

# Satisfaction with Job and Income Among Older Individuals Across European Countries

Eric Bonsang · Arthur van Soest

Accepted: 26 May 2011 / Published online: 24 June 2011  
© Springer Science+Business Media B.V. 2011

**Abstract** Using data on individuals of age 50 and older from 11 European countries, we analyze two economic aspects of subjective well-being of older Europeans: satisfaction with household income, and job satisfaction. Both have been shown to contribute substantially to overall well-being (satisfaction with life or happiness). We use anchoring vignettes to correct for potential differences in response scales across countries. The results highlight a large variation in self-reported income satisfaction, which is partly explained by differences in response scales. When differences in response scales are eliminated, the cross-country differences are quite well in line with differences in an objective measure of purchasing power of household income. There are common features in the response scale differences in job satisfaction and income satisfaction. French respondents tend to be critical in both assessments, while Danish and Dutch respondents are always on the optimistic end of the spectrum. Moreover, correcting for response scale differences decreases the cross-country association between satisfaction with income and job satisfaction among workers.

**Keywords** Anchoring vignettes · Response scale differences · Ageing

## 1 Introduction

Labour market and living conditions of older individuals have become key policy issues in all European countries. Poverty is more prevalent among the elderly than among other age groups, particularly in several Southern European countries (Tsakloglou 1996). Lack of

---

**Electronic supplementary material** The online version of this article (doi:[10.1007/s11205-011-9879-5](https://doi.org/10.1007/s11205-011-9879-5)) contains supplementary material, which is available to authorized users.

---

E. Bonsang (✉)  
ROA, Maastricht University, Netspar, Meteor, Maastricht, The Netherlands  
e-mail: E.Bonsang@maastrichtuniversity.nl

A. van Soest  
Netspar, Tilburg University, RAND, Tilburg, The Netherlands  
e-mail: A.H.O.vanSoest@uvt.nl

economic resources makes elderly people vulnerable to poor quality of life (Grundy 2006). Downward income mobility is larger among older age groups, particularly among widows and those with an unemployment history, suggesting policies to strengthen the social safety-net and to protect against unemployment and its consequences for economic welfare (Zaidi et al. 2005). Population ageing has led to more pressure on pension and old age benefit systems, and policies aimed at increasing the labour force participation of older individuals are required in order to preserve the sustainability of pension systems and old age social security. To design such policies, it is important to assess the determinants of retirement. Since job satisfaction is an important factor driving retirement decisions (Kosloski et al. 2001), this makes it particularly relevant to study job satisfaction among older workers.

In this paper, using data on individuals of age 50 and older from 11 European countries, we analyze two economic aspects of subjective well-being of older Europeans: satisfaction with household income, and job satisfaction. Both contribute substantially to overall well-being (satisfaction with life or happiness). For example, Ferrer-i-Carbonell and Van Praag (2002) and Van Praag et al. (2003) analyze how satisfaction with life of adult Germans is determined by satisfaction with domains of life (satisfaction with job, finances, housing, health, leisure, and the environment) and find that, together with health satisfaction, job satisfaction and satisfaction with the financial situation are the most important determinants. Similarly large effects of financial and job satisfaction on satisfaction with life are found for the general UK population (not only the older part) by Van Praag and Ferrer-i-Carbonell (2008, p. 91), though they find even larger effects of satisfactions with leisure-use and social life.

Satisfaction with household income has often been studied in the context of household equivalence scales; see, e.g., Van Praag and Van der Sar (1988), Van Praag and Warnaar (1997), Charlier (2002), or Van Praag and Ferrer-i-Carbonell (2008). The economic literature on satisfaction with life emphasizes the role of income (cf., e.g., Clark et al. 2008), but often analyzes the role of income for life satisfaction directly, without considering satisfaction with income (see, for example, Schyns 2002). An exception is the work of Van Praag and co-authors (e.g., Van Praag et al. 2003) who introduced a two-stage model where satisfaction with life is a function of satisfaction with several domains, including satisfaction with income or the financial situation, and where domain specific satisfaction variables are determined by socio-economic characteristics including income. This two stage approach shows the importance of income satisfaction relative to satisfaction with other domains. Van Praag and Ferrer-i-Carbonell (2008) also compare income satisfaction in several countries. Kapteyn et al. (2008) compare income satisfaction of the adult population in the US and the Netherlands. Hsieh (2004) analyzes income satisfaction among older Americans. We are not aware of studies that focus specifically on international comparison of income satisfaction among older individuals.

Job satisfaction has traditionally been studied in sociology and psychology, but has more recently also been shown to provide useful information about economic life that should not be ignored (Hamermesh 1977; Freeman 1978; Borjas 1979; Clark and Oswald 1996). For example, it appears to have predictive value for observable phenomena such as quit rates (Freeman 1978; Clark et al. 1998) or absenteeism (Clegg 1983). The determinants of job satisfaction have been studied extensively for populations of all adult workers; see, for example, Clark (1997), Clark et al. (1998), and Hamermesh (2001). Sousa-Poza and Sousa-Poza (2000) and Kristensen and Johansson (2008) compare job satisfaction and satisfaction with various job characteristics across countries. We do not

know of studies that focus specifically on international comparisons of job satisfaction among older workers.

An important issue underlying the cross-country comparison of self-reported well-being or satisfaction with different domains of life is that individuals from different countries or socio-demographic backgrounds may use different response scales, referred to as differential item functioning (DIF) in the psychology literature (Holland and Wainer 1993). Indeed, if individuals use the same scale, differences in self-reported satisfaction reflect “true” differences across countries or groups of individuals that may also affect behaviour, whereas if differences are due to response scale differences only, they do not influence behaviour and adjustments are required to compare true satisfaction across individuals. Van Praag et al. (2003) use panel data models with (quasi-) fixed effects, capturing persistent differences in response scales.<sup>1</sup> This allows them to identify how changes in satisfaction respond to changes in characteristics but does not help to identify cross-country differences in satisfaction levels that keep response scales constant. Specifically for the latter purpose, King et al. (2004) have proposed to use anchoring vignettes—respondents are asked to evaluate hypothetical situations described in the survey question. This additional information helps to identify interpersonal differences in response scales, even with cross-section data.

Anchoring vignettes have been used to analyze cross-country differences in various subjective measures of well-being, such as political efficacy (King et al. 2004), health (Salomon et al. 2004; Bago d’Úva et al. 2008a, b), life satisfaction (Angelini et al. 2011; Kapteyn et al. 2010), or work disability (Kapteyn et al. 2007). Kapteyn et al. (2008) use anchoring vignettes to compare income satisfaction between the Netherlands and US. They find that the distribution of self-reported income satisfaction differs substantially across countries, but correcting for response scale differences makes the distributions much more similar. Kristensen and Johansson (2008) analyse the job satisfaction across seven European countries using anchoring vignettes and find evidences of cultural differences in reporting job satisfaction. They show that correcting for such differences alters the country ranking.

The aim of this paper is to compare income and job satisfaction of older individuals (50+) across European countries correcting for differences in reporting styles of the respondent by using anchoring vignettes. The results of Bago d’Úva et al. (2008b) and Kapteyn et al. (2007) suggest that differences in reporting styles across countries and socio-economic groups are important for older age groups, though it is not clear whether they are systematically larger or smaller than for younger age groups.

The remainder of this paper is organized as follows. Section 2 presents the econometric model and motivates the use of anchoring vignettes. Section 3 presents the data and descriptive statistics. Estimation results are presented in Sect. 4. Section 5 presents some simulations of counterfactual distributions, showing how income and job satisfaction compare across countries when response scales are kept constant. Section 6 concludes.

## 2 The Model

The methodology of anchoring vignettes to measure subjective ordinal responses taking into account differences in the reporting styles across individuals was first introduced by

<sup>1</sup> They control for unobserved heterogeneity by including the individual means over time of the time-varying explanatory variables as additional controls.

King et al. (2004). We follow their parametric model, the conditional hopit (chopit) model. Define a latent self-satisfaction variable ( $s_i^*$ ) as:

$$s_i^* = X_i\beta + \varepsilon_i, \tag{1}$$

where  $X_i$  is a vector of explanatory variables such as country dummies, gender, years of education, and household income, and  $\beta$  is a vector of parameters. The error term  $\varepsilon_i$  is assumed to be standard normally distributed and independent of  $X_i$ . Reported satisfaction ( $s_i$ ) is a 5-point-scale ordered categorical variable based upon an underlying latent variable  $s_i^*$ :

$$s_i = j \quad \text{if } \tau_i^{j-1} < s_i^* \leq \tau_i^j, \tag{2}$$

For  $j = 1, \dots, 5$  and  $\tau_i^0 = -\infty, \quad \tau_i^5 = +\infty.$

If the thresholds between categories are the same for all respondents ( $\tau_i^j = \tau^j$  for all  $i, j$ ) then this gives the ordered probit model, a standard model for ordered response dependent variables. The main distinguishing feature compared to this standard case is that all thresholds can vary with observed respondent characteristics:

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

where the  $\gamma^j, j = 1, 2, 3, 4,$  are vectors of parameters. Without additional information,  $\gamma^1$  and  $\beta$  are not separately identified. Imposing  $\gamma^1 = 0$  leads to a generalized ordered probit model in which the distances between cut-off points can vary with  $X_i$ ; the exponential function is taken to guarantee that the distances are always positive. We are particularly interested, however, in allowing for non-zero  $\gamma^1$ , since this means that a change in the characteristics leads to a parallel shift in all cut-off points, with the intuition that some respondents use more positive evaluations than other respondents. To identify  $\gamma^1$ , additional information is used in the form of vignette evaluations  $V_i^k (k = 1, \dots, K)$ , where  $K$  is the number of different vignettes evaluated by the respondents. The vignette equivalence assumption implies that there exists a common “true” (objective) actual level of satisfaction  $\theta^k$  underlying the situation described by a given vignette  $k$ ; the vector of all these is denoted by  $\theta = (\theta^1, \dots, \theta^K)$  The vignette evaluations are modelled as follows:

$$\begin{aligned} V_i^{*k} &= \theta^k + v_i^k, \\ V_i^k &= j \quad \text{if } \tau_i^{j-1} < V_i^{*k} \leq \tau_i^j, \end{aligned} \tag{4}$$

where  $V_i^k$  is the evaluation of vignette  $k$  by respondent  $i$ , and the  $v_i^k$  are errors, assumed to be normally distributed with mean 0 and variance  $\sigma_v^2$ , independent of each other,  $\varepsilon_i$ , and  $X_i$ .<sup>2</sup>

