Central European Journal of Economic Modelling and Econometrics

Impact of Complementarity and Heterogeneity on Health Related Utility of Life

Michał Jakubczyk*

Submitted: 27.02.2009, Accepted: 15.05.2009

Abstract

This study aims at identifying determinants of health related quality of life in Poland, and in particular at verifying whether health domains are complements or substitutes and what the impact of heterogeneity of population on the health state valuation is. The paper uses data in panel structure coming from a survey conducted in Poland and consisting of 6700 valuations (after data cleaning) of EQ-5D health states with time trade-off method. Several econometric models are built in order to detect the impact of complementarity and heterogeneity. Random effects models as well as random parameters models estimated using Bayesian approach are used. The results show that health domains are complementary goods. Especially the lack of pain/discomfort is a complement to other health domains. Demographic factors influence how health state change impacts utility. These factors encompass sex, education, respondent's health state and even belief in life after death.

Keywords: health related quality of life, QALY, EQ-5D, complementarity, heterogeneity

JEL Classification: C11, C23, D61, I10.

*Institute of Econometrics, Warsaw School of Economics; e-mail: michal.jakubczyk@gmail.com

1 Introduction

The aim of this study is to detect the determinants of health related quality of life (HRQoL) in Poland, and in particular to quantify the impact of interactions within a health state description and between health state description and demographic factors on the utility of a health state. Economically speaking, the aim is to verify whether health domains should be treated as complements or substitutes and what the nature of heterogeneity of society with respect to health state valuation is.

The development of medicine and resulting availability of more expensive and efficient technologies as well as aging of society and growing consumer awareness raising the demand for health services combined with unavoidable budget constraint make a public regulator face a decision problem of choosing which health technologies should be provided (e.g. in Poland the reimbursement list or the basket of guaranteed services). This choice needs to be based on the clinical and economical criteria, so as to satisfy needs and meet the budget constraint at the same time, Gold *et al.* (1996), Drummond *et al.* (2005). The process of supporting this choice by providing necessary evidence coming from credible sources is called *health technology assessment* (HTA). Its role has been increasing in Poland and resulted in founding by the Ministry of Health the Agency of Health Technology Assessment.

HTA process should support the choice between alternative therapies for a single given illness, as well as the resource allocation between treatment of various illnesses. In the latter case competing technologies can differ in the measure of clinical effectiveness, e.g. some technologies can result in life prolongation, while others in the improvement of the quality of life with no impact on life expectancy. Therefore the necessity of making such comparisons requires a measure of clinical effectiveness that combines the longevity and quality of life. One of such measures that has been proposed in the literature and used in applications is QALY - quality-adjusted life years. The axiomatic rationale was given by Pliskin *et al.* (1980), this approach was then developed by Bleichrodt *et al.* (1997).

Making comparisons between health technologies based on societal preferences requires that the impact of health state change on QALY is quantified. This in turn demands assigning utility levels to all the health states. Due to possible differences between populations in different countries it is often argued that these studies should be performed separately for each country - to reflect the actual preferences. Such studies have been performed for various countries across the world, e.g. United Kingdom, Dolan *et al.* (1996); Japan, Tsuchiya *et al.* (2002); Germany, Greiner *et al.* (2004); or Netherlands, Lamers *et al.* (2006). This task for Polish population has been first performed by Golicki *et al.* (2009). They present the assignment of utility values to all the health states defined by EQ-5D form Brazier *et al.* (2007). In the modelling process used by Golicki *et al.* (2009) the stress was put on the impact of various domains of health states on utility (as the main purpose was to provide the decision maker with the operational results). The analysis of possible interactions within health state description was limited (no state specific variables were used).

Demographic factors were not included in their analysis.
The present paper aims at extending the analysis of Golicki *et al.* (2009) and verifying whether the factors omitted therein can influence the results. First the impact of interaction between specific health domains is quantified. From the economic perspective it is therefore analyzed whether the specific health domains are complements or substitutes. Secondly, the demographic factors are included - both as direct determinants of HRQoL and as terms that interact with health state descriptors. As it is often claimed in the literature that the assignment of utility to a health state should be done by the whole general population and not only by the affected by a given illness in order to provide an objective allocation Brazier *et al.* (2007), one of the aims of the present study is to detect whether the health state of the valuator impacts the assigned utility levels. The present study is based on the same data as Golicki *et al.* (2009), whose help is greatly acknowledged.
In the next section the methodology of data collection is briefly described, in particular the EQ-5D form and the methods of QALY measurement. In the third section the econometric approach is presented, i.e. the variables used in the modelling and models specification. The fourth section presents the results, and the last section summarizes.

2 Data

The data used in the present study came from the pioneer survey conducted in Poland in 2008 and described in Golicki *et al.* (2009). The survey included the valuation of several health states as well as demographic characteristics of the respondent. In the next subsection the EQ-5D form is described. In 2.2 the formal definition of QALY is presented along with methods of its measurement. In 2.3 the survey and data cleaning process are described.

