

Satisfaction with Life and Its Domains among Older Individuals across European Countries¹

Eric Bonsang, ROA, Maastricht University, and Netspar

Arthur van Soest, Tilburg University, RAND, and Netspar

October 2011

Very preliminary and incomplete

Abstract

Using data on individuals of age 65 and older from 11 European countries, we analyze the subjective well-being of older Europeans, decomposing it into several domains of well-being: financial, daily activities, social contacts and family life, and health. The four domains are measured with self-reports on satisfaction with income, daily activities and social contacts, and a health index constructed on the basis of a large number of health variables. We use domain specific anchoring vignettes to correct for differences in response scales across countries in the three subjective self-satisfaction scales.

We develop a new model in which genuine life satisfaction is a weighted mean of all three genuine domain satisfactions and (objectively measured) health. Each domain satisfaction report is based upon an extended ordered probit equation with country and socio-economic group specific threshold levels distinguishing very satisfied from satisfied, satisfied from neither satisfied, nor dissatisfied, etc. Life satisfaction reports are based upon comparing the genuine satisfaction level with the weighted means of the domain specific thresholds. The results show, among other things, that controlling for differences in response scales has a substantial effect on the comparison of the distribution of life satisfaction among European countries.

Key words: happiness, subjective well-being, anchoring vignettes, response scale differences, ageing

JEL codes: I30, J14, Z13

¹ Data collection and infrastructure for making data available to researchers was mainly funded by the European Commission through several SHARE related projects in the 5th and 6th framework programmes (CIT5-CT-2005-028857, QLK6-CT-2001- 00360; RII-CT- 2006-062193). Additional funding was provided by the US National Institute on Aging (grant numbers U01 AG09740-13S2; P01 AG005842; P01 AG08291; P30 AG12815; Y1-AG-4553-01; OGHA 04-064; R21 AG025169) and various national sources (see <http://www.share-project.org> for a full list of funding institutions).

1. Introduction

Economists have discovered happiness (or rediscovered) or at least research on subjective well-being and its economic correlates (see, e.g., Van Praag et al. 2003, Layard, 2005, or Clark et al., 2008). In the context of the ageing of the European population, well-being of older individuals in particular has become a key policy issue in all European countries. Particularly in Southern Europe, poverty is more prevalent among the elderly than among other age groups (Tsakoglou, 1996), and the lack of economic resources makes elderly people vulnerable to poor quality of life (Grundy, 2006).

Using data on individuals of age 65 and older from 11 European countries from the Survey of Health, Ageing and Retirement in Europe (SHARE), we analyze the subjective well-being of older Europeans, decomposing it into several domains of well-being: financial, daily activities, social contacts and family life, and health. All these four domains have been shown to contribute substantially to overall well-being (satisfaction with life or happiness). See, for example, Ferrer-i-Carbonell and van Praag (2002), Van Praag et al. (2003), and Van Praag and Ferrer-i-Carbonel (2008).

The four domains are measured with subjective self-reports on satisfaction with income, daily activities and social contacts, and a health index constructed on the basis of a large number of health variables measured on an objective scale. The self-assessments all use the same subjective response scale: a five point scale ranging from very dissatisfied to very satisfied (very dissatisfied; dissatisfied; neither satisfied nor dissatisfied; satisfied; very satisfied). An important issue underlying the cross-country comparison on such a subjective scale is that individuals from different countries or socio-demographic backgrounds may use different response scales. This phenomenon is referred to as differential item functioning (DIF) in the psychology literature (Holland and Wainer, 1993). Only if individuals use the same scales, differences in self-reported satisfaction reflect “true” differences across countries or groups of individuals. Van Praag et al. (2003) use panel data models with quasi-fixed effects, capturing persistent differences in response scales. This allows them to identify how changes in satisfaction respond to changes in characteristics but does not help to identify cross-country differences in satisfaction levels that keep response scales constant. Specifically for the latter purpose, King et al. (2004) have proposed to use anchoring vignettes – respondents are asked to evaluate hypothetical situations described in the survey question.

This additional information helps to identify interpersonal differences in response scales, even with cross-section data.

In this study, we use domain specific anchoring vignettes to correct for differences in response scales across countries in the three subjective self-satisfaction scales. The vignettes describe how well-off hypothetical persons are in each of the three given domains. Existing models use the vignettes to correct for response scale differences in one given domain of life or directly use vignettes on life satisfaction without considering the domains. See, for example, Bonsang and van Soest (2011a,b) for studies of the former type on satisfaction with family income and social contacts, or Angelini et al. (2011) and Kapteyn et al. (2010) for studies of life satisfaction of the latter type. The novelty of the current paper is that it combines vignettes and self-assessments in several domains to analyze life satisfaction as a whole, allowing for response scale differences in each of the subjectively measured domains.

We develop a new model in which genuine life satisfaction is a weighted mean of all genuine domain satisfactions. Each of the three domain satisfaction self-reports is based upon an extended ordered probit equation with country and socio-economic group specific threshold levels distinguishing very satisfied from satisfied, satisfied from neither satisfied, nor dissatisfied, etc. Life satisfaction reports are based upon comparing the genuine satisfaction level with the weighted means of the domain specific thresholds and the health index (assumed not to suffer from response scale differences since it is based upon measurements on objective scales). The equations for domain satisfaction, thresholds, and satisfaction with life are estimated jointly with maximum likelihood. The estimates are used to simulate counterfactual distributions of domain and life satisfaction, replacing the reporting thresholds by those of a fixed benchmark country.

The results confirm the existing finding in the literature that reports of satisfaction with income, daily activities, and social contacts are affected by differences in reporting scales, and that correcting for such differences changes the ranking of average satisfaction in these domains across countries. Moreover, each of the domains considered has an important effect on life satisfaction, and the threshold variation by domain also implies threshold variation in satisfaction with life self-reports. Controlling for these differences in response scales has a substantial effect on the comparison of the distribution of life satisfaction among European countries.

