Vili-Matti Ojala

Missing preferences in pairwise comparison matrices: a numerical study.

School of Science

Bachelor's thesis
Espoo, 04.05.2015

Thesis supervisor:

Prof. Raimo P. Hämäläinen

Thesis advisor:

PhD Matteo Brunelli

Aalto University
School of Science

AALTO UNIVERSITY
SCHOOL OF SCIENCE

ABSTRACT OF THE
BACHELOR'S THESIS

Author: Vili-Matti Ojala

Title: Missing preferences in pairwise comparison matrices: a numerical study.

Date: 04.05.2015 Language: English Number of pages: 3+19

Degree programme: Engineering Physics and Mathematics

Supervisor: Prof. Raimo P. Hämäläinen

Advisor: PhD Matteo Brunelli

In this thesis the concept of pairwise comparison matrices is introduced and how they are used in Analytic Hierarchy Process (AHP) is briefly covered. How to estimate weights from a pairwise comparison matrix is explained and the focus is on how to deal with missing comparisons. Harker and Shiraishi et al. proposed methods to estimate the weights despite the missing comparisons. How these methods work is explained and the performance of them is tested with numerical simulations. In the simulations pairwise comparison matrices are created and weights are estimated from them. Then they are compared to weights estimated from the same matrices after some comparisons have been removed. The order of the matrices and the percentage of missing comparisons are varied. The goal of the simulations is to observe how the missing comparisons affect the stability of the solution.

Keywords: Pairwise comparison matrices, missing preferences, Analytic Hierarchy Process (AHP), weights, Harker's method, Method of Shiraishi et al.

Contents

1 Introduction

When making a simple decision such as choosing what to eat for lunch, a solution often pops into our heads effortlessly. Our internal heuristic decision making process provides us with an answer, which we might accept without much thought. On the other hand when making a more important decision such as, deciding what apartment to buy, a good solution might not pop into our mind as easily. Then a more analytical approach is needed and dividing the problem into sub problems might be helpful. What are the criteria for a good apartment? How the available apartments fulfil these criteria? Thinking all of these things at once can be overwhelming. To aid in a situation such as this, many decision making tools have been developed and one of them is the Analytic Hierarchy Process (AHP)[1]. It is a multi criteria decision making tool, which provides a structured way to approach a decision and come to a conclusion. In this thesis, we focused on two steps in the AHP. In the first you compare the importance of every criterion against each other and then these comparisons are stored in a pairwise comparison matrix. In the next step weights are estimated from the matrix. These weights describe the importance of each criterion relative to the decision at hand. The weights can be later used in calculating the best alternative for the decision maker.
In some cases the decision maker might not be able to give an answer when asked to compare two criteria against each other. Missing comparisons cause problems when estimating the weights and there are several methods which have been developed to estimate the weights despite the missing comparisons. In this thesis we study two of them, Harker's method [2] and the method of Shiraishi et al. [3]. In simulations, these methods were applied on pairwise comparison matrices of different sizes and amounts of missing comparisons. The purpose of the simulation was to observe how much the the missing comparisons affect the estimation of the weights and how well the methods perform compared to each other.

2 Pairwise comparison matrix

Pairwise comparison matrices play an important role in the Analytic Hierarchy Process (AHP), which is a multi criteria decision-analysis tool developed by T.L Saaty in the 1970's [1]. AHP can be used to rank the options available to the decision maker, based on his or her preferences. For example, when deciding what apartment to buy, one might focus on criteria such as price, size, location and condition. Arranging the available apartments from best to worst requires estimating a score to each apartment based on these criteria. For example, one apartment might have excellent price and location, but the condition and size can be poor. Another option could have superb location and condition, but it might be small and expensive. How to say which one is the best alternative? AHP can give an answer to that question, but we will not describe the whole process here. Instead we will focus on the part of AHP, where we estimate the weights of the criteria from a pairwise comparison matrix. Weights w_1, w_2, ..., w_n $\in \mathbb{R}_+$ are used to describe the relative importance

of the chosen n criteria. In the previously introduced apartment problem $n = 4$. w_1 is the weight of price, w_2 is the weight of size, etc.. The more important a criterion is to the decision maker, the larger the corresponding weight w_i should be. [1]
If people are asked to write down the value of each weight and the score of each option is calculated using those weights, the decision makers might feel that the results do not represent their true preferences. Assigning weights to the criteria becomes increasingly difficult as their number grows [4]. To overcome these challenges, in AHP pairwise comparison matrices are used, where the entries are numerical comparisons of only two criteria at a time. A pairwise comparison matrix is a square matrix of size $n \times n$ where $a_{ij},\ i, j \in \{1, 2, ..., n\}$, is the entry on row i and column j. To fill a pairwise comparison matrix the decision maker is asked questions such as: "On a scale from $\frac{1}{9}$ to 9, how much more important is the criterion i compared to criterion j?" The answer to this question, noted as a_{ij}, gives insight into what the ratio of weights $\frac{w_i}{w_j} \approx a_{ij}$ is. The answers are saved into a pairwise comparison matrix $\mathbf{A}$ in the following manner [1]:

$$\mathbf{A} = \begin{pmatrix} 1 & a_{12} & a_{13} & \cdots & a_{1n} \\ 1/a_{12} & 1 & a_{23} & \cdots & a_{2n} \\ 1/a_{13} & 1/a_{23} & 1 & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1/a_{1n} & 1/a_{2n} & 1/a_{3n} & \cdots & 1 \end{pmatrix} \approx \begin{pmatrix} w_1/w_1 & w_1/w_2 & w_1/w_3 & \cdots & w_1/w_n \\ w_2/w_1 & w_2/w_2 & w_2/w_3 & \cdots & w_2/w_n \\ w_3/w_1 & w_3/w_2 & w_3/w_3 & \cdots & w_3/w_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_n/w_1 & w_n/w_2 & w_n/w_3 & \cdots & w_n/w_n \end{pmatrix}.$$

In AHP, only the upper triangular part $(a_{ij}, \forall i < j)$ of the matrix $\mathbf{A}$ is filled directly with the answers of the decision maker. The lower triangular part $(a_{ij}, \forall i > j)$ is filled with the inverse of the corresponding value from the upper triangular part. It is sensible to assume that if we know a_{ij}, then a_{ji} should be the inverse of a_{ij} [1]. For example if the decision maker says that price of an apartment is 2.5 times more important than the location, then location should be 0.4 times as important as price. By asking the decision maker to fill only the upper triangular part of the matrix, the number of required comparisons is $(n-1)n/2$. The diagonal entries of a pairwise comparison matrix are all equal to 1, because the importance of a criterion compared to itself must be 1.
If we fill the first row of the pairwise comparison matrix and we assume that the sum of the weights is one, then we have n variables $w_1, w_2, ..., w_n$ and n equations.

$$\begin{cases} \sum_{k=1}^{n} w_k = 1 \\ w_2 = w_1/a_{12} \\ w_3 = w_1/a_{13} \\ w_4 = w_1/a_{14} \\ \vdots \\ w_n = w_1/a_{1n} \end{cases}$$

Thus we can estimate the values of the weights $w_1, w_2, ..., w_n$, but the robustness of the solution is questionable. A single poorly thought answer from the decision

maker will have a large effect. In AHP the number of comparisons is $n(n-1)/2$ and the information in every row of the pairwise comparison matrix is used to estimate the weights. By using the information on each row, the robustness of the solution increases, because a single poorly thought answer from the decision maker does not have such a large impact on the results.
Weights can be estimated from the pairwise comparison matrix, for example, with the eigenvalue method. In this widely used method, the approximation of the weights is the principal eigenvector of $\mathbf{A}$. Therefore, the weights can be calculated by solving $\mathbf{w}$ from the following equation[1].

$$\mathbf{A}\mathbf{w} = \lambda_{max}\mathbf{w}, \tag{1}$$

where λ_{max} is the principal eigenvalue of $\mathbf{A}$. The first entry of $\mathbf{w}$ is the weight of the first criterion, the second entry is the weight of the second criterion, etc.. The weights are usually normalised so that the sum of the weights is one.

