A potential gender bias in assessing quality of life – a standard gamble experiment among university students

Leath Al Obaidi¹
Jörg Mahlich²,³

¹University of Nottingham, School of Economics, Nottingham, UK; ²Health Economics, Janssen KK, Tokyo, Japan; ³Düsseldorf Institute for Competition Economics (DICE), University of Düsseldorf, Düsseldorf, Germany

Correspondence: Jörg Mahlich
Health Economics, Janssen Pharmaceutical K.K., 5-2, Nishi-kanda 3-chome, Chiyoda-ku, Tokyo 101-0065, Japan
Email jmahlich@its.jnj.com

Background: There are several methodologies that can be used for evaluating patients’ perception of their quality of life. Most commonly, utilities are directly elicited by means of either the time-trade-off or the standard-gamble method. In both methods, risk attitudes determine the quality of life values.

Methods: Quality of life values among 31 Austrian undergraduate students were elicited by means of the standard gamble approach. The impact of several variables such as gender, side job, length of study, and living arrangements on the quality of life were identified using different types of regression techniques (ordinary least squares, generalized linear model, Betafit).

Results: Significant evidence was found that females are associated with a higher quality of life in all specifications of our estimations.

Discussion: The observed gender differences in quality of life can be attributed to a higher degree of risk aversion of women. A higher risk aversion leads to a higher valuation of given health states and a potential gender bias in health economic evaluations. This result could have implications for health policy planners when it comes to budget allocation decisions.

Keywords: quality of life, gender, risk aversion, standard gamble, students

Background
In some countries such as the UK, health technology assessment agencies evaluate new medical interventions like new pharmaceuticals in the framework of a cost utility analysis. While the additional costs of a given new intervention can, on the whole, be easily measured in monetary units, the benefit of this intervention is much harder to evaluate. A cost utility analysis employs gained quality adjusted life years as a benefit measure. One additional life year is weighted by its “quality” expressed in so called utilities drawing on the work of von Neumann and Morgenstern who developed an axiom based decision theory under uncertainty.¹ In clinical praxis there are different ways to elicit health related utilities. Most commonly, answers from standardized questionnaires such the EQ 5D from the EuroQol Group Association, or the short form 36 (SF36) are transformed into utilities by using empirically based algorithms.²⁻⁴ Alternatively, utilities can be directly elicited by means of either the standard-gamble (SG) or the time-trade-off method. The SG method is recommended for measurement of individuals’ preferences under uncertainty and to express the outcome of different therapeutic choices in utility values to be used in clinical decision analysis and health program evaluation.⁵ Based on normative expected-utility arguments, health researchers consider the SG method as the gold standard for utility measurement.⁶ As this method presents a decision problem in the framework of a risky lottery, it gives rise to a potential gender bias.
as men and women differ in their willingness to accept risks, with women being more risk averse than men. Rosen et al for instance found significant predictors of increasing risk aversion in multivariate modeling, namely white race ($P<0.01$), lower education ($P<0.05$), and female gender ($P=0.07$). Based on those findings we expect that females would report lower quality of life values when those values are elicited in the framework of an SG experiment.\(^7\) To test this hypothesis we conducted an SG experiment with economics students of the University of Vienna, Austria. We framed the experiment using study related stress as the background scenario. Stress and burn out symptoms among students are worldwide on the rise leading to an impairment of students’ quality of life.\(^8\)–\(^11\) For instance, the Austrian Ministry for Science and Research commissioned a survey looking at the situation of Austrian students since the early 1970s. The latest report was released in 2012 by the Institute for Advanced Studies.\(^12\) One of the key findings was that currently 45% of the students suffer from mental health problems attributable to one of the following reasons: pressure to perform, fear of failure, fears of taking an exam, existential fear, feelings of depression, low self-esteem, pressure of competition, contact difficulties (social isolation). Moreover, the authors also discovered female students to be more prone to stress than their male counterparts, which was observed in many descriptive surveys.\(^13\) However, the cited studies have not employed a decision analytic framework to elicit stress related quality of life values. In the subsequent analysis we try to identify if gender influences the observed variance in the quality of life between students using an SG experiment setting.

