USING BAYESIAN TECHNIQUES TO BUILD UP AN INCONSISTENCY FREE HEALTH STATUS INDEX

Juan M. Cabasés
Eduardo Sánchez

D.T.2004/05

USING BAYESIAN TECHNIQUES TO BUILD UP AN INCONSISTENCY FREE HEALTH STATUS INDEX

JUAN M. CABASÉS [a] AND EDUARDO SÁNCHEZ [a,*]

[a] *Department of Economics, Public University of Navarre, Navarre, Spain*

* Correspondence to: Departamento de Economía, Universidad Pública de Navarra. 31006 Pamplona, Spain. Tel. + 34 948 169 358; e-mail: eduardo.sanchez@unavarra.es

SUMMARY

Objectives

1- To obtain a set of values of health states of the EQ-5D based on self-related health VAS using linear and non-linear models Bayesian techniques. 2 - To analyse "logical consistency" in different models and to derive a model free from logical inconsistencies. 3 - To analyse and compare results of several models when using a priori sources of information.

Methods

We apply the usual models and transformations of these, in order to attain logical consistency of the value set. Models proposed are: linear model (1); linear with dummy variables (2) and two models with a logistic structure with different distributions of the coefficients to be estimated (3 and 4). For two of these models new dummies are added in order to obtain logical consistency (2B and 4B).

Results

We propose a modelling to guarantee consistency of values of the EQ-5D health states that may be applied to suitable samples at apparently low cost of fit. This model is non-linear, has distribution Gamma in the coefficients and specific dummy variables.
The introduction of priors may reduce the cost of forcing logical consistency.

KEY WORDS – Bayesian analysis; EQ-5D; Logical inconsistency

INTRODUCTION

Several models have been developed to obtain valuation sets for the health states of the EQ-5D. A common problem of these models is that of logical inconsistencies. Up to now, no model fully guarantees the consistency of the value set obtained for every possible sample.

This paper uses regression models to build up a value set of health states of the EQ-5D using Bayesian methods. We aim a model that guarantees consistency in the values of the health states. It is important to distinguish between "primary" inconsistencies, caused by the intrinsic limitations of respondents and "secondary" inconsistencies caused by methodological aspects in the measurement procedure [1]. This study analyses the secondary inconsistencies.

The EQ-5D is a generic instrument of health related quality of life. It is a two part instrument. Part 1 records self-reported problems on each of five domains: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each domain is divided into three levels of severity corresponding to no problem, some problem, and extreme problem. The combination of these levels defines a total of 243 health states. Part 2 records the subject's self-assessed health on a Visual Analogue Scale (VAS), it is a vertical 20 cm line on which the best and worst imaginable health states score 100 and 0 respectively [2]. It has been validated in Spanish by Badía X. *et al*. [3].

Bayesian methods are widely used in health care economic evaluations. The Bayesian perspective allows for a natural interpretation of the results in terms of probability, as well as the incorporation of *a priori* information in the analysis [4]. To estimate the Bayesian models Markov Chain Monte Carlo (MCMC) is used [5,6].

Aims

1.- To obtain a set of values of health states of the EQ-5D based on self-related health VAS from a random sample of the general population of Spanish region (Life

Conditions of Navarra Population 2001, N= 2477), using linear and non-linear models Bayesian techniques.

2. - To analyse "logical consistency" in different models and to derive a model free from logical inconsistencies.

3. - To analyse and compare results of several models when using a priori sources of information.

METHODS

Modelling

We explore different models to obtain an evaluation set for EQ-5D health states. We apply the usual models and transformations of these, in order to attain logical consistency of the value set. Models proposed are: linear model (1); linear with dummy variables (2); and two models with a logistic structure with different distributions of the coefficients to be estimated (3 and 4). For models 2 and 4 new dummies are added in order to obtain logical consistency (2B and 4B).

Linear Model with N3 (Model 1)

Badía *et al*. [3] produced a social value set for Spain. They justified the model used using three criteria: goodness of fit, parsimony, and the logical consistency of the values obtained. The last two criteria are the ones used by Devlin *et al*. [7] to select this model among the nine possibilities proposed by them. The model is expressed through the following equation:

$$VAS_{score\,i} = ß_0 + \mathbf{MO}_i ß_1 + \mathbf{SC}_i ß_2 + \mathbf{UA}_i ß_3 + \mathbf{PD}_i ß_4 + \mathbf{AD}_i ß_5 + \mathbf{N3}_i ß_6 + e_i$$

Model 1

$$VAS_{score}, ß_j \sim N(0, O_\mu), \quad j=0,\ldots,5 \qquad e_i \sim N(0, s^2_e) \quad s^2_e \sim Ga(a_e, b_e)$$

Variables $\mathbf{MO}_i$, $\mathbf{SC}_i$, $\mathbf{UA}_i$, $\mathbf{PD}_i$, $\mathbf{AD}_i$ represent the five dimensions of the EQ-5D. In addition, a dummy variable (N3) takes the value "1" if one of the dimensions is at level 3, and "0" otherwise. VAS is the value of the Visual Analogue Scale.

Linear Model with dummy variables N2 and N3 (Model 2)

Dummy variables are included to value the move between levels 1 and 2 as different from the move between levels 2 and 3. Two dummy variables are used for each dimension. In addition, a new variable N2 is included. This model was used by Greiner *et al*. [8]

$$VAS_{score\,i} = ß_0 + \mathbf{MO2}_i ß_1 + \mathbf{SC2}_i ß_2 + \mathbf{UA2}_i ß_3 + \mathbf{PD2}_i ß_4 + \mathbf{AD2}_i ß_5 + \mathbf{MO3}_i ß_6 + \mathbf{SC3}_i ß_7 + \mathbf{UA3}_i ß_8 + \mathbf{PD3}_i ß_9 + \mathbf{AD3}_i ß_{10} + \mathbf{N2}_i ß_{11} + \mathbf{N3}_i ß_{12} + e_i$$

Model 2

$$VAS_{score}, ß_j \sim N(0, O_\mu), \quad j=0,\ldots,12 \quad e_i \sim N(0, s^2_e) \quad s^2_e \sim Ga(a_e, b_e)$$

The variables used in this model take the following values:

MO2, SC2, UA2, PD2, AD2 = 1 if the score is 2; 0 otherwise.