2.1 Consistency

Only a perfectly rational person is able to give consistent answers which satisfy the equation of cardinal consistency for pairwise comparison matrices [1]:

$$a_{ik} = a_{ij}a_{jk}, \ \forall i, j, k. \tag{2}$$

In the context of AHP and pairwise comparison matrices, consistency means that all connected paths from one entry to another should give the same result. For example, the decision maker might state that: "The condition of the apartment is 2 times more important than location ($a_{43} = 2$), price is tree times more important than condition ($a_{14} = 3$) and price is six times more important than location (a_{13})=6" For this example we can see that equation (2) holds and therefore these three statements are consistent with each other:

$$6 = 3 \times 2.$$

In practise the decision maker is not expected to be able to give answers that are all completely consistent. For example, if he had changed only one of the answers even slightly, the statement would have been inconsistent, because equation (2) would no longer hold. Fortunately, approximations of the weights can be calculated even when the answers are inconsistent. In fact, inconsistency is expected and decision making tools like AHP are designed to deal with it.
Saaty showed that the consistency of an positive reciprocal matrix $\mathbf{A}$ is connected to the value of its principal eigenvalue λ_{max}. When $\mathbf{A}$ is consistent, it applies that $\lambda_{max} = n$, and $\lambda_{max} > n$ when there is inconsistency in the matrix. To measure the inconsistency of the pairwise comparison matrix Saaty introduced the consistency index (C.I) [1]:

$$\text{C.I} = \frac{\lambda_{max} - n}{n - 1} \tag{3}$$

and the consistency ratio CR:

$$\text{CR} = \frac{\text{C.I}}{\text{RI}(n)} = \frac{\frac{\lambda_{max}-n}{n-1}}{\text{RI}(n)}, \tag{4}$$

Table 1: RI values for matrices of order n [6].

n	3	4	5	6	7	8	9	10
RI(n)	0.5247	0.8816	1.1086	1.2479	1.3417	1.4057	1.4499	1.4854

where RI(n) is the average value of C.I of pairwise comparison matrices of size $n \times n$, where the entries are randomly generated from a scale from 1/9 to 9. Table 1 has RI values for matrices of different order. Small C.I and CR values are an indication of consistency. Saaty proposed that matrices with a CR value between 0 and 0.10 are considered good enough to evaluate weights from [1]. If the CR value of the matrix is larger than 0.10, it is worth considering to ask the decision maker to answer the questions again, in hope of more consistent answers. Inconsistent answers, might be, for example, a sign of lack of focus from the decision maker. In these cases the solution derived from the answers, might not represent accurately the preferences of the decision maker. How to measure inconsistency and how to determine which matrices are too inconsistent to evaluate weights from is still under research [5].

3 Missing pairwise comparisons

Missing comparisons are empty entries in a pairwise comparison matrix. There are several reasons why the pairwise comparison matrix might not be complete and Harker presented some examples [2]:

(i) The decision maker might not know the answer to a question due to lack of expertise or he might not have a clear opinion on the matter. Leaving the answer empty might provide better solution than forcing the decision maker to guess.

(ii) The decision maker does not want to explicitly tell their preferences. For example, when asked about the importance of profit compared to the safety of the employees.

(iii) The number of pairwise comparisons that have to be made by the decision maker can be overwhelming. There might not simply be enough time to answer all the questions.

Harker's examples makes it easy to understand that missing comparisons are not that rare and therefore understanding their effects is important. For example Ureña et al. [7] and Carmone et al. [8] have done research on missing comparisons and decision making. There are several methods of estimating the weights despite the missing preferences and in this thesis two methods are compared. The first was presented by Harker in 1987 and it will be from now on referred to as Harker's method [2]. The second method is developed by Shiraishi et al. in 1998 and from now on it will be referred to as method of Shiraishi et al.[3].
Short explanations of these methods will be presented later, but first we need to

introduce some notations. Let us call $\mathbf{A}_n(\mathbf{x}_k)$ the incomplete comparison matrix of order n with k missing comparisons. $\mathbf{x}_k$ is a vector, which contains the variables for missing comparisons. In the following example $n = 4$, $k = 3$. Pairwise comparisons a_{12}, a_{14}, a_{23} and their reciprocals $1/a_{12}, 1/a_{14}, 1/a_{23}$ are missing from $\mathbf{A}_n(\mathbf{x}_k)$. The missing preferences are replaced with $\mathbf{x}_3 = [x_{12}, x_{14}, x_{23}]$ and their reciprocals are replaced with $1/x_{12}, 1/x_{14}$ and $1/x_{23}$:

$$\mathbf{A}_n(\mathbf{x}_3) = \begin{pmatrix} 1 & x_{12} & a_{13} & x_{14} \\ 1/x_{12} & 1 & x_{23} & a_{24} \\ 1/a_{13} & 1/x_{23} & 1 & a_{34} \\ x/a_{14} & 1/a_{24} & 1/a_{34} & 1 \end{pmatrix}.$$

3.1 Harker's method

Harker's method modifies the incomplete pairwise comparison matrix $\mathbf{A}_n(\mathbf{x}_k)$ using simple rules into a quasi-reciprocal comparison matrix $\mathbf{C}$. The weights can then be estimated from $\mathbf{C}$ using the eigenvector method. The rules for creating the matrix $\mathbf{C}$ from $\mathbf{A}_n(\mathbf{x}_k)$ are [2]:

$$c_{ij} = \begin{cases} 0 & \text{if } a_{ij} \text{ is a missing comparison or its reciprocal, } \forall i \neq j \\ a_{ij} & \text{if } a_{ij} \text{ is a real number} > 0, \ \forall i \neq j \\ 1 + m_i & \text{where } m_i \text{ is the number of zeros in the row } i, \ \forall i = j. \end{cases}$$

Let us assume that pairwise comparison matrix $\mathbf{B}$ contains the true preferences of the decision maker:

$$\mathbf{B} = \begin{pmatrix} 1 & 1/3 & 1/4 & 1/6 \\ 3 & 1 & 3/4 & 1/2 \\ 4 & 4/3 & 1 & 2/3 \\ 6 & 2 & 3/2 & 1 \end{pmatrix}$$

The principal eigenvalue of $\mathbf{B}$ is $\lambda_{max} = 4 = n$, which means that $\mathbf{B}$ is consistent. Using the eigenvector method, the weights $\mathbf{w}$ can be solved from the following equation.

$$\mathbf{B}\mathbf{w} = \lambda_{max}\mathbf{w}, \tag{5}$$

where λ_{max} is the principal eigenvalue of $\mathbf{B}$. The principal eigenvector $\mathbf{w}$ is [1/6, 1/2, 2/3, 1]. Weights are usually normalised so that their sum is one. Therefore the normalised weights are $\mathbf{w} \approx [0.07, 0.21, 0.29, 0.43]$. These are the weights estimated from a complete pairwise comparison matrix and they will be compared to the weights estimated with Harker's method from the incomplete pairwise comparison matrix.

