The influence of background knowledge when measuring utility

A comparison between BMG students and Economic students

Michelle Hartmanns

302088mh

Erasmus University Rotterdam, School of Economics

Applied Economics, Health Economics

Bachelor Thesis

Supervisor: dr. A. Baillon

August 2010

Abstract:

The present study is done to evaluate whether economic background knowledge or health related background knowledge has an influence on the results when measuring utility.

The comparison is done with three groups of students (Economics, Policy & Management in Health care and Economics & Law) and their utilities for three different health states using the PE-method and the PLE2-method.

No significant results are found, but there are indications that respondents with economic background knowledge are less likely to give inconsistent answers in the questionnaire and therefore are less likely to be excluded from the analysis. Future studies with larger groups of respondents are needed to further test these indications.

Keywords:

Utility, background knowledge, economic knowledge, experts on the field, biases.

Introduction:

In health economics it is important to know how people feel about certain health states, which value they give those health states. This information can be used in multiple ways, for example to see whether some treatment, drugs or operation, is worth it's costs or simply to know which health state is preferred to another. For this, full health is considered the best option and death is often considered the worst.

Such measurements are called utility elicitation. The utility from a person for one particular health state is searched. There are many methods to do so, mostly used are the standard gamble method and the time trade off method.

All methods cause a lot of different biases that need to be excluded from the results. These biases are the result of different backgrounds (such as social-economic class), sex (male or female), age and more of such characteristics. Little differences between the respondents can cause major biases in the results of a study. To exclude such biases it is vital that the characteristics that cause the biases are known and registered during the study. The characteristics mentioned above are known to cause biases and are well studied.

The present study focuses on a different kind of characteristic, the level of economic background knowledge. When measuring utility there are a number of economic theories used and even the methods of measuring, standard gamble and time trade off for instance, are based on assumptions that are deducted from economic knowledge.

This makes it interesting to see how respondents with economic background knowledge answers questions in health related issues.

In this study we also use a group of BMG students (literally: policy and management in health care) who are familiar with health related questions. That can be seen as background knowledge of the subject.

Results could lead to the insight that this is yet another characteristic that researchers need to take in to account when measuring utility in the future. Alongside with age and sex, background knowledge could become a selecting criteria for these types of study.

Review of literature:

Not much is known about the influence of background knowledge on the measurement of utility. In the article of Rosen, Tsai and Downs (**ROSEN ET AL., 2003**[1]) is a significant difference found in risk attitude between different levels of education. This does not implicate though, that the same applies for the kind of education on the same level.

The difference in kind of study is examined in the article of Frank and Schulze (**FRANK ET AL., 1999**[2]). This study showed that economic students are significantly more corrupt than others (being non-economic students). The difference is due to self-selection, more than indoctrination during the study. This means for this study that we could expect a difference in the way economic and non-economic student answer the questions.

On the other hand, the BMG-students are used to these kinds of questionnaires and could therefore be considered 'experts from the field' as mentioned in the article of Fatas, Neugebauer and Tamborero (**FATAS ET AL., 2003**[3]). The experts in their study were politicians who were involved in decisions in relationship to economic political dilemmas. The research question in that study was whether prospect theory predicts the expert respondents better than rational choice theory (expected utility). The pool of expert respondents was compared to a pool of economic students who were considered non-experts.

Hypotheses

From the literature mentioned above I conclude that there are two hypotheses to work with. First it makes sense that BMG student, because of their education, could be called 'experts from the field' as mentioned in **FATAS ET AL., 2003**[3]. This assumption leads to the hypothesis that BMG students, when compared to others, give 'better' answers. Better is this case means more consistent with the theories and more rational. This leads to hypothesis 1 where we expect that the utilities derived from the questionnaire for non Economic students (being the BMG group) differ from the utilities for Economic students.

Hypothesis 1:

H_0 There is no difference between the utilities of Economic and non Economic students.
H_1 There is a difference between the utilities of Economic and non Economic students.

On the other hand, from another perspective economic students are more likely to be corrupt, as shown in **FRANK ET AL., 1999**[2]. In this present study this could mean that these students are more likely to think ahead and try to 'predict' the next question and make their answers depending on that.