MO3, SC3, UA3, PD3, AD3 = 1 if the score is 3; 0 otherwise.

N2=1 if the score is either 2 or 3 in one of the dimensions; 0 otherwise.

N3=1 if the score is 3 in one of the dimensions; 0 otherwise.

Non-Linear Model with dummy variables N2, N3 and normally distributed coefficients (Model 3)

Here we are suggesting a change in the structure of the model. The possibility of non-linearity is justified on the big variances generally observed in the parameter estimations. Moreover, if we take into account the differences between the estimated and observed VAS mean in the work by Greiner *et al*. [8], we can see that nearly all negative values are concentrated in a place where VAS has values higher than 50, and positive values where VAS has lower values. This could support the use of a non-linear structure in order to diminish that variance.

The structural form comes from logit models [9,10]. We are using only the functional form and not the estimation of the logit model because the interpretation of the results of the dependent variable in the logit model is done in terms of probability of success,

and this would not be adequate in this case. For example, an estimated of the dependent variable of 0,7 in a logit model would mean a probability of 70% of being in the upper bound of 100, whereas in our model this 0,7 would be the VAS_{score} expected value in the sample.

We transform VAS_{score} into a Beta distribution, dividing by 100 to bounding the value of VAS_{score} between zero and one.

The model is as follows:

$$VAS_{score\,i} = 1/ (1+ EXP - (ß_0 - \mathbf{MO2}_i ß_1 - \mathbf{SC2}_i ß_2 - \mathbf{UA2}_i ß_3 - \mathbf{PD2}_i ß_4 - \mathbf{AD2}_i ß_5 - \mathbf{MO3}_i ß_6 - \mathbf{SC3}_i ß_7 - \mathbf{UA3}_i ß_8 - \mathbf{PD3}_i ß_9 - \mathbf{AD3}_i ß_{10} - \mathbf{N2}_i ß_{11} - \mathbf{N3}_i ß_{12}))+ e_i$$

Model 3

$$VAS_{score} \sim Be\,(Y_i, 1-Y_i) \quad ß_j \sim N(0, O_\mu), \; j=0,\ldots,12 \quad e_i \sim N(0, s^2_e) \quad s^2_e \sim Ga\,(a_e, b_e)$$

Non-Linear Model with dummy variables N2, N3 and Gamma distributed coefficients (Model 4)

While maintaining the structural function of the previous model, a new one is developed where the dependent variable follows a Beta distribution. The coefficients – that in Bayesian Theory are random variables – follow a Gamma distribution (1,1). This means that are bounded to 0, therefore its value is always positive. Hence, the move from level 1 to 2 and from level 2 to 3 in any of the five dimensions of the EQ-5D will always have a diminishing effect upon the index value.

$$VAS_{score\,i} = 1/ (1+ EXP - (ß_0 - \mathbf{MO2}_i ß_1 - \mathbf{SC2}_i ß_2 - \mathbf{UA2}_i ß_3 - \mathbf{PD2}_i ß_4 - \mathbf{AD2}_i ß_5 - \mathbf{MO3}_i ß_6 - \mathbf{SC3}_i ß_7 - \mathbf{UA3}_i ß_8 - \mathbf{PD3}_i ß_9 - \mathbf{AD3}_i ß_{10} - \mathbf{N2}_i ß_{11} - \mathbf{N3}_i ß_{12}))+ e_i$$

Model 4

$$VAS_{score} \sim Be\,(Y_i, 1-Y_i) \quad ß_j \sim Ga(1, 1), \; j=0,\ldots,12 \quad e_i \sim N(0, s^2_e) \quad s^2_e \sim Ga\,(a_e, b_e)$$

Logical Consistency

One of the problems to face when calculating a value set for a small population is the non-completion of the *logical consistency* in the health states values. Logical consistency seems to be a desirable characteristic in the value data set modelling.

There are different definitions of inconsistency [1,11]. To our purpose, logical inconsistency exits when to a health state less severe a lower value is assigned.

We define logical consistency as:

Be two health states A and B described by 5 dimensions (a_1, a_2, a_3, a_4, a_5) and (b_1, b_2, b_3, b_4, b_5) where a_i, b_j are {1,2,3} for i,j =1,...,5:

If when i=j, dimensions either are in the same level or in different levels such that $b_j = a_i + 1$ or $b_j = a_i + 2$,

then the value set should give a higher value to health state A.

A necessary, although not sufficient, condition to satisfy this propriety is that the coefficients should have a diminishing effect on the index value, as happens in model 4. To guarantee sufficiency we introduce an additional restriction by changing dummy variables.

The new dummy variables are:

MO2, SC2, UA2, PD2, AD2 = 1 if the score is 2 or 3; else=0

MO3, SC3, UA3, PD3, AD3 = 1 if the score is 3; else=0

The new dummies guarantee that the coefficients associated with level 3 will always be higher in absolute terms than those associated with level 2, because level 3 adds the coefficient associated to variables **MO3, SC3, UA3, PD3, AD3,** to that of the variables **MO2, SC2, UA2, PD2, AD2,** respectively.

Both models 2 and 4 have been adapted to these new dummy variables. Results are presented as Model 2B that does not require negativity on the coefficients, and therefore does not comply with the propriety of consistency, and Model 4B that complies with it.