This paper is organized as follows. Section 2 discusses the use of anchoring vignettes in the context of response scale differences. Section 3 introduces the model for life satisfaction and domain satisfactions. Section 4 presents the data and descriptive statistics.

Estimation results are discussed in Section 5. Section 6 presents some simulations of counterfactual distributions, showing how life satisfaction compares across countries when response scales are kept constant. Section 7 concludes.

2. Anchoring Vignettes

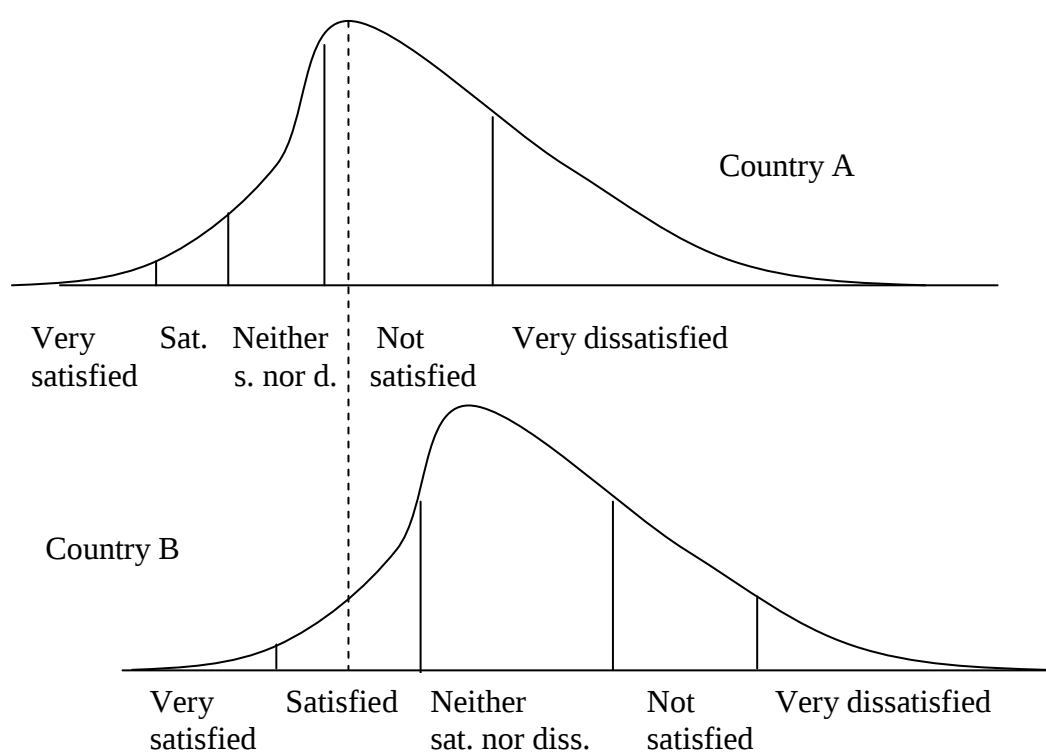
In this section we describe how vignettes can be used for identifying response scale differences, following, for example, King et al. (2004) and Kapteyn et al. (2010). Figure 1 (taken from Kapteyn et al., 2010) sketches distribution of a continuously distributed variable reflecting genuine satisfaction with a certain domain of life (income, social contacts, or daily activities, in our case) in two hypothetical countries. The density in country A is to the left of that in country B, so that people in country A are genuinely more satisfied than in country B. But the respondents in countries A and B use very different response scales. In the example in the figure, country B respondents attach more positive labels to given points on the satisfaction scale than country A respondents. Someone in country A with satisfaction indicated by the dashed line would report “not satisfied,” while a person in country B with the same genuine satisfaction would report “satisfied.” The frequency distributions of self-reports in the two countries would suggest that people in country B are *more* satisfied than those in country A—the opposite of what is implied by the genuine distributions. Correcting for the differences in the response scales (DIF, “differential item functioning,” in the terminology of King et al., 2004) is essential to compare the actual distributions of satisfaction in the two countries.

This is where the vignettes come in. A vignette briefly describes the relevant aspects of the life of a hypothetical person and is followed by a question asking the respondent to evaluate that person’s satisfaction with the given domain of life on the same five-point scale that was used for the self-report. Since the vignette descriptions are the same in the two countries, the vignette persons in the two countries have the same genuine satisfaction in the given domain. For example, respondents can be asked to evaluate the satisfaction with daily activities of a person whose genuine quality of daily activities is given by the dashed line. In country B, this will be evaluated as “satisfied.” In country A, the evaluation would be “not satisfied.” Since the actual level of satisfaction is the same in the two countries, the difference in the country evaluations must be due to DIF.

Using the scales in one country as the benchmark, the frequency distribution of evaluations in the other country can be adjusted by evaluating them on the benchmark scale. Once on the same scale, the corrected distribution of the evaluations can then be compared to

that in the benchmark country. In the example in the figure, this will lead to the correct conclusion that people in country A tend to be more satisfied than those in country B. The underlying assumptions are *response consistency* - a given respondent uses the same scale for self-reports and the vignette evaluations – and *vignette equivalence* – the vignette person’s genuine quality of the given domain of life is perceived as the same by respondents in countries A and B. We will apply the vignette approach to satisfaction with income, daily activities, and social contacts and family relations of the 65+ population in 11 European countries.

Figure 1. Comparing self-reported satisfaction in two countries in case of DIF



3. Model

The model consists of four parts: separate sub-models for genuine satisfaction, response scales, and vignette satisfactions in each of the three domains (income, social contacts, daily activities), and an ordered response equation for life satisfaction as a function of domain satisfactions and the health index.