The model consisting of Eqs. 1–4 is estimated by maximum likelihood, combining the information in the self-assessments with the information in the vignette evaluations. The likelihood contribution of a given respondent consists of a self-assessment part and a vignette part:

<sup>2</sup> The assumption that the  $v_i^k$  are mutually independent may be too strong. Moreover, unobserved heterogeneity in the thresholds may also lead to correlated vignette evaluations. Sensitivity checks of Kapteyn et al. (2007) suggest that allowing for a richer covariance structure of the errors is a statistically significant improvement but has no effect on the substantive results.

$$L(\beta, \theta, \gamma|s, V) = L_s(\beta, \gamma|s) \times L_V(\theta, \gamma|V), \quad (5)$$

where  $L_s(\beta, \gamma|s)$  is the likelihood component for the self-assessment:

$$L_s(\beta, \gamma|s) = \prod_{i=1}^N \prod_{j=1}^4 \left[ \Phi(\tau_i^j | X_i \beta) - \Phi((\tau_i^{j-1} | X_i \beta)) \right]^{I(s_i=j)}, \quad (6)$$

and  $L_V(\theta, \gamma|V)$  is the likelihood component for the vignette part:

$$L_V(\theta, \gamma|V) = \prod_{i=1}^N \prod_{k=1}^K \prod_{j=1}^4 \left[ \Phi(\tau_i^j | \theta^k, \sigma_v^2) - \Phi(\tau_i^{j-1} | \theta^k, \sigma_v^2) \right]^{I(V_i^{k=j})} \quad (7)$$

The parameters  $\gamma = (\gamma^1, \dots, \gamma^4)$  drive both components of the likelihood contributions, which is why the additional information in the vignette evaluations helps for identification. The main identifying assumptions in this model are twofold. The first is “response consistency:” a given respondent uses the same scales  $\tau_i^j$  for self-reports and vignettes. King et al. (2004) and Van Soest et al. (2011) found support for this hypothesis for vignettes on vision and drinking behaviour, by comparing vignette corrected self-reports and more objective measures. The second assumption is called “vignette equivalence”: there should be no systematic differences in the interpretation of a given vignette between respondents with different characteristics  $X_i$  (so that  $V_i^{*k}$  does not vary with  $X_i$ ).

### 3 Data and Descriptive Statistics

The empirical analysis is based on data from the COMPARE sample which is part of the second wave (2006–2007) of the Survey of Health, Ageing and Retirement in Europe (SHARE). SHARE includes rich information about health, employment, financial situation, family contacts, and social activities of a representative sample of the 50+ populations in a number of European countries (Börsch-Supan et al. 2005, 2008). The COMPARE sample consists of random subsamples of the complete SHARE samples in 11 countries. Respondents in these subsamples did the complete face to face SHARE interview and then completed a drop-off questionnaire with self-assessed satisfaction with various domains of life and with vignette evaluations for the same domains; see Van Soest (2008). SHARE respondents in the other subsamples got a completely different drop-off questionnaire. Response rates to the main survey and the drop-off were similar for the COMPARE sample and the remaining SHARE sample. The COMPARE sample includes about 7,000 individuals aged 50+ from eleven European countries: Belgium, Czech Republic, Denmark, France, Germany, Greece, Italy, the Netherlands, Poland, Spain, and Sweden.

#### 3.1 Income Satisfaction and Anchoring Vignettes

Objective measures of economic poverty across countries are typically based upon household income or household consumption expenditures corrected for purchasing power differences and differences in household composition. Such measures, however, may provide only a partial measure of poverty, since whether people can make ends meet may also depend on other factors such as access to cheap housing, availability of help from family, friends, or neighbours, or the availability of free public goods and services such as health care. A more general assessment of living standard is the answer to the income satisfaction question:

**Table 1** Distribution of reported own income satisfaction by country (in %)

	Very dissatisfied (%)	Dissatisfied (%)	Nor satisfied, neither dissatisfied (%)	Satisfied (%)	Very satisfied (%)
Belgium	3	16	25	47	8
Czech Republic	5	22	39	31	3
Denmark	1	3	15	58	22
France	5	22	40	30	4
Germany	3	12	25	52	9
Greece	17	17	37	20	9
Italy	5	16	33	42	4
Netherlands	1	6	16	60	17
Poland	14	31	32	21	3
Spain	6	24	28	39	3
Sweden	1	6	28	47	17
Total	5	15	28	43	9

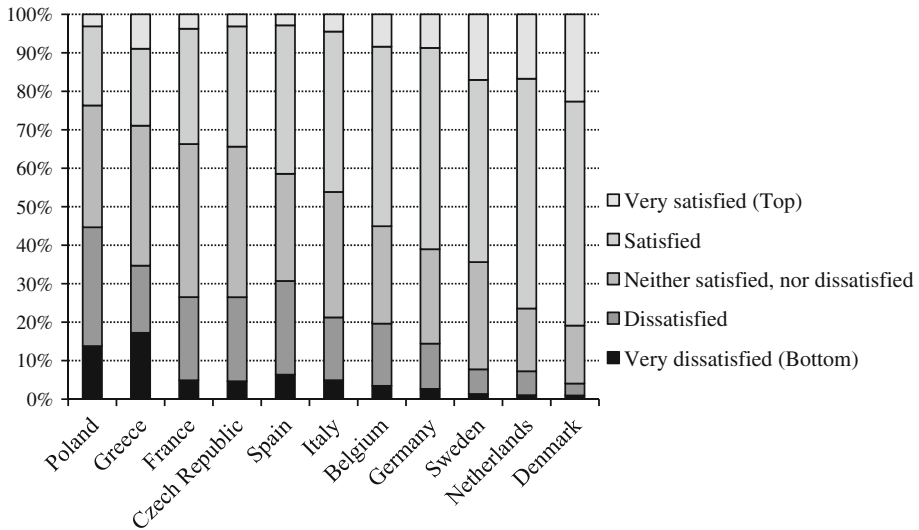
*How satisfied are you with the total income of your household?*

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

The distribution of income satisfaction among the aged 50+ individuals across countries is presented in Table 1. The ranking of the countries varies with the chosen cut-off point. For example, the percentage of satisfied/very satisfied individuals with their income is higher in Spain than in France, but the percentage of individuals being very dissatisfied or dissatisfied individuals is slightly lower in France than in Spain.

To compare the complete income satisfaction distributions and investigate whether an unambiguous ranking across subsets of countries can be obtained, Fig. 1 is presented. It is based upon the numbers in Table 1 and compares the cumulative distribution of reported satisfaction with income across countries by stacking percentages of each outcome. For example, the left hand bars indicate that in Poland, 14% are very dissatisfied, 45% are very dissatisfied or dissatisfied, 77% are at very dissatisfied, dissatisfied, or “neither satisfied nor dissatisfied,” etc. The countries are ranked on the basis of the latter percentages: Poland has the largest percentage at most “neither satisfied nor dissatisfied,” so that Polish respondents report the worst income satisfaction if we set the cut-off between “neither satisfied nor dissatisfied” and “satisfied”. The graph shows, however, that Poland does worst whichever cut-off we use. For example, the percentage very dissatisfied or dissatisfied is higher in Poland (45%) than in any other country. In other words, reported income satisfaction is unambiguously worse in Poland than in all other countries. Such an unambiguous ranking of pairs of countries is not always possible. For example, if the cut-off is put between satisfied and “neither satisfied nor dissatisfied,” Spain does better than France or the Czech Republic, but this reverses if the cut-off is between dissatisfied and “neither satisfied nor dissatisfied.” The figure also shows that Denmark, the Netherlands, and Sweden unambiguously rank first, second and third, respectively, followed by Germany and Belgium.

Figure 2 compares income satisfaction and equivalent monthly household income by country, using the modified OECD equivalence scale  $[1 + 0.5*(adult-1) + 0.3*child]$ , where *adult* is the number of adult (15 years and older) in the household and *child* is the



**Fig. 1** Distribution of reported income satisfaction by country

number of children (at most 14 years old)].<sup>3</sup> Like Table 1, this figure is based upon reported income satisfaction, and therefore does not take into account the fact that individuals from different countries may use different response scales. The horizontal axis gives the country-specific mean of equivalent monthly net household income corrected for purchasing power parity (PPP) differences, while the vertical axis gives the percentage of individuals who are satisfied or very satisfied with their income. The figure suggests a strong positive (and linear) relationship between income and income satisfaction, except that France does not seem to fit this relationship. While France has quite high household income, it performs poorly in terms of income satisfaction.

While the subjective income satisfaction measure has the advantage of encompassing many aspects of economic well-being, it has the drawback that it may suffer from differential item functioning (DIF): individuals in different countries may use different response scales and give different answers although they are economically equally well off. Vignettes describing hypothetical people in given economic circumstances are used in order to correct for these response scale differences. In the COMPARE sample, the vignette questions about income satisfaction are the following:

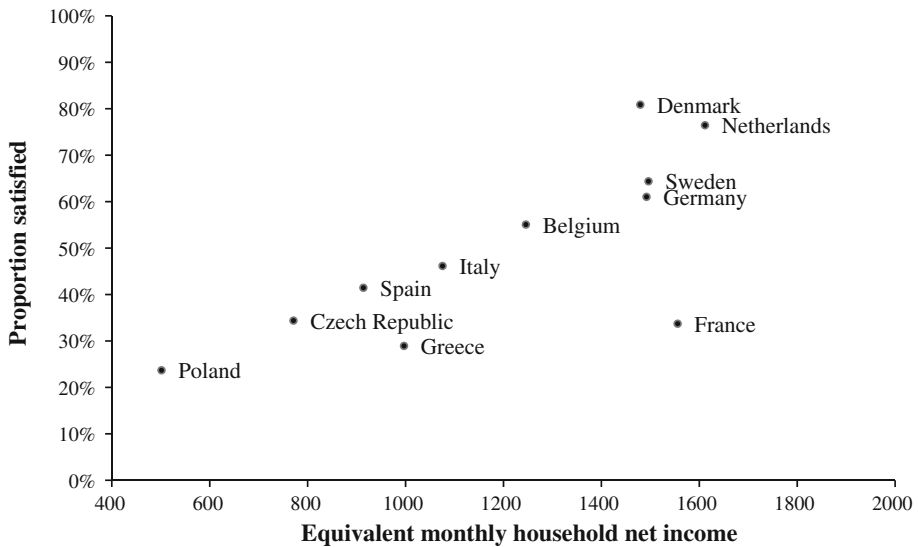
Vignette 1: *Jim is married and has two children; the total after tax household income of his family is €1,500 per month. How satisfied do you think Jim is with the total income of his household?*

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

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

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

<sup>3</sup> The equivalence scales are used in Figs. 2 and 4 only, and we therefore chose to use a simple equivalence scale common to all countries. Of course there are many alternative equivalence scales, including country specific ones, as in, for example, Van Praag and van der Sar (1988).