The economic theories and assumption underlying the questionnaire make that economic student are more likely to be able to foresee the thought behind the questions. One of the most important underlying assumption of the standard gamble theory is that of rational choice. People are expected to think rational and make the best choice available. Economic

students are more than familiar with this concept, they are taught about this concept since day one of their study.

This familiarity leads to recognition of the concept in questions asked and the economic students will try to 'live up to' the rules of rational choice. This leads to hypothesis 2 where we expect that Economic students give more rational answers, more concrete: we expect to see less inconsistent answers with the Economic students. The hypothesis has no sign, but we expect that Economic students have less inconsistencies.

Hypothesis 2:

H_0 There is no difference in consistency level between Economic and non Economic students.
H_1 There is a difference in consistency level between Economic and non Economic students.

Methods:

Participants:

The respondents of our questionnaire all study at the Erasmus University Rotterdam. We selected our respondents by asking friends and acquaintances to fill in the questionnaire. From there it become more or less a snowball-effect. Because of the need for BMG as well as economic students, we of course focused on those groups. We tried to keep the fraction of men and women representative for the focus group. In table 1 a summary is given of all the respondents and their characteristics.

n = 70		
Characteristic	Number of respondents	% of respondents
Sex		
Male	29	41,4%
Female	41	58,6%
Study		
BMG	43	61,4%
Economics	9	12,9%
Mr.Drs.	18	25,7%
Study phase		
Bachelor 1	13	18,6%
Bachelor 2	9	12,9%
Bachelor 3	35	50,0%
Master	13	18,6%

Table 1: basic characteristics of respondents

The average age of the respondents is 21,2 years and the median is 21. Despite of the outlier of 28, we consider the group as almost equally old. Together with the same University, this leads to a group with only differences in the kind of study.

All questionnaires are answered between June 18th and July 22nd 2009.

Study design:

The questionnaire dealt with three different methods and three different health states, leading to nine different questions asked to the respondents.

Utility measurement methods

- probability equivalence (PE)
- probability lottery equivalence I (PLE I)
- probability lottery equivalence II (PLE II)

These methods are all based on standard gamble. Standard gamble is one way to elicit utilities. Other way are for example the Time Trade Off method or the visual analogue scale. The standard gamble method searches for the indifferent point of people, when choosing between a situation with and a situation without risk.

The PE method aims to find the probability for which the respondent is indifferent between a certain health state for sure and a risky outcome. The respondent can choose between two options; the certainty of health state I or a chance at full health, though with the risk of dying. The PE methods find the p for which the respondent no longer can make a decision between the two options and is, therefore, indifferent. The situation of the PE method is schematically described in figure 1.

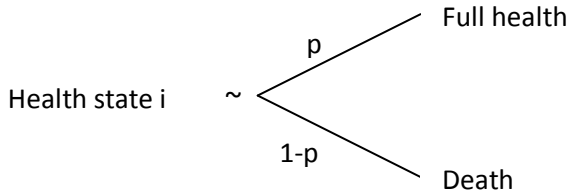

figure 1 - PE method

The PE method is called a riskless-risk method, because of the certainty given when choosing Health State i.

The second method used for this thesis is the PLE II. This method is introduced by a paper by Bleichrodt et al. in 2007 (**BLEICHRODT ET AL., 2007**[4]) and is called a risk-risk method. Questions concerning this method give the respondents the choice between a chance of x% of health state i and 1-x% of dying and the chance of p% of full health and 1-p% of dying. In the paper Bleichrodt et al. used for x the percentage of 35. For this theses we used 90%, to make sure the respondents could make a distinction between the two options.

The situation in the PLE II method is schematically described in figure 2.

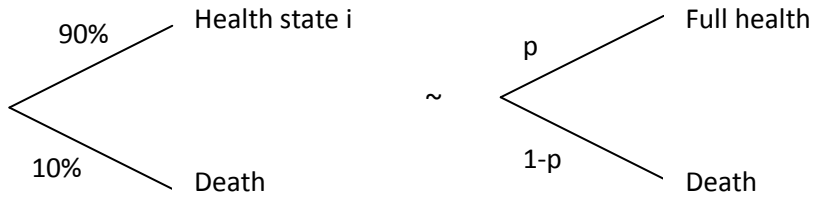

figure 2 – PLE II method

For this thesis PLE I method will not be used because of its complications when assuming expected utility, though it is recorded in the data. This method was discussed by Farquhar in 1984 (**FARQUHAR, 1984**[5]). Where in PLE II the first gamble is between health state i and death, Farquhar uses two health states i and j. The situation then looks like described in figure 3.

