2007

Study of ranking irregularities when evaluating alternatives by using some ELECTRE methods and a proposed new MCDM method based on regret and rejoicing

Xiaoting Wang
Louisiana State University and Agricultural and Mechanical College

Recommended Citation

Wang, Xiaoting, "Study of ranking irregularities when evaluating alternatives by using some ELECTRE methods and a proposed new MCDM method based on regret and rejoicing" (2007). *LSU Master's Theses*. 860.
https://digitalcommons.lsu.edu/gradschool_theses/860

STUDY OF RANKING IRREGULARITIES WHEN EVALUATING ALTERNATIVES BY USING SOME ELECTRE METHODS AND A PROPOSED NEW MCDM METHOD BASED ON REGRET AND REJOICING

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
Master of Science in Industrial Engineering

in

The Department of Industrial and Manufacturing Systems Engineering

by
Xiaoting Wang
B.S., Nanchang Institute of Aeronautical Technology, China, 1999
M.S., Northwestern Polytechnic University, China, 2002
August, 2007

Acknowledgements

I would like to express my deepest appreciation and respect to my major professor, Dr. Evangelos Triantaphyllou, for his guidance, help and encouragements throughout my entire research studies.

Very special thanks go to my co-chair professor, Dr. Warren Liao, for all his help and support to my study and research.

I would also like to specially thank Dr. Xiaoyue Jiang and Dr. Jianhua Chen for their help, encouragements and being the committee members.

It is really an honor to have all these professors in my committee.

Also I would like to thank my dearest husband for his tremendous support to my research and studies.

Table of Contents

Abstract

Multi-criteria decision-making (MCDM) is one of the most widely used decision methodologies in the sciences, business, and engineering worlds. MCDM methods aim at improving the quality of decisions by making the process more explicit, rational, and efficient. One intriguing problem is that oftentimes different methods may yield different answers to the same decision problem. Thus, the issue of evaluating the relative performance of different MCDM methods is raised. One evaluating procedure is to examine the stability of an MCDM method's mathematical process by checking the validity of its proposed rankings.

The ELECTRE II and III methods are two well-known MCDM methods and widely accepted in solving MCDM problems in civil and environmental engineering. However these two methods have never been studied in detail for the validity of their proposed rankings. Thus, the first aim of this thesis is to examine if these two methods suffer of any type of ranking irregularities and analyze the reasons of the phenomenon.

As the research results in this thesis revealed, the ELECTRE II and III methods do allow some types of ranking irregularities to happen. For instance, these two methods might change the indication of the best alternative for a given MCDM problem when one of the non-optimal alternatives was replaced by a worse one. The two methods were also evaluated in terms of two other ranking tests and they failed them as well. Two real-life cases are described to demonstrate the occurrence of rank reversals. Then reasons behind the phenomenon are analyzed. Next an empirical study and some real-life case studies were executed and discussed. The results of these examinations show that the rates of those ranking irregularities were rather significant in both the simulated decision problems and the real-life cases studied in this research.

However, some recent studies showed that rank reversals could also happen because people may reverse their preferences due to some emotional feelings, like regret and rejoicing. Thus this thesis proposes a new MCDM method which is based on regret and rejoicing. This new method is expected to satisfy a set of critical conditions.

1. Problem Description

Making all kinds of decisions is an indispensable part of our lives. From the ancient times to the modern age, people never stopped their efforts in seeking ways for making more reliable and scientifically sound decisions. For those daily life decision problems, one may quickly decide them just by using his/her personal preferences, experiences, and/or instincts. However, in many fields of engineering, business, government, and sciences, where decisions may be worth millions or billions of dollars, or decisions may have a significant impact on the welfare of the society, decision-making problems are usually too complex and anything but as easy.

For instance, many large companies and organizations face the problem of prioritizing a set of competing projects. Each one of these projects may have some short-term and long-term potential profits, costs and some negative or positive side effects. At the same time, there is a limited budget to be distributed among these projects. Some of the projects may not get funded at all. Besides these projects, the decision makers have also defined some criteria to be used to study these projects in depth. The problem underlined in the above situation is how to use all available information and assign priorities to these competing projects or decision alternatives. This is a pervasive problem in today's highly sophisticate world. When faced with such decision making problems, no single decision maker or group of decision makers can systematically consider all the available information simultaneously and reach the right decisions by just using their experiences or instincts. For such case people need to use valid decision analysis approaches and tools to analyze all the issues involved and eventually to reach the optimal decisions. They need to do so in a way that can be easily and objectively explained to others and be defended to a wide audience of stakeholders. This is how and why the field of decision sciences has emerged as an important scientific discipline for today's complex world.

In the past few decades, numerous decision-making methods and decision aid software packages have been proposed in the literature and are used in various areas. Among them, a class of methods known as multi-criteria decision-making (MCDM) is one of the most widely used decision methodologies in the sciences, business, and engineering worlds. MCDM methods aim at improving the quality of decisions by making the decision-making process more explicit, rational, and efficient. It is not a coincidence that a simple search (for instance, by using google.com) on the web under the key words "multi criteria decision making" returns more than ten million hits. Some applications of MCDM include the use in civil and environmental engineering [Zavadskas, *et al*., 2004; Hobbs and Meier, 2000], like water resources planning [Raj, 1995], in financial engineering [Zopounidis and Doumpos, 2000], like credit risk assessment [Doumpos, *et al*., 2002] and in some current problems like flood management for flood hazard mitigation [Jason, *et al*., 2007] and allocation of scarce homeland security resources for economic efficiency and maximum protection [Farrow and Valverde, 2005] etc.

Although MCDM has attracted the interest of researchers and practitioners for many years in a wide spectrum of areas, it is far from mature and there are still a lot of unsolved issues. One intriguing problem with MCDM methods is that oftentimes different methods may yield different answers when they are fed with exactly the same decision problem. Thus, the issue of evaluating the relative performance of different MCDM methods is naturally raised. This, in turn, raises the question of how can one evaluate the performance of different MCDM methods? Since it is practically impossible to know which one is the best alternative for a given decision problem, some kind of testing procedures need to be developed. One such procedure is to examine the stability of an MCDM method's mathematical process by checking the validity of its proposed rankings. The development of this procedure of evaluating the performance of different MCDM methods comes from some studies on some MCDM methods [Belton and Gear, 1983; Dyer and Wendell, 1985; Triantaphyllou, 2000]. In

these studies, the original additive AHP method, one of the most well-known MCDM methods, was found to allow some rank reversal problems to happen.

Rank reversal means that the ranking between two alternatives might be reversed after some variation occurs to the decision problem, like adding a new alternative, dropping an old alternative or replacing an old alternative by a worse one, etc. For example, two alternatives A_1 and A_2 may be initially ranked as $A_1 > A_2$ (i.e., A_1 is more preferable than A_2). After a new alternative A_3 is introduced into the decision problem and the alternatives are ranked again by using the same method, the ranking between A_1 and A_2 may be reversed and become $A_2 > A_1$. (For a detailed rank reversal example, please refer to Chapter 2). Usually such rank reversals are undesirable for decision-making problems when they are totally caused by the mathematical instabilities of the used method. Cases of rank reversals and some other types of ranking irregularities when the AHP is used have been reported by many researchers for a number of years [Belton and Gear, 1983; Dyer and Wendell, 1985; Triantaphyllou, 2000]. However there are very few studies which examine the validity of the ranking results from some other widely used MCDM methods.

The ELECTRE method is another type of well-known MCDM method, especially in Europe. Among different variants of the ELECTRE method, the ELECTRE II and III methods have been widely accepted in solving MCDM problems in the engineering world, like civil and environmental engineering [Hobbs and Meier, 2000]. Applications include the assessment of complex civil engineering projects, selection of highway designs, site selection for the disposal of nuclear waste and nuclear plant [Roy and Bouyssou, 1986], water resources planning [Raj, 1995] and solid waste management [Hokkanen and Salminen, 1997]. Though there have been so many publications which reported the applications of these two MCDM methods, the ELECTRE II and III methods have never been studied in detail for the validity of their proposed rankings. Thus, the first goal of this research is to examine if the ELECTRE II and III methods suffer of any type of ranking irregularity problems and analyze the reasons behind the phenomenon.

This thesis is organized as follows: the second chapter presents a literature review on MCDM and some past studies on rank reversals with the AHP methods. The third chapter discusses the three test criteria that are used in this thesis to test the performance of the ELECTRE II and III methods. The fourth chapter describes two examples of real-life decision problems for which rank reversals occurred under test criterion #1 by using the ELECTRE II and III methods. The fifth chapter presents a detailed analysis on the reasons for the rank reversal problems with the ELECTRE II and III methods. Next the sixth chapter discusses the results from an empirical study on randomly generated decision problems according to the three test criteria described in chapter 3. Then the results of some real-life case studies are described in the seventh chapter. In the eighth chapter, some ongoing studies on the influence of regret and rejoicing to MCDM problems are discussed and a new MCDM method is proposed. Finally, the concluding remarks are presented in the ninth chapter.

2. Introduction to MCDM and Literature Review

A typical MCDM problem 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, the performance values a_{ij} and the criteria weights w_j are viewed as the entries of a *decision matrix* as shown below. The a_{ij} element of the decision matrix represents the performance value of the *i-th* alternative in terms of the *j-th* criterion. The w_j value represents the weight of the *j-th* criterion. Data for MCDM problems can be determined by direct observation (if they are easily quantifiable) or by indirect means if they are qualitative [Triantaphyllou, *et al.*, 1994].

	C r i t e r i a			
	C_1	C_2	...	C_n
	$(w_1$	w_2	...	$w_n)$
Alternatives				
A_1	a_{11}	a_{12}	...	a_{1n}
A_2	a_{21}	a_{22}	...	a_{2n}
.	.	.	...	.
.	.	.	…	.
.	.	.	…	.
A_m	a_{m1}	a_{m2}	...	a_{mn}

From the early developments of the MCDM theories in the 1950s and 1960s, a plethora of MCDM methods have been proposed in the literature and new contributions are continuously coming forth in this area. There are also many ways to classify the existing MCDM methods. One of the ways is to classify MCDM methods according to the type of data they use. Thus, there are deterministic, stochastic, or fuzzy MCDM methods [Triantaphyllou, 2000]. Another way of classifying MCDM methods is according to the number of the decision makers involved in the decision process. Hence, there are single decision maker MCDM methods and group decision-making MCDM. For some representative articles in this area, see [George, *et al.*, 1992], [Hackman and Kaplan, 1974], and [DeSanctis and Gallupe, 1987].

2.1 Some Well-known MCDM Methods

Among the numerous MCDM methods, there are several prominent families that have enjoyed a wide acceptance in the academic area and many real-world applications. Each of these methods has its own characteristics, background logic and application areas. Next is a brief description of some of them.

2.1.1 The Analytic Hierarchy Process and Some of Its Variants

The Analytic Hierarchy Process (or AHP) method was developed by Professor Thomas Saaty [Saaty, 1980; Saaty, 1994; Saaty and Vargas, 2000]. This decision-making method can help people set priorities and choose the best alternatives by reducing complex decision problems to a system of hierarchies. Since its inception, it has evolved into several different variants and has been widely used to solve a broad range of multi-criteria decision problems [Vaidya and Kumar, 2006].

2.1.1.1 The Original Analytic Hierarchy Process

The AHP method uses the pairwise comparisons and eigenvector methods to determine the a_{ij} values and also the criteria weights w_j. The details about the pairwise comparisons and eigenvector methods can be found in [Saaty, 1980; Saaty, 1994; Saaty and

Vargas, 2000]. In this method, a_{ij} represents the relative value of alternative A_i when it is considered in terms of criterion C_j. In the original AHP method, the a_{ij} values of the decision matrix need to be normalized vertically. That is, the elements of each column in the decision matrix add up to 1. In this way, values with various units of measurement can be transformed into dimensionless ones. If all the criteria are *benefit* criteria (that is, the higher the score the better the performance is), then according to the original AHP method, the best alternative is the one that satisfies the following expression:

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

From the above formula, it can be seen that the original AHP method uses an additive expression to determine the final priorities of the alternatives in terms of all the criteria simultaneously. Next the revised AHP is introduced, which is also an additive variant of the original AHP method.

2.1.1.2 The Revised Analytic Hierarchy Process

The revised AHP model was proposed by Belton and Gear in [1983] after they first found a case of rank reversal that occurred when the original AHP was used. In their case, the original AHP method was used to rank three alternatives in a simple test problem. Then a fourth alternative, identical to one of the three alternatives, was introduced in the original decision problem without changing any other data. The ranking of the original three alternatives was changed after the revised problem was ranked again by the same method. The following is this rank reversal example from [Belton and Gear, 1983].

Suppose the decision matrix of a decision problem with three alternatives and three criteria is as follows:

	C r i t e r i a		
	C_1	C_2	C_3
	(1/3	1/3	1/3)
Alts.			
A_1	1	9	8
A_2	9	1	9
A_3	1	1	1

By using the original AHP method, the above decision matrix is normalized first by the column totals to get the relative data as follows:

	C r i t e r i a		
	C_1	C_2	C_3
	(1/3	1/3	1/3)
Alts.			
A_1	1/11	9/11	8/18
A_2	9/11	1/11	9/18
A_3	1/11	1/11	1/18

Then, it can be shown that the final AHP scores of the three alternatives are: (0.45, 0.47, 0.08). That is, $A_2 > A_1 > A_3$. Next, a new alternative A_4 which is identical to the existing alternative A_2 is added to the decision matrix. Now the normalized decision matrix is as follows:

	C r i t e r i a		
	C_1	C_2	C_3
	(1/3	1/3	1/3)
Alts.			
A_1	1/20	9/12	8/27
A_2	9/20	1/12	9/27
A_3	1/20	1/12	1/27
A_4	9/20	1/12	9/27

By using the same AHP method, now the final AHP scores of these alternatives are: (0.37, 0.29, 0.06, 0.29). That is, the four alternatives are ranked as $A_1 > A_2 = A_4 > A_3$. This result contradicts the previous one in which $A_2 > A_1$.