Rescaling

Using VAS as a dependent variable, regression models do not generate a bounded scale between 0 and 100. So the values of the value set are rescaled. To each value a rescaled value is assigned using the following equation:

$$X_{resc} = 100 \cdot (X - X_{33333}) / (X_{11111} - X_{33333})$$

X_{resc} Rescaled VAS value for the health state X. X is the value of a health state when estimated coefficients are applied.

Bayesian estimation

Bayesian estimation is done using simulation techniques MCMC [5]. Browne *et al*. [12] show how this method gives more efficient and robust estimations than the more classic Maximum Likelihood methods as Iterative Generalized Least Square (IGLS), or its restricted version (RIGSL), and they converge quicker. For the simulation, the statistical program WinBUGS 1.4 [13] has been used.

Informative prior

Except in the cases that will be specified later, models have been estimated with a relatively uninformative priors [14]. Where prior information has been included, we have used the estimated parameters of another random sample, the Navarra Health Survey 2000 (ENS2000, N=1495). When the parameters of Gamma distributions are needed (models 4 and 4B), then the transformation used is [15]: Ga(a,b), being x, y mean and standard deviation, respectively, $x = a/b$, and $y = a/b^2$.

Data

General Household Survey 2001 (ECV2001, N=2477) of Navarra (Northern developed Spanish region, 560.000 habitants) [16]. It is a random age and sex quota sampling, stratified by county and municipality of residence, of the adult population (over 15) of Navarra. The interviews took place during May and June 2001.

In addition, the Navarra Health Survey 2000 (ENS2000, N=1495) is used to see whether the models that best fit our initial sample (ECV2001) are the ones that best fit the other one.

RESULTS

Demographic characteristics of the sample

Table 1 shows the demographic characteristics of the sample studied (ECV2001) by age and sex, as well as the mean VAS.

Table 1 about here

In table 2 the number of observations included in the analysis (2389) and the reasons of the non-valid answers (88) is shown

Table 2 about here

Models

Table 3 shows the results of the models discussed, with coefficient estimates, means, standard deviations and 95% probability bayesian intervals. It can be seen that models 1, 2 have positive estimate coefficients while in model 3 are negative, which produce values higher than that of 11111 for some health states in the three models. Moreover, it can be seen that in models 1 and 2 the probability intervals are very high.

Table 3 about here

Our estimates in models 1, 2, and 3 do not guarantee logical consistency. There is, at least, an example of non-fulfilment of logical consistency in models 1 and 2 (21111 vs. 22111), and model 3 (21111 vs. 23111). This also happens in other models described in the literature whose aim is to estimate a tariff for the 243 health states but do not take into account the fulfilment of this propriety. As an example, in Dolan and Roberts's value set there is higher value for 33233 than for 33133 [17].

In table 4 models 2B and 4B are shown. Model 2B, despite including new dummy variables, does not fulfil with the rule of logical consistency due to the positive value of some of the estimated coefficients (21111 vs. 22111). On the other hand, it has to be noted that model 4B fulfils the rule due to its design.

Table 4 about here

Health States Values

The following equation has to be used to obtain the values for the health states for models 3, 4 and 4B

$$X = (1/ (1+ EXP - (Z))) \cdot 100$$

where Z is a linear function of the estimated coefficients and dummy variables for each state.

As an example, in model 4B the estimated VAS value for the state 21231 is:

$$Z = 1.\ 054 - 0.208 - 0.183 - (0.197 + 0.173) - 0.206 - 0.110 = -0.024$$

$$X_{21231} = (1/ (1+ EXP - (Z))) \cdot 100 = 49.39 \text{ (0.4939 in the 0,1 scale)}$$

An the rescaled value:

$$X_{resc\text{-}21231} = 100 \cdot (X - X_{33333})/ (X_{11111} - X_{33333}) = 55.70 \text{ (0.557 in the 0,1 scale)}$$

Once obtained the valuation set, the best adjusted model will be the one showing the lowest differences between the estimated and observed means for the 63 known health states. In table 5, the 63 health states observed in our sample are shown, with frequencies, VAS observed means, estimated values with model 4B, and the differences between estimated and observed values. Figure 1 shows the mean of the absolute differences between estimated and observed values for the 63 health states observed in the sample (ECV2001), through the estimated models without and with rescaling.

Tables 5 and Figure 1 about here

It can be seen that the sum of the absolute differences is bigger when health states are rescaled. If we take into account rescaled values, the logistic transformation seems to fit better than the linear one. When including restrictions, models increase absolute differences.

The models with the least differences are the linear one with new dummies (model 2B) without rescaling, and the logistic transformation where coefficients show a normal distribution for the rescaled VAS (model 3).

Figure 2 shows, in two graphs, the differences between estimated and observed values. The first one compares model 2 (linear with dummies) versus model 3 (logistic transformation with normally distributed coefficients). The second one shows model 2 versus model 4. In both cases, it can be seen how the linear regression with dummies (model 2) overestimates higher values and underestimates lower values. This is slightly amended when a different functional form is used to allow for more flexibility (models 3 and 4).

Figure 2 about here

In order to choose a model to estimate the best value set, the one which values are the more representative of the population analysed, another sample of the same population is used. In table 6 the observed 56 health sates (ESN2000), frequency, VAS observed mean, estimated values with model 4B, and the differences between estimated and observed values are shown. Figure 3 shows the mean of the absolute differences of the observed values in the second sample (ESN2000) trough the estimated models.

Tables 6 and Figure 3 about here

As it can be seen in figure 3, the linear model fits better over the non-rescaled values but do not guarantee consistency. Absolute differences with rescaled values seem to be lower when models do not have a linear structure. In spite of model 4B assuming the cost of introducing some restrictions to guarantee consistency, absolute differences are lower compared to model 1 and model 2B, with and without rescaling, and model 2 with rescaling.