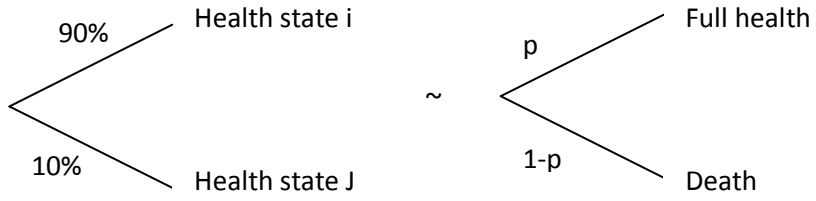

figure 3 – PLE I method

Here health state i is better than health state j and j is better than k.

Heath states used

To choose the health states used in the questionnaire, we used the EuroQol 5D instrument. This is easy way to show the condition of the illusive disease the respondent has to take in mind when answering the questions. It a standardized method, which allows applying a formula to calculate utility. We use it the other way around and try to find the personal utilities for each respondent.

The EuroQol uses five dimensions to represent the health state of a patient. Those dimensions are:

- mobility
- self-care
- usual activities
- pain/discomfort
- anxiety/depression

The descriptions of the health states we used, are enclosed as appendix 1.

We named the health states X, Y and Z, mostly to make a good distinction between the health states and the questions, which were numbered A to I. For the health states applies:

X > Y > Z

Calculating the utility values

The results of the questionnaire are the probabilities for which the respondents are indifferent in the various questions. These need to be calculated in to utilities. The two different methods ask for two different ways of calculating those utilities.

PE-method

Because the PE method compares a risky option to a riskless option, the utility is equal to the probability found. This leads to the following equation:

$$U(x) = p_x$$

The prospect (p:H,K), with health state H with probability p and health state K with probability (1-p), is represented by

$$p*u(H) + (1-p)*u(K)$$

We assume that u(Full health) = 1 and u(Death) = 0.

All this has as consequence, that when comparing health state X with a probability of full health or death (X ~ (p:Full health, Death)), we get the following:

$$u(X) = p*1 + (1-p)*0 = p$$

This explains the PE-method and the earlier mentioned equation of $u(X)=p_x$.

PLE II-method

Things are a little bit more complicated when dealing with the PLE II method. As discussed earlier this method is a risk-risk method. This means that we no longer compare a health state for certain and a uncertain outcome. Both options involve risk. In the questionnaire the probability of 90% (0.90) was used.

(0.90:X, Death) ~ (p:Full health, Death)

This implies that

$$0.90*u(X) + 0.10*0 = p*1 + (1-p)*0$$

And hence:

$$u(X) = p/0.90$$

This equation can now be used to calculate the utility values when considering the PLE II-method.

Results:

Excluded respondents

Not all respondents were valid to be taken in to account during the analyses. The respondents excluded were excluded due to inconsistencies in their answers. These inconsistencies were determined in SPSS by using an condition to select cases. Most inconsistencies consist of the fact the proper order of health states, as mentioned before X>Y>Z (or X>=Y>=Z), does not show in the utilities found for that respondent.

In two cases the respondent was excluded from the analysis because he or she answered the question in such way that the found utility for health state X in the PLE II method exceeded the value 1.00. Since this was the value given to 'full health', the elicited utility does not correspond with the description of the health states.

Some respondents argued that living beneath their current standard, being full health, will be equally bad for them. The utilities elicited show that by equal utilities with two or more health states. These respondents were not excluded, since this is a reasonable explanation and does not indicate, in contrast to the earlier mentioned inconsistency, that the respondent does not understand the question.

This leads to the conclusion that some respondents should not be taken in to consideration for the analysis. This is only true for the independent sample t-test though, since the crosstabs-test (Chi squared and Lamba) is based on the inconsistencies. The crosstabs-test are used to find out whether there is a difference in number of inconsistencies between economic students, being group 2 (Economic) and 3 (Mr.Drs.), and non-economic students, being group 1 (BMG).

An overview of all respondents excluded, ten in total, can be found in table 11 in appendix 3.