Later the above ranking abnormality was defined as a rank reversal. According to Belton and Gear the root for this inconsistency is the fact that the relative values of the alternatives for each criterion sum up to 1. So instead of having the relative values of the alternatives sum up to 1, they proposed to divide each relative value by the maximum value of the relative values. 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 (2-1), but now the normalization is different.

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

The revised AHP was sharply criticized by Saaty in [1990]. After many debates and a heated discussion (e.g., [Dyer, 1990a; and 1990b], [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]. However, even earlier, the revised AHP method was found to suffer of some other ranking problems even without the introduction of identical alternatives [Triantaphyllou and Mann, 1989]. In that study and also in [Triantaphyllou, 2000; and 2001], it was found that most of the problematic situations of the AHP methods are caused by the required normalization (either by dividing by the sum of the elements or by the maximum value in a vector) and the use of an additive formula on the data of the decision matrix for deriving the final preference values of the alternatives. However, in the core step of one of the MCDM methods known as the Weighted Product Model (WPM) [Bridgeman, 1922; Miller and Starr, 1969], the use of an additive formula is avoided by using a multiplicative expression. This brought the development of a multiplicative version of the AHP method, known as the multiplicative AHP.

2.1.1.3 The Multiplicative Analytic Hierarchy Process

The use of multiplicative formulas in deriving the relative priorities in decision-making is not new [Lootsma, 1991]. A critical development appears to be the use of multiplicative formulations when one aggregates the performance values a_{ij} with the criteria weights w_j. In the WPM method, 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, the following formula is used ([Bridgeman, 1922; Miller and Starr, 1969]) in order to compare two alternatives A_K and A_L:

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

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.

Based on the WPM method, Barzilai and Lootsma [1994] and Lootsma [1999] proposed the multiplicative version of the AHP method. According to this method, the relative performance values a_{ij} and criteria weights w_j are not processed according to formula (2-1), but the WPM formula (2-3) is used instead. Furthermore, one can use a variant of formula (2-3) to compute preference values of the alternatives that in turn, can be used to rank them. The preference values can be computed as follows:

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

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.

From formula (2-3), it can be seen that not only the use of an additive formula is avoided in the multiplicative AHP, but also the negative effects of normalization can also be eliminated by using the multiplicative formula. These properties of the multiplicative AHP are demonstrated theoretically in [Triantaphyllou, 2000]. In that study, it was also proved that most of the ranking irregularities which occurred when the additive variants of the AHP method were used will not occur with the multiplicative AHP method.

2.1.2 The ELECTRE Methods

Another prominent role in MCDM methods is played by the ELECTRE approach and its derivatives. This approach was first introduced in [Benayoun, *et al.*, 1966]. The main idea of this method is the proper utilization of what is called "outranking relations" to rank a set of alternatives. The ELECTRE approach uses the data within the decision problems along with some additional threshold values to measure the degree to which each alternative outranks all others. Soon after the introduction of the first version known as ELECTRE I [Roy, 1968], this approach has evolved into a number of other variants. Today two of the most widely used versions are known as ELECTRE II [Roy and Bertier, 1971, 1973] and ELECTRE III [Roy, 1978]. Another variant of the ELECTRE approach is the TOPSIS method [Hwang and Yoon, 1981]. Since the ELECTRE approach is more complicated than the AHP approach, a detailed description about the process of ELECTRE II and III methods is given in Chapter 4 by ways of two numerical examples.

Compared with the simple process and precise data requirement of the AHP methods, ELECTRE methods are able to apply more complicated algorithms to deal with the complex and imprecise information from the decision problems and use these algorithms to rank the alternatives. The ELECTRE algorithms look reliable and in neat format. People believe that the process of this approach could lead to an explicit and logical ranking of the alternatives. However at it will be shown later, this may not always be the case.

2.1.3 Multi-Attribute Utility Analysis

In contrast to the above two approaches, multi-attribute utility analysis (MAUA) is another type of a systematic method for identifying and analyzing various alternatives and factors in order to arrive at a rational decision [Keeney and Raiffa, 1976; Kirkwood, 1997]. This MCDM approach transfers the performance value of an alternative under each decision criterion into a utility value according to a corresponding utility scale for that criterion and assigns a relative value for each criterion to stand for its importance (i.e., the weight of each criterion). The utility is a numerical value between 0 and 1 and it represents the preferability of the alternative under that decision criterion. Considering the importance of each criterion, the utility of each decision alternative under each criterion is multiplied by the weight of that criterion. Then the total utility of each decision alternative could be calculated by summing up the weighted utility values under each decision criterion. Next the alternatives are ranked in terms of their aggregated utilities.

One of the key assumptions under the above utility model is that the decision makers (DMs) are "Rational Individuals" which devoid of psychological influences or emotions [Luce, 1992]. Under this assumption, it is expected that DMs will always want to make choices that can maximize the utilities of the chosen alternatives and the utilities of the alternatives are independent with each other. However, behavioral scientists have proved that it is not always appropriate to relate decision rationality to utility maximization. Examples demonstrating systematic violations of the independence and the utility maximization principle can be found in [Allais, 1988] and [Ellsberg, 1961].

2.2 Literature Review on Rank Reversals with the AHP Methods

The AHP method has been widely used in many real-life decision problems. Thousands of AHP applications have been reported in edited volumes and books (e.g., Golden, *et al.*, 1989, Saaty and Vargas, 2000) and on websites (e.g., www.expertchoice.com). However, many researchers have also criticized the AHP method for some of its problems. One such key problem is rank reversals. As mentioned previously, Belton and Gear in [1983] first reported the problem of rank reversals with the AHP. Their rank reversal example demonstrated that the ranking of alternatives might be altered by the addition (or deletion) of non-optimal alternatives (please refer to the rank reversal example in Section 2.1.1.2). This phenomenon inspired some doubts about the reliability and validity of the original AHP method. Soon after the first report, some other types of ranking irregularities with the original AHP method were also found. Dyer and Wendell in [1985] studied rank reversals when the AHP was used and near copies of some alternatives were considered in the decision problem. In [2000] Triantaphyllou reported another type of rank reversal with the additive AHP methods in which the indication of the optimal alternative may change when one of the non-optimal alternatives is replaced by a worse one. Next in [2001] Triantaphyllou reported two new cases of ranking irregularities with the additive AHP methods. One is that the ranking of the alternatives may be different when all the alternatives are compared two at a time and also simultaneously. Another case is that the ranking of the alternatives may not follow the transitivity property when the alternatives are compared two at a time.

Many researchers have also put a lot of effort in explaining the reasons behind the rank reversals and study how to avoid them. Belton and Gear in [1983] proposed the revised AHP method in order to preserve the ranking of the alternatives under the presence of identical alternatives. Saaty in [1987] pointed out that rank reversals were due to the inclusion of duplicates of the alternatives. Therefore he suggested that people should avoid the introduction of similar or identical alternatives. However, other cases were later found in which rank reversal occurred without the introduction of identical alternatives [Triantaphyllou, 2000]. Dyer in [1990a] indicated that the sum to unity normalization of priorities makes each one dependent on the set of alternatives being compared. He also claimed that the resulted individual priorities are thus arbitrary, as arbitrary sets of alternatives may be considered in the decision problem. Stam and Silva, in [1997] revealed that if the relative preference statements about alternatives were represented by judgment intervals (that is, the pairwise preference judgments are uncertain (stochastic)), rather than single values, then the rankings resulting from the traditional AHP analysis based on the single judgment values may be reversed and therefore are incorrect. Based on this statement, they developed some multivariate statistical techniques to obtain both point estimates and confidence intervals for the occurrence of certain types of rank reversal probabilities with the AHP method.

3. Some Test Criteria for Evaluating MCDM Methods

From the above chapter, it can be seen that most of the past research studies on examining the validity of ranking results from MCDM methods concentrated on the AHP method. There are very few studies that explore the reliability and validity of some other MCDM methods, like the widely used ELECTRE II and III methods. Does that mean decision makers can trust these methods without any questioning of the validity of their answers? The answer is absolutely "No". Usually, decision makers undertake some kind of a sensitivity analysis to examine how the decision results will be affected by changes in some of the uncertain data in a decision problem. For example, is the ranking of the alternatives stable or easily changeable under different set of criteria weights? By this process, decision analysts may better understand a decision problem. It is not safe to accept an MCDM method as being accurate all the time.

In [Triantaphyllou, 2000], three test criteria were established to evaluate the performance of MCDM methods by testing the validity of their ranking results. These test criteria are as follows:

Test Criterion #1:
An effective MCDM method should not change the indication of the best alternative when a non-optimal alternative is replaced by another worse alternative (given that the relative importance of each decision criterion remains unchanged).

Suppose that an MCDM method has ranked a set of alternatives in some way. Next, suppose that a non-optimal alternative, say A_k, is replaced by another alternative, say $A_k^{/}$, which is less desirable than A_k. Then, according to test criterion #1 the indication of the best alternative should not change when the alternatives are ranked again by the same method. The same should also be true for the relative rankings of the rest of the unchanged alternatives.

Test Criterion #2:
The rankings of alternatives by an effective MCDM method should follow the transitivity property.

Suppose that an MCDM method has ranked a set of alternatives of a decision problem in some way. Next, suppose that this problem is decomposed into a set of smaller problems, each defined on two alternatives at a time and the same number of criteria as in the original problem. Then, according to this test criterion all the rankings which are derived from the smaller problems should satisfy the transitivity property. That is, if alternative A_1 is better than alternative A_2, and alternative A_2 is better than alternative A_3, then one should also expect that alternative A_1 is better than alternative A_3.

The third test criterion is similar to the previous one but now one tests for the agreement between the smaller problems and the original un-decomposed problem.

Test Criterion #3:
For the same decision problem and when using the same MCDM method, after combining the rankings of the smaller problems that an MCDM problem is decomposed into, the new overall ranking of the alternatives should be identical to the original overall ranking of the un-decomposed problem.

As before, suppose that an MCDM problem is decomposed into a set of smaller problems, each defined on two alternatives and the original decision criteria. Next suppose that the rankings of the smaller problems follow the transitivity property. Then, according to this test criterion when the rankings of the smaller problems are all combined together, the

new overall ranking of the alternatives should be identical to the original overall ranking before the problem decomposition.

In this research, these three test criteria were used to evaluate the performance of the ELECTRE II and the ELECTRE III methods. Both of them failed in terms of each one of these three test criteria. Next two rank reversal examples which occurred with the ELECTRE II and III methods under the first test criterion are demonstrated. The other two test criteria were also applied as they had been stated above.

4. Illustration of Rank Reversals with the ELECTRE II and III Methods

For most ELECTRE methods, there are two main stages. These are the construction of the outranking relations and the exploitation of these relations to get the final ranking of the alternatives. Different ELECTRE methods may differ in how they define the outranking relations between the alternatives and how they apply these relations to get the final ranking of the alternatives. This is true with the ELECTRE II and III methods. However, the essential difference between these two methods is that they use different types of criteria. ELECTRE II uses the true criteria where no thresholds exist and the differences between criteria scores are used to determine which alternative is preferred. In this preference structure, the indifference relation is transitive [Rogers, *et al.*, 1999]. The criteria used by ELECTRE III are pseudo-criteria which involve the use of two-tiered thresholds. One is the indifference threshold q, below which the decision maker shows clear indifference, and the other one is the preference threshold p, above which the decision maker is certain of strict preference [Rogers, *et al.*, 1999]. The situation between the above two is regarded as weak preference for alternative a over alternative b which indicates the decision maker's hesitation between indifference and strict preference [Rogers, *et al.*, 1999]. The following two rank reversal examples demonstrate how both of the two methods work and how the rank reversals may happen when using them to rank a set of decision alternatives.

