

Do Danes and Italians Rate Life Satisfaction in the Same Way? Using Vignettes to Correct for Individual-Specific Scale Biases*

Viola ANGELINI,[†]Danilo CAVAPOZZI,[‡]Luca CORAZZINI,[§]Omar PACCAGNELLA[¶]

August, 2008

Abstract

Self-reported life satisfaction is highly heterogeneous across similar countries. We show that this phenomenon can be largely explained by the fact that individuals adopt different scales and benchmarks in evaluating themselves. Using a cross sectional dataset on individuals aged 50 and over in ten European countries, we compare estimates from an Ordered Probit in which life satisfaction scales are invariant across respondents with those from a Hopit model in which vignettes are used to correct for individual-specific scale biases. We find that variations in response scales explain a large part of the differences found in raw data. Moreover, the cross countries ranking in life satisfaction dramatically depends on scale biases.

JEL classification: C42, D12, I31, J14.

Keywords: Life satisfaction, scale biases, vignette, counterfactuals.

*We thank Luigino Bruni and Benedetto Gui for useful comments and suggestions. This paper uses data from SHARE. The SHARE data collection has been primarily funded by the European Commission through the 5th framework programme (project QLK6-CT-2001-00360) and the 6th framework programme (projects SHARE-I3, RII-CT-2006-062193, and COMPARE, CIT5-CT-2005-028857). Additional funding came from the U.S. National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, Y1-AG-4553-01 and OGHA 04-064, IAG BSR06-11, OGHA04-064).

[†]Department of Economic Studies “Marco Fanno”, University of Padua [email: viola.angelini@unipd.it].

[‡]Department of Economic Studies “Marco Fanno”, University of Padua [email: danilo.cavapozzi@unipd.it].

[§]Department of Economic Studies “Marco Fanno”, University of Padua; ISLA, Bocconi University [email: luca.corazzini@unipd.it].

[¶]Department of Economic Studies “Marco Fanno”, University of Padua [email: omar.paccagnella@unipd.it].

1 Introduction

Empirical studies on subjective well-being seek to identify the relation between life satisfaction (or other proxy of utility such as happiness) and economic, demographic, social and physical characteristics of individuals. In particular, in order to collect data on individuals' well-being, researchers use general surveys asking individuals to self-report their life satisfaction on a scale which goes from a minimum value to indicate very low satisfaction to a maximum value which expresses very high satisfaction.

There is no doubt about the importance of these studies. Starting from the famous Easterlin's paradox which states the existence of a reversed U-shaped relation between happiness and income in the US between 1946-1996 (Easterlin, 2001), social scientists have devoted a large quantity of research using self-reported data to identify the main determinants of life satisfaction. As an established result (see Bruni and Porta, 2005, Frey and Stutzer, 2002a, 2002b for surveys), "*money is not enough to make people happy*" and in addition to economic factors such as income,¹ low inflation rate,² being employed,³ there are non economic variables which have a significant and positive effect on life satisfaction: high education,⁴ good health conditions and being physically autonomous,⁵ being female,⁶ being white rather than black,⁷ being local rather than immigrant,⁸

¹Clark and Oswald (1994); Frey and Stutzer (2000); Gerdtham and Johannesson (2001); Di Tella et al. (2001, 2003); Clark (2003); Van Praag et al. (2003); Alesina et al. (2004); Blanchflower and Oswald (2004); Clark and Lelkes (2005); Easterlin (2005); Ferrer-i-Carbonell (2005); Van Praag and Baarsma (2005); Van den Berg and Ferrer-i-Carbonell (2005); Di Tella and MacCulloch (2006); Oswald and Powdthavee (2007).

²Di Tella et al. (2001, 2003); Alesina et al. (2004).

³Clark and Oswald (1994); Di Tella et al. (2001, 2003); Helliwell (2003); Alesina et al. (2004); Blanchflower and Oswald (2004); Clark and Lelkes (2005); Ferrer-i-Carbonell and Gowdy (2005); Haller and Hadler (2006); Oswald and Powdthavee (2007).

⁴Blanchflower and Oswald (1997); Frey and Stutzer (2000); Easterlin (2001); Schwarze and Harpfer (2003); Di Tella et al. (2003); Blanchflower and Oswald (2004); Alesina et al. (2004); Stutzer and Frey (2005, 2006).

⁵Clark and Oswald (1994); Winkelmann and Winkelmann (1998); Clark and Oswald (2002); Clark (2003); Helliwell (2003); Helliwell and Putnam (2004); Ferrer-i-Carbonell and Frijters (2004); Clark and Lelkes (2005); Blanchflower and Oswald (2005); Oswald and Powdthavee (2007).

⁶Shmotkin (1990); Di Tella, et al. (2001, 2003); Easterlin (2001, 2006); Biblarz and Gottainer (2000); Alesina et al. (2004); Blanchflower and Oswald (2004); Haller and Hadler (2006); Meier and Stutzer (2008).

⁷Davis (1984); Hughes and Thomas (1998); Alesina et al. (2004); Blanchflower and Oswald (2004); Graham (2005).

⁸Frey and Stutzer (2000); Stutzer (2004).

being involved in the community and volunteering,⁹ having frequent contacts with family, friends and neighbors,¹⁰ being married and having children,¹¹ praying and attending religious communities.¹²

A crucial methodological issue which limits cross-country comparability is that, when asked to self-report their life satisfaction, individuals, who are similar for both economic and non economic conditions, can use different benchmarks or scales in evaluating themselves (Winkelmann and Winkelmann, 1998; Clark and Oswald, 2002; Ferrer-i-Carbonell and Frijters, 2004; Senik, 2004; Clark et al., 2005). For instance, Clark et al. (2005) affirm that “*one worry regarding statistical analysis of subjective variables is that some people look at life pessimistically or optimistically, even though there is really no difference in their level of well-being*” (page 118). In psychometrics this phenomenon is called *differential item functioning* (henceforth DIF), defined as the inter-personal and inter-cultural variation in interpreting and using the response categories for the same question (Holland and Wainer, 1993).

On the one hand, scale biases can depend on group-specific characteristics such as socio-economic conditions (Frey and Stutzer, 2002b) or cultural connotations (Uchida et al., 2004; Diener and Suh, 2000; Inglehart and Klingemann, 2000). For instance, there is evidence showing that seemingly similar European countries are associated with highly dissimilar self-reported levels of life satisfaction (Inglehart and Rabier, 1986). Rather than being caused by the existence of objective differences in the conditions of inhabitants, this puzzling finding can be explained by “*different interpretations of numerical scales*” across societies (Frey and Luechinger, 2007, page 220) or with “*cultural differences in the norms that govern self-descriptions*” (Kahneman et al., 2004, page 430).

On the other hand, scale biases can be determined by individual-specific characteristics such as the psychological status of the individual (De Neve and Cooper, 1998), her personal interpre-

⁹Thoits and Hewitt (2001); Meier and Stutzer (2008).

¹⁰Björnskov (2003); Helliwell and Putnam (2004); Kahneman et al. (2004); Pichler (2006).

¹¹Clark and Oswald (1994); Waite (1995); Stack and Eshleman (1998); Diener and Suh (2000); Helliwell (2003); Alesina et al. (2004); Helliwell and Putnam (2004); Blanchflower and Oswald (2004, 2005); Dehejia et al.(2005); Kohler et al. (2005); Clark and Lelkes (2005); Haller and Hadler (2006).

¹²Cohen (2002); Helliwell (2003); Helliwell and Putnam (2004); Clark and Lelkes (2005); Dehejia et al. (2005); Haller and Hadler (2006).

tations of life satisfaction, or her intellectual skills in understanding questions and processing the information needed to answer (Frey and Stutzer, 2002b).

When panel data are available, conventional fixed or random effects (Kapteyn et al., 2007) or latent class techniques (Clark et al., 2005) can be used in order to take into account time-invariant DIF bias. However, these approaches present at least two main problems when used to analyze self-reported life satisfaction. First, they do not allow to control for DIF bias with cross sectional data in which respondents are not followed over time. Second, given the psychological nature of life satisfaction, the hypothesis of individuals' time-invariant scales which these methodologies are based on can be criticized. Indeed, the scale adopted by an individual to evaluate herself can vary over time according to her actual mood (Kahneman et al., 2003) and her socio-economic conditions.

In this paper, we use a vignettes methodology to assess whether individual and cross-country differences in self-reported life satisfaction are genuine or they just reflect individual-specific scale biases. Following this approach, individuals are presented with two categories of questions on life satisfaction. First, as in traditional questionnaires, individuals are asked to self-report their level of life satisfaction. Second, each individual is asked to evaluate the life satisfaction of one or more hypothetical persons described in particular situations or conditions and kept constant across interviewees (anchoring vignettes). In such a way it is possible to find a standard, that is an anchor to which the response categories of the survey questions will be attached. Therefore, by collecting individuals' evaluation of the anchoring vignettes it is possible to filter out the level of self-reported life satisfaction from the DIF bias measured by vignettes and enhance the comparability of subjective assessments across individuals.

Although vignettes have already been successfully used in several domains, such as political efficacy (King et al., 2004), health (Jürges, 2007; Salomon et al., 2004; Bago d'Uva et al., 2008), work disability (Kapteyn et al., 2007; van Soest et al., 2007) and job satisfaction (Kristensen and Johansson, 2008) to correct for individual-specific heterogeneity with cross sectional data, to our knowledge there is no empirical study applying this methodology to self-reported life satisfaction.