The sub-models for each of the three domains are standard conditional logit (chopit) models, following King et al. (2004) and many others. For each domain $d=1,2,3$, define a latent self-satisfaction variable (s_{id}^*) as:

$$s_{id}^* = X_{id}\beta_d + \varepsilon_{id}, \quad (1)$$

where X_{id} is a vector of explanatory variables such as country dummies, gender, years of education, household composition indicators, and variables that are specific to the domain considered, such as household income for income satisfaction and variables related to social participation for satisfaction with daily activities and satisfaction with social contacts (see below for details). β_d is a vector of parameters. The error terms ε_{id} are assumed to be standard normally distributed and independent of each other and of X_{id} . Reported satisfaction with domain d (s_{id}) is a 5-point-scale ordered categorical variable based upon an underlying latent variable s_{id}^* :

$$s_{id} = j \quad \text{if} \quad \tau_{id}^{j-1} < s_{id}^* \leq \tau_{id}^j, \quad (2)$$

$$j = 1, \dots, 5 \text{ and } \tau_{id}^0 = -\infty, \tau_{id}^5 = +\infty.$$

If the thresholds between categories are the same for all respondents ($\tau_{id}^j = \tau_d^j$ for all i, j, d) then this gives three standard ordered probit models. The main distinguishing feature compared to this standard case is that all thresholds can vary with respondent characteristics X_{id} :

$$\tau_{id}^1 = X_{id}\gamma_d^1, \\ \tau_{id}^j = \tau_{id}^{j-1} + \exp(X_{id}\gamma_d^j), j = 2, 3, 4, \quad (3)$$

where the $\gamma_d^j, j = 1, 2, 3, 4$, are parameter vectors. Without additional information, γ_d^1 and β_d are not separately identified. To identify γ_d^1 , the vignette evaluations $V_{id}^k (k=1, \dots, K)$ are used, where K is the number of different vignettes in each domain; in our case, $K=2$. The vignette equivalence assumption implies that there exists a common actual quality of life in domain d , θ_d^k underlying the situation described by a given vignette k . The vignette evaluations are modelled as follows:

$$V_{id}^{*k} = \theta_d^k + v_{id}^k, \\ V_{id}^k = j \quad \text{if} \quad \tau_{id}^{j-1} < V_{id}^{*k} \leq \tau_{id}^j, \quad (4)$$

where V_{id}^k is the evaluation of vignette k in domain d by respondent i , and the v_{id}^k are errors, assumed to be normally distributed with mean 0 and variance σ_{vd}^2 , independent of each other, ε_{id} , and X_{id} .

The equation for genuine life satisfaction s_i^* is specified as follows:

$$s_i^* = \sum_{d=1}^3 \omega_d s_{id}^* + \omega_4 H_i + \varepsilon_i \quad (5)$$

Here H_i is the constructed health index (see below) which is always observed. Life satisfaction is therefore modelled as a weighted mean of domain satisfactions and health. Plus an additive idiosyncratic error term ε_i , assumed to be normally distributed and independent of the explanatory variables (X_{id} , $d = 1, 2, 3$ and H_i) and the other error terms ε_{id} , $d = 1, 2, 3$.

Observed life satisfaction is linked to genuine life satisfaction as follows:

$$s_i = j \quad \text{if} \quad \tau_i^{j-1} < s_i^* \leq \tau_i^j \quad (6)$$

What remains is to specify the thresholds τ_i^j . In other words, how do differences in response scales for domain satisfactions translate into differences in response scales for life satisfaction? We interpret the thresholds as (domain) satisfaction levels of hypothetical persons who are on the category boundaries. Moreover, we assume that someone who is on the boundary between, for example, satisfied and very satisfied in each domain (including health), this person is also on the boundary between good and very good in terms of life satisfaction. This seems a natural assumption in line with the compensatory nature of (5). It implies that the thresholds for life satisfaction are given by:

$$\tau_i^j = \sum_{d=1}^3 \omega_d \tau_{id}^j + \omega_4 \tau_{i4}^j \quad (7)$$

The thresholds τ_{id}^j are the same domain specific thresholds as in equation (3). Unfortunately, we do not have vignettes on satisfaction with health, so that we cannot identify variation in τ_{i4}^j across respondents. Instead, we have to make an assumption, and in the current version of the paper we assume that there is no variation across respondents in τ_{i4}^j : $\tau_{i4}^j = \tau_4^j$, and the parameters τ_4^j , $j = 1, \dots, 4$ are parameters to be estimated. This assumption may lead to an underestimation of response scale differences in self-assessed life satisfaction (unless the variation in response scales in health would not be there or would go in the opposite direction of the variation of response scales in the other domains) implying that our corrections for response scale differences would go in the right direction but would be incomplete.

The model can be estimated with maximum likelihood. Equations (1)-(4) can be estimated separately for each domain $d=1,2,3$, but equations (5)-(7) cannot be estimated without the parameters in (1)-(3). Consistent estimates for (5)-(7) could be obtained by plugging in the domain specific estimates of (1)-(3) but to get the correct standard errors, we estimated (1)-(7) simultaneously.