4.1 An Example of Rank Reversal with the ELECTRE II Method

This example is based on a real-life case study where the ELECTRE II method was used to help find the best location for a wastewater treatment plant in Ireland [Rogers, *et al.*, 1999]. The decision problem is defined on 5 alternatives and 7 criteria. Note that here all the criteria are *benefit* criteria, that is, the higher the score the better the performance is. The decision matrix, that is, the performances of the alternatives A_i in terms of the criteria C_j, is as follows:

	C_1	C_2	C_3	C_4	C_5	C_6	C_7
A_1	1	2	1	5	2	2	4
A_2	3	5	3	5	3	3	3
A_3	3	5	3	5	3	2	2
A_4	1	2	2	5	1	1	1
A_5	1	1	3	5	4	1	5

The weights of the criteria are:

	C_1	C_2	C_3	C_4	C_5	C_6	C_7
Weight	0.0780	0.1180	0.1570	0.3140	0.2350	0.0390	0.0590

The ELECTRE methods are based on the evaluation of two indices, the *concordance* index and the *discordance* index, defined for each pair of alternatives. The concordance index for a pair of alternatives a and b measures the strength of the hypothesis that alternative a is at least as good as alternative b. The discordance index measures the strength of evidence against this hypothesis [Belton and Stewart, 2001]. There are no unique measures of concordance and discordance indices. In ELECTRE II, the concordance index $C(a, b)$ for each pair of alternatives (a, b) is defined as follows:

$$C(a,b) = \frac{\sum_{i \in Q(a,b)} w_i}{\sum_{i=1}^{m} w_i}. \tag{4-1}$$

Where $Q(a, b)$ is the set of criteria for which a is equal or preferred to (i.e., at least as good as) b, and w_i is the weight of the *i-th* criterion. For instance, the concordance indices for this example are as follows:

	A_1	A_2	A_3	A_4	A_5
A_1	1.0000	0.3730	0.4120	0.8430	0.5490
A_2	0.9410	1.0000	1.0000	1.0000	0.7060
A_3	0.9410	0.9020	1.0000	1.0000	0.7060
A_4	0.6670	0.3140	0.3140	1.0000	0.5490
A_5	0.8430	0.7650	0.7650	0.8820	1.0000

The discordance index $D\,(a,\,b)$ for each pair of alternatives $(a,\,b)$ is defined as follows:

$$D(a,b) = \frac{\max_i [g_i(b) - g_i(a)]}{\delta}. \tag{4-2}$$

Where $g_i(a)$ represents the performance of alternative a in terms of criterion C_i, $g_i(b)$ represents the performance of alternative b in terms of criterion C_i, and $\delta = \max_i | g_i(b) - g_i(a)|$ (i.e., the maximum difference on any criterion). This formula can only be used when the scores for different criteria are comparable. When the above formula is used, it turns out that the discordance indices for this example are as follows:

	A_1	A_2	A_3	A_4	A_5
A_1	0.0000	0.7500	0.7500	0.2500	0.5000
A_2	0.2500	0.0000	0.0000	0.0000	0.5000
A_3	0.5000	0.2500	0.0000	0.0000	0.7500
A_4	0.7500	0.7500	0.7500	0.0000	1.0000
A_5	0.2500	1.0000	1.0000	0.2500	0.0000

After computing the concordance and discordance indices for each pair of alternatives, two types of outranking relations are built by comparing these indices with two pairs of threshold values: $(C^*,\,D^*)$ and (C^-,D^-). The pair $(C^*,\,D^*)$ is defined as the concordance and discordance thresholds for the *strong* outranking relation and the pair $(C^-,\,D^-)$ is defined as the thresholds for the *weak* outranking relation where $C^* > C^-$ and $D^* < D^-$. Next the outranking relations are built according to the following two rules:

(1) If $C(a,\,b) \geq C^*$, $D(a,\,b) \leq D^*$ and $C(a,\,b) \geq C(b,\,a)$, then alternative a is regarded as strongly outranking alternative b.

(2) If $C(a,\,b) \geq C^-$, $D(a,\,b) \leq D^-$ and $C(a,\,b) \geq C(b,\,a)$, then alternative a is regarded as weakly outranking alternative b.

The values of $(C^*,\,D^*)$ and $(C^-,\,D^-)$ are decided by the decision makers for a particular outranking relation. They may be varied to yield more or less severe outranking relations. The higher the value of C^* and the lower the value of D^*, the more severe the outranking relation becomes. That is, the more difficult it is for one alternative to outrank another [Belton and Stewart, 2001].

For this example, two pairs of thresholds for the strong outranking relation and one pair of thresholds for the weak outranking relation were chosen to be as follows: C_1^*=0.85, D_1^*=0.50; C_2^*=0.75, D_2^*=0.25; and C^-=0.65, D^-=0.25. According to the above rules and the three pairs of thresholds, the outranking relations for this example were derived to be as follows:

	A_1	A_2	A_3	A_4	A_5
A_1	—			S^F	
A_2	S^F	—	S^F	S^F	
A_3	S^F		—	S^F	
A_4				—	
A_5	S^F			S^F	—

In the above notation S^F stands for the strong outranking relation. For example, A_1 S^F A_4 means that alternative A_1 strongly outranks alternative A_4. We use S^f (i.e., the superscript now

is low case "f"; not present on the above table) to stand for the weak outranking relation. The weak outranking relation would happen later in this example.

On the basis of the outranking relations, next the descending and ascending distillation processes are applied to obtain two complete pre-orders of the alternatives. The details of the distillation processes can be found in [Belton and Stewart, 2001] and [Rogers, *et al.*, 1999]. The descending pre-order is built up by starting with the set of "best" alternatives (those which outrank other alternatives) and going downward to the worse one. On the contrary, the ascending pre-order is built up by starting with the set of "worst" alternatives (those which are outranked by other alternatives) and going upward to the best one. The distillation results for this example are as follows: the pre-order from the descending distillation is $A_2 = A_5 > A_3 > A_1 > A_4$; the pre-order from the ascending distillation is $A_2 > A_5 = A_3 > A_1 > A_4$.

The last step is to combine the two complete pre-orders to get either a partial or a complete final pre-order. Whether the final product is a partial pre-order (not containing a relative ranking of all of the alternatives) rather than a complete pre-order depends on the level of consistency between the rankings from the two distillation procedures [Rogers, *et al.*, 1999]. The partial pre-order allows two alternatives to remain incomparable without affecting the validity of the overall ranking, which differentiates from the complete pre-order. A commonly used method for determining the final pre-order is to take the intersection of the descending and ascending pre-orders. The intersection of the two pre-orders is defined such that alternative *a* outranks alternative *b* if and only if *a* outranks or is in the same class as *b* according to the two pre-orders. If alternative *a* is preferred to alternative *b* in one pre-order but *b* is preferred to *a* in the other one, then the two alternatives are incomparable in the final pre-order [Rogers, *et al.*, 1999]. By following the above rules, the intersection of the two pre-orders for this example resulted in the following complete pre-order of the alternatives: $A_2 > A_5 > A_3 > A_1 > A_4$ and obviously A_2 is the optimal alternative at this point.

Next, alternative A_3 was randomly selected to be replaced by a worse one, say A_3', in order to test the stability of the ranking of the alternatives under the first test criterion. The new decision matrix now is as follows:

	C_1	C_2	C_3	C_4	C_5	C_6	C_7
A_1	1	2	1	5	2	2	4
A_2	3	5	3	5	3	3	3
A_3'	**2**	**4**	**2**	**4**	**2**	**1**	**1**
A_4	1	2	2	5	1	1	1
A_5	1	1	3	5	4	1	5

Please note that alternative A_3 is replaced by a less desirable one denoted as A_3' which is determined by subtracting the value 1 from the performance values of the original alternative A_3 (the subtracted value was selected randomly by a computer program to make sure it will cause the chosen alternative to become worse than the one it replaces).

The rest of the data are kept the same as before. The intermediate results during the ranking process are as follows:

The concordance indices are:

	A_1	A_2	A_3'	A_4	A_5
A_1	—	0.3730	0.6470	0.8430	0.5490
A_2	0.9410	—	1.0000	1.0000	0.7060
A_3'	0.5880	0	—	0.6860	0.2350
A_4	0.6670	0.3140	0.5690	—	0.5490
A_5	0.8430	0.7650	0.8040	0.8820	—

The discordance indices are:

	A_1	A_2	$A_3^{/}$	A_4	A_5
A_1	—	0.7500	0.5000	0.2500	0.5000
A_2	0.2500	—	0	0	0.5000
$A_3^{/}$	0.7500	0.5000	—	0.2500	1.0000
A_4	0.7500	0.7500	0.5000	—	1.0000
A_5	0.2500	1.0000	0.7500	0.2500	—

The outranking relations are:

	A_1	A_2	$A_3^{/}$	A_4	A_5
A_1	—			S^F	
A_2	S^F	—	S^F	S^F	
$A_3^{/}$			—	S^f	
A_4				—	
A_5	S^F			S^F	—

where S^f (in location (3, 4)) stands for the weak outranking relation.

When the descending and ascending distillation processes are applied again, the descending pre-order now is $A_2= A_5> A_3= A_1> A_4$ while the ascending pre-order is $A_2=A_5>A_3= A_1> A_4$. After combining the two pre-orders together, a new complete pre-order is got as follows: $A_2=A_5>A_3= A_1> A_4$. Now the best ranked alternatives are A_2 and A_5 together, a contradiction from the previous result which had A_2 as the only optimal alternative.

4.2 An Example of Rank Reversal with the ELECTRE III Method

This illustrative example is also based on a real-life decision problem which was defined on 11 alternatives and 11 decision criteria. The goal of this case is to choose the best waste incineration strategy for the eastern Switzerland region [Rogers, *et al.*, 1999]. In this example, the first test criterion reveals a case of rank reversal when the ELECTRE III method is used. The main data are as follows:

	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}
A_1	125	866	9.81	218	1.41	542	483	23	1.5	1	1
A_2	11,980	900	11.45	189	1.45	452	303	12	1.5	6	6
A_3	31,054	883	9.86	172	1.82	341	311	0	0	3	3
A_4	28,219	840	10.38	171	1.95	339	318	0	0	3	3
A_5	31,579	903	10.74	165	1.7	312	281	0	0	5	5
A_6	39,364	922	13.87	167	1.65	287	269	0	0	8	7
A_7	125	769	9.33	182	1.64	458	180	0	1.5	1	1
A_8	8,075	896	9.82	172	1.7	408	121	0	1.5	6	6
A_9	3,089	770	9.39	177	1.9	430	228	0	1	2	2
A_{10}	6,449	766	7.22	172	1.65	401	157	0	1	4	4
A_{11}	12,074	897	10.61	169	1.65	378	162	0	1	7	6

Please note that in this example, criteria C_2, C_6 and C_7 are *benefit* criteria, which means the higher the score of a given criterion is, the more preferable it is. The other criteria are *cost* criteria, which means the lower the score of a given criterion is, the more preferable it is.

The weights W, the indifference thresholds Q and the preference thresholds P of the criteria are as follows:

	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}
W	0.16	0.033	0.033	0.097	0.097	0.16	0.097	0.16	0.033	0.033	0.097
Q	$\pm 1,000$	10%	10%	± 5	10%	10%	10%	± 2	0.2	± 0	± 0
P	$\pm 2,000$	20%	20%	± 10	20%	20%	20%	± 4	0.4	± 1	± 1

*Please note that in this example, no veto threshold is specified as it is the case in the above cited paper.

Next the concordance index $C_i(a, b)$ calculated for each pair of alternatives (a, b) in terms of each one of the decision criteria according to the following formula:

$$C_i(a,b) = \begin{cases} 1, & if \quad z_i(a) + q_i(z_i(a)) \geq z_i(b) \\ 0, & if \quad z_i(a) + p_i(z_i(a)) \leq z_i(b) \end{cases} \qquad (4\text{-}3)$$

or by linear interpolation between 0 and 1 when $z_i(a)+q_i(z_i(a))<z_i(b)<z_i(a)+p_i(z_i(a))$, where $q_i(.)$ and $p_i(.)$ are the indifference and preference threshold values for criterion C_i [Belton and Stewart, 2001]. For instance, the concordance indices in terms of the first decision criterion, which is $C_1(a, b)$, are as follows:

	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}
A_1	—	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
A_2	0.00	—	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00
A_3	0.00	0.00	—	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00
A_4	0.00	0.00	1.00	—	1.00	1.00	0.00	0.00	0.00	0.00	0.00
A_5	0.00	0.00	1.00	0.00	—	1.00	0.00	0.00	0.00	0.00	0.00
A_6	0.00	0.00	0.00	0.00	0.00	—	0.00	0.00	0.00	0.00	0.00
A_7	1.00	1.00	1.00	1.00	1.00	1.00	—	1.00	1.00	1.00	1.00
A_8	0.00	1.00	1.00	1.00	1.00	1.00	0.00	—	0.00	0.37	1.00
A_9	0.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00	—	1.00	1.00
A_{10}	0.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00	0.00	—	1.00
A_{11}	0.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	—

The next step is to calculate the discordance index $D_i(a, b)$ for all the alternatives in terms of each one of the decision criteria according to the following formula:

$$D_i(a,b) = \begin{cases} 0, & if \quad z_i(b) \leq z_i(a) + p_i(z_i(a)) \\ 1, & if \quad z_i(b) \geq z_i(a) + t_i(z_i(a)) \end{cases} \qquad (4\text{-}4)$$

or by linear interpolation between 0 and 1 when $z_i(a)+p_i(z_i(a))<z_i(b)<z_i(a)+t_i(z_i(a))$, where $t_i(.)$ is the veto threshold for criterion C_i [Belton and Stewart, 2001]. If no veto threshold is specified, then $D_i(a, b)=0$ for all pairs of alternatives. For instance, in this example, since no veto thresholds are specified, the discordance indices in terms of each decision criterion are all equal to zero.

The next step is to calculate the overall concordance index $C(a, b)$ of all the alternatives by applying the following formula:

$$C(a,b) = \frac{\sum_{i=1}^{m} w_i C_i(a,b)}{\sum_{i=1}^{m} w_i}. \qquad (4\text{-}5)$$

Finally, the credibility matrix $S(a, b)$ of all the alternatives is calculated by applying the following formula:

$$S(a,b) = \begin{cases} C(a,b), & if \quad D_i(a,b) \leq C(a,b) \forall i \\ C(a,b) \prod_{i \in J(a,b)} \frac{(1-D_i(a,b))}{(1-C_i(a,b))}, & otherwise \end{cases} , \qquad (4\text{-}6)$$

where $J(a, b)$ is the set of criteria for which $D_i(a, b)>C(a, b)$. The credibility matrix is a measure of the strength of the claim that "alternative a is at least as good as alternative b". For this case, the credibility matrix is equal to the concordance matrix since no veto thresholds are assigned so the discordance matrices are all zero matrices, which results to $S(a, b) = C(a, b)$ and both are as follow in next page.