Let us assume that during the elicitation process two questions are left unanswered. By removing two answers from the matrix $\mathbf{B}$, we get an incomplete pairwise comparison matrix $\mathbf{A}_4(\mathbf{x}_2)$:

$$\mathbf{A}_4(\mathbf{x}_2) = \begin{pmatrix} 1 & x_{12} & 1/4 & 1/6 \\ 1/x_{12} & 1 & x_{23} & 1/2 \\ 4 & 1/x_{23} & 1 & 2/3 \\ 6 & 2 & 3/2 & 1 \end{pmatrix}.$$

By using Harker's rules we get:

$$\mathbf{C} = \begin{pmatrix} 2 & 0 & 1/4 & 1/6 \\ 0 & 3 & 0 & 1/2 \\ 4 & 0 & 2 & 2/3 \\ 6 & 2 & 3/2 & 1 \end{pmatrix}$$

The weights can be calculated by solving the principal eigenvector $\mathbf{w}$ from:

$$\mathbf{Cw} = \lambda_{max}\mathbf{w}. \tag{6}$$

The principal eigenvalue of $\mathbf{C}$ is $\lambda_{max} = 4 = n$ and the corresponding principal eigenvector $\mathbf{w} = [1/6,\ 1/2,\ 2/3,\ 1]$. This is exactly the same principal eigenvector than which were calculated from the complete pairwise comparison matrix $\mathbf{B}$. Harker's method was able to solve the eigenvector perfectly despite the missing pairwise comparisons. The reason why it was so successful is that the comparison matrix used in this example was consistent. A perfect scenario was used to keep the example simple and clear. Perfect consistency is not expected from a real decision makers and for this reason, the pairwise comparison matrices will not be consistent in the numerical study of the thesis.

3.2 Method of Shiraishi et al.

This method was presented in 1998 by Shiraishi et al. [3]. It is more complicated than Harker's method. The idea is to complete the pairwise comparison matrix $\mathbf{A}_n(\mathbf{x}_k)$ with values $\mathbf{x}_k^{'}$ which try to maximise the consistency of the matrix using a heuristic method. Then the weights can be estimated using the eigenvector method from this completed matrix noted as $\mathbf{A}_n^{'}(\mathbf{x}_k^{'})$.
The previously introduced consistency index:

$$\text{C.I} = \frac{\lambda_{max} - n}{n - 1} \tag{7}$$

can be used to measure the consistency of the pairwise comparison matrix. A small C.I value is an indication of consistency and it depends on the largest eigenvalue λ_{max}. To maximise the consistency, one could try to find values $\mathbf{x}_k^{'}$, which minimises the largest eigenvalue $\lambda_{max}(\mathbf{x}_k^{'})$, as Bozóki et al. have done [9]. Instead, Shiraishi et al. used a heuristic method to minimize λ_{max} [3]. By definition, the characteristic polynomial $P_A(\lambda)$ of a pairwise comparison matrix $\mathbf{A}$ of size $n \times n$ is:

$$P_A(\lambda) = \det(\lambda I - \mathbf{A}) = \lambda^n + c_1\lambda^{n-1} + ... + c_n, \tag{8}$$

where I is the identity matrix. The coefficient of the term λ^{n-3} is c_3 and it holds that $c_3 \leq 0$. Shiraishi et al. have shown in their paper [10] that c_3 has a connection to the matrix's consistency. The larger c_3 is, the smaller the C.I values tend to become. The connection is not theoretically proven, but empiric tests show that there is a strong relationship [10]. Other papers such as one made by Brunelli et al. [11] confirms this connection. The value of c_3 can be calculated using the following equation [10].

$$c_3 = \sum_{i<j<k} \{2 - (\frac{a_{ij}a_{jk}}{a_{ik}} + \frac{a_{ik}}{a_{ij}a_{jk}})\}. \tag{9}$$

Shiraishi et al. proposed to fill the missing comparisons in matrix $\mathbf{A}_n(\mathbf{x}_k)$ with positive values $\mathbf{x}_k^{'}$, which maximize the value of c_3 [3]. Using the method of Shiraishi et al. therefore requires solving the following optimisation problem:

$$\begin{aligned} &\text{maximise} \quad && c_3(\mathbf{x}_k^{'}) \\ &\text{subject to} \quad && x_{ij}^{'} > 0, \forall i, j \end{aligned} \tag{10}$$

The weights can then be estimated from $\mathbf{A}_n(\mathbf{x}_k^{'})$ using the eigenvalue method. This optimisation problem was solved in the simulation presented in this paper with an additional constraint:

$$x_{ij} \geq 0.0001, \forall i, j. \tag{11}$$

It was added because, without it, in some cases the variables would be given a value too close to zero, which led to a division with zero and an error. This restriction insured the proper function of the simulation and it did not have any effect to the results.

4 Methodology

Creation of the pairwise comparison matrices

Numerical simulations were done to observe the effects of missing comparisons. Next we will explain the steps in the simulations. First we create a set of weights $(w_1^*, w_2^*, ..., w_n^*)$, which represent the decision maker's accurate preferences on n criteria. The weights are randomly generated real numbers from a range from one to nine. Then these weights are used to create a pairwise comparison matrix $\mathbf{A}^*$ where $a_{ij}^* = w_i^*/w_j^*, \; \forall \; i, \; j$. This matrix is completely consistent and the answers describe the decision makers preferences accurately. However it is not a realistic pairwise comparison matrix elicited from a decision maker, since people are expected to give answers which can deviate from their true preferences so that a_{ij} is not necessarily an exact ratio of the weights, but an approximation of $w_i/w_j, \; \forall \; i, \; j$. To achieve a more realistic pairwise comparison matrix $\mathbf{A}$, we then multiply each value a_{ij}^* in the upper triangular part $(i < j, \forall i, j)$ of $\mathbf{A}^*$ with a different coefficient v_{ij}. Each v_{ij} is randomly sampled from a log-normal distribution so that $v_{ij} \sim lnN(\mu, \sigma), \; \forall i, j$. Then the lower triangular part of the matrix is remade by taking the reciprocal of the upper triangular part:

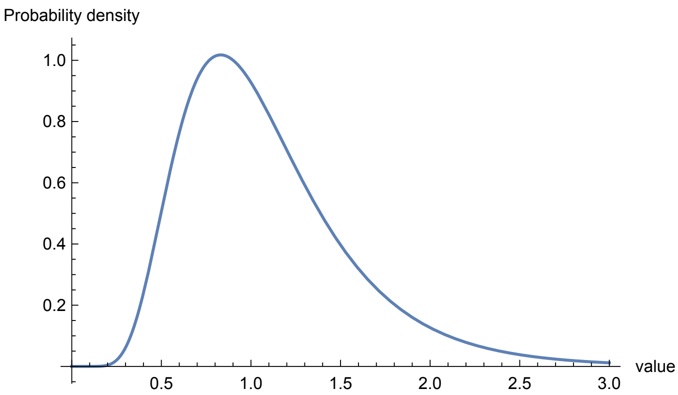

Figure 1: Probability density of the log-normal distribution with normal mean μ=0 and normal standard deviation σ=0.43.