Hypothesis 1

Independent sample t-test PE-method

To test whether the utilities of the two groups differ, we use the independent sample t-test. This test first compares the variances which are then used to compare the means. As shown in following table 2, the differences in means are not significant, since al values are above 0.05, the confidence level.

	Economics in study	N	Mean	Std. Deviation	Mean Difference	Sign. t-test equal means
Utility X PE	No economics	34	0.77485	0.210819	-0.071301	0.195
	Economics	26	0.84615	0.205861		
Utility Y PE	No economics	34	0.61647	0.233600	-0.020645	0.764
	Economics	26	0.63712	0.297678		
Utility Z PE	No economics	34	0.41912	0.281165	-0.017805	0.819
	Economics	26	0.43692	0.316149		

Table 2

Independent sample t-test PLE2-method

The same as mentioned above goes for the PLE2-method. The differences in means between the compared groups are not significant, since the values are between 0.640 and 0.774.

	Economics in study	N	Mean	Std. Deviation	Mean Difference	Sign. t-test equal means
Utility X PLE2	No economics	34	0.80327	0.225681	-0.024937	0.675
	Economics	26	0.82821	0.228305		
Utility Y PLE2	No economics	34	0.66830	0.235962	0.033899	0.640
	Economics	26	0.63440	0.323167		
Utility Z PLE2	No economics	34	0.45915	0.275899	0.021757	0.774
	Economics	26	0.43739	0.307730		

Table 3

Hypothesis 2

Independent sample t-test difference PE-method and PLE2-method

To test hypothesis 2 we compare the means of the difference in value between the PE-method and the PLE2-method for one subject. This new variable is an indication for inconsistencies, as both methods should elicit the same utility value. Again the independent sample t-test is used to see whether the two groups differ significantly. The results are not significant.

	Economics in study	N	Mean	Std. Deviation	Mean Difference	Sign. t-test equal means
Difference PEX-PLE2X	No economics	34	-0.02842	0.134740	-0.046364	0.194
	Economics	26	0.01795	0.136629		
Difference PEY-PLE2Y	No economics	34	-0.05183	0.168619	-0.054544	0.294
	Economics	26	0.00271	0.230484		
Difference PEZ-PLE2Z	No economics	34	-0.04003	0.121248	-0.039563	0.424
	Economics	26	-0.00047	0.225757		

Table 4

Crosstabulation Economics in study * Inconsistencies PE-method

Crosstabs are another way of testing whether the level of inconsistencies differ significantly between the two groups. This test compares the expected count and the real count. This leads to the Pearson Chi Square test. The higher this value, the more the real count differs from the expected count. In table 5, 6 and 7 the real counts and expected counts are given per group as well as in total. Table 5 shows that the 'No economics' group does have more inconsistencies than expected. The Chi Square test based on table 5 gave a value of 1.968 though, which is not high enough to speak about significant differences.

			Inconsistencies PE-method		Total
			No inconsistencies	Inconsistent	
Economics in study	No economics	Count	40	3	43
		Expected count	41.2	*1.8**	43.0
	Economics	Count	27	0	27
		Expected count	25.8	*1.2**	27.0
Total		Count	67	3	70
		Expected count	67.0	3.0	70.0

Table 5: Value of Pearson Chi-Square test is 1.968, value of Lambda is 0.000
** these cells have expected counts less than 5*

Crosstabulation Economics in study * Inconsistencies PLE2-method

In table 6 we also see that the 'No economics' group have more inconsistencies than expected, but the value of the Chi Square test is again to low to differ significantly. The expected count cells have a value less than 5, which indicates that the Pearson Chi Square test cannot be performed perfectly. The values are shown in italics.

			Inconsistencies PLE2-method		Total
			No inconsistencies	Inconsistent	
Economics in study	No economics	Count	37	6	43
		Expected count	38.7	*4.3**	43.0
	Economics	Count	26	1	27
		Expected count	24.3	*2.7**	27.0
Total		Count	63	7	70
		Expected count	63.0	7.0	70.0

Table 6: Value of Pearson Chi-Square test is 1.936, value of Lambda is 0.000
** these cells have expected counts less than 5*

Crosstabulation Economics in study * Overall inconsistencies

By overall inconsistencies we mean that the respondent has an inconsistency in one or in both methods. Again the table shows more counts than expected in the No economics*Inconsistent cell. The value of the Chi Square test is 4.020, the highest of the three Chi Square tests. The cell with the expected count for the inconsistencies in the Economics group has a value less than 5, so the Pearson Chi Square test cannot be performed as wanted.

