Complete Reference for this article:
Triantaphyllou, E., and K. Baig, (2004), **"The Impact of Aggregating Benefit and Cost Criteria in Four MCDA Methods,"** *IEEE Transactions on Engineering Management*, Vol. xx, No. xx, pp. xxx-xxx, in print.

The Impact of Aggregating Benefit and Cost Criteria in Four MCDA Methods

by Evangelos Triantaphyllou and Khalid Baig

Dept. of Industrial and Manufacturing Systems Engineering
Louisiana State University
3128 CEBA Building
Louisiana State University
Baton Rouge, LA 70803-6409

Email: trianta@lsu.edu
Web: http://www.imse.lsu.edu/vangelis

Last revision: July 21, 2004

ABSTRACT:
Multi-criteria decision analysis (MCDA) problems (also known as multi-criteria decision-making or MCDM) involve the ranking of a finite set of alternatives in terms of a finite number of decision criteria. Often times such criteria may be in conflict with each other. That is, an MCDA problem may involve both benefit and cost criteria at the same time. Although this is a frequent characteristic of many real-life MCDA problems, this subject has not received adequate attention in the literature. This paper examines the use of four key MCDA methods when two approaches for dealing with conflicting criteria are used. The two approaches are the benefit to cost ratio approach and the benefit minus cost approach. The MCDA methods used in this study are the weighted sum model, the weighted product model, and the analytic hierarchy process (AHP) along with some of its variants, including the multiplicative AHP. Not surprisingly, these two approaches for aggregating conflicting criteria may result in a different indication of the best alternative or ranking of all alternatives when they are used on the same problem. As it is demonstrated here, it is also possible for the two approaches to even result in the opposite ranking of the alternatives. An extensive empirical analysis of this methodological problem revealed that the previous phenomena might occur frequently on simulated MCDA problems. The WSM, the AHP, and the revised AHP performed in an almost identical manner in these tests. The contradiction rates in these tests were rather significant and became more dramatic when the number of alternatives was high. Although it may not be possible to know which ranking is the "correct" one, this study also theoretically proved that the multiplicative AHP is immune to these ranking inconsistencies.

KEY WORDS:
Decision criteria, conflicting criteria, preference disaggregation methods, benefit and cost criteria, reward and loss criteria, decision-making, multi-criteria decision analysis (MCDA), multi-criteria decision-making (MCDM), analytic hierarchy process (AHP), weighted sum model, multiplicative AHP.

1 Some Background Information

Multi-criteria decision analysis (MCDA) is one of the most widely used decision methodologies in the sciences, business, and engineering worlds. Some applications of MCDA in engineering include the use on flexible manufacturing systems [Wabalickis, 1988], layout design [Cambron and Evans, 1991], integrated manufacturing systems [Putrus, 1990], and the evaluation of technology investment decisions [Boucher and Mcstravic, 1991].

A typical problem in MCDA is concerned with the task of ranking a finite number of decision alternatives, each of which is explicitly described in terms of different characteristics (also often called attributes, decision criteria, or objectives) which have to be taken into account simultaneously. Usually, an MCDA method aims at one of the following four goals, or "*problematics*" [Roy, 1985], [Jacquet-Lagreze and Siskos, 2001]:

Problematic 1: Find the best alternative.
Problematic 2: Group the alternatives into well-defined classes.
Problematic 3: Rank the alternatives in order of total preference.
Problematic 4: Describe how well each alternative meets all the criteria simultaneously.

Many interesting aspects of MCDA theory and practice are discussed in [Hobbs, 2000; and 1986], [Hobs, et al., 1992], [Stewart, 1992], [Triantaphyllou, 2000], [Zanakis, et al., 1995], and [Zanakis, et al., 1998].

Another term that is also used frequently to mean the same type of decision models is *multi-criteria decision-making* (MCDM). It should be stated here that the term MCDM is also used to mean finding the best alternative in a continuous environment. This is a different type of decision problem than the one highlighted in the previous paragraph. In that setting the alternatives are not known a priori but they can be determined by calculating the values of a number of discrete and/or continuous variables.

All methods of multi-criteria decision analysis require information regarding the relative or absolute importance of each criterion. The main challenge of a multi-criteria problem is that, mathematically speaking, it is not well defined. A central problem is how to quantify all pertinent data. Even in the very special case of which one can know precisely the values of the different alternatives in terms of the decision criteria, it is not clear how to process the data. For such cases the main methodological problem is how to process data which may be expressed in different units.

More often than not there exist contradictions between the different decision criteria, in such a way that an action (alternative) *a* might be 'better' than an action *b* in terms of one criterion, and 'worse' in terms of another. Another major problem occurs when decision criteria can be grouped into two opposite categories, usually called the **"benefit"** and the **"cost"** criteria. Alternatively, benefit criteria may be called **"reward"** criteria and cost criteria **"regret"** or **"loss"** criteria. A benefit criterion means that the higher an alternative scores in terms of it, the better the alternative is. The opposite is considered true for the cost criteria.

This kind of decision problems with conflicting criteria is very common in engineering applications. It is hard to imagine a complex real-life engineering problem which does not involve the need to evaluate alternatives in terms of conflicting criteria. That means that many problems need a kind of conflicting criteria analysis in view of a finite number of possible alternatives.

This analysis can be done in terms of a number of ways. One way is to use criteria aggregation techniques that group the decision criteria into two sets: one set for the *"benefit"* criteria and another set for the *"cost"* criteria. Next the analyst uses the *"benefit"* criteria to compute an aggregated index. The same is done with the *"cost"* criteria. Then the alternatives are evaluated by using information from these two aggregated indexes. Methods that belong to this category are the Analytic Hierarchy Process and its variants (AHP) [Saaty, 1994], the weighted sum model [Fishburn, 1967], and the weighted product model [Bridgeman, 1922], [Miller and Starr, 1969].

A closely related idea is to use "outranking" methods. Such methods also group criteria into two categories as above. They also compute aggregated indexes for each one of the two categories and compare the alternatives two at a time in terms of these two aggregates. Finally, these methods propose a ranking of the alternatives. Such methods are the ELECTRE group of methods [Roy and Bouyssou, 1993] the TOPSIS method [Hwang and Yoon, 1981], and the PROMETHEE group of methods (for an overview please see [Figueira, De Smet and Barns, 2004].

A different way is to use an explicit trade-offs approach which is based on the so-called *"value functions"* [Keeney and Raiffa, 1976], [Keeney, 1992], [Kirkwood, 1997]. A value function attempts to map changes of values of performance of the alternatives in terms of a given criterion, into a dimensionless value. Some key assumptions are made in the process for transferring changes in values into these dimensionless quantities (see, for example, the fourth chapter of the book by Kirkwood [1997]). If the criterion is of the "benefit" type, then we would like this value to be high. The opposite is true if the criterion is of the "cost" type. The roots of this type of analysis can be found in [Edwards, 1977], [Edwards and Barron, 1994], [Edwards and Newman, 1986], and [Dyer and Sarin, 1979].

Although the type of decision analysis which is based on value functions has many supporters, it is not the subject of study in this paper. Perhaps an interesting research task is for one to compare value function based approaches with approaches that use an explicit aggregation of conflicting criteria into the "benefits" and "costs" groups. Such a study would require the analysis of many behavioral, psychological, as well as mathematical properties of such different multi-criteria decision analysis philosophies.

Finally, a third approach is based on "preference disaggregation approaches" (or PDA). Such methods use the notion of a reference set of alternatives. That is, it is assumed that the decision maker has access to a set of ranked alternatives and any new ranking should somehow be consistent with that ranking. Next, regression methods are applied to infer a model that was implicitly used to derive the reference ranking. This model is calibrated with the use of linear programming (LP) optimization models in order to achieve an adequate level of consistency. New alternatives are then ranked by using this derived model. An overview of this approach is given in [Jacquet-Lagreze and Siskos, 2001]. Some related recent work can be found in [Doumpos and Zopounidis, 2002; and 2004].

In [Triantaphyllou 2001] a number of MCDA methods were evaluated in terms of a number of evaluative criteria. In that study it was found that all of them do exhibit ranking irregularities. Such ranking irregularities, for instance, appear when the transitivity property is not observed. That is, when a large MCDA problem is decomposed into a number of small MCDA problems, each defined on two alternatives, then often times it is possible to have situations in which three alternatives, say A, B, and C, exhibit A > B, B > C, but A < C. Other evaluative tests do demonstrate similar ranking irregularities. Recently, in [Wang and Triantaphyllou, 2004] it is demonstrated that these problems do occur with some of the ELECTRE type of MCDA methods.

In common practice, when criteria can be divided into the *"benefit"* and *"cost"* groups, practitioners use two processing approaches for ranking the alternatives. The first approach is the **benefit to cost approach** and the second one is the **benefit minus cost approach** (see, for example, [Saaty, 1994], [Saaty and Vargas, 1994], [Bozeman and Melkers, 1993], [Augood, 1978], and [Kuwahara and Takeda, 1990]). These approaches are then combined with the standard MCDA methods and rank the alternatives of a given problem. The benefit to cost ratio approach and the benefit minus cost approach are mathematically two different scales (or functions) for quantifying the existing differences between two mathematical quantities. The former approach is a ratio scale and the latter, is an interval scale. Although both approaches are not expected to give the same numerical value, they give a relationship between two quantities, that is, an indication of which quantity is higher and which one is lower.

This paper demonstrates that when the two approaches for aggregating conflicting criteria into two groups (i.e., the benefits minus cost approach and the benefits to cost approach) are used on the same problem, even when using the same MCDA method, then one may derive very different rankings of the alternatives. Furthermore, an extensive empirical study revealed that this situation might occur rather intensively with random test problems. The only method that is invariant to these ranking disputes is the weighted product method (WPM) and a newer version of the AHP called the multiplicative AHP [Barzilai and Lootsma, 1994], [Lootsma, 1999]. However, in [Triantaphyllou and Mann, 1989] the WPM (and consequently the multiplicative AHP) method was found to be susceptible to some other methodological problems.

This paper is organized as follows. The second section briefly describes the four MCDA methods considered in this study. The third section illustrates, in terms of a number of demonstrative examples, the ranking disputes that can be generated when the previous two criteria aggregation approaches are used. The fourth section describes the way simulated test problems were generated on a computer. The fifth section analyzes the computational results. Finally, the paper ends with a conclusions section.

2 Brief Description of the MCDA Methods Used

The typical MCDA problem considered in this paper is comprised by a number, say m, of alternatives to be evaluated in terms of a number, say n, of decision criteria. The alternatives will be denoted as A_i (for $i = 1, 2, 3, ..., m$) and the criteria as C_j (for $j = 1, 2, 3, ..., n$). Each criterion is associated with a weight of importance, denoted as w_j (for $j = 1, 2, 3, ..., n$). In general, the higher the weight is, the more important the criterion is assumed to be. Usually these weights are normalized so they add up to one. Furthermore, when an alternative A_i is considered in terms of criterion C_j, the decision maker is assumed to know the corresponding performance value a_{ij}. That is, a_{ij} denotes how well alternative A_i meets criterion C_j. Different MCDA methods may impose different requirements on these performance values. When the data are qualitative or are expressed in different units of measurement, some methods (such as the AHP) require that these values be normalized and thus be transformed into dimensionless quantities. The above data are best summarized in a **decision matrix** as follows:

	Criteria			
	C_1	C_2	...	C_n
	$(w_1$	w_2	...	$w_n)$
Alts.				
A_1	a_{11}	a_{12}	...	a_{1n}
A_2	a_{21}	a_{22}	...	a_{2n}
.	.		.	
.	.		.	
.	.			.
A_m	a_{m1}	a_{m2}	...	a_{mn}

It should be clarified here that an entry a_{ij} in the previous decision matrix reflects the relative importance of alternative A_i in the set of all alternatives when they are evaluated in terms of criterion C_j. This is not to be confused with the entries of a judgment matrix (also known as a *pairwise matrix*) in which an entry a_{ij} reflects the relative importance of alternative A_i with alternative A_j when they are compared in terms of a single criterion. The decision matrix is of dimension $m\times n$ and its columns are usually the eigenvectors of a number of pairwise (judgment) matrices. The later type of matrices (i.e., pairwise/judgment) are always square (actually reciprocal ones).

In the previous decision matrix the decision criteria have not been separated into the benefit and cost groups. Moreover, the problems examined in this paper are deterministic in nature. Next, we present a brief description of the way the MCDA methods considered in this paper process a decision matrix.

