Age, Health and Life Satisfaction Among Older Europeans

Viola Angelini · Danilo Cavapozzi · Luca Corazzini · Omar Paccagnella

Abstract In this paper we investigate how age affects the self-reported level of life satisfaction among the elderly in Europe. By using a vignette approach, we find evidence that age influences life satisfaction through two counterbalancing channels. On the one hand, controlling for the effects of all other variables, the own perceived level of life satisfaction increases with age. On the other hand, given the same true level of life satisfaction, older respondents are more likely to rank themselves as "dissatisfied" with their life than younger individuals. Detrimental health conditions and physical limitations play a crucial role in explaining scale biases in the reporting style of older individuals.

Keywords Life satisfaction · Ageing · Scale biases · Vignettes · Counterfactuals

V. Angelini (✉)
Department of Economics, Econometrics and Finance, University of Groningen, Groningen,
The Netherlands
e-mail: v.angelini@rug.nl

V. Angelini
Netspar, Tilburg, The Netherlands

D. Cavapozzi · L. Corazzini
Department of Economics and Management "Marco Fanno", University of Padua, Padua, Italy
e-mail: danilo.cavapozzi@unipd.it

L. Corazzini
e-mail: luca.corazzini@unipd.it

L. Corazzini
ISLA, Bocconi University, Milan, Italy

O. Paccagnella
Department of Statistical Sciences, University of Padua, Padua, Italy
e-mail: omar.paccagnella@unipd.it

1 Introduction

In an ageing society, the well-being of the elderly is at the top of the public policy agenda. 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.[1] In economics, measures of the well-being of the elderly have focused mainly on observable and objective economic factors, such as income and wealth. While there is no doubt about the importance of these studies, income and wealth might not be an accurate proxy of individuals' overall welfare, which seems to be also significantly influenced by numerous non-economic factors.

A widely recognized methodology used to collect data on subjective well-being and to test the validity of economic models is to directly ask individuals to self-report their level of life satisfaction on an ordered scale. Given the individuals' answers, social scientists seek to identify the relation between these self-assessed measures and both economic and non-economic factors. As a consequence, economists, psychologists and sociologists have developed a growing interest in the measurement and determinants of life satisfaction (see Frey and Stutzer 2002a, b; van Praag and Ferrer-i-Carbonell 2004; Bruni and Porta 2005; Dolan et al. 2008 for surveys) in later life.

A crucial methodological issue which limits the validity of these studies is that, when asked to self-report their life satisfaction, individuals can adopt different benchmarks or scales in evaluating themselves (van Praag 1971; Winkelmann and Winkelmann 1998; Clark and Oswald 2002; Ferrer-i-Carbonell and Frijters 2004; Senik 2004; Clark and Lelkes 2005). This phenomenon is known as *differential item functioning* (henceforth DIF). In a nutshell, the determinants of the level of life satisfaction reported by respondents might also influence how they perceive the thresholds of the survey question. As a consequence, estimates are biased and do not reflect the genuine relationship between covariates and respondents' self-assessment of life satisfaction. The existence of scale biases might be caused by differences in demographic characteristics, socio-economic conditions (Frey and Stutzer 2002b) and cultural connotations (Uchida et al. 2004; Diener and Suh 2000; Inglehart and Klingemann 2000; Jürges 2007). In particular, when focusing on older people's responses, a reasonable conjecture is that scale biases are associated with their deteriorating health conditions (Rowe and Kahn 1998; Atchley 1999; Hilleras et al. 2001a, b; Easterlin 2003) and physical limitations (Dulin and Pachana 2005). Painful pathologies or mobility limitations might affect self-assessments both *directly* as genuine determinants of life satisfaction and *indirectly* by negatively biasing respondents' mood and the perception of the scale. Similarly, after controlling for health conditions, one might expect age to generate scale biases, per se. For instance, the "role" theory of the ageing process highlights the positive effects on both life satisfaction and respondents' reporting style of virtuous aspects like self-integration, insight, positive psychological traits including satisfaction and self-esteem which are increasing in age (Gove and Shin 1989).

Conducting analysis on panel data represents a potential solution of the scale bias problem. Indeed, when individuals' life satisfaction is followed over time, conventional fixed or random effects (Kapteyn et al. 2007) or latent class techniques (Clark and Lelkes 2005) allow filtering out time-invariant heterogeneity. Unfortunately, in addition to

---

[1] See Roostgard (2002) for a survey on the main institutional changes undertaken in Europe in fields affecting the welfare of aged persons.

objective problems of availability of panel data on individuals' life satisfaction, other important methodological issues question the validity of this solution. For instance, given the psychological nature of life satisfaction, the assumption of individuals' time-invariant scales required by studies based on panel data is arguable. The scale adopted by an individual may vary over time according to her actual mood (Kahneman et al. 2004) or her socio-economic conditions. Moreover, when the analysis focuses on the condition of aged persons, panel data analysis can be strongly influenced by attrition problems which limit the possibility of following the observation over time.

The vignette methodology allows for properly correcting cross sectional data from scale biases. 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. Following this approach, individuals are asked to self-report their level of life satisfaction, along with the assessment, on the same scale, of hypothetical people described in fixed situations or conditions (anchoring vignettes). The idea is that, since the description of the hypothetical person is invariant across respondents, answers to the vignette questions will differ only because of the individual heterogeneity in reporting styles. Therefore, the evaluation of the anchoring vignettes provides sufficient information to correct self-reported life satisfaction from individual specific scale biases.