Next the descending and ascending distillations [Belton and Stewart, 2001; Rogers, *et al.*, 1999] based on the credibility matrix are applied to construct two pre-orders for the

alternatives. The pre-order obtained from the descending distillation is as follows: $A_9> A_4> A_7> A_{10}> A_3= A_5= A_8= A_{11}> A_1> A_2> A_6$. The pre-order obtained from the ascending distillation is as follows: $A_1= A_7> A_9> A_4> A_{10}> A_2= A_5> A_3= A_{11}> A_8> A_6$. Then the two pre-orders are combined to get the final overall ranking of the alternatives as shown in Figure 1.

	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}
A_1	0.00	0.74	0.71	0.71	0.71	0.71	0.74	0.74	0.71	0.68	0.71
A_2	0.44	0.00	0.57	0.58	0.58	0.71	0.48	0.57	0.39	0.39	0.71
A_3	0.36	0.58	0.00	0.84	0.96	1.00	0.55	0.69	0.55	0.69	0.83
A_4	0.36	0.58	1.00	0.00	0.95	0.95	0.49	0.65	0.55	0.62	0.76
A_5	0.38	0.63	0.86	0.68	0.00	1.00	0.54	0.68	0.54	0.52	0.68
A_6	0.38	0.48	0.48	0.47	0.68	0.00	0.52	0.52	0.52	0.52	0.52
A_7	0.72	0.86	0.76	0.77	0.75	0.74	0.00	0.88	0.87	0.84	0.85
A_8	0.38	0.84	0.74	0.74	0.70	0.87	0.58	0.00	0.58	0.61	0.87
A_9	0.35	0.78	0.85	0.85	0.74	0.73	0.67	0.97	0.00	0.94	0.89
A_{10}	0.40	0.81	0.72	0.74	0.81	0.84	0.60	0.98	0.61	0.00	0.98
A_{11}	0.41	0.70	0.74	0.74	0.74	0.87	0.53	0.81	0.55	0.68	0.00

The way to combine the two pre-orders is the same as that of ELECTRE II, which has been described in the first example. The arrow line in Figure 1 means 'outrank'. For example, A_9 outranks A_4. Two alternatives are incomparable if there is no direct or indirect arrow line to link them together. For example, A_7 and A_9 are incomparable. Alternatives are indifferent if they are at the same level. For example, A_3 and A_{11} are indifferent with each other. It can be seen now that A_7 and A_9 are both located at the top level and they are incomparable with each other. Incomparability may be caused by the lack of the criterion information of the alternatives. This means that there is no clear evidence in favor of either A_7 or A_9. In real-life applications of ELECTRE II and III methods, the decision analysts will need to find more information about such alternatives and do a further study to decide which one is the best one. However, for the simplicity of the current test, A_7 and A_9 are both regarded as the best-ranked alternatives because both of them are ranked first in the final partial pre-order. As a result, the rest of the alternatives were regarded as the non-optimal ones.

Next, according to the first test criterion, we randomly selected one of the non-optimum alternatives; say alternative A_1 to be replaced by a worse one $A_1^/$ to test the reliability of the alternatives' ranking. Since the performance value of A_1 in terms of each criterion was [125 866 9.81 218 1.41 542 483 23 1.5 1 1], we subtracted [-2,000 0 0 0 0 90 0 0 0 0 0] (the subtracted value was selected randomly by a computer program to make sure it will make the chosen alternative to become worse than before) from A_1 to get $A_1^/$ which is less desirable than A_1. Please recall that the first criterion is a *cost* criterion which means the bigger a score the less desirable it is. The performance values of $A_1^/$ are [2,125 866 9.81 218 1.41 452 483 23 1.5 1 1]. The new decision matrix with the a_{ij} values of the alternatives (after alternative A_1 is replaced by $A_1^/$) is as follows:

	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}
$A_1^/$	**2,125**	**866**	**9.81**	**218**	**1.41**	**452**	**483**	**23**	**1.5**	**1**	**1**
A_2	11,980	900	11.45	189	1.45	452	303	12	1.5	6	6
A_3	31,054	883	9.86	172	1.82	341	311	0	0	3	3
A_4	28,219	840	10.38	171	1.95	339	318	0	0	3	3
A_5	31,579	903	10.74	165	1.7	312	281	0	0	5	5
A_6	39,364	922	13.87	167	1.65	287	269	0	0	8	7
A_7	125	769	9.33	182	1.64	458	180	0	1.5	1	1
A_8	8,075	896	9.82	172	1.7	408	121	0	1.5	6	6
A_9	3,089	770	9.39	177	1.9	430	228	0	1	2	2
A_{10}	6,449	766	7.22	172	1.65	401	157	0	1	4	4
A_{11}	12,074	897	10.61	169	1.65	378	162	0	1	7	6

The rest of the data are kept the same. When the previous steps are applied on the modified problem again, we get that the descending pre-order is $A_7 > A_9 > A_4 > A_{10} > A_3 = A_5 = A_8 = A_{11} > A_1 > A_2 > A_6$ and the ascending pre-order is $A_7 > A_1 = A_9 > A_4 > A_{10} > A_2 = A_5 > A_3 = A_{11} > A_8 > A_6$. The overall ranking of the alternatives is as shown in Figure 2. This time it turns out that the best-ranked alternative now is only A_7 which is different from the original conclusion which had A_7 and A_9 as the best-ranked alternatives.
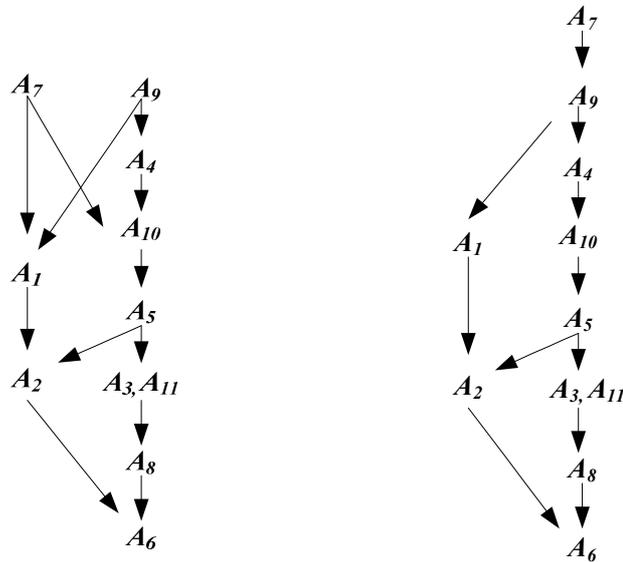
Figure 1. Ranking for the original example.   Figure 2. Ranking for the changed example.

Why did the above contradictions occur? When A_1 was replaced by a worse one, it is reasonable to assume that some alternatives which originally are ranked lower than A_1 may become more preferable than it. However, there is no legitimate reason why the optimal alternative should also be changed and why the original incomparable relation between two equally ranked alternatives should also be changed.

5. Analysis of the Rank Reversals with the ELECTRE II and III Methods

After analyzing the ranking processes of the ELECTRE II and III methods and some rank reversal cases which occurred when these methods were used, it was found that the main reason for the above rank reversals lies in the exploitation of the pairwise outranking relations. That is, the upward and downward distillation processes of ELECTRE II and ELECTRE III. The basic idea behind the distillation processes is to decide the rank of each alternative by the degree of how this alternative outranks all the other alternatives. When a non-optimal alternative in an alternative set is replaced by a worse one, the pairwise outranking relations related to it may be changed accordingly and the overall ranking of the whole alternative set, which depends on those pairwise outranking relations, may also be changed. The first change is reasonable when considering the fact that a non-optimal alternative has been replaced by a worse one. However, the second change is unreasonable and may cause undesirable rank reversals as in the examples presented in Chapter 3.

As it is demonstrated in the next chapter, where one decomposes a decision problem into smaller problems and analyzes them by using the ELECTRE II or III method, the rankings of the smaller problems may not follow the transitivity property. This fact along with the above rank reversal examples reveals that there is not an priori ranking of the alternatives when they are ranked by the ELECTRE II or III methods because the ranking of an individual alternative derived by these methods depends on the performance of all the other alternatives currently under consideration. This causes the ranking of the alternatives to depend on each other. Thus, it is likely that the optimal alternative may be different and the ranking of the alternatives may be distorted to some extent if one of the non-optimal alternatives in the alternative set is replaced by a worse one.

This can be further explained by means of a simple example. Given three alternatives: A_1, A_2, and A_3, suppose that originally A_1 strongly outranks A_3, A_2 weakly outranks A_3 and A_1 and A_2 are indifferent with each other. The ranking of these three alternatives will be $A_1 > A_2 > A_3$ when using the ELECTRE II method. Next, if the non-optimal alternative A_3 is replaced by a worse one, then A_2 may strongly outrank A_3 while A_1 is still strongly outranking A_3 and A_1 is still indifferent with A_2. Nothing is wrong so far. But now the ranking of the three alternatives will be $A_1 = A_2 > A_3$ by using the same method since both A_1 and A_2 now strongly outrank A_3 and they are indifferent with each other. It can be seen that A_1 and A_2 are ranked equally now because A_3 becomes less desirable. This is exactly what happened in the first example: A_2 and A_5 are ranked equally after A_3 has been replaced by a less desirable alternative. This kind of irregular situation is undesirable for a practical decision-making problem though it is reasonable in terms of the logic of the ELECTRE II method. It could leave the ranking of a set of alternatives to be manipulated to some extent.

The ranking irregularity in the above example is very likely to occur when using the ELECTRE II or III method to rank a set of alternatives. If the number of alternatives of a decision problem is more than 3, there will be more than C_3^2 (=6) pairwise outranking relations between them. Then the situation may become worse by totally changing the indication of the best ranked alternative. It was once pointed out in [Belton and Stewart, 2001] that the results of the distillations are dependent on the whole alternative set, so that the addition or removal of an alternative may alter some of the preferences between the remaining alternatives. A similar situation occurs with the PROMETHEE method which is another variant of the outranking method. In [Keyser and Peeters, 1996], it was pointed out that the complete pre-orders from the PROMETHEE method are based on an all-to-all comparison between the alternatives; adding or deleting an alternative can put the previous pre-orders upside down. From the study reported in this thesis, now it can be seen that a similar situation also occurs with the ELECTRE II and III methods. That is, even without the addition or removal of alternatives, the best ranked alternative might be altered and the

previous pre-order between the remaining alternatives might be changed to some degree by just replacing a non-optimal alternative by a worse one.

It must be pointed out here that there is another factor that may contribute to rank reversals. During the construction of the pairwise outranking relations, both ELECTRE II and III need to use a value or a threshold which is also dependent on the performance values of all the currently considered alternatives. For ELECTRE II, it is the parameter δ (i.e., the maximum difference of any criterion) in the discordance index formula. For ELECTRE III, it is the parameter λ used to decide the $\lambda - preference$ relations between the alternatives during the distillations. These δ and λ values may be altered when a non-optimal alternative is replaced by a worse one. Then the previous outranking relations between the other unchanged alternatives may be distorted to some degree, which finally may alter the indication of the best ranked alternative or the overall ranking of the alternatives. According to some experimental analysis, the above two factors may function together or separately to cause rank reversals.

From the above analysis, it can be seen that the ranking processes of the ELECTRE II and III are not reliable and robust enough to offer a firm answer to a decision problem. Usually, decision makers undertake some kind of sensitivity analysis to appreciate the sensitivity of the final rankings and the robustness of the ranking procedures to changes in the criteria weights and thresholds when they use ELECTRE methods to solve decision problems. However, the above ranking irregularities can warn decision analysts that they should be cautious in accepting the ranking recommendations of the ELECTRE methods even after a careful sensitivity analysis is undertaken.

6. An Empirical Study

This chapter describes an empirical study that focused on how often these ranking irregularities may happen under the ELECTRE II and III methods. Some computer programs were written in MATLAB in order to generate simulated decision problems and test the performance of ELECTRE II and III under the three test criteria described in Chapter 3. In these test problems, the number of the alternatives was equal to the following ten different values: 3, 5, 7, 9, 11, 13, 15, 17, 19, and 21. However, there is not a common range of the criteria for all the tests. Compared with the tests of ELECTRE II, a wider range of criteria for the tests of ELECTRE III was needed in these experiments in order to clearly show how the ranking irregularity rates under ELECTRE III will fluctuate with the increase on the number of the criteria. For the three tests of ELECTRE II, the number of criteria was equal to 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, and 31. Thus, a total of 150 different cases were examined with 10,000 randomly generated decision problems (in order to derive statistically significant results) per case. For the three tests of ELECTRE III, the number of criteria was equal to the odd numbers between 3 and 61. Thus, a total of 300 different cases were examined with 10,000 random decision problems per case. Each random decision problem was analyzed first by using the ELECTRE II or III method and then was analyzed again by using the same method after one of the non-optimal alternatives was replaced by a worse one or the whole decision problem was decomposed into smaller problems as described in the last two test criteria. Any occurred ranking irregularity was recorded. Figures 3 to 8 summarize these test results. In these figures, different curves correspond to cases with different numbers of alternatives; the X axis stands for the number of criteria and the Y axis is the rate of ranking irregularities that occurred in the 10,000 simulated decision problems.