2.1 The Weighted Sum Model

The Weighted Sum Model (WSM) is the simplest and the most commonly used method in MCDA. The basic principle behind this technique is the additive utility assumption. That is, if the performance of each alternative in terms of each criterion in the decision problem (i.e., the a_{ij} values) is measurable and is of the same unit where higher is better, then the alternative with the largest cumulative value is the best. For example, if all the criteria represent benefit, then the most preferred alternative is the one for which the preference value (denoted as P_i) satisfies the following expression [Fishburn, 1967](for the maximization case):

$$P^*_{WSM} = \max_i P_i = \max_i \sum_{j=1}^{n} a_{ij} w_j, \quad \text{for } i = 1, 2, 3, ..., m. \qquad (2\text{-}1)$$

where P^*_{WSM} is the WSM preference value of the best alternative. Next the alternatives can be ranked according to their P_i vales.

In WSM the performance measures of the alternatives must be both numerical and comparable and expressed in the **same unit**. A violation may occur, because of the assumption of additive utility, when dealing with multi-dimensional criteria.

If the data are not expressed in terms of the same unit, then one may employ an approach based on *tradeoffs* that leads to the definition of what is called *single dimensional value functions* (see, for example, Chapter 4 in [Kirkwood, 1997]). Next, the basic WSM formula (2-1) is applied to the transformed data. An alternative way to value functions is to normalize the data in which case the WSM is transformed into an additive AHP type of model as described in the following sub-sections. In this paper when we talk about WSM, we mean that all the data are expressed in the same unit from the beginning of the definition of the problem and that formula (2-1) is used.

2.2 The Analytic Hierarchy Process

Details about this method can be found in [Saaty, 1980; and 1994]). This method has attracted the interest of many researchers and practitioners alike. A central difference with the WSM method is that now the a_{ij} values of the decision matrix are normalized vertically. That is, the elements of each column in the decision matrix add up to one. In this way, values with units of measurement can be transformed into dimensionless ones. When all the criteria express some type of benefit, then according to the AHP the best alternative is the one that satisfies the following expression:

$$P^*_{AHP} = \max_i P_i = \max_i \sum_{j=1}^{n} a_{ij} w_j , \quad \text{for } i = 1, 2, 3, ..., m. \qquad (2\text{-}2)$$

Although this formula is similar to the one used by the WSM, it should be emphasized again that now the a_{ij} values have been normalized to add up to one. Saaty uses an approach for eliciting judgments that compare two decision-making items (i.e., a pair of alternatives or a pair of criteria) at a time in terms of a single criterion (for the case of comparing two alternatives) or how well two criteria meet a goal. That approach is based on the use of a scale for converting linguistic statements about the relative importance of decision-making items (i.e., alternatives or criteria) into numerical statements. More on this can be found in [Saaty, 1980; and 1994] or in [Triantaphyllou, 2000].

2.3 The Revised Analytic Hierarchy Process

The Revised Analytic Hierarchy Process (RAHP), was proposed by Belton and Gear [Belton and Gear, 1983]. These two authors demonstrated a case of ranking abnormality when the original AHP is used. After many debates and a heated discussion (e.g., [Dyer, 1990*a*; and 1990*b*], [Saaty, 1983; 1987; and 1990], and [Harker and Vargas, 1990]) Saaty accepted this variant and now it is also called the ***ideal mode*** AHP [Saaty, 1994]. According to this variant, the a_{ij} values of the decision matrix need to be normalized by dividing the elements of each column in the decision matrix by the largest value in that column. As before, the best alternative is given again by the additive formula (2-2), but now the normalization is different.

2.4 The Weighted Product Model

The Weighted Product Model (WPM) is a method that uses multiplication to rank alternatives instead of addition (which is used in the WSM, the AHP and its previous additive variant). Each alternative is compared with others in terms of a number of ratios, one for each criterion. Each ratio is raised to the power of the relative weight of the corresponding criterion. Generally, in order to compare two alternatives A_K and A_L, (where: $m \geq K, L \geq 1$) the following formula is used ([Bridgeman, 1922] and [Miller and Starr, 1969]):

$$R\left(\frac{A_K}{A_L}\right) = \prod_{j=1}^{n} \left(\frac{a_{Kj}}{a_{Lj}}\right)^{w_j} . \qquad (2\text{-}3)$$

If $R(A_K / A_L) \geq 1$, then A_K is more desirable than A_L (for the maximization case). Then the best alternative is the one that is better than or at least equal to all other alternatives.

The structure of the WPM eliminates any units of measure. Hence it is also called **<u>dimensionless analysis</u>**. Therefore, this method can be used for single and multi-dimensional decision problems.

2.5 The Multiplicative AHP

In [Barzilai and Lootsma, 1994] and [Lootsma 1999] a multiplicative version of the AHP was proposed. This method was further analyzed in [Triantaphyllou, 2000; and 2001]. According to that approach, the relative performance values a_{ij} and criteria weights w_j are not processed according to formula (2-2), but the WPM formula (2-3) is used instead.

Furthermore, one can use a variant of formula (2-3) to compute preference values that in turn, can be used to rank the alternatives. These preference values can be computed as follows:

$$P_{i,\,multi\text{ - }AHP} = \prod_{j=1}^{n} \left(a_{ij}\right)^{w_j}. \qquad (2\text{-}4)$$

Please note that if $P_i > P_j$, then $P_i/P_j > 1$, or equivalently, $P_i - P_j > 0$. That is, two alternatives A_i and A_j can be compared in terms of their preference values P_i and P_j by forming the ratios or, equivalently, the differences of their preference values.

3 Examples of the Use of the Two Ranking Approaches

The benefit to cost ratio approach and the benefit minus cost approach used in conjunction with various multi-criteria decision analysis methods may affect the choice of the best alternative when the criteria are grouped into the two classes as described earlier. These two approaches may also affect the ranking of all the alternatives. The next sub-section describes how these two approaches were combined with the four MCDA methods considered in this paper. After that, we present some numerical examples which demonstrate that the ranking of the alternatives may depend on which approach is combined with these MCDA methods.

3.1 Implementation of the Two Approaches for Aggregating Conflicting Criteria

In the following treatments we assume that an MCDA problem is defined both on benefit and cost criteria. As a convention, criteria C_1, ..., C_{n1} are assumed to be the benefit criteria, while criteria C_{n1+1}, ..., C_n (i.e., the remaining ones) are assumed to be the cost criteria (where $1 < n_1 < n$). Furthermore, now it will be assumed that the criteria weights have been normalized as follows:

$$\sum_{j=1}^{n_1} w_j = 1, \text{ and } \sum_{j=n_1+1}^{n} w_j = 1. \qquad (3\text{-}1)$$

When the WSM method is used with the benefit to cost approach, the alternatives are ranked according to their performance values P_i (for i = 1, 2, 3, ..., m). These *"Ratio"* performance values are computed according to an extension of the previous formula (2-1) as follows:

$$P_{i,WSM,Ratio} = \frac{\sum_{j=1}^{n_1} a_{ij} w_j}{\sum_{j=n_1+1}^{n} a_{ij} w_j}, \qquad \text{for } i = 1, 2, 3, ..., m. \qquad (3\text{-}2)$$

Similarly, under the benefit minus cost criteria, the previous preference values will be computed now as *"Difference"* performance values as follows:

$$P_{i,WSM,Diff} = \sum_{j=1}^{n_1} a_{ij} w_j - \sum_{j=n_1+1}^{n} a_{ij} w_j, \qquad \text{for } i = 1, 2, 3, ..., m. \qquad (3\text{-}3)$$

When the AHP (or its revised additive version as described in Section 2.3) is used, then the previous two formulas are used but now the a_{ij} elements are normalized accordingly.

The case of the WPM or the multiplicative AHP presents some special interest. When the previous formulas are combined with the WPM or the multiplicative AHP main formula given as (2-3) or (2-4), the following are derived:

For the benefit to cost criteria approach:

$$R\left(\frac{A_K}{A_L}\right) = \frac{\prod_{i=1}^{n1} a_{Ki}^{w_i}}{\prod_{i=n1+1}^{n} a_{Ki}^{w_i}} \times \frac{\prod_{j=n1+1}^{n} a_{Lj}^{w_j}}{\prod_{j=1}^{n1} a_{Lj}^{w_j}}. \qquad (3\text{-}4)$$

Similarly, for the benefit minus cost criteria approach:

$$D\left(\frac{A_K}{A_L}\right) = \prod_{i=1}^{n_1} a_{Ki}^{w_i} \times \prod_{j=n_1+1}^{n} a_{Lj}^{w_j} - \prod_{i=n_1+1}^{n} a_{Ki}^{w_i} \times \prod_{j=1}^{n_1} a_{Lj}^{w_j}. \qquad (3\text{-}5)$$

It should be stated here that the last formula expresses the difference (and hence the notation $D(A_K/A_L)$)

between the performance measures of the two alternatives A_K and A_L under the WPM or the multiplicative AHP method. From the last two formulas it can easily be seen that when alternative A_K is ranked better than alternative A_L according to formula (3-4), then the same will be true according to formula (3-5) and vice-versa. This is true because if in formula (3-4) the numerator is greater than the denominator, then the previous ratio will be greater than one and the previous difference will be positive and vice-versa. This is summarized next as Theorem 1.

Theorem 1:

When the WPM or the multiplicative AHP method is used, then the benefit to cost and benefit minus cost approaches always yield the same ranking of the alternatives.

The previous theorem indicates that under the WPM or the multiplicative AHP method the two approaches for dealing with conflicting criteria are perfectly consistent with each other. The numerical examples in the following sub-section demonstrate that this property does not hold true with the WSM and the two additive AHP models described in sub-sections 2.2 and 2.3, respectively.

3.2 Extensive Numerical Examples

The main points in this sub-section are best illustrated in terms of some numerical examples. Thus, five numerical examples are described next. One example for each of the WSM method, the AHP, the revised AHP, the WPM method and the multiplicative AHP.

Example 3.1 (the WSM case):

Let us consider the data depicted in the following decision matrix. In these data it is assumed that the first two criteria are benefit criteria, while the last two are cost ones.

	Criteria			
	Benefit		**Cost**	
	C_1	C_2	C_3	C_4
Alts.	(8/13	5/13)	(6/13	7/13)
A_1	96	83	75	7
A_2	63	5	56	9
A_3	72	30	32	48
A_4	11	4	27	9
A_5	77	21	17	11

Next, we use the benefit to cost approach with the WSM method. Thus, formula (3-2) is applied as follows:

$$P_{1,WSM} = \frac{96 \ x \ (8/13) + 83 \ x \ (5/13)}{75 \ x \ (6/13) + 7 \ x \ (7/13)} = 2.37 \ .$$

Similarly, working as above we get:

$P_{2,\ WSM} = 1.32$, $P_{3,\ WSM} = 1.38$, $P_{4,\ WSM} = 0.48$, and $P_{5,\ WSM} = 4.03$.

However, when using the **benefit minus cost approach** (i.e., formula (3-3)) the corresponding preference values become:

$P_{1,WSM}$ = {96H(8/13) + 83H(5/13)} - {75H(6/13) + 7H(7/13)} = 52.62.

Similarly, working as above we get:

$P_{2,\ WSM} = 10.00$, $P_{3,\ WSM} = 15.23$, $P_{4,\ WSM} = -9.00$, and $P_{5,\ WSM} = 41.70$.

The previous results are next used to derive the rankings of the five alternatives which are summarized as follows:

Alternative	Ranking Under the Benefit to Cost Ratio	Ranking Under the Benefit Minus Cost
A_1	2	1
A_2	4	4
A_3	3	3
A_4	5	5
A_5	1	2
	or: $A_5 > A_1 > A_3 > A_2 > A_4$	or: $A_1 > A_5 > A_3 > A_2 > A_4$

That is, the two rankings differ by the indication of the best alternative (only).

Example 3.2 (the AHP case):

The data for this example are different than the ones for the previous example in order to better illustrate the main points of interest. If the same data were used throughout these examples, the magnitude of the ranking differences may had not been so dramatic. These data are presented in the following decision matrix:

	Criteria			
	Benefit		**Cost**	
	C_1	C_2	C_3	C_4
Alts.	(8/13	5/13)	(6/13	7/13)
A_1	7/31	8/21	7/27	9/33
A_2	4/31	2/21	5/27	9/33
A_3	6/31	4/21	5/27	5/33
A_4	8/31	1/21	8/27	3/33
A_5	6/31	6/21	7/27	7/33

The above data have been normalized according to the standard AHP practice. That is, the entries in each column, and the criteria weights within each of the two groups, add up to one. Next the two aggregation criteria approaches are applied as before.

As earlier, the benefit to cost ratio approach is applied and it yields the following results:

$$P_{1,\,AHP} = \frac{7/31\ x\ (8\ /\ 13) + 8/21\ x\ (5\ /\ 13)}{7/27\ x\ (6\ /\ 13) + 9/33\ x\ (7\ /\ 13)} = 1.15\ .$$

Similarly, working as above we get:

$P_{2,\,AHP} = 0.53$, $P_{3,\,AHP} = 1.25$, $P_{4,\,AHP} = 1.08$, and $P_{5,\,AHP} = 1.06$.