The vignette methodology has been successfully applied to several domains which can be potentially affected by scale biases. In particular, it has been used in researches based on cross-sectional data and studying political efficacy (King et al. 2004), health (Salomon et al. 2004; Bago d'Uva et al. 2008), work disability (Kapteyn et al. 2007; van Soest et al. 2006) and job satisfaction (Kristensen and Johansson 2008). As far as we know, Angelini et al. (2008) are the first to apply this methodology to a cross-country study of self-reported life satisfaction. Using data from the 2006 wave of the Survey of Health, Ageing and Retirement in Europe (SHARE), they show how the high variability in the self-reported levels of life satisfaction across European countries is mainly caused by differences in the scales and benchmarks adopted by individuals to evaluate themselves. In particular, when differences in reporting styles are not taken into account, once controlled for several dimensions such as economic, demographic, health and social conditions, Danes and Italians result to be the populations with the highest and the lowest levels of life satisfaction respectively. On the other hand, when the vignette methodology is applied to correct for individual-specific biases, 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.

Using the richness of the SHARE vignette data, in this paper we aim at further investigating the relationship between age, health conditions and life satisfaction among the elderly. We find evidence that age affects the self-reported level of life satisfaction in two opposite ways. Controlling for the effects of all other variables, the own perceived level of life satisfaction increases with age; however, at the same time, given the same true level of life satisfaction, older respondents are more likely to rank themselves as "dissatisfied" with their life than younger individuals. The resulting effect is that both in the self-reported data and after controlling for potential differences in individual reporting styles, age variations in the level of life satisfaction are rather small, in particular for respondents younger than 75 years. Moreover, coherently with our initial conjecture of an interplay between age and health conditions, we find that mobility limitations and other pathologies play an important role in explaining scale biases in the reporting style of older individuals.

This paper is organized as follows. In Sect. 2, we specify the econometric model based on the vignettes (the Hopit model), as well as its main properties and statistical assumptions. In Sect. 3, we describe the dataset used in our analysis, along with some descriptive statistics of the variables used in the regressions. Section 4 reports and comments the estimates of the Hopit model. A comparison with an Ordered Probit model in which scales are assumed to be constant is also proposed. In Sect. 5, we discuss some possible limitations of our approach. Section 6 ends the paper, summarizing the main findings.

2 The Econometric Model

We adopt the econometric specification that was first introduced by King et al. (2004) and is usually referred to as the hierarchical ordered probit (Hopit) model. It mainly consists of two elements: the self-assessment component and the vignette component.

Let us denote with Y_i^* the life satisfaction perceived by individual $i = 1, \ldots, n$ and assume that it is a linear function of the observed variables X_i and a normally distributed error term ε_i :

$$
\begin{aligned}
Y_i^* &= X_i\beta + \varepsilon_i; \\
\varepsilon_i | X_i &\sim N(0, 1),
\end{aligned}
\tag{1}
$$

with parameter vector β.

As in the standard ordered probit model, we do not observe directly Y_i^* but only the answers to a question on life satisfaction recorded as an ordered categorical variable, which goes from 1 ("very dissatisfied") to 5 ("very satisfied"). In particular,

$$
Y_i = j \quad \text{if } \tau_i^{j-1} < Y_i^* \leq \tau_i^j, \quad j = 1, \ldots, 5.
\tag{2}
$$

The main difference between the Hopit and the standard ordered probit model is that the thresholds τ_i^j are individual-specific:

$$
\begin{aligned}
\tau_i^0 &= -\infty; \quad \tau_i^5 = \infty \\
\tau_i^1 &= X_i\gamma^1;
\end{aligned}
\tag{3}
$$

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

Unfortunately, using only the self-assessment component does not allow to separately identify the parameters in β and γ. To do so, we need to use the information provided by the vignettes.

Let us denote with $Z_{il}^*, l = 1, 2$ how respondent i perceives the actual level of life satisfaction of the person described in vignette l. We assume that

$$
\begin{aligned}
Z_{il}^* &= \theta_l + v_{il}; \\
v_{il} &\sim N(0, \sigma_v^2),
\end{aligned}
\tag{5}
$$

where θ_l is the actual level of life satisfaction described in the vignette l and v_{il} is a stochastic component assumed to be independent of ε_i. The requirement that θ_l does not vary over i is what is referred to in the literature as the *vignette equivalence* assumption, according to which the situation described in the vignettes is perceived by each respondent in the same way.

Even in this case, we do not observe Z_{il}^* but only the ordered responses to vignette questions on the same 5-item scale used for self-assessments:

$$Z_{il} = j \quad \text{if } \tau_i^{j-1} < Z_{il}^* \le \tau_i^j, \quad j = 1, \ldots, 5. \tag{6}$$

Note that the same set of thresholds is also found in Eq. 2, which implies that we assume the same reporting styles to be used both for the self-assessment and the vignette evaluations. This hypothesis is known as the *response consistency* assumption.

In this set-up using the same set of thresholds in (2) and in (6) allows connecting the self-assessment component and the vignette component. This implies that the information relevant to estimate Eqs. 1 and 5 in the sample should be combined to estimate the common set of parameters in the threshold Eqs. 3 and 4. Following King et al. (2004), the joint estimation is carried out via conditional maximum likelihood and implemented by the STATA module *gllamm*.[2]

3 Data

In this paper we use data from the 2006 wave of the Survey of Health, Ageing and Retirement in Europe (SHARE). SHARE is an interdisciplinary survey on ageing that is run every two years and collects extensive information on health, socioeconomic status and family interactions of individuals aged 50 and over in a host of European countries. We present evidence for eleven countries for which vignette data were collected. These countries range from the North (Sweden and Denmark) through Central Europe (Germany, The Netherlands, Belgium, France, Poland and Czech Republic) to the South (Greece, Italy and Spain). Data are collected by face-to-face, computer-aided personal interviews (CAPI), supplemented by a self-completion paper and pencil questionnaire, which collects self-assessments and vignette evaluations as part of the COMPARE project. We select only those respondents who have answered both to the self-evaluation question and to at least one vignette and our final sample is composed by 7,320 individuals. Table 1 summarizes their main characteristics. Respondents are prevalently females (55%) and 64 years old on average. About 78% of them live with a partner, even though this percentage falls to 54% for individuals older than 75 years. More than 50% are retired from work (86% among 76+ respondents) and only 30% of them are still working. As expected, bad health conditions are more frequent among older age groups; in particular, the percentage of respondents who have one or more limitations with activities of daily living (adl) or instrumental activities of daily living (iadl) among the 76+ is five times higher than that of respondents younger than 55 years.