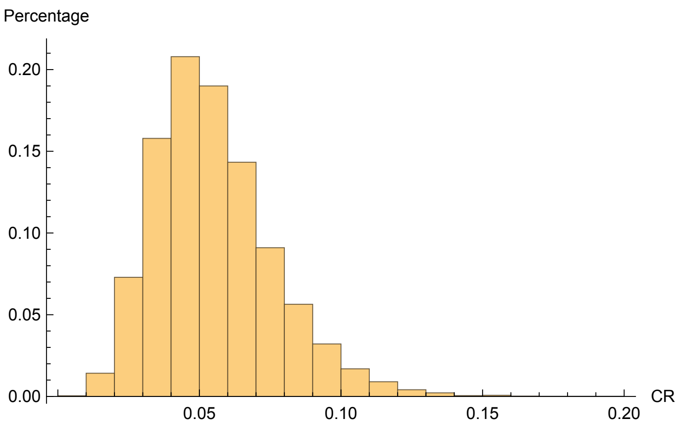

Figure 2: Histogram of the consistency ratios (CR) of 10 000 pairwise comparison matrices of size 7×7.

$$\mathbf{A}^* = \begin{pmatrix} 1 & a_{12}^* & \cdots & a_{1n}^* \\ 1/a_{12}^* & 1 & \cdots & a_{2n}^* \\ \vdots & \vdots & \ddots & \vdots \\ 1/a_{1n}^* & 1/a_{2n}^* & \cdots & 1 \end{pmatrix} \Rightarrow \mathbf{A} = \begin{pmatrix} 1 & v_{12}a_{12}^* & \cdots & v_{1n}a_{1n}^* \\ 1/(v_{12}a_{12}^*) & 1 & \cdots & v_{2n}a_{2n}^* \\ \vdots & \vdots & \ddots & \vdots \\ 1/(v_{1n}a_{1n}^*) & 1/(v_{2n}a_{2n}^*) & \cdots & 1 \end{pmatrix}.$$

The normal mean μ is selected to be zero and the normal standard deviation σ is 0.43. The probability density of $lnN(0, 0.43)$ is plotted in Figure 1. This distribution was selected because it developed pairwise comparison matrices, which are inconsistent, but rarely exceed the 0.1 limit for consistency ratio proposed by Saaty[1].
As a result of this procedure we have a positive reciprocal pairwise comparison matrix $\mathbf{A}$, where the entries simulate answers elicited from a real person. Using the described method, ten thousand pairwise comparison matrices of size 7×7 were created. Figure 2 is a histogram of consistency ratio values of those matrices.

4.1 Comparing the two methods

Applying the eigenvector method to **A**, we get weights $w_{1,}, w_2, ..., w_n$. They are then ranked based on their value. The largest weight gets rank one, the second largest gets rank two, etc. As a result we get a vector $\mathbf{r} = [r(1), r(2), ..., r(n)], r(k) \in \{1, 2, ..., n\}, k \in \{1, 2, ..., n\}$, where $r(k)$ is the ranking of the k-th weight. Let us take an example, where $n = 4$ and $\mathbf{r} = [4, 2, 1, 3]$. This means that the third weight is the largest and it is estimated to be the most important criterion to the decision maker. The index of the weight, which is calculated from the full pairwise comparison matrix **A** and which ranking is one, is noted as $m, m \in \{1, 2, ..., n\}$ and the corresponding weight is referred to as the m-th weight. In this example $r(3) = 1$ and therefore $m = 3$. The corresponding weight is noted as w_m. The ranking of the largest weight $r(m)$ is then compared to the rankings created by Harker's method $r_H(m)$ and method of Shiraishi et al. $r_S(m)$, when they are used on the incomplete comparison matrix $\mathbf{A}_n(\mathbf{x}_k)$. It is important to understand that the m-th weight, which was the largest when estimated from the complete matrix **A**, might not get the same rank when the weights are estimated from $\mathbf{A}_n(\mathbf{x}_k)$. We save the values $r_H(m)$ and $r_S(m)$ and repeat the process with new pairwise comparison matrices. As a result we get statistics on how many times the methods are able to rank the largest weight w_m correctly to rank one, but we also know what rank they give it when they fail. The lower rank it gets, the worse the method is considered to be performing.

To show how this works, suppose that we constructed 100 pairwise comparison matrices of size 4×4 and removed 2 random entries from each one. The full matrices had originally $n(n-1) = 6$ comparisons and removing 2 of them, means that 33 % of the comparisons are missing. Methods of Harker and Shiraishi et al. were then applied to the incomplete matrices in the previously described manner and the results are shown in Table 2.

Table 2: Results on 100 matrices of size 4×4 with 33 % of the comparisons missing.

	Percentage of instances		Percentage of instances
$r_H(m) = 1$	75%	$r_S(m) = 1$	73%
$r_H(m) = 2$	19%	$r_S(m) = 2$	23%
$r_H(m) = 3$	5%	$r_S(m) = 3$	3%
$r_H(m) = 4$	1%	$r_S(m) = 4$	1%

In this example the method of Harker was able to correctly give the rank 1 to the weight w_m 75% of the times despite the missing comparisons. It failed in 25% of the times and in these cases it gave the largest weight w_m second rank 19 % of the cases, third rank 5% of the cases and fourth rank 1% of the cases. Method of Shiraishi et al. performed similarly and therefore a much larger sample sizes has been used in this study to draw conclusions about the differences in the performance of these methods.

Table 3: In the simulations three different size of matrices were used and each was subjected to a low percentage of missing comparisons and a high percentage of missing comparisons.