Priors

Table 7 shows the results from models 2, 2B, 4 and 4B: coefficient estimates, mean, standard deviation and 95%, bayesian probability intervals, using ESN2000 coefficient estimates and standard deviation as priors. Figure 4 shows the means of the absolute

differences between estimated and observed VAS values, in models 2, 2B, 4 and 4B using ESN2000 estimates as prior information.

Tables 7 and Figure 4 about here

For models 2 and 2B, we introduce as priors the mean and standard deviation of the estimated coefficients in models 2 and 2B for ESN2000 sample. For models 4 and 4B, the priors used are the parameters of the Gamma distributions obtained through the mean and standard deviation of the estimated coefficients in models 4 and 4B for ESN2000 sample.

As it can be seen in figure 4, the restrictions imposed by models 4 and 4B have a cost in terms of wider means of the sum of absolute differences, as compared with models 2 and 2B. But the introduction of priors may reduce this cost, as shown in model 4B in figure 1 and figure 4.

DISCUSSION

We have developed several models to obtain a set of values of the 243 health states of the EQ-5D looking for consistency. As we have seen, linearity do not guarantee the fulfilment of consistency, even when using our proposed new dummies. This is shown in model 2B, where coefficient estimates can be negative and can lead to logical inconsistencies.

In this analysis, logit models cannot be used because our dependent variable has to be interpreted as a probability of success of the occurrence of an event. Nevertheless, as it has been shown in this study, its functional form can be applied to our dependent variable.

On the usability of variables N2 and N3, it is suggested that maybe are not needed, because dummy variables and structural changes of the model should reflect the true variations in the levels of different dimensions. Once models 2 and 3 are examined without those variables, no true gains are obtained on the differences between VAS estimated and observed values in the sample.

The use of another randomised sample of the same population allows us to check whether the models estimates are representative of the population. When comparing figure 1 versus figure 3, our inconsistency free model (model 4B) fits even better when its coefficients are applied to another sample of the same population. But we should be cautious, since we cannot tell that the model chosen is the one that fits the best for that population.

Nevertheless, to apply the local estimations is a better option than to import coefficients from other populations (see table 8).

As we have seen, the introduction of priors may reduce the cost of forcing logical consistency, as shown in model 4B in figure 4 as compared with the same model in figure 1. This may be due to the fact that more information from the same population gives more accurate fit by reducing the sample error.

One advantage of our proposed model 4B is that, even if we had first order inconsistencies, the model does not allow any inconsistencies in the set of values, and could use all the existing information, even that from inconsistent respondents.

This is an exploratory study. We have used a sample of self-assessed VAS. However, to get a value set of the EQ-5D health states useful for CUA, a sample of valuations obtained from general population from hypothetical health states, using accepted scaling techniques, should be used.

However, in this paper, we propose a modelling to guarantee consistency of values of the EQ-5D health states that may be applied to suitable samples at apparently low cost of fit. This model is non-linear, has distribution Gamma in the coefficients and specific dummy variables.

ACKNOWLEDGEMENTS

The authors acknowledge financial support from the EuroQol Group, and thank Emilio Domínguez and Idoia Gaminde for their comments and suggestions. An earlier version of the paper was presented at 21th Plenary Meeting of the EuroQol Group, Chicago, Illinois, USA, 16th-18th September 2004.

REFERENCES

1. Dolan P, Kind P. Inconsistency and health state valuations. *Social Sci Med* 1996; **42**(4): 609-615.

2. Brooks R. EuroQol: the current state of play. *Health Policy* 1996; **37**: 53-72.

3. Badía X, Roset M, Monserrat S, Herdman M, Segura A. La versión española del EuroQol: descripción y aplicaciones. *Medicina clínica* (Barc.) 1999; **112** (Supl 1): 79-86.

4. O'Hagan A, Stevens JW. Assessing and comparing costs: how robust are the bootstrap and methods based on asymptotic normality?. *Health Economics*, 2003; **12**: 33-49.

5. Gilks WR, Richardson S, Spiegelhalter DJ. *Markov Chain Monte Carlo in practice*. Chapman and Hall: London 1996.

6. Browne W. *Applying MCMC methods to multilevel models. Statistics*, Bath, University of Bath 1998.

7. Devlin NJ, Hansen P, Kind P, Williams A. Logical inconsistencies in survey respondents` health state valuation – a methodological challenge for estimating social tariffs. *Health Economics* 2003; **12**: 529-544.

8. Greiner W, Weijnen T, Nieuwenhuizen M, Oppe S, Badia X, Busschbach J, Buxton M et al. A single European currency for EQ-5D health states. Results from a six-country study. *Eur J Health Econom*, 2003; **4**:222-231.

9. Greene WH. *Econometric Analysis*, third edition. Macmillan: New York University 1998.

10. Novales A. *Econometría*. Segunda edición. McGraw-Hill: Madrid 2000.

11. Ohinmaa A, Sintonen H. Inconsistencies and modeling of the Finnish EuroQol (EQ-5D) preference values. In *EuroQol Plenary Meeting 1998 Discussion Papers*, Greiner W, J-M Graf vd, Schulenburg Piercy J (eds).

12. Browne W, Draper D, Goldstein H, Rasbash J. Bayesian and likelihood methods for fitting multilevel models with complex level-1 variation. *Computational Statistics and Data Analysis* 2002; **39**: 203-225.

13. Spiegelhalter S, Thomas A, Best N, Lunn D. WinBUGS user manual", Biostatistic Unit, Cambrige, UK 2003. http://www.mrc-bsu.cam.ac.uk/bugs

14. Gill J. *Bayesian methods for the social and behavioural sciences*. Chapman and Hall: Boca Raton, Florida 2002.

15. Novales A. *Estadística y Econometría*. McGraw-Hill: Madrid 1997.

16. Instituto de Estadística de Navarra. *Encuesta de condiciones de vida de la población Navarra en 2001*. Gobierno de Navarra: Navarra 2003.