In our analysis we use data from the 2006-2007 wave of the Survey of Health, Ageing and Retirement in Europe (SHARE), an innovative and rich survey collecting cross sectional micro data on health, socio-economic, demographic conditions of aged persons. Considering the demographic situation of the “Old Europe,” a world region characterized by a steadily increasing life expectancy (passing from 70 to 79 between 1960 and 2008), a (almost) null birth rate and the highest proportion of people over the age of 65 (Eurostat, 2005), it is easy to understand the growing interest of both policy makers and social scientists for the conditions of aged persons.

As concrete consequences, not only important institutional changes in the pension, health and security systems have been undertaken by European countries to improve aged persons’ well-being, but also economists, psychologists and sociologists have produced a large amount of research studying the determinants of life satisfaction in later life.

The level of life satisfaction in later life depends on the trade-off between counterbalancing effects. On the one hand, there are at least three characteristics of oldness which suggest a negative relation between oldness and life satisfaction (George, 2006). First, oldness is associated with deteriorating health conditions. Second, later life is generally associated with weak and risky economic conditions with costs for health and assistance which are increasing in age and pensions which are constant over time. Third, oldness is associated with several sources of social exclusion. Physical problems, retirement and social stigmas reduce aged persons’ active participation to social life. Even within the basic unit of social life, the family, aged persons experience social exclusion through widowhood and children approaching adulthood.

On the other hand, the “role” theory of the aging process suggests that self-integration, insight, and positive psychological traits such as satisfaction and self-esteem are increasing in age and positively influence the level of self-reported life satisfaction (Gove et al. 1989).

Although in modern societies one can expect the negative aspects of oldness to be stronger than the positive elements, the empirical evidence supports the existence of a significant and positive relation between life satisfaction and age in later life (Mroczek and Kolarz, 1998; Charles et al.,

2001; Yang, 2008; Clark and Oswald, 1996).

Concerning the determinants of life satisfaction in later life, it has been found that poor health conditions¹³ and physical limitations¹⁴ strongly reduce life satisfaction. Active social relations¹⁵ as well as being married rather than single¹⁶ increase life satisfaction significantly. Moreover, retired aged persons who use their time in job and voluntary activities report significantly high levels of life satisfaction relatively to those who are more inactive.¹⁷ Income and wealth have a positive effect on aging people with the effect of the latter being apparently stronger.¹⁸ Empirical evidence of the effects of retirement on life satisfaction of aged persons is mixed. There are several studies finding that retirees are less satisfied with their lives than workers.¹⁹ However, more recent contributions find opposite results once controlled for economic variables and retirement conditions.²⁰

In our paper, we present new and complete analysis of the relation between life satisfaction in later life and a large number of economic and non economic variables.

The main results of our paper can be summarized as follows. We compare estimates from an Ordered Probit model in which scales are assumed to be constant across individuals with those from a Hopit model in which vignettes are used to correct for the DIF bias. According to the Ordered Probit, once controlled for economic, demographic, health and social conditions, Danes and Italians result to be the most and the least satisfied with life respectively. However, by correcting for scale biases, the ranking across countries dramatically changes. The difference in self-reported life satisfaction between Danes and Italians disappears and the Netherlands and Czech Republic respectively replace Denmark and Italy in the ranking of life satisfaction. A formal likelihood ratio test strongly rejects the ordered probit not allowing for response scale variation against the more

¹³Rowe and Kahn (1998); Atchley (1999); Hilleras et al. (2001a, 2001b); Easterlin (2003).

¹⁴Dulin and Pachana (2005).

¹⁵Diener and Seligman (2004); Chan and Lee (2006); Cheng and Chan (2006).

¹⁶Easterlin (2003); Bierman et al. (2006).

¹⁷Butrica and Schaner (2005).

¹⁸Wiatrowski (1993); Easterlin (2001, 2003); Saris (2001); Diener and Biswas-Diener (2002); Diener and Seligman (2004); Arendt (2005); Arthaud-day and Near (2005); Calvo et al. (2007).

¹⁹Kutner et al. (1956), Atchley and Robinson (1982); Bosse et al. (1987); de Grace et al. (1994); Charles (2004).

²⁰Jackson et al. (1993); Easterlin (2003).

general Hopit model. We find evidence that the thresholds significantly depend on the explanatory variables used in the regressions. We also present results from counterfactual simulations to study how the percentage of satisfied individuals in a country change when the response scale of other countries is imposed to its inhabitants. When the Danish scale is used, more than 95 percent of respondents in all countries would rate themselves as satisfied with their own life. The picture considerably changes when using the Italian response scale. Indeed, the proportion of satisfied individuals according to the Italian scale drops significantly in all countries. Thus, variations in response scales explain a large part of the differences found in raw data. Concerning the determinants of individuals' well-being, once controlled for individual-specific scale biases, we find that life satisfaction in later life is positively and significantly correlated with being married, being female, being not retired and having not experienced unemployment, having frequent (almost weekly) contacts with sons, parents and grandchildren, enjoying good health conditions, being active in the community and volunteering, high economic wealth. We also confirm the existence of a non linear relation between age and life satisfaction.

The rest of the paper proceeds as follows. In section 2, we describe the dataset and we present descriptive statistics of the variables used in our analysis placing particular emphasis on the structure of the vignettes. In section 3, we specify the econometric model based on the vignettes (the Hopit model) showing its properties and stating the statistical assumptions it is based on. In section 4, we compare estimates from an Ordered Probit model in which scales are assumed to be constant across individuals with those from the more general Hopit model and we present formal tests to evaluate their estimation performances. In section 5, by using the Hopit model, we estimate the country-specific scales of life satisfaction and we present results from counterfactual simulations to study how the percentage of satisfied individuals in a country change when the response scale of other countries are imposed to its inhabitants. Finally, in section 6 we discuss our findings and we conclude.

2 Data and Descriptive Statistics

Our sample is drawn from the wave 2006-2007 of the Survey of Health, Ageing and Retirement in Europe (SHARE). SHARE is a unique and multidisciplinary dataset that contains a large amount of information on both the economic and non economic conditions of individuals aged 50 and over. As shown by Table 1, in our analysis we use a large number of individuals' characteristics which can be ideally classified into ten different categories: demographic conditions, education, employment, civil status, family backgrounds, health status, social activities, politics, religion and financial status.

After the personal interview (CAPI, Computer Assisted Personal Interview), as part of the COMPARE project, a subset of respondents are asked to fill out additional questions on a paper and pencil questionnaire, which focuses on self assessments and vignette evaluations. In particular, respondents are asked to rate their life satisfaction using the question “*How satisfied are you with your life in general?*”. This self-evaluation is followed by two anchoring vignettes that allow us to implement the methodology proposed by King et al. (2004) to control for individual heterogeneity in reporting styles. In particular, the following vignettes are used in our analysis:

1. John is 63 years old. His wife died 2 years ago and he still spends a lot of time thinking about her. He has 4 children and 10 grandchildren who visit him regularly. John can make ends meet but has no money for extras such as expensive gifts to his grandchildren. He has had to stop working recently due to heart problems. He gets tired easily. Otherwise, he has no serious health conditions. How satisfied with his life do you think John is?
2. Carry is 72 years old and a widow. Her total after tax income is about € 1,100²¹ per month. She owns the house she lives in and has a large circle of friends. She plays bridge twice a week and goes on vacation regularly with some friends. Lately she has been suffering from arthritis, which makes working in the house and garden painful. How satisfied with his life do you think Carry is?

²¹The value is PPP-adjusted.

Both for the self-evaluation and the vignettes respondents answer using the following 5-point scale: “very dissatisfied”, “dissatisfied”, “neither satisfied, nor dissatisfied”, “satisfied”, “very satisfied”.

As a result, SHARE is a unique data-source to investigate our research question since it (i) enables tight cross countries comparisons by administering in all countries the same internationally-standardized questionnaire and (ii) provides additional information crucial to detect actual differences in life satisfaction filtering out heterogeneity in the response styles used in self-assessments.

Overall, our sample contains 5,606 individuals living in Sweden, Denmark, Germany, The Netherlands, Belgium, France, Spain, Italy, Poland and Czechia. Figure 1 reports the proportion of respondents who rate themselves as either satisfied or very satisfied with their life. As documented in previous studies (Inglehart and Rabier, 1986; Kanhneman et al., 2004), life satisfaction is characterized by clear cross countries heterogeneity. The sharpest differential derives from the comparison between Denmark and Italy. While in the former country the proportion of individuals satisfied with their life is higher than 90 percent, in the latter it is slightly above 60 percent. In general, the picture shows that Denmark has the highest level of life satisfaction, followed by Sweden, The Netherlands and Germany, while Italy is the country with the lowest self-reported satisfaction.

With the exception of Poland and Czechia, these differences are somewhat striking in that they refer to a set of countries comparable in many dimensions, such as quality of life, social and economic inclusion and activism of the welfare state in protecting societies from adverse economic shocks. As anticipated in the introduction, a rationale for this evidence is that individuals are affected by *differential item functioning* (DIF) biases in evaluating themselves.

The nature of the dataset allows us to study the geographical dimension of the heterogeneity in self reporting life satisfaction. Nevertheless, many other factors, such as education, employment and family background, have a strong effect on self evaluations. From a methodological point of view, the hypothesis of dealing with individual-specific reporting styles casts doubts on the reliability of self assessments as a measure for life satisfaction and leaves open the question about how to control

for this source of bias.

The aim of our analysis is to remove individual-specific heterogeneity in reporting styles and then to attach individual self-assessments to a comparable response-scale allowing for meaningful comparisons across individuals. As pointed out before, our estimation strategy relies on the fact that in our sample each respondent is asked to evaluate the same hypothetical situations described in the vignettes reported above.