**Fig. 2** Household income (PPP corrected) and reported income satisfaction of the 50+ across COMPARE countries

The amounts used for net household income<sup>4</sup> in the above vignettes, i.e., €1,500 and €3,000, are the amounts used in the vignette questions in France, Belgium and the Netherlands in which purchasing power of one euro was almost identical. In other countries, PPP adjusted amounts were used in local currencies.<sup>5</sup> The underlying assumption here, necessary for vignette equivalence, is that the living standard that income satisfaction is trying to measure is not affected by the distribution of income in the country of residence. This distribution may affect the answers to the income satisfaction question, but only because it changes the social norms and therefore the response scales, not because it makes someone genuinely better or worse off.<sup>6</sup> The chosen amounts (€1,500 and €3,000) place vignettes 1 and 2 between the 20th and 25th and between the 70th and 75th percentiles of the actual equalized income distribution pooled over all countries. Because of the large cross-country differences in real incomes, the country specific positions vary from the lowest to the highest decile.

Tables 2 and 3 display the distribution of responses to the two vignette questions by country. As expected, the income satisfaction assigned to Vignette 1 is always much lower than for Vignette 2. For both vignettes, there are substantial differences across countries,

<sup>4</sup> We ignore here that income tax is collected in different ways across countries (In France, for example, tax is not withheld on earnings like in most other countries, but paid afterwards).

<sup>5</sup> The amounts in vignette 1 were CK 24,000 in the Czech Republic, DKK 14,200 in Denmark, €1,550 in Germany, €1,200 in Greece, €1,450 in Italy, PZ 3,300 in Poland, €1,300 in Spain and SK 15,400 in Sweden. The amounts in vignette 2 were always twice as high. As pointed out by a referee, the different degrees of rounding might have effects on the responses, but we do not think this is a major issue.

<sup>6</sup> Kapteyn et al. (2008) make the opposite assumption that the living standard is purely relative, and therefore use vignettes with multiples of country specific median incomes. Which assumption is better seems to depend on the interpretation of the living standard concept one is trying to measure.



**Table 2** Distribution of reported income satisfaction Vignette 1 by country (in %)

	Very dissatisfied (%)	Dissatisfied (%)	Nor satisfied, neither dissatisfied (%)	Satisfied (%)	Very satisfied (%)
Belgium	9	42	30	18	1
Czech Republic	7	29	33	26	5
Denmark	12	48	24	14	2
France	12	48	28	11	1
Germany	5	41	36	18	1
Greece	34	54	12	0	0
Italy	17	37	27	16	4
Netherlands	3	40	38	18	0
Poland	2	13	24	50	11
Spain	13	39	17	23	8
Sweden	13	53	23	10	1
Total	11	40	28	19	3

pointing at systematic differences in response styles across European countries. For example, the low-income vignette in Table 2 is rated as satisfactory or very satisfactory by about 61% of the older individuals in Poland, by only 12% in France, 11% in Sweden and by no one in Greece. The high-income vignette in Table 3 is rated as “very satisfied” by 52% of older individuals in Poland, compared to only 14% in Greece.

### 3.2 Job satisfaction and anchoring vignettes

Job satisfaction is measured in the COMPARE survey by a single satisfaction question asked to all respondents (ages 50 and over):

*How satisfied are you with your daily activities (for example, your job, if you work)?  
Very dissatisfied/Dissatisfied/Neither satisfied, nor dissatisfied/Satisfied/Very satisfied*

For this paper, we only consider the responses of 50–64 year old respondents who do paid work; satisfaction with other daily activities is beyond the scope of the current study. Table 4 presents the frequency distributions in each country. On average, older workers are satisfied with their job: 80% of the workers in the total sample report either “satisfied” or “very satisfied.” The differences across countries are substantial, however.

Figure 3, constructed in the same way as Fig. 1, presents the cumulative distribution of job satisfaction by country. Again, Denmark outperforms all other countries, followed by Sweden and the Netherlands. At the other end of the country ranking, we find Greece, France, and the Czech Republic. Interestingly, the ranking of Poland depends crucially on the cut-off point: looking at the proportion of satisfied or very satisfied individuals, Poland does quite well and ranks fourth, but Poland is also the country with the lowest proportion of very satisfied workers.

This cross-country ranking in job satisfaction is largely consistent with the international comparisons including younger workers of Sousa-Poza and Sousa-Poza (2000) based on data on Work Orientations from the 1997 International Social Survey Program (ISSP) and

**Table 3** Distribution of reported income satisfaction Vignette 2 by country (in %)

	Very dissatisfied (%)	Dissatisfied (%)	Nor satisfied, neither dissatisfied (%)	Satisfied (%)	Very satisfied (%)
Belgium	2	6	13	43	35
Czech Republic	0	2	5	46	46
Denmark	0	4	17	57	22
France	1	9	20	52	17
Germany	5	2	10	48	35
Greece	1	11	30	43	14
Italy	4	4	15	57	20
Netherlands	0	0	6	50	43
Poland	1	4	4	38	52
Spain	1	5	15	49	30
Sweden	1	4	18	52	26
Total	2	4	13	49	32

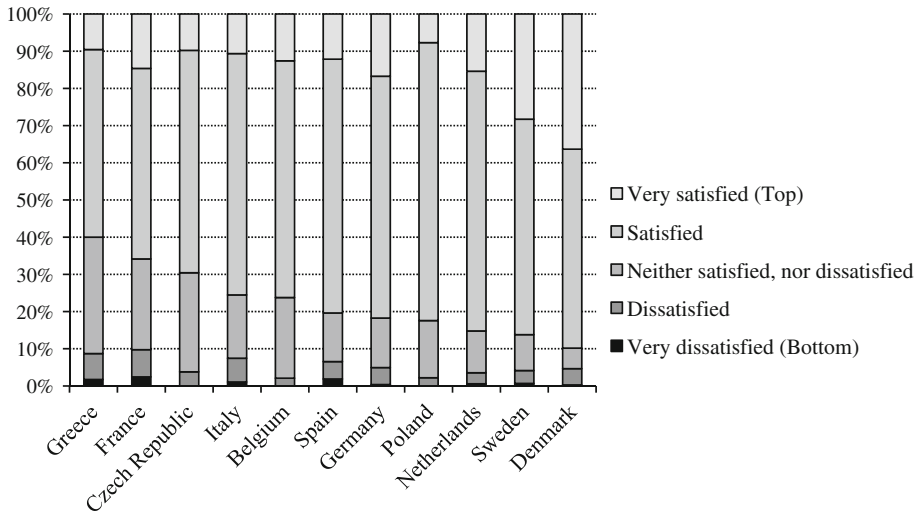
**Table 4** Distribution of reported own job satisfaction by country (in %)

	Very dissatisfied (%)	Dissatisfied (%)	Nor satisfied, neither dissatisfied (%)	Satisfied (%)	Very satisfied (%)
Belgium	0	2	22	63	13
Czech Republic	0	4	27	60	10
Denmark	0	4	6	53	36
France	2	7	23	52	15
Germany	0	5	13	65	17
Greece	2	7	31	50	10
Italy	1	6	17	65	11
Netherlands	1	3	11	70	15
Poland	0	2	15	75	8
Spain	2	5	13	68	12
Sweden	1	3	10	58	28
Total	1	4	15	61	19

Kristensen and Johansson (2008) from data collected in seven European countries in 2004. In line with our study, they find that Northern countries, especially the Danes, are the most satisfied with their job while the French and Greeks rate their job satisfaction quite low.

To correct for potential differences in response scales in the job satisfaction assessments, each respondent younger than 65 years in the COMPARE sample also got two job satisfaction vignettes, describing hypothetical workers with given job characteristics.<sup>7</sup> They are asked to rate the job satisfaction of these hypothetical workers on the same scale used to measure their own job satisfaction. The following two vignette questions are asked:

<sup>7</sup> Respondents of age 65 or older got vignettes on other daily activities.



**Fig. 3** Distribution of reported job satisfaction by country

*Vignette 1: Mike works full-time, five days per week; in principle, he can organize his work in his own way but is still often under a lot of pressure to meet deadlines. He works for a big company and feels that his job is quite secure. How satisfied do you think Mike is with his job?*

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

*Vignette 2: Sally works four days per week and does not experience her job as stressful; she has little say over what she is doing, this is decided by her boss. She feels it is a very secure job. How satisfied do you think Sally is with her job?*

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

These vignettes only describe a subset of all possible job characteristics (hours of work, whether the job is stressful, control over activities, job security) but not, for example, the wage. Ideally, vignettes should be complete, but there is a trade off between being as complete as possible and the drawbacks of long stories that many respondents will not read seriously. Whether the current vignettes are sufficient remains a topic of future research.

Tables 5 and 6 present the frequency distributions of the job satisfaction vignette assessments by country. The job in Vignette 2 is seen as less satisfactory than the job in Vignette 1. Differences across countries are again substantial. Danish respondents are quite positive about the first vignette in particular (with 78% evaluating it as satisfied or very satisfied), while Spanish respondents are very critical of this vignette (52% satisfied or very satisfied). On the other hand, the Swedes are particularly critical about the job in Vignette 2.