Our measure of subjective well-being is obtained by the question:

"How satisfied are you with your life in general?"

Respondents' self-assessments are measured according to the scale "Very dissatisfied", "Dissatisfied", "Neither satisfied nor dissatisfied", "Satisfied", "Very satisfied".

In order to investigate the relationship between age and life satisfaction, we divide respondents into five age classes (younger than 56 years, between 56 and 60, between 61 and 65, between 66 and 75 and older than 75 years). However, it is worth noting that working with cross-sectional data does not allow disentangling between age, cohort and period effects. Figure 1 reports the answers provided by respondents to the self-assessment

---

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

Table 1 Description of the variables included in the regressions

Variable	Description	Mean	SD
SE	Dummy = 1 if the person lives in Sweden	0.062	0.241
DK	Dummy = 1 if the person lives in Denmark	0.132	0.338
NL	Dummy = 1 if the person lives in The Netherlands	0.067	0.250
BE	Dummy = 1 if the person lives in Belgium	0.113	0.316
FR	Dummy = 1 if the person lives in France	0.049	0.217
GR	Dummy = 1 if the person lives in Greece	0.074	0.262
IT	Dummy = 1 if the person lives in Italy	0.093	0.290
ES	Dummy = 1 if the person lives in Spain	0.068	0.252
PL	Dummy = 1 if the person lives in Poland	0.071	0.257
CZ	Dummy = 1 if the person lives in Czech Republic	0.119	0.324
DE	Dummy = 1 if the person lives in Germany (baseline)		
Age55	Dummy = 1 if the person is aged 55 or less	0.232	0.422
Age5660	Dummy = 1 if the person is aged 56–60	0.204	0.403
Age6165	Dummy = 1 if the person is aged 61–65	0.171	0.377
Age6675	Dummy = 1 if the person is aged 66–75	0.255	0.436
Age76	Dummy = 1 if the person is aged 76 or more (baseline)		
Male	Dummy = 1 if the person is male	0.447	0.497
Spouse	Dummy = 1 if the person has a cohabiting partner	0.781	0.414
Low_edu	Dummy = 1 if the ISCED code is at most 1	0.270	0.444
Med_edu	Dummy = 1 if the ISCED code is either 2 or 3	0.497	0.500
High_edu	Dummy = 1 if the ISCED code is at least 4 (baseline)		
Retired	Dummy = 1 if retired from work	0.507	0.500
Not_at_work	Dummy = 1 if homemaker, unemployed or disabled	0.192	0.394
Employed	Dummy = 1 if employee or self-employed (baseline)		
adl	Dummy = 1 if at least one limitation with activities of daily living	0.090	0.286
iadl	Dummy = 1 if at least one limitation with instrumental adl	0.140	0.347

question. Self-evaluations of life satisfaction seem to be substantially homogenous across age groups, apart from the very old individuals. The proportion of individuals satisfied with their life is about 80 percent until 75 years, then it falls to 70%.

In order to implement the methodology introduced in the previous section, this self-evaluation is followed by two anchoring vignettes:

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 € 1,100[3] 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
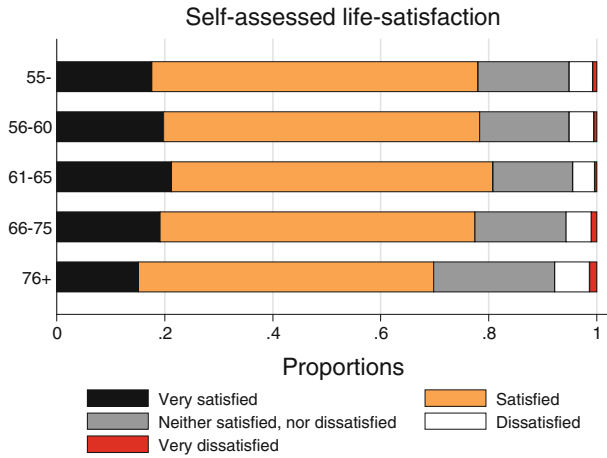
---

[3] The value is PPP-adjusted.

Self-assessed life-satisfaction

Fig. 1 Life satisfaction self-assessments by age group
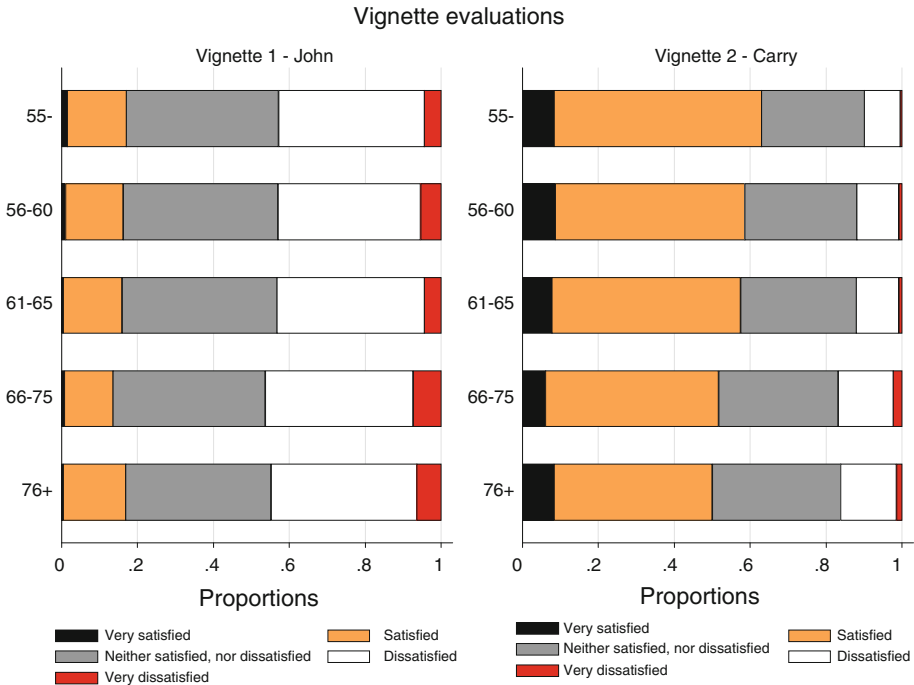
Vignette evaluations