Analysing the distribution of vignettes evaluations at a pure descriptive level may help understanding how they can be used to purge self assessments from individual heterogeneity in reporting styles. Figure 2 presents the proportion of respondents that rate the persons described in the two vignettes (John and Carry) as satisfied with their life for each country. The figure shows that response scales are not invariant across countries.²² While in Denmark and Germany more than 20 percent of respondents describe John as satisfied with his life, the proportion drops to less than 10 percent for France and Italy. We also find that Spain, Poland and Czechia adopt reporting styles closer to those of Germany, whereas the remaining countries are more in line with France and Italy. Cross countries differences are also found in the assessment of the second vignette. As before, while Danish respondents are more likely to consider Carry as satisfied with her life, the French are more reluctant to use the most extreme labelling of vignette evaluation. Remarkably, while nearly 80 percent of the Danes find Carry satisfied with her life, this probability falls to less than 40 percent for the French and lies between 50 and 60 percent for the remaining countries.

To summarize, we find evidence suggesting that individuals living in different countries may adopt different reporting styles in life satisfaction self-assessments. Relying on such evaluations may then lead to misleading conclusions and emphasize differences that are mainly due to incomparable response scales rather than true differences in life satisfaction. As an example, although Figure 1 shows that the Danes are more likely to describe themselves as satisfied with their life than the

²²However, it should be kept in mind that the evidence reported in these figures does not exclude the explanation that cross-country variability is actually due to uncontrolled differences in terms of other socioeconomic characteristics related to the country of residence.

French, in Figure 2 we find that they are also more likely to use the modalities at the top of the life satisfaction scale when asked to evaluate the well-being of the same hypothetical persons described in the vignettes. Our estimation method will exploit the variability in vignette evaluations to assess to what extent the differences in Figure 1 are genuine or they just reflect differences in the response scales used by respondents.

3 The econometric model

Anchoring vignettes were first introduced by King et al. (2004) for analyzing ordinal survey responses taking into account individual differences in the interpretation of the survey questions. Vignettes are indeed a new tool for enhancing self-report data comparability across individuals. Under the assumption that the extent of the situation described in the vignettes is perceived by respondents in the same way (*vignette equivalence*), variability in vignette evaluations is only due to the different reporting styles adopted. Hence, if the same response style is used for both self-ratings and vignette evaluations (*response consistency*), the additional information provided by vignettes acts as an *anchor* to adjust the self-assessments of different individuals according to a homogenous classification allowing for inter-personal comparisons. Our econometric specification has been introduced by King et al. (2004) and it is usually referred to as the Hopit model. It mainly consists of two components modelling self-assessments and vignette evaluations as standard multinomial ordered variables.

Let us denote with Y_i^* the life satisfaction perceived by individual $i = 1, \dots, n$ and assume that it is the result of the linear specification

$$Y_i^* = X_i\beta + \varepsilon_i; \tag{1}$$

$$\varepsilon_i \sim N(0, 1); \quad \varepsilon_i \perp X_i, \tag{2}$$

where X_i includes observed covariates, β is a set of parameters to estimate and ε_i is a stochastic com-

ponent normally distributed and encompassing unobserved factors relevant for the determination of life satisfaction levels.

Although we do not observe Y_i^* , SHARE questionnaire picks up its discrete counterpart Y_i , which is the answer to the life satisfaction self-assessment and it is recorded as a multinomial ordered variable taking on values 1 ("very dissatisfied"), ..., 5 ("very satisfied"). In particular,

$$Y_i = j \text{ if } \tau_i^{j-1} < Y_i^* \leq \tau_i^j, \quad j = 1, \dots, 5. \quad (3)$$

The thresholds τ_i^j are individual-specific and are given by

$$\begin{aligned} \tau_i^0 &= -\infty; \quad \tau_i^5 = \infty; \\ \tau_i^1 &= X_i \gamma^1; \end{aligned} \quad (4)$$

$$\tau_i^j = \tau_i^{j-1} + \exp(X_i \gamma^j), \quad j = 2, 3, 4. \quad (5)$$

Allowing the thresholds to vary across individuals entails that respondents with the same perceived life satisfaction Y^* may report different self-assessments Y because of different sets of cut-off points τ^j used. This is the main difference with respect to standard ordered probit specifications, which formally state that the same thresholds are utilized by the whole population of interest.

Using self-reports on life satisfaction alone is not enough to separately identify the parameters in β and γ . In fact, if a covariate in X influences both the self-assessment Y^* and the thresholds τ^j , we are not able to distinguish these two effects on the basis of the information conveyed by a single self-evaluation. To achieve this goal we need at least two *distinct* evaluations in which the same response scale is adopted.

In our sample each respondent is asked to answer two vignette questions. We denote with Z_{il}^* , $l = 1, 2$ the variable indicating how the actual level of the domain of interest described in the

vignette l is perceived by respondent i . We assume that

$$\begin{aligned} Z_{il}^* &= \theta_l + \nu_i; \\ \nu_i &\sim N(0, \sigma_v^2), \end{aligned} \tag{6}$$

where θ_l is the actual level of the domain of interest described in the vignette l . The requirement of *vignette equivalence* assumes that the extent of the situation described in the vignettes is perceived by respondents in the same way and formally restricts θ_l to not vary over i .

Although the actual perceived value Z_{il}^* is unobserved, SHARE questionnaire gathers respondent evaluations of vignette questions according to the same 5-point scale used for self-assessments. As a result, we are provided with the multinomial ordered response Z_{il} such that

$$Z_{il} = j \text{ if } \tau_i^{j-1} < Z_{il}^* \leq \tau_i^j, \quad j = 1, \dots, 5. \tag{7}$$

It is worth noting that the same set of thresholds is also found in equation (3) and this follows from the *response consistency* hypothesis claiming that the same reporting styles are used for both self-assessments and vignette evaluations.

In this set-up the specifications modelling self-assessment and vignette ordered responses are connected via the utilization of the same set of thresholds. This amounts to say that the information relevant to estimate equation (1) and equation (6) in the sample should be combined to estimate the common set of parameters showing up in the threshold equations (4) and (5). Following King et al. (2004), the joint estimation is carried out via conditional maximum likelihood and implemented by the STATA module *gllamm*.²³

²³For further details, see Rabe-Hesketh et al. (2004) and Rabe-Hesketh and Skrondal (2008).

4 Results

In our estimates we consider a large number of factors associated with life satisfaction: demographic characteristics (age and gender), socio-economic variables (employment, income and wealth and education), health (number of chronic diseases, arthritis, limitations with mobility, symptoms, ADL, IADL, obesity and having been diagnosed with depression) and social relationships (marital status, family bonds and extra-familiar activities), as well as country dummies.

In Table 3 we present the results for the self-assessment equation of life satisfaction with individual-specific response scales (second column), comparing it with a baseline model not allowing for any threshold variation across respondents (first column). The equation in the latter model is almost identical to an ordered probit model, not taking into account potential differences in reporting styles: in the presence of scale biases across countries or socio-economic groups, the parameter estimates of this model will reflect both true well-being effects and the effects of reporting heterogeneity. The results for the threshold equations are presented in columns 3 to 6. The estimates show that the thresholds significantly depend on a number of variables, such as country dummies, age, education, employment and marital status, several health conditions (chronic diseases and symptoms of diseases, arthritis, mobility problems and having been diagnosed with depression), social activities and wealth. Indeed a formal likelihood-ratio test strongly rejects the model not allowing for response scale variation against the more general model that does allow for correction of the DIF bias.²⁴

We first focus on the impact of different response scales on cross countries variations in self-reported well-being. In the model without scale biases correction, Denmark is the country with the highest reported level of life satisfaction, while Italy is the one with the lowest. As regards the other countries, Sweden comes second in the ranking of self-reported well-being and The Netherlands third, followed by Belgium. Together with Italy, France and Czech Republic are the only countries with a level of life satisfaction that is lower than in Germany. When we correct for the DIF bias, the

²⁴ $\chi_{156}^2 = 1103.177$, $p - value = 0.000$.

ranking of countries changes dramatically. The most striking result is that now the life satisfaction of the Danes is not significantly different from that of the Italians.²⁵ The happiest country becomes the Netherlands, followed by Sweden and France. Czech Republic, on the other hand, reports the lowest level of life satisfaction. The remaining countries do not significantly differ from Germany.

The parameter estimates for the other variables are sensible and consistent with the literature. The relation between age and life-satisfaction is non-linear: life satisfaction increases up to age 85 and then gradually decreases. Note that the age-profile correcting for the DIF bias is similar to that found with the model not accounting for threshold variation across respondents but it is steeper, pointing to more marked age differences. However, with cross-sectional data we cannot distinguish between age, cohort and time effects. Women seem to be happier than men and married individuals are more satisfied with their lives than unmarried ones, whose level of self-reported well-being is not significantly different from that of divorced and widowers. As regards family bonds, parents who have a child with whom they have less than weekly contacts²⁶ are the most dissatisfied with their life. In other words, it is better not having any child than having a child and not being in contact with her. Having a living parent or a grandchild is significantly and positively correlated with life satisfaction only if the contacts with her are frequent (at least once a week).

Not surprisingly, health problems are associated with lower levels of self-reported well-being: in particular, reporting at least two symptoms of diseases, limitations with mobility, ADL, IADL and having being diagnosed with affective or emotional disorders, all are significantly and negatively related to life satisfaction.

We also classify individuals by their employment status as employed (either employee or self-employed), retired or out of the labour force (that includes homemakers who never worked and the disabled). Our results show that the retired are worse off than the employed but are more satisfied with their life than those out of the labour force. For a given employment status, taking part to social activities, such as voluntary or charity work, caring for a sick person, educational courses and

²⁵ $\chi_1^2 = 1.98$, $p - value = 0.159$.