### 3.3 Explanatory Variables

In addition to country dummies, the regressors in the econometric model include socio-demographics such as gender, age, marital status, years of education, dummies for

**Table 5** Distribution of reported job satisfaction Vignette 1 by country (in %)

	Very dissatisfied (%)	Dissatisfied (%)	Nor satisfied, neither dissatisfied (%)	Satisfied (%)	Very satisfied (%)
Belgium	0	8	25	62	5
Czech Republic	0	7	35	52	7
Denmark	0	5	17	64	14
France	0	6	35	56	4
Germany	0	4	23	68	5
Greece	1	11	29	43	17
Italy	0	12	30	54	4
Netherlands	0	5	22	64	8
Poland	1	3	21	70	4
Spain	0	16	32	51	1
Sweden	1	12	30	54	3
Total	0	7	25	60	7

**Table 6** Distribution of reported job satisfaction Vignette 2 by country (in %)

	Very dissatisfied (%)	Dissatisfied (%)	Nor satisfied, neither dissatisfied (%)	Satisfied (%)	Very satisfied (%)
Belgium	1	6	34	49	10
Czech Republic	0	7	34	46	13
Denmark	0	7	36	43	14
France	0	10	38	44	7
Germany	0	9	30	55	6
Greece	2	8	32	32	26
Italy	4	12	41	39	3
Netherlands	0	8	41	46	5
Poland	0	12	29	56	3
Spain	0	10	33	52	5
Sweden	3	14	46	32	4
Total	1	9	36	45	10

employment status, and the logarithm of net household income last month, adjusted for PPP differences across countries.<sup>8</sup> We also include two health indicators: the numbers of self-reported symptoms and chronic diseases. See [Appendix](#), Table 9, for variable definitions and sample statistics, revealing large differences across countries in many of the explanatory variables.

The job satisfaction model also includes variables describing job conditions, such as workload, recognition, job security, monthly net labour income and usual hours worked per week. Job conditions are measured by asking whether respondents strongly agree, agree, disagree, or strongly disagree with the statements: “My job is physically demanding”;

<sup>8</sup> Missing household incomes were imputed using, among other variables, an alternative measure of household income as one of the predictors. See [Appendix](#) for details.

“I am under constant time pressure due to a heavy workload”; “I have very little freedom to decide how I do my work”; “I have an opportunity to develop new skills”; “I receive adequate support in difficult situations”; “I receive the recognition I deserve for my work”; “My job promotion prospects/prospects for job advancement are poor”; “My job security is poor”. For each statement, a dummy is created which is equal to one either when the respondent agrees or strongly agrees for positive job characteristics or when the respondent disagrees or strongly disagrees for negative job characteristics. See [Appendix](#), Table 10 for details and sample statistics, again showing large differences across countries.

## 4 Estimation Results

### 4.1 Income Satisfaction

Table 7 presents the parameter estimates of the main equation for the model with identical thresholds for everyone (the baseline model, column (i); these estimates are virtually identical to those of a simple ordered probit model) and the estimates of the (conditional) hopit model [column (ii) to column (vi)] taking account of differences in response scales (DIF). The results for the baseline model are in accordance with most findings in the literature. As expected, household income has a strong positive effect on income satisfaction, while household size has a substantial negative effect. In terms of equivalence scales, the estimates imply that an increase in family size from one to two household members would require an increase in household income of almost 29% to keep income satisfaction constant<sup>9</sup>—an estimated equivalence scale of 1.29. This is comparable to the results of Van Praag and Van der Sar (1988, Table 3), who find equivalence scales between 1.15 and 1.35 for eight out of nine countries. The estimate of Van Praag and Ferrer-i-Carbonell (2008, Table 3.1.4) for the UK is 1.31—also very similar to what we find.

Conditional on income (and other covariates), higher educated individuals are more satisfied with their income. This is consistent with results of Kapteyn et al. (2008), who point out it may be due to the fact that higher educated people have higher permanent income, or to the fact that our measure of income is imperfect so that education is a proxy for the deviation between self-reported income and actual income. The estimated effect of an additional year of education is about the same as the effect of a 2% rise in household income.

Women tend to report higher income satisfaction than men. Age has a positive effect,<sup>10</sup> while poor health (number of symptoms and number of chronic diseases) reduces income satisfaction. Keeping other variables constant, we find no significant differences in income satisfaction between workers, retirees, or disability benefit recipients, but unemployed individuals experience a significantly lower income satisfaction than workers, while inactive persons are more satisfied than workers.

Country dummies indicate that, conditional on income and other covariates, French respondents report the lowest income satisfaction level while Danish respondents report the highest level. Interestingly, keeping the other covariates constant, Polish respondents

<sup>9</sup> The formula to derive the equivalence scale is the following:  $\frac{y^N}{y^1} = N^{-\beta_{\log\text{house}} / \beta_{\log\text{income}}}$ . Where  $y^N$  is the income that a household with  $N$  individuals should have to have the same income satisfaction as a single household with an income of  $y^1$ .

<sup>10</sup> Adding age squared (in all equations) hardly improved the fit and did not change any of the substantive results. We therefore present the specification with a linear age term only, which is easier to interpret.

**Table 7** Baseline model and Hopit model of income satisfaction

Model	Baseline (i) $\beta$	Hopit (ii) $\beta$	(iii) $\gamma^1$	(iv) $\gamma^2$	(v) $\gamma^3$	(vi) $\gamma^4$
Constant	–	–	6.522*** (0.458)	–0.300 (0.309)	–0.617** (0.298)	–0.392* (0.230)
Woman	0.088*** (0.028)	0.169*** (0.036)	0.022 (0.036)	0.046 (0.029)	0.017 (0.027)	–0.007 (0.020)
Age	0.020*** (0.002)	0.019*** (0.003)	–0.003 (0.003)	–0.002 (0.002)	0.004** (0.002)	0.003** (0.001)
Years of education	0.011*** (0.004)	0.025*** (0.005)	–0.008 (0.005)	0.020*** (0.004)	–0.000 (0.004)	–0.001 (0.003)
Log(household size)	–0.313*** (0.038)	–0.516*** (0.049)	–0.203*** (0.048)	0.016 (0.037)	–0.016 (0.036)	0.000 (0.028)
Log(household income)	0.862*** (0.038)	1.296*** (0.049)	0.324*** (0.047)	0.038 (0.037)	0.043 (0.036)	0.098*** (0.028)
Number of symptoms	–0.094*** (0.009)	–0.063*** (0.011)	0.037*** (0.011)	0.005 (0.009)	–0.015* (0.008)	–0.015** (0.007)
Number of chronic diseases	–0.035*** (0.011)	–0.049*** (0.014)	–0.018 (0.014)	0.007 (0.011)	–0.011 (0.011)	0.009 (0.008)
<i>Labour force status:</i>						
Working	–	–	–	–	–	–
Retired	0.063 (0.041)	0.037 (0.052)	–0.009 (0.052)	–0.006 (0.042)	0.013 (0.039)	–0.013 (0.029)
Unemployed	–0.341*** (0.077)	–0.317*** (0.099)	0.152 (0.095)	–0.128* (0.075)	–0.023 (0.076)	–0.003 (0.059)
Disabled	–0.019 (0.072)	0.070 (0.092)	0.129 (0.091)	0.022 (0.068)	–0.085 (0.074)	–0.079 (0.055)
Inactive	0.178*** (0.055)	0.045 (0.069)	–0.021 (0.068)	–0.075 (0.053)	–0.012 (0.052)	–0.081** (0.040)
<i>Country:</i>						
Belgium	–0.010 (0.051)	0.122* (0.064)	0.109 (0.068)	0.096* (0.054)	–0.090* (0.049)	–0.068* (0.036)
Czech Republic	0.020 (0.054)	–0.082 (0.070)	–0.103 (0.075)	–0.033 (0.062)	0.072 (0.050)	–0.041 (0.040)
Denmark	0.626*** (0.050)	0.903*** (0.066)	0.093 (0.068)	0.132** (0.053)	–0.037 (0.050)	0.045 (0.033)
France	–0.565*** (0.066)	–0.153* (0.085)	0.146* (0.085)	0.169*** (0.064)	0.156*** (0.057)	0.015 (0.053)
Germany	–	–	–	–	–	–
Greece	–0.463*** (0.062)	0.685*** (0.080)	1.159*** (0.072)	–0.010 (0.060)	0.129** (0.057)	–0.336*** (0.057)
Italy	–0.108* (0.057)	0.319*** (0.073)	0.493*** (0.071)	–0.100 (0.062)	–0.006 (0.054)	0.062 (0.041)

**Table 7** continued

Model	Baseline (i) $\beta$	Hopit (ii) $\beta$	(iii) $\gamma^1$	(iv) $\gamma^2$	(v) $\gamma^3$	(vi) $\gamma^4$
Netherlands	0.338*** (0.061)	0.165** (0.079)	-0.506*** (0.103)	0.237*** (0.073)	0.001 (0.060)	0.030 (0.040)
Poland	0.071 (0.068)	-0.270*** (0.091)	-0.318*** (0.098)	-0.064 (0.079)	-0.045 (0.069)	0.043 (0.051)
Spain	-0.096 (0.065)	0.230*** (0.083)	0.274*** (0.084)	0.146** (0.066)	-0.251*** (0.067)	-0.037 (0.048)
Sweden	0.217*** (0.062)	0.571*** (0.080)	0.164** (0.081)	0.176*** (0.062)	0.046 (0.058)	-0.118*** (0.044)
Log-likelihood	-26,914	-25,786				

\*, \*\*, \*\*\* Means that the coefficient is significantly different from zero at the 10%, 5%, 1%-level respectively. Number of observations: 7,069. See Table 9 (Appendix) for variable definitions

report about the same level of income satisfaction as German respondents. The fact that Polish respondents report low income satisfaction (Table 1) is therefore mainly explained by the characteristics of the Polish respondents, particularly their low income and large family size.

