EconomEtica

Centro interuniversitario per l'etica economica
e la responsabilità sociale di impresa
promosso dalla Fondazione Italiana Accenture

N.44 November 2012

Leonardo Becchetti
Luisa Corrado
Paola Samà

Inside the Life Satisfaction Blackbox

Working papers

Inside the Life Satisfaction Blackbox

Leonardo Becchetti* Luisa Corrado† Paola Samà‡

This Daft: November 21, 2012

Abstract

We propose a measure of life satisfaction, alternative to the standard synthetic cognitive wellbeing question, based on the specific contribution of eleven life satisfaction sub-components (including satisfaction about the past, life opportunities, hope for the future, vitality, control over one's on life, meaning of life). The alternative measure is either estimated as a latent factor, obtained as a simple unweighted average from the above mentioned sub-components, or extracted with principal component analysis. We document that the new dependent variable fits much better standard socio-demographic controls and corrects the "Danish life satisfaction bias" in the direction suggested by the vignette approach. These findings do not reject our theoretical assumption that the alternative measures derived from the life satisfaction sub-components are less noisy and less culturally biased and therefore perform better than the standard self-reported life satisfaction. The straightforward policy advice of the paper is to introduce the above mentioned sub-components (similarly to what happens with sub-questions used to calculate the General Health Questionnaire score) in an additional question to measure more effectively subjective wellbeing.

Keywords: life satisfaction, country bias, measurement error, multiple indicators.
JEL: I30, I31.

*Department of Economics, Law and Institutions, University of Rome Tor Vergata (Italy). E-mail: becchetti@economia.uniroma2.it

†Department of Economics, Law and Institutions, University of Rome Tor Vergata (Italy) and University of Cambridge (UK), CreMic and CIMF. E-mail: luisa.corrado@uniroma2.it; lc242@econ.cam.ac.uk. Luisa Corrado gratefully acknowledges the Marie-Curie Intra European fellowship 039326.

‡Department of Economics, Law and Institutions, University of Rome Tor Vergata (Italy).

1 Introduction

The investigation of the determinants of life satisfaction has boomed in recent years due to the availability of worldwide information on subjective wellbeing at individual levels in many well-known surveys (such as the German Socioeconomic Panel, the British Household Panel Survey and the Gallup Survey). The topic is of particular importance for at least four reasons. First, it provides an alternative independent source (beyond experimental evidence) for testing previously undemonstrated assumptions on human preferences (or social norms) affecting subjective wellbeing which are at the basis of all theoretical economic models. Second, it provides precious evidence which widens the range of factors affecting life satisfaction beyond the dimension of observed choices helping us to understand the importance, among others, of relative comparisons, hedonic adaptation, experienced utility and relationship between expectations and realizations (Frey and Stutzer, 2002a and 2002b). Third, it sheds lights on so far unexplored important aspects of economic reality (i.e. the measurement of the shadow value of non market goods)[1] with relevant policy insights. Fourth, it provides information and evidence for the debate on the reform of wellbeing indicators: if straightforward happiness maximization is not a good idea for various reasons, happiness studies may provide stimulating insights on what currently adopted wellbeing indicators may have left behind.

In spite of this great potential the validity of life satisfaction literature findings is challenged by many methodological problems related to interpersonal and inter-country comparability of the standardly adopted measure (self declared subjective wellbeing) which lacks of cardinality. The vignette approach is a recent attempt to overcome the problem (King and Wand, 2007; Corrado and Weeks, 2010). The approach corrects for individual scale heterogeneity using differences across individuals in evaluating a common situation (the vignette) with the same response categories as the self-assessment question. As it is well known, even this approach has limits since the two hypotheses on which it hinges (vignette equivalence and response consistency[2]) are often rejected by empirical tests (Bago d'Uva *et al.*,2009; Ferrer-i-Carbonell *et al.*, 2010; Corrado and Weeks, 2010).

The original contribution of our paper is in the definition of a theoretical framework which aims at improving upon standard subjective wellbeing measures and predicts the superiority of three alternative measures of life satisfaction in terms of capacity of reducing the dependent variable noise and cultural biases captured by country dummies. The approach is based on the measure of 11 life satisfaction sub-components. Our main argument is that, when asked to formulate their life satisfaction score, individuals intuitively weight different sub-components (evaluation of past life, opportunities for the future, overall meaning of their own life, vitality, etc.). Since the operation is not easy, the general abstract life satisfaction question incorporates much more noise and measurement error than a latent variable which may be extracted by using direct answers to each of the above mentioned implicit sub-components. A second argument supporting our main assumption is that the sub-component questions are much more straightforward and easy to answer when they are formulated (as in the SHARE database we use in this paper) on a 1-4 range in which any number is associated to an adjective whose meaning can be grasped immediately. On the contrary, in the standard 0-10 life satisfaction questions there is no verbal correspondence for each of the scale numerical values. We finally postulate that an additional advantage of the sub-component approach is that country specific cultural biases (also due to the different nuances of the translation of the "life satisfaction" term in different languages) tend to be much larger on the general questions than when averaging sub-components or extracting from them the error-free latent life satisfaction factor. We test our hypothesis on data from the SHARE database where, to the standard life satisfaction question - "*How satisfied are you with your life, all things considered?"* with responses on a scale from 0 (completely dissatisfied) to 10 (completely satisfied) - an additional question on life satisfaction sub-component is added. The question relates to 11 items:

1. *How often do you think your age prevents from doing the things you would like to do?*

2. *How often do you feel that what happens to you is out of control?*

[1] The main contributions in this field are those valuing air pollution (Welsch, 2002 and Luechinger, 2007), terrorist activity (Frey et al., 2009), noise nuisance (van Praag and Baarsma, 2005) and flood disasters (Luechinger and Raschky, 2009).

[2] Vignette equivalence requires that the scenarios in the vignettes are perceived with no systematic differences across respondents. Response consistency requires that individuals use the response category in the self-assessment question in the same way as when they evaluate hypothetical scenarios in the vignettes.

3. *How often do you feel left out of things?*
4. *How often do you feel that you can do the things that you want to do?*
5. *How often do you feel that family responsibilities prevent you from doing what you want to do?*
6. *How often do you feel that shortage of money stops you from doing the things that you want to do?*
7. *How often do you look forward to another day?*
8. *How often do you feel that your life has meaning?*
9. *How often on balance, do you look back to your life with a sense of happiness?*
10. *How often do you feel full of energies these days?*
11. *How often do you fell that life is full of opportunities?*

For each item answers are given on a 1-4 scale with an adjective (often, sometimes, rarely, never) being matched to any value.[3]

Our findings document that the three alternative approaches (estimated latent life satisfaction regressing the standard 0-10 answer on the eleven life satisfaction sub-components, unweighted average of the eleven life satisfaction sub-components, extraction of the first principal component from principal component analysis on the sub-components) improve by large the goodness of fit of our baseline life satisfaction estimate with respect to the use of the standard life satisfaction question. The adjusted R-squared grows by around 15 points (20 points when sub-components are interacted with socio-demographic controls and country dummies to calculate predicted life satisfaction) and the AIC and BIC scores confirm the improvement. These findings support our theoretical hypothesis that the three alternative measures reduce the dependent variable noise. We further document that our three approaches correct in the expected direction the well-known Danish cultural bias in life satisfaction answers (Inglehart and Klingemann, 2000, Eurobarometer 2002; Corrado and Weeks, 2010; Kapteyn, Smith, and Soest, 2009) which we find in our data as well.[4] Our approaches therefore provide in this respect results similar to the vignette approach without requiring the two limiting assumptions of vignette equivalence and response consistency.

The straightforward policy advice stemming from our paper is that surveys should introduce an additional question including the above mentioned sub-components in order to have a better measure of subjective wellbeing. Since an additional question with the eleven sub-points is enough to achieve the goal, our results indicate that the trade-off between improving quality of data and enriching surveys with more precise questions on different well-being dimensions is clearly in favour of such decision. What can be also noted is that what we propose for life satisfaction is akin to the approach followed for the construction of another wellbeing index (the General Health Questionnaire[5] score in the BHPS) which is used to measure emotional prosperity and is the average of 12 mental distress sub-questions.

The paper is divided into five sections (including introduction and conclusions) and organized as follows. In the second section we illustrate our theoretical framework and the two hypotheses to be tested. In the third section we discuss descriptive findings and present our econometric specification. In the fourth section we present and discuss econometric findings and illustrate several robustness checks. The fifth section concludes.

[3] In order to harmonise the response scale of the 11 sub-components with the response scale of the synthetic question on life satisfaction we have re-ordered the response scale of the five sub-components denoting positive dimensions of well-being (questions 4 and 7-11) as follows: never=1, rarely=2,sometimes=3,often=4. For the five sub-components denoting negative dimensions of well-being (questions 1-3 and 5-6) the response scale are left unchanged: 1=often, 2=sometimes, 3=rarely, 4=never.

[4] Corrado and Weeks (2010) examine the use of vignettes to correct for the different use of response scales when rating life satisfaction. They show that these additional questions can, under certain conditions, be used to correct for the resultant biases in model parameters. The bias is found especially for top ranked countries such as Denmark thereby confirming that country rankings reflect not just the true variation in life satisfaction but a different use of the response scales and more optimistic evaluations of life of certain countries and cultures (see also Kapteyn, Smith, and Soest, 2009).

[5] See Golderberg and Williams (1988).

2 Theoretical framework

We conceive the "true" cognitive measure of subjective wellbeing for the $i-th$ individual as a latent variable which is a weighted average of J different components (vitality, evaluation over past life, outlook at the future, money and leisure satisfaction, being in control over one's own life, meaning of life, etc.):

$$LS_i^* = \boldsymbol{\omega}_i^{'}\mathbf{Z}_i^* \tag{1}$$

where LS_i^* is the true overall life satisfaction for the $i-th$ individual, while $\mathbf{Z}_i^*=\{Z_{ij}\}$ is a $J \times 1$ column vector in which subcomponents have ω-weights - $\boldsymbol{\omega}_i=\{\omega_{ij}\}$ is a $J \times 1$ vector of parameters - measuring their specific impact on the synthetic life satisfaction evaluation.[6]

Our assumption is that, when individuals are directly asked the standard life satisfaction question (LS), the random component is larger due to the higher difficulty of *i*) understanding the more general question (in itself and comparatively across countries due to the different language nuances); *ii*) matching a different more intuitive verbal evaluation to any numerical value of the response scale and *iii*) averaging its different components without explicitly mentioning them.[7] We therefore consider our dependent variable,[8] the standard life satisfaction question, as characterised by measurement error within the classical errors-in-variables framework, and characterised by a fully observed continuous dependent variable.[9] Hence, when the standard question is formulated, surveys capture the following variable:

$$LS_i = LS_i^* + v_i \tag{2}$$

where v_i represents the measurement error and $E(v_i) \neq 0$. More specifically, we assume that v_i has a country specific (v_c) and an individual specific (ε_i) bias

$$v_i = v_c + \varepsilon_i \tag{3}$$

Conversely, when individuals are asked about the individual components we obtain

$$Z_{ij} = Z_{ij}^* + \tilde{v}_{ij} \tag{4}$$

where

$$\tilde{v}_{ij} = \widetilde{v}_{cj} + \widetilde{\varepsilon}_{ij}$$

is also a measurement error which captures individual bias $\widetilde{\varepsilon}_{ij}$ (i.e. due to a misunderstanding of the specific Z question or to a difficulty of the individuals of evaluating correctly his/her situation) and country-specific bias $\widetilde{v}_{cj}$ (i.e. due to cultural and linguistic differences in the way the life satisfaction sub-component question is understood in different countries or to strategic answering that is a social/cultural tendency to overestimate or underestimate own levels of life satisfaction for each sub-component).