			Overall inconsistencies		Total
			No inconsistencies	Inconsistent	
Economics in study	No economics	Count	34	9	43
		Expected count	36.9	6.1	43.0
	Economics	Count	26	1	27
		Expected count	23.1	*3.9**	27.0
Total		Count	60	10	70
		Expected count	60.0	10.0	70.0

Table 7: Value of Pearson Chi-Square test is 4.020, value of Lambda is 0.000

** these cells have expected counts less than 5*

Interpretation:

Hypothesis 1

There is no significant evidence found that would lead to reject H_0. The significant levels of the independent sample t-tests a never below 0.05 (the confidence level is 5%).

Hypothesis 2

There is no significant evidence found that would led to reject H_0. The values of the Pearson Chi Square test and the value of Lambda are too low to do so.

Discussion:

The findings of this study are not significant. This means for hypothesis 1 that the utility values of economic and non-economic students don't differ a lot. For studies on this field done in the future this means that this aspect of respondent does not need to be taken in to account.

In the case of hypothesis 2, the inconsistency levels, the results show no significance either. The crosstabs do show the tendency that non economic students have more inconsistencies in their answers than economic students. This fact is a hint in the direction that economic students do have the tendency to 'predict' the questions and the ability to answer in line with the underlying assumptions, mostly that of rational choice. However, with these study results it is not possible to give scientific significant proof of this 'hint'.

We expect that a broader study with more respondents in each group would indeed show a significant difference in the inconsistency levels of both groups. The Chi Square test in this study is not always as accurate as hoped, since some cells of expected counts have a value less than 5. This problem would likely be solved with more respondents.

As explained in the review of literature there are two forces we need to consider for the differences between Economic and BMG students. First of all BMG students can be called experts from the field since they are used to health related questionnaires, so we expect they answer the questions better. On the other Economic students are more known with the underlying assumption and therefore we expect them to do better. In this study it is not clear which 'force' has more effect on the final results. A third test group, control group, is needed to test that. It is possible that the two forces mentioned work in opposite directions and therefore eliminate each other's effect.

In this study the group called 'economic students' actually consists two subgroups, students studying just Business and Economics and students who combine that study with Dutch Law, the so called 'Mr.Drs. students'. The combining of these groups is based on the fact that both groups have the same level of economic background in comparison with BMG students, which are considered the non economic students.

The combining of these groups is done after the questionnaire. In the original dataset these groups are mentioned differently. It could be interesting to test the differences between these two groups as well, but again the groups need to be larger to do so.

When comparing Economic students and Mr.Drs. students the influence of Law could be tested as well. It is possible that students with a la background are more 'sensitive' in answering these kind of questions.

In the questionnaire we used three methods to elicit the utilities of the respondents. In this study only two of those are used. The means of the variable 'Difference PE and PLE2' are very small though and no other significant differences between the two methods are found in this study.

For the exclusion of students we use a few criteria. First of all students who's utilities are not in line with the assumption X>Y>Z were excluded, but utilities like X=Y=Z (or X=Y>Z or X>Y=Z) are not excluded. In other words, we use the criteria X>=Y>=Z instead of only X>Y>Z. You could argue that those students should be excluded as well. We choose not to do so, because it is possible that people find one or more health states equally bad, considering their current health state.

During the questionnaire we tried to make sure that the respondents understood the questions. For some that wasn't easy though. This makes it hard to distinct which students simply did not understand the question(s) and which student rationally chose a percentage. Making more notes of the comments people made, could have made this task easier. The comments could give an indication of what the respondents were thinking during their questionnaire. Excluding based on the comments would have to be done manually though, which could lead to further biases.

Conclusion:

Both null hypotheses cannot be rejected, since no evidence for a difference was found. As earlier said in the discussion further studies with more respondents could lead to the insight that there is a difference after all. When such differences are found, this would mean that in further studies the influence of background knowledge of economics needs to be taken in to account when selecting respondents or when judging the results.

Acknowledgements:

I would like to thank Dr. Aurélien Baillon for all his patient with me and his guidance along the rocky road, Karen van Hedel for setting up and performing the questionnaire with me, Alex Verhof for refreshing my SPSS knowledge and last but not least al the respondents for their time and effort.