Order of the matrix Number of comparisons	$n = 4$ 6	$n = 7$ 21	$n = 10$ 45
Low: number of missing comparisons and their percentage	1 17%	4 19%	9 20 %
High: number of missing comparisons and their percentage	3 50%	10 48%	20 44%

5 Results

The simulations were made with Wolfram Mathematica version 10.0.2.0 [1]. Matrices of size 4×4, 7×7 and 10×10 were used. The amount of comparisons required from the decision maker to fill each matrix is $n(n-1)/2$, where n is the order of the matrix. Therefore the amount of comparisons available for removal from these matrices are 6, 21 and 45, respectively. Certain percentages of these comparisons were removed to test how the missing comparisons affect the estimation of the weights. Two levels of missing comparisons were chosen. There was a high amount of missing comparisons, which was 44%-50% of the available comparisons and a low amount, which was 17%-20%, depending on the size of the matrix. See Table 3 for details. The percentages of missing comparisons are not exactly the same for every size of matrices, but they were chosen so that they would be close to each others. For example it is impossible to remove 50% of the comparisons from a matrix that has 21 comparisons. For that reason the percentage of missing comparisons in the 7×7 matrix is 48% and not exactly the same as in 4×4 matrix. The simulations were conducted so that both methods were used on all three different size matrices with a low level of missing comparisons and a high percentage of missing comparisons.
The scenarios from Table 3 were each simulated 50 000 times as described in the methodology section. The statistics of the simulations can be found on Table 4 in the Appendix A, and the main results are summarised in Figure 3. Both methods performed very similarly when the percentage of missing comparisons was low: 17%, 19% and 20%. Based on those results, neither of the methods can be considered to be better or worse. Small differences in the performance did occurred when the amount of missing comparisons was high: 50%, 48% and 44%. In these cases Harker's method was able to place the m-th weight correctly to first rank more often than the method of Shiraishi et al. On the other hand, it also placed the m-th weight more frequently incorrectly to the last rank compared to the method of Shiraishi et al. Therefore neither of the methods can be considered to be clearly better or worse than the other, based on their accuracy, but there is one advantage in using Harker's method. It was more than nine times faster to compute than the method of Shiraishi et al. In Harker's method the incomplete comparison matrix $\mathbf{A}_n(\mathbf{x}_k)$ is

[1] The orignal script was made by Matteo Brunelli and it was modified by Vili-Matti Ojala.

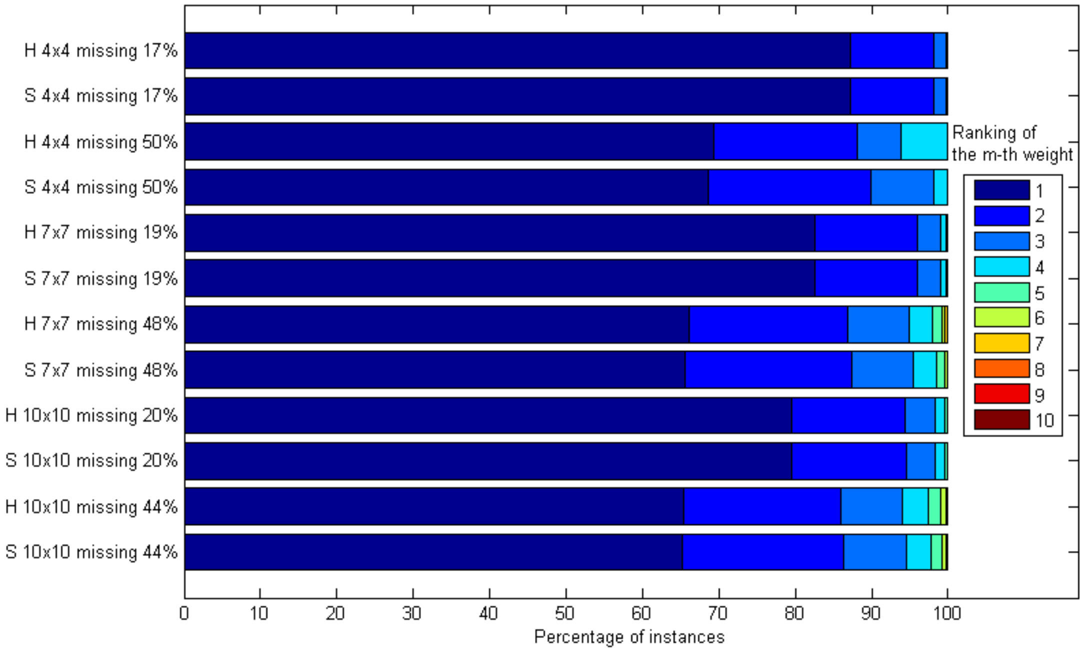

Figure 3: The percentage of instances of each rank was given to the m-th weight, when each scenario was simulated 50 000 times. S stands for method of Shiraishi et al. and H stands for Harker's method. The size of the matrix is indicated with $n \times n$ and the amount of missing comparisons is shown with percentages. Small ranks are a sign of accurate estimations of the m-th weight. This chart was made from Table 4, which can be found in the Appendix.

modified using simple rules, but in the method of Shiraishi et al. the optimisation problem displayed in equation 10, must be solved. Solving this problem took by far the most time compared to any other part of the simulations.

6 Discussion

The pairwise comparison matrices used in this thesis were randomly generated and they were made inconsistent by multiplying their entries with different coefficients v_{ij} randomly sampled from a log-normal distribution so that $v_{ij} \sim lnN(\mu, \sigma)$, $\forall i, j$, where $\sigma = 0.43$ and $\mu = 0$. Are pairwise comparison matrices produced in this way good approximations of those elicited from real persons? Maybe not exactly, but they can still be used to test the performance of method of Shiraishi et al. and Harker's method. If more research on this subject were made in the future, one might want to compare how the inconsistency of the pairwise comparison matrices affect the results. The inconsistency can be altered, by changing the distribution of from which parameters v_{ij} are drawn from. To create more realistic pairwise comparison matrices, one might want to change the level of inconsistency depending of the size of the matrix. According to Bozóki et al., the larger the matrix is, the more

inconsistent peoples answers tend to be [12].
The performance of the methods were measured by the differences in the ranking of the weights estimated using full information of a complete pairwise comparison matrix and the weights estimated from the same matrix, but with some of the comparisons missing. We assumed that the weights estimated from the complete matrix were more accurate representations of the decision makers true preferences, compared to the weights estimated from the matrix with missing comparisons. Changes in the ranking of the weights were considered to be a failure. The reasoning behind this was that, more information should give more accurate results. In reality this is not always the case. Filling a pairwise comparison matrix can be a challenging task to perform, and not all of the comparisons might be accurate representations of the decision maker's true preferences. These poorly thought comparisons will affect the estimated weights in an undesirable way. The estimated weights would represent more accurately the decision makers preferences, if we could ignore the poorly thought comparisons all together and use only the accurate comparisons. Unfortunately distinguishing the accurate answers from the inaccurate is hard or impossible, but surely if some of the answers are randomly removed, like in the simulations in this thesis, in some cases the bad answers are removed and the weights can be accurately estimated from the remaining good answers. Therefore in some cases, when the methods ranked the m-th weight to some other place than first, it is not necessarily a bad thing for the decision maker. Perhaps the ranking made with less information happened to represent more accurately the true preferences of the decision maker.