The benefit minus cost approach yields:

$P_{1,\,AHP}$ = {7/31 H 8/13 + 8/21 H 5/13} - {7/27 H 6/13 + 9/33 H 7/13} = 0.53.

Similarly, working as above we get:

$P_{2,\,AHP} = 0.33$, $P_{3,\,AHP} = 0.35$, $P_{4,\,AHP} = 0.34$, and $P_{5,\,AHP} = 0.44$.

The previous results are next used to derive the rankings of the five alternatives which are summarized as follows:

Alternative	Ranking: Under the Benefit to Cost Ratio	Ranking: Under the Benefit Minus Cost
A_1	2	1
A_2	5	5
A_3	1	3
A_4	3	4
A_5	4	2
	or: $A_3 > A_1 > A_4 > A_5 > A_2$	or: $A_1 > A_5 > A_3 > A_4 > A_2$

It is remarkable to observe that in this illustrative example there is not even any resemblance in the rankings derived by using both approaches except that both approaches rank A_2 as the worst alternative.

Example 3.3 (the Revised AHP Case):

For the revised AHP case we also use a new data set as presented in the following decision matrix:

	Criteria: Benefit		Criteria: Cost	
	C_1	C_2	C_3	C_4
Alts.	(8/13	5/13)	(6/13	7/13)
A_1	4/7	1	1	2/7
A_2	4/7	2/8	4/9	2/7
A_3	1	3/8	1	1
A_4	4/7	1	7/9	2/7

The previous data have been normalized according to the revised AHP requirement. That is, in each column each entry is divided by the largest element, while the criteria weights are normalized as in the previous example.

When the benefit to cost ratio approach is applied as before, the previous data yield the following preference values:

$$P_{1,\,RAHP} = \frac{4/7 \text{ x } (8/13) + 1 \text{ x } (5/13)}{1 \text{ x } (6/13) + 2/7 \text{ x } (7/13)} = 1.20\,.$$

Similarly, working as above we get:

$P_{2,RAHP} = 1.25$, $P_{3,\,RAHP} = 0.76$, and $P_{4,\,RAHP} = 1.44$.

The benefit minus cost approach yields the following results:

$P_{1,\,RAHP} = \{4/7 \text{ H } 8/13 + 1 \text{ H } 5/13\} - \{1 \text{ H } 6/13 + 2/7 \text{ H } 7/13\} = 1.35$.

Similarly, working as above we get:

$P_{2,\,RAHP} = 0.81$, $P_{3,\,RAHP} = 1.76$, and $P_{4,\,RAHP} = 1.25$.

As before, the above results can be summarized as follows:

	Rankings	
Alternative	Under the Benefit to Cost Ratio	Under the Benefit Minus Cost
A_1	3	2
A_2	2	4
A_3	4	1
A_4	1	3
	or: $A_4 > A_2 > A_1 > A_3$	or: $A_3 > A_1 > A_4 > A_2$

The rankings of the alternatives, as derived by using the two approaches, are almost the complete opposite! The best two alternatives in one approach are the worst two alternatives in the other.

Next we observe that according to Theorem 1, the WPM and the multiplicative AHP do not exhibit any ranking contradictions when these two conflicting criteria aggregation approaches are used. However, for the sake of completeness of the presentation two numerical examples are briefly described when the WPM and the multiplicative AHP are used. For easier comparison, the data for these examples are now the same as the ones for Example 3.2.

Example 3.4 (the WPM case):

Under the WPM method one does not have to normalize the data. As a matter of fact, it can be easily seen (for instance, by comparing the computations between Examples 3.4 and 3.5) that the end results are essentially the same whether the data are normalized or not. Thus, one can view the data to be as follows (i.e., before normalization):

	Criteria			
	Benefit		**Cost**	
	C_1	C_2	C_3	C_4
	(8/13	5/13)	(6/13	7/13)
Alts.				
A_1	7	8	7	9
A_2	4	2	5	9
A_3	6	4	5	5
A_4	8	1	8	3
A_5	6	6	7	7

When one uses the benefit to cost approach with the WPM method, then formula (3-4) yields for the pair of the alternatives A_1 and A_2:

$$R\left(\frac{A_1}{A_2}\right) = \frac{7^{8/13} \quad 8^{5/13}}{7^{6/13} \quad 9^{7/13}} \times \frac{5^{6/13} \quad 9^{7/13}}{4^{8/13} \quad 2^{5/13}} = 2.059 > 1, \quad \text{or } A_1 > A_2.$$

Similarly, working as above we get:

$$R\left(\frac{A_1}{A_3}\right) = 0.896 < 1,\ R\left(\frac{A_1}{A_4}\right) = 1.206 > 1,\ R\left(\frac{A_1}{A_5}\right) = 1.073 > 1,$$

$$R\left(\frac{A_2}{A_3}\right) = 0.435 < 1,\ R\left(\frac{A_2}{A_4}\right) = 0.586 < 1,\ R\left(\frac{A_2}{A_5}\right) = 0.521 < 1,$$

$$R\left(\frac{A_3}{A_4}\right) = 1.347 > 1,\ R\left(\frac{A_3}{A_5}\right) = 1.198 > 1,\ R\left(\frac{A_4}{A_5}\right) = 0.889 < 1.$$

From the above results, it can be easily derived that the implied ranking is as follows:

$A_3 > A_1 > A_5 > A_4 > A_2$.

Next we apply the benefit minus cost approach (by using formula (3-5)) as follows:

$D\left(\frac{A_1}{A_2}\right) = \{(7^{8/13} \quad 8^{5/13}) \times (5^{6/13} \quad 9^{7/13})\} - \{(7^{6/13} \quad 9^{7/13})(4^{8/13} \quad 2^{5/13})\} = 26.007 > 1$, Similarly, working as above we get:

$$D\left(\frac{A_1}{A_3}\right) = -4.298 < 0,\ D\left(\frac{A_1}{A_4}\right) = 5.949 > 0,\ D\left(\frac{A_1}{A_5}\right) = 3.496 > 0,$$

$$D\left(\frac{A_2}{A_3}\right) = -19.905 < 0,\ D\left(\frac{A_2}{A_4}\right) = -10.216 < 0,\ D\left(\frac{A_2}{A_5}\right) = -19.722 < 0,$$

$$D\left(\frac{A_3}{A_4}\right) = 6.241 > 0,\ D\left(\frac{A_3}{A_5}\right) = 5.935 > 0,\ D\left(\frac{A_4}{A_5}\right) = -3.138 < 0.$$

From the above results it can easily be derived that the implied ranking is identical with the previous one (i.e., $A_3 > A_1 > A_5 > A_4 > A_2$) as it should be according to Theorem 1 in Section 3.1 but different than the AHP results of the same data in Example 3.2.

Example 3.5 (the Multiplicative AHP case):

The application of the multiplicative AHP is very similar to the case of the WPM method. The only difference is that now the entries of the decision matrix are normalized by dividing the elements in each column by the largest element of that column. Thus, the decision matrix becomes as follows:

	C r i t e r i a			
	Benefit		**Cost**	
	C_1	C_2	C_3	C_4
	(8/13	5/13)	(6/13	7/13)
Alts.				
A_1	7/8	8/8	7/8	9/9
A_2	4/8	2/8	5/8	9/9
A_3	6/8	4/8	5/8	5/9
A_4	8/8	1/8	8/8	3/9
A_5	6/8	6/8	7/8	7/9

When one uses the benefit to cost approach with the multiplicative AHP method, then formula (3-4) yields for the pair of the A_1 and A_2 alternatives:

$$R\left(\frac{A_1}{A_2}\right) = \frac{(7/8)^{8/13} \quad (8/8)^{5/13}}{(7/8)^{6/13} \quad (9/9)^{7/13}} \times \frac{(5/8)^{6/13} \quad (9/9)^{7/13}}{(4/8)^{8/13} \quad (2/8)^{5/13}} = 2.059 > 1,$$

One may observe that it should had been expected that the WPM and the multiplicative AHP would produce identical results when the ratio formula is used. This occurs because the results for the WPM case and also for the multiplicative AHP case are divided by a constant that cancels out in the ratio formulas. Similarly with the above, the rest of the $R(A_i/A_j)$ values are identical to those in the example for the WPM method.

When we apply the benefit minus cost approach (i.e., by using formula (3-5)) the multiplicative AHP also yields results identical (as far as the positive/negative sign is concerned) to those for the WPM method. This occurs because now the results for the multiplicative AHP method are exactly the same as those for the WPM case multiplied by a positive quantity which is due to the presence of the normalization step under the multiplicative AHP method. For this numerical example this positive quantity is equal to: $(1/8)^{8/13}(1/8)^{5/13} \times (1/8)^{6/13} (1/9)^{7/13}$.

The startling variations which are exhibited in the final results in Examples 3.1, 3.2, and 3.3 lead one next to study this phenomenon in more depth in terms of an extensive empirical analysis as described in the following section. This empirical analysis studies how often such ranking discrepancies occur in simulated MCDA problems.

Table I (part *a*): *Computational Results for the WSM Model.*

Size of Problem		Differences in Rankings					Size of Problem		Differences in Rankings				
Alts.	Crit.	Rate 1	Rate 2	Rate 3	Rate 4	Rate 5	Alts.	Crit.	Rate 1	Rate 2	Rate 3	Rate 4	Rate 5
3	3	6.02	7.57	7.73	11.2	5.53	3	9	3.42	4.51	4.59	6.66	3.09
5	3	10.45	14.01	14.97	31.68	6.17	5	9	5.7	8.4	8.73	19.85	3.39
7	3	13.84	19.8	21.97	54.76	6.39	7	9	8.22	12.07	12.9	36.59	3.68
9	3	16.91	24.62	28.52	73.16	6.55	9	9	9.96	15.31	16.67	53.2	3.78
11	3	20.38	29.51	35.6	86.42	6.49	11	9	11.89	18.69	20.93	68.6	3.87
13	3	22.43	33.56	42.07	93.58	6.52	13	9	13.62	21.82	24.98	80.94	3.88
15	3	23.83	37.31	48.53	97.39	6.53	15	9	14.64	24.74	29.04	89.05	3.90
17	3	26.11	40.61	55.1	98.93	6.46	17	9	15.65	27.38	32.63	93.91	3.87
19	3	27.63	43.84	61.49	99.68	6.52	19	9	17.66	30	36.79	96.91	3.87
21	3	28.24	46.46	67.35	99.87	6.53	21	9	18.15	32.22	40.33	98.59	3.86
3	5	4.56	6.06	6.15	9	4.37	3	11	2.76	3.76	3.79	5.62	2.92
5	5	7.93	11.19	11.74	25.9	4.74	5	11	5.79	7.7	7.94	18.28	3.24
7	5	10.43	15.87	17.25	46.18	4.88	7	11	7.44	11.03	11.62	34.04	3.38
9	5	13.57	20.13	22.58	64.16	5.07	9	11	9.26	13.83	14.97	49.14	3.51
11	5	15.77	23.98	27.86	78.43	5.14	11	11	11.04	17.13	18.94	64.95	3.52
13	5	18.53	27.87	33.38	88.68	5.14	13	11	12.32	20.07	22.64	77.47	3.50
15	5	19.53	30.79	37.96	94.33	5.10	15	11	14.08	22.82	26.11	86.39	3.54
17	5	21	34.01	43.15	97.52	5.16	17	11	14.61	25.15	29.62	92.18	3.52
19	5	22.41	36.97	48.29	99.01	5.12	19	11	15.6	27.48	32.96	96.03	3.51
21	5	23.6	39.64	53.34	99.66	5.09	21	11	17.03	30.09	37.12	98.01	3.51
3	7	3.71	4.88	4.98	7.23	3.85	3	13	2.63	3.51	3.55	5.24	2.80
5	7	6.78	9.37	9.78	22.03	4.14	5	13	4.59	6.93	7.1	16.66	2.96
7	7	9.23	13.47	14.46	40.51	4.27	7	13	6.75	9.96	10.45	31.14	3.12
9	7	11.4	17.15	18.97	57.9	4.28	9	13	9.13	13.51	14.55	48.48	3.08
11	7	13.9	20.79	23.61	72.8	4.34	11	13	9.66	15.86	17.42	62.34	3.21
13	7	15.24	23.82	27.82	83.82	4.38	13	13	11.16	18.9	21.15	74.98	3.21
15	7	17	27.36	32.74	91.69	4.42	15	13	12.25	21.2	24.25	84.14	3.23
17	7	18.33	30.31	36.94	95.88	4.42	17	13	13.29	23.56	27.46	90.6	3.22
19	7	19.12	32.71	41.11	98.04	4.34	19	13	14.54	25.99	30.88	95	3.26
21	7	20.04	35.37	45.66	99.22	4.34	21	13	15.37	27.78	33.54	97.28	3.23

