce the number of parameters; this way, it is not necessary to estimate the criteria weights.

Step 2:
The main properties of the data set are defined through this step. These properties are largely determined by the value functions assessed in the previous step. However, generating random data based on these value functions does not guarantee a consistent data set. For this reason, in the current step, a matrix of excluding values for every possible data combination is developed. This matrix is assessed according to the following formula:

$$E\left(i_1, i_2, \ldots i_n\right) = \begin{cases} 0 \ \text{ if } \exists k : \left| \sum_{j=1}^{n} b_j x_j^{i_j} - y^{*k} \right| \le D_e \\ 1 \ \text{ otherwise} \end{cases} \qquad i_j = 1, 2, \ldots, \alpha_i \tag{6.17}$$

The matrix $E(.)$ is able to determine if any data combination $(i_1, i_2, \ldots, i_n)$ is consistent, thus $E(i_1, i_2, \ldots, i_n) = 0$, or inconsistent thus $E(i_1, i_2, \ldots, i_n) = 1$.

Step 3:
The last step refers to the data generation process according to the aforementioned properties and assumptions. This process may be considered as a type of Monte Carlo simulation analysis. Analytically, the procedure consists of the following steps:

1. Generation of a set of random numbers $(v_1, v_2, \ldots, v_n)$, which corresponds to the satisfaction of a fictitious customer for each one of the defined satisfaction criteria. These numbers are generated randomly, i.e. $v_j \sim U(1, \alpha_j)$, respecting the selected satisfaction levels.
2. If the previous data combination is inconsistent, that is $E(i_1, i_2, \ldots, i_j) = 1$, these numbers are rejected and a new random data set is generated. In the opposite case, the optimal level of the overall value function y^{*m} is calculated. In order to achieve the maximum consistency between y^{*m} and the data combination $(v_1, v_2, \ldots, v_n)$ the following is applied:

$$\left| \sum_{j=1}^{n} b_j x_j^{*v_j} - y^{*m} \right| = \min_k \left| \sum_{j=1}^{n} b_j x_j^{*v_j} - y^{*k} \right| \tag{6.18}$$

3. The values $(y^{*m}, v_1, v_2, \ldots, v_n)$ are added in the data set and the previous steps are repeated starting with the generation of a new set of random numbers $(v_1, v_2, \ldots, v_n)$. The algorithm ends when the desired data set size is reached.