Allowing for DIF substantially modifies the estimates of the satisfaction equation (column (ii) in Table 7). The likelihood-ratio test strongly rejects the constrained model of no DIF against the more general model allowing for DIF (LR = 2256; 84 degrees of freedom) The coefficient on household income is much higher once we control for DIF, suggesting that individuals with higher income are more “demanding”—they evaluate a given income as less satisfactory than low income individuals with the same other characteristics. The effect of family size also increases, and this approximately compensates the increased income effect so that the equivalence scale does not change much compared to the baseline model—a two person household needs 32% more than a one person household according to the model with DIF, compared to 29% in the baseline model. The effects of education and gender are also much higher than in the baseline model, suggesting that women and higher educated individuals use more negative response scales. On the other hand, the effects of other socio-economic variables (age, employment status, health) do not change much or even decrease.

Many of the socio-economic characteristics significantly affect the thresholds, particularly the first threshold [see column (iii)]. The differences between effects on income satisfaction in the two models can be explained by the effects of the same background variables on the thresholds. For example, income has a positive effect on the first threshold, implying that higher income respondents will more often assess a given income as very unsatisfactory. This is in line with the notion that higher income makes people more demanding; see, for example, Van Praag and van der Sar 1988, who find that the (stated) income required to achieve a given utility level increases with actual income. Our model specification implies that a shift in the first threshold also leads to a parallel shift in all other thresholds, and our estimates of the income coefficients in  $\gamma^1$ ,  $\gamma^2$ ,  $\gamma^3$  and  $\gamma^4$  imply that higher income respondents are more critical at all cut-off points, not only the first.

Thresholds also significantly depend on the country dummies. Italians, for example, uses higher thresholds (i.e., tend to give more negative assessments) than Germans throughout the scale. As was already clear from Tables 2 and 3, Greek respondents tend to

give quite negative vignette evaluations, translating into an unusually high first threshold. As a consequence, the coefficients on the country dummies in the income satisfaction equation turn out to be quite different in the hopit and the baseline model. Polish respondents tend to evaluate the vignettes quite positively, and when this is corrected for, they are worse off than respondents in any other country with the same income and other characteristics. The ranks of the Czech Republic, the Netherlands, and Germany also worsen substantially when correcting for DIF—respondents in all these countries use relatively optimistic evaluation scales and are worse off when this is corrected for. The opposite is found for Greek respondents: for given income and other characteristics, they are in 10th place in the model without DIF, but correcting for their very negative evaluations moves them to 2nd place. Correcting for DIF also improves the position of Italy and Spain.

#### 4.2 Job Satisfaction

Table 8 presents the results for the ordered probit model [column (i) and (iii)] and the hopit model [column (ii) and (iv)] for job satisfaction among 50–64 year-old workers. The first two columns show the results without taking into account job conditions other than hours worked and earnings, while the last two columns add a richer set of job characteristics.<sup>11</sup> As for income satisfaction, a likelihood-ratio test strongly rejects the constrained model without DIF against the more general model allowing for DIF for both specifications (LR = 256.2;  $df = 68$ ) for the model without the set of job characteristics (in either Eq. 1 or Eq. 3) and LR = 302.0;  $df = 100$  for the specification including them (in Eqs. 1 and 3).

The ordered probit model suggests that, keeping individual and job characteristics constant, women report to be more satisfied with their job than men. This is in accordance with many other studies on job satisfaction (Clark 1997; Kaiser 2007). Once DIF is corrected for, however, the difference between women and men is not significant anymore, suggesting that women report being more satisfied with their job because they have different response scales. A reason for this may be that they have lower work expectations than men and are therefore less demanding (Phelan 1994).

Age has a significant positive effect on job satisfaction in both models. Note also that the age effect may reflect a selection process if less satisfied workers retire earlier than more satisfied workers. Years of education has no significant effect on job satisfaction whichever model is considered. Health symptoms have a significant negative effect on job satisfaction in both models. Their effect is lower when the larger set of job characteristics is included in the model, since health problems are associated with unattractive (reported) job characteristics.

Higher earnings have a positive effect on job satisfaction, but this effect is insignificant when more job conditions are included, suggesting that attractive job characteristics (that are correlated with high wages) are more important than the wage itself. Clark and Oswald (1996) find a negative relationship between working hours and job satisfaction in the general adult UK population. All our models suggest that, keeping monthly earnings constant, there is no significant relation between job satisfaction and working hours of older workers in Europe.

<sup>11</sup> Estimates of the parameters determining the thresholds are not presented to save space. They are available in the online appendix.



**Table 8** Job satisfaction among 50–64 year-old workers: Baseline model and Hopit model

	$\beta$		$\beta$	
	Baseline (i)	Hopit (ii)	Baseline (iii)	Hopit (iv)
Woman	0.169*** (0.061)	0.099 (0.075)	0.141** (0.062)	0.069 (0.079)
Age	0.027*** (0.008)	0.023** (0.010)	0.024*** (0.008)	0.019* (0.011)
Years of education	-0.003 (0.008)	0.010 (0.010)	-0.013 (0.009)	0.000 (0.011)
Number of symptoms	-0.141*** (0.025)	-0.146*** (0.031)	-0.096*** (0.026)	-0.099*** (0.032)
Number of chronic diseases	0.025 (0.029)	0.027 (0.036)	0.040 (0.030)	0.044 (0.038)
Log(yearly net earnings)	0.262*** (0.080)	0.196** (0.098)	0.117 (0.083)	0.034 (0.105)
Log(working hours)	-0.065 (0.093)	-0.005 (0.113)	0.048 (0.096)	0.120 (0.120)
<i>Work conditions (dummy variables)</i>				
Job is physically demanding	-	-	-0.025 (0.060)	-0.064 (0.076)
Heavy workload	-	-	-0.138** (0.060)	-0.198*** (0.076)
Freedom at work	-	-	0.090 (0.067)	0.016 (0.084)
Can develop new skills	-	-	0.270*** (0.068)	0.323*** (0.085)
Adequate support in difficult situations	-	-	0.329*** (0.068)	0.310*** (0.085)
Recognition for the job	-	-	0.533*** (0.071)	0.516*** (0.088)
Job advancement opportunity	-	-	0.256*** (0.062)	0.320*** (0.079)
Job Security	-	-	0.182** (0.072)	0.179** (0.090)
<i>Country:</i>				
Belgium	-0.101 (0.117)	-0.056 (0.144)	-0.133 (0.121)	-0.100 (0.154)
Czech Republic	-0.041 (0.122)	-0.033 (0.149)	-0.037 (0.125)	-0.058 (0.158)
Denmark	0.489*** (0.094)	0.190* (0.114)	0.425*** (0.096)	0.106 (0.121)
France	-0.357** (0.141)	-0.218 (0.171)	-0.252* (0.146)	-0.120 (0.184)
Germany	-	-	-	-

**Table 8** continued

	$\beta$		$\beta$	
	Baseline (i)	Hopit (ii)	Baseline (iii)	Hopit (iv)
Greece	-0.577*** (0.127)	-0.628*** (0.152)	-0.431*** (0.131)	-0.501*** (0.162)
Italy	-0.221 (0.136)	0.142 (0.168)	-0.191 (0.139)	0.202 (0.177)
Netherlands	0.054 (0.115)	0.118 (0.142)	-0.048 (0.119)	-0.001 (0.152)
Poland	0.226 (0.162)	0.286 (0.205)	0.141 (0.166)	0.200 (0.217)
Spain	-0.143 (0.132)	0.194 (0.166)	-0.185 (0.135)	0.174 (0.176)
Sweden	0.277** (0.118)	0.622*** (0.150)	0.186 (0.121)	0.569*** (0.159)
Log-likelihood	-5,713	-5,584	-5,595	-5,444

\*, \*\*, \*\*\* Means that the coefficient is significantly different from zero at the 10%, 5%, 1%-level respectively. Number of observations: 1,737. See Table 10 for variable definitions (Appendix)

The final two columns show that most job characteristics significantly affect job satisfaction with the expected sign. The magnitudes of some of the coefficients change when DIF is controlled for, but signs and significance levels do not change much. A heavy workload has a negative effect while the opportunity to develop new skills, receiving adequate support in difficult situations, recognition for the job, job advancement opportunities, and job security all have a positive influence on job satisfaction. The largest impact on overall job satisfaction comes from recognition for the job and from receiving support in difficult situations. Opportunities for developing new skills and future job advancement are also important. This may seem surprising given the fact that the sample consists of older workers who are approaching retirement age. Whether the job is physically demanding and (in the hopit model) freedom at work have no significant effect. These results support the hypothesis that non-pecuniary job characteristics are important for job satisfaction, confirming findings for broader age groups (Clark 2005; Skalli et al. 2008).

The coefficients of the country dummies reflect *ceteris paribus* differences between respective countries and Germany, keeping constant individual characteristics and job characteristics [earnings and hours only in columns (i) and (ii), or the larger set of job characteristics in columns (iii) and (iv)]. Some of them are strongly significant and which ones these are varies across the four model specifications. Correcting for differences in response scales mainly affects the position of Denmark, Sweden, and France. Compared to Germans, Danish and French workers tend to use the more positive and more negative responses, respectively (cf. Tables 5 and 6); once this is taken into account in the models with DIF, their job satisfaction levels are not significantly different from those of German workers with the same characteristics. Swedish workers evaluate a given job more negatively than German workers (cf. Tables 5 and 6) and when this is corrected for in the

models with DIF, their job satisfaction levels are actually higher than those of similar Germans.<sup>12</sup>

In the final model in the last column of Table 8, the only countries which are significantly different from Germany are Greece and Sweden. In all other countries, keeping response scales, individual characteristics, and the rich set of job characteristics constant, job satisfaction levels are not significantly different from those in Germany. Greek workers are less satisfied than Germans with similar jobs. Only Swedish workers are significantly more satisfied, possibly pointing at attractive unobserved job characteristics that are particularly relevant in Sweden, such as a more positive attitude towards older workers than in other countries. This would be in line with Wadensjö (2006), who argues that Swedish firms are willing to share the responsibility of society to increase employability of older workers and sees this as one of the explanations of the success of the Swedish partial retirement program.