Table I (part *b*): *Computational Results for the WSM Model.*

Size of Problem		Differences in Rankings					Size of Problem		Differences in Rankings				
Alts.	Crit.	Rate 1	Rate 2	Rate 3	Rate 4	Rate 5	Alts.	Crit.	Rate 1	Rate 2	Rate 3	Rate 4	Rate 5
3	15	2.6	3.45	3.48	5.13	2.48	3	19	2.07	2.8	2.83	4.17	2.23
5	15	4.8	6.78	7	16.22	2.73	5	19	4.15	5.82	5.97	14.01	2.37
7	15	6.25	9.42	9.86	29.61	2.91	7	19	5.48	8.27	8.61	26.09	2.63
9	15	7.62	11.9	12.68	43.65	2.99	9	19	7.55	11.07	11.78	41.28	2.60
11	15	9.61	15.05	16.38	59.96	2.95	11	19	8.57	13.53	14.63	56.03	2.65
13	15	10.84	17.24	19.22	71.98	3.01	13	19	10.04	15.86	17.42	68.39	2.69
15	15	11.93	19.89	22.61	81.9	3.00	15	19	10.42	18.06	20.16	78.66	2.71
17	15	12.33	22.46	25.92	89.49	2.99	17	19	10.88	20.12	22.78	86.39	2.71
19	15	13.12	24.36	28.49	93.81	3.03	19	19	11.55	22.35	25.82	91.69	2.70
21	15	14.69	26.61	31.81	96.84	3.02	21	19	12.12	24.05	28.24	94.96	2.70
3	17	2.31	3.05	3.07	4.54	2.36	3	21	2.43	2.96	3	4.41	2.02
5	17	4.04	6.12	6.23	14.7	2.64	5	21	3.73	5.44	5.58	13.05	2.38
7	17	6.29	9.12	9.53	28.65	2.69	7	21	5.41	8.26	8.62	26.1	2.44
9	17	7.25	11.29	12.04	42.06	2.80	9	21	6.66	10.64	11.25	40.07	2.45
11	17	8.9	14.09	15.29	56.94	2.84	11	21	7.81	13.01	13.95	54.3	2.55
13	17	9.99	16.56	18.36	69.8	2.87	13	21	9.66	15.31	16.67	67.27	2.52
15	17	10.88	19.08	21.47	80.66	2.88	15	21	9.82	17.26	19.13	76.2	2.55
17	17	12.31	21.2	24.21	87.35	2.84	17	21	11.11	19.2	21.63	84.07	2.58
19	17	12.56	23.19	26.95	92.39	2.87	19	21	12.06	21.56	24.71	90.51	2.57
21	17	13.22	25.11	29.6	95.69	2.84	21	21	11.84	23.16	26.88	94.26	2.57

4 Computational Experiments

A computer program was written in FORTRAN in order to generate simulated decision problems. As data we used random numbers in the range 9 to 1 in order to be consistent with the basic scale used by AHP which assumes any two performance values to be in the interval [9, 1/9]. The number of alternatives was equal to 3, 5, 7, ..., 21. The number of the benefit and cost criteria (within each group) was also equal to 3, 5, 7, ..., 21. That is we consider the problems to have the same number of benefit and cost criteria. We did not consider cases with different numbers of benefit and cost criteria because that would have dramatically increased the number of the test problem cases by introducing more problem parameters. However, some pilot tests indicate that the contradiction rates (which as described in the following section are rather high) would be even higher when the numbers of cost and benefit criteria are not equal with each other.

With our experimental design approach we captured a rather wide range of sizes of MCDA problems. In some actual problems sizes might be outside this range (i.e., of higher value since the number of criteria cannot be less than 3), but it is believed that the trends revealed in this extensive simulation study will hold true as well. These ranges have also been used in all the computational experiments run by the first author and his research associates (see, for instance, [Triantaphyllou, Lootsma, Pardalos, and Mann, 1994], [Triantaphyllou, and Mann, 1994*a*; and 1994*b*], and [Triantaphyllou, 2000]). Thus each test problem had the same number of benefit and cost criteria. Therefore, the total number of criteria was equal to 6, 10, 14, ..., 42. Each test problem was examined by applying the WSM, the AHP, and the revised AHP as was the case in the illustrative examples described in the previous section. Since the multiplicative AHP is perfectly consistent under both criteria aggregation approaches, this method was not considered in this empirical study.

Since we had ten cases of different numbers of criteria and alternatives, we considered 100 (i.e., 10x10) different combinations. For each such combination we run 10,000 random replications and each problem was solved by using the previous three MCDA methods as in the previous illustrative examples. The sample size of 10,000 was large enough to ensure statistically significant results. Some of the computational results are presented as Table I (parts *a* and *b*). The rest of the computation results are available to interested readers from the first author.

For each test problem, the two rankings derived by using the benefit to cost and benefit minus cost approach were analyzed in five different ways. The first way was to see whether the two rankings agreed in the indication of the best alternative (since in many MCDA problems the interest is in the identification of the best alternative only). The percentage of times the two approaches yielded a different indication of the best alternative is denoted as *"Rate 1"* in Table I (parts *a* and *b*). For instance, in Table I (part *a*) when the WSM method is used and the number of criteria and alternatives are equal to 3 and 13, respectively, *"Rate 1"* is equal to 2.63. This means that 2.63% of the test problems with 3 cost and 3 benefit criteria and 13 alternatives resulted in a different indication of the best alternative when the two ranking approaches were combined with the WSM method.

The second way of comparing the two rankings was the number of different rankings divided by the corresponding number of alternatives. This is indicated as *"Rate 2"* in Table I. As with the previous case, for test problems with the number of cost (or benefit) criteria and alternatives equal to 3 and 13, respectively, *"Rate 2"* is equal to 3.51. This means that, on the average, the two rankings were different by 3.51 when the difference was measured as the number of cases the two rankings were different from each other and then by dividing by the number of alternatives. For instance, suppose that five alternatives were evaluated and one ranking was equal to (1, 2, 3, 4, 5) while the second ranking was equal to (1, 2, 5, 4, 3). Then the value recorded was equal to: (1 + 1) / 5 = 0.40 (since these two rankings differ in two ranks only).

The third way was to calculate the sum of the absolute differences of the two rankings. This was recorded as *"Rate 3"* in Table I. For instance, for the previous illustrative example with the two rankings of (1, 2, 3, 4, 5) and (1, 2, 5, 4, 3), the corresponding value is: $(0 + 0 + |3-5| + 0 + |5-3|) / 5 = 0.80$.

The fourth way (denoted as *"Rate 4"* in Table I) was to record the number of times the two rankings were different from each other, without considering the magnitude of the individual differences. That is, when two rankings are evaluated, this rate would be equal to 0 if the two rankings are identical or equal to 1, otherwise (i.e., it is binary valued). For instance, for the pair of rankings (1,2,3,4,5) and (1,2,5,4,3) the rate is equal to 1. Similarly with above, for test problems with 3 cost and 3 benefit criteria and 13 alternatives this rate is equal to 5.24% when the WSM method is used.

The fifth way is to consider a weighted measure for expressing differences in ranking discrepancies. In this paper we call it *"Rate 5"*. According to this measure one may wish to assign more significance to discrepancies on top rankings and less significance to discrepancies on lower rankings. Although one may use any kind of weights to achieve the previous intentions, in [Ray and Triantaphyllou, 1998] it was suggested to use the weights (n, n-1, n-2, …, 2, 1). For instance, if the two rankings are (3, 2, 1, 5, 4) and (1, 2, 3, 4, 5) then the weighted absolute differences are: 5|3-1| + 4|2-2| + 3|1-3| + 2|5-4| + 1|4-5| = 20. That is, the following formula is used to express this weighted difference of two rankings denoted as $Rank^{(1)}$ and $Rank^{(2)}$:

$$\sum_{i=1}^{n}(n+1-i)\times|\,Rank_i^{(1)}-Rank_i^{(2)}\,|\,.$$

Next one may wish to normalize this sum with the largest possible value it can have. The largest such value is when the two ranking vectors are opposite of each other. Thus, the maximum sum is equal to:

$$\sum_{i=1}^{n}(n+1-i)\times|\,i-(n+1-i)\,|\,.$$

For the previous example that maximum value will be equal to:

5|5-1| + 4|4-2| + 3|3-3| + 2|2-4| + 1|1-5| = 36.

Therefore, the mathematical formula for the fifth rate is the ratio of the previous two formulas:

$$\frac{\sum_{i=1}^{n}(n+1-i)\times|\,Rank_i^{(1)}-Rank_i^{(2)}\,|}{\sum_{i=1}^{n}(n+1-i)\times|\,(i-(n+1-i)\,|}\,.$$

Thus, for this illustrative example the value is equal to 20/36 = 0.556. Finally, it should be stated here that some theoretical results on the domain values the above discrepancy measures may assume, are discussed in [Ray and Triantaphyllou, 1999].

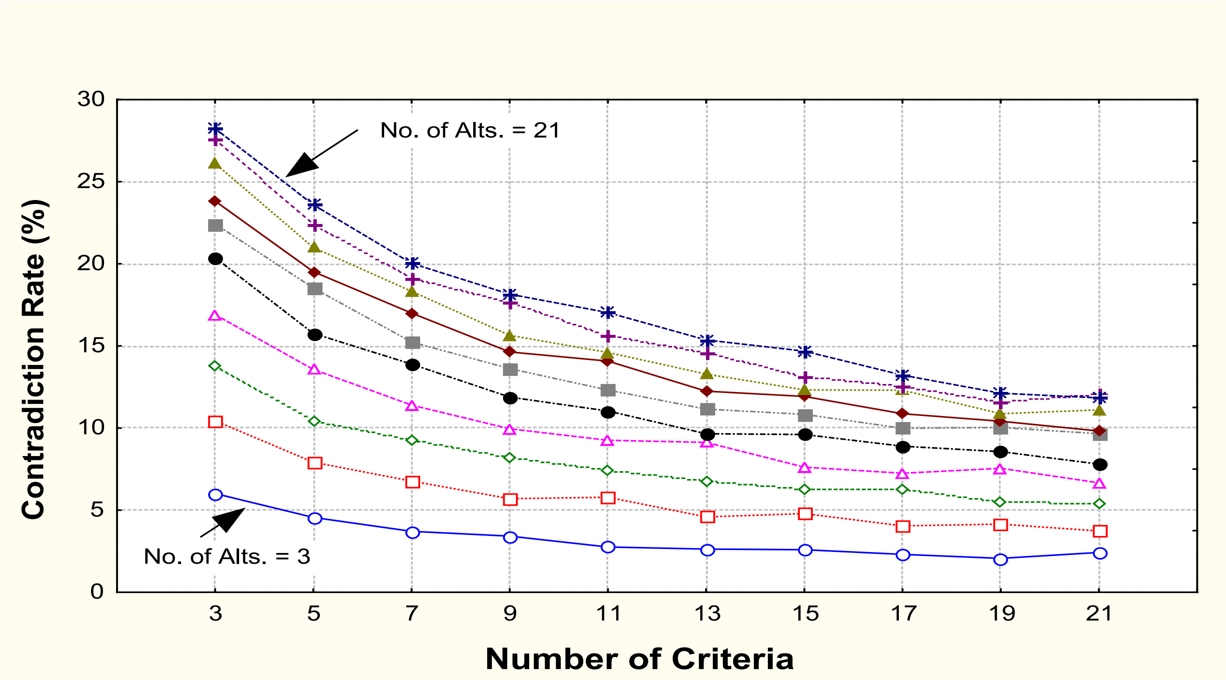

Figure 1: Rate the two rankings point to different alternative as being the best one (i.e., *"Rate 1"*). The WSM case. The different curves represent cases where the number of alternatives is equal to 3, 5, 7, 9, …, 17, 19, and 21 from the bottom and up in that order.