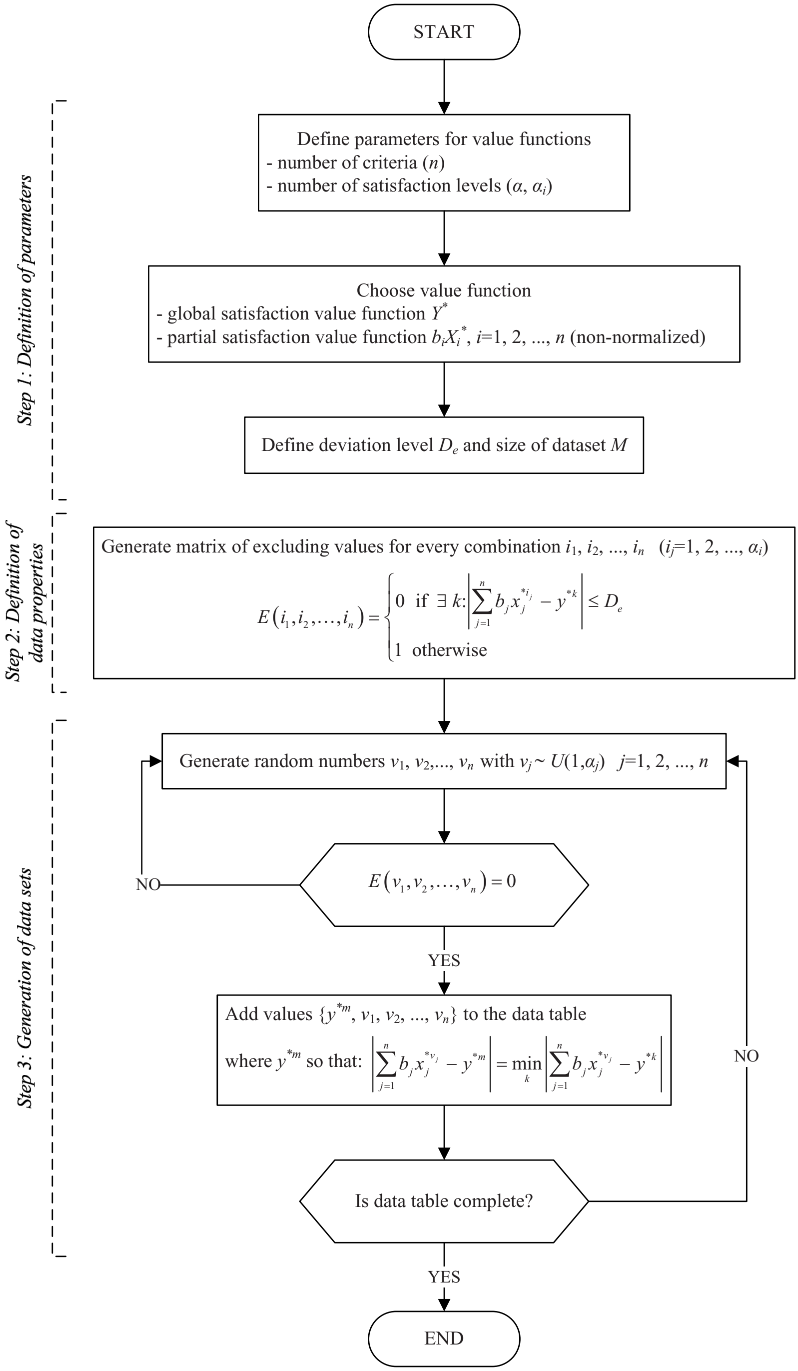

Fig. 6.9 Data generation algorithm

Using the presented algorithm, 16 different customer satisfaction data sets have been generated. These data sets are characterized by four different properties, as shown in Table 6.2. For each one of these properties (deviation level, number of customers, number of criteria, and number of satisfaction scales), 2 different values have been chosen, based on a series of pre-tests that are able to discriminate the results estimated by the MUSA method and the Logit-Probit analysis. This is the main reason for the large difference of the chosen deviation levels (5% and 40%). Moreover, the selected size of the data sets (500 and 1000 customers) is similar to the number of customers participating in real-world satisfaction surveys.

Table 6.2 Properties of the generated data sets

Deviation level (D_e)	Number of customers (M)	Number of criteria (n)	Number of satisfaction levels (α=α_i)
0.05-0.40	500-100	3-5	3-5

Finally, it should be noted that different families of value functions have been defined for each one of these generated data sets. Similarly, these sets of value functions present different characteristics, concerning the weights of the criteria and the shape of the assumed curve:

- The coefficient of variation may be used in order to measure differences in the distribution of criteria weights. For the defined value functions, the coefficient of variation for the selected weights ranges in [0.43, 0.73] with an average of 0.53.
- As noted in section 4.3.3, the average demanding indices are able to indicate the shape of a value function. For the defined value functions, these indices have the maximum possible range (i.e. [–1, 1]), with an average of –0.01.

Finally, for reasons of simplicity and without loss of generality, an equal number of satisfaction levels have been assumed for the selected overall and marginal value functions, i.e. $\alpha = \alpha_i \ \forall \ i$.

6.4.2 Simulation Results

The generated customer satisfaction data sets are used in order to compare the evaluation results provided by the MUSA method, as well as other alternative models. The presented results do not focus on the analysis of customer satisfaction, but rather on how these models behave for each one of the experimental data sets.