²⁶ Any kind of contact, either in person, by phone or mail.

social clubs, substantially increases the well-being of the over 50.

The overall satisfaction of the elderly is positively and significantly correlated with wealth, while income does not seem to play any role. Our explanation for this result is that what really matters for the well-being of the elderly is the saving they have accumulated throughout their whole life and not current income. Even controlling for income and wealth, the less educated are the most dissatisfied with their life.

Note that one might argue that there may be problems of reverse causality: for example, individuals who are satisfied with their life may be more likely to get married or to have success in life and accumulate more wealth. However, we are not interested in precisely establishing causality. Our aim is to identify the factors with which the well-being of the elderly is associated and identify whether these relations are genuine or they just reflect differences in response scales.

In Table 4 we estimate the same specifications as before but introducing in the model two additional sets of control variables related to the political and religious background of the respondent.²⁷ The results show that life satisfaction generally increases with the frequency of praying, the most satisfied being the elderly who pray more than once a day. This evidence is consistent with the findings of Clark and Lelkes (2004), who also show that people become happier the more often they attend church and the more often they pray. However, there seems to be no difference in self-reported life satisfaction between those who never pray and those who pray less than once a week. As for politics, even if we control for income, wealth and occupational status, life satisfaction increases when moving from left to right. The estimates on the other variables remain qualitatively unchanged.

4.1 Counterfactuals

We now present results of counterfactuals simulations to assess the relevance of the DIF bias across countries. In particular, we simulate the number of persons in a country who would report to be

²⁷France is not included in the sample because in this country the questions about religion and politics were not asked for privacy reasons.

satisfied with their life if they were asked to evaluate themselves according to the scale of response estimated for another country. In Figure 3 we compare the proportions of individuals satisfied with their life if all the respondents used the Danish and Italian response scales respectively. The simulations show that using the Danish thresholds more than 95 percent of respondents in all countries would rate themselves as satisfied with their own life. This means that, given the same true level of life satisfaction, the Danes are more likely to rank themselves high in the 5-point response scale for life satisfaction. The picture considerably changes when using the Italian response scales. Now cross countries heterogeneity is much more evident: in Poland the proportion of individuals who are satisfied with their life drops from about 97 percent to 60 percent and similar differences can be seen in all countries. A similar conclusion can be reached by looking at Figure 4. On the basis of the estimates of the parameters in the threshold equations, we are able to associate to each individual the cut-off points used in equation (3). Figure 4 shows the medians of the individual-specific thresholds by country. We report medians rather than means because they are more robust to the presence of outliers.²⁸ The fact that the third threshold, which is the one that determines whether someone is satisfied with her own life, is lower for Denmark than for the other countries implies that the Danes will be more likely to report that they are satisfied *ceteris paribus*. Thus, variations in response scales explain a large part of the differences found in raw data.

5 Conclusion

In analyzing data of a Eurobarometer survey, Kahneman et al. (2004) notice that there is an implausible large difference in the self-reported level of life satisfaction between Denmark and France. Indeed, while 64 percent of the Danish interviewees affirm to be “very satisfied” with their lives, only 16 percent of the French did so. As suggested by authors, the puzzling differences in the level of self-reported life satisfaction across seemingly similar countries can be caused by the fact that the scales adopted by respondents to rate their satisfaction can be affected by individual-specific

²⁸ Anyway, we take variability into account in the estimates of the Hopit model.

biases.

Using the 2006-2007 wave of SHARE, a cross-sectional dataset containing demographic, physical, social and economic information of individuals aged over 50 living in 10 European countries, we estimate and measure cross countries differences in individuals' scales of life satisfaction. In particular, we compare estimates from an Ordered Probit model in which scales are assumed to be constant across country with those from a Hopit model in which vignettes are used to correct for country specific scale biases.

Our results are astonishing. According to the Ordered Probit that does not account for differences in reporting styles, once controlled for economic, demographic, health and social conditions, Danes and Italians result to be the most and the least satisfied with life respectively. However, by correcting for the DIF bias, the ranking across countries dramatically change. The difference in self-reported life satisfaction between Danes and Italians disappears and the Netherlands and Czech Republic respectively replace Denmark and Italy in the ranking of life satisfaction. Interestingly, according to the Ordered Probit specification Danes are significantly more satisfied with their life than the French. On the contrary, an opposite and highly significant result is found when we estimate the Hopit model. This amounts to say that the assumption of invariant response scales across countries is far from being innocuous.

A formal likelihood ratio test strongly rejects the ordered probit not accounting for threshold variation across respondents against the more general Hopit model. We find evidence that the thresholds significantly depend not only on country dummies but also on age, education, employment status, health, social activities and wealth. We also present results from counterfactual simulations to study how the percentage of satisfied individuals in a country change when the response scale of other countries is imposed to its inhabitants. When the Danish scale is used, more than 95 percent of respondents in all countries would rate themselves as satisfied with their own life. The picture considerably changes when using the Italian response scale. Indeed, the proportion of satisfied individuals according to the Italian scale drops significantly in all countries. Thus, variations in

response scales explain a large part of the differences found in raw data.

To our knowledge, this is the first study which applies the vignettes methodology to estimate and measure cross countries differences in the scale of self-reported life satisfaction. Our results inform the empirical literature studying the determinants of life satisfaction (or happiness) about the importance of correcting for the DIF bias to produce more reliable cross countries comparisons.

As a proposal for further research, future waves of SHARE data will allow us to combine the vignettes approach with panel data techniques to control for time variant scales and to investigate to what extent individual reporting styles are stable over time.

References

Alesina, A., R. Di Tella and R. MacCulloch (2004). Inequality and Happiness: are Europeans and Americans Different? *Journal of Public Economics*, 88, 2009–42.

Arendt, J.N. (2005). Income and "Outcomes" for Elderly: Do the Poor Have a Poorer Life? *Social Indicators Research*, 70, 327-347.

Arthaud-day, M. and Near, J. (2005). The Wealth of Nations and the Happiness of Nations: Why "Accounting" Matters. *Social Indicators Research*, 74, 511-548.

Atchley, R.C. (1999). Continuity and adaptation in aging: Creating positive experiences. Baltimore, MD: John Hopkins University Press.

Atchley R. C. and Robinson, J. L. (1982). Attitudes toward Retirement and Distance from the Event. *Research on Aging*, 4, 299-313.

Bago d’Uva, T., van Doorslaer, E., Lindeboom, M., O’Donnell, O. (2008). Does Reporting Heterogeneity Bias the Measurement of Health Disparities? *Health Economics* 17, 351-375

Biblarz, T. J. and Gottainer, G. (2000). Family Structure and Children’s Success: A Comparison of Widowed and Divorced Single-Mother Families. *Journal of Marriage and the Family*, 62, 533-548.

Bierman, A., Fazio, E.M. and Milkie, M.A. (2006). A Multifaceted Approach to the Mental

Health Advantage of the Married: Assessing How Explanations Vary by Outcome Measure and Unmarried Group. *Journal of Family Issues*, 27, 554-582.

Bjornskov, C. (2003). The Happy Few: Cross-Country Evidence on Social Capital and Life Satisfaction. *Kyklos*, 56, 3-16

Blanchflower, D. G. and Oswald A. J. (2004) Well-being over time in Britain and the USA. *Journal of Public Economics*, 88, 1359-86.

Bosse, R., Aldwin, C. M., Levenson, M. R., and Ekerdt, D. J. (1987). Mental Health Differences among Retirees and Workers: Finding from the Normative Aging Study. *Psychology and Aging*, 2, 383-389.

Bruni, L. and Porta, P.L. (2005). *Economics and Happiness. Framing the Analysis*. Oxford: Oxford University Press.

Calvo, E., Haverstick, K. and Sass, S.A. (2007). What Makes Retirees Happier: A Gradual or 'Cold Turkey' Retirement? WP Center for Retirement Research at Boston College n.18.

Chan, Y. and Lee, R. (2006). Network Size, Social Support and Happiness in Later Life: A Comparative Study of Beijing and Hong Kong. *Journal of Happiness Studies*, 7, 87-112.

Charles, K.K. (2004). Is Retirement Depressing? Labor Force Inactivity and Psychological Well-Being in Later Life. In Polachek, S.W. (Ed-) "Accounting for Worker Well-Being" (pp.269-299), *Research in Labor Economics*, vol. 23.

Amsterdam, San Diego and Oxford: Elsevier, JAI.

Charles, S.T., Reynolds, C.A. and Gatz, M. (2001). Age-Related Differences and Change in Positive and Negative Affect over 23 Years. *Journal of Personality and Social Psychology*, 80, 136-151.

Cheng, S.T. and Chan, A.C.M. (2006). Relationship With Others and Life Satisfaction in Later Life: Do Gender and Widowhood Make a Difference? *Journal of Gerontology B Psychological Sciences Social Sciences*, 61, 46-53.

Clark, A. (2003). Unemployment as a Social Norm: Psychological Evidence from Panel Data.

Journal of Labor Economics, 21, 289-322.

Clark, A. and Lelkes, O. (2005) Deliver us from Evil: Religion as Insurance. PSE, Paris, November

Clark, A, Etilè, F., Postel-Vinay, F., Senik, C. and Van der Straeten, K. (2005). Heterogeneity in Reported Well-Being: Evidence from Twelve European Countries. *Economic Journal*, 115, C118-C132.

Clark, A. and Oswald, A. (1994) Unhappiness and Unemployment. *Economic Journal*, 104, 648-659.

Clark, A.. and Oswald, A (1996). Satisfaction and Comparison Income. *Journal of Public Economics*, 61, 359-381.

Clark, A. and Oswald, A. (2002) Well-being in panels. DELTA, mimeo.