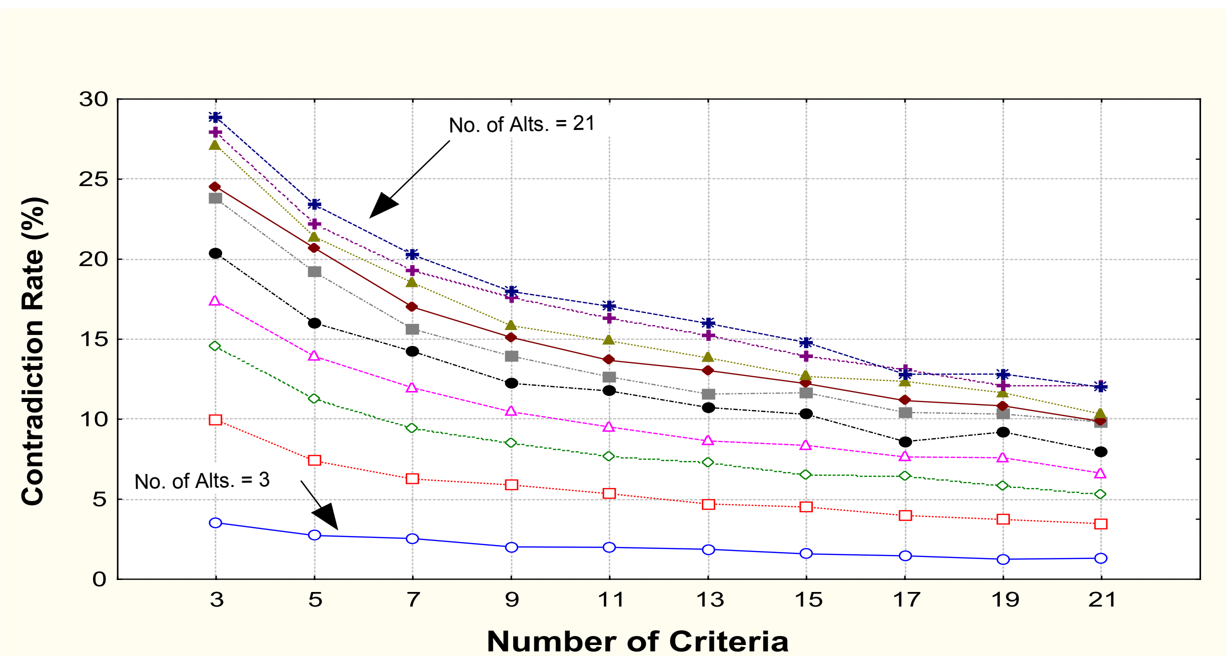

Figure 2: Rate the two rankings point to different alternative as being the best one (i.e., *"Rate 1"*). The AHP case. The different curves represent cases where the number of alternatives is equal to 3, 5, 7, 9, …, 17, 19, and 21 from the bottom and up in that order.

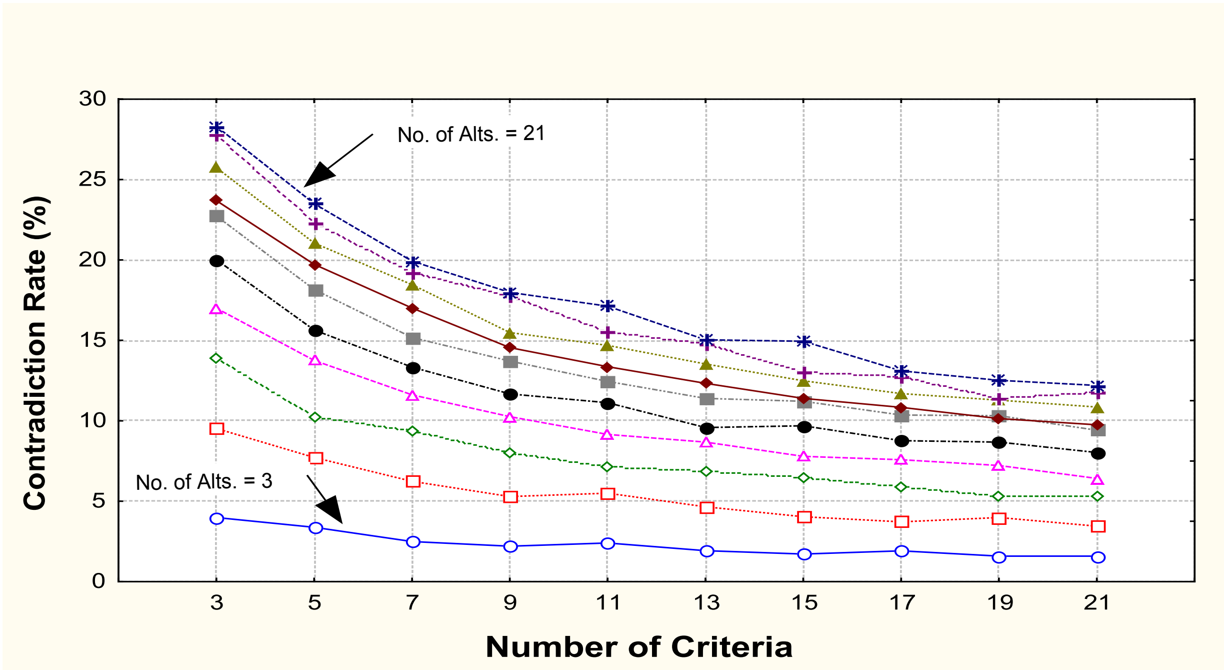

Figure 3: Rate the two rankings point to different alternative as being the best one (i.e., *"Rate 1"*). The revised AHP case. The different curves represent cases where the number of alternatives is equal to 3, 5, 7, 9, …, 17, 19, and 21 from the bottom and up in that order.

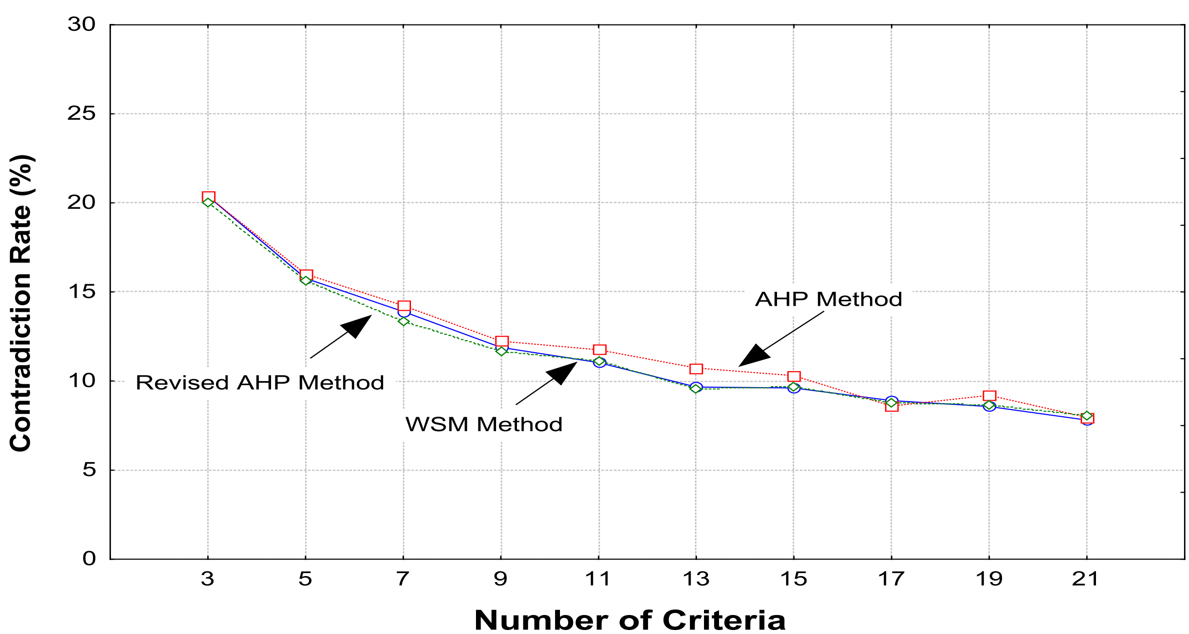

Figure 4: Comparison of the three MCDA Methods in terms of Rate 1 when a problem has 11 alternatives.

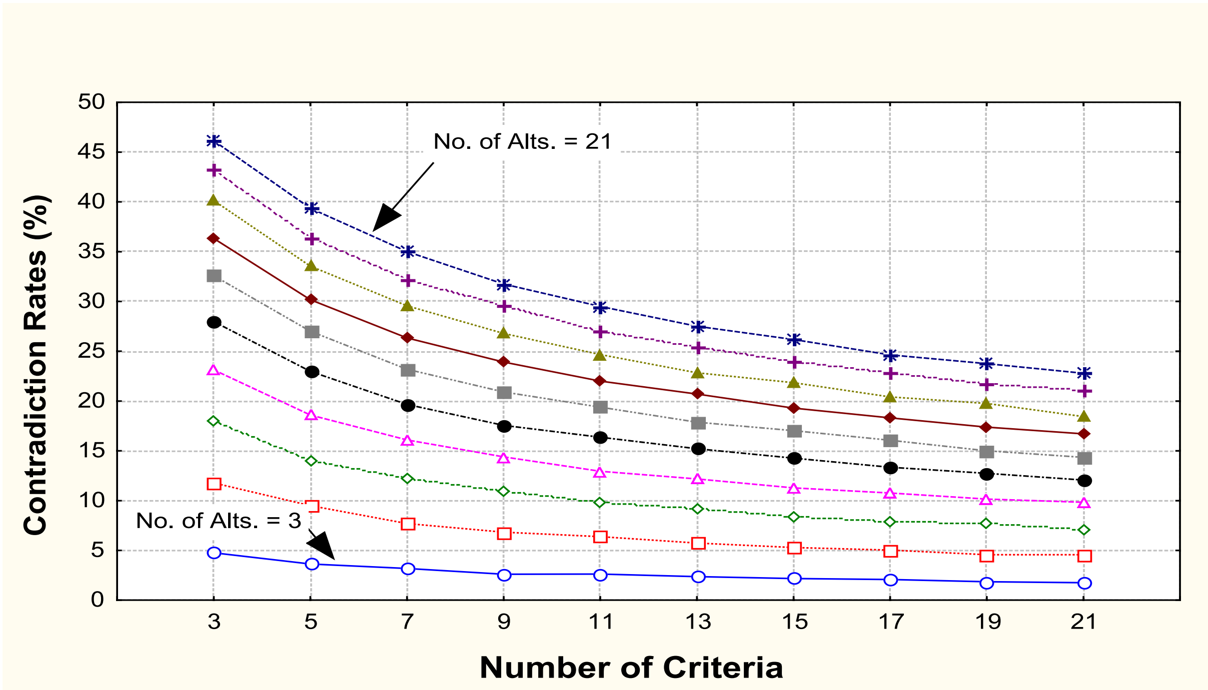

Figure 5: Number of differences in the two rankings (i.e., *"Rate 2"*). The AHP case. The different curves represent cases where the number of alternatives is equal to 3, 5, 7, 9, …, 17, 19, and 21 from the bottom and up in that order.

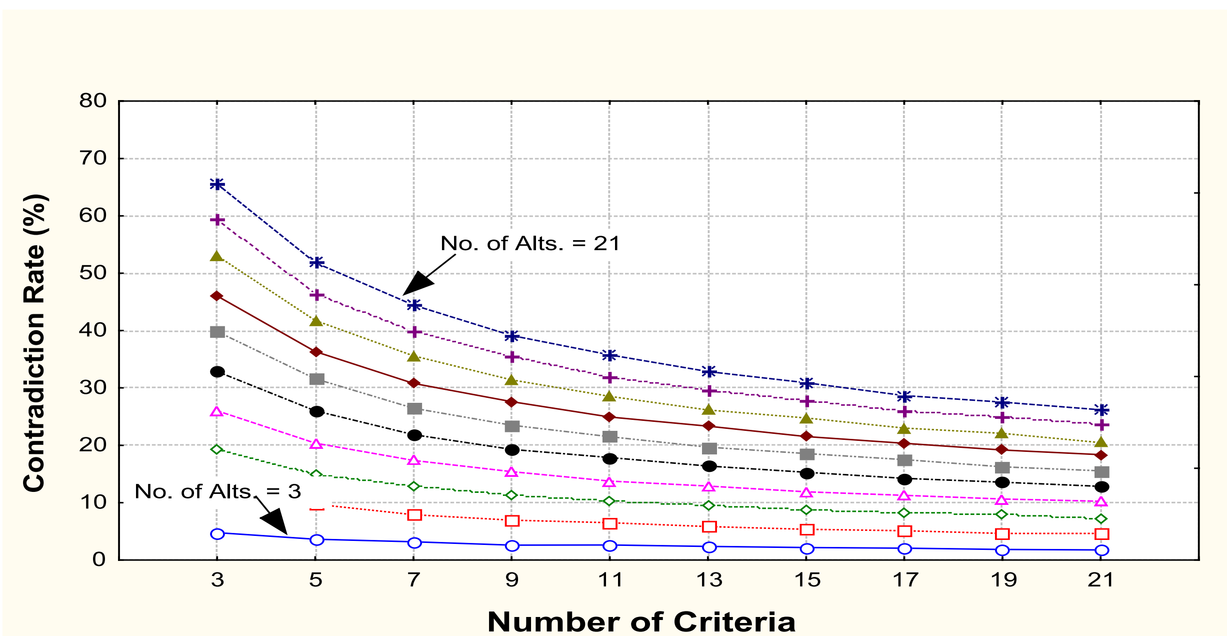

Figure 6: Absolute differences in the two rankings (i.e., *"Rate 3"*). The AHP case. The different curves represent cases where the number of alternatives is equal to 3, 5, 7, 9, …, 17, 19, and 21 from the bottom and up in that order.