## 5 Counterfactuals

To understand the implications of our approach we simulate the distribution of income or job satisfaction in each country using different thresholds—the thresholds that the average respondent in the benchmark country (Germany)<sup>13</sup> would use instead of the actual thresholds used by the respondent. The latter simulation (own thresholds) almost exactly<sup>14</sup> reproduces the observed distribution of reported satisfaction levels in each country, presented in Tables 1 and 4 and Figs. 1 and 3. The simulation of interest, however,—using each country’s own parameters in the satisfaction equation but using the threshold parameters for Germany—produces a counterfactual distribution without observational equivalent. Comparing these counterfactual simulations across countries shows how much of the difference between each country and the benchmark country remains when differences due to DIF are eliminated.

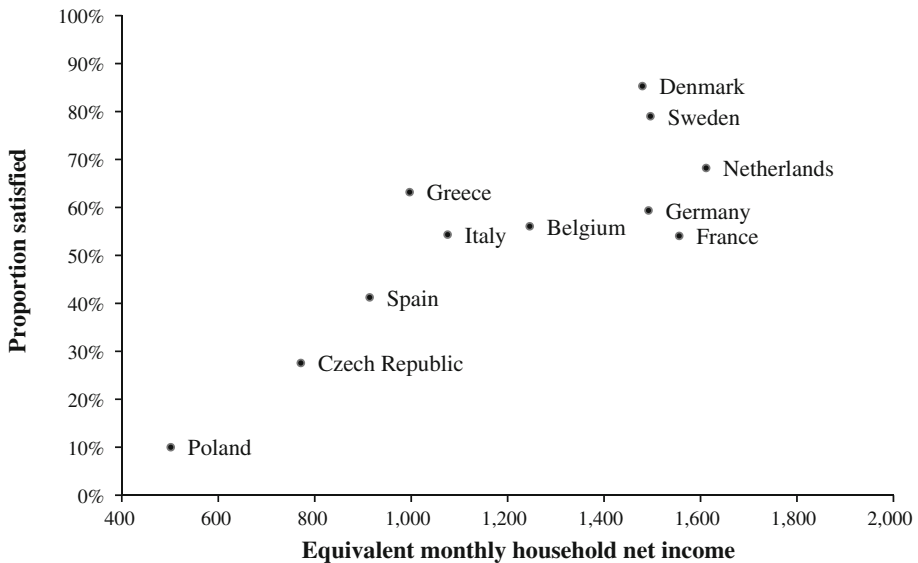
### 5.1 Income Satisfaction

Figure 4 is similar to Fig. 2 but uses the counterfactual simulation to construct the values along the vertical axis. It presents, for each country, the proportion of individuals who would report being satisfied or very satisfied with their income if they would use German benchmark thresholds. The horizontal axis gives the corresponding equivalent monthly household income, as in Fig. 2. Compared to Fig. 2, income satisfaction France is now much more in accordance with income satisfaction in other countries with a similar income level. The low proportion of individuals reporting satisfied with their income in France that we saw in Fig. 2 apparently was partly due to DIF. Greece moves from a relatively low

<sup>12</sup> The changes in the country ranking when correcting for DIF can be compared with results of Kristensen and Johansson (2008) for all workers. Similar to what we find, they find that the ranking of France improves, while that of Denmark worsens. Different from our findings, however, correcting for DIF substantially improves job satisfaction in the Netherlands and Greece and worsens it in Spain. Our other countries are not in their data set.

<sup>13</sup> For each respondent, we replace the thresholds by thresholds of the average German respondent (i.e., with the average individual characteristics of the German sample).

<sup>14</sup> The fit is not exact due to finite sample errors, simulation errors, and, possibly, the fact that the model may not fit the data perfectly well.



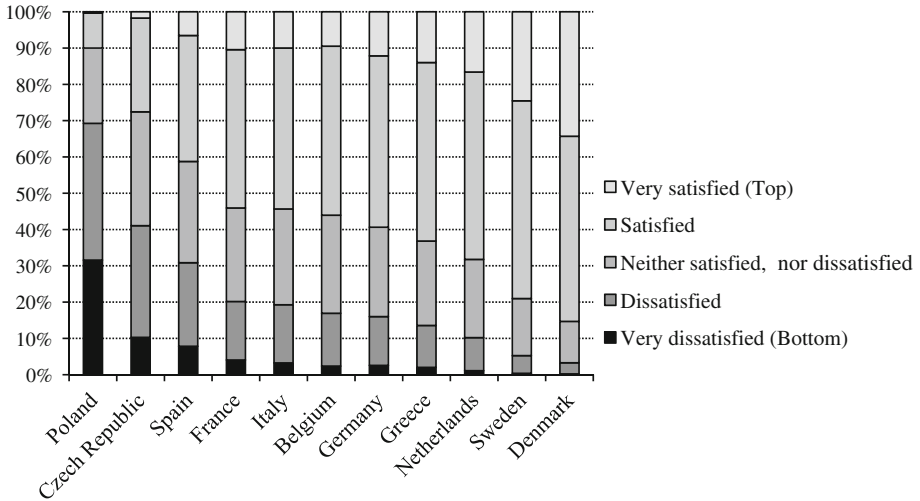
**Fig. 4** Household income (PPP corrected) and income satisfaction among the 50+ individuals across COMPARE countries using German thresholds

satisfaction (given its actual income level) to a relatively high satisfaction country. Correcting for response scale differences makes the difference between Poland and the other countries even larger than before. All in all, the correction brings the ranking of the countries more in line with the ranking of their income levels. The Spearman rank correlation coefficient is equal to 0.66 when DIF is taken into account while it is equal to 0.64 in the raw data; the Pearson correlation coefficient increases from 0.74 to 0.84 when we control for DIF.

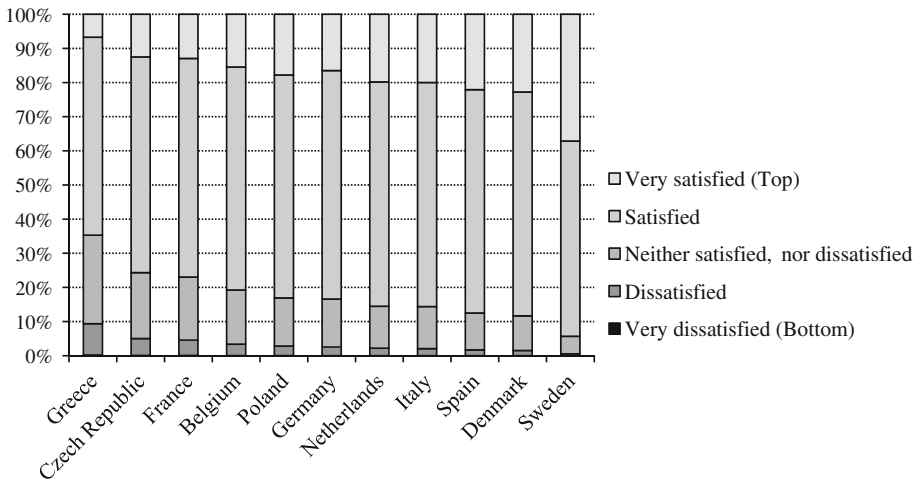
Figure 5 presents the complete counterfactual cumulative income satisfaction distribution for all countries using German benchmark thresholds. It confirms that correcting for DIF has important effects on the country ranking. First, the ranking between Sweden and the Netherlands is reversed—a consequence of correcting for the fact that Swedish respondents tend to assess vignettes with a given income level more negatively than Dutch respondents. Second, there is hardly any difference left between Belgium, Italy and Germany once DIF is eliminated. As in Fig. 4, one of the most salient changes due to eliminating DIF is France. Using German scales, French respondents would be much more satisfied with their incomes than their actual reports (based upon the French scales) suggest, and France becomes an “average country.” As expected given the estimation results and Fig. 4, Greece does much better after the correction than before correcting for DIF. Finally, the cumulative distribution function of income satisfaction in Spain no longer crosses that of the Czech Republic. Spain does unambiguously better than the Czech Republic.

## 5.2 Job Satisfaction

The counterfactual cumulative distributions of job satisfaction assuming that all individuals use the German benchmark thresholds are presented in Fig. 6. It is based upon the



**Fig. 5** Predicted distribution of income satisfaction using German thresholds



**Fig. 6** Predicted distribution of job satisfaction using German thresholds

final model in Table 8 [column (iv)], including the rich set of job characteristics. The country ranking differs substantially from the one in Fig. 3. Once differences in response scales are eliminated, Sweden becomes the country with the highest level of job satisfaction, with Denmark in second place, but at substantial distance. Greece is the country with worst job satisfaction in both figures, but the difference with the other countries is much larger once DIF is corrected for. As for income satisfaction, job satisfaction in France increases when German rather than French thresholds are used.

Accounting for DIF reduces the cross-country association between job and income satisfaction: the cross-country rank correlation between country specific percentages of

working respondents younger than 65 who are (at least) satisfied with their income and with their jobs decreases from 0.80 for reported satisfaction to 0.43 for the counterfactual rates using the German thresholds.<sup>15</sup> An interpretation is that response scales in different domains are positively correlated: respondents who tend to give negative evaluations in one domain will often do the same in another domain. For example, French respondents assign low satisfaction to the income vignettes as well as the job satisfaction vignettes compared to respondents in other countries. This illustrates that correcting for DIF may also be important to analyze the relation between satisfaction levels in various domains of life.

## 6 Conclusion

This paper analyses two important components of economic well-being among the 50+ in 11 European countries: satisfaction with household income and job satisfaction. The first one is important in order to assess the overall economic welfare of the elderly. The results highlight a large variation in self-reported income satisfaction. The lowest is found in Poland and the highest in Denmark. Differences across countries are partly explained by differences in response scales. Once these differences are eliminated, the cross-country differences are much better in line with differences in an objective measure of purchasing power of household income. Correcting for differences in response scales also alters the ranking across countries. The most striking change is for France, where respondents tend to use negative assessments more often than in other countries.

An important motivation for this paper is that how a country compares to other countries in terms of living standard is an important input for public policy on old age social security and pensions and combating poverty and social exclusion among the older part of the population. We have shown that it matters whether the country comparison is done with or without correcting for response scale differences (DIF). So should policy makers use the cross-country comparison with or without corrections for DIF? Under the assumptions that we have made, the answer is a clear yes: assuming differences in vignette evaluations purely reflects differences in the way terms like “very satisfied” and “not satisfied” are used, correcting self-assessments for such differences seems a good thing if the aim is to compare genuine living standards. This leads to the conclusion that living standard comparisons come much closer to objective comparisons of equivalized and PPP corrected average household incomes than the subjective income satisfaction reports would suggest.