2.1 EQ-5D form

Before employing quantitative methods the term *health state* needs to be defined operationally. The standard approach is to characterize this state by a set of parameters describing quantitatively the functioning of a person in selected aspects. This approach relates to the functional definition of health according the which "*...health is the state of optimal feasibility to performed valued tasks...*", Parsons (1964). In other words health relates to the possibility of fulfilling by the person roles and tasks determined by the socialization process.
The form that has been gaining popularity in recent years is EQ-5D developed by the interdisciplinary group of researchers - EuroQol Group, Rabin and de Charro (2001). In this form the health is described with five domains: mobility, self-care, usual activities, pain/discomfort, anxiety/depression, see Table 1. The first three domains relate to functioning of the person - abilities or participation skills, the last two to the

impairment of organism functioning. The health state in each domain is characterized on one of three levels, denoted 1, 2, 3, with 1 denoting no problems in a given domain, 2 - moderate problems, and 3 - severe problems. The more detailed descriptions are presented in Table 1. Each health state is then defined by five elements of a set $\{1, 2, 3\}$, e.g. the combination 11111 (in abbreviated notation) denotes the state of a full health. Due to its simplicity, i.e. small number of domains and levels resulting in only 243 health states (additionally "death" and "unconsciousness" are included, but they are not used in this study), EQ-5D can be widely used in population-based surveys and has become a popular tool in HTA, Brazier *et al.* (2007).

2.2 QALY and its measurement

Resource allocation between health technologies asks for comparing health states, and more precisely - comparing finite sequences of health states describing the life of a given duration with varying health. It is assumed that the decision maker has preferences in the space of such sequences. *Per analogiam* to the work of Neumann and Morgenstern (1944), it can be asked whether these preferences can be represented by the maximization of the expected utility function defined on health state sequences taking into account the longevity and quality of life.

Pliskin *et al.* (1980) gave the axioms on the preference structure that guarantee the existence of such a representation using quasi-additive utility function. Then the sequence of n health states: q_1, q_2, $\ldots$, q_n, each lasting for one year, is assigned the total utility $u(q_1) + u(q_2) + \ldots + u(q_n)$, where $u(\cdot)$ is the function representing the utility of a year in a given health state, normalized so that the utility of a full-health state is equal to 1, and death - 0. This total utility is called *quality-adjusted life years*, QALY. In particular, t years in health state q is represented by $tu(q)$ QALYs. Less restrictive axiomatization allowing for the above representation was given by Bleichrodt *et al.* (1997). For discussion on the plausibility of this approach see e.g. Culyer and Newhouse (2000).

The representation of preferences and normalization allow to determine the utility of any health state by comparing it to some combination of full-health and death. Two popular methods are: the standard gamble method and time trade-off (see Dolan *et al.* (2003), Brazier *et al.* (2007)), the latter was used by Golicki *et al.* (2009) and so in the present study. In this method the respondent is asked to compare the health state q lasting for 10 years to the full-health lasting for t years. The respondent is asked (*implicite* by a formalized procedure of successive questions) to determine such t (between 0 and 10 years) that she is indifferent between the two alternatives. By the definition of QALY the following equality holds $10u(q) = t$, and so $u(q) = \frac{t}{10} \in [0, 1]$. Some health states may be perceived by the respondent to be worse than death, then the above equality does not hold for any $t \in [0, 10]$. In such a case the respondent is asked to determine such t, that she is indifferent between an immediate death and t years in state q followed by $(10 - t)$ years in full health. Then $0 = tu(q) + (10 - t)$, and so $u(q) = -\frac{10-t}{t}$.

In the literature the utility of health states worse than death is further normalized, so as it is greater or equal to -1. The normalization used in the present study is of the form: $u^*(q) = \frac{u(q)}{1-u(q)} = -\frac{10-t}{10}$ for $u(q) < 0$, where $u^*(q)$ is the normalized value. Therefore for each health state q the assigned QALY is equal to $\frac{t}{10} \in [0,1]$ for health states better than death and $-\frac{10-t}{10} \in [-1,0]$ for health states worse than death. Figure 1 illustrates the time trade-off approach for health states better and worse than death.

2.3 Data collection

The data were collected with a survey study among 321 persons. The data were collected in eight medical centers in Poland, among visitors. The survey was designed so as to be representative for Poland with respect to age and gender.
A total of 44 health states was used in the study. The selection of these was based on the results of Lamers *et al.* (2006) providing lowest estimation errors. Most respondents assigned utility to 23 health states. Some respondents valuated fewer states due to logistic reasons. Altogether the initial dataset consisted of 7351 observations. The data have a panel structure with multiple observations (various health states) relating to one respondent. Additionally each respondent has been asked for demographic features: sex, age, education, accommodation, belief in life after death; her own health state (according to EQ-5D description).
As in all empirical studies there are data-quality issues present. These problems can arise on various steps of data collection process. First of all some respondents can have preferences violating the QALY approach axioms. The extreme example is the inconsistency with Pareto-dominance, i.e. the same respondent can assign a lower utility to a health state that is objectively better (i.e. better according to some domains, and not worse according to any). Secondly the respondent might have got tired during the survey (which lasted about 2 hours), therefore her later response may not be credible (i.e. due to misunderstanding of a health state description or time trade-off procedure). Finally there might have occurred typos during the survey or data entry into the database.
Each of the above type of errors would require a different theoretical approach and assumptions. In order to simplify the data-cleaning process a two step procedure was assumed. The aim of the first step was to erase all the non credible data resulting e.g. from typos or question misunderstanding. To that purpose for each health state separately all the observations which differed from the mean utility by more than 3 standard deviations were excluded. After this step the dataset contained 7282 observations. The aim of the second step was to remove respondents whose valuations are not credible due to internal inconsistencies. All the respondents who had more than 10 strong inconsistencies, i.e. pairs of healths states violating the Pareto-dominance by more than 0.2 (i.e. Pareto-dominated state was given utility greater by more than 0.2 than the Pareto-dominant state), were eliminated. There were 26 such respon-

dents. The order of the above steps is purposeful. One mistyped valuation could lead to many internal inconsistencies and to the removal of respondent whose valuations are otherwise credible. The final dataset consisted of 6701 observations for 295 respondents. Table 2 presents the characteristics of the data used with respect to the utilities assigned to various health states before and after data cleaning.