We assume, however, that the bias disappears when using individual components as far as the number of components increases and the number of interviewed individuals grows so that $E\mathbf{Z}_i \simeq E\mathbf{Z}_i^*$ where $\mathbf{Z}_i = \{Z_{ij}\}$.

[6] For simplicity and without lack of generality we assume that the weights are common to each individual. The idea of happiness fundamentals which are common to all individuals in different countries is, in some way, supported by the empirical literature showing that determinants of subjective wellbeing are quite similar across different countries and time periods (Becchetti *et al.*, 2010).

[7] In the context of attitudinal surveys where observed responses are often discrete, the disjunction between what is observed and the underlying latent measurement error in the dependent variable is generally understood as arising from an error in either recording or reporting of a response. Corrado and Weeks (2010) analyse different solution methods to correct for response scale heterogeneity when responses are discrete.

[8] If measurement error affects one or more explanatory variables, this will generate biased and inconsistent parameter estimates, with a general tendency towards attenuation. Kreider (1999) discusses the problem of measurement error for self-reported health and in particular work disability in the context of models of labour force participation. However, the focus here is the impact of likely overreporting of disability on parameter estimates associated with one or more explanatory variables whereas our focus is on measurement error affecting the dependent variable.

[9] Life satisfaction is ordinal, so that its panel estimation would require something like a ordered probit or conditional fixed effect logit (as in Clark, 2003). However, as Ferrer–i–Carbonell and Frijters (2004) argue, Cardinal estimation seems to perform just as well as ordinal estimation when life satisfaction is measured on the 0-10 scale (Ferrer–i–Carbonell and Frijters, 2004).

As stressed by Bound *et al.* (2001), in presence of exogenous determinants of the error ridden variables[10] or, in some cases, multiple indicators of the same outcome, it is possible to use them as instruments to infer the "true" value of life satisfaction. Hence, using $\mathbf{Z}_i$ as the set of instruments for LS_i^* we assume that $E(LS_i^* \mid \mathbf{Z}_i^*)$ is a strict linear function of $\mathbf{Z}_i^*$ so that these multiple indicators are orthogonal to the error implying $Cov(\mathbf{Z}_i^*, \tilde{v}_i) = 0$ and there is no measurement error, hence $E(\tilde{v}_i) = 0$.

The measurement error term of the synthetic question LS_i does not go to zero as far as the number of interviewed individuals grows implying $E(v_i) \neq 0$ for at least three reasons. First, the life satisfaction term is more abstract than a straightforward question on its components (vitality, evaluation of part life etc.). Second, it requires a quick calculus of (1) and of the weights of the individual $\mathbf{Z}_i$ components which is not easy and intuitive. It is for instance highly likely that cultural differences affect the more abstract life satisfaction question, while they cancel out when using the more straightforward set of $\mathbf{Z}_i$ questions. Finally, the sub-component questions are much more straightforward and easy to answer when they are formulated on a 1-4 range since in this instance any number is associated to an adjective whose meaning can be grasped immediately. On the contrary, in the 1-10 scale life satisfaction question there is no verbal correspondence for each of the scale values. This is why the error term, v_i, may be significantly different from zero and different across countries when we adopt the synthetic question on life satisfaction. Hence, a more articulated set of questions on the set of $\mathbf{Z}_i$ components may produce a much richer and accurate measure of cognitive subjective wellbeing.

To understand the implications of measurement errors when we adopt a synthetic answer on life satisfaction suppose that we estimate the following relationship for the true level of life satisfaction:

$$LS_i^* = \beta_0 + \boldsymbol{\theta}' \mathbf{d}_i + \boldsymbol{\beta}' \mathbf{X}_i + \xi_i \tag{5}$$

$$i = 1, ..., N \quad c = 1, ..., C-1 \tag{6}$$

where β_0 denotes a constant, $\mathbf{d}_i = \{d_{ic}\}$ is a $(C-1) \times 1$ vector of country dummies, with $d_{ic} = 1$ if individual i is resident in country c and $\boldsymbol{\theta}$ is the corresponding $(C-1) \times 1$ vector of coefficients. $\mathbf{X}_i = \{X_{ik}\}$ with $k = 1, ..., K$ denotes a $(K-1) \times 1$ vector of controls and $\boldsymbol{\beta}$ is the corresponding $K \times 1$ vector of coefficients. Finally, ξ_i is an individual error term which is normally distributed with $E(\xi_i) = 0$.

If the synthetic answer for life satisfaction is observed with error we get:

$$LS_i = \beta_0 + \boldsymbol{\theta}' \mathbf{d}_i + \boldsymbol{\beta}' \mathbf{X}_i + \xi_i + v_i \tag{7}$$

Notice that the error term is now $e_i = \xi_i + v_i$. This means that if we estimate the model, we would see the error in measurement appearing in the new error term. If the measurement error is correlated with the independent variables this also implies a violation of the assumption that the conditional expectation of the error should be zero.

By the OLS assumptions ξ_i should be uncorrelated with the covariates and consequently should not present any problem.[11] If the measurement error, v_i, is independent from the regressors $Cov(\mathbf{X}_i, v_i) = 0$ but there is measurement error, $E(v_i) \neq 0$, the estimate of the coefficients will still be unbiased but measured with less precision. However, if both $E(v_i) \neq 0$ and $Cov(\mathbf{X}_i, v_i) \neq 0$, the endogeneity between v_i and $\mathbf{X}_i$ induced by the measurement error implies that the estimates of the coefficients may also be biased.

The other main problem comes from inference since $var(\xi_i + v_i) = \sigma_\xi^2 + \sigma_v^2 > \sigma_\xi^2$. The last inequality means that the estimated variance is larger than when using the true life satisfaction measure, which means that inference is liable to type I error. Collecting more data in the form of multiple components for the dependent variable could be a solution also to this problem since more observations imply a better estimator of variance, and consequently reduces errors in inferences.

[10] All the above mentioned methods involve introducing external information. A number of authors have suggested instrumental variable estimators that use third or higher moments of the variables as instruments for X or Y (Cragg, 1997; Dagenais and Dagenais, 1997; Lewbel, 1997). Alternatively, Wald (1940) suggested an estimator which involves grouping the data. However, unless some external information can be used to form groups (i.e. an instrument) is available, the resulting estimator will typically be no less biased than OLS (Pakes, 1982).

[11] In a non-linear regression model, such as a probit model, the effects of measurement error are more severe. If the dependent variable is binary, measurement error takes the form of misclassification errors; some observations where the variable is truly a 1 will be misclassified as a 0 and vice versa. In this case the measurement error is negatively correlated with the true variable. This can lead to coefficient estimates that are biased and inconsistent.

In this instance, the SHARE questions allow us to measure some fundamental components of the life satisfaction evaluation such as: vitality, a negative evaluation of the past life, a positive look at the future, absence of lack of monetary constraints, sensation that life is meaningless, feeling in control of one's own life, sense of not being left out, perception of having time beyond family duties, freedom of choice. As it is clear from these attributes, the components of life satisfaction in the survey include an outlook on the past and on the future, money and leisure satisfaction, vitality and control over one's own life plus an eudaemonic definition of life satisfaction (meaning of life). This is why they can be considered quite richer than the simple standard cognitive synthetic question. An additional advantage of the question on sub-components is that a unique and different verbal modality is attached to each numerical value of the ordinal scale. This makes the answer easier and more intuitive and may help to counteract the measurement error, v_i.

What we propose is the use of the *estimated* latent life satisfaction $\widehat{LS}_i^*$ based on the following specification:

$$LS_i^* = \alpha_0 + \boldsymbol{\alpha}^{'}\mathbf{Z}_i + \eta_i \tag{8}$$

where $\boldsymbol{\alpha}$ is a $J \times 1$ vector of coefficients. Here, the use of multiple components allows to counteract the measurement error in the sub-components as defined in (4) implying $E(\eta_i) = 0$. We use the predicted value of life satisfaction:

$$\widehat{LS}_i^* = \hat{\alpha}_0 + \hat{\boldsymbol{\alpha}}^{'}\mathbf{Z}_i \tag{9}$$

as the dependent variable of our benchmark estimate in (5) which is a more correct measure of LS_i^* than the observed synthetic question $LS_i = LS_i^* + v_i$ under the assumption that the weights in (7) are correctly estimated i.e. $\widehat{\alpha}_{ij} = \omega_{ij}\ \forall i, j$.

Alternatively, our second approach consists of simply averaging the sub-components to obtain:

$$\overline{LS}_i = \frac{\sum_{j=1}^{J} Z_{ij}}{J}$$

which is equal to the previous measure under the assumption that $\omega_{ij} = 1$. The limit of this second approach is that it is contradicted by empirical evidence in case the ω-weights are different from one. The advantage is that it has less noise if, for some reasons (i.e. endogeneity), we think that the estimated weights do not coincide with actual weights ($\widehat{\alpha}_{ij} \neq \omega_{ij}$).

A third approach which may overcome the problem of overlaps and correlations among the different components of life satisfaction is the principal component approach. On the basis of the approach we will extract the first orthogonal factor accounting for the higher share of the variance and use such variable as dependent variable in our estimate (see section 3).

2.1 Hypothesis testing

Given our assumptions a first testable hypothesis is that the goodness of fit of the model using the latent variable $\widehat{LS}_i^*$ is better than that of the model using the declared life satisfaction LS_i. In fact, in presence of measurement error in a model for declared life satisfaction, LS_i, the variance of residual grows, implying that both the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) will be higher, while the adjusted R^2 will be lower, indicating a poorer performance.[12]

A second testable hypothesis related to country dummies is that our approach should be effective in reducing country bias (measured by the significance of country dummies in vignette responses) in the expected direction.

[12]The Akaike Information Criterion (AIC) is defined as $AIC = n\ln(\frac{RSS}{n}) + 2k$ and the Bayesian Incformation Criterion (BIC) as $BIC = n\ln(\frac{RSS}{n}) + n\ln(k)$ where n denotes the number of observations, k is the number of parameters and RSS denotes the Residual Sum of Squares. In presence of measurement error the Residual Sum of Squares will be higher, hence both the AIC and BIC will be higher indicating a poorer fit of the model. The adjusted R^2 defined as $\overline{R}^2 = (1 - \frac{RSS}{TSS})\frac{n-1}{n-k-1}$ will be, instead, lower.