Figures 3 and 4 describe how often rank reversal happened with the ELECTRE II and III methods under test criterion #1 in this empirical study. That is, how often the indication of the best alternative is changed when a non-optimal alternative is replaced by another worse alternative (given that the relative importance of each decision criterion remains unchanged). The basis of any ELECTRE method is to decide the pairwise outranking relations between the alternatives. Given n alternatives, when a non-optimal alternative was replaced by a worse one, the number of pairwise outranking relations that might be changed is at most $(n-1)$. This indicates that the higher the number of the alternatives is the more possible becomes that the variation of a single alternative may have a noticeable influence on the outranking relations. It is the change of the outranking relations that results in the rank reversals. This is why the rank reversal rates usually increase with the increase on the number of alternatives.

It should be clarified here that even if a case passed test criterion #1, this does not mean that this case is immune to the rank reversal situation described in test criterion #1. When applying test criterion #1, one non-optimal alternative needs to be picked up and replaced by a worse one. Which non-optimal alternative will be selected and how worse it could be to trigger the rank reversal to happen were all randomly chosen by the program. When replacing a non-optimal alternative by a worse one, the program only makes the selected non-optimal alternative to be worse than before to a certain degree to test if the change is enough to trigger a rank reversal to occur. If no rank reversal happens, the case will be released and marked as having passed test criterion #1. It is not possible to test all the possibilities in terms of a given single case. Therefore, even if a case passed test criterion #1 in a single experiment, this does not mean that it is immune to the type one rank reversal.

Figures 5 and 6 depict how the ranking irregularity rates of ELECTRE II and III varied with the increase on the number of alternatives and the number of criteria in terms of test criterion #2. One can see from these figures that the rates generally increase with the increase on the number of the alternatives. This happens because the higher the number of alternatives is, the higher is the number of smaller problems that a decision problem was

decomposed into, and then the more likely it is for a contradiction between the smaller problems to happen.

In order to apply test criterion #3, 10,000 random decision problems whose rankings follow the transitivity property by using the ELECTRE II or III method must be generated and then be examined under test criterion #3. However, as one can see from Figures 5 and 6, when the number of alternative is up to 7 or 9, the rankings of the random decision problems almost never follow the transitivity property when the number of criteria is in some range. It is difficult to find 10,000 random decision problems per case that can be used in the third series of tests. Thus, only the cases where the number of alternatives was equal to 3, 5 or 7 were tested. Figures 7 and 8 show how often the ranking irregularity will happen to these cases under test criterion #3. The reason why this rate increases with the number of alternatives is the same as that of the experiments under test criterion #2.

What is the relationship between these ranking irregularity rates with the number of the decision criteria? From these figures one can see that, in general, the ranking irregularity rates will first increase with an increase on the number of the criteria but then decrease when the number of criteria increases beyond a certain value for each case. The reason for this tendency can be explained as follows. First, please recall that the pairwise outranking relations between each pair of alternatives are decided by the concordance and discordance indices which are computed by using their performance values under each criterion. For a fixed number of alternatives, if the number of criteria is increased to a certain large value, the pairwise outranking relations between the alternatives and the subsequent ranking of the alternatives will tend to become more stable than before.

A different type of experiments was run as well. The goal now was to examine if there are any explicit connections between the results under the three test criteria, especially between test criterion #1 and test criterion #2. The experimental tests were executed as follows. First, a large number (i.e., 10,000) of randomly generated decision problems were examined by using the ELECTRE II or III method in terms of test criterion #1 and the random test problems were divided into two groups. One group had the problems that passed test criterion #1 and the other group had those that did not pass test criterion #1. Next, the problems within each of these two groups were examined in terms of the test criterion #2 and the rates of how often they passed or failed to pass this test criterion were recorded and plotted for each one of the two groups.

Next, a test process similar to the above one was performed. Again, a large number of randomly generated decision problems were examined by using the ELECTRE II or III method but now the process started by first testing for behavior under the test criterion #2. The problems were divided into two groups indicating passing or not passing this test criterion. Then the problems within each one of these two groups were examined in terms of the test criterion #1 and the rates of how often they will pass or fail to pass this test criterion are recorded and plotted as before for each one of these two groups. Similar tests as the above ones were also performed between test criterion #1 and test criterion #3. From the above experimental test results, no clear tendency was found to indicate that failure in one test criterion would have a tendency to lead to failure in terms of another test criterion. That is, not any explicit connection between the results under the three different test criteria was found from these types of experiments.
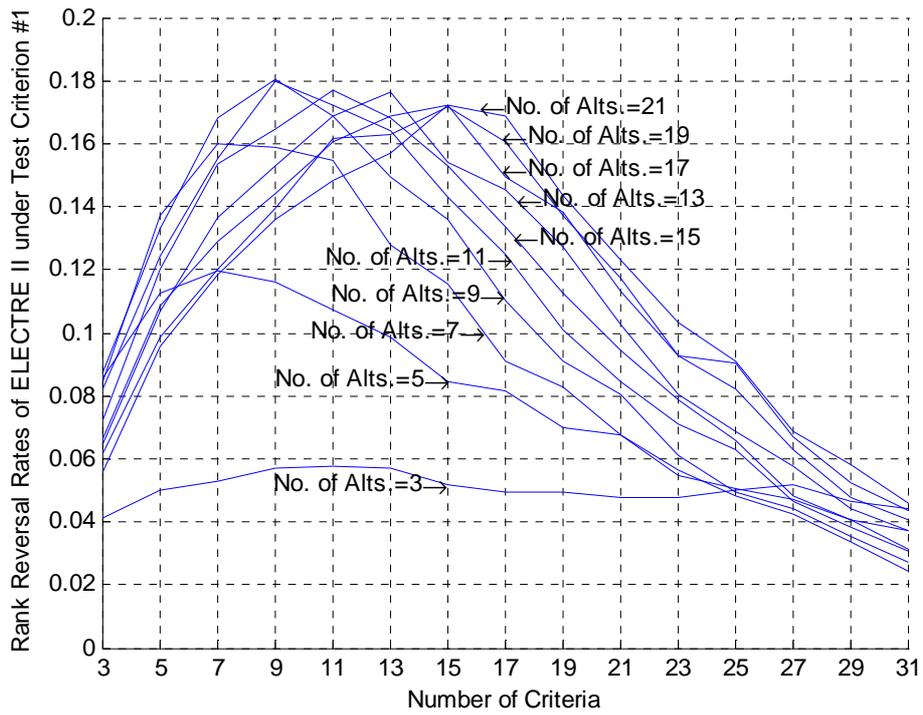
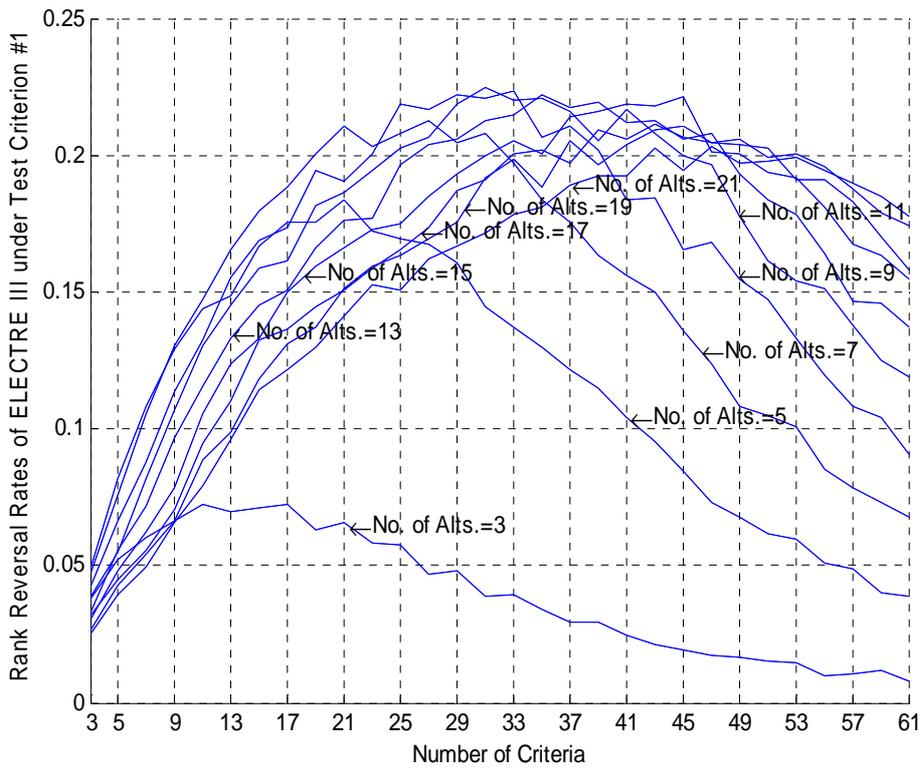Figure 3. Rank Reversal Rates of ELECTRE II under Test Criterion #1.

Figure 4. Rank Reversal Rates of ELECTRE III under Test Criterion #1.
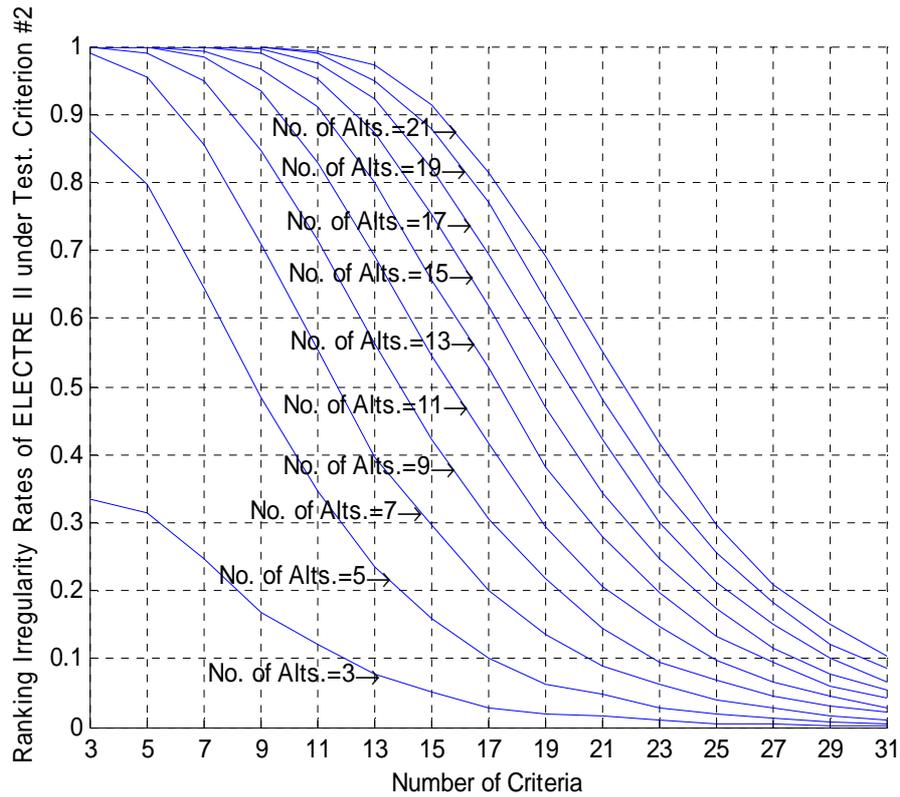
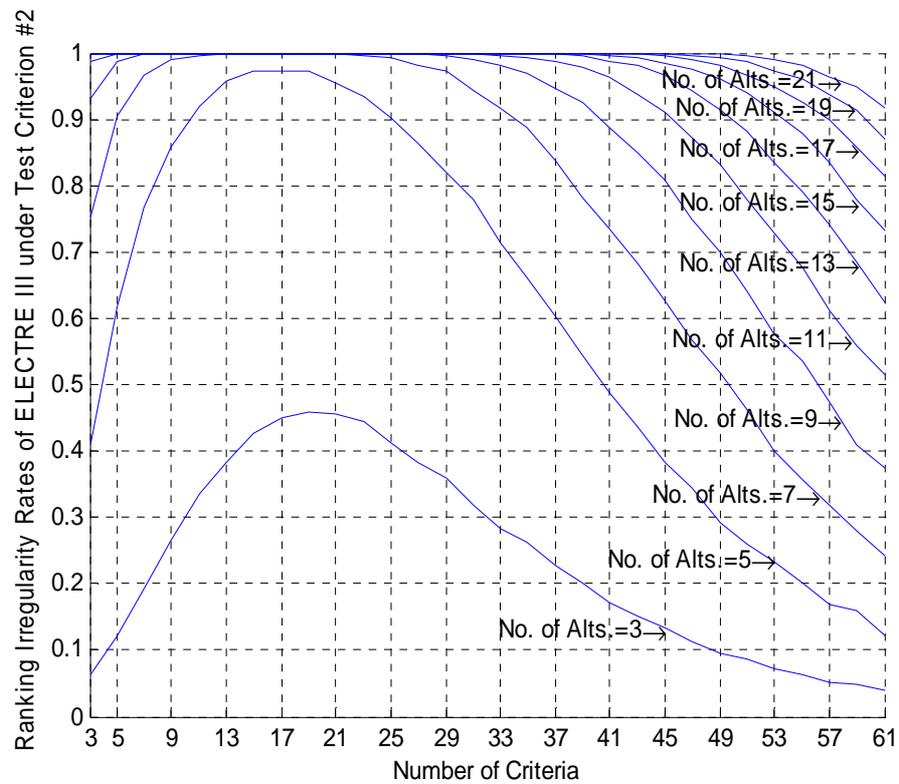Figure 5. Ranking Irregularity Rates of ELECTRE II under Test Criterion #2.