Cohen, A. (2002) The importance of spirituality in well-being for Jews and Christians. *Journal of Happiness Studies*, 3, 287-310.

Davis, J.A. (1984) New Money, an Old Man/Lady and 'Two's Company': Subjective Welfare in the NORC General Social Surveys, 1972-1982. *Social Indicators Research*, 15, 319-50.

de Grace, G. R., Joshi, P., Pelletier, R. and Beaupre, C. (1994). Consequences Psychologiques de la Retraite en Fonction du Sexe et du Niveau Occupationnel Anterieur. *Canadian Journal on Aging*, 13, 149-168.

De Neve, K.M and Cooper, H. (1999). The Happy Personality: a Meta-Analysis of 137 Personality Traits of Subjective Wellbeing. *Psychological Bulletin*, 125, 197-229.

Dehejia, R., DeLeire, T. and Luttmer, E.F.P.(2007). Insuring Consumption and Happiness Through Religious Organizations. *Journal of Public Economics*, 91, 259-279.

Di Tella, R. and MacCulloch, R. (2006). Some Uses of Happiness Data in Economics. *Journal of Economic Perspectives*, 20, 25-46.

Di Tella, R., MacCulloch, R. and Oswald, A. (2001). Preferences over Inflation and Unemployment: Evidence from Surveys of Happiness. *American Economic Review*, 91, 335-341.

Di Tella, R., MacCulloch, R. and Oswald, A. (2003). The Macroeconomics of Happiness. *Review of Economics and Statistics*, 85, 809-827.

Diener, E. and Biswas-Diener, R. (2002). Will Money Increase Subjective Well-Being? *Social Indicators Research* 57,119-169.

Diener, E. and Seligman, M.E.P. (2004). Beyond Money: Toward an Economy of Well-Being. *Psychological Science in the Public Interest*, 5, 1-31.

Diener, E. and Suh, E.M. (2000). *Culture and Subjective Well-Being*. Cambridge MA: MIT Press.

Dulin, P.L. and Pachana, N.A. (2005). Older Adult Mood Functioning: Developmental Changes and Factors Associated with Positive and Negative Outcomes." In Clark, A.V. (Ed.) "Mood State and Health" (pp.175-193). Hauppauge, NY: Nova Biomedical Books.

Easterlin, R. (1995) Will raising the incomes of all increase the happiness of all, *Journal of Economic Behaviour and Organization*, 27, 35-47

Easterlin, R. (2001) Income and happiness: towards a unified theory. *Economic Journal*, 111, 465-84

Easterlin, R. (2003). Happiness of Women and Men in Later Life: Nature, Determinants, and Prospects. In Sirgy, M.J., Rahtz, D. and Samli, A.C. (Eds.) "Advances in Quality-of-Life Theory and Research" (pp.13-26), Dordrecht, The Netherlands: Kluwer Academic Publishers.

Easterlin, R. (2005). Diminishing Marginal Utility of Income? Caveat Emptor. *Social Indicators Research*, 70, 243-255.

Ferrer-i-Carbonell, A. (2005) Income and well-being: an empirical analysis of the comparison income effect. *Journal of Public Economics*, 89, 997-1019

Ferrer-i-Carbonell, A. and Frijters, P. (2004) How important is methodology for the estimates of the determinants of happiness. *Economic Journal*, 114, 641-659

Ferrer-i-Carbonell, A. and Gowdy, J. M. (2005) Environmental awareness and happiness. *Rensselaer Working Papers in Economics*, Number 0503, April 2005

Ferrer-i-Carbonell, A. and Van Praag, B. M. S. (2002) The subjective costs of health losses due to chronic diseases. An alternative model for monetary appraisal. *Health Economics*, 11(8), 709-22.

Frey, B.S. and Luechinger, S. (2007). Concepts of Happiness and their Measurement. In Meyerhoff, J. and Schwarze, R. (Eds.) "Jahrbuch Ökologische Ökonomik" (pp. 129-237), Vol. 5: Soziale Nachhaltigkeit. Marburg, DE: Metropolis Verlag.

Frey, B.S. and Stutzer, A. (2002a). Happiness and Economics; How the Economy and Institutions Affect Well-Being. Princeton, NJ.

Frey, B.S. and Stutzer, A. (2002b). What Can Economists Learn from Happiness Research? *Journal of Economic Literature*, 40, 402-435.

Frey, B.S. and Stutzer, A. (2000). Happiness, Economy and Institutions. *Economic Journal*, 110, 918-938.

Gerdtham, U. and Johannesson, M. (2001). The relationship between happiness, health, and socio-economic factors: results based on Swedish microdata. *Journal of Socio-Economics*, 30, 553-557.

George, L.K. (2006). Perceived Quality of Life. In Binstock, R.H. and George, L.K. "Handbook of Aging and the Social Sciences" (pp.320-336), San Diego, CA: Elsevier.

Gove, W.R., Ortega, S.T. and Style, C.B.(1989). The Maturational and Role Perspectives on Aging and Self through the Adult Years: An Empirical Evaluation. *American Journal of Sociology*, 94, 1117-1145.

Graham, C. (2005). The Economics of Happiness: Insights on Globalization from a Novel Approach. *World Economics*, 6, 41-55.

Haller, M. and Hadler, M. (2006). How Social Relations and Structures Can Produce Happiness and Unhappiness: an International Comparative Analysis. *Social Indicators Research*, 75, 169-216.

Helliwell, J. F. (2003). How's Life? Combining Individual and National Variables Explain Subjective Well-Being. *Economic Modelling*, 20, 331-360.

Helliwell, J.F. and Putnam, R.D. (2004). The social context of well-being. *Philosophical Trans-*

actions of the Royal Society of London, Series B, 359,1435-1446

Hilleras, P.K., Agüero-Torres, H. and Winblad, B. (2001a). Factors Influencing Well-Being in the Elderly. *Current Opinion in Psychiatry*, 14, 361-365.

Hilleras, P.K., Jorm, A.F., Herlitz, A.. and Winblad, B. (2001b). Life Satisfaction Among the Very Old: A Survey on a Cognitively Intact Sample Aged 90 Years or Above *International Journal of Aging and Human Development*, 52, 71-90.

Holland, P. W. and Wainer, H. (1993). *Differential item functioning*. Hillsdale, NJ: Erlbaum.

Hughes, M. and Thomas, M.E. (1998). The Continuing Significance of Race Revisited: A Study of Race, Class and Quality of Life in America, 1972-1996. *American Sociological Review*, 63, 785-795.

Inglehart, R., and Klingemann, H.D. (2000). Genes, culture, democracy and happiness. In Diener, E. and Suh, E.M. (Eds.) "Subjective well-being across cultures" (pp. 165–183). Cambridge, MA: MIT Press.

Inglehart, R. and Rabier, J. (1986). Aspirations Adapt to Situations – But Why are the Belgians So Much Happier than the French?. In Andrews, F.M. (Ed.) "Research on the Quality of Life" (pp. 1-56) Ann Arbor, MI: Inter-University Consortium for Political and Social Research.

Jackson, J. S., Chatters, L. M., and Taylor, R. J. (1993). *Aging in Black America*. Newbury Park, CA, Sage.

Jürges, H. (2007). True Health vs Response Styles: Exploring Cross-Country Differences in Self-Reported Health. *Health Economics*, 16, 163-178.

Kahneman, D., Krueger, A.B., Schkade, D., Schwarz, N. and Stone, A. (2003). A Survey Method For Characterizing Daily Life Experience: The Day Reconstruction Method (DRM). Mimeo, Princeton University.

Kahneman, D., Krueger, A.B., Schkade, D., Schwarz, N. and Stone, A. (2004). Towards National Well-Being Accounts. *American Economic Review, Papers and Proceedings*, 94, 429 - 434.

Kapteyn, A., Smith, J.P. and Van Soest, A. (2007). Vignettes and Self-Reports of Work Dis-

ability in the U.S. and the Netherlands. *American Economic Review*, 97, 461-473.

King, G., Murray, C., Salomon, J. and Tandon, A. (2004). Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research. *American Political Science Review*, 98, 567-583.

Kohler, H.P., Behrman, J. R., and Skytthe, A. (2005). Partner + Children = Happiness? The Effects of Partnerships and Fertility on Well-Being. *Population and Development Review*, 31, 407-445.

Kristensen, N. and Johansson, E. (2008). New Evidence on Cross-Country Differences in Job Satisfaction Using Anchoring Vignettes. *Labour Economics*, 15, 96-117.

Kutner, B., Fanshel, D., Togo, A., and Langer, S. (1956). *Five Hundred over Sixty*. Russel Sage Foundation, New York.

Meier, S. and Stutzer, A. (2008). Is Volunteering Rewarding in Itself? *Economica*, 75, 39-59.

Mroczek, D.K. and Kolarz, C.M. (1998). The Effect of Age on Positive and Negative Affect: A Developmental Perspective on Happiness. *Journal of Personality and Social Psychology*, 75, 1333-1149.

Oswald, A. and Powdthavee, N. (2007). Obesity, Unhappiness, and The Challenge of Affluence: Theory and Evidence. *Economic Journal*, 117, F441-454.

Pichler, F (2006). Subjective Quality of Life of Young Europeans. Feeling Happy but Who Knows Why? *Social Indicators Research*, 75, 419-444.

Rabe-Hesketh, S. and Skrondal, A. (2008), *Multilevel and Longitudinal Modeling Using STATA* (2ed.), STATA Press.

Rabe-Hesketh, S., Skrondal, A. and Pickles, A. (2004). *GLLAMM Manual*, Working Paper 160, U.C, Berkeley Division of Biostatistics, Working Paper Series.