4. Data

The sample

The empirical analysis is based on data from the COMPARE sample, which is part of the second wave (2006-2007) of the Survey of Health, Ageing and Retirement in Europe (SHARE). SHARE includes extensive survey information on health, employment, financial situation, family and activities of a representative sample of the 50+ populations in 15 European countries (Börsch-Supan et al, 2005, 2008). The COMPARE sample consists of random subsamples of the complete SHARE samples in 11 countries. Respondents in these subsamples did the complete face to face SHARE interview, and were then asked to complete a drop-off questionnaire with self-assessment evaluations on satisfaction with different domains of life and to evaluate satisfaction with the same domains of life for hypothetical individuals described in the survey questions (the vignettes); see Van Soest (2008). SHARE respondents in the other subsamples got a completely different drop-off questionnaire. Response rates to the main survey and the drop-off were similar for the COMPARE sample and the remaining SHARE sample. The COMPARE sample includes 7,513 individuals aged 50+ from eleven European countries: Belgium, Czech Republic, Denmark, France, Germany, Greece, Italy, the Netherlands, Poland, Spain, and Sweden. We focus on non-working individuals of ages 65 and older (3,200 individuals). Among them, 2,799 respondents have answered to the questions about satisfaction with life, with household income, with social contacts, and with daily activities, along with the corresponding vignettes. After discarding 313 observations with missing or unreliable values for the explanatory variables used in the analysis, our final sample includes 2,486 individuals.

Self-assessments and vignettes

Satisfaction with life is measured using the following question:

Self-assessment: *How satisfied are you with your life in general?*

Very dissatisfied/ Dissatisfied/ Neither satisfied, nor dissatisfied/ Satisfied/ Very satisfied

These self-assessments are discussed extensively for the complete COMPARE samples (ages 50 and over) in Angelini et al. (2011). They also use vignettes on life satisfaction. We do not use the vignettes on life satisfaction, only those on satisfaction with domains of life. One of the purposes of our analysis is to analyze whether this leads to similar cross-country comparisons as the satisfaction with life as a whole vignettes.

Satisfaction with household income is measured using the following question:

Self-assessment: *How satisfied are you with the total income of your household?*

Very dissatisfied/ Dissatisfied/ Neither satisfied, nor dissatisfied/ Satisfied/ Very satisfied

The vignette questions about income satisfaction are the following:

Vignette 1: *Jim is married and has two children; the total after tax household income of his family is €1,500 per month. How satisfied do you think Jim is with the total income of his household?*

Very dissatisfied/ Dissatisfied/ Neither satisfied nor dissatisfied/ Satisfied/ Very satisfied

Vignette 2: *Anne is married and has two children; the total after tax household income of her family is €3,000 per month. How satisfied do you think Anne is with the total income of her household?*

Very dissatisfied/ Dissatisfied/ Neither satisfied nor dissatisfied/ Satisfied/ Very satisfied

The amounts used for net household income in the above vignettes, i.e. 1,500€ and 3,000€, are the amounts used in the vignette questions in France, Belgium and the Netherlands in which purchasing power of one euro was almost identical. In other countries, PPP adjusted amounts were used in local currencies.²

Bonsang and van Soest (2011a) study the self-assessments and vignettes on satisfaction with income for the complete population using the same data source. Here we only consider the 65+.

² The amounts in vignette 1 were 24,000CK in the Czech Republic, 14,200DK in Denmark, 1,550€ in Germany, 1,200€ in Greece, 1,450€ in Italy, 3,300PZ in Poland, 1,300€ in Spain and 15,400SK in Sweden. The amounts in vignette 2 were always twice as high. As pointed out by a referee, the different degrees of rounding might have effects on the responses, but we do not think this is a major issue.

Satisfaction with social contacts is measured using the following question:

Self-assessment: *How satisfied are you with your social contacts (with family, friends, etc.)?*

Very dissatisfied/ Dissatisfied/ Neither satisfied, nor dissatisfied/ Satisfied/ Very satisfied

The vignettes for satisfaction with social contacts are formulated as follows:

Vignette 1: *John is single, but gets on well with his relatives and has a large circle of friends. They often go out together to attend sporting events or to have a meal. How satisfied do you think John is with his social contacts (family, friends, etc.)?*

Very dissatisfied/ Dissatisfied/ Neither satisfied, nor dissatisfied/ Satisfied/ Very satisfied

Vignette 2: *Mary has been married for many years. Lately she has spent little time with her husband and they have been quarrelling more. They seem to prefer spending time with others rather than with each other. Both of them have many friends. How satisfied do you think Mary is with her social contacts (family, friends, etc.)?*

Very dissatisfied/ Dissatisfied/ Neither satisfied, nor dissatisfied/ Satisfied/ Very satisfied

Using the same data source, Bonsang and van Soest (2011b) study self-assessments and vignettes for social contacts for the complete 50+ population. Here we focus on the 65+.

The self-assessment and vignette questions about satisfaction with daily activities are as follows:

Self-assessment: *How satisfied are you with your daily activities (for example, your job, if you work)?*

Very dissatisfied/ Dissatisfied/ Neither satisfied, nor dissatisfied/ Satisfied/ Very satisfied

Vignette 1: *Mike has been retired for five years. He quit his job as soon as he could. He enjoys spending most of his time with friends and family and watches TV when he sometimes gets bored. How satisfied do you think Mike is with his daily activities?*

Very dissatisfied/ Dissatisfied/ Neither satisfied, nor dissatisfied/ Satisfied/ Very satisfied

Vignette 2: *Sally has been retired for five years. Although she enjoys spending time with her children and grandchildren, she still misses the contacts with her colleagues and would have liked to keep working. How satisfied do you think Sally is with her daily activities?*

Very dissatisfied/ Dissatisfied/ Neither satisfied, nor dissatisfied/ Satisfied/ Very satisfied

The health index

In order to control for health we construct a health index following the strategy of Bound et al. (1999) by using objective health variables to predict self-reported overall health. This health index attenuates the reporting bias in the self-reported measure of health and does not suffer from multi-collinearity. The health index is obtained by estimating an ordered probit model with self-reported overall health as the dependent variable and a large set of detailed objective health indicators as explanatory variables. From the estimation results, we compute the predicted latent variable and define it as our health index. The objective health indicators include a set of dummies related to chronic diseases diagnosed by a doctor (17 chronic diseases), a list of 12 symptoms, a list of 10 limitations with activities of daily living (ADL) and 13 limitations with instrumental activities of daily living (IADL). We also include as explanatory variables the body-mass index (BMI), a measure of grip strength, and three measures related to cognitive functioning (word recall test score, fluency test score, and numeracy test score).