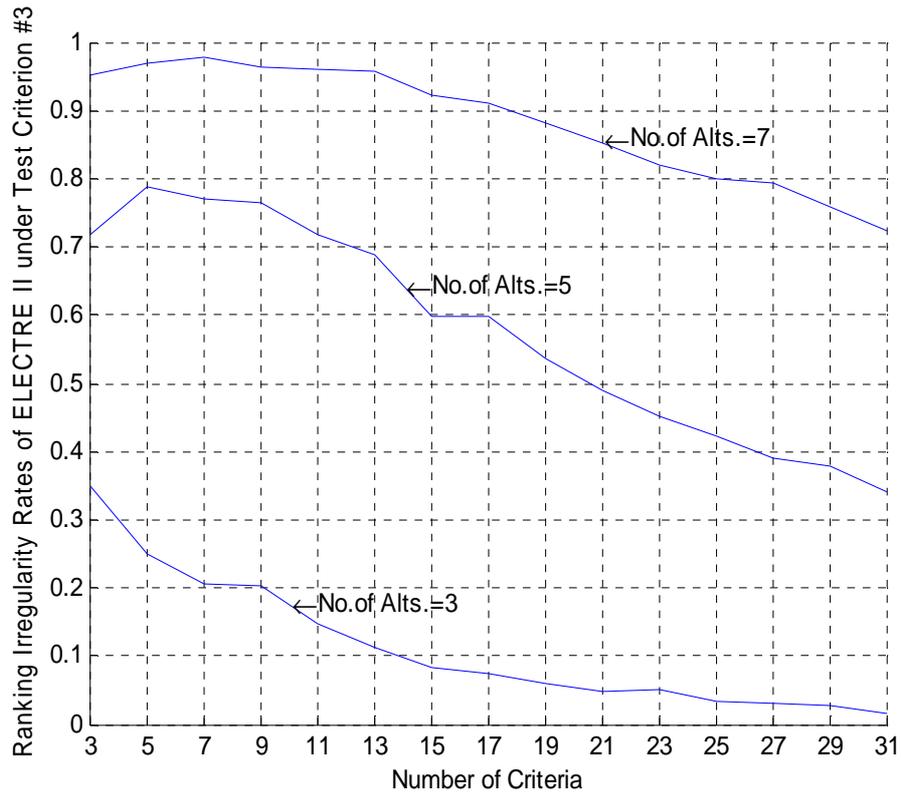Figure 6. Ranking Irregularity Rates of ELECTRE III under Test Criterion #2.

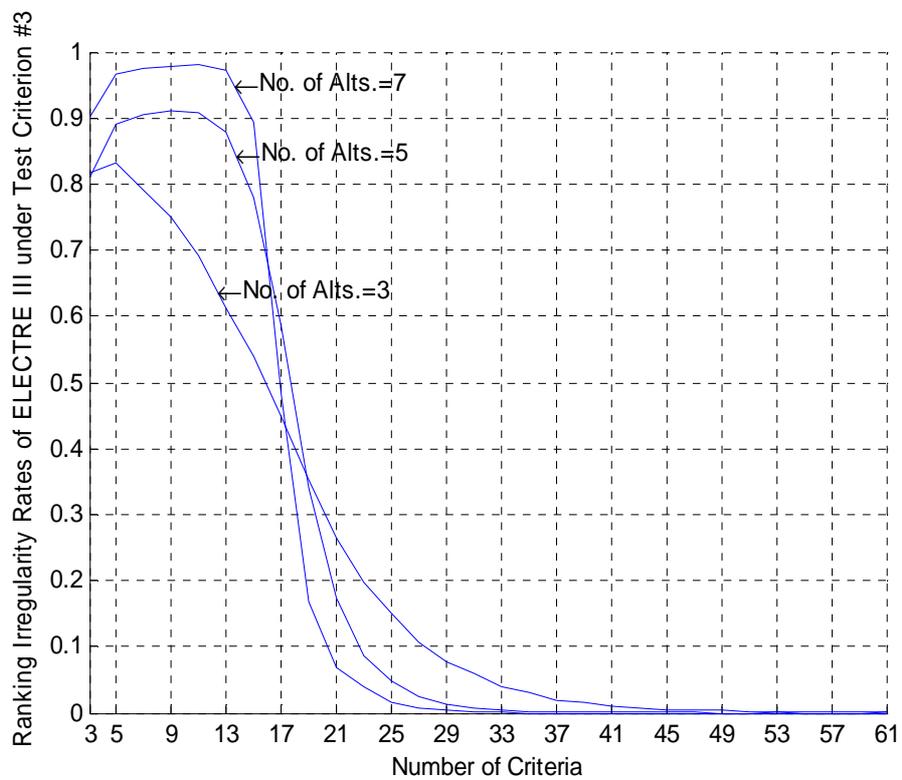Figure 7. Ranking Irregularity Rates of ELECTRE II under Test Criterion #3.

Figure 8. Ranking Irregularity Rate of ELECTRE III under Test Criterion #3.

23

7. Some Real-life Case Studies

The previous computational results revealed that the ranking irregularities studied in this research may occur frequently in simulated decision problems. This raised the question whether the same could also be true with real-life decision problems. In order to enhance our understanding of this situation, ten real-life cases were studied. These cases were selected randomly from the published literature. That is, no special screening was performed. The only requirement was to be able to extract the numerical data needed to form a decision matrix and the weights of the criteria. It is better if threshold values could be given in the published case to avoid the inconvenience with the using of the newly defined thresholds. In these experiments, the required thresholds for case 1 to case 8 have been specified in the referenced publications. For the last two cases, the thresholds were specified appropriately according to the score range of each criterion. After getting the data, every case was tested by using the ELECTRE II or III method as in the referenced publication. Then the three types of ranking irregularities were recorded whenever they occurred. Please refer to Table I for the summary of the experimental results. Actually, the two examples presented in Chapter 4 are among these 10 tested cases.

Under test criterion #1, that is, when replacing one of the non-optimal alternatives by a worse one, there are mainly two types of rank reversal situations:

1. The optimal alternatives of the changed decision problem are partially different from that of the original problem. The number of the optimal alternatives of the changed problem is more or less than that of the original problem. For example, in terms of case 7, originally the optimal alternative is A_8. Next the optimal alternatives may become A_7 and A_8 under test criterion #1; for case 2, originally the optimal alternatives are A_6 and A_3, and then it may be just A_6 after one of the non-optimal alternatives was replaced by a worse one.

2. The optimal alternative of the new problem is totally different from that of the original problem. For example, for case 8, originally the optimal alternative is A_9, and then it becomes A_7 when one of the non-optimal alternatives was replaced by a worse one; for case 6, originally the optimal alternative is A_4, it may become A_{10} and A_{18} under the first test criterion #1.

In terms of the same case, the above two situations might both happen or just one of them happened in the tests. The emphasis is that the indication of the best alternative had been changed for those cases if any of the two situations occurred to them. Then one can conclude that rank reversals occurred to those cases and they failed to pass test criterion #1. From Table I, it can be seen that 6 out 10 cases failed to pass test criterion #1. Also, 9 out 10 cases failed to pass test criterion #2. For the only case which could be tested under test criterion #3, it failed to pass it too.

Table I. Summary of Case Studies.

Case number	Reference	Domain of application and method used	Size of decision problem		Did it fail T. C. #1?	Did it fail T. C. #2?	Did it fail T. C. #3?
			No. of alternatives	No. of criteria			
1	Hokkanen, J., and P. Salminen, [1997a]	Choosing a solid waste management system (ELECTRE III)	22	8	Yes	Yes	—
2	Belton, V., and T.J. Stewart, [2001]	Business location problem (ELECTRE III)	7	6	No	Yes	—
3	Rogers, M., and M. Bruen, [1996]	Environmental appraisal (ELECTRE II)	9	9	No	No	Yes
4	Rogers, M.G., M. Bruen, and L.-Y. Maystre, [1999]	Site selection for a wastewater treatment plant (ELECTRE II)	5	7	Yes	Yes	—
5	Raj, P.A., [1995]	Water resources planning (ELECTRE II)	27	6	Yes	Yes	—
6	Buchanan, J., P. Sheppard, and D.V. Lamsade, [1999]	Project ranking (ELECTRE III)	5	5	No	Yes	—
7	Hokkanen, J., and P. Salminen, [1997b]	Choosing a solid waste management system (ELECTRE III)	11	8	Yes	Yes	—
8	Rogers, M.G., M. Bruen, and L.-Y. Maystre, [1999]	Choosing a waste incineration strategy (ELECTRE III)	11	11	Yes	Yes	—
9	Poh, K.L., and B.W. Ang, [1999]	Choosing an alternative fuel system for land transportation (ELECTRE II)	4	6	No	Yes	—
10	Leyva-López, J.C., and E. Fernández-González, [2003]	Selection of an alternative electricity power plant (ELECTRE III)	6	6	Yes	Yes	—

* "T. C." stands for "Test Criterion".
* If a case failed under test criterion #2, it will not be able to get an overall ranking of the alternatives from the smaller problems. That means it will not be able to apply test criterion #3 to that case. The symbol "—" was used in the corresponding cell to stand for this situation.

8. Some Ongoing Research on Regret and Rejoicing

Usually the problem of rank reversal is undesirable in decision-making. If a method does allow it to happen, the validity of the method should be questioned, like the ELECTRE II and III method and the AHP method. However, some recent studies showed that it is not always unreasonable to have such rank reversals happening with MCDM problems [Kujawski, 2005]. The critical question is to be able to distinguish why it happens. When a method exhibits rank reversals, is it because it accurately captured the way rational humans deal with decision making and their preferences' change or is it because the method has some kind of numerical instabilities/defects? Let us put it more clearly through a metaphor: suppose a method is like a photo camera or X-ray image taking machine. One takes a photo or takes an X-ray image of the body and sees something strange in that image, like some very bright spots. Do these bright spots exist in reality or are the result of some kind of hardware defects?

For different MCDM methods and decision models, the answer to the above question could be very different. As it was discussed above, the reasons for the rank reversal problems with ELECTRE II and III methods and the additive AHP method lie in these methods' own mathematical instabilities/defects. However, rank reversals could also happen because people may reverse their preference due to their emotional feelings, like regret and rejoicing. Next this point is illustrated by some vivid examples.

8.1 The Influence of Regret and Rejoicing on MCDM Problems

First, let us see how rejoicing may play a vital role in real-life decision making process. Suppose one is planning to buy a new car and a dealer offers two cars, say car A1 and car A2. In this hypothetical scenario car A1 is cheaper than car A2 but car A2 is of better quality than car A1. Then, one may decide to buy car A1 because it is cheaper. Next, suppose that besides the above two cars, the dealer introduces a third car A3 (let us call it a phantom alternative) which may not even be at stock at that dealership but it has been publicized by the media. This third car A3 is much more expensive than the previous two cars but it is of slightly better quality than car A2. Knowing this situation about the third car, the perspective buyer may shift his/her preference and now choose car A2 instead of car A1 without actually changing anything regarding the two cars and the importance of the two evaluation criteria: cost and quality. This may happen because when comparing car A2 with car A3, the buyer feels very happy for getting a great deal by paying much less money to buy an almost equal quality car A2. Thus for this example, it is this *rejoicing* feeling that makes one unintentionally to reverse his/her preference between cars A1 and A2.

Except rejoicing, another type of emotional feeling which can greatly influence people's preference in decision-making is *regret*. This type of emotional feeling comes from the fact that humans often base their choices on comparisons across the alternatives under consideration and relative to "what might have been" under another choice [Plous, 1993; Hastie and Dawes, 2001]. For example, given are two alternatives A_1 and A_2 which have been evaluated in terms of three criteria. Assume that by using one MCDM method, the overall performance value of A_1 is better than that of A_2 but the individual performance value a_{1k} of alternative A_1 under some criterion C_k is worse than that of alternative A_2. Then the decision maker (DM) who chooses A_1 and forgoes A_2 may experience a certain level of regret because the criterion value a_{1k} is worse than a_{2k}. This regret feeling could be so strong that he/she may regret to have chosen A_1 instead of A_2. In order to avoid the above situation, the decision maker would want to anticipate the regret feeling and consider it in the decision-making process by making some tradeoffs for a more balanced alternative.

From the above examples, it can be seen that making a choice/decision, no matter what kind of, can be an intensive emotional experience. When making decisions, except those cognitive considerations about the decision problems themselves, sometimes people also need to consider some intense emotional factors, like regret and rejoicing. Psychologically speaking humans often behave based on a combination of reasons and emotions. Thus it is natural that decisions could be made by the mind and also by the heart instead of by a complete rational mind which is dissociated from psychological feelings. Meanwhile, the above car example indicates that consideration of regret and rejoicing might be able to offer decision makers the flexibility to change their preferences and choices in a rational way. Thus, rank reversals could be justified as natural consequences of rational preference changes when there are no mathematical drawbacks involved in the ranking process.

8.2 Some Past Studies on Regret and Rejoicing

Fortunately, there have been many studies on regret and rejoicing in the literature. In [Sugden, 1985], regret is defined as "the painful sensation of recognizing that 'what is' compares unfavorably with 'what might have been'." The converse experience of a favorable comparison between the two is called "rejoicing". Some experimental studies confirm that for most individuals regret has the greater impact [Mellers, 2000]. In related research, regret is also the one that has received most of the attention. The following is a simple introduction to some regret models.