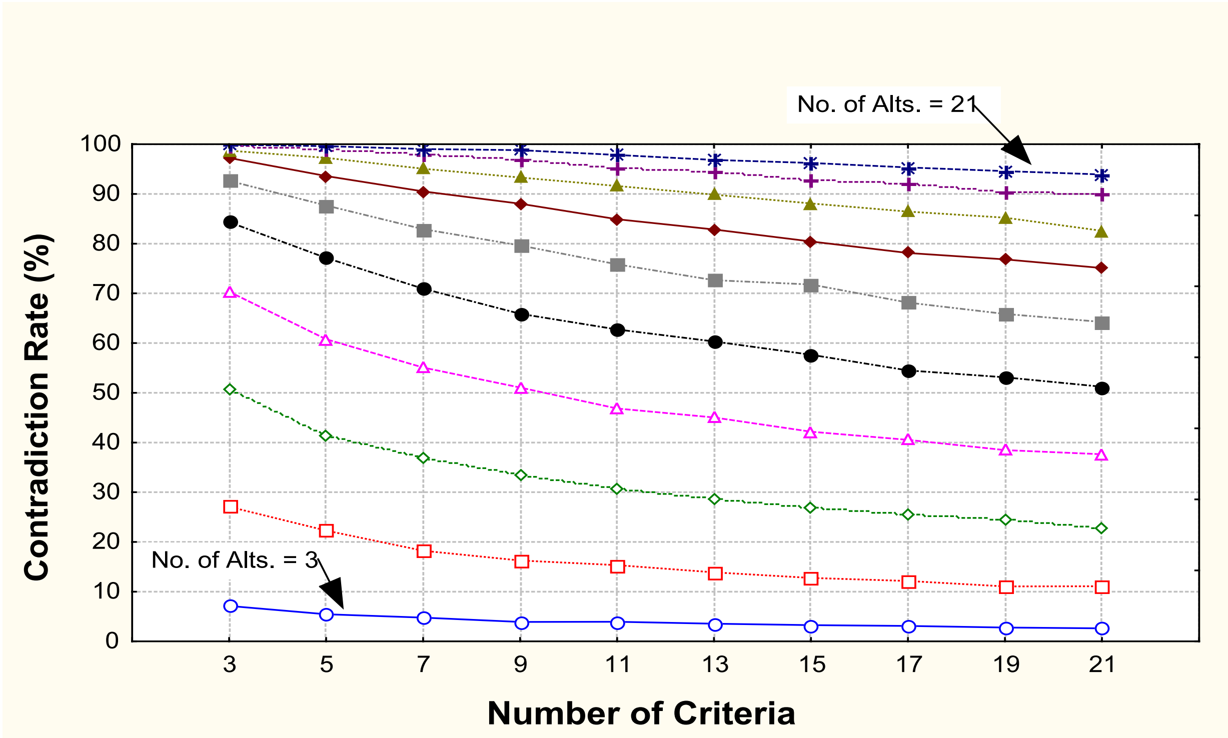

Figure 7: Number of times the two rankings differ with each other (i.e., *"Rate 4"*). The AHP case. The different curves represent cases where the number of alternatives is equal to 3, 5, 7, 9, …, 17, 19, and 21 from the bottom and up in that order.

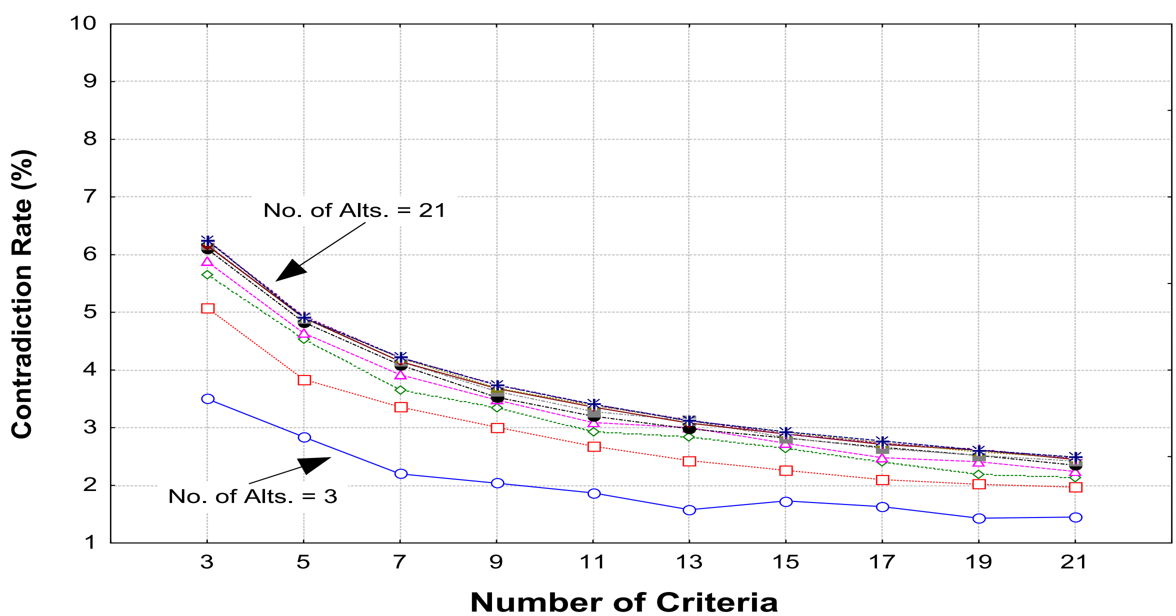

Figure 8: Weighted differences in the two rankings (i.e., *"Rate 5"*). The AHP case. The different curves represent cases where the number of alternatives is equal to 3, 5, 7, 9, …, 17, 19, and 21 from the bottom and up in that order.

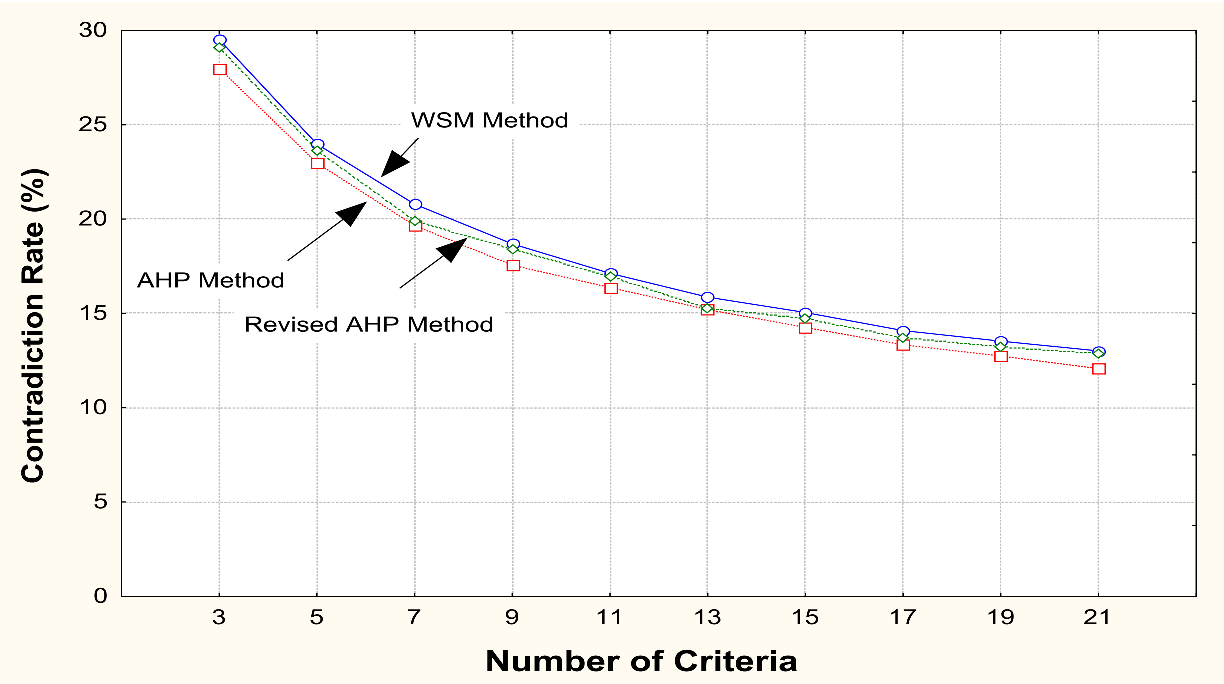

Figure 9: Comparison of the three MCDA Methods in terms of Rate 2 when a problem has 11 alternatives.

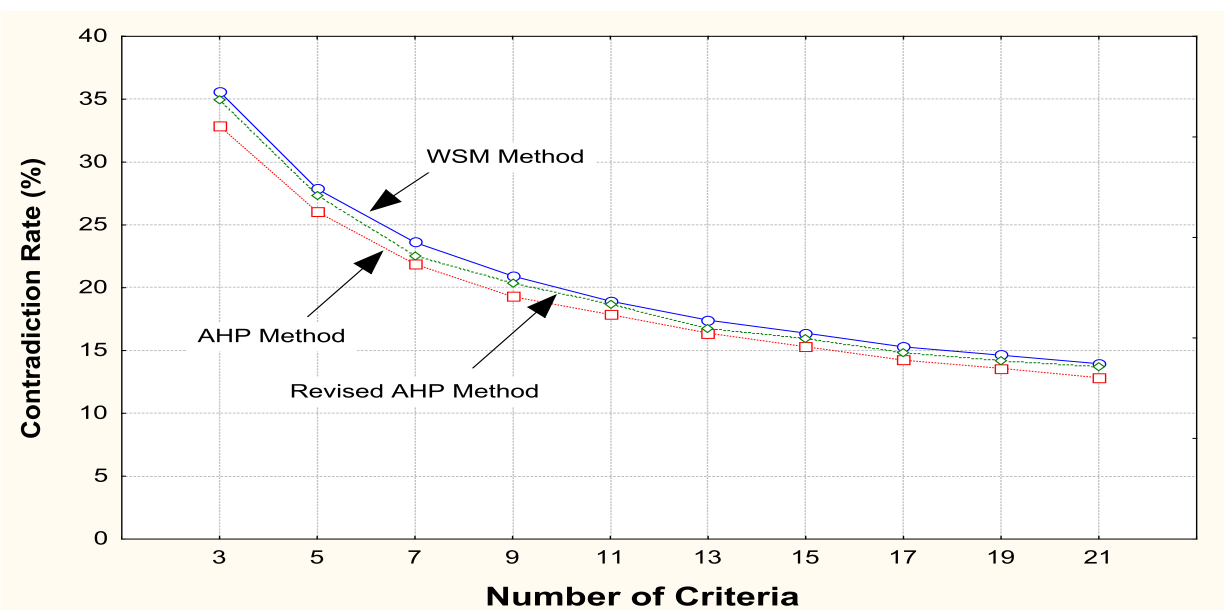

Figure 10: Comparison of the three MCDA Methods in terms of Rate 3 when a problem has 11 alternatives.

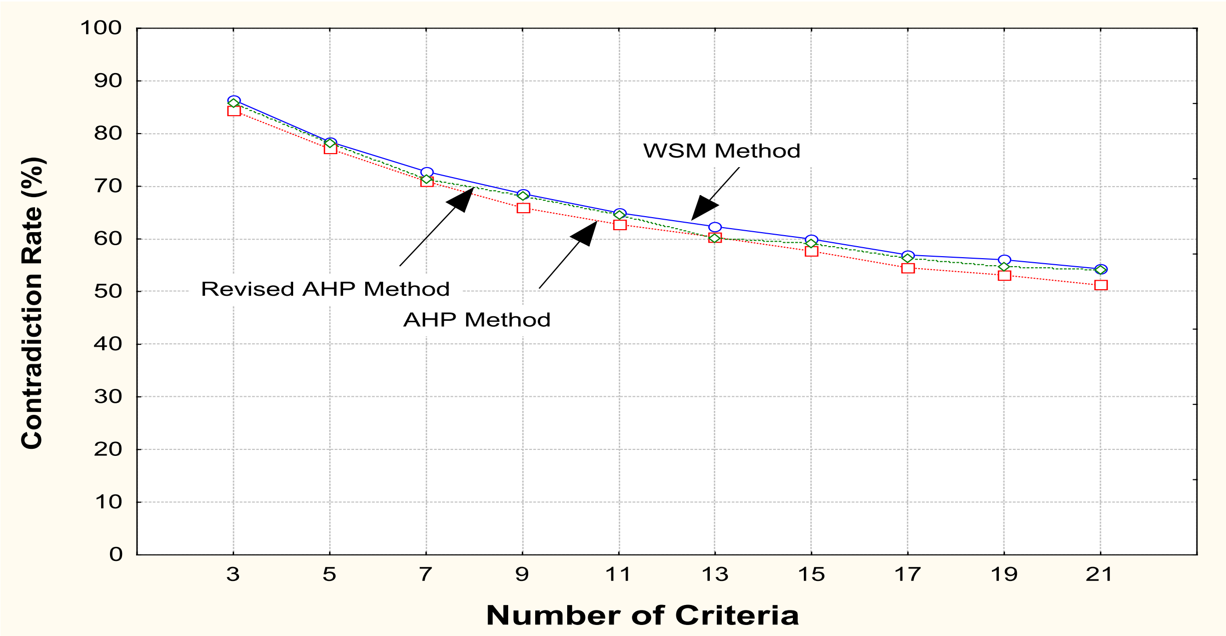

Figure 11: Comparison of the three MCDA Methods in terms of Rate 4 when a problem has 11 alternatives.