References:

1. Rosen, A.B., Tsai, J.S. & Dowsn, S.M. (2003). Variations in Risk attitude across Race, Gender, and Education. Medical Decision Making, 2003; 23; 511
2. Frank B. & Schulze G.G. (1999). Does economics make citizens corrupt?. Journal of Economic Behavior & Organization, Vol. 43 (2000) 101-113
3. Fatas, E., Neugebauer, T. & Tamborero, P. (2003). How Politicians Make Decisions: A Political Choice Experiment. Journal of Economics, Vol. 92 (2007), No. 2, pp. 167-196
4. Bleichrodt, H., Abellan-Perpiñan, J.M., Pinto-Prades, J.L. & Mendez-Martinez, I. (2007). Resolving Inconsistencies in Utility Measurement Under Risk: Tests of Generalizations of Expected Utility. Management Science, Vol. 53, No. 3, March 2007, pp. 469-482
5. Farquhar, P.H. (1984). Utility Assessment Methods. Management Science, Vol. 30, No. 11, November 1984, pp. 1283-1300

Appendix 1

Health state X

EuroQol dimension	State
Mobility	Some problems walking around.
Self-care	Some problems performing self-care, such as washing and dressing.
Daily activities	No problems with daily activities, such as work, sport or study.
Pain/discomfort	Moderate pain or discomfort
Anxiety/depression	Moderately anxious or depressed

Table 8: Health state X

Health state Y

EuroQol dimension	State
Mobility	Some problems walking around.
Self-care	Some problems performing self-care, such as washing and dressing.
Daily activities	Unable to do daily activities, such as work, sport or study.
Pain/discomfort	Moderate pain or discomfort
Anxiety/depression	Moderately anxious or depressed

Table 9: Health state Y

Health state Z

EuroQol dimension	State
Mobility	Confined to bed.
Self-care	Some problems performing self-care, such as washing and dressing.
Daily activities	Unable to do daily activities, such as work, sport or study.
Pain/discomfort	Moderate pain or discomfort
Anxiety/depression	Moderately anxious or depressed

Table 10: Health state Z

Appendix 2

In this appendix the English translation of the Dutch questions are given. All respondents answered the questions in Dutch.

Note that the respondents were first given an example question before answering the questions of a certain method. These examples had different percentages than shown here, mainly to simplify the question, and were 99% for full health and 1% for death, or just the other way around. The versions were given to the respondents randomly. The figures shown in this appendix are examples of how the 'real' questions started out, with the percentages 50%-50%.

Example question PE-method

Imagine you are suffering from a disease which, in a few weeks, will decline your quality of life so that you will be in health state X. Aside from the effect it has on your quality of life, it also has an effect on the length of your life. In this health state you will live for 43 more years. You can choose to undergo a medical treatment that might stop the reduction in quality of life. The treatment has two possible outcomes, your health will restore your health and you will live for 43 years in full health, or you could die within a few weeks. The choice has two options: will you treat the disease or not.