There is an alternative interpretation of the differences in vignette evaluations, however. If, for example, goods are publicly provided (free of charge) in one country and not in another, or poor households can do more with a given income in one country than in another country, because of differences in, e.g., housing subsidies or health insurance, then a given income amount may lead to different living standards in different countries. In that case vignette equivalence would not be satisfied and our corrections would take away genuine differences in living standards. We do not think this can explain much of our results—for example the fact that French respondents give negative assessments would then suggest that the French get less public support than similar countries, which seems implausible. A similar conclusion is drawn by Kapteyn et al. (2008) on the basis of comparing evaluations of vignettes with low and high incomes. Moreover, the tendency to

<sup>15</sup> The Pearson correlation decreases from 0.75 to 0.24.

give less positive evaluations in France is also found for other subjective well-being measures such as life satisfaction (Angelini et al. 2011), further supporting the notion of cultural differences in thresholds.

Older workers in Europe are generally satisfied with their jobs. Cross-country differences are not as large as for income satisfaction. Being able to develop new skills and having job advancement opportunities contribute substantially to job satisfaction, though recognition for the job is the most important factor. Keeping job characteristics as well as response scales constant, Swedish workers are more satisfied than workers in all other countries considered, possibly due to a more positive attitude of employers towards older workers in Sweden than elsewhere. Sweden remains the country where job satisfaction among older workers is highest if cross-country variation in job characteristics is taken into account and only the response scales are kept constant. The raw data, however, do not reveal this, since the actual job satisfaction reports are also affected by response scale variation, leading to lower reported satisfaction in Sweden and higher satisfaction in Denmark, for example. Like for income satisfaction, correcting for response scale differences changes the ranking of the countries. Now that financial incentives for early retirement have been or are being removed, and other factors like job characteristics and job satisfaction are gaining importance for the decision to work longer, this seems an important message for national policy makers who compare the situation in their own country to that in other countries. Whereas looking at the raw data would suggest that Denmark is the European role model for job satisfaction of older workers, Sweden becomes the best performing country when controlling for the Danish tendency to use positive scales and the Swedish tendency to be more negative.

There are common features in the response scale differences in job satisfaction and income satisfaction. French respondents tend to be critical in both assessments, while Danish and Dutch respondents are always on the optimistic end of the spectrum. The tendency to give negative evaluations in France and the Danish tendency to use very positive qualifications seems rather general; Angelini et al. (2011) and Bonsang and Van Soest (2011), for example, also find it for satisfaction with life and social contacts, respectively, suggesting that the correction corrects for differences in cultural norms relevant for reporting behavior. As a consequence, correcting for DIF decreases the cross-country association between average income and job satisfaction among workers younger than 65.

The fact that correcting for DIF brings subjective and objective evaluations closer to each other can be seen as support for the validity of the vignettes approach as a tool for improving cross-country comparisons. It is in line with the finding of King et al. (2004) that correcting for DIF using anchoring vignettes increases the cross-country correlation between objective and subjective measures of health. Still, more work is needed to test the validity of the vignette approach in the domains considered and establish the robustness of the results. The main underlying assumptions are response consistency and vignette equivalence, which have been studied in other domains (e.g., Van Soest et al. 2011) but not for income and job satisfaction. Response consistency requires that respondents evaluate the hypothetical situations on the same scale that they use to evaluate themselves; this could be violated, for example, if self-assessments are affected by social desirability bias but vignette evaluations are not. We do not think this is particularly problematic in our case. Vignette equivalence means that respondents in different countries interpret the vignettes in the same way. As discussed above, this is not an innocuous assumption, particularly in the context income satisfaction, but we have also explained why we think our

results are not due to violation of vignette equivalence. Still, validating the use of vignettes and testing these assumptions remains an important issue for future research.

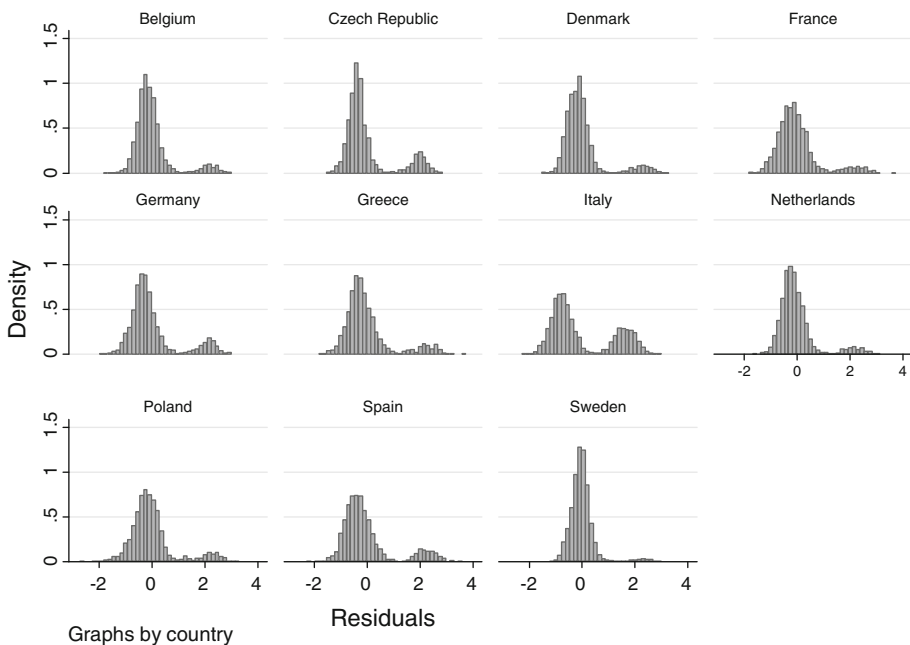
**Acknowledgments** We are grateful to two anonymous referees, Teresa Bago d'Uva, Didier Fouarge, Hendrik Juerges, Raymond Montizaan, and participants of the final COMPARE conference in Brussels for useful comments. This paper was written as part of the project COMPARE, funded by the European Commission through its 6th framework (project number CIT5-CT-2005-028857). Data collection and infrastructure for making data available to researchers was mainly funded by the European Commission through several SHARE related projects in the 5th and 6th framework programmes (CIT5-CT-2005-028857, QLK6-CT-2001-00360; RII-CT-2006-062193). Additional funding was provided by the US National Institute on Aging (grant numbers U01 AG09740-13S2; P01 AG005842; P01 AG08291; P30 AG12815; Y1-AG-4553-01; OGHA 04-064; R21 AG025169) and various national sources (see <http://www.share-project.org> for a full list of funding institutions).

## Appendix

See Tables 9 and 10.

### Construction of a household income measure

The measure of after-tax household income contains a substantial number of missing values and unreliable outliers. To get a reliable measure of after tax household income, we applied the following procedure. First, we ran a regression of log of household income last month, excluding the country-specific first and last percentile, on standard explanatory variables [country dummies, education, age, gender, log household size, employment status, and health status (numbers of chronic diseases and reported symptoms)]. Figure 7



**Fig. 7** Distribution of the residuals of an OLS on log(hh income)



**Table 9** Descriptive statistics. Income satisfaction sample

	All countries	Belgium	Czech Republic	Denmark	France	Germany	Greece	Italy	Netherlands	Poland	Spain	Sweden
Woman	54%	54%	58%	54%	55%	53%	52%	54%	52%	55%	53%	54%
Age	64.1	65.0	64.3	64.0	64.6	64.5	64.9	64.5	61.4	62.3	63.6	65.8
Year of education	11.0	11.6	11.4	13.1	11.9	12.5	8.7	8.2	11.4	9.2	7.7	11.1
Household size	2.2	2.0	2.1	2.0	2.0	2.1	2.3	2.5	2.2	2.9	2.7	1.9
Monthly household income (in Euros, PPP corrected)	1,878	1,883	1,177	2,210	2,311	2,266	1,607	1,864	2,542	899	1,635	2,198
Number of symptoms <sup>a</sup>	1.8	1.9	2.1	1.5	1.9	1.7	1.5	1.8	1.1	2.6	1.6	1.7
Number of chronic diseases <sup>b</sup>	1.6	1.6	1.8	1.7	1.4	1.4	1.4	1.8	1.1	2.1	1.7	1.5
<i>Labour force status:</i>												
Working	29%	21%	27%	43%	28%	28%	33%	18%	40%	20%	27%	37%
Retired	52%	52%	68%	49%	59%	55%	41%	54%	36%	58%	34%	60%
Unemployed	3%	6%	3%	2%	3%	7%	0%	1%	0%	5%	4%	2%
Disabled	4%	4%	2%	5%	2%	3%	1%	2%	5%	14%	4%	2%
Inactive	11%	17%	0%	1%	8%	8%	25%	25%	18%	3%	31%	0%
Number of observations	7,069	810	861	936	347	1088	470	650	484	515	456	452

All 50+ individuals

<sup>a</sup> The respondent is asked to report whether he/she suffers from the following symptoms: (1) Pain in your back, knees, hips or any other joint; (2) Heart trouble or angina, chest pain during exercise; (3) Breathlessness, difficulty breathing; (4) Persistent cough; (5) Swollen legs; (6) Sleeping problems; (7) Falling down; (8) Fear of falling down; (9) Dizziness, faints or blackouts; (10) Stomach or intestine problems, including constipation, air, diarrhoea; (11) Incontinence or involuntary loss of urine; (12) Other symptoms, not yet mentioned

<sup>b</sup> Chronic disease corresponds to a list of conditions of the respondent diagnosed by a doctor: (1) A heart attack including myocardial infarction or coronary thrombosis or any other heart problem including congestive heart failure; (2) High blood pressure or hypertension; (3) High blood cholesterol; (4) A stroke or cerebral vascular disease; (5) Diabetes or high blood sugar; (6) Chronic lung disease such as chronic bronchitis or emphysema; (7) Asthma; (8) Arthritis, including osteoarthritis, or rheumatism; (9) Osteoporosis; (10) Cancer or malignant tumour, including leukaemia or lymphoma, but excluding minor skin cancers; (11) Stomach or duodenal ulcer, peptic ulcer; (12) Parkinson's disease; (13) Cataracts; (14) Hip fracture or femoral fracture; (15) Other conditions, not yet mentioned