Table 6.3 presents a summary of the simulation results for the MUSA method. The fitting level of the MUSA method is rather high, since AFI_1 ranges between 87.9% and 99.1%, with an average of 94.5% for the generated data sets. This justifies the ability of the MUSA method to effectively evaluate a value system for the set of customers. However, *ASI* appears to have smaller values. Although this index has an average of 75.1%, there are particular data sets where the MUSA method is not able to achieve a high level of stability. Since *ASI* is calculated from the results of the post-optimality analysis, these data sets refer to the cases where customers' judgments do not appear homogenous. This probably indicates a comparative relation among the criteria weights, given the variability observed in the post-optimality analysis table, and it is caused by the chosen high deviation level D_e. Finally, in order to examine if the MUSA method is able to accurately estimate the defined experimental parameters, the hit rate ability of the model is calculated. In this case, the hit rate is defined as the average absolute deviation between the initially assumed and the finally estimated criteria weights. As shown in Table 6.3, the estimation accuracy of the MUSA is relatively high, since the hit rate ranges between 82% and 99.2%, with an average of 93.9%. However, it should be noted that the increase of the number of parameters for the value functions, increases the degrees of freedom of the MUSA method, and thus the hit rate is decreased.

Table 6.3 Simulation results for the MUSA method

Index	Statistics	Value
AFI_1	Range	0.879-0.991
	Average	0.945
ASI	Range	0.120-0.986
	Average	0.751
Hit rate	Range	0.820-0.992
	Average	0.939

Another important objective of the experimental analysis is to examine the influence of the parameters of the MUSA method to the fitting and the stability level of the estimated results. For this reason, a series of one-way ANOVA analyses have been performed in order to analyze the influence of each parameter of the experiment to the calculated *AFI* and *ASI* indices. Tables 6.4 and 6.5 present the summary results for this analysis of variance, from where the following points raise:

- The chosen deviation level of the experiment does not affect *ASI*, but influences *AFI*. This is more or less expected, since D_e determines the consistency of the satisfaction judgments and therefore it is strongly related with the fitting ability of the MUSA method.
- The size of the data set (number of customers) does not seem to affect the fitting and stability level of the MUSA method.
- Similarly, both *AFI* and *ASI* are not influenced by the chosen number of criteria and the number of satisfaction levels. However, these parameters may have a greater impact on the stability level (p-value less than 10%).

- The distribution of criteria weights, as measured by the average coefficient of variation, seems to affect *ASI* because large differences in the selected criteria weights generate data sets with heterogeneous customer preferences, which lead to an increased variability in the post-optimality analysis.

Table 6.4 Summary results for one-way ANOVA (*AFI*)

Factors	*SS*	*df*	*MS*	*F*	*p*-value
Deviation level	0.013	1	0.013	32.477	0.000
Number of customers	0.000	1	0.000	0.000	0.985
Number of criteria	0.000	1	0.000	0.011	0.917
Number of satisfaction levels	0.000	1	0.000	0.011	0.917
Distribution of weights	0.001	3	0.000	0.206	0.890

Table 6.5 Summary results for one-way ANOVA (*ASI*)

Factors	*SS*	*df*	*MS*	*F*	*p*-value
Deviation level	0.036	1	0.036	0.501	0.491
Number of customers	0.000	1	0.000	0.000	0.993
Number of criteria	0.202	1	0.202	3.385	0.087
Number of satisfaction levels	0.202	1	0.202	3.385	0.087
Distribution of weights	0.618	3	0.206	5.920	0.010

Similarly to the previous analysis, the generated customer satisfaction data sets have been used in the conditional probability models. As presented in section 2.3.1, the estimated parameters of these models include the threshold values of the dummy dependent variable y^* (overall satisfaction) and the coefficients of the independent variables x_i (marginal satisfaction) in the regression formula (this formula relates y^* and x_i in a weighted sum expression). Since the conditional probability models have a different philosophy (assumptions, interpretation of parameters, etc.) compared to the MUSA method, a straightforward comparison of the results provided by this approaches is not possible. However, the thresholds of the dummy dependent variables may indicate the shape of the overall value function, while the regression coefficients represent a measure of the relative importance for the satisfaction criteria.