3 Econometric approach

In this section the econometric approach used in the analysis is presented - in subsection 3.1 the explanatory variables, and in 3.2 the specification of models built and estimation procedure.

3.1 Explanatory variables

Data used in the modelling came from 295 respondent and in total consisted of 6701 observations, i.e. health state valuations. Most of the respondents valued 23 health states. For each observation a utility value u^* is defined (see subsection 2.2). As an explained variable, the loss of utility to the full health is used, i.e. $1 - u^* \in [0, 2]$.
Basic explanatory variables relate to the valued health state. For each of five domains two binary variables are defined $d_{i,2}$, $d_{i,3}$, $i \in \{1, 2, 3, 4, 5\}$, denoting whether a given health state in i-th domain is valued on level 2 or 3 respectively. All models contain a constant, therefore variables $d_{i,1}$ are excluded. Because health state 11111 is not valuated, the constant can be interpreted as a utility loss due to non-perfect health.
Additionally variables allowing for interactions between domains in utility determination were used. Based on the literature (Tsuchiya *et al.* (2002), Greiner *et al.* (2004), Lamers *et al.* (2006)) the following variables were used: $N2$ ($N3$) - binary variables denoting whether any domain is valued at level 2 (level 3); $I2$ ($I3$, $D1$) - number of domains valuated at level 2 (level 3; not level 1) minus one; $I2^2$; $I3^2$; $D1^2$. These variables allow for detecting complementarity (substitution) between health domains, i.e. whether improving health in a given domain has greater (lower) positive impact on utility when the health state in the other domain is better. Moreover the pairwise product variables of $d_{i,3}$ were included in order to detect complementarity between specific domains, denoting $dd_{i,j} = d_{i,3} \times d_{j,3}$.
In this study available demographic variables were used: male - binary variable denoting sex; age - measured in years; Warsaw - indicator of respondents coming from Warsaw; country - indicator of respondents from countryside; education - indicator of higher education respondents; faith - indicator of strong belief in life after death. In order to detect the impact of demography on the relation between health state and utility, the product variables between demographic binary variables and $d_{i,2}$, $d_{i,3}$ were included.
The last group of explanatory variables aim at detecting the impact of respondent self health state valuation (also accordingly to EQ-5D form) on the results. Variables

$sd_{i,k}$ were included, where $sd_{i,k} = 1$ for observations in which respondent considered own health state in domain i ($i = 1, \ldots, 5$) on level 2 or 3, and the valuated health state had in domain i value k ($k = 2, 3$). These variables allow e.g. to detect whether respondents who themselves fell pain/discomfort think that it reduces the utility less/more than other respondents. It thus relates to the issue whose preferences should be used in utility assignment.

3.2 Estimation procedure

Four models of increasing complexity were built. In all models the final set of explanatory variables consisted of independent variables statistically significant at $\alpha = 0.05$. The hypotheses of the simultaneous significance of all omitted variables were tested where possible (not always due to near co-linearity problems).
Model M1 used only variables $d_{i,j}$. The aim of this model is to detect the simple impact of health domains on utility. As the aim of this study is to detect interactions between health domains or between demographics and health domains, this model is treated as a reference point for models M2–M4. The results of Golicki *et al.* (2009) were not cited here as in the present paper slightly different data-cleaning procedure was used.
Model M2 additionally included variables allowing for interaction between domains. It aims at detecting complementarity of health domains. Model M3 additionally used demographic variables, allowing for heterogeneity among respondents in utility determinants. Models M1–M3 were built as random effects models. Model M4 took full account of the heterogeneity assuming random parameters (Maddala (2001)). In model M4 only explanatory variables relating to health states were used as the demographic factors were included in the randomness of the parameters.
Models M1–M3 were estimated using GRETL 1.6.5. Model M4 was estimated with WinBugs 1.4 in Bayesian approach with non-informative priors using Markov Chain Monte Carlo (see e.g. Casella (2004), Geyer (1992)).

4 Results

In the following subsections the results of the estimation is presented, starting with the model with health domains only (M1), then including interaction terms (M2) and demographic variables (M3), and ending with random parameters model (M4).

4.1 Model M1 - impact of health domains

Model M1 consisted only of variables describing health states directly, without any interactions between domains or demographics. The results of the estimation are presented in the Table 4.
The R^2 equals to 40.05%. Breusch-Pagan test indicates the presence of individual

effects (null hypothesis rejected at p<0.001).
The model is intuitive in the sense that parameters for all domains are positive (worsening of health state reduces utility) and the parameters are higher for each domain for level 3 (severe problems with domain reduce utility more than moderate problems).
Pain/discomfort (domain 4) has the biggest impact on utility, and especially the extreme pain. The second most influential factor is complete lack of mobility (domain 1 on level 3). The less negative impact on utility is related to anxiety/depression (domain 5). Setting this domain on level 2 does not change the utility significantly.
It is worth noticing that the constant in this model is positive and statistically significant. It means that the first worsening of any health domain from perfect health results in greater utility loss than this assigned to this specific domain. The subsequent worsening results solely in an utility loss related to specific domains. Economically speaking health domains are complements, i.e. for worse level of one domain the improvement of the other domain results in smaller utility gain. This is a rationale to include more explanatory variables capturing interactions between domains and move from model M1 to model M2.