Fig. 2 Vignette evaluations by age group

been suffering from arthritis, which makes working in the house and garden painful. How satisfied with her life do you think Carry is?

Analysing the distribution of the 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. Indeed Fig. 2 presents how respondents rate the level of life

satisfaction of the persons described in the two vignettes (John and Carry) according to the age groups defined before. While vignette 1 shows ratings invariant across age groups, the figure for vignette 2 seems to indicate that response scales cannot be invariant across age groups, because of a decreasing pattern over age. The youngest respondents in our sample are more likely to consider Carry as satisfied or very satisfied with her life, about 10 percent higher than the oldest ones. Note also that there is consistency in the responses to the vignette questions, since in each age group Carry is rated as more satisfied with her life than John.

Even though individuals living in different countries may adopt different reporting styles in life satisfaction self-assessments (Angelini et al. 2008), the same evidence does not appear when conditioning on the age of these respondents. In the next section our estimation method will exploit the variability in vignette evaluations to assess to what extent the small differences in Fig. 1 are genuine or they hide larger differences in the response scales used by respondents.

4 Results

The richness of information collected in SHARE allows us to include a large number of variables in our model specification: demographics (gender and age), education, employment status (being at work, retired or out of work), marital status and physical health status (adl and iadl limitations), as well as country dummies. Age is included in the model through dummy variables generated from the underlying age class. We choose this specification because it is more flexible and it allows us to keep an agnostic view on the shape of the relationship between age and life satisfaction.[4]

We estimate both the Hopit model described in Sect. 2 and a baseline model that does not allow for any threshold variation across respondents, which is an ordered probit model that does not take into account potential differences in reporting styles. Results are showed in Table 2: column 1 reports the estimates of the model not accounting for reporting style heterogeneity, column 2 those of the Hopit model, while the estimates for the threshold equations of the Hopit model are reported in columns 3 to 6. For both specifications, standard errors are robust to arbitrary correlation at the household level.

The main result is that age is positively related to the self-evaluation of the life satisfaction and this is in line with the conclusions of Yang (2008), who shows that the overall levels of life satisfaction increase with age, *ceteris paribus*. However, using the estimates of the Hopit model of Table 2, we find the presence of a trade-off between counterbalancing effects for life satisfaction in later life, as discussed in the Introduction. On the one hand, the own perceived level of life satisfaction increases with age; on the other hand, the individual thresholds that determine whether an individual is satisfied with her life are shifted up with age (that is, given the same true level of life satisfaction, older respondents are more likely to rank themselves as "dissatisfied" with their life than younger individuals). This result can be explained by looking at the Hopit estimates of the self-reported level of life satisfaction (column 2) and the threshold equations (columns 3–6). In our sample, respondents younger than 56 years are the most dissatisfied, followed by respondents in the 56–60 year class, while there are no statistical differences between all age classes of respondents older than 60 years. As regards the estimates for the threshold equations, there are a few significant variables, which also have a negative sign and are

---

[4] We have estimated the same model also with a quadratic in age instead of age dummies but our specification fits the data better according to the AIC criterion and the Pseudo R^2.