3 Descriptive finding and benchmark model specification

We use data from the second wave of SHARE including interviews run between 2006 and 2007[13] which is the only wave including the crucial information on the eleven happiness sub-components. Table 1 provides the list of variables, Figure 1 reports the histogram of self-reported life satisfaction, while Table 2 illustrates some descriptive findings of our sample.

Self-reported life satisfaction has the usual right skewed distribution with a mean of 7.54 and around 60 percent of respondents declaring a self-reported life satisfaction above 6 (Figure 1). All answers to the 11 life satisfaction sub-questions have averages between 2.5 and 3.5 (the range is 1-4) with the lowest average for the sub-components related to age preventing activities and lack of money (items 1 and 5 of the 12 sub-questions, see introduction) (Table 2). The sample is almost perfectly balanced in terms of gender characteristics (females are 49.9 percent), while the average years in education are 10.5. Half of the sample is between 50 and 60 year old. Average household size is 2.25 (Table 2).

Our baseline estimate of life satisfaction is

$$LS_i = \beta_0 + \boldsymbol{\theta}^{'}\mathbf{d}_i + \boldsymbol{\beta}^{'}\mathbf{X}_i + e_i \tag{10}$$

As it is well known the dependent variable (self reported life satisfaction) is ordinal, so that its estimation would require something like an ordered logit or probit (as in Clark, 2003). However, as Ferrer–i–Carbonell and Frijters (2004) argue, cardinal estimation seems to perform just as well as ordinal estimation in this context.[14]

Regressors are those standard in this literature and include a gender dummy (taking value one if the respondent is male), years of education, household size, the number of children and the number of grandchildren. Marital status is measured by five dummies (married, divorced, separated, widowed, living with regular partner) with single status as the reference category. Age is controlled for with six age class dummies with age 50-55 as the reference category. Four dummies (big city, large town, small town and suburbs with rural as the reference category) capture characteristics of the place of residence.

The SHARE database gives us the opportunity of controlling for a large number of health factors. Such factors concern various kinds of physical disabilities and a number of reported illnesses. We measure them synthetically with three variables which sum many specific single items in the three domains (see Table 1 for details). We finally add a set of variable measuring voluntary work, religious attendance, participation to sport and social activity, helping in families and inheritance leaving.

As it is well known the SHARE database has a very high number of individuals who refuse to report income and many missing values for other important variables. We therefore follow an approach which is standard in previous empirical studies on this dataset by using Christelis' data on imputed gross total household income included in the Share database[15] and calculated following the Fully Conditional Specification methods (FCS) of Van Buuren *et al.* (2006). The imputations[16] are country specific in the sense that they are made separately for each country[17] and the sample is representative of the population aged fifty and above. The main scope of this procedure is to generate the distribution of the missing value of a specific variable conditional on the value of the observed values of other non-missing variables in the dataset. The SHARE database provides data obtained with this procedure by creating five imputed datasets. We therefore end up having five different values (one for each iteration) of

[13] Last Release 2.5.0: May 24th, 2011 available at http://cdata8.uvt.nl/sharedatadissemination/releases/show/w2_250/All+CAPI+modules/stata.

[14] See also Ferrer-i-Carbonell (2004 and 2008) on this point.

[15] The dataset used is "sharew2_rel2-5-0_imputations".

[16] A key aspect of the FCS method is that it operates under the missing at random (MAR) assumption where the missingness of each variable depends only on other variables in the system and not on the values of the variable itself. In the iteration process, the initial conditions of the first iteration are derived by imputing the first variable in the system based only on the variables that are never missing (age, gender and geographic location), then the variables in the second iteration are calculated based on the first and the non-missing variables, in order to achieve a complete set of values for these initial conditions. In this calculation the fully imputed demographic variables are used as predictors for the economic variables; in the imputation of a specific wave, large part of information that comes from other waves is taken into account. The imputation in SHARE also allows an initial burn-in period in order to decrease the dependence of the chain on the initial values. Five burn-in iterations are used by evaluating the Gelman-Rubin criterion from the seventh iteration on. For more details see Christelis (2011).

[17] Ireland is excluded from this procedure.

the imputed variables. In what follows we propose estimates using just one of them while performing robustness checks using the other four iterated values. Variables with imputed variables in our specification are the log of household income and a number of other variables characterised by item non-response: the number of children, the number of grandchildren, the number of rooms in the main residence, whether the respondent lives in a big city, in suburbs/outskirts of a big city, in a large town, in a small town or in a rural area or village.[18]

3.1 Econometric findings

In the benchmark estimate where the dependent variable is the 0-10 standard life satisfaction question the adjusted R-squared is .217 and the AIC and BIC criteria are equal to 112,824.6 and 112,924.4 respectively (Table 3, column 1). The model includes many controls but the Variance Inflation Factor (VIF) shows that there are no significant problems of multicollinearity.[19] The log of imputed household income and education years are positive and significant as expected. Consistently with the empirical literature we find that being married and living with a partner increase significantly life satisfaction with respect to the single status. The household size has a negative sign proxying presumably for the impact of household size on the individual portion of gross total household income. Living in big cities impacts positively while the age dummy effects grow with age.[20] The number of grandchildren affects positively self declared life satisfaction while the number of children does not. All of the three variables indicating health problems are negative and significant while those measuring social life (participation to sport, helping those in the family, voluntary work) are mostly positive and significant. Inheritance transmission during life is also positively associated to our dependent variable. Most country dummies are significant and are expected to include two components: country specific omitted variables affecting life satisfaction (such as climate, institutions and cultural effects) and heterogeneity in life satisfaction scales (country bias). We will try to identify country bias in what follows by testing whether country dummies are significant when the dependent variable is the respondents' evaluation of the same vignette.

In order to produce estimates from the first alternative approach we estimate the latent life satisfaction factor with the following specification

$$LS_i^* = \alpha_0 + \boldsymbol{\alpha}^{'}\mathbf{Z}_i + \eta_i \tag{11}$$

in which the standard life satisfaction variable is regressed on the eleven life satisfaction sub-components (the Z-variables) described in the introduction and $\boldsymbol{\alpha} = \{\alpha_j\}$ with $j = 1, ..., 11$ denotes the corresponding 11×1 vector of coefficients. Note that the correlation matrix of the different happiness components displays a maximum correlation between the *future good* and *opportunities* variables (around .63). Other strong correlations are between *vitality* and, respectively, *future good* (around .54) and *opportunities* (56 percent) (Table 4).

All the regressors are strongly significant as expected and the adjusted R-squared is around 39 percent (Table 5, column 1). The VIF shows that there are no multicollinearity problems in this estimate. The most important component is future perspectives (*future good*) but also the evaluation of the past (*past good*) is strongly significant thereby confirming that life satisfaction is the product of a weighted average of different sub-components including an evaluation of the present, the past and future life perspectives. To give an idea of the magnitude of these effects, when the model is re-estimated as an ordered logit a unit increase in future perspectives adds a 3.3 percent to the likelihood of declaring the highest level of life satisfaction while positive evaluation of the past only 2.4 percent (Table 5, column 3).

The predicted value of the regression in (2):

$$\widehat{LS}_i^* = \widehat{\alpha}_0 + \widehat{\boldsymbol{\alpha}}^{'}\mathbf{Z}_i \tag{12}$$

is then used as dependent variable in the baseline model in (1) which becomes

[18] The imputed datasets are available from SHARE at http://cdata8.uvt.nl/sharedatadissemination/releases/show/w2_250/Generated+Variables/Imputations/stata.

[19] As it is well known the VIF (variance inflation factor) formula is $1/1 - R(x)$ where $R(x)$ is the R-squared obtained by regressing each independent variable on all other independent variables (Marquardt, 1970). If $R(x)$ is low (tends to zero) the VIF test is low (equal to one).

[20] Since our sample is made by people aged above 50 this apparently surprising result may capture the ascending part of the U-shaped relationship between age and happiness (see among others Clark et al., 1996 and Frijters and Beatton, 2008).

$$\widehat{LS}_i^* = \beta_0 + \boldsymbol{\theta}' \mathbf{d}_i + \boldsymbol{\beta}' \mathbf{X}_i + \tilde{e}_i \tag{13}$$

The baseline model with the modified dependent variable has a much better goodness of fit (from .217 to .342). The AIC and BIC (76,647.54 and 76,747.01) are also considerably improved (Table 3, column 2).

When comparing the sign and significance of the regressors between standard and alternative models we find that: *i*) life satisfaction is not increasing with age anymore; *ii*) the magnitude of income and the significant marital status variables (married and with regular partner) is reduced even though the regressors remain significant; *iii*) house size becomes significant; *iv*) the significance of geographical dummies changes. Magnitudes and signs of all the other variables remain remarkably stable.

The first alternative method has several limits. First, it still uses in the first stage the dependent variable whose limit we want to overcome. Second, the estimated coefficients in the regression used to calculate the predicted latent life satisfaction variable may be biased by omitted variables, endogeneity or multicollinearity (even though we documented that the last problem is not severe).

We therefore test the robustness of our theoretical hypotheses with other two alternatives. The first is a simple unweighted average of the life satisfaction sub-components. We are aware that, in this way, we overcome the limit of the latent life satisfaction estimate even though, by using an unweighted average, we assume quite restrictively that the different sub-components have unit weights.

The estimate of the baseline model with the sub-components unweighted average dependent variable provides a goodness of fit which is very close to that of our benchmark alternative approach in terms of R-squared (34.2 percent), while improving further in terms of AIC and BIC (35,813.61 and 35,813.49 respectively) (Table 3, column 3).

The other alternative which avoids the arbitrary choice of equal weights is the extraction of a principal component from the life satisfaction sub-components. The approach has the additional advantage of correcting for correlation and potential multicollinearity among different life satisfaction sub-components (i.e. the answer to the meaning of one's own life may be correlated with feeling in control, not feeling left out, having a good perspective on the future, etc.). The principal component analysis documents that the first extracted component accounts for 37 percent of the variability of the selected variable. The first component has its strongest correlation with the sub-questions about future perspectives (.38), life opportunities (.37) and vitality (.36) (Tables 6 and 7). The Kaiser-Meyer-Olin measure of sampling adequacy (.76) (Kaiser and Rice, 1974) excludes that the selected variables have too little in common to warrant a factor analysis.

When using the first principal component as dependent variable (our third alternative method) we find that the goodness of fit is around .35 (Table 3, column 4) with significance and signs of regressors being very close to those of the two previous approaches.

The comparison of the goodness of fit among the standard model and our three alternatives in terms of AIC and BIC values tells us that the best model is the one in which the dependent variable is the unweighted average of sub-components followed by the one in which we use the predicted life satisfaction estimated on the eleven sub-components. The ranking of the models in terms of adjusted R-squared is, however, different since all of the three models are very close and outperform by far the standard one with the unweighted average doing slightly worse. The reason for the different ranking is that the unweighted average model has by far the smallest residual sum of squares (which is the crucial factor for AIC and BIC), but also a much smaller total sum of squares (which is the factor on which progress in goodness of fit is scaled for when using adjusted R-squared).