Methods

Standard-gamble

We have operationalized the gamble as follows: the students were asked to choose between two alternatives A and B. Alternative A is to stay in the current state of health (with the current stress level) with certainty. Alternative B is a lottery with two outcomes. Outcome B1 is an absolutely stress free state of health that occurs with probability $p$ while outcome B2 represents an intense stress level (“burnout scenario”) that occurs with probability $1-p$. Outcome B2 is characterized by the following stress symptoms: fatigue, anorexia, insomnia, attentiveness disorder, self-reproaches, obliviousness, inability to make decisions, anxiety attacks, testiness, and listlessness.\(^14\) It was also mentioned that high stress levels have a negative impact on the cardiovascular system and subsequently on mortality;\(^15\) that stress is related to adverse lifestyle and diabetes;\(^16\) or leads to people adopting bad habits such as smoking or heavy drinking as a coping behavior.\(^17\)

The students were asked to state probability $p^*$ that indicates indifference between the two alternatives A and B. We then used $p^*$ to interpret the utility level of the current state of health. The gamble is illustrated in Figure 1. We made use of a so called probability wheel as a visual aid to the respondents.

Note that the worst outcome in this gamble is not death but a very severe degree of stress. To compare the utility values elicited from this gamble with other indications one would need to calibrate the results. However, for our purpose the absolute values are not important as we only want to understand influencing factors of the quality of life values.

Thirty-one undergraduate economics students of the University of Vienna participated in this SG experiment. They were recruited among attendees of the health economics class in summer term 2013 that was held by one of the authors. The gamble was conducted between May and August 2013. Informed consent was obtained from all participants.

Regression analysis

To identify factors that influence the observed quality of life we employed multivariate regression techniques to estimate the following equation:

$$QoL = f(X), \quad (1)$$

with QoL as quality of life index and X being a vector of explanatory variables. Assuming a linear relationship we can specify:

$$QoL = a + bX \quad (2)$$

In the base case scenario we used ordinary least squares (OLS). As the dependent variable QoL is an index limited
between 0 and 1, some statisticians have proposed estimation methods other than OLS for this particular case. Papke and Wooldridge suggested a generalized linear model to accommodate for proportional data. Their approach makes use of the logit transformation of the response variable and the binomial distribution. Yet, Ferrari and Cribari-Neto came up with the Betafit regression. Their approach assumes that the proportion follows a beta distribution. The beta distribution is bounded between 0 and 1 and models heteroscedasticity in such a way that the variance is largest when the average proportion is near 0.5. The main advantage of beta regression over OLS regression is that it is more robust in estimating covariate effects.

We estimated Equation 2 with all three methods using Stata 12 (StataCorp LP, College Station, TX, USA). We included the following explanatory variables that were believed to influence the stress level:

- Gender: in the Austrian student survey the researchers found female students to be more prone to stress than their male fellow students, which confirmed older results gathered from physicians or results from the Global Ageing Survey. Also recent large scale cardiovascular study also found that women had a significantly higher level of mental stress than men. The question is if higher stress levels are also reflected in lower utilities when elicited in an SG setting.
- Living with parents: as trouble with parents is a major cause of students’ stress level, we expected students who still live at home with their parents to suffer from more stress than those who live on their own.
- Part-time job to finance studies: more and more students are working during their studies, which might increase their stress level. On the other hand, a part-time job affects income which in turn positively influences quality of life at least up to a certain level, although some researchers even suggest an inverse relationship between income and quality of life.
- Studying more than one subject: some students study not just one subject but sometimes two or more, as they believe that this will give them an edge in the labor market. It goes without saying that we expect a higher stress level to be associated with this strategy.
- Number of study semesters: we do not have a particular assumption in mind that relates to the length of the studies to the quality of life, but one might think of two conflicting effects: freshman face more stress because of the transitional nature of a new stage. On the other hand, the longer a student is at university, the more pressure they might be under from parents, not to mention the exam related stress they face which intensifies at the later stage of a student’s life.
- Expected number of semesters to graduation: this information is much related to the previous question. We expect that stress becomes stronger the closer a student is to their graduation.

Results

In Table 1 we show the headline statistics from our sample. From the 31 respondents (15 female and 16 male undergraduate economics students), all respondents had finished at least one semester at the university. The highest number of semesters studied by a participant was 12. About 35% of the students study a second degree. Fifty percent of the students work while they are studying. Eighty-seven percent of the respondents live on their own. In the last row of the table we present the quality of life values as assessed by the SG method. The value has a mean of 0.57 with a standard deviation of 0.25.

Subgroup analysis

We report some subgroup results for the quality of life values in Table 2: female students have a higher QoL value than their male classmates; people who live with their parents have a lower mean in QoL-value than people who live away from their parents. Students who do not work have a lower quality of life than their peers who do work.