4.2 Model M2 - domains complementarity

Model M2 additionally includes interaction variables between domains allowing for more thorough complementarity detection than M1. We decided to use interaction variables between specific domains in order to account for possible variation in complementarity. In this respect this work differs from Golicki *et al.* (2009), who used only general interaction terms: $I2$, $I2^2$, $I3$, $I3^2$, $D1$, $D1^2$, $N3$. Table 5 presents results of the modelling.
The R^2 slightly increased to 40.47%. Breusch-Pagan test indicates the presence of individual effects (null hypothesis rejected at p<0.001).
All omitted variables were individually not significant. Due to near co-linearity the simultaneous significance of all omitted variables was not tested. Instead three subgroups of variables were tested: $\{d_{1,2}, d_{3,2}, d_{4,2}, d_{5,2}, N2, N3, I2, I2^2, I3, I3^2, D1\}$, $\{dd_{1,2}, dd_{1,3}, dd_{1,5}, dd_{2,3}, dd_{2,5}\}$, and $\{dd_{1,2}, dd_{1,3}, dd_{1,5}, dd_{2,3}, dd_{3,5}\}$. In all cases the sets of variables were non-significant with p respectively equal to 0.72, 0.99 i 0.95. Therefore the list of variables in Table 5 was considered to be final.
Model M2 additionally proves the complementarities between health domains, especially between pain/discomfort (domain 4) and other domains, especially strong with bed confinement (domain 1). As the parameters by the product variables are negative, the improvement of one domain (of domains 1, 2, 3, 5) results in greater improvement if the person does not feel an extreme pain or discomfort. The results prove that health domains are complements and that this complementarity can be directly attached to specific domains (namely pain/discomfort).

4.3 Model M3 - impact of demographic factors

Model M3 additionally includes demographic variables allowing for heterogeneity of population. Table 6 presents results of the modelling.
The R^2 further increased to 46.19%. Breusch-Pagan test indicates the presence of individual effects (null hypothesis rejected at p<0.001). Null hypothesis of Hasuman test that the GLS estimator is consistent was not rejected, though on the border of statistical significance, p*=0.052.
All omitted variables were individually not significant. Due to near co-linearity the simultaneous significance of all omitted variables was not tested. Instead three subgroups of variables were tested: $\{male, edu, Warsaw, country, faith\}$, $\{d_{5,2}, N2, N3, I2, I2^2, I3, I3^2, D1, D1^2\}$, $\{dd_{1,2}, dd_{1,3}, dd_{1,5}, dd_{2,3}, dd_{2,4}, dd_{2,5}\}$, $\{dd_{1,2}, dd_{1,3}, dd_{1,5}, dd_{2,3}, dd_{2,4}, dd_{3,5}\}$, $\{male_{1,2}, male_{1,3}, male_{2,2}, male_{3,2}, male_{4,2}, male_{5,2}, male_{5,3}\}$, $\{warsaw_{1,2}, warsaw_{1,3}, warsaw_{2,2}, warsaw_{2,3}, warsaw_{3,2}, warsaw_{4,2}, warsaw_{5,2}, warsaw_{5,3}\}$, $\{country_{1,2}, country_{1,3}, country_{2,2}, country_{2,3}, country_{3,2}\}$, $\{country_{4,2}, country_{4,3}, country_{5,2}, country_{5,3}\}$, $\{edu_{1,2}, edu_{1,3}, edu_{2,2}, edu_{3,2}, edu_{3,3}, edu_{4,2}, edu_{5,2}, edu_{5,3}\}$, $\{faith_{2,2}, faith_{2,3}, faith_{3,2}, faith_{4,2}, faith_{5,2}, faith_{5,3}\}$, $\{sd_{2,2}, sd_{2,3}, sd_{3,2}, sd_{3,3}, sd_{4,2}, sd_{4,3}, sd_{5,3}\}$. In all cases the sets of variables were non-significant with p respectively equal to 0.962, 0.619, 0.938, 0.942, 0.53, 0.621, 0,684, 0.264, 0.342, 0.272, 0.829. Therefore the list of variables in Table 6 was considered to be final.
The signs of parameters from M2 did not change, neither the individual domains nor the interaction variables. New results relate to the impact of demographic factors. The general conclusions are as follows:

- demography does not influence indirectly the utilities, the exception is the age - the older assign higher utility values; instead demography influences the impact of domains on utility;
- demography influences mostly the impact of level 3 of health domains on utility - demography matters for severe health impairment;
- in terms of number of interactions the most influential demographic factor is the belief in life after death - it influences domains 1, 3 and 4 (mobility, usual activities, pain/discomfort);
- the fact of living in Warsaw influences the results, which is a limitation for the present study, as the inhabitants of Warsaw are over-represented in the sample (Table 3).

The specific conclusions are:

- men attach less utility loss to health worsening in domains 3 and 4 (usual activities and pain/discomfort), more in domain 2 (self-care);

- respondents with higher education attach more importance to domains 2 and 4 (self-care and pain/discomfort);
- faith in life after death reduces the negative impact of health worsening in domains 1, 3 and 4 (mobility, usual activities, pain/discomfort);
- respondents who themselves feel impairment in domain 1 (mobility) attach less importance to this domain in terms of utility; on the contrary for domain 5 (anxiety/depression).