Finally, we aim to check whether our approach can correct for the country bias. We tackle this issue in the most conservative and simple way. We first average the values of the two life satisfaction vignettes included in the SHARE[21] and then regress the variable on the country dummies with/without socio-demographic controls. France is the omitted reference country. The Danish dummy is the highest in magnitude and significance (around .627,

[21]The first vignette is *"John is 63 years old. His wife died 2 years ago and he still spends a lot of time thinking of her. He has 4 children and 14 grandchildren who visit him regularly. John can make ends meet but cannot make for extra such as expensive gifts for his grandchildren. He has had to stop working recently due to heart problems. He gets tired easily. Otherwise he has not serious health conditions."* The second vignette is *"Carry is 72 years old and a widow. Her total after tax income is about € 1,100 per month. She owns the house she lives in and has a large circle of friends. She plays bridge twice a week and goes on vacation regularly with some friends. Lately she has been suffering from arthritis, which makes working in the house and garden painful."*

t-stat 14.08), followed the Czech dummy (.474, t-stat 10.50) and the German dummy (.332, t-stat 8.42) indicating that respondents from these three countries overevaluate the common vignette situations in terms of life satisfaction *vis à vis* the French who are the reference category. The Danish dummy result is consistent with what found in the vignette literature as commented in the introduction (Inglehart and Klingemann, 2000, Eurobarometer 2002; Corrado and Weeks, 2010; Kapteyn, Smith, and Soest, 2009). This gives us enough confidence on the fact that a cultural bias exists at least for this country.

We therefore check whether our three approaches correct country biases in the expected direction. The inspection of the country dummies in the first column of Table 3 (standard life satisfaction estimate) compared with those in the other three columns of the same Table 3 (our three alternative approaches) documents that all of the three approaches correct the Danish effect in the desired direction. The Danish dummy is in fact .798 in the standard life satisfaction estimate. It falls to .321 under our first alternative method (life satisfaction predicted on the eleven sub-components), to .135 when using the second alternative method (unweighted average of sub-components) and to .454 when using the third alternative method (principal component analysis). Confidence intervals of the Danish dummy from the three alternative methods do not overlap with those of the standard life satisfaction estimate. Note as well that both the Czech and German dummies are corrected in the expected direction (reduction of the positive magnitude) in five out of six cases by the three alternative approaches.

3.2 Robustness check and discussion

We perform several robustness checks to control whether our main findings are robust to perturbations of the benchmark model. First of all we want to control their sensitivity to the imputation variables. We therefore report for simplicity only goodness of fit statistics (and not full regression estimates) considering imputed variable values from the other four iterations. The results are very close to those of the first iteration, consistently with what found in the literature using the same data (Table 8).

Since the number of observations in the second model is slightly lower than that in the first model due to some missing values on the life satisfaction sub-component questions which do not match with missing values on baseline regressors (21,680 against 22,494), we repeat the first estimate with exactly the same valid observations of the second and find that our conclusions remain unchanged (Table 9).

In a third robustness check to our first approach we re-estimate the latent life satisfaction factor by assuming that the impact of the eleven sub-components is not the same according to different countries or crucial sociodemographic factors. More specifically, we interact the sub-components with all country dummies, age classes and gender according to the following specification.

$$LS_i^* = \alpha_0 + \sum_{c=1}^{C} d_{ic}\boldsymbol{\alpha}_c^{'}\mathbf{Z}_i + \sum_{a=1}^{A} d_{ia}\boldsymbol{\alpha}_a^{'}\mathbf{Z}_i + d_{ig}\boldsymbol{\alpha}_g^{'}\mathbf{Z}_i + \tilde{\eta}_i \tag{14}$$

where d_{ic} $(c = 1, ..., C)$ denote the country dummies, d_{ia} $(a = 1, ..., A)$ denote the age dummies and d_{ig} denotes the gender dummy. The goodness of fit of the estimate jumps to .40 (highest among all models) even though the AIC is almost unchanged with respect to Table 3 column 2 where we use predicted life satisfaction from (2) (Table 11). This indicates that there is a clear trade-off with the capacity of correcting country dummies in the right direction. The result is consistent with the fact that overparametrisation improves goodness of fit (as it generally does) at the cost of creating noise on the coefficient values since values of the predicted life satisfaction component used as dependent variable are affected by many insignificant interacted variables (while the 11 sub-components in the simple model in Table 5 are all significant).

In another robustness check we re-estimate the benchmark model from Table 3 excluding the health variables which may be suspected of endogeneity. Last, we eliminate from our specification all the variables imputed with the Cristelis *et al.* (2011) approach to check whether our findings are sensitive to such imputation. Our main findings are robust to these changes (Tables 12 and 13).

Consider, finally, that a peculiarity of our work is that we are comparing models in terms of alternative dependent variables and not, as it usually occurs, nested or non nested models on the basis of differences in the considered set of regressors. An observationally equivalent interpretation of our findings could therefore be that the selected regressors are spurious and that the "true" set of life satisfaction determinants could, in principle, yield

a superior goodness of fit when using the standard life satisfaction dependent variable. In such case the superior goodness of fit of the alternative model does not tell *per se* that our alternative dependent variables capture better factors affecting life satisfaction.

What could be said against this interpretation is that we use regressors (marital status, income, gender, education, etc.) which are standard in the life satisfaction literature. Furthermore, their sign and magnitude does not vary much between the standard and the alternative models. Hence, it is much more reasonable to assume that the considered regressors are the true observable determinants of life satisfaction and that our alternative dependent variable can be measured with less bias and noise than the standard one as we assume in our theoretical framework. Last but not least, we demonstrate that the superior goodness of fit of the alternative models is robust to several changes in the set of regressors. On such basis is hard to imagine an alternative set of observable and "true" life satisfaction determinants that we did not consider which could justify the observationally equivalent interpretation of our result which we mention above.

4 Conclusion

The standard life satisfaction question used in surveys is likely to suffer from serious problems of abstraction, complexity of calculus and cultural bias. Abstraction depends on the fact that its 0-10 scale prevents intuitive correspondence with verbal modalities. Complexity of calculus originates from the problem that the overall life satisfaction evaluation is implicitly derived from a weighted sum of sub-components affecting it (i.e. money satisfaction, sense of life, outlook at the past, perspectives on the future, vitality, etc.). Cultural bias depends on the fact that different linguistic nuances in the meaning of the term may enhance differences in answers across individuals from different countries which do not depend on true differences in life satisfaction.

The point we raise in our paper is that the richness of direct and simpler information on the life satisfaction subquestions (answers on 1-4 scale on each item with correspondence between an adjective and each numerical value) may significantly reduce these three problems thereby improving goodness of fit and reducing the noise component of country dummies. We articulate our alternative strategy under three different approaches (estimation of the latent life satisfaction regressing the standard life satisfaction variable on the above mentioned sub-components, simple unweighted average of the subcomponents, extraction of the first principal component among the subcomponents with principal component analysis).

Our findings do not reject the above mentioned hypotheses. The goodness of fit is greatly enhanced under all of the three alternative approaches. The well known Danish cultural bias, which we find also in our data consistently with similar findings in the vignette literature, is corrected in the desired direction by all of our three approaches.

What our results suggest is that the use of a small set of less abstract and comprehensive life satisfaction subquestions increases the share of subjective wellbeing which is accounted for by observable life events. Since this improvement can be obtained by merely adding one demand to standard surveys (hence a reasonable cost more than compensated by the documented benefits) our straightforward political advice is that new life satisfaction surveys should all contain such demand. The suggestion we make is very close to what is currently done to calculate a (mental) wellbeing index (the General Health Questionnaire score) which has been used in alternative to self-declared life satisfaction as a proxy of the subjective wellbeing (Golderberg and Williams, 1988). The index is the unweighted average of 12 mental distress questions and therefore closely follows one of our alternative approaches.

An important element which should be taken into account is that our findings are obtained on a database (SHARE) which includes only individuals aged from 50 and above. Future research should verify whether our findings are equally valid when younger age cohorts are included. This will however not be possible until the additional question on life satisfaction subcomponents is added. Additional reflection should be made on whether the range of questions applied to the 50+ sample are also applicable to the younger cohorts or whether a different set of questions should be considered.

References

[1] 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.

[2] Becchetti, L., Massari, R. and P. Naticchioni (2010): Why has Happiness Inequality Increased? Suggestions for Promoting Social Cohesion. mimeo.

[3] Bound, J., Brown, C. and N. Mathiowetz (2001): Measurement Error in Survey Data, in J. Heckman and E. Leamer, eds, *Handbook of Econometrics*,Vol. 5.

[4] Clark, A. E., Oswald, A.J. and P.B. Warr (1996): Is Job Satisfaction U-shaped in Age?, *Journal of Occupational and Organizational Psychology,* 69, pp. 57-81.

[5] Corrado, L. and M. Weeks (2010): Identification Strategies in Survey Response Using Vignettes. Cambridge Working Papers in Economics 1031, Faculty of Economics, University of Cambridge.

[6] Christelis D. (2011): Imputation of Missing Data in Waves 1 and 2 of SHARE, SHARE Working Paper Series 01-2011.

[7] Cragg, J. G. (1997): Using Higher Moments to Estimate the Simple Errors-in-Variables Model, *RAND Journal of Economics*, 28, pp. 571-91.

[8] Dagenais, M. G. and D. L. Dagenais (1997): Higher Moment Estimators for Linear Regression Models with Errors in the Variables, *Journal of Econometrics*, 76, pp. 193-221.

[9] Eurobarometer (2002): Public Opinion in the European Union. Available at http://europa.eu.int/comm/publicopinion=archives=eb=eb57=eb57en:pdf.

[10] Frey, B. S. and A. Stutzer (2000): Happiness, Economy and Institutions, *The Economic Journal*, 110, pp. 918-938.

[11] Frey, B. and A. Stutzer (2002a): What can Economists Learn from Happiness Research, *Journal of Economic Literature*, 40, pp. 402-435.

[12] Frey. B. S. and A. Stutzer (2002b): *Happiness and Economics. How the Economy and Institutions Affect Well-being*, Princeton University Press. Princeton. New Jersey.

[13] Frey, B.S., Luechinger, S. and A. Stutzer (2009): The Life Satisfaction Approach to the Value of Public Goods: The Case of Terrorism, *Public Choice*, 138, pp.317-345.

[14] Ferrer-i-Carbonell, A. and P. Frijters (2004): How Important is Methodology for the Estimates of the Determinants of Happiness? *The Economic Journal*, 114. pp. 641-659.

[15] Ferrer-i-Carbonell, A., Van Praag, B.M.S. and I. Theodossiou (2011): Vignette Equivalence and Response Consistency: The Case of Job Satisfaction, IZA DP No. 6174.

[16] Frijters, P. and T. Beatton (2008): The Mystery of the U-shaped Relationship between Happiness and Age. NCER Working Paper Series 26, National Centre for Econometric Research.

[17] Golderberg, D. and P. Williams (1988): A User's guide to the General Health Questionnaire. Windsor, UK: NFER-Nelson.