**Table 10** Descriptive statistics. Job satisfaction sample

	All countries	Belgium	Czech Republic	Denmark	France	Germany	Greece	Italy	Netherlands	Poland	Spain	Sweden
Woman	47%	44%	47%	53%	46%	52%	34%	45%	45%	43%	36%	55%
Age	55.3	54.9	54.6	56.1	55.1	55.6	56.5	55.8	53.5	53.5	55.3	56.7
Year of education	13.1	13.5	12.5	14.3	13.4	14.2	11.1	12.0	13.1	11.3	11.1	13.0
Number of symptoms	1.1	1.2	1.6	1.1	1.2	1.2	0.7	1.0	0.8	1.3	0.7	1.2
Number of chronic diseases	0.9	1.0	1.3	1.0	0.9	0.9	0.6	1.0	0.7	1.1	0.9	0.9
Monthly net labour earnings (in Euros, PPP corrected)	1,505	1,650	871	1,550	1,946	1,766	1,778	1,380	1,711	628	1,472	1,529
Working hours per week	38.9	35.5	44.0	36.5	38.1	39.0	43.4	36.5	34.4	45.0	39.7	39.8
<i>Job conditions (dummy variables):</i>												
Job is physically demanding	46%	46%	59%	40%	38%	46%	56%	54%	40%	59%	44%	34%
Heavy workload	52%	44%	49%	52%	33%	70%	57%	56%	41%	49%	46%	52%
Freedom at work	73%	73%	60%	84%	78%	74%	50%	70%	84%	53%	60%	88%
Can develop new skills	70%	61%	67%	82%	60%	74%	49%	63%	84%	45%	61%	84%
Adequate support in difficult situation	72%	74%	61%	82%	56%	74%	55%	66%	76%	67%	75%	80%
Recognition for the job	73%	71%	57%	80%	54%	77%	63%	66%	79%	78%	75%	79%
Job advancement opportunity	33%	49%	28%	30%	41%	28%	34%	35%	47%	27%	29%	25%
Job security	81%	87%	72%	88%	89%	83%	73%	80%	73%	64%	86%	86%
Number of observations	1,737	143	184	344	82	263	115	94	169	91	107	145

All 50–64 year-old workers

Dummy variables for job conditions are equal to one either when the respondent disagrees or strongly disagrees (for little freedom at work, poor job advancement opportunities, and poor job security) or when the respondent agrees or strongly agrees (for the other job characteristics)

presents the distribution of the residuals of this regression by country. It appears clearly that we have a high proportion of outliers in many countries (especially in Italy, Czech Republic). Based upon these results, we chose thresholds for each country to define observations as outliers.

In a second step, we used the valid measures of current household income to run a second regression including the standard explanatory variables and another measure of household income based on the information about personal income and the income of other household members received last year. Finally, we replaced the unreliable or missing values of the general household income by the prediction of the model using the other household income measure. This method has the advantage of providing information about household income for almost all observations in the SHARE sample.

## References

- Angelini, V., Cavapozzi, D., Corazzini, L., & Paccagnella, O. (2011). Do Danes and Italians rate life satisfaction in the same way? Using vignettes to correct for individual-specific scale biases. *Social Indicators Research*, forthcoming.
- Bago d'Uva, T., O'Donnell, O., & Van Doorslaer, E. (2008a). Differential health reporting by education level and its impact on the measurement of health inequalities among older Europeans. *International Journal of Epidemiology*, 37(6), 1375–1383.
- Bago d'Uva, T., Van Doorslaer, E., Lindeboom, M., & O'Donnell, O. (2008b). Does reporting heterogeneity bias the measurement of health disparities? *Health Economics*, 17(3), 351–375.
- Bonsang, E., & Van Soest, A. (2011). Satisfaction with social contacts among older Europeans. *Social Indicators Research*, forthcoming.
- Borjas, G. (1979). Job satisfaction, wages and unions. *Journal of Human Resources*, 14, 21–40.
- Börsch-Supan, A., Brügiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., et al. (2008). *Health, ageing and retirement in Europe (2004–2007)*. Mannheim: MEA.
- Börsch-Supan, A., Brügiavini, A., Jürges, H., Mackenbach, J., Siegrist, J., & Weber, G. (2005). *Health, ageing and retirement in Europe—first results from the survey of health, ageing and retirement in Europe*. Mannheim: MEA.
- Charlier, E. (2002). Equivalence scales for the former West-Germany. *Review of Income and Wealth*, 48(1), 99–126.
- Clark, A. E. (1997). Job satisfaction and gender: Why are women so happy at work? *Labour Economics*, 4(4), 341–372.
- Clark, A. E. (2005). What makes a good job? Evidence from OECD countries. In S. Bazen, C. Lucifora, & W. Salverda (Eds.), *Job quality and employer behaviour*. Basingstoke: Palgrave Macmillan.
- Clark, A. E., Frijters, P., & Shields, M. (2008). Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. *Journal of Economic Literature*, 46(1), 95–144.
- Clark, A. E., Georgellis, Y., & Sanfey, P. (1998). Job satisfaction, wage changes and quits: Evidence from Germany. *Research in Labor Economics*, 17, 95–121.
- Clark, A. E., & Oswald, A. J. (1996). Satisfaction and comparison income. *Journal of Public Economics*, 61, 359–381.
- Clegg, C. W. (1983). Psychology of employee lateness, absence and turnover: A methodological critique and an empirical study. *Journal of Applied Psychology*, 68, 88–101.
- Ferrer-i-Carbonell, A., & 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, 709–722.
- Freeman, R. B. (1978). Job satisfaction as an economic variable. *American Economic Review Papers and Proceedings*, 68(2), 135–141.
- Grundy, E. (2006). Ageing and vulnerable elderly people: European perspectives. *Ageing and Society*, 26, 105–134.
- Hamermesh, D. S. (1977). Economic aspects of job satisfaction. In O. C. Ashenfelter & W. E. Oates (Eds.), *Essays in labour market analysis* (pp. 53–72). New York: Halsted Press.
- Hamermesh, D. S. (2001). The changing distribution of job satisfaction. *Journal of Human Resources*, 36(1), 1–30.
- Holland, P. W., & Wainer, H. (1993). *Differential item functioning*. Hillsdale, NJ: Lawrence Erlbaum.

- Hsieh, C.-M. (2004). Income and financial satisfaction among older adults in the United States. *Social Indicators Research*, 66, 249–266.
- Kaiser, L. (2007). Gender-job satisfaction differences across Europe: An indicator for labor market modernization. *International Journal of Manpower*, 28, 75–94.
- 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(1), 461–473.
- Kapteyn, A., Smith, J. P., & Van Soest, A. (2008). Are Americans really less happy with their incomes? RAND labor and population working paper WP-591.
- Kapteyn, A., Smith, J. P., & Van Soest, A. (2010). Life satisfaction. In E. Diener, J. E. Helliwell, & D. Kahneman (Eds.), *International differences in subjective well being*. Oxford: Oxford University Press.
- 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(1), 567–583.
- Kosloski, K., Ekerdt, D., & DeViney, S. (2001). The role of job-related rewards in retirement planning. *Journal of Gerontology: Psychological Sciences*, 56B(3), 160–169.
- Kristensen, N., & Johansson, E. (2008). New evidence on cross-country differences in job satisfaction using anchoring vignettes. *Labour Economics*, 15, 96–117.
- Phelan, J. (1994). The paradox of the contented female worker: An assessment of alternative explanations. *Social Psychology Quarterly*, 57(2), 95–107.
- Salomon, J., Tandon, A., & Murray, C. (2004). Comparability of self rated health: Cross sectional multi-country survey using anchoring vignettes. *British Medical Journal*, 328(7434), 258–260.
- Schyns, P. (2002). Wealth of nations, individual income and life satisfaction in 42 countries: A multilevel approach. *Social Indicators Research*, 60, 5–40.
- Skalli, A., Theodossiou, I., & Vasileiou, E. (2008). Jobs as Lancaster goods: Facets of job satisfaction and overall job satisfaction. *Journal of Socio-Economics*, 37(5), 1906–1920.
- Sousa-Poza, A., & Sousa-Poza, A. A. (2000). Well-being at work: A cross-national analysis of the levels and determinants of job satisfaction. *The Journal of Socio-Economics*, 29, 517–538.
- Tsakloglou, P. (1996). Elderly and non-elderly in the European Union: A comparison of living standards. *Review of Income and Wealth*, 42(3), 271–291.
- Van Praag, B. M. S., & Ferrer-i-Carbonell, A. (2008). *Happiness quantified—a satisfaction calculus approach*. Oxford: Oxford University Press.
- Van Praag, B. M. S., Frijters, P., & 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., & Van der Sar, N. L. (1988). Household cost functions and equivalence scales. *Journal of Human Resources*, 23, 193–210.
- Van Praag, B. M. S., & Warnaar, M. F. (1997). The cost of children and the use of demographic variables in consumer demand. In M. Rosenzweig & O. Stark (Eds.), *Handbook of population and family economics* (pp. 241–274). Amsterdam: Elsevier North-Holland.
- Van Soest, A. (2008). Enhancing international comparability using anchoring vignettes. In A. Börsch-Supan, et al. (Eds.), *Health, ageing and retirement in Europe (2004–2007)* (pp. 351–355). Mannheim: MEA.
- Van Soest, A., Delaney, L., Harmon, C., Kapteyn, A., & Smith, J. P. (2011). Validating the use of vignettes for subjective threshold scales. *Journal of the Royal Statistical Society Series A*, forthcoming.
- Wadensjö, E. (2006). Part-time pensions and part-time work in Sweden. *European Papers on the New Welfare*, 6, 29–45.
- Zaidi, A., Frick, J. R., & Büchel, F. (2005). Income mobility in old age in Britain and Germany. *Ageing and Society*, 25, 543–565.