Model M3 is not operational. As it hinges the utility change on demography (and even belief in life after death), it would require in practice to use large number of explanatory variables, and more importantly would differentiate the optimality of health technology depending on non-medical factors. It would be thus not politically correct and probably unethical. It is thus necessary to build another model, which would on one hand allow for heterogeneity of population and on the other provide universal parameters for the whole population. Therefore in the next subsection the random parameters model M4 is presented.

4.4 Model M4 - random parameters model

Due to the reasoning presented in the end of the above subsection, a random parameters model was estimated. This model assumes that regression parameters are randomly and normally distributed in population around a common mean (Maddala (2001)), i.e.

$$Y_{i,j} = \Big(\alpha_0 + \epsilon_{i,0}\Big) + \sum_{k=1}^{5}\sum_{l=2}^{5}\Big(\alpha_{k,l} + \epsilon_{i,k,k}\Big)d_{k,l}(j) + \ldots + \eta_{i,j}, \qquad (1)$$

where $Y_{i,j}$ is the valuation of j-th health state by i-th respondent; $d_{k,l}(j)$ is the indicator of k-th domain being on level l-th in j-th health state; α_0 and $\alpha_{k,l}$ are the population means of parameters and $\epsilon_{i,0}$ and $\epsilon_{i,k,l}$ are the individual random disturbances in these parameters; finally $\eta_{i,j}$ is the random term. It is assumed that $\epsilon_{i,0}$ and $\epsilon_{i,k,l}$ are independent normal random variables with zero mean and fixed variance for all respondents. Therefore this model at the same time allows for heterogeneity in the population, and for obtaining operational values. As this model can fit to the data very well estimating the random values of regression parameters, the number of explanatory variables was reduced. In M4 only the basic $d_{i,j}$ variables were used along with pairwise products $d_{i,3} \times d_{j,3}$. With this approach the impact of individual domains as well as specific complementarities can be detected.
The model was estimated in the Bayesian approach with non-informative priors on all the parameters: α_0, $\alpha_{k,l} \sim N(0, 1000)$, $\mathrm{Var}(\epsilon_{i,0})$, $\mathrm{Var}(\epsilon_{i,k,l}) \sim \Gamma$(mean=1, variance=1000). The estimation was conducted in Markov Chain Monte Carlo approach using WinBugs software. Using this approach no significance testing was done, instead

95% confidence intervals were calculated using the percentile method and variables for which these intervals excluded zero were interpreted to be statistically significant. In this approach additionally the heterogeneity of the population with respect to individual parameters can be calculated. This heterogeneity is measured by the (estimated) variance of the regression parameters, e.g. $\mathrm{Var}(\epsilon_{i,0})$, $\mathrm{Var}(\epsilon_{i,k,l})$. The standard deviations of these parameters along with all other results are presented in Table 7.
The results are similar to those obtained in M2. Model M4 shows slightly more negative impact of level 3 in all domains. All the domains are statistically significant (a minor change comparing to M2). The domain 4 (pain/discomfort) is a complement with domains 1, 2 and 5 (mobility, self-care, anxiety/depression), though not with domain 3 (usual activities), which is another minor change comparing to M2.
The heterogeneity of parameters in population (measured as a ratio of standard deviation to the mean) is similar for all parameters related to individual domains.
A bigger diversity is present with respect to the interaction parameters.

5 Concluding remarks

The study aimed at quantifying the impact of health state description and demographic factors on the health related quality of life in Poland. The basic findings of this study are as follows.
All the domains used in EQ-5D on all levels impact the utility of health. Only in one model one variable was statistically insignificant (Table 5).
Health domains are complementary goods but the degree of complementarity differs between domains. Especially the lack of pain/discomfort is a complement to other health domains. This finding motivates using more domain-specific variables than traditionally used (e.g. Tsuchiya *et al.* (2002), Golicki *et al.* (2009)).
Demographic factors influence the impact of health state on utility. These factors encompass sex, education, respondent's health state and even belief in life after death. In particular men attach less utility loss to health worsening in usual activities and pain/discomfort, more in self-care. Respondents with higher education attach more importance to self-care and pain/discomfort.
It was found that faith in life after death reduces the negative impact of health worsening in domains mobility, usual activities and pain/discomfort. This may result from the fact that believers reject (perhaps subconsciously) the idea of time trade-off experiment and reluctantly reduce the longevity of life (perceiving it not morally proper) or claim the health state is worse than death. It would yield higher utility assignments to all the health states. It may be an argument for restricting the TTO-based surveys to subpopulations agreeing with the assumptions of the experiment.
Other finding is that respondents who themselves feel impairment in mobility attach less importance to this domain in terms of utility; on the contrary for anxiety/depression - this can be interpreted that a person can overcome mobility impairment when it happens, which is not the case for anxiety/depression. Therefore

conducting the survey in general populations (as opposed to actually being in given health states) can change the results in different directions.

The present study presents some limitations. First of all the data were collected in a non-representative manner. The visitors in hospital can be biased in some direction. Moreover it was detected that a large percentage of respondents coming from Warsaw can impact the results (Table 7). The last issue is the data quality. The cleaning was performed, but it was not based on any theoretical assumptions, and the data remaining still contain some values that may be result of a mistake - very high maximal utilities assigned to bad health states (Table 2). No further cleaning was performed in order not to reduce the dataset.