Table 2 Hopit model, determinants of life satisfaction

Coefficient	Baseline	Self assessments	Threshold equations			
			τ_1	τ_2	τ_3	τ_4
SE	0.436***	0.567***	0.283**	0.098	−0.088	−0.226***
	(0.072)	(0.095)	(0.124)	(0.077)	(0.058)	(0.050)
DK	0.664***	0.204***	−0.460***	0.097	−0.037	−0.144***
	(0.054)	(0.076)	(0.143)	(0.087)	(0.048)	(0.035)
NL	0.349***	0.665***	0.491***	−0.248***	0.077	0.016
	(0.063)	(0.107)	(0.121)	(0.090)	(0.053)	(0.048)
BE	0.144**	0.016	0.286***	0.005	−0.213***	−0.263***
	(0.057)	(0.074)	(0.110)	(0.072)	(0.047)	(0.039)
FR	−0.127*	0.346***	0.431***	0.007	0.092	−0.041
	(0.070)	(0.105)	(0.128)	(0.084)	(0.058)	(0.054)
GR	−0.363***	−0.335***	0.510***	−0.170**	0.054	−0.564***
	(0.069)	(0.084)	(0.113)	(0.078)	(0.053)	(0.053)
IT	−0.415***	−0.081	0.705***	−0.133*	−0.161***	0.000
	(0.060)	(0.082)	(0.111)	(0.073)	(0.053)	(0.044)
ES	−0.059	0.057	0.513***	−0.013	−0.367***	−0.085*
	(0.073)	(0.091)	(0.131)	(0.086)	(0.066)	(0.047)
PL	−0.114*	−0.144*	0.285**	−0.074	−0.095	−0.143***
	(0.066)	(0.087)	(0.130)	(0.088)	(0.059)	(0.049)
CZ	−0.294***	−0.493***	−0.079	−0.054	0.072	−0.113***
	(0.052)	(0.074)	(0.120)	(0.079)	(0.045)	(0.039)
Age55	−0.241***	−0.459***	−0.211*	0.098	−0.113**	−0.050
	(0.064)	(0.079)	(0.110)	(0.079)	(0.053)	(0.045)
Age5660	−0.093	−0.247***	−0.142	0.077	−0.086*	−0.061
	(0.058)	(0.072)	(0.101)	(0.074)	(0.048)	(0.042)
Age6165	0.008	−0.065	−0.221**	0.133*	−0.051	−0.006
	(0.054)	(0.069)	(0.098)	(0.068)	(0.044)	(0.039)
Age6675	−0.020	0.017	0.118	−0.029	−0.040	−0.021
	(0.049)	(0.061)	(0.080)	(0.060)	(0.038)	(0.036)
Male	−0.117***	−0.102***	0.010	0.005	−0.007	0.013
	(0.025)	(0.034)	(0.052)	(0.036)	(0.023)	(0.019)
Spouse	0.313***	0.357***	0.047	−0.022	0.003	0.044*
	(0.035)	(0.044)	(0.063)	(0.043)	(0.029)	(0.025)
Low_edu	−0.180***	−0.277***	−0.030	−0.015	−0.006	−0.056*
	(0.045)	(0.057)	(0.079)	(0.055)	(0.038)	(0.032)
Med_edu	−0.103***	−0.123***	−0.005	−0.030	0.026	0.019
	(0.035)	(0.046)	(0.066)	(0.044)	(0.030)	(0.023)
Ret	−0.111**	−0.183***	0.047	−0.010	−0.045	−0.064**
	(0.044)	(0.057)	(0.086)	(0.058)	(0.039)	(0.031)
Naw	−0.341***	−0.461***	0.058	−0.040	−0.049	−0.097***
	(0.044)	(0.057)	(0.089)	(0.060)	(0.040)	(0.032)
adl	−0.495***	−0.417***	0.236***	−0.104*	−0.053	0.008
	(0.059)	(0.068)	(0.078)	(0.060)	(0.046)	(0.043)
iadl	−0.390***	−0.381***	0.066	0.005	−0.034	−0.091**
	(0.049)	(0.056)	(0.068)	(0.048)	(0.037)	(0.036)
Constant			−3.837***	0.405***	0.218***	0.730***
			(0.170)	(0.086)	(0.061)	(0.052)

Table 2 continued

Coefficient	Baseline	Self assessments	Threshold equations			
			τ_1	τ_2	τ_3	τ_4
Vignette equations parameters						
θ_1	−1.860*** (0.079)	−2.034*** (0.101)				
θ_2	−0.775*** (0.076)	−0.945*** (0.098)				
σ_v^2	1.099	1.032				
Log-likelihood	−25424.945	−24758.520				
Sample size	7320	7320				

The first column refers to a baseline Hopit specification not allowing for threshold variation across individuals

mostly related to the younger age classes. This means that the threshold for discriminating between dissatisfaction and satisfaction with life is shifted up for older respondents.

Estimates of country dummies confirm the results of Angelini et al. (2008). Correcting for DIF, Dutch and Swedish people are the most satisfied, while Czech respondents are the most dissatisfied. All other individual variables in the self-assessment equation of the Hopit model are significant and consistent with the literature. Women seem to be happier than men and being married is associated with a higher level of self-reported life satisfaction. Life satisfaction increases with the level of education and, as regards occupation, respondents who are at work are more satisfied than those who do not currently work (both because they are retired or because they are out the labour force for other reasons). In other words, people with a high socio-economic status are happier than the others. Not surprisingly, life satisfaction is strongly and negatively related to the presence of health problems. This can also be seen in Fig. 3, which shows that *ceteris paribus* having limitations with instrumental activities of daily living (IADL) is associated with lower levels of life satisfaction for all age groups. The variables that mostly affect the thresholds are the country of residence, age and health variables. The other characteristics included in the regression are significantly correlated with individuals' level of life satisfaction but not with the thresholds.

The findings on the relationship between age and life satisfaction in this section are not in contrast with the descriptive statistics of the previous section. In the Hopit model, parameters on age variables are estimated conditioning on all the other explanatory factors. Figure 1 does not account for these differences. Indeed, older respondents are more likely to suffer from bad health conditions, live alone or be less educated than younger respondents. Summing up all these conditions, older respondents have an overall level of life satisfaction smaller than young respondents. This relationship reverses when we analyse the age effect net of all the other explanatory variables included in the model.

It is worth noting that the baseline model (column 1) is not able to stress such a strong age effect on life satisfaction and many parameter estimates are different in terms of magnitude and significance with respect to the Hopit model. Indeed, the baseline model can be strongly rejected against the model allowing for the correction of DIF by a formal Wald test.[5]