Roostgard, T. (2002). Caring for children and older people in Europe – A comparison of European policies and practice. *Policy studies*, 23, 51-68.

Rowe, J.W. and Kahn, R.L. (1998). *Successful Aging*. New York: Pantheon.

- Salomon, J.A., Tandon, A., Murray, C.J.L. (2004). Comparability of Self-Rated Health: Cross Sectional Multi-Country Survey Using Anchoring Vignettes. *British Medical Journal*, 328, 258-261.
- Saris, W.E. (2001). The Relationship Between Income and Satisfaction: The Effect of Measurement Error and Suppressor Variables." *Social Indicators Research*, 53, 117-136.
- Senik, C. (2004). When Information Dominates Comparison: Learning from Russian Subjective Panel Data. *Journal of Public Economics*, 88, 2099-2123.
- Shmotkin, D. (1990). Subjective Well-Being Is a Function of Age and Gender: A Multivariate Look for Differentiated Trends. *Social Indicators Research*, 23, 201-30.
- Schwarze, J. and Härpfer, M. (2007). Are People Inequality Averse, and Do They Prefer Redistribution by the State? Evidence From German Longitudinal Data on Life Satisfaction. *Journal of Socio-Economics*, 36, 233-249.
- Stack, S. and Eshleman, J.R. (1998). Marital Status and Happiness: A 17-nation Study. *Journal of Marriage and the Family*, 60, 527-536.
- Stutzer, A. (2004). The role of income aspirations in individual happiness. *Journal of Economic Behavior and Organization*, 54, 89-109.
- Stutzer, A. and Frey, B. S. (2006). Does marriage make people happy, or do happy people get married? *Journal of Socio-Economics*, 35, 326-347.
- Thoits, P.A. and Hewitt, L.N. (2001). Volunteer Work and Well-Being. *Journal of Health and Social Behavior*, 42,115-131.
- Uchida, Y., Norasakkunit, V. and Shinobu, K. (2004). Cultural Constructions of Happiness: Theory and Empirical Evidence. *Journal of Happiness Studies*, 5, 223-239.
- Van Den Berg, B. and Ferrer-i-Carbonell, A. (2007). Monetary valuation of informal care: the well-being valuation method. *Health Economics*, 16, 1227-1244.
- Van Praag, B. M. S., Frijters, P. and Ferrer-i-Carbonell, A. (2003). The anatomy of subjective well-being. *Journal of Economic Behavior and Organization*, 51, 29-49.
- Van Praag, B. M. S. and B.E. Baarsma (2005). Using Happiness Surveys to Value Intangibles:

the Case of Airport Noise. *Economic Journal*, 115, 224-246.

Waite, L.J. (1995). Does Marriage Matter? *Demography*, 32, 483-508.

Wiatrowski, W. (1993). Factors Affecting Retirement Income. *Monthly Labor Review*, 116, 25-35.

Winkelmann, L. and R. Winkelmann (1998). Why Are the Unemployed so Unhappy? Evidence from Panel Data. *Economica*, 65, 1-15.

Yang, Y. (2008). Social Inequalities in Happiness in the United States, 1972 to 2004: An Age-Period-Cohort Analysis. *American Sociological Review*, 73, 204-226.

A Appendix

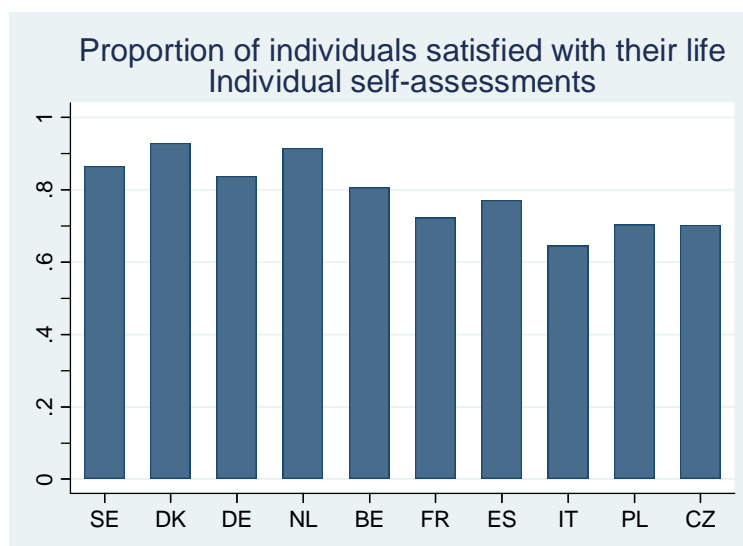


Figure 1: Life-satisfaction self-assessments by country.

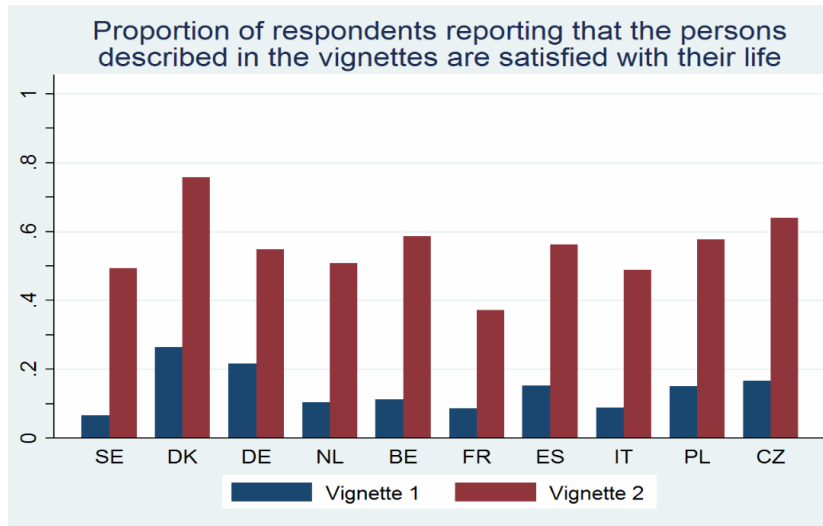


Figure 2: Vignette evaluations by country.

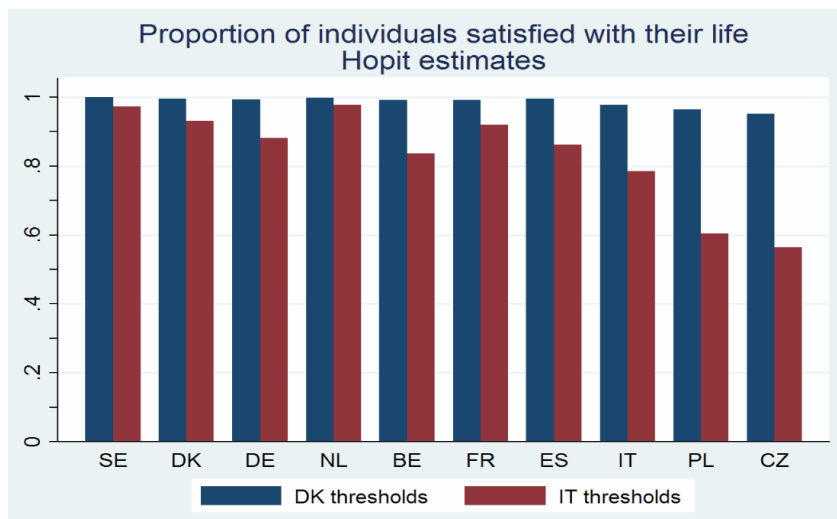
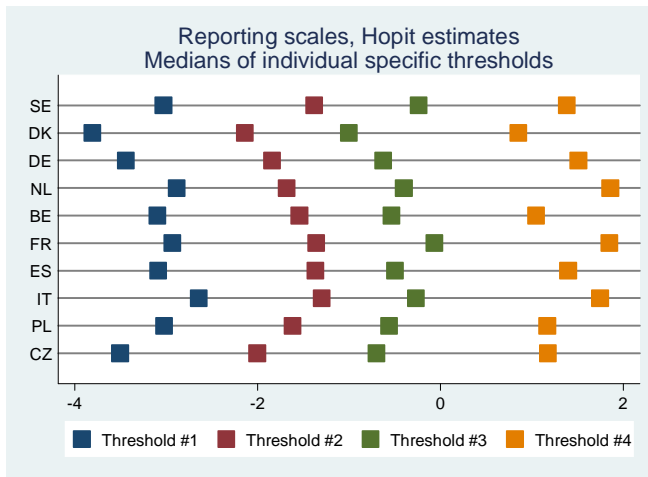


Figure 3: Counterfactuals: Danish and Italian thresholds.



Country	Medians of the thresholds			
	τ_1	τ_2	τ_3	τ_4
SE	-3.03	-1.38	-0.24	1.38
DK	-3.81	-2.14	-1	0.85
DE	-3.44	-1.84	-0.63	1.51
NL	-2.89	-1.69	-0.4	1.86
BE	-3.1	-1.54	-0.54	1.04
FR	-2.93	-1.36	-0.07	1.85
ES	-3.09	-1.37	-0.5	1.4
IT	-2.65	-1.3	-0.27	1.75
PL	-3.02	-1.62	-0.56	1.17
CZ	-3.51	-2.01	-0.7	1.17

Figure 4: Medians of individual-specific thresholds by country

Table 1: Description of the variables included in the regressions.