17. Dolan P, Roberts J. Modelling Valuation for Eq-5d Health States An Alternative Model Using Differences in Valuations. *Medical Care* 2002; **40**(5): 442-446.

Table 1. Demographic characteristics and VAS means of the sample studied (ECV2001) (N = 2477)

Variables		%	VAS means
Age	15-34	34.7	80.8
	35-44	16.9	76.9
	45-64	26.0	74.4
	65 +	22.4	69.7
Sex	Male	47.1	77.0
	Female	52.9	75.4

Table 2. Data used for modelling (ECV2001)

Non valid answers	**88 (3.56%)**
No VAS data	81
No data on the 5 dimensions	2
No data on 2 dimensions	1
No data on 1 dimension	4
Valid answers	**2389 (96,44 %)**
TOTAL	**2477**

Table 3. Results from the model regression on the VAS values. Estimated coefficients, mean, standard deviation and 95% bayesian probability interval in models 1,2,3 and 4.

	Model 1		Model 2		Model 3		Model 4	
	mean (sd)	*I.B. 95%*	*mean (sd)*	*I.B. 95%*	*mean (sd)*	*I.B. 95%*	*mean (sd)*	*I.B. 95%*
Constant	80.49(23.6)	(33,126)	80.85(24.29)	(33,128)	1.055(0.03)	(0.98,1.12)	1.054(0.03)	(0.98,1.12)
MO2	-9.14(94.0)	(-195,177)	-9.35(94.15)	(-194,179)	0.282(0.14)	(-0.01,0.57)	0.226(0.12)	(0.02,0.47)
SC2	3.02(130.2)	(-249,257)	2.20(160.6)	(-309,311)	-0.191(0.25)	(-0.69,0.30)	0.127(0.10)	(0.00,0.40)
UA2	-5.84(106)	(-212,198)	-5.79(115.4)	(-228,219)	0.249(0.18)	(-0.11,0.61)	0.204(0.12)	(0.01,0.49)
PD2	-5.38(91.4)	(-186,173)	-6.05(89.72)	(-182,170)	0.204(0.14)	(-0.07,0.48)	0.212(0.10)	(0.02,0.43)
AD2	-8.30(52.8)	(-110,97)	-6.41(87.72)	(-174,164)	0.249(0.14)	(-0.02,0.52)	0.245(0.11)	(0.03,0.48)
MO3			-4.92(421.8)	(-836,810)	0.516(4.44)	(-1.77,2.47)	0.398(0.34)	(0.01,1.25)
SC3			-0.10(383)	(-752,738)	-0.072(2.05)	(-1.87,1.64)	0.325(0.28)	(0.00,1.06)
UA3			-21.26(370.5)	(-752,706)	0.397(0.81)	(-1.19,2.03)	0.332(0.27)	(0.01,1.00)
PD3			-10.3(259.5)	(-521,502)	0.342(0.53)	(-0.67,1.44)	0.243(0.17)	(0.00,0.64)
AD3			-11.92(266)	(-534,517)	0.508(0.54)	(-0.54,1.61)	0.322(0.20)	(0.01,0.76)
N2			-4.32(98.0)	(-198,191)	0.203(0.15)	(-0.10,0.50)	0.193(0.10)	(0.01,0.41)
N3	-2.98(108.9)	(-217,209)	0.32(271.3)	(-538,541)	-0.007(0.57)	(-1.16,1.08)	0.170(0.13)	(0.00,0.49)
Sigma	1.28(0.70)	(0.5,3.2)	1.33(0.8243)	(0.53,3.45)	0.183(0.01)	(0.16,0.20)	0.182(0.01)	(0.16,0.20)

Table 4. Results from the model regression on the VAS values. Estimated coefficients, mean, standard deviation and 95% bayesian probability interval in models 2B and 4B.

	Model 2B		Model 4B	
	mean (sd)	*I.B. 95%*	*mean (sd)*	*I.B. 95%*
Constant	80.85(24.27)	(33.7,129)	1.054(0.03)	(0.98,1.12)
MO2	-8.42(95.07)	(-195,178)	0.208(0.11)	(0.01,0.46)
SC2	2.31(159.9)	(-304,313)	0.108(0.09)	(0.00,0.35)
UA2	-9.90(114.8)	(-234,213)	0.183(0.12)	(0.00,0.45)
PD2	-5.8(89.61)	(-181,171)	0.196(0.10)	(0.01,0.41)
AD2	-6.68(87.97)	(-177,163)	0.224(0.11)	(0.02,0.45)
MO3	-1.78(411.2)	(-794,795)	0.355(0.32)	(0.01,1.19)
SC3	-1.34(393.1)	(-773,757)	0.290(0.26)	(0.00,0.97)
UA3	-1.63(371.8)	(-734,726)	0.279(0.24)	(0.00,0.91)
PD3	-3.73(259.6)	(-514,510)	0.173(0.13)	(0.00,0.50)
AD3	-6.40(270.8)	(-537,532)	0.218(0.16)	(0.00,0.59)
N2	-3.97(98.5)	(-197,184)	0.206(0.11)	(0.01,0.43)
N3	0.047(270.6)	(-530,533)	0.110(0.09)	(0.00,0.34)
Sigma	1.33(0.8244)	(0.53,3.45)	0.183(0.00)	(0.16,0.20)

Table 5.

The health states (63) observed in the sample (ECV2001), frequencies, VAS observed means, estimated values with model 4B, and the differences between estimated and observed values.