---

[5] $\chi^2_{88} = 902.78$, *p* value = 0.000. We also test the joint significance of all the coefficients but the constant in the threshold equations separately for each threshold and the null hypothesis is always rejected.
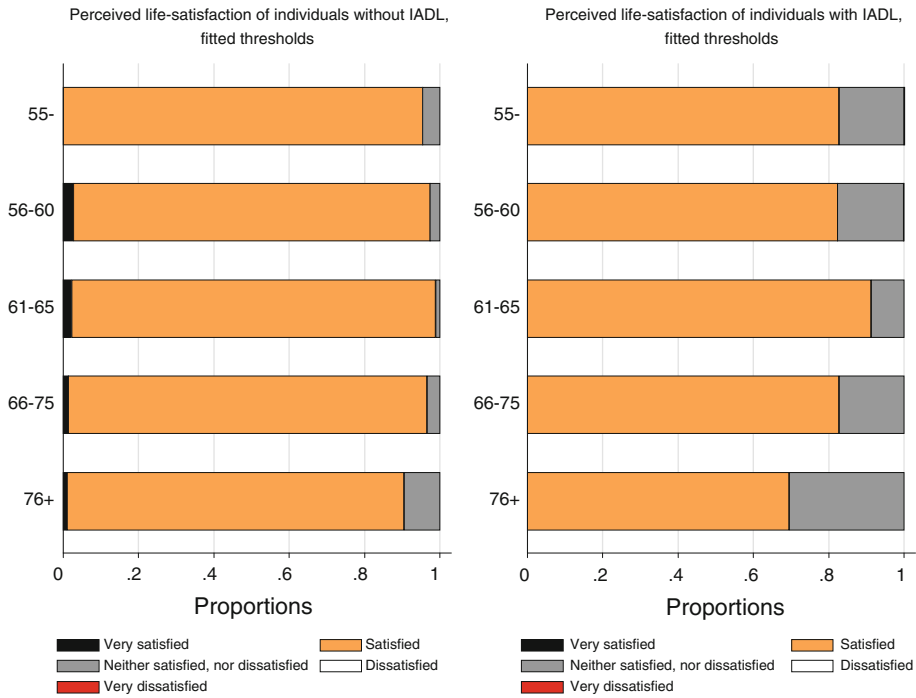
Fig. 3 Model predictions, life-satisfaction and the presence of IADL

4.1 Counterfactuals

In this section we simulate counterfactuals in order to assess the relevance of the DIF bias across age groups. Indeed the main objective of the vignette methodology is to estimate the DIF for each respondent and correct for it. The procedure is very easy. The analyst can choose a benchmark and then compute adjusted distributions of the observed variable for all respondents, using the benchmark scale instead of the respondent's own scale. Usually the benchmark is defined as the scale of one country. However, as we want to compare life satisfaction across socio-economic groups, the flexibility of the vignette approach allows us to define other benchmark scales. In particular, we first use as a benchmark the scales of a specific age class and we calculate how many people in the age class A would report to be satisfied with their life if they used the response scales of respondents in the age class B. Then, we repeat the same exercise using as a benchmark the reporting styles of the persons reporting limitations with IADL.

In Figs. 4 and 5 we compare the proportions of men (women) satisfied with their life if all men (women) used respectively the response scales of men (women) younger than 56 years and the response scale of men (women) older than 75 years. Using the response scales of the youngest group in our sample, the proportion of satisfied individuals becomes larger than when we use the response scales of the oldest respondents in our sample, in particular for women. This means that, given the same true level of life satisfaction, older respondents are more likely to rank themselves low in the 5-point scale for life satisfaction than younger individuals. Figure 6 shows that what is true for older respondents also holds for respondents with IADL limitations.
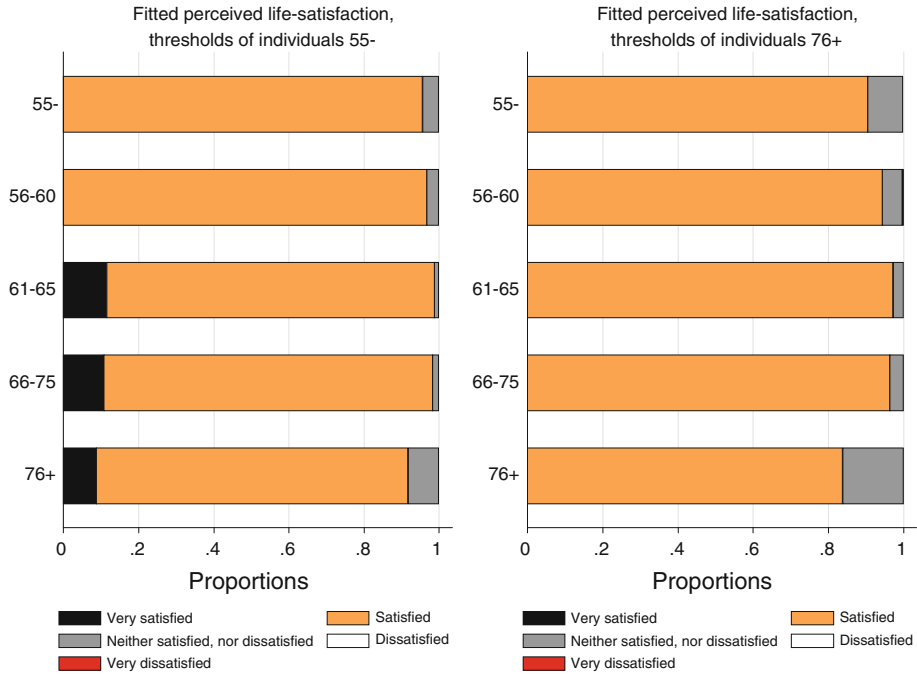
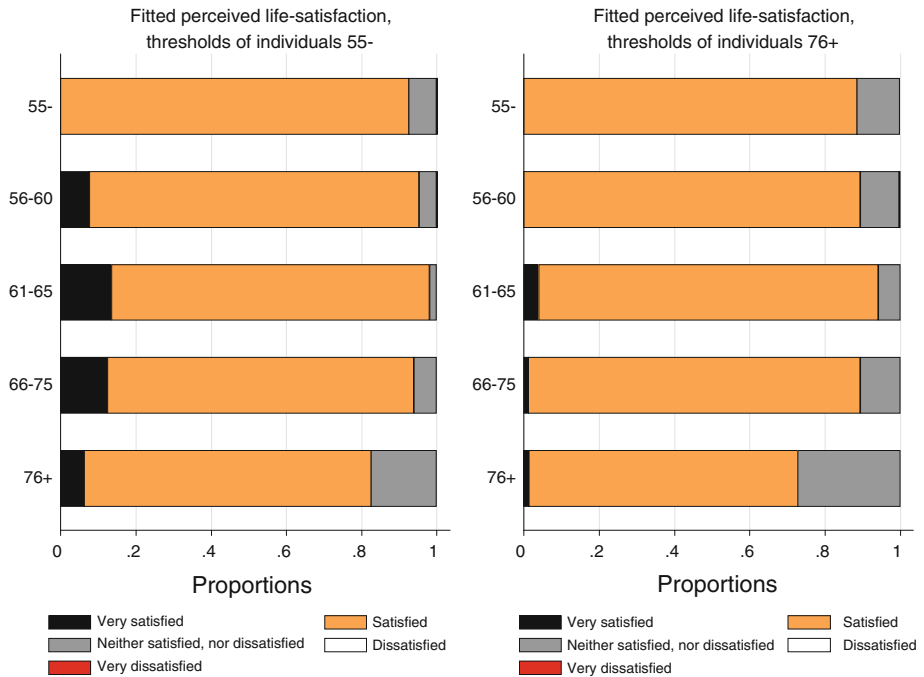**Fig. 4** Counterfactual simulations: age-specific thresholds, men