Regression analysis

We conducted the regression analysis to identify drivers of quality of life. The results are displayed in Table 3.

<table>
<thead>
<tr>
<th>Table 1 Student characteristics and quality of life values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Item</strong></td>
</tr>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Male (%)</td>
</tr>
<tr>
<td>Female (%)</td>
</tr>
<tr>
<td>Live with parents (%)</td>
</tr>
<tr>
<td>Part-time job to finance studies (%)</td>
</tr>
<tr>
<td>Study of more than one subject (%)</td>
</tr>
<tr>
<td>Semester (mean, [standard deviation])</td>
</tr>
<tr>
<td>Expected semesters to graduation (mean, [standard deviation])</td>
</tr>
<tr>
<td>Quality of life (mean, [standard deviation])</td>
</tr>
</tbody>
</table>
is also significant in all specifications. The small sample size drives the low $R^2$ value of 15% in model 1.

### Discussion

In this work we introduce the SG method with an examination of students at the University of Vienna. We consider our most striking finding to be that quality of life values differ significantly between male and female students. This is in contrast to other studies conducted using different methods to assess quality of life or happiness, who did not find significant gender differences or even reported higher stress and lower quality of life values. In our study, however, female students seem to value given health states more highly than men, which can probably be attributed to differences in their respective utility functions. Experimental economists confirm these findings and found that male gender, parental education, and height positively correlate with the willingness to take risks. Shorter individuals, older persons, and women on the other hand are less willing to take risks. Other researchers have come to the same conclusion with regard to female gender being a significant predictor of risk aversion. One explanation that was put forward is the potential overconfidence of men about their decisions, as men believe they can predict possible outcomes more accurately than women. In more familiar health economic terms, one could hypothesize that women are more subject to scale compatibility than men in that they focus more on the outcome that involves the bad-outcome probability in SGs.

This finding gives rise to a potential gender bias in the assessment of quality of life. Risberg et al define a gender bias in medicine to be when one assumes sameness and/or equity between women and men when there are genuine differences to consider. Then, because in the SG method participants’ risk attitudes drive the utility assessment, a gender bias has a big impact in many ways. The risk attitude is directly related to the curvature of the von Neumann–Morgenstern utility functions: risk neutral individuals have linear utility functions, while risk seeking individuals have convex utility functions and risk adverse individuals have concave utility functions. If women are really more risk averse than men, their utility functions would be more concave and show diminishing marginal utilities, this has the following far reaching implications:

1. As women have a preference for the status quo, they systematically value given state of health more than men.
2. For interventions that lead to a restoration of the perfect health state (utility level 1), men always benefit more than women regardless of their initial state of health before treatment.
3. For interventions that lead to a marginal improvement of a patient’s state of health, the marginal benefit is larger for women in states of health that are close to death (utility level 0). On the other hand, from a certain threshold

### Table 2 Subgroup results for quality of life values

<table>
<thead>
<tr>
<th>Item</th>
<th>QoL value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.51</td>
</tr>
<tr>
<td>Female</td>
<td>0.64</td>
</tr>
<tr>
<td>Live with parents</td>
<td>0.54</td>
</tr>
<tr>
<td>Live without parents</td>
<td>0.58</td>
</tr>
<tr>
<td>Part-time job to finance studies</td>
<td>0.61</td>
</tr>
<tr>
<td>No part-time job</td>
<td>0.54</td>
</tr>
</tbody>
</table>

Abbreviation: QoL, quality of life.