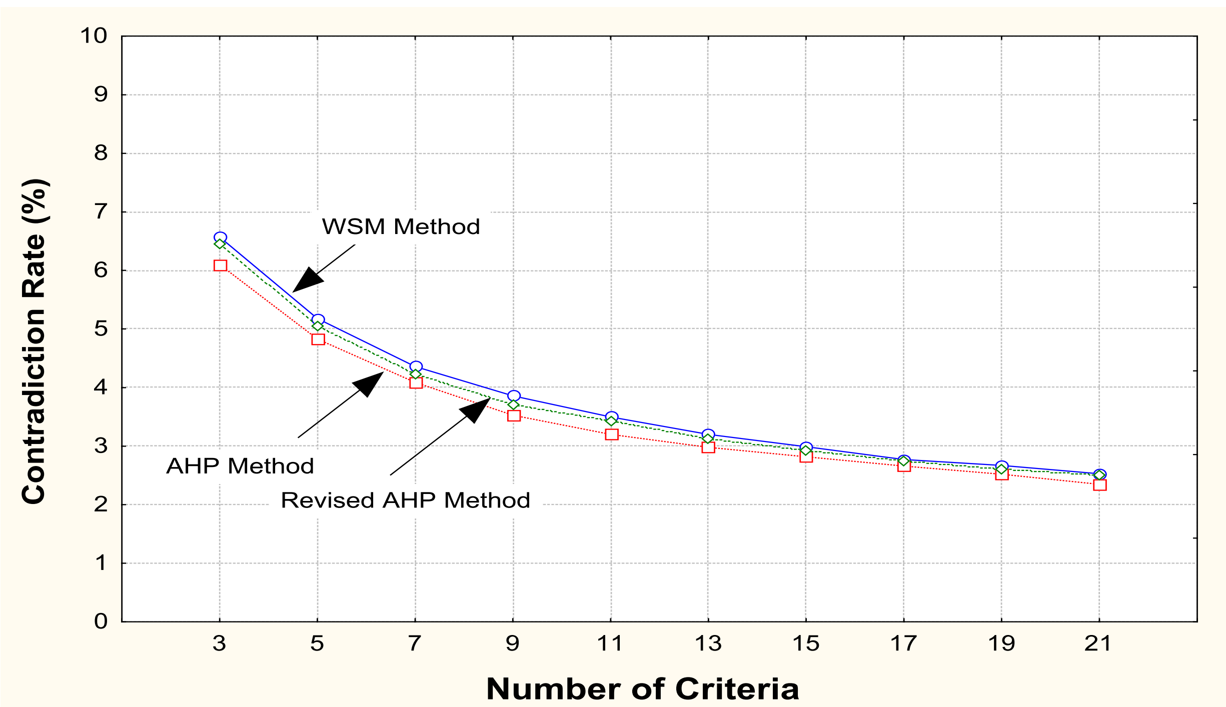

Figure 12: Comparison of the three MCDA Methods in terms of Rate 5 when a problem has 11 alternatives.

5 Analysis of the Computational Results

The computational results also suggest that the three methods (i.e., the WSM, the AHP, and the revised AHP) performed almost identically with each other. This is graphically illustrated in Figures 1, 2, 3, and 4. The patterns exhibited in these three figures are almost the same, although they correspond to different MCDA methods. The results in these four figures describe the contradiction rates when the focus is on the indication of the best alternative (i.e., *"Rate 1"*).

A similar situation holds true for the other remaining four contradiction rates. For reason of simplicity, plots are presented only for the AHP case, since the patterns for the WSM and the revised AHP are very similar with those for the AHP. This is evident in Figure 4 which is representative of all cases. Detailed results are presented as Figures 5, 6, 7, and 8. In all these figures the top curve corresponds to problems with 21 alternatives while each one successively lower curve corresponds to problems with 19, 17, 15, etc. alternatives. Thus, the bottom curve corresponds to problems with 3 alternatives.

A closer examination of these figures reveals a number of interesting points. First, with reference to Rate 1, the following observations can be made (see also Figures 1, 2, 3, and 4). These contradiction rates range from 3% to 30%. The higher the number of the alternatives, the higher is the percentage of different indications of the best alternative. This was anticipated because the more the alternatives are, the more likely is for the ranking procedures to disagree in the indication of the best alternative.

However, the opposite is true with the number of decision criteria in an MCDA problem. This likely happens because with a greater number of criteria and uniformly distributed input data in the interval [9, 1], the best alternatives tend to assume similar values. Thus, the same alternative is likely to be identified as the best when both ranking approaches are used and the number of criteria is high.

The above situation also holds true when differences in the entire ranking (i.e., *"Rate 2"*) are concerned (see also Figure 5). A similar situation also occurs with the results for *"Rate 3"* (see also Figure 6). Only the intensities of the previous contradictions differ.

Please recall that the case of *"Rate 4"* (i.e., the number of times the two ranking approaches yielded different rankings) is depicted in Figure 7. This figure reveals that these contradiction rates are very dramatic. For instance, for test problems with 9 alternatives, the contradiction rates ranged from 70% (when the number of cost or benefit criteria is equal to 3) to almost 50% (when the number of criteria is equal to 21). For test problems with more than 11 alternatives the contradiction rates are between 100% and 50%, while for test problems with more than 19 alternatives the contradiction rates are always above 90%. The same results also indicate that the numbers of alternatives and criteria in a test problem play a similar role as in the previous results.

Figure 8 describes the results for the weighted measure given as "*Rate 5*". Please recall that this measure has been normalized by dividing the individual results by the maximum possible value for that case (see also the formulas at the end of Section 4 describing this weighted measure).

The computational results in this study strongly suggest that the two conflicting criteria aggregation approaches can make a substantial difference on the way the alternatives in an MCDA problem are ranked. The more the alternatives in a problem are, the higher is the possibility to have a significant discrepancy between the two rankings. For the case of *"Rate 4"* these differences can be profoundly dramatic.

Finally, Figures 9, 10, 11, and 12 present results when all the three additive MCDA methods (i.e., the WSM, revised AHP, and the original AHP) are compared with each other. These figures are similar to Figure 4 which presents comparative results regarding Rate 1. In all these figures the number of alternatives was kept equal to 11 while the number of criteria was equal to 3, 5, 7, …, 21. These figures demonstrate that in our computational experiments the three additive MCDA models (i.e., the WSM, the revised AHP and the original AHP) performed in almost the same manner. One may observe that in these last four figures it appears that the WSM performed slightly worse than the revised AHP which, in turn, performed slightly worse than the original AHP. The differences are statistically very small to have any real practical significance.

On the other hand, the multiplicative AHP would always generate contradiction rates equal to zero, as it is proven theoretically that it is always consistent and the two criteria aggregation methods always generate identical rankings of the alternatives. Please recall that in previous studies (i.e., in [Triantaphyllou, 2000; and 2001]) by the first author it was found that these three MCDA methods generated considerable ranking reversals under certain ranking reversal tests, while the multiplicative AHP was always perfectly consistent in terms of those ranking reversal tests.

6 Concluding Remarks

This paper does not claim that the process of dividing the criteria into two groups ("benefit" and "cost") is good or bad. It simply uses this grouping to derive two aggregated measures of performance of the alternatives. One measure is based on all the "benefit" criteria while the other is based on all the "cost" criteria. Next the two measures are processed according to the two approaches and the overall rankings of the alternatives are derived. This study found that this might yield contradictory results when a number of MCDA methods are used. This implies that the two criteria aggregation approaches, when benefit and cost criteria are simultaneously present in an MCDA problem, are indeed questionable when they are used with some additive MCDA methods.

Other MCDA methods, such as the ELECTRE and TOPSIS, were not examined because their structure dictates only a single way for aggregating criteria into the two categories. Some previous studies [Triantaphyllou, 2000; and 2001], [Wang and Triantaphyllou, 2004] have found that these methods suffer from certain ranking irregularities. Another group of methods, namely the PROMETHEE methods, have not been examined in this paper. However, since all these methods also use additive formulas, we believe that they too would exhibit ranking irregularities.

All the MCDA methods that suffer of the various ranking irregularities do so because of two factors. First, they somehow normalize the data when new alternatives are considered or existing ones are eliminated or modified and secondly, they use additive formulas in their computations. Such mathematical operations may alter the relative strength of some of the alternatives and thus alter their ranking.

In this paper we also demonstrated theoretically that the above problems with the two criteria aggregation methods, when benefit and cost criteria are simultaneously present in an MCDA problem, do not occur when two multiplicative MCDA methods are used. These methods are the WPM and the multiplicative AHP. We also note that this does not prove that these two methods are "perfect", but this consistency in their rankings is to their credit.

The above theoretical analyses and empirical results revealed that the two conflicting criteria aggregation approaches described in this paper may yield dramatically different rankings in many test MCDA problems. If for a given MCDA problem, the two rankings are identical, then the decision maker may feel more comfortable with it. This may be more the case if the corresponding final preference values (i.e., the P_i values) strongly discriminate among the alternatives. Otherwise, any results must be taken with some degree of skepticism. Finally, it should be stated at this point that the criteria aggregation approaches should be applied when criteria of both types (i.e., benefit and cost type) are present simultaneously in an MCDA problem. This is true regardless of the number of criteria, as long as criteria from both types are present.

The previous analyses reinforce the conclusions of previous studies which strongly support the use of the multiplicative AHP method. It should be stated here that either one of the two aggregation methods is applicable when the multiplicative AHP is used (as, according to this study, both methods are equivalent to each other under this particular MCDA method). The multiplicative AHP was also perfectly consistent when it was examined in terms of some ranking tests as reported in [Triantaphyllou, 2000; and 2001]. The investigation in [Triantaphyllou, 2001] also revealed cases with ranking irregularities (ranking reversals) when the WSM, AHP, and the revised AHP are used on simulated test problems or on some real-life MCDA problems.

Although the WPM and the multiplicative AHP are perfectly consistent in terms of the two ranking approaches analyzed in the present study or the ranking studies reported in [Triantaphyllou, 2001], these methods may still fail. When one considers problems in which all the criteria can be expressed in the same unit of measurement (say in U.S. dollars), then the results derived with the application of the WSM are the most widely accepted ones. However, when the WSM and the WPM (which can be viewed as an early version of the multiplicative AHP) are tested under the assumption of having the same unit of measurement, then in [Triantaphyllou and Mann, 1989] it was shown that these two methods may yield very different rankings.

Another alternative aggregation approach might be to compare each one of the *"benefit"* criteria with each one of the *"cost"* criteria and derive relative weights between the two types of criteria. Such an approach would certainly give the decision analyst the opportunity to incorporate more information into the decision-making process. It is noticeable that a similar approach plays a key role in the Analytic Hierarchy Process (AHP). However, that may introduce some additional challenges. For example, what happens if individual pairwise comparisons are inconsistent with each other? How to combine the individual comparisons into a unified result on the weights of the criteria? How to deal with the potentially high (actually, quadratic) number of comparisons?

In summary, the main findings of this study are as follows:

(1) For the decision problem of selecting the best alternative or ranking the alternatives when conflicting (i.e., "benefit" and "cost") criteria are present, it could make the difference which criteria aggregation method (i.e., the ratio of benefit to cost or the difference of benefit minus cost) one uses and which is the MCDA method used. In particular, if the original or the revised AHP are used, then the two criteria aggregation methods may yield significantly different results.
(2) The contradiction rates between the two rankings are more dramatic for problems with many alternatives but a few criteria.
(3) The three additive MCDA methods (i.e., the WSM, the revised AHP and the original AHP) performed almost the same way with the original AHP slightly better than the other two and the WSM slightly worse than the other two.
(4) There is no way to know which is the "right" ranking and which is not.
(5) The multiplicative AHP (and the WPM; an early version of the multiplicative AHP) is immune to any ranking reversals as both criteria aggregation methods always yield identical results. It is proved theoretically that these two criteria aggregation methods are perfectly consistent in terms of the ranking tests performed in this study. This result reinforces the merit of using the multiplicative AHP and it further supports similar results reported in [Triantaphyllou, 2000; and 2001] where some additive versions of the AHP and the WSM methods were tested for ranking reversals. In those studies all the criteria were considered to be either of the benefit or the cost type.