Health States	Frequency	Observed VAS [1]	Estimated VAS [2]	Difference [2] –[1]
11111	1699	80.87	74.15	-6.72
11112	116	69.97	65.10	-4.87
11113	23	66.52	57.31	-9.21
11121	196	70.60	65.74	-4.87
11122	62	62.54	60.52	-2.03
11123	10	52	52.46	0.46
11131	16	73.68	59.08	-14.61
11132	1	50	53.56	3.56
11133	1	20	48.12	28.12
11211	7	69.85	66.04	-3.82
11212	1	50	60.84	10.84
11221	12	63.33	61.51	-1.83
11222	7	65	56.07	-8.93
11223	1	50	47.88	-2.12
11233	2	52.5	43.58	-8.92
12111	2	90	67.70	-22.30
12112	1	70	62.61	-7.39
12211	1	70	63.58	-6.42
12213	1	10	50.09	40.09
12221	2	92.5	58.92	-33.58
13322	1	40	36.72	-3.28
21111	41	68.02	65.46	-2.57
21112	5	57	60.22	3.22
21121	45	61.08	60.89	-0.20
21122	10	46.5	55.43	8.93
21123	1	50	47.24	-2.76
21131	8	56.25	53.95	-2.30
21132	1	50	48.35	-1.65
21133	1	60	42.94	-17.06
21211	10	62	61.21	-0.79
21221	20	52	56.46	4.46
21222	19	60.26	50.88	-9.38
21231	1	50	49.39	-0.61
21232	5	41	43.80	2.80
21233	1	25	38.53	13.53
21311	1	80	51.65	-28.35
21312	1	75	46.04	-28.96
21321	1	20	46.74	26.74
21332	1	15	37.08	22.08
22111	4	81.25	62.98	-18.27
22121	4	72.5	58.29	-14.21
22123	1	70	44.56	-25.44
22133	1	60	40.32	-19.68
22211	3	56.66	58.62	1.95
22212	1	70	53.09	-16.91
22221	12	53.75	53.79	0.04
22222	7	47.14	48.18	1.04
22223	3	50	40.10	-9.90
22231	4	40	46.69	6.69
22232	2	50	41.16	-8.84
22233	1	40	36.00	-4.00
23221	1	50	43.81	-6.19
23233	1	30	29.62	-0.38
23311	1	80	41.77	-38.23
23321	1	70	37.08	-32.92
23333	1	30	24.14	-5.86
31122	1	75	43.84	-31.16
31311	1	47	42.82	-4.18
32221	1	55	42.21	-12.79
32321	1	40	35.58	-4.42
33212	1	50	34.70	-15.30
33311	1	40	33.46	-6.54
33332	1	50	21.72	-28.28
Sum of absolute differences	2389			713.57
Mean				11.33

Figure 1. Mean of absolute differences between estimated and observed values in the sample (ECV2001) through the estimated models without and with rescaling.

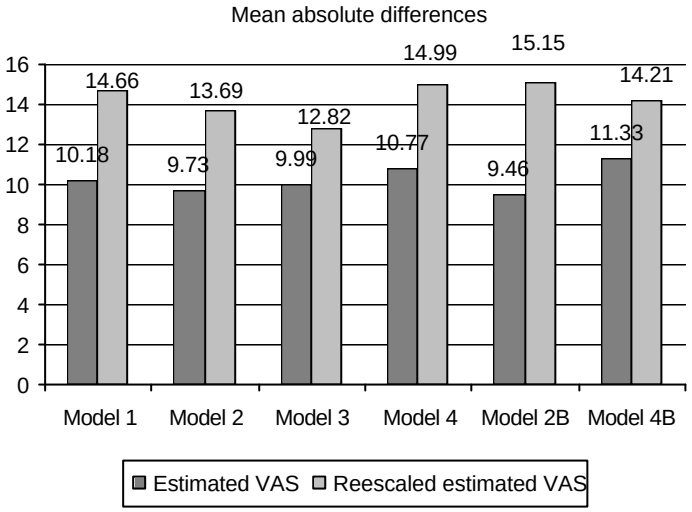
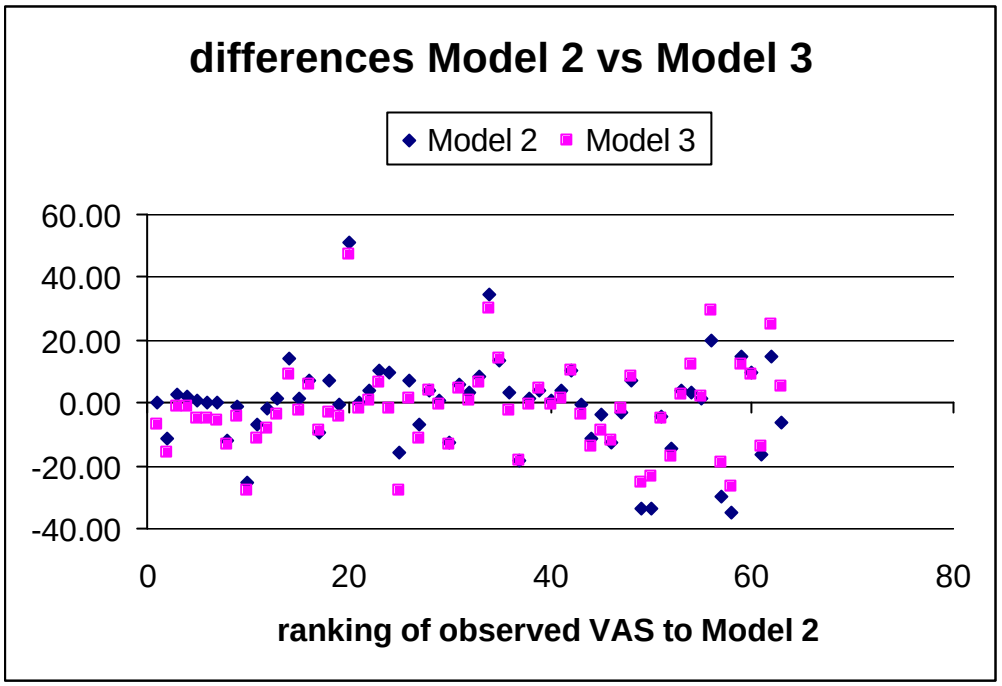