### Table 3 Determinants of quality of life

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model (1)</th>
<th>Model (2)</th>
<th>Model (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>GLM</td>
<td>Betafit</td>
</tr>
<tr>
<td></td>
<td>$P &gt; Z$</td>
<td>$P &gt; Z$</td>
<td>$P &gt; Z$</td>
</tr>
<tr>
<td></td>
<td>Coefficient (standard error)</td>
<td>Coefficient (standard error)</td>
<td>Coefficient (standard error)</td>
</tr>
<tr>
<td>Gender (female =1)</td>
<td>0.155 (1.82)$^*$</td>
<td>0.656 (2.03)$^{**}$</td>
<td>0.630 (2.00)$^{**}$</td>
</tr>
<tr>
<td>Live with parents =1</td>
<td>−0.119 (−0.89)</td>
<td>−0.517 (−1.04)</td>
<td>−0.388 (−0.93)</td>
</tr>
<tr>
<td>Part-time job =1</td>
<td>0.066 (0.60)</td>
<td>0.287 (0.68)</td>
<td>0.348 (0.78)</td>
</tr>
<tr>
<td>More than one subject =1</td>
<td>0.076 (0.68)</td>
<td>0.335 (0.78)</td>
<td>0.278 (0.61)</td>
</tr>
<tr>
<td>Semester</td>
<td>−0.063 (−1.45)</td>
<td>−0.270 (−1.62)</td>
<td>−0.279 (−1.79)</td>
</tr>
<tr>
<td>Expected semesters to graduation</td>
<td>0.068 (1.16)</td>
<td>0.291 (1.33)</td>
<td>0.320 (1.55)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.327 (1.32)</td>
<td>−0.755 (−0.81)</td>
<td>−1.038 (−1.13)</td>
</tr>
<tr>
<td>Prob $&gt; F = 0.35$</td>
<td>0.201</td>
<td>0.418</td>
<td>0.258</td>
</tr>
<tr>
<td>$R^2 = 0.15$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Significant at $^*10\%$, $^{**}5\%$.

Abbreviations: OLS, ordinary least squares; GLM, generalized linear model; Prob, probability.
health level men always benefit more from an incremental improvement.

The different shapes of the utility functions would also impact allocation decisions in health care. A social planner with a given budget, who wants to maximize the utility of the society, would discriminate between men and women. Depending on the nature of the disease and the intervention, in most cases he would allocate more of his budget to the treatment of men, this is because women enjoy already higher quality of life values and investing in women’s health care would result only in marginal utility gains. Our results therefore cast some doubt on the use of “unisex” tariffs in health economic evaluations of health states, when men and women have different quality of life interpretations of certain health dimensions. On the other side, some economists suggest that gender differences in risk aversion are not significant when it comes to decisions on pension plans and an Israeli study found female students even less risk averse than their male classmates. Further research should certainly explore this issue in more detail in the health care context.

Another confounder beside gender that drives risk aversion is cognitive ability. Although this was not part of our study, previous literature found individuals with higher cognitive ability are significantly more willing to take risks. If this argument is thought through, smarter people would benefit more from most medical interventions as well. Eventually, smart men would receive the highest share of the health care budget if allocation decisions were solely based on utility maximization. A recent study from Thailand indeed showed that male respondents with a high level of education have a higher willingness to pay for an increase of their quality of life.

Risk aversion can also explain utility variation across countries. The Japanese utilities for instance are smaller than those of other countries for mild health states but, for worse states, above all other models. Compared to other countries, the Japanese society is one of the most uncertainty avoiding countries on earth as Dutch sociologist Geert Hofstede puts it. The observed utilities for Japanese therefore perfectly correspond with the high risk aversion of the Japanese society. Hofstede explains the high risk aversion of Japanese by the threat of natural disasters such as earthquakes or tsunamis. Japanese people had to learn to prepare themselves for any uncertain situation.

When it comes to the other variables we analyzed, we found numerical evidence that a part-time job does not hamper the quality of life. The opposite seems to be true; the better income asserts a positive impact on the quality of life. With regard to the length of study, we found some numerical evidence that the quality of life decreases with the number of years studied and the closer students get to graduation. In line with our expectations, we discovered that living with parents does decrease the quality of life, while studying more than one subject increases the quality of life. The latter result is somewhat contrary to expectations as we have expected a higher stress level of dual degree students. The major limitation of our study is clearly the small sample size. Due to our small sample size, most observations do have a rather explanatory character and future research should validate our findings in a confirmatory study design using a broader sample. On the other hand, a sample size of 31 is not as small as one is inclined to believe; Schwappach and Boluarte found in their survey of 18 German studies that elicited utilities for health economic evaluations that the majority of the papers worked with sample sizes in the range of 20–40 respondents. Another issue is the generalizability of the results since our sample is very homogeneous and probably not representative for the general population.

Conclusion
Based on a small experiment we conclude that gender specific differences in risk attitudes lead to a systematic gender bias in eliciting utilities. Health care planners should be aware of this bias when making allocation decisions.

Disclosure
JM is employed at Janssen KK, LAO has no conflict of interest.

References
5. Gafni A. The standard gamble method: what is being measured and how it is interpreted. Health Serv Res. 1994;29(2):207–224.