[18] Hu, Y.J., Stewart-Brown, S., Twig, L. and S. Weich (2007): Can the 12-item General Health Questionnaire be Used to Measure Positive Mental Health?, *Psychological Medicine*, 37(7), pp. 1005-1013.

[19] Inglehart, R. F. and H.-D. Klingemann (2000): Subjective Well-Being Across Cultures, in *Genes, Culture, Democracy and Happiness*, pp. 165-184. MIT Press.

[20] Kapteyn, A., J.P. Smith and A. Van Soest (2007): Vignettes and Self-Reports of Work Disability in the United States and the Netherlands, *American Economic Review*, 97(1), 461-472.

[21] Kaiser, H. and J. Rice (1974): Little Jiffy, Mark IV, *Journal of Educational and Psychological Measurement,* 34, pp. 111-117.

[22] King, G. and J. Wand (2007): Comparing Incomparable Survey Responses: New Tools for Anchoring Vignettes, *Political Analysis*, 15, pp. 46–66.

[23] Lewbel, A. (1997): Constructing Instruments for Regressions with Measurement Error when no Additional Data are Available, with an Application to Patents and R&D, *Econometric*a, 65, pp. 1201-1213.

[24] Luechinger, S. and P. Raschky (2009): Valuing Flood Disasters using the Life Satisfaction Approach, *Journal of Public Economics*, Elsevier, 93(3-4), pp. 620-633.

[25] Marquardt, D.W (1970): Generalized Inverses, Ridge Regression, Biased Linear Estimation, and Nonlinear Estimation, *Technometrics*, 12, pp. 591-612.

[26] Pakes, A. (1982): On the Asymptotic Bias of Wald-type Estimators of a Straight Line when both Variables are Subject to Error, *International Economic Review*, 23, pp. 491-497.

[27] SHARE (2011): Survey of Health, Ageing and Retirement in Europe, Guide to Release 2.5.0 Waves 1 and 2. Mannheim Research Institute for the Economics of Ageing.

[28] Van Buuren, S., Brand, J.P.L., Groothuis-Oudshoorn, C.G.M. and D.B. Rubin (2006): Fully Conditional Specification in Multivariate Imputation, *Journal of Statistical Computation and Simulation*, 76(12), pp. 1049-1064.

[29] Van Praag, B.M.S. and B.E. Baarsma (2005): Using Happiness Surveys to Value Intangibles: the Case of Airport Noise, *The Economic Journal*, 115, pp. 224–246.

[30] Wald, A. (1940): The Fitting of Straight Lines if both Variables are Subject to Error, *Annals of Mathematical Statistics*, 11, pp. 284-300.

[31] Welsch, H. (2002): Preferences over Prosperity and Pollution: Environmental Valuation Based on Happiness Surveys, *Kyklos*, 55(4), pp. 473-494.

Table 1: **VARIABLE LEGEND**

VARIABLE	DESCRIPTION
Female	*Dummy var. =1 if respondent is female; =0 otherwise.*
Log income	*Log of household total gross income. Its value is equal to the sum over all household members of the individual-level values of: annual net income from employment and self-employment (in the previous year);* *Annual public old age/early or pre-retirement/disability pension (or sickness benefits); Annual public unemployment benefit or insurance, public survivor pension from partner; Annual war pension, private (occupational) old age/early retirement/disability pension, private(occupational) survivor pension from partner's job, public old age supplementary pension/public old age/public disability second pension, secondary public survivor pension from spouse or partner, occupational old age pension from a second and third job; Annual public and private long-term insurance payments; Annual life insurance payment, private annuity or private personal pension, private health insurance payment,alimony, payments from charities received; Income from rent. Values of the following household level variables are added: Annual other hhd members'net income; Annual other hhd members'net income from other sources; Household bank accounts, government and corporate bonds, stocks/shares; mutual funds. (imputed as in Christelis, 2011)*
Education years	*Years the respondent has been in full time education.*
Household size	*Household size.*
Age class	*Respondent's age class: =1 if respondent's age<55; =2 if resp.'s age=[55,59]; =3 if resp.'s age=[60,64]; =4 if resp.'s age=[64,69]; =5 if resp.'s age =[69,74]; =6 if resp.'s age =[74,79]; = 7 if age>79.*
Leaving inheritance	*Respondent's answer to the question: including property and other valuables,what are the chances that you or your husband/wife/partner will leave an inheritance totaling 50,000 euro or more? The possible answers range from 0 to 100.*
Married	*Dummy =1 if the respondent lives with spouse.*
Registered partnership	*Dummy =1 if the respondent lives with a partner.*
Widowed	*Dummy =1 if the spouse is died.*
Divorced	*Dummy =1 if respondent is divorced.*
Separated	*Dummy =1 if the respondent lives separated from spouse.*
Single	*Dummy =1 if the respondent lives as a single.*
N.of children	*Respondent's number of children (imputed as in Christelis, 2011).*
N.of grandchildren	*Respondent's number of grandchildren (imputed as in Christelis,2011).*
Hrooms	*Number of rooms in the main residence (imputed as in Christelis, 2011).*
Big city	*Dummy =1 if the respondent lives in a big city (imputed as in Christelis, 2011).*
Suburbs	*Dummy =1 if the respondent lives in suburbs/outskirts of a big city (imputed as in Christelis, 2011).*
Large town	*Dummy var.=1 if the respondent lives in a large town (imputed as in Christelis, 2011).*
Small town	*Dummy =1 if the respondent lives in a small town (imputed as in Christelis, 2011).*
Rural area	*Dummy =1 if the respondent lives in a rural area or village (imputed as in Christelis,2011).*
Long-term illness	*Dummy =1 if the respondent declares any long-term health problems, illness, disability or infirmity. Survey question: some people suffer from chronic or long-term health problems. By long-term we mean it has troubled you over a period of time or is likely to affect you over a period of time. Do you have any long-term health problems, illness, disability or infirmity?*
No limited activities	*Dummy =1 if the respondent has not been limited because of a health problem in activities people usually do. Survey question: for the past six months at least,to what extent have you been limited because of a health problem in activities people usually do?*
Numb illnesses	*It is the sum of illnesses the respondent is currently being treated for or bothered (A heart attack including myocardial infarction or coronary thrombosis or any other heart problem including congestive heart failure; high blood pressure or hypertension; high blood cholesterol; a stroke or cerebral vascular disease diabetes or high blood sugar; chronic lung disease such as chronic bronchitis or emphysema; asthma; arthritis, including osteoarthritis, or rheumatism; osteoporosis; cancer or malignant tumor, including leukaemia or lymphoma, but excluding minor skin cancer; stomach or duodenal ulcer, peptic ulcer; Parkinson disease; cataracts; hip fracture or femoral fracture; Alzheimer disease, dementia, organic brain syndrome, senility or any other serious memory impairment; benign tumor).*
Life satisfaction	*Respondent degree of life satisfaction. Survey question: On a scale from 0 to 10 where 0 means completely dissatisfied and 10 means completely satisfied, how satisfied are you with your life?*
Age no prevent	*Respondent degree of statements that have used to describe their lives or how they feel. Survey question: How often do you think your age prevents from doing the things you would like to do ? For each item answers are given on a 1-4 scale where an adjective (often, sometimes, rarely, never) is matched to any value.*
No out control	*Respondent degree of statements that have used to describe their lives or how they feel. Survey question: How often do you feel that what happens to you is out of control? For each item answers are given on a 1-4 scale where an adjective (often, sometimes, rarely, never) is matched to any value.*
No feel left out	*Respondent degree of statements that have used to describe their lives or how they feel. Survey question: How often do you feel left out of things? For each item answers are given on a 1-4 scale where an adjective (often, sometimes, rarely, never) is matched to any value.*
Fred. choice	*Respondent degree of statements that have used to describe their lives or how they feel. Survey question: How often do you feel that you can do the things that you want to do? For each item answers are given on a 1-4 scale where an adjective (often, sometimes, rarely, never) is matched to any value.*
No fam.responsibility	*Respondent degree of statements that have used to describe their lives or how they feel. Survey question: How often do you feel that family responsibilities prevent you from doing what you*

(Continued on next page)

(Continued from previous page)

	want to do?. For each item answers are given on a 1-4 scale where an adjective (often, sometimes, rarely, never) is matched to any value.
No lack money	*Respondent degree of statements that have used to describe their lives or how they feel. Survey question: How often do you feel that shortage of money stops you from doing the things that you want to do?. For each item answers are given on a 1-4 scale where an adjective (often, sometimes, rarely, never) is matched to any value.*
Life meaningful	*Respondent degree of statements that have used to describe their lives or how they feel. Survey question: How often do you feel that your life has meaning? For each item answers are given on a 1-4 scale where an adjective (often, sometimes, rarely, never) is matched to any value.*
Past good	*Respondent degree of statements that have used to describe their lives or how they feel. Survey question: How often on balance, do you look back to your life with a sense of happiness? For each item answers are given on a 1-4 scale where an adjective (often, sometimes, rarely, never) is matched to any value.*
Vitality	*Respondent degree of statements that have used to describe their lives or how they feel. Survey question: How often do you feel full of energies these days? For each item answers are given on a 1-4 scale where an adjective (often, sometimes, rarely, never) is matched to any value.*
Opportunities	*Respondent degree of statements that have used to describe their lives or how they feel. Survey question: How often do you fell that life is full of opportunities? For each item answers are given on a 1-4 scale where an adjective (often, sometimes, rarely, never) is matched to any value.*
Voluntary	*Dummy =1 if respondent has done voluntary or charity work in the last month.*
Religion attendance	*Dummy =1 if respondent has taken part in activities of a religious organization (church, synagogue, mosque etc.) in the last month.*
Political participation	*Dummy =1 if the respondent has taken part in a political or community-related organization in the last month.*
Help to family	*Dummy =1 if the respondent has provided help to family,friends or neighbors in the last month.*
Cared for sick	*Dummy =1 if the respondent has cared for a sick or disabled adult in the last month.*
Attended education	*Dummy =1 if the respondent has attended an educational or training course in the last month.*
Sport social	*Dummy =1 if the respondent has gone to a sport, social or other kind of club in the last month.*