This paper does not aim at providing value set of utilities to be used in applied research as it would be obviously politically incorrect (and operationally difficult) to vary the provided medical therapy according to such features as address or faith in life after death. The results of this paper can instead be a starting point for further theoretical research on health state utility valuation, e.g. the impact of assumptions of TTO experiment on results for specific subpopulations. Moreover, next research could concentrate on the construction of models that can fit the data better - as the present models indicate that there are strong interdependencies between variables, data mining approach could be useful. More thorough data cleaning process could also be implemented.

References

[1] Bleichrodt, H., M. Wakker, and M. Johannesson (1997), Characterizing QALYs by Risk Neutrality, *Journal of Risk and Uncertainty*, 15, 107–114.

[2] Brazier, J., J. Ratcliffe, J. Alomon, and A. Tsuchiya (2007), *Measuring and Valuing Health Benefits for Economic Evaluation*. Oxford University Press.

[3] Casella G., C. P. R. (2004), *Monte Carlo Statistical Methods*. Springer-Verlag, Nowy Jork.

[4] Culyer, A., and J. Newhouse (eds.) (2000), *Handbook of Health Economics*. Elsevier Science.

[5] Dolan, P., C. Gudex, P. Kind, and A. Williams (1996), The time trade-off method: results from a general population study, *Health Economics*, 5, 141–154.

[6] Dolan, P.and Stalmeier, P. (2003), The validity of time trade-off values in calculating QALYs: constant proportional time trade-off versus the proportional heuristic, *Journal of Health Economics*, 22, 445–458.

[7] Drummond, M., M. Sculpher, G. Torrance, B. O'Brien, and G. Stoddart (2005), *Methods for the economic evaluation of health care programmes. Third edition.* Oxford University Press, Oxford.

[8] Geyer, C. (1992), Practical Markov Chain Monte Carlo, *Statistical Science*, 7, 473–483.

[9] Gold, M., J. Siegel, L. Russell, and M. Weinstein (eds.) (1996), *Cost-Effectiveness in Health and Medicine*. Oxford University Press.

[10] Golicki, D., M. Jakubczyk, M. Niewada, W. Wrona, and J. Busschbach (2009), Valuation of EQ-5D Health States in Poland: First TTO-Based Social Value Set in Central and Eastern Europe, *Value in Health*, published online, DOI: 10.1111/j.1524-4733.2009.00596.x.

[11] Greiner, W., C. Claes, J. Busschbach, and J.-M. Graf von der Schulenburg (2004), Validating the EQ-5D with time trade off for the German population, *European Journal of Health Economics*, 6, 124–130.

[12] Lamers, L. M., J. McDonnell, P. Stalmeier, P. Krabbe, and J. Busschbach (2006), The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies, *Health Economics*, 15, 1121–1132.

[13] Maddala, G. S. (2001), *Introduction to Econometrics. Third Edition.* John Wiley & Sons, Ltd.

[14] von Neumann, J., and O. Morgenstern (1944), *Theory of Games and Economic Behavior*. Princeton University Press.

[15] Parsons, T. (1964), *Social Structure and Personality.* The Free Press, Nowy Jork.

[16] Pliskin, J., D. Shepard, and M. Weinstein (1980), Utility Functions for Life Years and Health Status, *Operations Research*, 28, 206–224.

[17] Rabin, R., and F. de Charro (2001), EQ-5D: a measure of health status from the EuroQol Group, *Annales Medicales*, 33, 337–343.

[18] Tsuchiya, A., S. Ikeda, N. Ikegami, S. Nishimura, I. Sakai, T. Fukuda, C. Hamashima, A. Hisashige, and M. Tamura (2002), Estimating an EQ-5D population value set: the case of Japan, *Health Economics*, 11, 341–353.

Appendix - Tables and Figures

Table 1: Health domains and their levels used in the EQ-5D form.

Mobility	
1	I have no problems in walking about
2	I have some problems in walking about
3	I am confined to bed
Self-care	
1	I have no problems with self-care
2	I have some problems washing or dressing myself
3	I am unable to wash or dress myself
Usual activities	
1	I have no problems with performing my usual activities (e.g. work, study, housework, family or leisure activities)
2	I have some problems with performing my usual activities
3	I am unable to perform my usual activities
Pain/discomfort	
1	I have no pain or discomfort
2	I have moderate pain or discomfort
3	I have extreme pain or discomfort
Anxiety/depression	
1	I am not anxious or depressed
2	I am moderately anxious or depressed
3	I am extremely anxious or depressed

Table 2: Health states utilities in the data set.