Other Explanatory variables

In addition to country dummies, the domain satisfaction models includes socio-demographic characteristics such as gender, age, reported years of education, household size, and marital status (a dummy for living with a partner, married or not). Income is measured by the logarithm of reported monthly net household income last month, adjusted by PPP.³ Wealth is taken into account by the logarithm of household total net worth (the sum of all financial and real assets, minus liabilities). We included several variables related to family ties: the number of children, a dummy for individuals having a co-residing child and the (log of the) number of annual contacts with children. To measure the involvement of the older individuals in non-professional activities, we added a set of dummies related to different types of activities: “Doing voluntary or charity work”, “caring for a sick or disabled adult”, “providing help to

³ Outliers and missing incomes are imputed using an alternative income measure (last year’s income of all household members) as one of the predictor variables. An appendix with details is available upon request from the authors.

friends or neighbours”, “attending an educational or training course”, “going to a sports club, a social club or another kind of club”, “taking part in activities of a religious organization”, and “taking part in a political or community-related organization”. In the literature, such activities are seen as an important aspect of social capital (see Helliwell and Putnam, 2004), but they can also be a source of social contacts.

5. Estimation Results

In this section we present the estimation results for a parsimonious specification of the model: we include country dummies in all equations (since we focus on cross-country differences), but use other covariates only where they are specifically relevant for satisfaction in a given domain. Thresholds are allowed to be country specific, but shifts across countries are assumed to be the same for all thresholds in a given domain, that is, the parameter vectors driving distances between the first and other thresholds $\gamma_d^j, j = 2, 3, 4$, are zero except for the intercept terms. This is to reduce the total number of parameters to be estimated (which is still 128 in our “parsimonious” specification).

Satisfaction with daily activities

Table 1 presents the parameters of equation (1) for satisfaction with daily activities using a standard ordered probit model (first column) and using our chopit model (second column); both models use the same explanatory variables; the standard ordered probit model assumes thresholds are constant across countries while in the chopit model, they are country specific.

The largest difference the two columns is in the country dummies. This is not surprising, given that the difference between the specifications refers to the country dummies. The benchmark country is Germany, so the coefficients should be interpreted as differences with Germany, keeping the other variables in the equations constant. Column 1 shows that older individuals from the Netherlands report the highest satisfaction report the highest satisfaction with their daily activities, followed by Denmark. The Netherlands remains by far the best country in terms of daily activities in the chopit model, but Denmark falls back behind Sweden and Italy. Particularly Italy improves substantially when response scales are kept constant. This is because the Danes use quite optimistic response scales, easily reporting being satisfied or very satisfied (as is apparent from the raw data on their vignette evaluations) while the Italians are very hesitant to do so. Correcting for this makes the Danes less and the

Italians look more satisfied. On the opposite end of the ranking, respondents from Poland report being the least satisfied among the selected countries, followed by France, *ceteris paribus* (column 1). Polish respondents remain the least satisfied with their daily activities if response scale differences are controlled for, but the position of France improves, since the French tend to use less positive responses. Spain drops a few places to last but one. These shifts in the country ranking are in line with what we saw in the vignette evaluations (Tables 2 and 3).

Table 1. Equation (1) for satisfaction with daily activities

	Model without DIF		Model with DIF	
	par.	t-val	par.	t-val
Constant	2.869*	42.295	2.940*	40.582
Greece	0.155*	7.310	-0.080*	2.084
Belgium	0.016	0.939	-0.120*	3.726
Denmark	0.346*	18.162	0.131*	3.439
France	-0.123*	4.898	0.011	0.217
Italy	0.081*	3.638	0.240*	5.954
Netherlands	0.903*	33.547	0.971*	18.079
Poland	-0.395*	19.846	-0.554*	14.412
Spain	0.018	0.710	-0.192*	3.692
Sweden	0.133*	5.930	0.183*	4.583
Czech Rep	-0.100*	5.538	-0.180*	5.433
Female	-0.078*	7.815	-0.087*	8.513
Age	-0.017*	21.197	-0.017*	20.030
Educ. years	0.014*	10.217	0.011*	7.627
ln fam size	0.095*	3.314	0.077*	2.673
Charity	-0.117*	6.865	-0.104*	5.950
Caring	-0.256*	11.395	-0.191*	8.176
Friend	0.250*	14.048	0.257*	13.956
Training/educ	0.184*	7.016	0.167*	6.097
Sports	0.342*	24.869	0.329*	23.083
Religion	0.077*	5.050	0.053*	3.384
Political	0.217*	7.205	0.286*	8.721
Couple	0.000	0.008	-0.013	0.590
Children	-0.020*	5.562	-0.021*	5.568
Residing child.	0.008	0.325	0.021	0.843
ln contact	-0.002	0.807	0.000	0.055

Notes: Model without DIF: model with constant thresholds; Model with DIF: model with country specific thresholds;

*: significant at two-sided 5 percent level.

See Appendix for variable definitions

For the other variables, the results are similar in the two columns. Keeping other variables constant, satisfaction with daily activities is significantly larger for men than for women, decrease significantly with age, and increases with years of education and family size. Van Praag et al. (2003) estimate a similar equation for leisure activities of the complete adult population. They find the same gender difference but quite different age and education

patterns, suggesting that the determinants of satisfaction with daily activities for the 65+ may be rather different from the factors that drive satisfaction with leisure for a general adult population.