The simulation results for the ordered Probit and Logit analysis are presented in Tables 6.6 and 6.7, where a summary of t-test statistics is given for each one of the aforementioned parameters (the p-value represents the probability of error under the hypothesis of accepting the values of the estimated parameters). Overall, it appears that the fitting ability of these models is satisfactory, since in most of the generated data sets the p-value is small ($p < 0.0001$). However, in many cases (almost 40% of the generated data sets) the Probit and Logit models are not able to achieve a high fitting level. These cases do not only concern data sets where the

MUSA method is not able to provide reliable results, but they also refer to data sets where *AFI* and *ASI* indicators of the MUSA method have relatively high values. Finally, it should be noted that the estimated results of the Probit analysis do not differ significantly from those provided by the Logit analysis.

Table 6.6 Simulation results for the Probit model

Parameters	Index	Statistics	Value
Thresholds	t-test	Range	0.000-26.040
		Average	14.581
	p-value	Range	0.000-1.000
		Average	0.125
Regression coefficients	t-test	Range	–5.878-19.410
		Average	5.793
	p-value	Range	0.000-1.000
		Average	0.292

Table 6.7 Simulation results for the Logit model

Parameters	Index	Statistics	Value
Thresholds	t-test	Range	0.000-22.709
		Average	13.351
	p-value	Range	0.000-1.000
		Average	0.125
Regression coefficients	t-test	Range	–5.417-18.246
		Average	5.584
	p-value	Range	0.000-1.000
		Average	0.272

Apart from the fitting and stability analyses, the prediction table of global satisfaction (see Figure 6.5) may also be used in order to compare the estimated results of the MUSA method and the conditional probability models. To this end, formula (6.12) is used to calculate the *OPL* for each one of the generated data sets. A summary of the analysis, regarding the prediction ability of these alternative models, is given in Table 6.8, where it should be noted that for all the examined data sets, the *OPL* of the MUSA method is higher compared to the other models.

Table 6.8 Overall prediction level for alternative models

Model	Range	Average
MUSA	0.701-1.000	0.885
Probit analysis	0.622-1.000	0.784
Logit analysis	0.622-1.000	0.764

Furthermore, in case of highly consistent and homogenous data sets, a high prediction index appears for all alternative approaches. In general, the most important differences between the prediction levels achieved by the MUSA method and the Logit-Probit analysis concerns data sets with greater number of estimated parameters (e.g. number of criteria, number of satisfaction levels, etc.). Moreover, it seems that the size of the data sets does not affect the *OPL*, while the prediction index between the Probit and the Logit analysis is similar for all the data sets of the experiment.

The main result of the presented experimental comparison analysis is the high prediction ability of all alternative models, although *OPL* is slightly higher for the MUSA method. However, the fitting and stability level of the MUSA method is significantly higher compared to the conditional probability models for all data sets of the experiment. Moreover, in case of inconsistent and non-homogenous data, poor stability results may appear for all alternative approaches.

The presented experiment may be considered as a pilot analysis, since a larger number of data sets is required, in order to increase the reliability of the findings. Moreover, additional parameters and desired properties of the generated data sets may be examined (e.g. parameters of the MUSA method). The presented results examine the effect of several parameters to the fitting and the stability level of the MUSA method using one-way ANOVA analysis. For this reason, future research may focus on other alternative customer satisfaction evaluation models or examine how several combinations of these parameters may affect the reliability of the results. Finally, it should be noted that the development of an unbiased data generation process for satisfaction judgments is rather difficult, since it requires a strong assumption about the preference model of the customers. In the presented experiment, this assumption appears through the assumed value functions during the first step of the simulation process.