Figure 2

Differences between estimated and observed values for 63 health states ranked by observed mean. Comparison of model 2 (linear with dummies) versus model 3 (logistic transformation with normally distributed coefficients)



Differences between estimated and observed values for 63 health states ranked by observed mean. Comparison of model 2 versus model 4

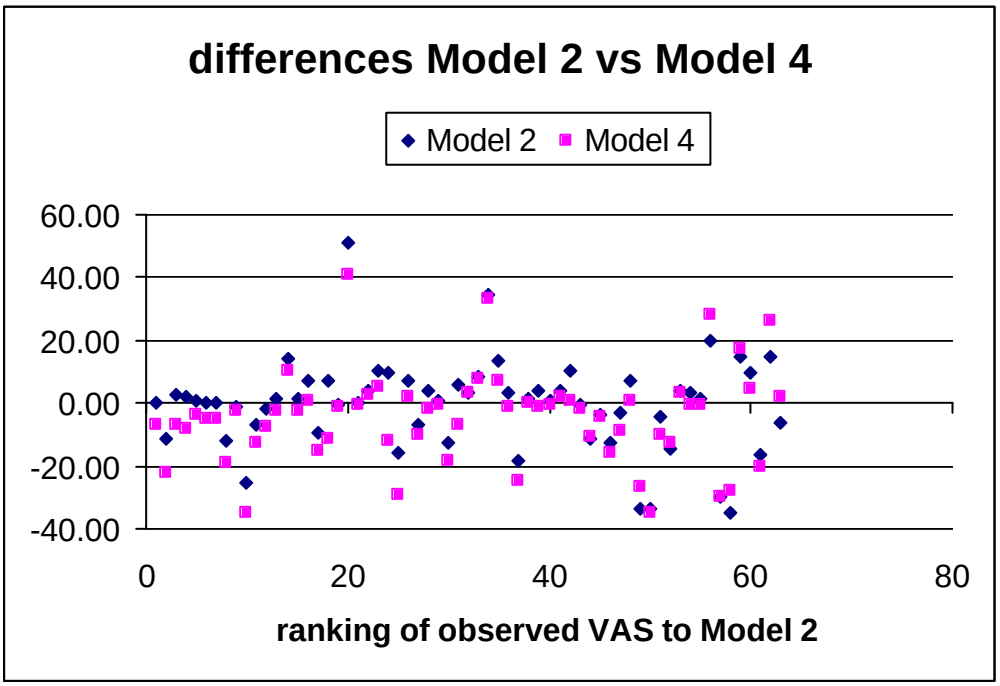


Table 6.

Health states (56) observed in the sample (ESN2000), frequencies, VAS observed means, estimated values with model 4B, and the differences between estimated and observed values.

Health States	Frequency	Observed VAS [1]	Estimated VAS [2]	Difference [2] –[1]
11111	1035	84.50	74.15	-10.34
11112	57	71.95	65.10	-6.85
11113	2	50	57.31	7.31
11121	141	71.30	65.74	-5.56
11122	22	64.09	60.52	-3.58
11123	1	70	52.46	-17.54
11131	14	60.57	59.08	-1.49
11132	4	72.50	53.56	-18.94
11211	10	69.50	66.04	-3.46
11212	2	45	60.84	15.84
11221	12	62.08	61.51	-0.58
11222	4	42.50	56.07	13.57
11231	3	53.33	54.59	1.26
11322	1	40	46.35	6.35
12111	1	50	67.70	17.70
12211	1	60	63.58	3.58
12212	1	78	58.23	-19.77
12213	1	50	50.09	0.09
12221	1	50	58.92	8.92
12222	2	65	53.39	-11.61
12231	1	30	51.90	21.90
13232	1	25	39.21	14.21
21111	23	71.52	65.46	-6.06
21112	9	58.89	60.22	1.33
21121	34	64.56	60.89	-3.67
21122	11	58.64	55.43	-3.20
21123	2	55	47.24	-7.76
21131	7	53.86	53.95	0.09
21132	1	50	48.35	-1.65
21133	1	50	42.94	-7.06
21211	9	65.56	61.21	-4.34
21221	23	53.48	56.46	2.98
21222	9	57.22	50.88	-6.34
21231	8	53.75	49.39	-4.36
21232	3	51.67	43.80	-7.86
21332	1	30	37.08	7.08
22121	1	75	58.29	-16.71
22211	6	56.67	58.62	1.95
22212	1	50	53.09	3.09
22221	6	48.33	53.79	5.45
22222	4	42.50	48.18	5.68
22223	1	60	40.10	-19.90
22231	1	60	46.69	-13.31
22232	3	43.33	41.16	-2.17
22233	1	40	36.00	-4.00
22323	1	25	33.60	8.60
22332	1	70	34.60	-35.40
22333	1	20	29.84	9.84
23222	1	60	38.38	-21.62
23321	2	40	37.08	-2.92
23322	1	40	32.01	-7.99
23331	1	10	33.13	23.13
31111	2	82.50	54.32	-28.18
31232	1	20	35.34	15.34
32311	1	50	40.20	-9.80
33322	1	40	24.82	-15.18
Sum of absolute differences	1495			524.51
Mean				9.37

Figure 3. The mean of the absolute differences between estimated and observed values in the sample (ESN2000, n=1495) through the estimated models without and with rescaling.