Health state	Original data					Cleaned data				
	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max
11112	171	0.896	0.212	-1	1	155	0.917	0.135	0.325	1
11113	171	0.656	0.425	-1	1	151	0.704	0.33	-0.35	1
11121	149	0.88	0.206	-0.525	1	135	0.909	0.144	0.3	1
11122	173	0.826	0.287	-0.875	1	155	0.861	0.184	-0.025	1
11131	149	0.286	0.619	-0.975	1	137	0.327	0.608	-0.975	1
11133	170	0.195	0.648	-1	1	157	0.244	0.621	-1	1
11211	170	0.9	0.168	0	1	152	0.923	0.115	0.45	1
11312	147	0.685	0.362	-0.875	1	132	0.733	0.259	-0.15	1
12111	148	0.901	0.168	0.05	1	135	0.916	0.136	0.425	1
12121	150	0.853	0.203	0	1	136	0.868	0.175	0.275	1
12211	149	0.849	0.178	0.275	1	137	0.861	0.167	0.325	1
12222	170	0.727	0.356	-1	1	152	0.771	0.253	-0.175	1
12223	149	0.527	0.462	-0.875	1	135	0.543	0.438	-0.85	1
13212	150	0.615	0.403	-0.875	1	136	0.677	0.304	-0.475	1
13311	150	0.49	0.513	-0.85	1	138	0.517	0.496	-0.85	1
13332	170	-0.071	0.655	-1	1	156	-0.008	0.643	-1	1
21111	170	0.915	0.14	0.025	1	154	0.928	0.105	0.525	1
21133	170	0.202	0.635	-1	1	157	0.244	0.616	-1	1
21222	149	0.76	0.259	-0.6	1	136	0.779	0.222	0.025	1
21232	170	0.287	0.631	-1	1	157	0.32	0.609	-1	1
21312	170	0.549	0.479	-0.9	1	155	0.598	0.417	-0.85	1
21323	149	0.417	0.554	-0.9	1	137	0.445	0.529	-0.9	1
22112	170	0.783	0.306	-1	1	153	0.805	0.233	-0.05	1
22121	149	0.803	0.262	-0.825	1	136	0.819	0.206	0.125	1
22122	170	0.754	0.311	-1	1	153	0.785	0.24	-0.05	1
22222	319	0.663	0.405	-1	1	285	0.73	0.272	-0.55	1
22233	150	0.058	0.62	-0.95	1	138	0.081	0.624	-0.95	1
22323	172	0.296	0.595	-1	1	157	0.321	0.576	-1	1
22331	171	0.071	0.657	-1	1	157	0.118	0.645	-1	1
23232	149	0.046	0.627	-1	1	137	0.057	0.638	-1	1
23313	149	0.129	0.616	-0.95	1	137	0.169	0.598	-0.925	1
23321	173	0.293	0.598	-1	1	158	0.326	0.586	-1	1
23333	169	-0.204	0.626	-1	1	155	-0.19	0.63	-1	1
32211	149	0.464	0.559	-0.9	1	136	0.483	0.543	-0.9	1
32223	149	0.187	0.587	-0.925	1	137	0.205	0.585	-0.925	1
32232	171	-0.05	0.65	-1	1	156	-0.002	0.64	-1	1
32313	171	0.024	0.653	-1	1	157	0.07	0.644	-1	1
32331	169	-0.11	0.627	-1	1	155	-0.069	0.627	-1	1
32333	172	-0.295	0.597	-1	1	158	-0.291	0.601	-1	1
33212	149	0.278	0.6	-0.975	1	137	0.299	0.584	-0.975	1
33232	150	-0.183	0.6	-1	1	138	-0.169	0.61	-1	1
33321	148	0.033	0.648	-1	1	136	0.063	0.642	-1	1
33323	150	-0.15	0.606	-1	1	138	-0.163	0.608	-1	1
33333	318	-0.362	0.542	-1	1	292	-0.382	0.554	-1	1

Table 3: Variables characteristics (for clarity interaction variables are omitted here).

Variable	Mean	SD	Min	Max
Y	0.602	0.625	0	2
$d_{1,2}$	0.397	0.489	0	1
$d_{1,3}$	0.26	0.439	0	1
$d_{2,2}$	0.414	0.493	0	1
$d_{2,3}$	0.277	0.448	0	1
$d_{3,2}$	0.339	0.473	0	1
$d_{3,3}$	0.376	0.484	0	1
$d_{4,2}$	0.341	0.474	0	1
$d_{4,3}$	0.336	0.472	0	1
$d_{5,2}$	0.372	0.483	0	1
$d_{5,3}$	0.329	0.47	0	1
$N2$	0.869	0.337	0	1
$N3$	0.676	0.468	0	1
$I2$	0.994	1.141	0	4
$I2^2$	2.29	3.788	0	16
$I3$	0.902	1.13	0	4
$I3^2$	2.091	3.821	0	16
$D1$	2.441	1.406	0	4
$D1^2$	7.937	6.036	9	16
male	0.479	0.5	0	1
age	42.862	15.769	18	86
Warsaw	0.628	0.483	0	1
country	0.147	0.354	0	1
edu	0.432	0.495	0	1
faith	0.299	0.458	0	1

Table 4: Model M1 - results. All of the domains on all levels impact utility, except for $d_{5,2}$.

Variable	Estimate	Error	t statistics	p
const	0.0775840	0.0206046	3.7654	0.0002
$d_{1,2}$	0.0592913	0.0114886	5.1609	0.0000
$d_{1,3}$	0.320809	0.0144161	22.2535	0.0000
$d_{2,2}$	0.0428250	0.0122047	3.5089	0.0005
$d_{2,3}$	0.225800	0.0150717	14.9817	0.0000
$d_{3,2}$	0.0701070	0.0130860	5.3574	0.0000
$d_{3,3}$	0.216513	0.0142578	15.1855	0.0000
$d_{4,2}$	0.0536540	0.0117188	4.5785	0.0000
$d_{4,3}$	0.452945	0.0121231	37.3622	0.0000
$d_{5,3}$	0.197406	0.0106865	18.4725	0.0000

Table 5: Model M2 - results. Health domains are complementary goods - parameters associated with $dd_{i,j}$ variables are negative.