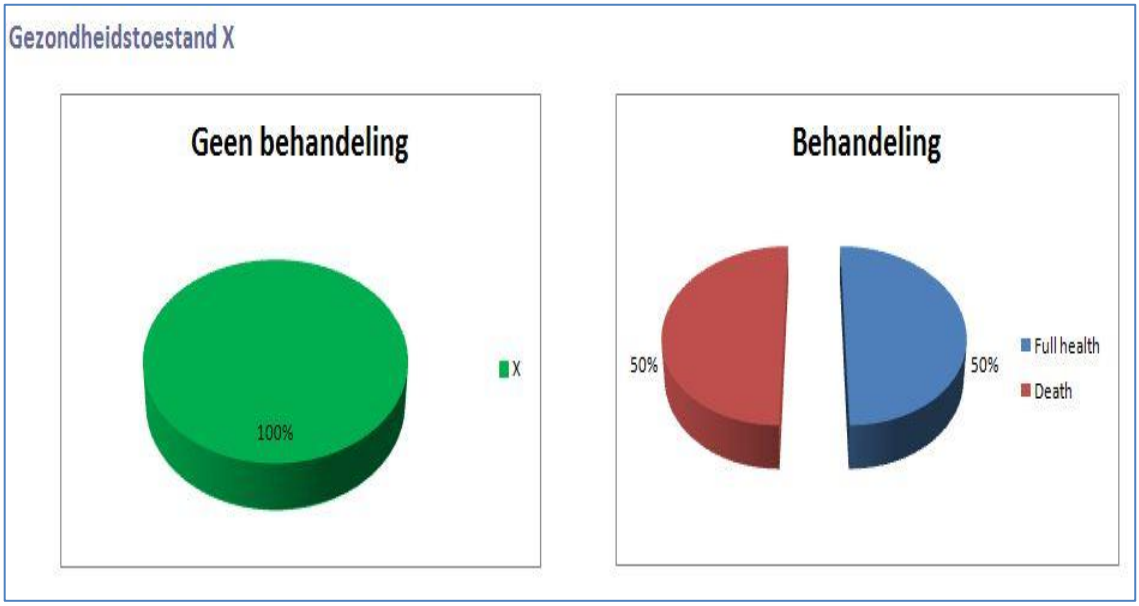

figure 4 – PE-method question

Example question PLE II-method

Imagine that you are suffering from disease that influences your health. The exact influence is unknown though. The doctors give you a 90% chance to live for 43 more years in health state X and a 10% chance to die within a few weeks. You can choose to undergo a treatment which might stop the reduction in quality of life. This treatment can restore your health and you will live for 43 more years in full health, or you can die within a few weeks. The choice has two options: will you treat the disease or not.

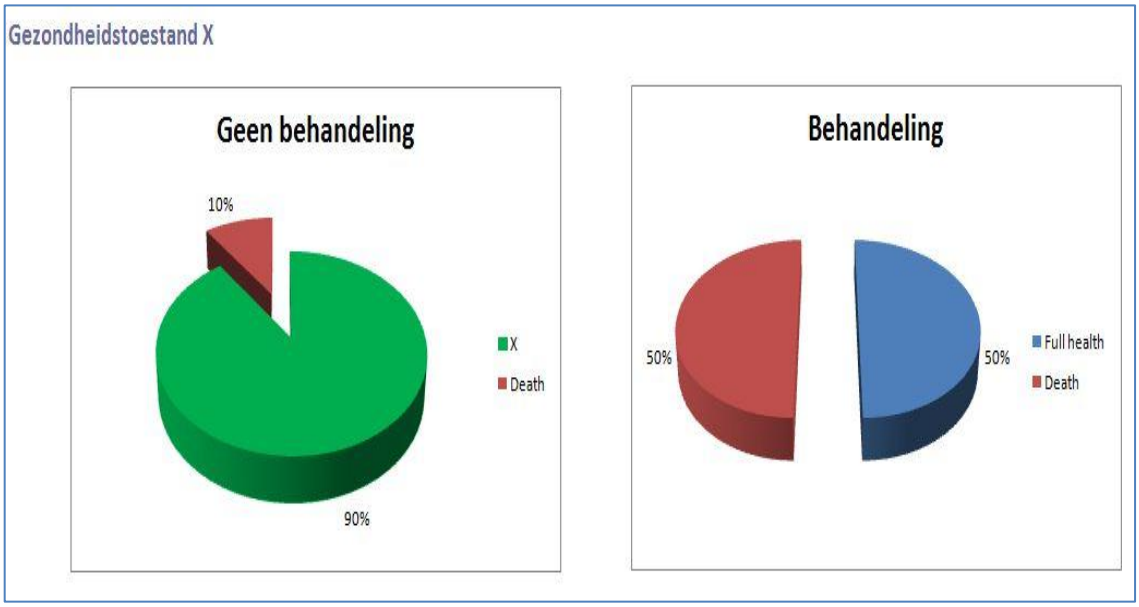

figure 5 – PLE II-method question

Appendix 3

Respondents excluded

Respondent number	Method	Reason	Study
1	PLE II	X>1,00	BMG
4	PE	X>Z>Y	BMG
5	PE	X>Z>Y	BMG
9	PLE II	Y>X>Z	BMG
14	PLE II	X>Z>Y	BMG
17	PLE II	Y>X>Z	BMG
30	PLE II	X>1,00	BMG
39	PLE II	X>Z>Y	Economics
43	PLE II	Y>X>Z	BMG
48	PE	Z>X>Y	BMG

Table 11: respondents excluded