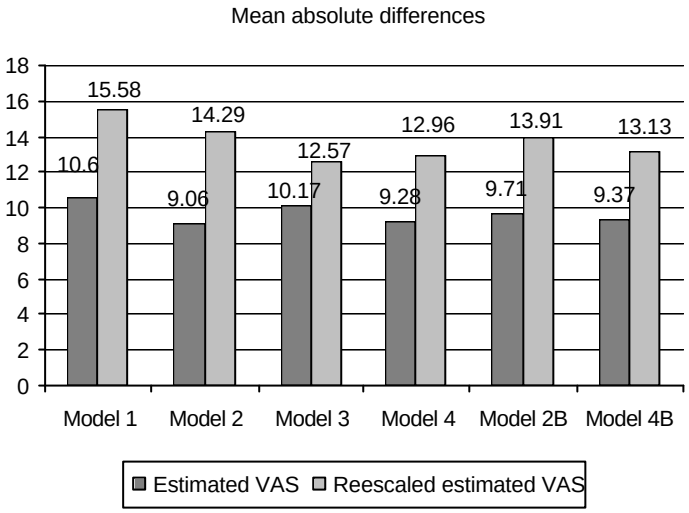


Table 7. Results from the model regression on the VAS values. Models 2, 2B, 4 and 4B. Coefficient estimates, mean, standard deviation and 95%, bayesian probability intervals, using ESN2000 coefficient estimates and standard deviation as priors.

	Model 2		**Model 4**		**Model 2B**		**Model 4B**	
	mean (sd)	*I.B. 95%*	*mean (sd)*	*I.B. 95%*	*mean (sd)*	*I.B. 95%*	*mean (sd)*	*I.B. 95%*
Constant	82.16(18.9)	(45.01,119)	1.067(0.03)	(0.99,1.13)	82.15(18.9)	(45,119)	1.068(0.03)	(0.99,1.14)
MO2	-6.49(67.0)	(-136,127)	0.141(0.12)	(0.00,0.42)	-7.16(67.3)	(-139,127)	0.134(0.12)	(0.00,0.42)
SC2	1.09(122.2)	(-236,237)	0.052(0.08)	(0.00,0.29)	0.83(121.6)	(-232,237)	0.041(0.06)	(0.00,0.25)
UA2	-7.76(82.0)	(-167,153)	0.188(0.14)	(0.00,0.51)	-9.43(81.7)	(-170,148)	0.170(0.14)	(0.00,0.49)
PD2	-4.42(64.3)	(-129,119)	0.112(0.09)	(0.00,0.33)	-4.13(64.4)	(-129,120)	0.104(0.09)	(0.00,0.33)
AD2	-5.04(64.1)	(-127,120)	0.109(0.11)	(0.00,0.37)	-4.57(63.9)	(-128,119)	0.103(0.10)	(0.00,0.36)
MO3	-4.11(318.6)	(-634,613)	0.252(0.28)	(0.00,1.04)	-2.71(312.9)	(-601,606)	0.213(0.27)	(0.00,0.95)
SC3	-5.17(296.4)	(-587,566)	0.221(0.25)	(0.00,0.914)	-4.68(302)	(-597,577)	0.173(0.22)	(0.00,0.79)
UA3	-22.39(261.8)	(-538,488)	0.211(0.24)	(0.00,0.89)	-3.64(262)	(-523,512)	0.137(0.19)	(0.00,0.69)
PD3	-11.61(188.7)	(-382,365)	0.141(0.16)	(0.00,0.56)	-6.78(187.8)	(-375,370)	0.069(0.11)	(0.00,0.40)
AD3	-13.24(194.7)	(-394,372)	0.232(0.19)	(0.00,0.69)	-9.32(197.4)	(-395,381)	0.173(0.17)	(0.00,0.60)
N2	-7.97(65.2)	(-135,122)	0.371(0.11)	(0.14,0.57)	-7.87(65.4)	(-135,118)	0.389(0.11)	(0.15,0.59)
N3	3.28(185.3)	(-360,364)	0.087(0.12)	(0.00,0.43)	3.15(184.9)	(-360,365)	0.041(0.08)	(0.00,0.29)
Sigma	1.33(0.82)	(0.533,3.45)	0.182(0.01)	(0.16,0.20)	1.33(0.824)	(0.53,3.45)	0.182(0.00)	(0.16,0.20)

Figure 4. Means of the absolute differences between estimated and observed VAS values, in models 2, 2B, 4 and 4B using ESN2000 estimates as prior information.

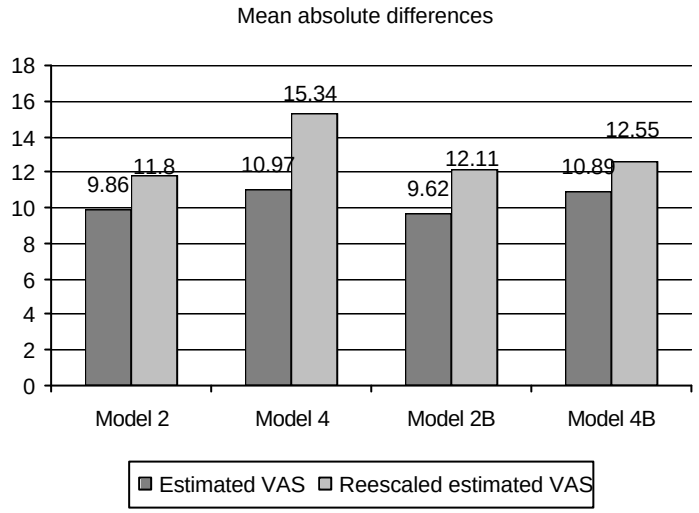


Table 8. Means of the absolute differences between estimated and observed VAS values obtained by Model 2 in the sample ECV2001 and Greiner et al (2003)

		ECV2001	**ECV2001** Reescaled	**ESN2000**	**ESN2000** Reescaled
Model 2 in the sample ECV2001	Sum of absolute difference	613.22	862.41	507.52	800.21
	Mean	9.73	13.69	9.06	14.29
Estimated by Greiner et al. (Model 2)	Sum of absolute difference	1018.04	1075.14	787.85	833.07
	Mean	16.16	17.07	14.07	14.88