Figure 1: Distribution of Self Reported Life satisfaction

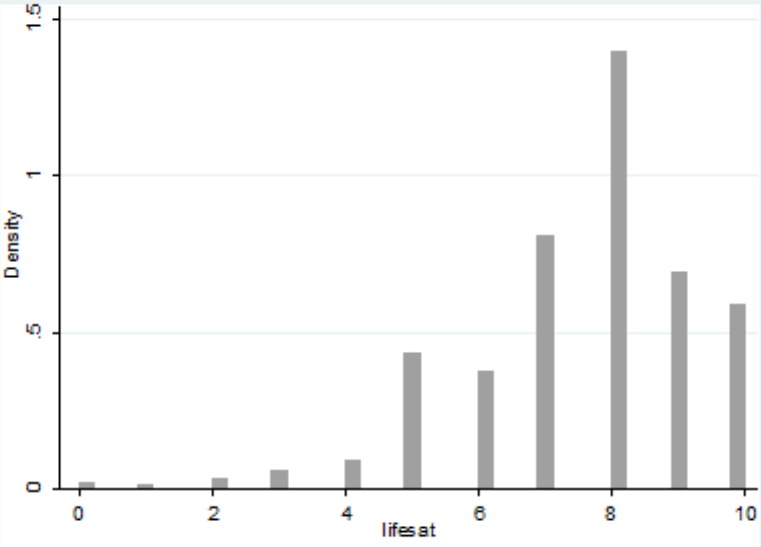

Table 2: **DESCRIPTIVE STATISTICS**

	Imputation n.1				
Variable	**Mean**	**Max**	**Min**	**S.dev.**	**N.**
Life satisfaction	7.54	10.00	0.00	1.78	32412
Female	0.49	1.00	0.00	0.50	33280
Married	0.71	1.00	0.00	0.45	33254
Log income	10.60	15.32	3.00	1.42	32957
Education years	10.50	25.00	0.00	4.28	32712
Household size	2.25	14.00	1.00	1.08	33280
Age class	3.64	7.00	1.00	1.91	33271
Leaving inheritance	0.58	1.00	0.00	0.43	31428
Married	0.71	1.00	0.00	0.45	33254
Widowed	0.15	1.00	0.00	0.35	33254
Divorced	0.07	1.00	0.00	0.25	33254
Separated	0.01	1.00	0.00	0.11	33280
Registered partner	0.01	1.00	0.00	0.12	33254
Sociability	0.10	0.88	0.00	0.13	32517
Voluntary	0.12	1.00	0.00	0.33	32517
Religion attendance	0.11	1.00	0.00	0.31	32517
Political participation	0.04	1.00	0.00	0.20	32517
Help to family	0.17	1.00	0.00	0.38	32517
Cared for sick	0.07	1.00	0.00	0.26	32517
Attended education	0.07	1.00	0.00	0.26	32517
Sport social	0.20	1.00	0.00	0.40	32517
Age no prevent	2.63	4.00	1.00	1.03	32504
No out control	2.84	4.00	1.00	0.96	32339
No feelleftout	3.05	4.00	1.00	0.96	32400
Fred choice	3.23	4.00	1.00	0.89	32458
No fam. Resp. prevent	3.03	4.00	1.00	0.97	32458
No lack money	2.56	4.00	1.00	1.10	32467
Life meaningful	3.55	4.00	1.00	0.72	32265
Past good	3.38	4.00	1.00	0.76	32172
Vitality	3.15	4.00	1.00	0.86	32486
Opportunities	3.09	4.00	1.00	0.87	32290
Future good	3.07	4.00	1.00	0.88	32077
Long-term illness	0.48	1.00	0.00	0.50	33166
Limitedactivities	0.43	1.00	0.00	0.50	33166
Numbillnesses	1.41	13.00	0.00	1.44	33280

Table 3: **THE DETERMINANTS OF LIFE SATISFACTION UNDER STANDARD AND ALTERNATIVE DEPENDENT VARIABLES**

VARIABLES	Self reported latent life satisfaction	Predicted life satisfaction	Life satisfaction average 11 subcomponents	Principal component: 11 subcomponents
Female	0.022	−0.033	−0.016	−0.066
	(0.021)	(0.027)	(0.013)	(0.049)
Log income	0.113***	0.086***	0.037***	0.131***
	(0.022)	(0.013)	(0.006)	(0.022)
Education years	0.024***	0.027***	0.012***	0.048***
	(0.008)	(0.006)	(0.003)	(0.011)
Household size	−0.037**	−0.025***	−0.016***	−0.034**
	(0.014)	(0.007)	(0.004)	(0.013)
Age class 55-59	0.058**	0.079***	0.033***	0.100***
	(0.022)	(0.015)	(0.008)	(0.028)
Age class 60-64	0.092*	0.118***	0.048**	0.149**
	(0.047)	(0.033)	(0.017)	(0.065)
Age class 65-69	0.150***	0.118***	0.053**	0.124*
	(0.045)	(0.036)	(0.018)	(0.069)
Age class 70-74	0.191**	0.095*	0.034	0.028
	(0.068)	(0.050)	(0.025)	(0.092)
Age class 75-79	0.204***	0.052	0.010	−0.110
	(0.063)	(0.048)	(0.025)	(0.086)
Age class above 79	0.211**	−0.058	−0.056*	−0.417***
	(0.073)	(0.057)	(0.028)	(0.101)
Leaving inheritance	0.317***	0.262***	0.119***	0.431***
	(0.045)	(0.025)	(0.011)	(0.042)
Married	0.479***	0.185***	0.067***	0.309***
	(0.079)	(0.034)	(0.016)	(0.060)
Widowed	0.079	0.034	0.016	0.049
	(0.065)	(0.040)	(0.019)	(0.072)
Divorced	−0.010	−0.060	−0.025	−0.048
	(0.082)	(0.058)	(0.026)	(0.100)
Separated	−0.104	−0.029	−0.017	−0.011
	(0.099)	(0.071)	(0.033)	(0.121)
Registered partner	0.533***	0.201***	0.078***	0.354***
	(0.080)	(0.042)	(0.019)	(0.071)
N. of children	0.014	−0.011	−0.008**	−0.010
	(0.009)	(0.007)	(0.004)	(0.013)
N.of grandchildren	0.011**	0.006**	0.003**	0.008*
	(0.005)	(0.002)	(0.001)	(0.004)
Hrooms	0.020	0.026**	0.009**	0.037**
	(0.012)	(0.008)	(0.004)	(0.014)
Big city	0.105**	0.041	0.016	0.080
	(0.047)	(0.031)	(0.016)	(0.058)
Suburbs	0.021	0.055**	0.024*	0.100**
	(0.066)	(0.023)	(0.011)	(0.041)
Large town	0.121**	0.084**	0.041**	0.147**
	(0.055)	(0.035)	(0.016)	(0.063)
Town	0.123*	0.106*	0.053*	0.185**
	(0.058)	(0.049)	(0.024)	(0.085)
Long-term illness	−0.206***	−0.142***	−0.068***	−0.267***
	(0.029)	(0.020)	(0.010)	(0.036)
Limited activities	−0.543***	−0.434***	−0.232***	−0.882***
	(0.048)	(0.036)	(0.018)	(0.068)
Numb. Illnesses	−0.133***	−0.100***	−0.051***	−0.196***
	(0.010)	(0.009)	(0.005)	(0.016)
Voluntary	0.103***	0.103***	0.057***	0.212***
	(0.029)	(0.020)	(0.008)	(0.032)
Religion attendance	0.157**	0.110***	0.044**	0.182**
	(0.053)	(0.034)	(0.019)	(0.062)
Political participation	0.152***	0.087**	0.040**	0.173**
	(0.044)	(0.036)	(0.015)	(0.062)
Help to family	0.103***	0.087***	0.037***	0.197***
	(0.029)	(0.016)	(0.009)	(0.032)
Cared for sick	−0.053*	−0.039*	−0.043***	−0.059
	(0.028)	(0.020)	(0.010)	(0.039)
Attended education	0.023	0.027	−0.001	0.057
	(0.038)	(0.025)	(0.013)	(0.045)
Sport social	0.115***	0.122***	0.061***	0.251***
	(0.030)	(0.024)	(0.012)	(0.043)
Austria	0.490***	0.291***	0.116***	0.388***

(Continued on next page)

(Continued from previous page)

	(0.033)	(0.027)	(0.012)	(0.047)
Belgium	0.226***	0.042***	−0.008 * *	−0.040***
	(0.011)	(0.007)	(0.003)	(0.013)
Czech Rep.	−0.346***	−0.443***	−0.226***	−0.796***
	(0.035)	(0.033)	(0.015)	(0.055)
Switzerland	0.846***	0.419***	0.173***	0.678***
	(0.028)	(0.015)	(0.007)	(0.024)
Spain	0.245***	−0.010	−0.024	−0.093
	(0.041)	(0.032)	(0.015)	(0.057)
Germany	0.340***	0.236***	0.104***	0.310***
	(0.016)	(0.007)	(0.003)	(0.012)
Greece	−0.216***	−0.386***	−0.253***	−0.787***
	(0.037)	(0.027)	(0.013)	(0.046)
Denmark	0.798***	0.321***	0.135***	0.454***
	(0.046)	(0.028)	(0.013)	(0.048)
Italy	0.059	−0.209***	−0.161***	−0.511***
	(0.035)	(0.030)	(0.015)	(0.054)
Netherlands	0.553***	0.582***	0.251***	0.876***
	(0.012)	(0.006)	(0.003)	(0.010)
Poland	−0.316***	−0.078 * *	−0.053***	−0.135 * *
	(0.053)	(0.032)	(0.015)	(0.058)
Sweden	0.590***	0.064*	0.016	0.049
	(0.059)	(0.036)	(0.016)	(0.060)
Constant	5.54 * **	6.22 * **	2.57 * **	−1.84 * **
	(0.26)	(0.17)	(0.07)	(0.29)
Observations	30325.000	29414.000	30427.00	29414.000
R-squared	0.217	0.342	0.342	0.353
Log-Likelihood	−56400.000	−38312.000	−17895.00	−55924.000
AIC	112824.600	76647.540	35813.610	111871.800
BIC	112924.400	76747.010	35913.490	111971.300

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, *p<0.1
Reference Categories: Age class:50-54; Marital Status:Single;Urban area:Rural; Country:France.