Variable	Description	Mean	Std. Dev.
<i>Country of residence</i>			
SE	dummy=1 if the person lives in Sweden	0.064	0.245
DK	dummy=1 if the person lives in Denmark	0.149	0.356
NL	dummy=1 if the person lives in The Netherlands	0.062	0.241
BE	dummy=1 if the person lives in Belgium	0.092	0.290
FR	dummy=1 if the person lives in France	0.056	0.229
IT	dummy=1 if the person lives in Italy	0.113	0.317
ES	dummy=1 if the person lives in Spain	0.064	0.244
PL	dummy=1 if the person lives in Poland	0.087	0.281
CZ	dummy=1 if the person lives in Czechia	0.148	0.355
DE	dummy=1 if the person lives in Germany (baseline)		
<i>Demographics</i>			
male	dummy=1 if the person is male	0.447	0.497
age	$(age - 65) / 100$	-0.007	0.097
age2	$[(age - 65) / 100]^2$	0.010	0.012
<i>Education</i>			
low_edu	dummy=1 if the ISCED code is at most 1	0.161	0.368
med_edu	dummy=1 if the ISCED code is either 2 or 3	0.602	0.490
high_edu	dummy=1 if the ISCED code is at least 4 (baseline)		
<i>Employment</i>			
retired	dummy=1 if retired from work	0.528	0.499
not_at_work	dummy=1 if homemaker, unemployed or disabled	0.173	0.378
employed	dummy=1 if employee or self-employed (baseline)		
<i>Civil status</i>			
spouse	dummy=1 if the person has a cohabiting partner	0.777	0.416
widowed	dummy=1 if widow or widower	0.125	0.330
divorced	dummy=1 if the person is divorced	0.065	0.247
single	dummy=1 if the person has never been married (baseline)		
<i>Family background</i>			
child	dummy=1 if the person has at least one living child	0.911	0.285
child_weekly	dummy=1 if the person has almost weekly contacts with her child(ren)	0.881	0.324
parent	dummy=1 if the person has at least one parent alive	0.265	0.441
parent_weekly	dummy=1 if the person has almost weekly contacts with her parent(s)	0.229	0.420
gchild	dummy=1 if the person has at least one grandchild	0.640	0.480
gchild_weekly	dummy=1 the person has almost weekly contacts with her grandchild(ren)	0.161	0.367
(See the next page)			

Variable	Description	Mean	Std. Dev.
<i>Health status</i>			
obese	dummy=1 if the person is obese	0.192	0.394
chronic	dummy=1 if the person has at least two chronic diseases	0.447	0.497
arthritis	dummy=1 if the person is affected by arthritis	0.222	0.415
symptoms	dummy=1 if the person has at least two symptoms of diseases	0.439	0.496
mobility	dummy=1 if the person has mobility limitations	0.312	0.463
adl	dummy=1 if at least one limitation with activities of daily living	0.094	0.292
iadl	dummy=1 if at least one limitation with instrumental adl	0.135	0.341
depression	dummy=1 if diagnosed with affective or emotional disorders	0.085	0.279
<i>Social activities</i>			
volunteer	dummy=1 if done voluntary or charity work	0.138	0.345
caregiving	dummy=1 if cared for a sick person or provided informal help	0.227	0.419
training	dummy=1 if attended an educational course or gone to a social club	0.263	0.441
community	dummy=1 if taken part in political or religious activities	0.123	0.329
no_activities	dummy=1 if the person has not done any social activity (baseline)		
<i>Politics</i>			
left_wing	dummy=1 if the person is left-wing in politics	0.181	0.385
right_wing	dummy=1 if the person is right-wing in politics	0.252	0.434
center	dummy=1 if the person is centrist (baseline)		
<i>Religion</i>			
rel_more_day	dummy=1 if the person prays more than once a day	0.086	0.281
rel_day	dummy=1 if the person prays once a day	0.186	0.389
rel_week	dummy=1 if the person prays at least once a week	0.152	0.359
rel_less_week	dummy=1 if the person prays less than once a week	0.163	0.369
rel_nev	dummy=1 if the person never prays (baseline)		
<i>Financial status</i>			
income	$\operatorname{arcsinh}(\text{household income})^*$, where income is PPP-adjusted and in euros	10.092	1.892
wealth	$\operatorname{arcsinh}(\text{household wealth})^*$, where wealth is PPP-adjusted and in euros	11.647	3.940

* $\operatorname{arcsinh}(x) = \ln(x + \sqrt{x^2 + 1})$

Table 2: Hopit model, determinants of life satisfaction. The first column refers to a baseline Hopit specification not allowing for threshold variation across individuals

	Baseline	Self assessments	γ_1	γ_2	γ_3	γ_4
SE	0.324*** (0.074)	0.414*** (0.104)	0.474*** (0.125)	-0.018 (0.077)	-0.047 (0.060)	-0.291*** (0.053)
DK	0.622*** (0.057)	0.053 (0.079)	-0.285** (0.130)	-0.009 (0.078)	-0.086* (0.047)	-0.177*** (0.036)
NL	0.127* (0.074)	0.460*** (0.112)	0.694*** (0.120)	-0.356*** (0.085)	0.066 (0.057)	0.011 (0.050)
BE	0.102 (0.064)	-0.038 (0.086)	0.341*** (0.111)	-0.007 (0.069)	-0.207*** (0.054)	-0.283*** (0.043)
FR	-0.192** (0.077)	0.287*** (0.107)	0.544*** (0.126)	-0.024 (0.080)	0.065 (0.059)	-0.069 (0.057)
IT	-0.456*** (0.065)	-0.076 (0.087)	0.783*** (0.104)	-0.122* (0.069)	-0.196*** (0.052)	-0.001 (0.046)
ES	-0.002 (0.077)	0.165 (0.105)	0.347*** (0.133)	0.122 (0.081)	-0.333*** (0.068)	-0.055 (0.052)
PL	-0.011 (0.067)	-0.06 (0.088)	0.379*** (0.114)	-0.114 (0.074)	-0.126** (0.054)	-0.163*** (0.045)
CZ	-0.247*** (0.056)	-0.448*** (0.075)	-0.088 (0.113)	-0.051 (0.072)	0.081* (0.043)	-0.119*** (0.037)
male	-0.159*** (0.033)	-0.133*** (0.044)	-0.018 (0.056)	0.05 (0.036)	-0.025 (0.026)	0.001 (0.022)
age	1.175*** (0.275)	1.787*** (0.363)	0.773* (0.469)	-0.558* (0.302)	0.723*** (0.224)	-0.009 (0.186)
age2	-2.619* (1.516)	-4.463** (1.996)	-5.686** (2.686)	2.385 (1.659)	-0.895 (1.206)	1.716* (1.012)
low_edu	-0.074 (0.060)	-0.177** (0.081)	-0.039 (0.094)	-0.041 (0.062)	0.008 (0.049)	-0.035 (0.042)
med_edu	-0.041 (0.040)	-0.02 (0.054)	-0.062 (0.069)	-0.003 (0.044)	0.060* (0.033)	0.039 (0.026)
retired	-0.056 (0.051)	-0.121* (0.069)	0.124 (0.088)	-0.031 (0.055)	-0.094** (0.041)	-0.046 (0.034)
not_at_work	-0.258*** (0.052)	-0.399*** (0.069)	0.05 (0.089)	-0.028 (0.057)	-0.080* (0.042)	-0.082** (0.035)
spouse	0.310*** (0.095)	0.369*** (0.124)	0.396** (0.164)	-0.188* (0.099)	-0.062 (0.072)	0.005 (0.064)
widowed	0.119 (0.104)	0.128 (0.135)	0.345* (0.178)	-0.132 (0.109)	-0.092 (0.079)	-0.067 (0.071)
divorced	-0.028 (0.109)	0.001 (0.141)	0.307 (0.188)	-0.144 (0.114)	-0.066 (0.083)	0.011 (0.074)
child	-0.214** (0.107)	-0.409*** (0.138)	0.081 (0.160)	-0.112 (0.121)	-0.105 (0.083)	-0.04 (0.072)
child_weekly	0.231*** (0.089)	0.399*** (0.113)	-0.257* (0.136)	0.254** (0.107)	0.059 (0.070)	0.029 (0.061)

(See the next page)

	Baseline	Threshold equations				
		Self assessments	γ_1	γ_2	γ_3	γ_4
parent	-0.081 (0.084)	-0.173 (0.108)	0.106 (0.138)	-0.06 (0.091)	0.06 (0.066)	-0.224*** (0.059)
parent_weekly	0.135 (0.086)	0.279** (0.111)	-0.025 (0.140)	0.014 (0.091)	-0.016 (0.067)	0.211*** (0.060)
gchild	0.027 (0.042)	0.066 (0.056)	-0.003 (0.068)	0.031 (0.043)	-0.019 (0.033)	0.021 (0.027)
gchild_weekly	0.098** (0.046)	0.103* (0.061)	-0.093 (0.076)	0.008 (0.049)	0.065* (0.037)	0.03 (0.030)
obese	0.059 (0.039)	0.025 (0.052)	-0.053 (0.066)	-0.008 (0.043)	0.019 (0.031)	0.015 (0.026)
chronic	-0.112*** (0.037)	-0.072 (0.049)	0.145** (0.060)	-0.104*** (0.039)	0.038 (0.029)	0.017 (0.024)
arthritis	-0.071* (0.041)	-0.065 (0.053)	-0.122* (0.067)	0.031 (0.044)	0.098*** (0.032)	-0.03 (0.028)
symptoms	-0.349*** (0.038)	-0.348*** (0.050)	-0.056 (0.063)	0.063 (0.040)	0 (0.030)	-0.038 (0.025)
mobility	-0.144*** (0.042)	-0.128** (0.055)	0.098 (0.069)	0.001 (0.046)	-0.064* (0.034)	-0.027 (0.029)
adl	-0.394*** (0.060)	-0.304*** (0.076)	0.128 (0.090)	-0.013 (0.061)	-0.068 (0.047)	0.01 (0.046)
iadl	-0.158*** (0.054)	-0.174** (0.068)	-0.033 (0.082)	0.001 (0.056)	0.035 (0.041)	-0.052 (0.039)
depression	-0.321*** (0.056)	-0.317*** (0.070)	0.275*** (0.078)	-0.158*** (0.056)	-0.043 (0.043)	-0.069* (0.040)
volunteer	0.151*** (0.049)	0.274*** (0.070)	-0.199** (0.089)	0.114** (0.052)	0.066* (0.039)	0.044 (0.032)
caregiving	0.104*** (0.039)	0.133** (0.053)	-0.074 (0.068)	0.078* (0.042)	-0.034 (0.031)	0.02 (0.025)
training	0.175*** (0.038)	0.262*** (0.053)	-0.095 (0.069)	0.076* (0.042)	-0.011 (0.031)	0.059** (0.025)
community	0.148*** (0.049)	0.079 (0.067)	-0.046 (0.085)	0.006 (0.053)	-0.052 (0.040)	0.019 (0.032)
income	0.017** (0.009)	0.016 (0.011)	0.001 (0.012)	0 (0.008)	-0.004 (0.007)	-0.002 (0.006)
wealth	0.017*** (0.004)	0.022*** (0.005)	-0.004 (0.006)	-0.001 (0.004)	0.008** (0.003)	0.006** (0.003)
constant			-3.610*** (0.286)	0.480*** (0.150)	0.271** (0.111)	0.703*** (0.097)
Log-likelihood	-18855.232	-18303.643				
Sample size	5606	5606				