7 Conclusion

The percentage of missing comparisons and the order of the matrix both affected the reliability of the estimation of the weights. When estimating weights from a matrix of order n, the amount of comparisons required from the decision maker is $n(n-1)/2$ and the amount of estimated weights is n. Therefore the amount of information per weight in a large matrix, such as 10×10 is far larger than in a small matrix that is for example 4×4. The amount of pairwise comparisons elicited from the decision maker for these matrices, are 45 and 6, respectively. This makes the amount of comparisons, which can be used to estimate the weights 4.5 and 1.5 per weight, respectively. Therefore one might think that removing 50% of the comparisons on both matrices would make it relatively more harder to estimate the 4 weights accurately from the small matrix compared to estimating the 10 weights from the larger one, because there would be still on average 2.25 comparisons left to estimate one weight in the larger one and on average only 0.75 comparisons in the smaller one. This might be the case, but the results of the simulation do not support this hypothesis, because the larger the matrix was, the less accurate the methods became. This might be explained by the fact that, when the order of the matrix increases, there are more incorrect ranks that can be given to the m-th weight. In a 4×4 matrix there are 3 possible wrong answers and in a 10×10 matrix there are 9.

The way the performance was measured placed restrictions on what effects of the missing comparisons could be observed. In this thesis it was measured in a narrow way. Keeping track on what rank these methods give to the m-th rank, is a demonstrative way to measure performance, but it is surely not the only way. Perhaps, measuring the error of every weight would yield more comprehensive results.
Neither method can be considered to be clearly better than the other based on these results. Both performed similarly, when the amount of missing comparisons was less than 20%. Some differences did become visible, when the amount of missing comparisons was increased. Harker's method got more extreme results, when the percentage of missing comparisons was over 44%, compared to the method of Shiraishi et al.. Harker's method ranked the m-th weight to the first rank and to the last rank more frequently. A possible explanation for this phenomenon is that the methods behave differently in some cases when the amount of missing comparisons is larger than $n-1$. Let us take an example where the complete pairwise comparison matrix is

$$\mathbf{A} = \begin{pmatrix} 1 & 1.24 & 1.27 & 0.70 \\ 0.81 & 1 & 0.86 & 0.68 \\ 0.78 & 1.17 & 1 & 0.46 \\ 1.42 & 1.47 & 2.17 & 1 \end{pmatrix}.$$

The weights estimated from $\mathbf{A}$ are $\mathbf{w} \approx [0.32, 0.26, 0.15, 0.27]$ and the ranking is then $\mathbf{r} = [1, 3, 4, 2]$. Three comparisons are then removed and we get:

$$\mathbf{A}_4(\mathbf{x}_3) = \begin{pmatrix} 1 & x_{12} & x_{13} & x_{14} \\ 1/x_{12} & 1 & 0.86 & 0.68 \\ 1/x_{13} & 1.17 & 1 & 0.46 \\ 1/x_{14} & 1.47 & 2.17 & 1 \end{pmatrix}.$$

There are no pairwise comparisons in the matrix $\mathbf{A}_4(\mathbf{x}_3)$, which would give information about the first weight. In other words, it is not connected to the other weights with pairwise comparisons. The two methods give different results when faced with this situation. Weights estimated from the incomplete matrix $\mathbf{A}_4(\mathbf{x}_3)$ with the method of Harker are $\mathbf{w} \approx [0, 0.27, 0.26, 0.47]$ and therefore $\mathbf{r}_H = [4, 2, 3, 1]$. Because the first weight was estimated to be zero, it was given rank four, instead of the rank one it actually should get. Method of Shiraishi et al. behaves differently when faced with the matrix $\mathbf{A}_4(\mathbf{x}_3)$. It will complete the matrix before estimating the weights and we get

$$\mathbf{A}_4^{'}(\mathbf{x}_3^{'}) = \begin{pmatrix} 1 & 0.57 & 3.2 & 1.62 \\ 1.71 & 1 & 0.86 & 0.68 \\ 0.31 & 1.17 & 1 & 0.46 \\ 0.68 & 1.47 & 2.17 & 1 \end{pmatrix}.$$

Weights estimated from $\mathbf{A}_4^{'}(\mathbf{x}_3^{'})$ are $\mathbf{w} \approx [0.25, 0.20, 0.20, 0.35]$ and the ranking $\mathbf{r}_S = [2, 3, 4, 1]$. Method of Shiraishi et al. ranked the first weight incorrectly to second place, but it did manage better than Harker's method, which placed it to the fourth place.

When the missing comparisons happen to be placed so that one of the weights is not connected to the others, Harker's method will estimate it to be zero and give it the last rank. Method of Shiraishi et al. completes the matrix before estimating the weights and it can give the unconnected weight other values than zero. This gives an advantage to the method of Shiraishi et al., when the unconnected weight is the m-th weight. Harker's method gives it the last rank, but the method of Shiraishi et al. can give it other ranks. This might explain why Harker's method gave more frequently the last rank to the m-th weight.
On the other hand, when the unconnected weight is not the m-th weight, it can give an advantage to Harker's method. The unconnected weight gets automatically the last rank, which reduces the competition for the first rank. Harker's method should do better especially in cases where two or more of the largest weights estimated from the complete matrix have similar values. For example when weights estimated from a full comparison matrix are [0.30, 0.32, 0.15, 0.23], and the m-th weight is therefore the second weight. When three or more comparisons are removed and the first weight happens to be unconnected from the rest, Harker's method will give it the rank four. This should make it more likely for the second weight to be estimated correctly as the largest one, because its best competitor got the last rank. The method of Shiraishi et al. can give the unconnected weight other value than zero, therefore the m-th weight has a lower chance of being ranked to first place, compared to Harker's method. This might explain why Harker's method was more likely to give the m-th weight its correct rank one, when the percentage of missing comparisons were high.
The fact, that the largest differences between the results of these methods were in cases with matrices of order 4 with 50% missing comparisons and matrices of order 7 with 48% missing comparisons, support this explanation. In these cases the amount of missing comparisons were larger than $n-1$, which is the minimum amount of comparisons that needs to be removed to have an unconnected weight. The same phenomenon is also visible in the case with matrices of order 10 with 44% missing comparisons. The reason why it is not as strongly visible in the larger matrix compared to the smaller ones is that in a matrix with 45 comparisons, from which 20 are missing, it is less likely for any weight to be unconnected, compared to a 4×4 matrix with 6 comparisons, from which 3 are missing.
Having an unconnected weight is not a realistic scenario. If this happens during an elicitation process, it would be wise to pressure the decision maker to make more comparisons, to ensure that there is enough information to estimate all of the weights. Therefore allowing weights to be unconnected in the simulations, is not a preferable situation. Simulations where the comparisons are removed in a way, which leaves all of the weights connected would produce more realistic results. Perhaps then there would be no significant difference in the results between the performance of the two methods.
The method of Shiraishi et al. was considerably more computationally demanding, which is an disadvantage, but in practice the time difference is irrelevant. Both methods can estimate the weights from a missing comparisons matrix within seconds, when the order of the matrix is less than one hundred.

References

[1] T. Saaty. A scaling method for priorities in hierarchial structures. *Journal of Mathematical Psychology*, 15(3):234–281, 1977.