As expected, satisfaction with daily activities is positively related to participation in non-professional activities, especially helping friends or neighbours, attending training or educational courses, going to a sport or social club, and, to a lesser extent, taking part in religious activities. Although not directly comparable, this is broadly in line with Chen (2001) who finds a strong positive and significant effect of leisure activities on satisfaction with life of the 60+ in Taiwan. Caring for a sick or disabled person or doing voluntary work or work for a charity have a negative effect on satisfaction with daily activities. These activities may be seen as work rather than leisure activities. Family-related variables, on the other hand, only have small effects. Only the number of children (not co-residing with their parents) is significant and has a negative effect.

Satisfaction with social contacts and family relations

Table 2 presents the parameters of equation (1) for satisfaction with social contacts, for the same two specifications not allowing and allowing for country specific thresholds. Column 1 shows that Sweden, Denmark and Germany report the highest satisfaction with social contacts while the Greek, Dutch and Italian respondents give the lowest ratings, keeping other variables constant. Controlling for response scale differences (column 2) again lowers the rank of Denmark and improves the situation of Italy. It suggests that social contacts in the Netherlands, given the other variables such as participation in social activities, are of lesser quality than in any other country considered. These shifts are qualitatively similar to what we.

Other than for the coefficients on the country dummies, the results are similar in the two columns. The only exception is years of education – this has a significant positive impact without controlling for response style differences, but the effect becomes negative and insignificant when response scale differences are allowed for but they control for permanent income. Van Praag and Ferrer-i-Carbonell (2008) find a negative effect of years of education for the complete adult population of the UK. Keeping all other factors constant, women are significantly more satisfied with social contacts than men, and satisfaction with social contacts rises with age. Both results are in line with those of Van Praag and Ferrer-i-Carbonell (2008). The age results differs, however, from that of Motel-Klingebiel et al. (2004) who find

a negative association between age and satisfaction with social relationships for the general adult (25+) population in five countries.

Table 2. Equation (1) for satisfaction with social contacts

	Model without DIF		Model with DIF	
	par.	t-val	par.	t-val
Constant	1.209*	16.423	1.234*	15.791
Greece	-0.490*	21.032	-0.455*	9.885
Belgium	-0.214*	11.989	-0.343*	9.098
Denmark	0.088*	4.637	-0.126*	3.275
France	-0.363*	14.687	-0.309*	6.146
Italy	-0.411*	18.851	-0.260*	6.003
Netherlands	-0.487*	16.921	-0.701*	14.354
Poland	-0.065*	2.835	-0.078+	1.658
Spain	-0.139*	4.953	-0.246*	4.176
Sweden	0.165*	7.381	0.127*	2.790
Czech Rep	-0.213*	9.913	-0.159*	3.818
Female	0.110*	10.351	0.095*	8.654
Age	0.006*	7.200	0.007*	7.763
Educ. years	0.003*	2.336	-0.002	1.237
ln fam size	0.203*	7.467	0.191*	7.039
Charity	0.209*	13.131	0.196*	12.081
Caring	0.068*	3.119	0.090*	4.016
Friends	0.021	1.190	0.040*	2.205
Training/Educ.	0.040	1.637	0.063*	2.529
Sports	0.187*	13.377	0.185*	13.184
Religion	0.138*	7.975	0.137*	7.846
Political	0.215*	6.291	0.274*	7.971
Couple	-0.087*	4.256	-0.079*	3.839
Children	-0.042*	10.341	-0.045*	10.905
Residing child.	-0.103*	4.779	-0.119*	5.645
ln contacts	0.086*	31.956	0.084*	30.601

Notes: Model without DIF: model with constant thresholds; Model with DIF: model with country specific thresholds;

*: significant at two-sided 5 percent level.

See Appendix for variable definitions

Contacts with children have a strong and significant positive effect given the number of children and co-residence with children or not. Somewhat surprisingly, these latter variables have a negative effect, perhaps because it is the effect keeping number of contacts and other variables constant that we are estimating. Non-professional activities play an important positive role: Other than for satisfaction with daily activities, doing charity or voluntary work or caring for a sick or disabled person is significantly positively associated with satisfaction with social contacts. Participating in a sports or social club and taking part in political or community-related organizations also have a significant positive effect on satisfaction with social contacts, and so do helping friends or neighbours or attending training or an education program.

Income satisfaction

Table 3 presents the parameter estimates of the main equation (1) for income satisfaction. not accounting (column 1) and accounting (column 2) for differences in response scales. Country dummy estimates in column 1 indicate that, conditional on income and other covariates the Greek elderly report the lowest income satisfaction level followed by the French and the Italians. The position of all these three countries improves substantially when response scale differences are incorporated in the model (column 2). Particularly for the Greeks the unusually low position in column (1) seem mainly due to their pessimistic responses. Danish and Swedish respondents report the highest levels of income satisfaction, before as well as after correction for response scale differences. Interestingly, keeping the other covariates constant, the financial situation of Polish respondents is worse than in any other country. This is hidden by optimistic Polish response scales in the model in column 1. These results on the country ranking are quite different from the results we found in Bonsang and van Soest (2011a) for the complete 50+ populations, but the differences between the two models are qualitatively similar.