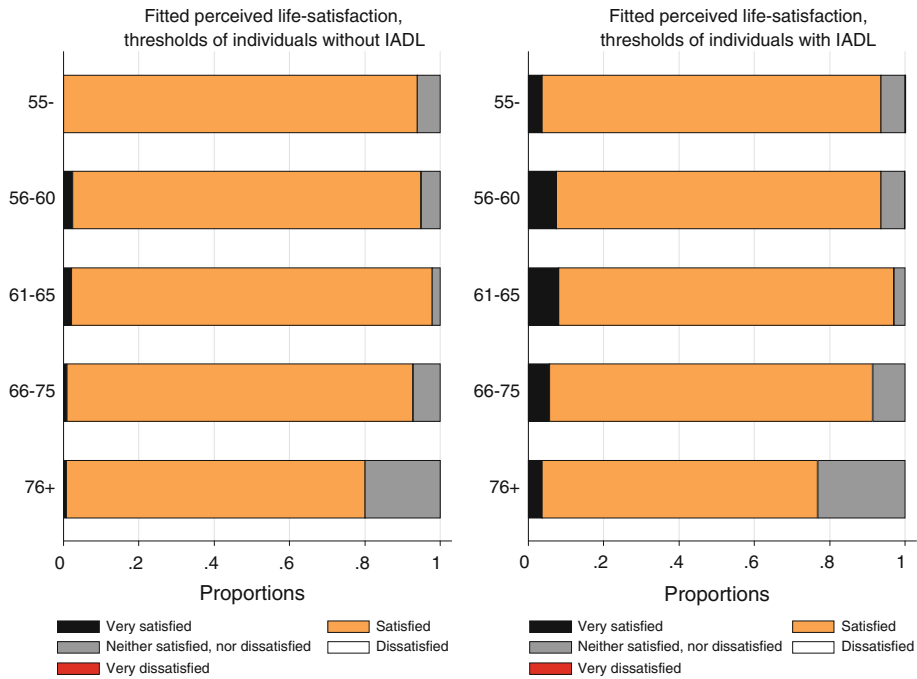**Fig. 5** Counterfactual simulation: age-specific thresholds, women

Fig. 6 Counterfactual simulation: thresholds varying with the presence of IADL

5 Discussion and *Caveats*

Our approach crucially relies on the assumptions of response consistency and vignette equivalence.

As regards vignette equivalence, it is important to note that there is consistency in the responses to the vignette questions, since in each country Carry is always rated as more satisfied with her life than John. Of course, "we still need to be careful of question wording, question order, accurate translation of different items, sampling design, interview length, social background of the interviewer and the respondent" (King et al. 2004, p. 199) because missing cultural differences across subsets of respondents might threaten this approach. Nonetheless, since both the self-assessments and the vignettes are measured with DIF, the fact that the same vignette is not interpreted in the same way in different cultures by different people is not a problem in itself. What would be a problem is if the nature of the DIF differed for the vignettes and the self-assessments, but we can reasonably assume that this difference is close to zero because this approach requires the same person (with the same biases!) to answer both the self-assessment and the vignette questions.

Testing the response consistency assumption requires the availability of an objective measure of the construct of interest, which is hard to find in the context of life satisfaction. In fact, given the multidimensionality of life satisfaction, we should take into account a variety of aspects of the life of individuals to devise a meaningful indicator. Indeed, self-evaluations are widely used in empirical research because they summarize in a single factor all this information, mostly unobserved to researchers. The existence of such objective indicators would question the necessity of collecting self-assessments as well.

Most importantly, the main concern in the literature about the validity of the response consistency assumption (see also Datta Gupta et al. 2010, and Bago d'Uva et al. 2009) is that, in the case of work limiting health problems or drinking behaviour, individuals might have clear incentives to misreport their actual condition due to social norms. As an example, working age individuals who are out of the labour force or unemployed might have an incentive to misreport their disability status to rationalize their labour market condition (justification bias) but not the disability status of the individuals described in the vignettes.

In the case of life satisfaction it is not clear which incentives could lead individuals to adopt different reporting styles when rating themselves or vignettes. The finding in social-psychology that assessments of oneself and others might differ is only due to the fact that respondents have less information about others. However, vignettes provide respondents with all the necessary information to evaluate the persons there described. In addition, the SHARE questionnaire explicitly asks respondents to evaluate vignettes according to their own preferences. In other words, our survey instrument is designed in order to support the validity of the response consistency assumption by requiring the same response scale for self-evaluations and vignette ratings.