Note: Standard errors in parentheses, *** p-value<0.01, ** p-value<0.05, * p-value<0.1

Table 3: Hopit model, determinants of life satisfaction when accounting for political and religious background.

	Baseline	Self	Threshold equations			
		assessments	γ_1	γ_2	γ_3	γ_4
SE	0.342*** (0.077)	0.412*** (0.109)	0.502*** (0.136)	-0.022 (0.083)	-0.080 (0.063)	-0.290*** (0.055)
DK	0.612*** (0.060)	0.050 (0.083)	-0.241* (0.139)	-0.017 (0.084)	-0.102** (0.049)	-0.174*** (0.038)
NL	0.106 (0.077)	0.462*** (0.119)	0.745*** (0.128)	-0.378*** (0.090)	0.046 (0.060)	0.023 (0.052)
BE	0.101 (0.068)	-0.073 (0.091)	0.342*** (0.122)	0.004 (0.075)	-0.238*** (0.057)	-0.288*** (0.046)
IT	-0.502*** (0.072)	-0.193** (0.098)	0.859*** (0.121)	-0.154* (0.080)	-0.231*** (0.058)	-0.053 (0.051)
ES	-0.021 (0.086)	0.089 (0.117)	0.410*** (0.155)	0.110 (0.094)	-0.415*** (0.077)	-0.081 (0.058)
PL	-0.093 (0.075)	-0.194* (0.099)	0.526*** (0.131)	-0.170* (0.088)	-0.183*** (0.061)	-0.219*** (0.051)
CZ	-0.244*** (0.059)	-0.463*** (0.079)	-0.096 (0.123)	-0.050 (0.078)	0.080* (0.045)	-0.124*** (0.039)
male	-0.137*** (0.036)	-0.093* (0.049)	-0.065 (0.065)	0.077* (0.042)	-0.024 (0.029)	0.021 (0.024)
age	1.084*** (0.302)	1.548*** (0.399)	0.754 (0.527)	-0.457 (0.339)	0.591** (0.242)	-0.211 (0.202)
age2	-2.772* (1.676)	-3.015 (2.228)	-4.681 (2.983)	1.762 (1.865)	0.431 (1.317)	1.653 (1.101)
low_edu	-0.126* (0.068)	-0.163* (0.091)	-0.077 (0.110)	-0.003 (0.072)	0.040 (0.056)	-0.009 (0.048)
med_edu	-0.063 (0.042)	-0.023 (0.058)	-0.054 (0.077)	0.000 (0.049)	0.043 (0.035)	0.059** (0.027)
retired	-0.057 (0.055)	-0.122* (0.074)	0.146 (0.097)	-0.058 (0.061)	-0.069 (0.044)	-0.051 (0.036)
not_at_work	-0.273*** (0.057)	-0.414*** (0.075)	0.110 (0.099)	-0.063 (0.064)	-0.062 (0.046)	-0.102*** (0.038)
spouse	0.326*** (0.105)	0.406*** (0.136)	0.408** (0.196)	-0.194* (0.117)	-0.030 (0.081)	-0.008 (0.069)
widowed	0.174 (0.116)	0.214 (0.150)	0.360* (0.214)	-0.156 (0.131)	-0.048 (0.090)	-0.053 (0.077)
divorced	0.056 (0.120)	0.100 (0.155)	0.408* (0.218)	-0.222* (0.133)	-0.026 (0.093)	-0.009 (0.079)
child	-0.237** (0.117)	-0.437*** (0.151)	0.119 (0.182)	-0.121 (0.137)	-0.130 (0.092)	-0.043 (0.077)
child_weekly	0.215** (0.098)	0.392*** (0.123)	-0.275* (0.156)	0.237* (0.122)	0.081 (0.077)	0.065 (0.065)
parent	-0.035 (0.092)	-0.125 (0.118)	0.074 (0.154)	-0.042 (0.102)	0.065 (0.072)	-0.203*** (0.064)
parent_weekly	0.099 (0.094)	0.229* (0.121)	-0.054 (0.153)	0.024 (0.099)	-0.005 (0.073)	0.181*** (0.065)

(See the next page)

	Baseline	Self				
		assessments	γ_1	γ_2	γ_3	γ_4
gchild	0.032 (0.045)	0.074 (0.061)	-0.059 (0.075)	0.064 (0.048)	-0.011 (0.036)	0.015 (0.029)
gchild_weekly	0.088* (0.049)	0.152** (0.067)	-0.074 (0.087)	0.001 (0.056)	0.079** (0.039)	0.067** (0.032)
obese	0.052 (0.043)	0.014 (0.057)	-0.010 (0.075)	-0.040 (0.050)	0.026 (0.034)	0.006 (0.029)
chronic	-0.122*** (0.040)	-0.074 (0.053)	0.107 (0.067)	-0.074* (0.043)	0.036 (0.031)	0.019 (0.026)
arthritis	-0.089* (0.046)	-0.079 (0.059)	-0.138* (0.076)	0.037 (0.050)	0.110*** (0.035)	-0.030 (0.031)
symptoms	-0.346*** (0.041)	-0.372*** (0.054)	-0.009 (0.071)	0.036 (0.045)	-0.018 (0.032)	-0.053** (0.027)
mobility	-0.118** (0.046)	-0.102* (0.060)	0.079 (0.076)	0.001 (0.050)	-0.058 (0.036)	-0.011 (0.031)
adl	-0.374*** (0.068)	-0.233*** (0.086)	0.178* (0.106)	0.000 (0.073)	-0.100* (0.053)	0.026 (0.051)
iadl	-0.151** (0.060)	-0.183** (0.076)	-0.069 (0.098)	-0.009 (0.068)	0.075 (0.046)	-0.042 (0.043)
depression	-0.274*** (0.065)	-0.294*** (0.081)	0.220** (0.096)	-0.132* (0.068)	-0.044 (0.051)	-0.077* (0.045)
volunteer	0.147*** (0.053)	0.245*** (0.077)	-0.215** (0.103)	0.103* (0.061)	0.073* (0.043)	0.041 (0.035)
caregiving	0.094** (0.042)	0.122** (0.057)	-0.085 (0.078)	0.086* (0.048)	-0.047 (0.034)	0.025 (0.027)
training	0.184*** (0.041)	0.236*** (0.057)	-0.081 (0.075)	0.062 (0.047)	-0.025 (0.033)	0.048* (0.026)
community	0.104* (0.055)	0.029 (0.076)	-0.002 (0.101)	0.006 (0.064)	-0.080* (0.046)	-0.007 (0.035)
income	0.013 (0.009)	0.014 (0.012)	0.012 (0.014)	-0.005 (0.009)	-0.006 (0.007)	0.002 (0.006)
wealth	0.018*** (0.004)	0.023*** (0.006)	-0.004 (0.007)	0.000 (0.005)	0.007* (0.004)	0.005* (0.003)
left_wing	-0.130*** (0.045)	-0.137** (0.059)	0.000 (0.075)	0.026 (0.048)	-0.053 (0.036)	0.004 (0.030)
right_wing	0.064 (0.040)	0.114** (0.055)	0.026 (0.067)	0.007 (0.044)	-0.012 (0.032)	0.025 (0.026)
rel_more_day	0.138** (0.070)	0.190** (0.094)	-0.071 (0.115)	-0.011 (0.080)	0.007 (0.057)	0.117** (0.047)
rel_day	0.121** (0.052)	0.153** (0.070)	-0.195** (0.097)	0.047 (0.062)	0.091** (0.041)	0.059* (0.034)
rel_week	0.072 (0.053)	0.132* (0.072)	-0.098 (0.095)	-0.003 (0.062)	0.063 (0.043)	0.091*** (0.035)
rel_less_week	0.015 (0.049)	0.009 (0.066)	-0.029 (0.090)	-0.049 (0.059)	0.065* (0.039)	0.020 (0.032)
constant			-3.646*** (0.325)	0.548*** (0.166)	0.254** (0.123)	0.611*** (0.105)
Log-likelihood	-15970.158	-15495.803				
Sample size	4801	4801				

Note: Standard errors in parentheses, *** p-value<0.01, ** p-value<0.05, * p-value<0.1