One of the earliest regret models is known as the minimax regret model which was introduced by Savage [Savage, 1951] and was first axiomatized by Milnor [Milnor, 1954]. The minimax regret is a strategy for decision-making under uncertainty whereby the DM chooses the alternative with the minimum worst possible outcome in order to minimize regret [Kujawski, 2005]. This model defines regret as the difference between the actual performance value of each decision alternative and the best possible value among all alternatives for each ultimate state of nature. Suppose the utility value of an alternative A_i under a state of nature S_k is u_{ik}. Then, the decision maker who chooses A_i will experience a level of regret R_{ik} for the state of nature S_k where R_{ik} is defined as $R_{ik}=\max_j(u_{jk})-u_{ik}$. Next the DM would first determine the possible highest level of regret that could occur to each decision alternative, and then chooses the alternative with the minimum of these maximum regret values [Zeelenberg, 1999]. Because this model decides the selection of alternatives totally by their regret values, it may lead to irrational choices. Such as, a small disadvantage in a single decision criterion, no matter how large/small its importance is, may eliminate alternatives with more preferable performance values under more important criteria [Kujawski, 2005]. Given this undesirable property, the minimax regret model has not been used widely.

Later, Loomes and Sugden and also Bell proposed a regret theory (referred as RT-B/LS) simultaneously in 1982 for rational decision-making under uncertainty [Loomes and Sugden, 1982; Bell, 1982 and 1985]. In the RT-B/LS model, regret is defined as a psychological reaction that is caused by comparing an outcome under one state with the payoff one could have had by making a different choice under the same state. The RT-B/LS model assumes that the levels of regret and rejoicing depend on the difference of the utilities between what is and what could have been. For example, the associated level of regret when comparing utility value u_{ik} with utility value u_{jk} is defined as follows [Kujawski, 2005]:

$$R(u_{ik}, u_{jk}) = \begin{cases} R(u_{jk} - u_{ik}), & u_{ik} < u_{jk} \\ 0, & otherwise. \end{cases}$$

Where u_{ik} is the classical utility of the *i-th* alternative in terms of the *j-th* criterion, and *R(.)* is a non-decreasing regret function.

However, it was illustrated in [Kujawski, 2005] that the RT-B/LS model exhibited intransitivity under pairwise comparisons and inconsistencies with some empirical evidence. To solve these problems, a new regret model called the Reference-Dependent Regret Model (RDRM) was proposed for deterministic decision making. Kujawski illustrated that, in general, a person's level of regret when he/she chooses a multi-attribute alternative often depends explicitly on the absolute values of the utilities of the chosen and forgone alternatives (i.e., alternatives that were considered but not chosen) rather than simply their differences. Thus, in his RDRM model, the anticipated regret when choosing u_{ik} and forgoing u_{jk} is defined as follows:

$$R(u_{ik}, u_{jk}) = \begin{cases} G(1-u_{ik}) - G(1-u_{jk}), & u_{ik} < u_{jk} \\ 0, & \textit{otherwise.} \end{cases}$$

Where the utility value is between 0 and 1, and $G(.)$ is the regret-building function which measures the level of regret referenced to the maximum possible utility normalized to 1 and is defined as follows in [Kujawski, 2005]:

$$G(x) = \begin{cases} \dfrac{1}{1+(B/x)^{2 \times S \times (B+x)}}, & x > 0 \\ 0, & \textit{otherwise.} \end{cases}$$

The two parameters B and S in the definition of $G(.)$ are determined by querying the decision maker about the levels of regret that he/she experiences for each criterion [Kujawski, 2005]. It can be seen that regret occurs when a given decision alternative is compared with another one which has at least one better criterion value. It is claimed that the RDRM model ensures the transitive pairwise rankings of three alternatives with any number of criteria ≥ 3 because it satisfies a special property of additive transitivity [Kujawski, 2005].

Among the previous regret models, the minimax and the RT-B/LS regret models are originally developed for decision making under uncertainty. However, both of them can be tailored to be used in deterministic decision-making problems by identifying the states of nature with the criteria of a given MCDM problem. As stated in [Kaliszewski and Michalowski, 1998], the notion of regret becomes meaningful in deterministic multi-criteria decision problems if a notion of state is equated to a notion of attribute, and a state /attribute matrix conveys regret type information (for instance, the difference between ideal and actual values of attributes). For example, the notion of regret in the RDRM model is defined as follows by tailoring Bell's [1982] notion of anticipated regret for decision-making under uncertainty: in the process of choosing a deterministic alternative, a rational individual may decide to trade off some benefits and forgo the alternative with the highest total value for a more balanced alternative in order to reduce his/her level of anticipated regret [Kujawski, 2005].

It needs to be noted that the effect of anticipated regret/rejoicing is different from the experienced emotions. In deterministic decision-making situations, decision makers do not have to experience the emotions in order to be influenced by them. Rather, they can predict the emotional consequences of different decision outcomes in advance, and opt for the choices that minimize the possibility of negative emotions. Please also note that both the RT-B/LS and the RDRM models may still exhibit rank reversal problems, but the occurrence of a rank reversal is defended as rational for the above two models because they reflect an individual's change of preference in response to rational emotions like the anticipated regret/rejoicing.

8.3 The Need to Consider Regret and Rejoicing in MCDM

From the above discussion it is clear that the introduction of regret and rejoicing into the decision-making process is based on two key assumptions: 1) people experience the

sensations of regret and rejoicing which can influence their current decision making; 2) while making decisions people may try to anticipate and take into account feelings like regret and rejoicing [Loomes and Sugden, 1982]. Therefore, building an MCDM model that incorporates these emotional effects not only provides for a better description of the human behavior in decision making, but also offers the DMs the flexibility to trade off some economic benefits explicitly in order to gain a state of psychological satisfaction, for prescriptive purposes [Bell, 1985].

However, the study of how to incorporate the notion of regret and rejoicing in deterministic MCDM models is just at its very beginning. Besides several tentative works on this direction [Kujawski, 2005; Kaliszewski and Michalowski, 1998], till now there has not been a comprehensive MCDM method that can incorporate the notion of regret and rejoicing systematically in the MCDM models for both *benefit* and *cost* criteria. Meanwhile, such a model should be able to rank decision alternatives in a mathematically stable way. To satisfy this requirement, the model should not suffer of the ranking irregularity problems which happen with the additive AHP and ELECTRE II and III methods when emotional factors are not present to justify those ranking problems. Thus, how to design such an effective regret and rejoicing based MCDM model and to which degree these emotional factors should be considered in decision-making problems is the subjects of following considerations.

One may think that the RDRM regret model could fill in this void. However, there are some constraints about this model which refrain it from being used in general MCDM problems. First of all, utility is an elusive concept for many people. It is not always easy for decision makers to find a proper utility function that can appropriately transform criteria values with different units into unitless and additive utility values. To solve a decision problem, using the original criteria values might be more direct and convenient for decision makers. Second, in the RDRM model, the level of regret not only depends on the utilities of the chosen and forgone alternatives but also on the maximum possible utilities. For the notion of utility, the maximum value is fixed and is always equal to 1. However, for a general MCDM problem, the maximum possible values are not always well-defined. Then the ranking result from this method might be arbitrary. One may think that, instead of using the maximum possible criteria values, we can use the maximum criteria values for all currently considered alternatives in the RDRM model. However, in this way the ranking result from this method may still be unstable. This happens because the maximum values might be changed when adding or deleting alternatives, like the normalization factor with the additive AHP method might be changed when adding or deleting alternatives. Thus, a new MCDM model which does not have the above constraints, while has the ability to deal with the effects of regret and rejoicing is long due.

8.4 A Potential Regret and Rejoicing Based MCDM Method

As it has been described previously, one significant factor in people's decision making process is their capacity to anticipate feelings of regret and rejoicing and these anticipated feelings may strongly influence people's choices. Therefore, the goal of the next step is to design a new MCDM model that can incorporate the behavioral notion of anticipated regret and anticipated rejoicing systematically in the MCDM modeling framework. Meanwhile, this model should have some essential properties which current MCDM methods do not possess.

Based on some ongoing studies, the new MCDM model is expected to satisfy the following conditions:
1) Besides the usual benefit and cost criteria, the new model should be able to incorporate the effects of regret and rejoicing for decision makers who value these emotional factors in multi-criteria decision making situations.

2) The determination of regret and rejoicing effects should be consistent with the fundamentals of behavioral science on this subject.

3) Rank reversals may occur only as results of readjusting the effects of regret and/or rejoicing when the set of the alternatives is altered.

4) The model should be able to deal with qualitative and quantitative criteria expressed in different units of measurement.

5) The model should not exhibit cyclic preferences when tested under certain special test problems.

From condition (1) it follows that the effects of regret and rejoicing need to be determined for both benefit and cost criteria. Condition (2) means that the determination of the regret/rejoicing effects should follow some of the fundamental studies on people's emotional feelings in behavioral science. Condition (3) may have the following implication: if one temporarily ignores the regret / rejoicing effects, then the remaining part of an effective MCDM model should not exhibit any undesirable rank reversals when the set of the alternatives is altered. The effects of regret and rejoicing may be ignored if, for instance, their presence could be considered negligible when compared to the usual performance values of the alternatives under the benefit and cost criteria. Condition (4) dictates that an MCDA model should be able to transform measurements expressed in different units into dimensionless ones or deal with such different measures in computationally valid ways (i.e., not to "add oranges to apples"). Condition (5) means that the new model should not allow cyclic preference to happen to symmetric decision problems.

As described before, rank reversals may happen with some additive models (such as the ELECTRE II and III models and the additive AHP models) when one considers benefit and cost criteria only (i.e., without the regret or rejoicing effects). While some previous studies [Triantaphyllou, 2000 and 2001] had found that the multiplicative AHP are immune to these ranking problems. Thus, instead of addition the new model will be based on multiplication. In this way no matter how the decision matrix is normalized, the ratios of the alternatives' performance values will be kept the same because the normalization factor will be cancelled off by using the multiplicative formula. Thus, for decision making problems, all variation left is due to the regret and rejoicing effects and that could be justifiable as it would not be due to any mathematical artifacts.

Based on the above points, after considering the anticipated regret and rejoicing for both benefit and cost criteria for a given MCDM problem which has m alternatives and n decision criteria, the equation for computing the final overall performance of each alternative by using the multiplicative formula and the benefit to cost approach to deal with conflicting criteria could be as follows:

$$P_i^* = \frac{P_i^B J_i^B J_i^C}{P_i^C R_i^B R_i^C} = \frac{\prod_{k=1}^{n_1} a_{ik}^{w_k} \times \prod_{k=1}^{n_1} J_{ik}^{w_k} \times \prod_{k=n_1+1}^{n} J_{ik}^{w_k}}{\prod_{k=n_1+1}^{n} a_{ik}^{w_k} \times \prod_{k=1}^{n_1} R_{ik}^{w_k} \times \prod_{k=n_1+1}^{n} R_{ik}^{w_k}}.$$

Where, w_k is the weight of the *k-th* decision criterion; a_{ik} is the performance value of the *i-th* alternative in terms of the *k-th* criterion.

R_{ik} is the anticipated regret of the *i-th* alternative in terms of the *k-th* criterion;

J_{ik} is the anticipated rejoicing of the *i-th* alternative in terms of the *k-th* criterion;

J_i^B is the aggregated anticipated rejoicing of alternative A_i under the benefit criteria;

R_i^B is the aggregated anticipated regret of alternative A_i under the benefit criteria;

P_i^B is the aggregated performance value of the alternative A_i under the benefit criteria; Similarly, J_i^C, R_i^C and P_i^C has the corresponding meaning as the above ones but now in terms of the cost criteria. Furthermore,

$$R_{ik}^{w_k} = \prod_{l=1}^{n}[R(u_{ik}, u_{lk})]^{\frac{w_k}{n-1}} \text{ and } J_{ik}^{w_k} = \prod_{l=1}^{n}[J(u_{ik}, u_{lk})]^{\frac{w_k}{n-1}}.$$

Where $R(u_{ik}, u_{lk})$ and $J(u_{ik}, u_{lk})$ are the anticipated regret and rejoicing for a decision maker when he/she chooses the *i-th* alternative and forgoes the *j-th* alternative in terms of the *k-th* criterion. Meanwhile, it is assumed that the first n_1 criteria are benefit criteria and the $(n_1 + 1) - th$ criterion to the $n - th$ criterion is cost criterion.

In the above equation, a key issue is how to measure regret and rejoicing. One commonly used way is to measure them by using a continuous regret function as in the RT/B-LS and RDRM model. However, the approach of using such a function may have some fundamental weaknesses. First of all, the definition of such functions involves the determination of certain customizing parameters (like the parameters *B* and *S* in the RDRM model), as not all decision makers may behave in exactly the same way. Furthermore, it is not always clear how such parameters may be determined and whether such functions and their parameters should change from one criterion to another criterion within the same decision problem and the same decision maker.

Please recall that decision criteria may be quantitative (such as cost, age, weight, volume, etc) or qualitative (such as desirability, aesthetic appeal, style, etc). Regret and rejoicing are definitely qualitative aspects in decision problems. For qualitative aspects an approach proposed by Saaty (as part of the AHP method) [Saaty, 1980; Saaty, 1994] has received widespread attention for dealing with qualitative criteria. That approach is based on a ratio scale. That is, a decision maker is asked to select a linguistic statement that best describes his/her assessment of the relative importance of two alternatives when they are considered in terms of a single criterion at a time. In this way, the decision maker has to make *n*(*n*-1)/2 pairwise comparisons. For each such comparison, the decision maker selects the best linguistic statement from a small number of statements (9 to be exact) that best describes a given pairwise comparison. This is the concept of the Saaty scale. Each linguistic statement is also associated with a numerical value in a way that attempts to reflect the natural importance of them. The numerical values are the numbers 1, 3, 5, 7, and 9 and their intermediate values when the decision maker feels that the best answer lies between two successive linguistic choices. The reciprocals of the previous numbers are used too.