6 Conclusions

Ageing is one of the greatest social and economic challenges of the twenty-first century for European societies. Europe is the continent with the highest proportion of people aged 65 or over and the proportion of older individuals is steadily rising. Accounting for this ageing process in Europe, it is easy to understand the growing interest of policy makers and scientists on having good social indicators to assess quality of life in the overall population and in population subgroups, as people aged 65 or over.

A limit of previous studies 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. In this paper, we analyze the relationship between age and life satisfaction in later life. In order to overcome the scale bias, we apply the vignette methodology to data from the Survey of Health, Ageing and Retirement in Europe (SHARE), a unique and innovative multidisciplinary dataset containing a large amount of information on both the economic and non economic conditions of individuals aged 50 and over.

Our main results can be summarized as follows. First, by comparing estimates from a standard Ordered Probit with those from a model in which vignettes are used to correct for the DIF bias (Hopit model), we find a significant effect of scale biases on estimated coefficients. Indeed, a formal Wald test strongly rejects the ordered probit against the more general Hopit model. Moreover, thresholds significantly depend on the explanatory variables used in the regressions. Second, both looking at the raw self-reported data and after controlling for potential differences in individual reporting styles, age variations in the level of life satisfaction are rather small. Larger variations are observed only for the oldest old (76+ individuals). Third, we gain a deeper understanding of the reasons for such age invariance. On the one hand, controlling for the effects of all other variables, the own level of life satisfaction increases with age; on the other hand, the individual thresholds that determine whether a respondent is "satisfied" with her life are shifted up with age; that is, given the same true level of life satisfaction, older respondents are more likely to rank themselves as "dissatisfied" with their life than younger individuals. The two effects work

in the opposite directions. Finally, we highlight health problems and physical limitations as potential sources of scale biases for older individuals. As reasonably expected, detrimental health conditions affect the self-assessments both directly by reducing the level of life satisfaction and indirectly by pessimistically biasing the reporting style of respondents.

References

Angelini, V., Cavapozzi, D., Corazzini, L., & Paccagnella, O. (2008). *Do Danes and Italians rate life satisfaction in the same way? Using vignettes to correct for individual-specific scale biases*. ISLA working paper, 31.

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

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.

Bago d'Uva, T., Lindeboom, M., O'Donnell, O., & van Doorslaer, E. (2009). *Slipping anchor? Testing the vignettes approach to identification and correction of reporting heterogeneity*. HEDG working paper 09/30, Health and Econometrics Data Group, The University of York.

Bruni, L., & Porta, P. (2005). *Economics and happiness. Framing the analysis*. Oxford: Oxford University Press.

Clark, A., & Lelkes, O. (2005). *Deliver us from evil: Religion as insurance*. Paris: PSE.

Clark, A., & Oswald, A. (2002). *Well-being in panels*. Mimeo: DELTA.

Datta Gupta, N., Kristensen, N., & Pozzoli, D. (2010). External validation of the use of vignettes in cross-country health studies. *Economic Modelling, 27*, 854–865.

Diener, E., & Suh, E. M. (2000). Culture and subjective well-being. Cambridge, MA: MIT.

Dolan, P., Peasgood, T., & White, M. (2008). Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. *Journal of Economic Psychology, 29*, 94–122.

Dulin, P. L., & Pachana, N. A. (2005). Older adult mood functioning: Developmental changes and factors associated with positive and negative outcomes. In A. V. Clark (Ed.), *Mood state and health* (pp. 175–193). Hauppauge, NY: Nova Biomedical Books.

Easterlin, R. (2003). Happiness of women and men in later life: Nature, determinants, and prospects. In M. J. Sirgy, D. Rahtz, & A. C. Samli (Eds.), *Advances in quality-of-life theory and research* (pp. 13–26). Dordrecht, The Netherlands: Kluwer.

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

Frey, B. S., & Stutzer, A. (2002a). Happiness and economics: How the economy and institutions affect well-being. Princeton, NJ: Princeton University Press.

Frey, B. S., & Stutzer, A. (2002b). What can economists learn from happiness research? *Journal of Economic Literature, 40*, 402–435.

Gove, W., & Shin, H. (1989). The psychological well-being of divorced and widowed men and women, an empirical analysis. *Journal of Family Issues, 10*, 122–144.

Hilleras, P. K., Aguero-Torres, H., & 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., & 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.

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

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., & Stone, A. (2004). A survey method for characterizing daily life experience: The day reconstruction method. *Science, 306*, 1776–1780.

Kapteyn, A., Smith, J. P., & Van Soest, A. (2007). Vignettes and self-reports of work disability in the U.S. and the Netherlands. *American Economic Review, 97*, 461–473.

King, G., Murray, C., Salomon, J., & Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review, 98*, 567–583.

Kristensen, N., & Johansson, E. (2008). New evidence on cross-country differences in job satisfaction using anchoring vignettes. *Labour Economics, 15*, 96–117.

Rabe-Hesketh, S., & Skrondal, A. (2008). *Multilevel and longitudinal modeling using Stata* (2nd ed.). STATA Press.

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

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., & 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.

Senik, C. (2004). When information dominates comparison: Learning from Russian subjective panel data. *Journal of Public Economics, 88*, 2099–2123.

Uchida, Y., Norasakkunit, V., & Shinobu, K. (2004). Cultural constructions of happiness: Theory and empirical evidence. *Journal of Happiness Studies, 5*, 223–239.

van Praag, B. M. S. (1971). The welfare function of income in Belgium: An empirical investigation. *European Economic Review, 2*, 337–369.

van Praag, B. M. S., & Ferrer-i-Carbonell, A. (2004). *Happines quantified: A satisfaction calculus approach*. Oxford: Oxford University Press.

Van Soest, A., Kapteyn, A., Andreyeva, T., & Smith, J. (2006). *Self-reported disability and reference groups. Wr-409, RAND labour and population*. Working paper series.

Winkelmann, L., & Winkelmann R. (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.