Variable	Estimate	Error	t statistics	p
const	0.103285	0.0193555	5.3362	0.0000
$d_{1,3}$	0.310046	0.0161354	19.2153	0.0000
$d_{2,2}$	0.0358967	0.0145038	2.4750	0.0133
$d_{2,3}$	0.220521	0.0190446	11.5792	0.0000
$d_{3,3}$	0.203441	0.0151047	13.4687	0.0000
$d_{4,3}$	0.516226	0.0162177	31.8309	0.0000
$d_{5,3}$	0.193589	0.0153760	12.5903	0.0000
$D1^2$	0.00979189	0.00120009	8.1593	0.0000
$dd_{1,4}$	-0.106377	0.0230555	-4.6139	0.0000
$dd_{2,4}$	-0.0493196	0.0242180	-2.0365	0.0417
$dd_{3,4}$	-0.0771489	0.0248271	-3.1074	0.0019
$dd_{4,5}$	-0.0805201	0.0219205	-3.6733	0.0002

Table 6: Model M3 - results. Demography influences the results.

Variable	Estimate	Error	t statistics	p
const	0.272920	0.0472430	5.7769	0.0000
$d_{1,2}$	0.0661780	0.0139302	4.7507	0.0000
$d_{1,3}$	0.408030	0.0191921	21.2603	0.0000
$d_{2,2}$	0.0677800	0.0127161	5.3302	0.0000
$d_{2,3}$	0.178652	0.0211380	8.4517	0.0000
$d_{3,2}$	0.0477160	0.0132342	3.6055	0.0003
$d_{3,3}$	0.357602	0.0266375	13.4247	0.0000
$d_{4,2}$	0.0531800	0.0115045	4.6225	0.0000
$d_{4,3}$	0.643903	0.0255593	25.1925	0.0000
$d_{5,3}$	0.228516	0.0144903	15.7703	0.0000
age	-0.00514541	0.00100812	-5.1039	0.0000
$male_{2,3}$	0.0764778	0.0236088	3.2394	0.0012
$male_{3,3}$	-0.0607473	0.0208544	-2.9129	0.0036
$male_{4,3}$	-0.144343	0.0197561	-7.3063	0.0000
$warsaw_{3,3}$	-0.0662245	0.0223816	-2.9589	0.0031
$warsaw_{4,3}$	-0.0646595	0.0197796	-3.2690	0.0011
$sd_{1,2}$	-0.0626838	0.0266677	-2.3506	0.0188
$sd_{1,3}$	-0.0597870	0.0295454	-2.0236	0.0431
$sd_{5,2}$	0.0344571	0.0164706	2.0920	0.0365
$country_{3,3}$	-0.125372	0.0304401	-4.1186	0.0000
$edu_{2,3}$	0.0588264	0.0211162	2.7858	0.0054
$edu_{4,3}$	0.115116	0.0198438	5.8011	0.0000
$faith_{1,2}$	-0.0461623	0.0224909	-2.0525	0.0402
$faith_{1,3}$	-0.158516	0.0275549	-5.7527	0.0000
$faith_{3,3}$	-0.0966855	0.0215023	-4.4965	0.0000
$faith_{4,3}$	-0.102571	0.0218141	-4.7020	0.0000
$dd_{1,4}$	-0.0894151	0.0223832	-3.9947	0.0001
$dd_{3,4}$	-0.0896234	0.0226222	-3.9617	0.0001
$dd_{4,5}$	-0.0846357	0.0216478	-3.9097	0.0001

Table 7: Model M4 - results. Variables for which 95% confidence intervals do not contain zero are bolded.

Variable	**Mean in population**	**Error of mean estimation**	**SD of parameter in population**
const	0.03935	0.01278	0.16
$\boldsymbol{d_{1,2}}$	0.03993	0.008381	0.03711
$\boldsymbol{d_{1,3}}$	0.3653	0.02752	0.3135
$\boldsymbol{d_{2,2}}$	0.05665	0.009229	0.03173
$\boldsymbol{d_{2,3}}$	0.2246	0.0224	0.1596
$\boldsymbol{d_{3,2}}$	0.04464	0.009758	0.03325
$\boldsymbol{d_{3,3}}$	0.229	0.0196	0.2353
$\boldsymbol{d_{4,2}}$	0.04993	0.008822	0.03784
$\boldsymbol{d_{4,3}}$	0.5666	0.02871	0.446
$\boldsymbol{d_{5,2}}$	0.04692	0.01035	0.04596
$\boldsymbol{d_{5,3}}$	0.2594	0.01792	0.1956
$dd_{1,2}$	-0.004763	0.02107	0.07348
$dd_{1,3}$	-0.02021	0.02694	0.2633
$\boldsymbol{dd_{1,4}}$	-0.08023	0.02567	0.272
$dd_{1,5}$	-0.005132	0.01884	0.1197
$dd_{2,3}$	0.05057	0.02612	0.07848
$\boldsymbol{dd_{2,4}}$	-0.09008	0.02237	0.1499
$dd_{2,5}$	0.02316	0.02504	0.1856
$dd_{3,4}$	-0.05391	0.02968	0.3024
$dd_{3,5}$	-0.01526	0.02046	0.1289
$\boldsymbol{dd_{4,5}}$	-0.1029	0.02139	0.2315

Figure 1: Time trade-off method for states better (left) and worse (right) than death.