The first three comments are supported by the empirical results derived in this study, while the last two ones are supported by the theoretical issues discussed in this paper. This study further promotes the use of the multiplicative AHP versus the additive versions of the AHP and the WSM method. The findings of this paper reinforce the belief that the results of MCDA methods should not be taken literally but should be dealt with as decision support tools. Clearly, more research in this area is required.

Acknowledgements
The authors are very appreciative for the thoughtful and constructive comments made by the three anonymous reviewers of this paper and also the Associate Editor of the journal assigned to this paper.

REFERENCES

1. Augood, D.R., “A New Approach to R&D Evaluation,” IEEE Transactions on Engineering Management, Vol. EM-22, No. 1, February, pp. 2-10, 1978.

2. Barzilai, J., and F.A. Lootsma, "Power Relations and Group Aggregation in the Multiplicative AHP and SMART," *Proceedings of the Third International Symposium on the AHP*, George Washington University, Washington, DC, U.S.A., pp. 157-168, 1994.

3. Belton, V., and Gear, T., "On a Short-coming of Saaty's Method of Analytic Hierarchies", *Omega*, pp. 228-230, December 1983.

4. Bozeman, B., and J. Melkers, *Evaluating R&D Impacts: Methods and Practice,* Kluwer Academic Publishers, Boston, Mass., U.S.A., 1993.

5. Boucher, T.O., and Mcstravic, E.L., "Multi-attribute Evaluation within a Present Framework and its Relation to the Analytic Hierarchy Process", *The Engineering Economist*, 37, pp. 55-71, 1991.

6. Bridgeman, P. W., *Dimensionless Analysis*, Yale University Press, New Haven, CT, U.S.A., 1922.

7. Cambron, K.E., and Evans, G.W., "Layout Design using the Analytic Hierarchy Process," *Computers and Industrial Engineering*, Vol. 20, No. 2, pp. 221-229, 1991.

8. Doumpos, M. and C. Zopounidis, (2002), Multicriteria Decision Aid Classification Methods, Kluwer Academic Publishers, Dordrecht, The Netherlands.

9. Doumpos, M. and C. Zopounidis, (2004), “Developing Sorting Models Using Preference Disaggregation Analysis: An Experimental Investigation, *European Journal of Operational Research,* (in print).

10. Dyer, J.S., "Remarks on the Analytic Hierarchy Process," *Management Science*, Vol. 36, No. 3, pp. 249-258, 1990*a*.

11. Dyer, J.S., 'A Clarification of "Remarks on the Analytic Hierarchy Process",' *Management Science*, Vol. 36, No. 3, pp. 274-275, 1990*b*.

12. Dyer, J.S., and R.K. Sarin, “Measurable Multiattribute Value Functions,” *Operations Research*, Vol. 27, pp. 810-822, 1997.

13. Edwards, W., “How to Use Multiattribute Utility Measurement for Social Decision Making,” IEEE Trans. On Systems, Man, and Cybernetics, Vol. SMC-7, pp. 326-340, 1979.

14. Edwards, W., and F.H. Barron, “SMARTS and SMARTER: Improved Simple Methods for Multiattribute Utility Measurement,” *Organizational Behavior and Human Decision Processes*, Vol. 60, pp. 306-325, 1994.

15. Edwards, W., and J.R. Newman, "Multiattribute Evaluation," in H.R. Arkes and K.R. Hammond (Eds.), *Judgment and Decision Making: An Interdisciplinary Reader,* Cambridge University Press, Cambridge, U.K., pp. 13-37, 1986.

16. Figueira, J., Y. De Smet, and J.-P. Brans, (2004), "MCDA Methods for Sorting and Clustering Problems: Promethee Tri and Promethee Cluster, Working paper, Faculty of Economics, University of Coimbra, Coimbra, Portugal.

17. Fishburn, P. C., "Additive Utilities with Incomplete Product Set: Applications to Priorities and Assignments", *Operations Research*, 1967.

18. Harker, P.T., and L.G. Vargas, "Reply to 'Remarks on the Analytic Hierarchy Process'," *Management Science*, Vol. 36, No. 3, pp. 269-273, 1990.

19. Hobbs, B.F., and P.M Meier, Energy Decisions & The Environment: A Guide to the Use of Multicriteria Methods, International Series in Operations Research & Management Science, Kluwer Academic Publishers, Norwell, MA, 2000.

20. Hobbs, B.F., V. Chankong, W. Hamadeh, E. Stakhiv, "Does choice of multicriteria method matter? An experiment in water resource planning," *Water Resources Research*, Vol. 28, pp. 1767-1779, 1992.

21. Hobbs, B.F., "What can we learn from experiments in multiobjective decision analysis," *IEEE Trans. On Systems Management and Cybernetics,* Vol. 16, pp. 384-394, 1986.

22. Hwang, C.L., and K. Yoon, (1981), Multiple Attribute Decision Making: Methods and Applications, Springer-Verlag, New York, NY, U.S.A.

23. Jacquet-Lagreze, E., and Y. Siskos, (2001), "Preference Disaggregation: 20 Years of MCDA Experience," Invited Review, *European Journal of Operational Research,* 130, pp. 233-245.

24. Kirkwood, C.W., *Strategic Decision Making: Multiobjective Decision Analysis with Spreadsheets,* by Duxbury Press, ISBN: 0534516920, Belmont, Calif., U.S.A., 1997.

25. Kuwahara, Y., and Y. Takeda, "A Managerial Approach to Research and Development Cost-Effectiveness Evaluation," *IEEE Transactions on Engineering Management*, Vol. 37, No. 2, May, pp. 134-138, 1990.

26. Lootsma, F.A., *Multi-Criteria Decision Analysis via Ratio and Difference Judgment."* Kluwer Academic Publishers, Applied Optimization Series, Vol. 29, Dordrecht, The Netherlands, 1999.

27. Miller, D. W. and Starr, M. K., "Executive Decisions and Operations Research, Prentice-Hall, Inc., Englewood Cliffs, NJ, U.S.A., 1969.

28. Putrus, P., "Accounting for Intangibles in Integrated Manufacturing (non-financial justification based on the analytic hierarchy process)," *Information Strategy*, Vol. 6, pp. 25-30, 1990.

29. Ray, T., and E. Triantaphyllou, "Procedures For the Evaluation of Conflicts in Rankings of Alternatives," *Computers and Industrial Engineering*, Vol. 36, No. 1, pp. 35-44, 1999.

30. Ray, T., and E. Triantaphyllou, "Evaluation of Rankings With Regard to the Possible Number of Agreements and Conflicts," *European Journal of Operational Research*, Vol. 106, No. 1, pp. 129-136, 1998.

31. Roy, B., (1985), Methodologie Multicritiere d'Aide a la Decision. Econometrica, Paris, France.

32. Saaty, T.L., *Fundamentals of Decision Making and Priority Theory with the AHP,* RWS Publications, Pittsburgh, PA, U.S.A., 1994.

33. Saaty, T.L., (1990), 'An Exposition of the AHP in Reply to the Paper "Remarks on the Analytic Hierarchy Process",' *Management Science*, Vol. 36, No. 3, pp. 259-268, 1990.

34. Saaty, T.L., "Rank Generation, Preservation, and Reversal in the Analytic Hierarchy Process," *Decision Sciences,* Vol. 18, pp. 157-177, 1987.

35. Saaty, T.L., "Priority Setting in Complex Problems", *IEEE Transactions on Engineering Management*, Vol. 30, No. 3, pp. 140-155, 1983.

36. Saaty, T.L., *The Analytic Hierarchy Process*, McGraw-Hill, New York, NY, U.S.A., 1980.

37. Saaty, T.L., and L.G. Vargas, *Decision Making in Economic, Political, Social and Technological Environments with the AHP,* RWS Publications, Pittsburgh, PA, U.S.A., 1994.

38. Stewart, T.J., (1992), *"*A critical survey of the status of multiple criteria decision making theory and practice*,"* *OMEGA - The International Journal of Management Science* 20: 569-586, 1992.

39. Triantaphyllou, E., *Multi-Criteria Decision Making: A Comparative Study*, 320 pages, Kluwer Academic Publishers, Boston, Mass., U.S.A., 2000.

40. Triantaphyllou, E., "Two New Cases of Rank Reversals when the AHP and Some of its Additive Variants are Used that do not Occur with the Multiplicative AHP," *Multi-Criteria Decision Analysis,* (May 2001) Vol. 10, pp. 11-25, 2001.
(also available as a PDF file from: http://www.imse.lsu.edu/vangelis).

41. Triantaphyllou, E., Lootsma, F.A., Pardalos, P.M., and Mann, S.H. "On the Evaluation and Application of Different Scales for Quantifying Pairwise Comparisons in Fuzzy Sets", *Multicriteria Decision Analysis*, Vol. 3, pp. 1-23, 1994.
(also available as a PDF file from: http://www.imse.lsu.edu/vangelis).

42. Triantaphyllou, E., and S.H. Mann, "A Computational Evaluation of the AHP and the Revised AHP when the Eigenvalue Method is Used under a Continuity Assumption," *Computers and Industrial Engineering,* Vol. 26, No. 3, pp. 609-618, 1994*a*.
(also available as a PDF file from: http://www.imse.lsu.edu/vangelis).

43. Triantaphyllou, E. and S.H., Mann, "Some critical issues in making decisions with pairwise comparisons," *Proceedings of the 3rd International Symposium on the Analytic Hierarchy Process*, George Washington University, Washington, DC, U.S.A., July, 1994*b*.
(also available as a PDF file from: http://www.imse.lsu.edu/vangelis).

44. Triantaphyllou, E. and Mann S. H., "An Examination of the Effectiveness of Multi-Dimensional Decision-Making Methods: A Decision-Making Paradox", *Decision Support Systems,* Vol. 5, pp. 302-312, 1989.
(also available as a PDF file from: http://www.imse.lsu.edu/vangelis).

45. Wabalickis, R.N., "Justification of FMS with the Analytic Hierarchy Process," *Journal of Manufacturing Systems*, Vol. 17, pp. 175-182, 1988.

46. Wang, X., and E. Triantaphyllou, (2004), "Some Ranking Irregularities When the ELECTRE Method is Used for Decision-Making," *Proceedings of the Industrial Engineering Research Conference (IERC)*, Houston, TX, May 15-19, (included in the Conference's CD).

47. Zanakis, S., T. Mandakovic, S.K. Gupta, S. Sahay, and S. Hong, (1995), "A Review of Program Evaluation and Fund Allocation Methods within the Service and Government Sectors," *Socio-Economic Planning Sciences,* Vol. 29, No. 1, pages 59- 79.

48. Zanakis, S., A. Solomon, N. Wishart, and S. Dublish, (1998), "Multi-Attribute Decision Making: A Comparison of Select Methods," *European Journal of Operational Research,* No. 107, pages 507-529.

BIOGRAPHICAL INFORMATION ABOUT THE AUTHORS:

Dr. **Evangelos Triantaphyllou** received the Dual MS in Environment and Operations Research (OR) in 1985, the MS in Computer Science in 1988, and the Dual Ph.D. in Industrial Engineering and Operations Research in 1990, all from Penn State University. Prior that, he had received the BS degree in Architectural Engineering from the Nat'l Techn'l University in Athens, Greece, in 1983. Currently he is a Professor in the Industrial Engineering Department and an Adjunct Professor in the Computer Science Department at the Louisiana State University (LSU). He has also served as an Interim Associate Dean for the College of Engineering from 2000-2001. He is the author of the best seller book "Multi-Criteria Decision Making: A Comparative Study" (Kluwer, 2000). His research is in multi-criteria decision making/analysis (MCDM/MCDA), data mining, and operations research (OR). He has published in top journals in the above areas and served/serves on various editorial boards and in the organization of international conferences. In 1998 he received the prestigious Research Award by the OR Division of the IIE and a number of other awards for excellence in research, teaching, and service. Many of his graduate students received highly prestigious awards as well. His research has been supported by many federal and state agencies and also the private sector. More details about his academic activities can be found at his personal website at LSU at: http://www.imse.lsu.edu/vangelis

Mr. **Khalid Baig** received the MS in Industrial Engineering from the Louisiana State University in 1996. His Major Professor was Professor Evangelos Triantaphyllou. A year later he received the MS degree in Systems Science from the Computer Science Department at LSU. Before that, he had received the BS degree with Honors in Production Engineering from the University of Bombay, India. Currently, he works as a software developer for a private computer consultancy company.