[2] P. Harker. Alternative modes of questioning in the analytic hierarchy process. *Mathematical Modelling*, 9(3-5):353–360, 1987.

[3] S. Shiraishi, T. Obata, and M. Daigo. Properties of a positive reciprocal matrix and their application to AHP. *Journal of the Operation Research Society of Japan*, 41(3):404–414, 1998.

[4] G. Miller. The magical number seven plus or minus two: some limits on our capacity for prosessing information. *The Psychological Review*, 63(2):81–97, 1956.

[5] S. Bozóki and T. Rapcsák. On Saaty's and Koczkodaj's inconsistencies of pairwise comparison matrices. *Journal of Global Optimization*, 42(2):157–175, 2008.

[6] J. Alonso and M. Lamata. Consistency in the analytic hierarchy process: a new approach. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 14(4):445–459, 2006.

[7] R. Ureña, F. Chiclana, J. Morente-Molinera, and E. Herrera-Viedma. Managing incomplete preference relations in decision making: a review and future trends. *Information Sciences*, 302:14 – 32, 2015.

[8] F. Carmone, A. Kara, and S. Zanakis. A monte carlo investigation of incomplete pairwise comparison matrices in AHP. *European Journal of Operational Research*, 102(3):538 – 553, 1997.

[9] S. Bozóki, J. Fülöp, and L. Rónyai. On optimal completion of incomplete pairwise comparison matrices. *Mathematical and Computer Modelling*, 52(1–2):318 – 333, 2010.

[10] T. Obata, S. Shiraishi, M. Daigo, and N. Nakajima. Assessment for an incomplete comparison matrix and improvement of an inconsistent comparison: computational experiments. *International Symposium on the Analytic Hierarchy Process*, pages 200–205, 1999.

[11] M. Brunelli, L. Canal, and M. Fedrizzi. Inconsistency indices for pairwise comparison matrices: a numerical study. *Annals of Operations Research*, 211(1):493–509, 2013.

[12] S. Bozóki, L. Dezső, A. Poesz, and J. Temesi. Analysis of pairwise comparison matrices: an empirical research. *Annals of Operations Research*, 211(1):511–528, 2013.

Appendix A Table of results

Table 4: The percentage of instances of each rank was given to the m-th weight, when each scenario was simulated 50 000 times. S stands for method of Shiraishi et al. and H stands for Harker's method.

		Missing		The percentage of cases this rank was given to the m-th weight									
Method	Matrix	Comparisons	Rank	1	2	3	4	5	6	7	8	9	10
H	4 x 4	17 %		87,198	10,862	1,716	0,224	-	-	-	-	-	-
S	4 x 4	17 %		87,124	10,944	1,696	0,236	-	-	-	-	-	-
H	4 x 4	50 %		69,376	18,71	5,738	6,176	-	-	-	-	-	-
S	4 x 4	50 %		68,526	21,298	8,332	1,844	-	-	-	-	-	-
H	7 x 7	19 %		82,53	13,366	3,138	0,738	0,192	0,032	0,004	-	-	-
S	7 x 7	19 %		82,486	13,524	3,058	0,746	0,154	0,028	0,004	-	-	-
H	7 x 7	48 %		66,018	20,89	7,924	3,118	1,2	0,398	0,452	-	-	-
S	7 x 7	48 %		65,596	21,762	8,01	3,06	1,128	0,364	0,08	-	-	-
H	10 x 10	20 %		79,568	14,848	3,896	1,164	0,38	0,104	0,03	0,008	0,002	0
S	10 x 10	20 %		79,594	14,982	3,8	1,124	0,374	0,096	0,024	0,004	0,002	0
H	10 x 10	44 %		65,398	20,502	8,168	3,402	1,566	0,644	0,208	0,076	0,014	0,022
S	10 x 10	44 %		65,248	21,034	8,248	3,27	1,422	0,57	0,128	0,07	0,006	0,002

Appendix B Finnish summary

Tehdessämme arkisia valintoja, kuten päätämme, mitä syömme lounaaksi, ratkaisu tulee usein vaivattomasti mieleen. Toisinaan eteemme tulee kuitenkin niin vaikeita valintoja, että päätöksenteko ei onnistukaan yhtä helposti. Tällainen tilanne voisi olla vaikka asunnon osto. Markkinoilla on eri vaihtoehtoja, ja niiden asettaminen paremmuusjärjestykseen ei ole yksiselitteistä. Yksi asunto saattaa olla edullinen ja tilava, mutta sen sijainti ja kunto ovat huonoja. Toisen sijainti ja kunto saattavat olla hyviä, mutta asunto voi olla kallis ja ahdas. Jotta päätöksentekijä voisi valita itselle parhaan asunnon, hänen täytyy päättää miten painottaa valittuja kriteerejä hintaa, kokoa, sijaintia ja kuntoa. Tällaisia tilanteita varten on kehitetty useita työkaluja, ja yksi suosituimmista on analyyttinen hierarkiaprosessi (AHP). Sitä käyttämällä päätöksentekijä pystyy lähestymään ongelmaa analyyttisesti ja ohjatusti. AHP:n tavoitteena on selvittää päätöksentekijän mieltymykset ja niitä käyttämällä laskea pisteet jokaiselle vaihtoehdolle. Suurimman pistemäärän saanut vaihtoehto on oletettavasti päätöksentekijälle paras valinta.

Sovelletaan AHP:tä edellä mainittuun asunnon valintaan. Tällöin päätöksentekijältä kysytään kysymyksiä, kuten “Miten paljon tärkeämpi hinta on verrattuna kokoon?” Tähän päätöksentekijä antaa luvun väliltä 1/9 ja yhdeksän. Hänen täytyy verrata jokaista kriteeriä keskenään. Nämä luvut tallennetaan parivertailu matriisiin $\mathbf{A}$. Sen vaaka- ja pystyrivin alkioiden määrä on sama kuin valittujen kriteerien määrä. Tässä esimerkissä $\mathbf{A}$:ssa on neljä riviä ja saraketta. AHP:ssa estimoidaan jokaiselle kriteerille painokerroin. Tämä tapahtuu esimerkiksi ominaisvektorimenetelmällä, jossa otetaan $\mathbf{A}$:n suurinta ominaisarvoa vastaava ominaisvektori ja skaalataan sen alkiot siten, että niiden summa on yksi. Tuloksena saadaan vektori $\mathbf{w}$, jonka alkiot ovat suoraan valittujen kriteerien painokertoimet. Niillä voidaan selvittää päätöksentekijälle paras vaihtoehto. Painokertoimien arvot ovat välillä nolla ja yksi. Mitä suurempi kriteerin kerroin on, sitä tärkeämpi kyseisen kriteerin uskotaan olevan päätöksentekijälle.