Table 4: **CORRELATION MATRIX**

	Lifesat	Age no prev	No out control	No felleftout	Fred of choice	No fam resp.	No lack money	Life meaningful	Past good	Vitality	Opportunity	Future good
Lifesat	1.00											
Age no prev	0.33	1.00										
No out control	0.33	0.43	1.00									
No felleftout	0.38	0.39	0.53	1.00								
Fred of choice	0.31	0.25	0.21	0.25	1.00							
No fam resp.	0.16	0.14	0.19	0.21	0.07	1.00						
No lack money	0.33	0.23	0.21	0.26	0.18	0.28	1.00					
Life meaningful	0.42	0.25	0.26	0.32	0.33	0.06	0.17	1.00				
Past good	0.37	0.17	0.18	0.24	0.23	0.09	0.20	0.44	1.00			
Vitality	0.42	0.40	0.34	0.35	0.36	0.06	0.17	0.44	0.33	1.00		
Opportunity	0.45	0.34	0.28	0.32	0.37	0.08	0.25	0.46	0.38	0.56	1.00	
Future good	0.50	0.36	0.31	0.35	0.37	0.09	0.29	0.49	0.40	0.54	0.63	1.00

Table 5: **THE IMPACT OF SUBJECTIVE WELLBEING SUB-COMPONENTS ON SELF DECLARED LIFE SATISFACTION**

	OLS	**Ordered LOGIT**	**Ordered LOGIT (marginal effects)***
VARIABLES	*Life satisfaction*	*Life satisfaction*	*Life satisfaction*
Life satisfaction			
Age no prevent	0.080***	0.134***	0.009***
	(0.012)	(0.013)	(0.015)
No out control	0.104***	0.145***	0.102***
	(0.019)	(0.027)	(0.175)
No felleftout	0.189***	0.232***	0.016***
	(0.022)	(0.028)	(0.028)
Fred of choice	0.089 * *	0.117 * *	0.008 * *
	(0.033)	(0.047)	(0.004)
No fam resp.prevent	0.072***	0.110***	0.008***
	(0.012)	(0.017)	(0.002)
No lack money	0.223***	0.303***	0.021***
	(0.028)	(0.032)	(0.004)
Life has meaning	0.285***	0.321***	0.023***
	(0.048)	(0.051)	(0.005)
Past good	0.255***	0.338***	0.024***
	(0.044)	(0.050)	(0.005)
Vitality	0.167***	0.216***	0.015***
	(0.029)	(0.037)	(0.003)
Opportunity	0.157***	0.210***	0.015***
	(0.026)	(0.034)	(0.003)
Future good	0.367***	0.472***	0.033***
	(0.030)	(0.031)	(0.006)
Constant	1.38 * **	–	–
	(0.37)	–	–
Observations	31185.000	31185.000	31185.000
R-squared	0.388	–	–
Pseudo R-squared	–	0.127	–
Log-Likelihood	−54388.000	−50699.000	–
AIC	108799.700	101421.400	101421.400
BIC	108899.900	101521.500	101521.500

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Reference Categories: Age class:50-54; Marital Status:Single; Urban area:Rural; Country:France.

Table 6: **PRINCIPAL COMPONENT ANALYSIS (PCA)**

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	4.10	2.78	0.37	0.37
Comp2	1.32	0.31	0.12	0.49
Comp3	1.01	0.22	0.09	0.59
Comp4	0.80	0.08	0.07	0.66
Comp5	0.72	0.04	0.07	0.72
Comp6	0.68	0.09	0.06	0.78
Comp7	0.58	0.06	0.05	0.84
Comp8	0.52	0.06	0.05	0.88
Comp9	0.46	0.02	0.04	0.93
Comp10	0.44	0.08	0.04	0.97
Comp11	0.37	-	0.03	1.00

Table 7: **CORRELATIONS WITH THE FIRST FIVE PRINCIPAL COMPONENTS**

Variable	Comp1	Comp2	Comp3	Comp4	Comp5
Age no prevent	0.30	0.23	-0.36	0.19	-0.27
No out control	0.29	0.38	-0.39	-0.23	0.08
No feel left out	0.31	0.33	-0.25	-0.26	0.14
Fred. choice	0.27	-0.15	0.00	0.57	0.68
No fam. resp.	0.12	0.55	0.48	-0.05	0.36
No lack money	0.22	0.34	0.50	0.31	-0.44
Life has meaning	0.33	-0.27	0.11	-0.31	0.18
Past good	0.28	-0.23	0.36	-0.53	0.03
Vitality	0.36	-0.18	-0.16	0.11	-0.10
Opportunity	0.37	-0.24	0.07	0.13	-0.19
Future good	0.38	-0.20	0.08	0.09	-0.20

Table 8: **THE DETERMINANTS OF DIFFERENCES IN EVALUATING VIGNETTES**

	OLS	OLOGIT	
VARIABLES	Average-vignettes	Vignette 1	Vignette 2
Female	0.017	0.026	0.046
	(0.016)	(0.047)	(0.047)
Log income	−0.028***	−0.076***	−0.051*
	(0.009)	(0.029)	(0.028)
Education years	−0.002	−0.015 **	0.009
	(0.002)	(0.007)	(0.007)
Household size	−0.003	−0.011	0.001
	(0.009)	(0.027)	(0.028)
Age class 55-59	−0.072***	−0.068	−0.243***
	(0.025)	(0.076)	(0.077)
Age class 60-64	−0.044*	−0.010	−0.183 **
	(0.026)	(0.078)	(0.080)
Age class 65-69	−0.121***	−0.232***	−0.312***
	(0.029)	(0.086)	(0.088)
Age class 70-74	−0.155***	−0.241 **	−0.423***
	(0.032)	(0.095)	(0.097)
Age class 75-79	−0.154***	−0.244 **	−0.441***
	(0.036)	(0.108)	(0.110)
Age class above 79	−0.124***	−0.163	−0.389***
	(0.036)	(0.109)	(0.110)
Leaving inheritance	0.016	0.047	0.048
	(0.019)	(0.056)	(0.057)
Married	−0.067*	−0.235*	−0.140
	(0.040)	(0.121)	(0.123)
Widowed	−0.089 **	−0.235*	−0.223
	(0.045)	(0.135)	(0.137)
Divorced	−0.107 **	−0.397***	−0.111
	(0.047)	(0.143)	(0.147)
Separated	−0.075	−0.212	−0.168
	(0.084)	(0.258)	(0.260)
Registered partner	−0.041	−0.210	−0.010
	(0.074)	(0.225)	(0.225)
N. of children	0.012	0.044*	0.011
	(0.008)	(0.023)	(0.024)
N. of grandchildren	−0.005	−0.000	−0.020*
	(0.004)	(0.011)	(0.011)
Hrooms	0.003	0.016	−0.006
	(0.006)	(0.016)	(0.017)
Big city	−0.019	−0.077	−0.025
	(0.026)	(0.079)	(0.078)
Suburbs	−0.000	−0.117	0.120
	(0.024)	(0.073)	(0.073)
Large town	0.006	0.018	0.030
	(0.023)	(0.068)	(0.070)
Small town	0.041*	0.021	0.165 **
	(0.022)	(0.065)	(0.065)
Long-term illness	−0.068***	−0.212***	−0.089
	(0.019)	(0.057)	(0.057)
Limited activities	0.039 **	0.111*	0.058
	(0.019)	(0.057)	(0.058)
Numb illnesses	−0.012 **	−0.014	−0.041 **
	(0.006)	(0.019)	(0.019)
Voluntary	−0.010	−0.061	−0.018
	(0.024)	(0.071)	(0.072)
Religion attendance	0.017	0.094	0.009
	(0.025)	(0.076)	(0.076)
Political participation	−0.018	−0.039	0.019
	(0.037)	(0.111)	(0.114)
Help to family	−0.029	−0.109*	−0.018
	(0.021)	(0.061)	(0.063)
Cared for sick	−0.007	0.023	−0.013
	(0.029)	(0.085)	(0.087)
Attended education	0.057*	0.065	0.166*
	(0.030)	(0.089)	(0.091)
Sport social	−0.010	−0.122 **	0.051
	(0.019)	(0.058)	(0.059)
Belgium	0.213***	0.032	0.920***

(Continued on next page)

(Continued from previous page)

	(0.041)	(0.124)	(0.122)
Czech Rep.	0.474***	1.064***	1.101***
	(0.045)	(0.136)	(0.134)
Spain	0.075	−0.220	0.512***
	(0.049)	(0.150)	(0.144)
Germany	0.332***	1.045***	0.506***
	(0.039)	(0.119)	(0.115)
Greece	0.137***	0.384***	0.137
	(0.048)	(0.147)	(0.142)
Denmark	0.627***	1.402***	1.540***
	(0.045)	(0.135)	(0.133)
Italy	−0.040	−0.431***	0.260 * *
	(0.044)	(0.135)	(0.128)
Netherlands	0.135***	0.528***	0.221*
	(0.045)	(0.135)	(0.134)
Poland	0.257***	0.338 * *	0.794***
	(0.046)	(0.140)	(0.138)
Sweden	0.147***	0.071	0.556***
	(0.051)	(0.153)	(0.149)
Constant	3.30 * **	−	−
	(0.05)	−	−
Observations	7154.000	7134.000	7131.000
R-squared	0.090	−	−
Pseudo R-squared	−	0.035	0.027
Log-Likelihood	−6747.000	−8406.000	−8438.000

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Reference Categories: Age class:50-54; Marital Status:Single; Urban area:Rural; Country:France.

Table 9: **ROBUSTNESS CHECK OF TABLE 3 ESTIMATES WITH THE FOUR ALTERNATIVE IMPUTATIONS (GOODNESS OF FIT ONLY)**

VARIABLES	Self reported latent life satisfaction	Predicted life satisfaction	Life satisfaction average 11 subcomponents	Principal component: 11 subcomponents
Imputation 2				
Observations	30325	29411	30427	29411
R-squared	0.216	0.342	0.342	0.353
AIC	112843	76645.56	35823.81	111860.7
BIC	112942.8	76745.03	35923.69	111960.1
Imputation 3				
Observations	30328	29414	30430	29414
R-squared	0.217	0.342	0.341	0.353
AIC	112826.2	76666.37	35836.53	111890.9
BIC	112926	76765.84	35936.41	111990.4
Imputation 4				
Observations	30316	29404	30418	29404
R-squared	0.217	0.342	0.342	0.353
AIC	112809.2	76626.13	35807.17	111842.6
BIC	112909.1	76725.6	35907.04	111942.1
Imputation 5				
Observations	30330	29416	30432	29416
R-squared	0.216	0.342	0.342	0.353
AIC	112856.5	76644.37	35819.32	111875
BIC	112956.3	76743.84	35919.2	111974.5

Table 10: **ROBUSTNESS CHECK OF TABLE 3 WITH THE SAME NUMBER OF OBSERVATIONS (GOODNESS OF FIT ONLY)**

VARIABLES	Self reported latent life satisfaction	Predicted life satisfaction	Life satisfaction average 11 subcomponents	Principal component: 11 subcomponents
Imputation 1				
Observations	29354	29354	29354	29354
R-squared	0.218	0.341	0.342	0.352
Log-Likelihood	-54462	-38221	-17149	-55796
AIC	108947.2	76466.86	34321.46	111616
BIC	109046.7	76566.31	34420.91	111715.4

Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 11: **BENCHMARK MODEL ESTIMATED WITH EXTENDED PREDICTED LIFE SATISFACTION (11 sub-components interacted with country, gender, age and education dummies)**