Table 3. Equation (1) for satisfaction with Income

	Model without DIF		Model with DIF	
	par.	t-val	par.	t-val
const inco	-4.604*	24.883	-4.464*	25.113
Greece	-0.476*	15.400	-0.014	0.304
Belgium	0.075*	2.685	0.080	1.893
Denmark	0.573*	19.695	0.609*	14.267
France	-0.177*	4.891	0.080	1.484
Italy	-0.146*	4.534	0.011	0.239
Netherland	0.007	0.167	-0.074	1.068
Poland	-0.107*	2.851	-0.411*	7.933
Spain	-0.019	0.514	-0.124*	2.448
Sweden	0.286*	7.355	0.413*	6.908
Czech Rep	0.148*	4.173	-0.048	1.040
Female	-0.067*	4.910	-0.061*	4.502
Age	0.017*	15.651	0.017*	14.685
Educ years	0.000	0.167	0.004*	2.060
ln hhsize	-0.352*	18.451	-0.356*	17.543
ln hhinc	0.651*	30.299	0.629*	30.202
ln wealth	0.036*	15.651	0.031*	13.038

Notes: Model without DIF: model with constant thresholds; Model with DIF: model with country specific thresholds;
 *: significant at two-sided 5 percent level.
 See appendix for variable definitions.

As expected, household income has a strong positive effect on income satisfaction, while household size has a substantial negative effect. In terms of equivalence scales, the chopit model estimates imply that an increase in family size from one to two household members would require an increase in household income of about 48% to keep income satisfaction constant. This is larger than what we found for the complete 50+ population in Bonsang and van Soest (2011a). Conditional on income (and other covariates), higher educated individuals are significantly more satisfied with their income according to the chopit model, in line with the notion that higher educated people may have higher permanent income. A similar interpretation may apply to the positive and significant effect of financial wealth. Women tend to report lower income satisfaction than men with the same income. Age has a positive significant effect.

Threshold Parameters

Table 4 presents the estimates of the parameters in the threshold equations (3) for the three satisfaction domains. Again, Germany is the benchmark country. A positive coefficient for a given country dummy means that respondents in that country use higher thresholds than German respondents, implying that they less easily call someone satisfied, that is, tend to give fewer positive or optimistic evaluations than Germans for a given genuine quality of life in the given domain. The table shows, for example, that Italian and French respondents tend to use the more negative qualifications more often. Correcting for this will make them look better. This is in line with what we already concluded in the discussion of Tables 1, 2 and 3. Polish and Spanish respondents, on the other hand, tend to use lower thresholds than German respondents, implying that they more easily give optimistic qualifications. Correcting for such differences makes them look worse compared to Germany, in line with what we saw comparing the two sets of estimates in Tables 1, 2 and 3. Something similar applies to Denmark, Belgium, the Netherlands, and the Czech Republic, though not always to the same extent and not for all domains. Only for Greece the results are very different for the three domains: Greek respondents tend to use positive qualifications for satisfaction with daily activities, but use by far the least positive qualifications for satisfaction with income. (The latter was also found by Bonsang and van Soest (2011a) for the complete 50+ population.). For the other countries, there is much more similarity across domains, suggesting that the use of more positive or more negative qualifications could be related to cultural differences that are not domain specific.

Table 4. Estimates of threshold parameters

	satisfaction with ...					
	daily activities		social contacts		household income	
	par.	t-val.	par.	t-val.	par.	t-val.
gamma 1						
Greece	-0.090*	7.007	0.025	1.627	0.180*	11.395
Belgium	-0.055*	5.302	-0.046*	3.988	-0.004	0.264
Denmark	-0.086*	7.244	-0.075*	6.288	0.018	1.419
France	0.050*	3.186	0.016	1.027	0.103*	5.799
Italy	0.068*	5.357	0.059*	4.603	0.066*	4.585
Netherlands	0.005	0.293	-0.081*	5.633	-0.028	1.339
Poland	-0.071*	5.531	-0.006	0.427	-0.177*	10.650
Spain	-0.097*	5.887	-0.038*	2.120	-0.052*	3.335
Sweden	0.006	0.491	-0.006	0.440	0.049*	3.024
Czech Rep	-0.045*	4.207	0.004	0.286	-0.064*	4.634
gamma 2	-0.081*	4.130	-0.209*	9.595	-0.248*	12.227
gamma 3	-0.249*	9.743	-0.315*	10.252	-0.204*	7.262
gamma 4	0.513*	28.879	0.635*	29.967	0.559*	23.891

Life satisfaction and vignette equations

Table 5 contains the remaining parameter estimates of the model allowing for differences in response scales. The most interesting of these are the weights for the various domains. As expected, all of them are significantly positive, as is the health index. This implies that the three domains (daily activities, financial well-being, and social contacts and family life), as well as health contribute significantly to overall well-being. Satisfaction with daily activities and social contacts have much higher weights than satisfaction with income – suggesting that, for this age group, satisfaction with income is less important than these other two domains. Although the health index is defined in a completely different way than the domain satisfactions, the way in which it is normalized implies that it has almost the same variance as satisfaction with income, so that its weight in Table 5 can be compared to the other weights. It is surprisingly low – much lower still than the weight of satisfaction with income.

The other parameters are auxiliary parameters that are presented for completeness. The ranking of the vignette dummies for each domain are in line with what we would expect on the basis of the vignette descriptions and the raw data. For daily activities and social contacts, vignette 1 describes someone with a better quality of life in the given domain than vignette 2, while the reverse is true for the income vignettes (vignette 2 has twice the income of vignette 1). The standard deviations of the errors in the vignette evaluations are all smaller than 1 – the standard deviation of the domain satisfactions in equation (1). This is plausible

since the errors in equation (1) capture both noise and idiosyncratic variation in genuine quality of life in the given domain, while the latter is missing in the vignette equation (4).