AHP:ssa oletetaan, että parivertailumartiisi on täysi, eli päätöksentekijältä on saatu tarvittavat parivertailut. Todellisuudessa osa parivertailusta voi puuttua. Näin voi käydä esimerkiksi, jos päätöksentekijällä ei ole tarvittavaa asiantuntemusta vastata kysymyksiin. Käytetään jatkossa merkintää $\mathbf{A}(\mathbf{x}_k)$ kuvaamaan matriisia, josta on poistettu k parivertailua. AHP:ta ei voi soveltaa sellaisenaan puutteelliseen parivertailumatriisin, ja tämän takia on kehitetty useita metodeja, joilla painokertoimet voidaan estimoida $\mathbf{A}(\mathbf{x}_k)$:sta puuttuvista parivertailuista huolimatta, ja tässä työssä on käytetty niistä kahta. Ensimmäinen on Harkerin menetelmä ja toinen on Shiraishin menetelmä. Harkerin menetelmässä, $\mathbf{A}(\mathbf{x}_k)$ muokataan yksinkertaisten sääntöjen mukaan apumatriisiksi $\mathbf{C}$ ja tästä matriisista voidaan estimoida painokertoimet ominaisvektorimenetelmällä. Shiraishin menetelmä on hiukan monimutkaisempi. Siinä puuttuvat parivertailut yritetään täyttää siten, että matriisista tulisi mahdollisimman johdonmukainen, kun puuttuvat parivertailut täydennetään. Johdonmukaisuudella tarkoitetaan sitä, että jos päätöksentekijän mielestä hinta on kolme kertaa tärkeämpi kuin koko ja, jos koko on kaksi kertaa tärkeämpi kuin kunto, silloin hinnan pitäisi olla kuusi kertaa kuntoa tärkeämpi. Jos oletetaan näin olevan ja ei tiedetä, miten tärkeä koko on verrattuna kuntoon, voidaan sen suuruus päätellä näiden kriteerien tärkeydestä verrattuna hintaan. Jos päätöksentekijä on johdonmukainen, koon pitäisi olla kaksi kertaa kuntoa tärkeämpi.

Harkerin ja Shiraishin menetelmiä käytettiin tässä opinnäytetyössä tehdyissä numeerisissa simulaatioissa. Simulaatioiden tarkoituksena oli selvittää, miten paljon puuttuvat parivertailut häiritsevät painokertoimien estimointia ja miten nämä kaksi menetelmäaa suoriutuvat toisiinsa verrattuina. Simulaatioissa täydestä parivertailumatriisista estimoitiin painokertoimet, ja suurin kerroin tallennettiin. Sitten samasta matriisista poistettiin joitakin parivertailuja, minkä jälkeen Harkerin ja Shiraishin menetelmien avulla estimoitiin uudet kertoimet. Jos sama kerroin oli edelleen suurin, katsottiin, että menetelmä toimi hyvin ja sille tallennettiin tilastoon numero yksi. Jos kyseinen kerroin ei ollut enää suurin, tallennettiin sen sijainti kertoimien suuruusjärjestyksessä. Esimerkiksi, jos suurin painokerroin putoaa parivertailujen poiston jälkeen kolmannelle sijalle, merkattiin tilastoon numero kolme. Mitä useammin menetelmät saivat merkinnän yksi, sitä paremmin niiden katsottiin suoriutuvan.

Simulaatioissa käytettiin kolmea erityyppistä matriisia. Niiden koot olivat 4×4, 7×7 ja 10×10. Niistä poistettiin joko noin 20 % parivertailuista tai noin 50 %. Poistettavien parivertailujen määrä on ennalta valittu, mutta poistettavien alkioiden sijainti on satunnainen. Kullekin koeasetelmalle luotiin 50 000 matriisia ja Harkerin ja Shiraishin menetelmiä käytettiin, kuten edellisessä kappaleessa kuvattiin.

Puuttuvat parivertailut heikensivät molempien menetelmien kykyä estimoida painokertoimet tarkasti, erityisesti kun matriisien koko kasvoi. Tämä voidaan selittää sillä, että suuremmissa matriiseissa on enemmän painokertoimia ja niiden suuruusjärjestys voi vaihtua helpommin, kuin jos painokertoimien lukumäärä on pienempi. Molemmat menetelmät suoriutuivat hyvin samanlaisesti. Kumpaakaan ei voi pitää selvästi toista parempana, sillä erot olivat hyvin pieniä, erityisesti kun puuttuvien parivertailujen määrä oli alle 20 %. Joitakin eroja kuitenkin syntyi, kun puuttuvien parivertailujen määrä oli suuri ja erityisesti kun matriisit olivat pieniä. Näissä tapauksissa Harkerin menetelmät estimoi useammin oikean kertoimen suurimmaksi, kuin Shiraishin menetelmä. Toisaalta näissä tapauksissa Harkerin menetelmä estimoi kertoimen virheellisesti kaikkein pienimmäksi useammin kuin Shiraishin menetelmä. Shiraishin menetelmä siis estimoi kertoimen harvemmin täysin oikein, mutta epäonnistuessaan se ei tehnyt yhtä suuria virheitä.

Selitys menetelmien tulosten eroihin saattaa löytyä erikoistapauksista, joissa jonkin painokertoimen kaikki parivertailut on poistettu. Tällaista kerrointa voidaan kutsua irralliseksi. Jos asuntoesimerkin parivertailumatriisista poistetaan kaikki kolme vertailua, joissa esiintyy hinta, on hinnan kerroin tällöin irrallinen. Tällöin ei ole olemassa mitään informaatiota, josta hinnan painokerroin voitaisiin mielekkäästi estimoida. Harkerin menetelmä arvioi tällaisessa tilanteessa hinnan painokertoimen

nollaksi, joka johtaa siihen, että sen uskotaan olevan pienin kerroin. Tämä on huonoin mahdollinen tulos, sillä hinnan kerroin oli todellisuudessa suurin. Shiraishin menetelmällä on tässä tilanteessa etu, sillä se voi päätyä toisenlaiseen tulokseen. Se täyttää matriisin, ennen kuin se arvio kertoimet, mikä johtaa siihen, että se kykenee estimoimaan irralliselle painokertoimelle jonkin muun arvon kuin nollan.
Syy siihen miksi, Hiraishin menetelmä onnistui estimoimaan kertoimen useammin täysin oikein, saattaa johtua myös irrallisista kertoimista. Jos toiseksi suurin kerroin on irrallinen, Harkerin menetelmä arvioi sen nollaksi. Tällöin oikea painokerroin tulee todennäköisemmin arvioiduksi suurimmaksi, sillä sen pahin kilpailija arvioitiin pienimmäksi kertoimeksi. Shiraishin menetelmä taas voi arvioida irrallisen kertoimen joksikin muuksi kuin nollaksi, joten kilpailu suurimman kertoimen paikasta on hiukan kovempi kuin Harkerin menetelmällä.

Kumpikaan metodeista ei ollut selvästi parempi kuin toinen. Eroa syntyi todennäköisesti vain siitä syystä, että poistettavat painokertoimet valittiin täysin satunnaisesti. Jos parivertailut poistettaisiin siten, että yksikään painokerroin ei jää irralliseksi, erot menetelmien välillä todennäköisesti pienenisivät. Harkerin menetelmällä on yksi etu, nopeus. Sen käyttäminen oli yli yhdeksän kertaa nopeampaa kuin Shiraishin menetelmän.