Following the same logic as with Saaty's scale, a similar approach could be used for assessing the regret and rejoicing values of the alternatives in terms of a single criterion. Where a decision maker is asked to select a linguistic statement that best describes his/her assessment of the regret and rejoicing when he/she chooses one alternative and forgoes another alternative in terms of a single criterion at a time. More studies on this key issue are under development.

9. Concluding Remarks

Although MCDM plays a critical role in many real-life problems, it is hard to accept an MCDM method as being accurate all the time. The present research results on ELECTRE II and III methods complement previous ones and reveal that even more MCDM methods suffer of ranking irregularities. The ELECTRE methods are widely used today in practice. However, the mathematical processes of these methods are not as exact and stable as they were assumed to be. The ranking irregularities found in this thesis should function as a warning for people in accepting ELECTRE's recommendations without questioning their validity. One great news is that a journal paper on these studies has been published [Wang and Triantaphyllou, 2008] and has already attracted some attention from researchers in this area.

As discussed previously, it is unacceptable to have rank reversals happen with the ELECTRE II and III methods and the additive AHP method because they are resulted from these methods' own mathematical artifacts. However, the introduction of the notions of regret and rejoicing could be able to offer decision makers the flexibility to change their preferences and choices in a rational way. Thus, rank reversals could be justified as natural consequences of rational preference changes when there are no mathematical drawbacks involved in the ranking process. As discussed in Chapter 8, some new studies on the notion of regret and rejoicing are under development. If these studies are carried out successfully, hopefully we will be able to offer a new effective MCDM method to the literature and more importantly to the users of MCDM methods. This method will not only have the ability to incorporate the effects of regret and rejoicing but also be able to refrain the occurrence of those previously discussed ranking irregularity problems when the effects of regret and rejoicing is negligible. Decision makers will have the flexibility to decide their own regret/rejoicing levels and the importance of these emotional factors in their decision-making process.

Another direction for future research is to define more test criteria against which existing and future MCDM methods can be evaluated. Some interesting work in this area has been conducted. For example, in [Kujawski, 2005], the author proposed three properties for a desirable MCDM approach. They are about independence of dominated alternatives, no imposed rank reversal and negative side effects associated with inferior substitutions. In [Kujawski, 2005], it is claimed that the RDRM model satisfies the above three properties. According to the first property the RDRM model preserves the ranking of two alternatives A_i and A_j with ranking $A_i > A_j$ when a new alternative dominated by A_i is introduced or an old alternative dominated by A_j is dropped. However, as our newly studies prove, this claim is not always true. A research note is under preparation to demonstrate this point algebraically and also by means of a numerical example [Wang, *et al.*, 2007]. The implication of the new finding is that the common practice of dropping dominated alternatives may lead to misleading conclusions if the analysis involves regret or rejoicing effects. Clearly, it is a fascinating area of research and it is of paramount significance to both researchers and practitioners in the multi-criteria decision-making field.

References

Allais, M., (1988), The general theory of random choices in relation to the invariant cardinal utility function and the specific probability function, *Risk, Decision and Rationality*, in Munier, editor, pp. 231-291, Dordrecht, The Netherlands: Reidel.

Barzilai, J., and F.A. Lootsma, (1994), "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.

Bell, D.E., (1982), "Regret in Decision Making under Uncertainty," *Operations Research*, Vol.30, pp. 961-981.

Bell, D.E., (1985), "Disappointment in decision making under uncertainty," *Operations Research*, Vol. 33, pp.1-27.

Belton, V., and A.E. Gear, (1983), "On a Short-Coming of Saaty's Method of Analytic Hierarchies," Omega, Vol. 11, pp. 228-230.

Belton, V., and T.J. Stewart, (2001), "Chapter 8: Outranking Methods," *Multiple Criteria Decision Analysis: An Integrated Approach*, Kluwer Academic Publishers, Boston, MA, U.S.A.

Benayoun, R., B. Roy, and N. Sussman, (1966), "Manual de Reference du Programme Electre," Note De Synthese et Formaton, No.25, Direction Scientifque SEMA, Paris, France.

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

DeSanctis, G., and R.B. Gallupe, (1987), "A foundation for the study of group decision support systems," *Management Science*, Vol. 33 (5), pp. 589-609.

Dyer, J.S., (1990a), "Remarks on the Analytic Hierarchy Process," *Management Science*, Vol. 36, pp.249-258.

Dyer, J.S., (1990b), 'A Clarification of "Remarks on the Analytic Hierarchy Process",' Management Science, Vol. 36, pp. 274-275.

Dyer, J.S., and R.E. Wendell, (1985), "A Critique of the Analytic Hierarchy Process," Technical Report 84/85-4-24, Department of Management, the University of Texas at Austin, Austin, TX, U.S.A.

Doumpos, M., K. Kosmidou, G. Baourakis, and C. Zopounidis (2002), "Credit risk assessment using a multicriteria hierarchical discrimination approach: A comparative analysis," *European Journal of Operational Research*, Vol. 138, pp. 392-412.

Ellsberg, D. (1961), "Risk, ambiguity and the Savage axioms," *Quarterly Journal of Economics*, Vol. 75, pp. 643-649.

Farrow, S., and L.J. Valverde, (2005), Federal Decision-Making for Homeland Security: Mapping the Descriptive/Normative Divide, Society for Risk Analysis Annual Meeting 2005, M11.2.

George, J.F., A.R. Dennis, and J.F. Nunamaker, (1992), An Experimental Investigation of Facilitation in an EMS Decision Room, Group Decision Negotiation, Vol. 1, pp. 57-70.

Golden B., E. Wasil, and P. Harker, (1989), The Analytic Hierarchy Process: applications and studies, Springer, Berlin, German.

Hackman, J.R., and R.E. Kaplan, (1974), Interventions into group process: An approach to improving the effectiveness of groups, *Decision Science*, Vol. 5, pp. 459-480.

Harker, P.T., and L.G. Vargas, (1990), "Reply to 'Remarks on the Analytic Hierarchy Process'," *Management Science*, Vol. 36 (3), pp. 269-273.

Hastie, R., and R.M. Dawes, (2001), *Rational choice in an uncertain world*, Sage Publications, Thousand Oaks.

Hobbs, B.F., and P. Meier, (2000), *Energy Decisions and the Environment: A Guide to the Use of Multicriteria Methods*, Kluwer Academic Publishers, Boston, MA, U.S.A.

Hokkanen, J., and P. Salminen, (1997), "Choosing a solid waste management system using multi-criteria decision analysis," *European Journal of Operational Research*, 98, pp. 19-36.

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

Jason K. L., J. Hartmann, K. W. Li, Y. An, A. Asgary, (2007), "Multi-Criteria Decision Support Systems for Flood Hazard Mitigation and Emergency Response in Urban Watersheds", *Journal of the American Water Resources Association*, Vol. 43 (2), pp. 346–358.

Kaliszewski, I., and W. Michalowski, (1998), "Establishing Regret Attitude of a Decision Maker within the MCDM Modeling Framework," IR-98-070, International Institute for Applied Systems Analysis.

Keeney, R., and H. Raiffa, (1976), *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, New York: Wiley, U.S.A.

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

Kujawski, E., (2005), "A reference-dependent regret model for deterministic tradeoff studies," *Systems Engineering*, Vol. 8, pp. 119-137.

Loomes, G., and R. Sugden, (1982), "Regret Theory: an Alternative Theory of Rational Choice Under Uncertainty," *The Economic Journal*, Vol. 92, pp.805-824.

Lootsma, F.A., (1991), "Scale Sensitivity and Rank Preservation in a Multiplicative Variant of the AHP and SMART," Technical Report 91-67, Faculty of Technical Mathematics and Informatics, Delft University of Technology, Delft, the Netherlands.

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

Luce, R.D., (1992), "Where does subjective expected utility fail descriptively?" *Journal of Risk and Uncertainty*, Vol. 5, pp. 5-27.

Loomes, G., and R. Sugden, (1982), "Regret Theory: an Alternative Theory of Rational Choice Under Uncertainty," *The Economic Journal*, Vol. 92, pp.805-824.

Mellers, B.A., (2000), "Choice and the relative pleasure of consequences", *Psychological bulletin*, Vol. 126, pp. 910–924.

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

Milnor, J., (1954), Games against nature. In R.M. Thrall, C.H. Coombs, and R.L. Davis, editors, *Decision Processes*, pp. 49-60, Wiley and Chapman & Hall, New York and London.

Plous, S., (1993), *The psychology of judgment and decision making*, McGraw-Hill, New York, NY, U.S.A.

Raj, P.A., (1995), "Multi-criteria Methods in River Basin Planning- A Case Study," *Water Science and Technology*, 31(8), pp. 261–272.

Rogers, M.G. and M.P. Bruen, (1999), Applying ELECTRE to an option choice problem within an environmental appraisal – three case studies from the Republic of Ireland, Chapter 10 in Meskens, N. and Roubens, M. (Eds.), Advances in Decision Analysis, *Kluwer Academic Publishers*, Dordrecht, The Netherlands.

Rogers, M.G., M.P. Bruen, and L.-Y. Maystre, (1999), "Chapter 3: The Electre Methodology," *Electre and Decision Support*, Kluwer Academic Publishers, Boston, MA, U.S.A.

Roy, B., (1968), "Classement et choix en presence de points de vue multiples: La methode ELECTRE," R.I.R.O 8, pp. 57-75.

Roy, B., (1978), "ELECTRE III: Un algorithme de classements fonde sur une representation floue des preference en presence de criteres multiples," Cahiers de CERO, Vol. 20, No.1, pp. 3-24.

Roy, B., and P. Bertier, (1971), "La methode ELECTRE II: Une methode de classement en presence de critteres multiples," SEMA (Metra International), Direction Scientifique, Note de Travail No. 142, Paris, 25p.

Roy, B., and P. Bertier, (1973), "La methode ELECTRE II: Une methode au media-planning," In Ross, M. (ed.), Operational Research 1972, North-Holland Publishing Company, pp. 291-302.

Roy, B. and D. Bouyssou, (1986), "Comparison of Two Decision-Aid Models Applied to a Nuclear Plant Sitting Example", *European Journal of Operations Research*, Vol. 25, pp. 200-215.

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

Saaty, T.L., (1983), Axiomatic Foundations of the Analytic Hierarchy Process, *Management Science*, Vol. 32, pp. 841-855.

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

Saaty, T.L., (1990), 'An Exposition of the AHP in Reply to the Paper "Remarks on the Analytic Hierarchy Process",' *Management Science*, Vol.36 (3), pp. 259-268.

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

Saaty, T.L., and L. Vargas, (2000), *Models, Methods, Concepts and Applications of the Analytic Hierarchy Process*, Kluwer Academic Publishers, Boston, MA, U.S.A.

Stam, A., and A.P.D. Silva, (1997), "Stochastic Judgments in the AHP: the measurement of rank reversal probabilities," *Decision Sciences*, Vol.28, pp.655-688.

Sugden, R., (1985), "Regret, recrimination and rationality," *Theory and Decision*, Vol. 19, pp. 77-99.

Savage, L.J., (1951), "The theory of statistical decision," *Journal of American Statistical Association*, Vol. 46, pp.55-67.

Triantaphyllou, E., (2000), *Multi-Criteria Decision Making Methods: A Comparative Study*, Kluwer Academic Publishers, Boston, MA, U.S.A.

Triantaphyllou, E., (2001), "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 issue) Vol. 10, pp. 11-25.

Triantaphyllou, E., and S.H. Mann, (1989), "An Examination of the Effectiveness of Multi-Dimensional Decision-Making Methods: A Decision-Making Paradox," *International Journal of Decision Support Systems*, No. 5, pp. 303-312.

Triantaphyllou, E., F.A. Lootsma, P.M. Pardalos, and S.H. Mann, (1994), "On the Evaluation And Application of Different Scales For Quantifying Pairwise Comparisons in Fuzzy Sets," Journal of Multi-Criteria Decision Analysis, Vol. 3 (3), pp. 133-155.

Vaidya, O.S., and S. Kumar, (2006), "Analytic hierarchy process: An overview of applications," European Journal of Operational Research, Vol. 169, pp.1-29.

Wang, X., and E. Triantaphyllou, (2008), "Ranking Irregularities When Evaluating Alternatives by Using Some ELECTRE Methods," *Omega*, Vol. 36, pp. 45-63.

Wang, X., E. Triantaphyllou, and E. Kujawski, (2007), A Research Note on the paper "A Reference-Dependent Regret Model for Deterministic Tradeoff Studies", working paper.

Zavadskas, E.K., L. Ustinovičius, and A. Stasiulionis, (2004), "Multicriteria valuation of commercial construction projects for investment purposes," *Journal of Civil Engineering and Management*, Vilnius: Technika, 2004, t.X, Nr. 2, pp. 151-166.

Zeelenberg, M., (1999), "Anticipated regret, expected feedback and behavioral decision making," *Journal of Behavioral Decision-Making*, Vol.12, pp. 93-106.

Zopounidis, C., and M. Doumpos, (2000), *Intelligent Decision Aiding Systems Based on Multiple Criteria for Financial Engineering*, Kluwer Academic Publishers, Boston, MA, U.S.A.

Vita

Ms. Xiaoting Wang is from Shaanxi province in China. She received her Bachelor of Science degree in mechatronics engineering at Nanchang Institute of Aeronautical Technology, Nanchang, China, in 1999, and a master's degree in mechatronics engineering at Northwestern Polytechnic University, Xi'an, China, in 2002. Xiaoting enrolled in the doctoral program in engineering science in Louisiana State University in the spring of 2003. Her current research activities mainly focus on the theory and applications of multi-criteria decision making (MCDM).