Table 5. Estimates of life satisfaction equation and other parameters

Weights in life satisfaction

daily activities		social contacts		household income		health	
par.	t-val.	par.	t-val.	par.	t-val.	par.	t-val.
0.855*	22.570	0.592*	18.487	0.350*	11.649	0.167*	5.965

vignette dummies

	daily activities		social contacts		household income	
	par.	t-val.	par.	t-val.	par.	t-val.
Vignette 1	1.910*	85.173	2.188*	95.319	0.929*	37.257
Vignette 2	1.115*	45.349	0.897*	43.232	2.128*	82.585
St. deviation	0.620*	74.378	0.652*	75.526	0.793*	74.782

error terms

Threshold coefficients health index

	par.	t-val.
Threshold 1	-2.538*	3.236
Threshold 2	-4.242*	7.046
Threshold 3	-2.818*	7.356
Threshold 4	0.106	0.379

Life Satisfaction

Figures 2 and 3 show simulated distributions of life satisfaction across countries on a five point scale using the estimates of the model accounting for differences in response scales. Figure 2 uses each country's own thresholds, in line with what people actually report. Figure 3 is a counterfactual simulation where each respondent uses German thresholds, irrespective of their country of residence. Cross-country differences in Figure 2 can be due to genuine differences in life satisfaction or to differences in response scales. The latter differences are neutralized in Figure 3 so that the differences in Figure 3 reflect genuine differences in life satisfaction only. Comparing Figures 2 and 3 shows the influence of correcting for response scale differences. By construction, the distribution for Germany is the same in the two figures, and all changes should be interpreted as in comparison to Germany.

The countries in each figure are ordered from least satisfied to most satisfied. (To be precise, from the lowest percentage to the highest percentage satisfied or very satisfied). Polish respondents are worse off in both figures. The Dutch report the highest satisfaction levels, but this is partly due to the positive way in which they use the scales. Correcting for this by giving them the German scales worsens their position in the ranking to fourth place. Similarly, Spain also falls substantially when the fact that Spanish respondents tend to use

positive evaluations (see Table 4) is corrected for. Spain falls from sixth to tenth place. On the other hand, the rankings of France and Italy improve – since these are the countries where respondents tend to give more critical evaluations of the vignettes.

Figure 2. Predicted life satisfaction using national thresholds.

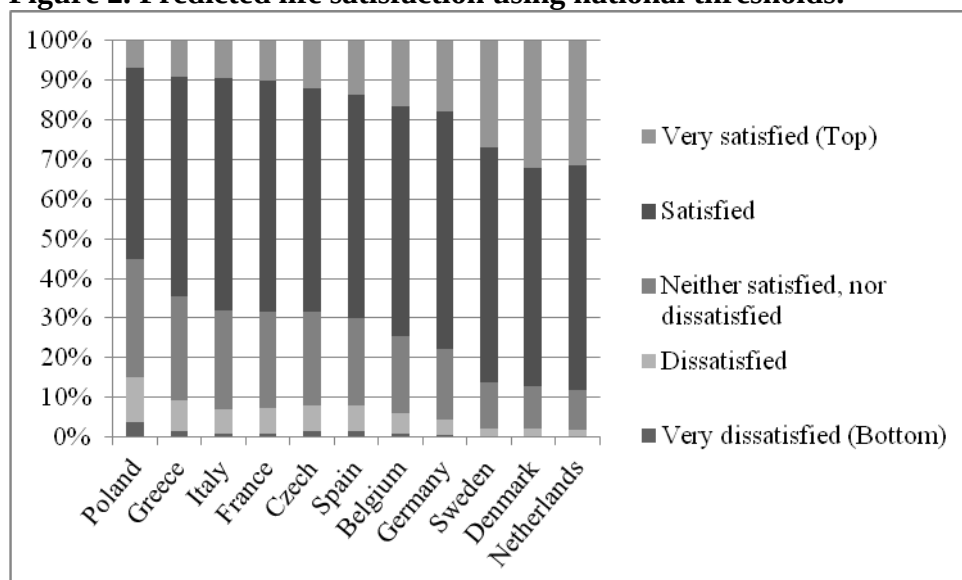
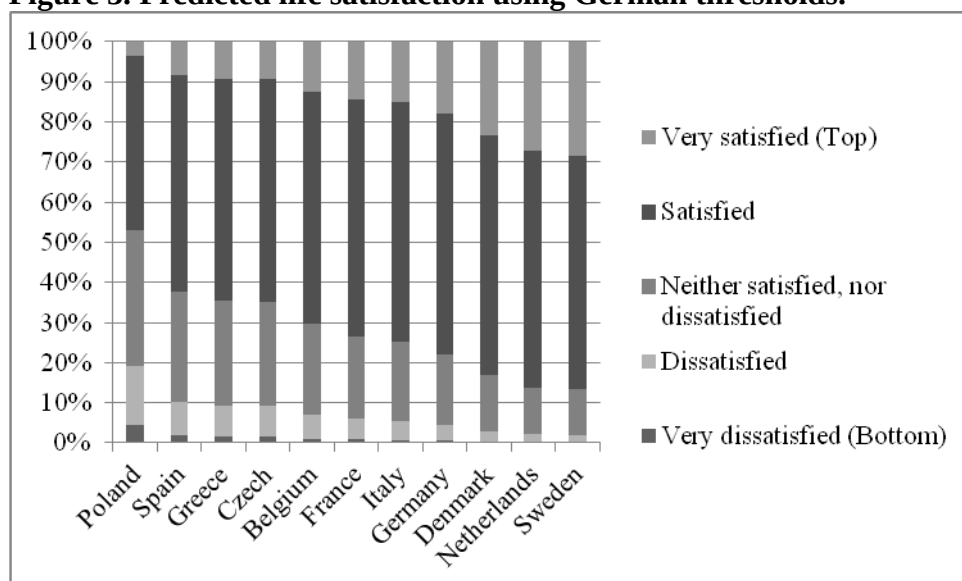


Figure 3. Predicted life satisfaction using German thresholds.



Conclusions to be completed

References to be completed