VARIABLES	Happypred	**VARIABLES**	Happypred
Female	−0.053∗	*Number grandchildren*	0.007 ∗ ∗
	(0.027)		(0.003)
Log income	0.083∗∗∗	*Hrooms*	0.025 ∗ ∗
	(0.015)		(0.008)
Education years	0.025∗∗∗	*Big city*	0.042
	(0.006)		(0.026)
Household size	−0.027∗∗∗	*Suburbs*	0.055 ∗ ∗
	(0.006)		(0.019)
Age class 55-59	0.089∗∗∗	*Large town*	0.081 ∗ ∗
	(0.017)		(0.035)
Age class 60-64	0.126∗∗∗	*Small town*	0.091 ∗ ∗
	(0.036)		(0.041)
Age class 65-69	0.125∗∗∗	*Voluntary*	0.102∗∗∗
	(0.038)		(0.023)
Age class 70-74	0.097∗	*Religion attendance*	0.107∗∗∗
	(0.046)		(0.032)
Age class 75-79	0.045	*Political participation*	0.063∗
	(0.044)		(0.032)
Age class above 79	−0.055	*Help to family*	0.081∗∗∗
	(0.051)		(0.016)
Leaving inheritance	0.250∗∗∗	*Austria*	0.502∗∗∗
	(0.032)		(0.029)
Married	0.201∗∗∗	*Belgium*	0.280∗∗∗
	(0.044)		(0.008)
Widowed	0.041	*Czech Rep.*	−0.316∗∗∗
	(0.040)		(0.033)
Divorced	−0.028	*Switzerland*	0.881∗∗∗
	(0.055)		(0.020)
Separated	−0.005	*Spain*	0.241 ∗ ∗
	(0.072)		(0.039)
Registered partner	0.220∗∗∗	*Germany*	0.352∗∗∗
	(0.046)		(0.007)
N. of children	−0.012	*Greece*	−0.172∗∗∗
	(0.008)		(0.029)
Long-term illness	−0.140∗∗∗	*Denmark*	0.879∗∗∗
	(0.021)		(0.037)
Limited activities	−0.427∗∗∗	*Italy*	0.811 ∗ ∗
	(0.044)		(0.033)
Numb. illnesses	−0.096∗∗∗	*Netherlands*	0.583∗∗∗
	(0.010)		(0.006)
Cared for sick	−0.036	*Poland*	−0.359∗∗∗
	(0.021)		(0.042)
Attended education	0.029	*Sweden*	0.715∗∗∗
	(0.025)		(0.041)
Sport social	0.121∗∗∗		
	(0.027)		
Constant	6.068∗∗∗		
	(0.219)		
Observations	–		29414.000
R-squared	–		0.404
AIC	–		76365.050
BIC	–		76464.520
Log-Likelihood	–		−38171.520

Robust standard errors in parentheses
*** p<0.01,** p<0.05, * p<0.1
Reference Categories: Age class:50-54; Marital Status:Single;Urban area:Rural; Country:France.

Table 12: **THE DETERMINANTS OF LIFE SATISFACTION UNDER STANDARD AND ALTERNATIVE DEPENDENT VARIABLES (base model from table 3 estimated without health variables)**

VARIABLES	Self reported latent life satisfaction	Predicted life satisfaction	Life satisfaction average 11 subcomponents	Principal component: 11 subcomponents
Female	−0.019	−0.064∗	−0.031 ∗ ∗	−0.127 ∗ ∗
	(0.023)	(0.030)	(0.014)	(0.055)
Log income	0.126∗∗∗	0.097∗∗∗	0.042∗∗∗	0.152∗∗∗
	(0.026)	(0.016)	(0.007)	(0.026)
Education years	0.034∗∗∗	0.034∗∗∗	0.016∗∗∗	0.061∗∗∗
	(0.009)	(0.008)	(0.004)	(0.013)
Household size	−0.033∗	−0.021∗∗∗	−0.014∗∗∗	−0.027 ∗ ∗
	(0.015)	(0.007)	(0.004)	(0.011)
Age class 55-59	0.004	0.039∗	0.012	0.021
	(0.024)	(0.018)	(0.009)	(0.035)
Age class 60-64	−0.005	0.042	0.010	−0.001
	(0.049)	(0.037)	(0.019)	(0.073)
Age class 65-69	−0.005	0.001	−0.008	−0.107
	(0.057)	(0.045)	(0.023)	(0.089)
Age class 70-74	−0.029	−0.072	−0.053	−0.303 ∗ ∗
	(0.082)	(0.060)	(0.031)	(0.115)
Age class 75-79	−0.067	−0.154 ∗ ∗	−0.097∗∗∗	−0.518∗∗∗
	(0.078)	(0.058)	(0.031)	(0.111)
Age class above 79	−0.111	−0.308∗∗∗	−0.185∗∗∗	−0.913∗∗∗
	(0.084)	(0.060)	(0.032)	(0.110)
Leaving inheritance	0.367∗∗∗	0.299∗∗∗	0.139∗∗∗	0.505∗∗∗
	(0.051)	(0.029)	(0.013)	(0.050)
Married	0.451∗∗∗	0.162∗∗∗	0.055∗∗∗	0.262∗∗∗
	(0.079)	(0.035)	(0.016)	(0.059)
Widowed	0.020	−0.013	−0.007	−0.044
	(0.064)	(0.042)	(0.020)	(0.076)
Divorced	−0.064	−0.099∗	−0.047∗	−0.125
	(0.078)	(0.054)	(0.023)	(0.092)
Separated	−0.123	−0.038	−0.024	−0.028
	(0.092)	(0.068)	(0.032)	(0.114)
Registered partner	0.521∗∗∗	0.187∗∗∗	0.073 ∗ ∗	0.325∗∗∗
	(0.089)	(0.051)	(0.025)	(0.088)
N. of children	0.018∗	−0.008	−0.006	−0.005
	(0.009)	(0.007)	(0.004)	(0.013)
N. of grandchildren	0.003	0.000	−0.000	−0.003
	(0.004)	(0.002)	(0.001)	(0.005)
Hrooms	0.029 ∗ ∗	0.033∗∗∗	0.013 ∗ ∗	0.051∗∗∗
	(0.013)	(0.010)	(0.005)	(0.017)
Big city	0.150 ∗ ∗	0.077 ∗ ∗	0.034∗	0.152 ∗ ∗
	(0.050)	(0.032)	(0.017)	(0.059)
Suburbs	0.053	0.078 ∗ ∗	0.037 ∗ ∗	0.147 ∗ ∗
	(0.069)	(0.027)	(0.013)	(0.050)
Large town	0.142 ∗ ∗	0.100 ∗ ∗	0.050 ∗ ∗	0.180 ∗ ∗
	(0.061)	(0.041)	(0.019)	(0.073)
Small town	0.150∗	0.126∗	0.063 ∗ ∗	0.226 ∗ ∗
	(0.070)	(0.059)	(0.029)	(0.103)
Voluntary	0.131∗∗∗	0.123∗∗∗	0.068∗∗∗	0.251∗∗∗
	(0.031)	(0.024)	(0.010)	(0.039)
Religion attendance	0.147 ∗ ∗	0.102 ∗ ∗	0.041∗	0.166 ∗ ∗
	(0.055)	(0.035)	(0.019)	(0.065)
Political participation	0.147∗∗∗	0.083 ∗ ∗	0.039 ∗ ∗	0.165 ∗ ∗
	(0.047)	(0.038)	(0.016)	(0.065)
Help to family	0.108∗∗∗	0.089∗∗∗	0.039∗∗∗	0.200∗∗∗
	(0.029)	(0.017)	(0.009)	(0.034)
Cared for sick	−0.071∗	−0.054 ∗ ∗	−0.050∗∗∗	−0.089∗
	(0.034)	(0.023)	(0.012)	(0.043)
Attended education	0.038	0.038	0.006	0.079
	(0.034)	(0.026)	(0.014)	(0.048)
Sport social	0.157∗∗∗	0.156∗∗∗	0.078∗∗∗	0.320∗∗∗
	(0.028)	(0.023)	(0.011)	(0.039)
Austria	0.486∗∗∗	0.284∗∗∗	0.112∗∗∗	0.369∗∗∗
	(0.036)	(0.029)	(0.013)	(0.051)
Belgium	0.197∗∗∗	0.018∗	−0.022∗∗∗	−0.089∗∗∗
	(0.011)	(0.008)	(0.004)	(0.015)
Czech Rep.	−0.499∗∗∗	−0.562∗∗∗	−0.290∗∗∗	−1.035∗∗∗
	(0.039)	(0.033)	(0.015)	(0.054)

(Continued on next page)

(Continued from previous page)

Switzerland	0.937***	0.487***	0.208***	0.812***
	(0.029)	(0.016)	(0.007)	(0.026)
Spain	0.267***	0.004	−0.015	−0.063
	(0.046)	(0.036)	(0.017)	(0.065)
Germany	0.270***	0.184***	0.075***	0.205***
	(0.019)	(0.011)	(0.005)	(0.018)
Greece	−0.092*	−0.293***	−0.205***	−0.605***
	(0.042)	(0.030)	(0.014)	(0.052)
Denmark	0.721***	0.264***	0.104***	0.340***
	(0.055)	(0.032)	(0.015)	(0.056)
Italy	0.053	−0.215***	−0.164***	−0.524***
	(0.037)	(0.031)	(0.015)	(0.055)
Netherlands	0.526***	0.559***	0.237***	0.824***
	(0.013)	(0.008)	(0.005)	(0.014)
Poland	−0.486***	−0.211***	−0.121***	−0.400***
	(0.046)	(0.027)	(0.013)	(0.048)
Sweden	0.536***	0.021	−0.007	−0.037
	(0.068)	(0.041)	(0.018)	(0.070)
Constant	4.910***	5.746***	2.325***	−2.782***
	(0.284)	(0.187)	(0.083)	(0.318)
Observations	30334.000	29422.000	30436.000	29422.000
R-squared	0.159	0.256	0.245	0.252
AIC	115002.200	80282.000	39996.020	116151.900
BIC	115102.000	80381.480	40095.900	116251.400
Log-Likelihood	−57489.000	−40129.000	−19986.000	−58064.000

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Reference Categories: Age class:50-54; Marital Status:Single; Urban area:Rural; Country:France.

Table 13: **THE DETERMINANTS OF LIFE SATISFACTION UNDER STANDARD AND ALTERNATIVE DEPENDENT VARIABLES (base model from Table 3 estimated without health and social activity variables)**

VARIABLES	Self reported latent life satisfaction	Predicted life satisfaction	Life satisfaction average 11 subcomponents	Principal component: 11 subcomponents
Female	−0.016	−0.062∗	−0.031∗	−0.125 ∗ ∗
	(0.024)	(0.030)	(0.014)	(0.057)
Log income	0.131∗∗∗	0.102∗∗∗	0.044∗∗∗	0.161∗∗∗
	(0.026)	(0.017)	(0.007)	(0.027)
Education years	0.037∗∗∗	0.037∗∗∗	0.017∗∗∗	0.068∗∗∗
	(0.009)	(0.008)	(0.004)	(0.013)
Household size	−0.034 ∗ ∗	−0.024∗∗∗	−0.015∗∗∗	−0.032 ∗ ∗
	(0.014)	(0.007)	(0.004)	(0.011)
Age class 55-59	0.005	0.036∗	0.011	0.014
	(0.025)	(0.019)	(0.010)	(0.036)
Age class 60-64	0.003	0.045	0.012	0.003
	(0.049)	(0.037)	(0.019)	(0.073)
Age class 65-69	0.007	0.008	−0.004	−0.098
	(0.056)	(0.043)	(0.023)	(0.087)
Age class70-74	−0.020	−0.069	−0.050	−0.300 ∗ ∗
	(0.080)	(0.059)	(0.030)	(0.112)
Age class 75-79	−0.073	−0.162 ∗ ∗	−0.100∗∗∗	−0.540∗∗∗
	(0.075)	(0.056)	(0.030)	(0.107)
Age class above 79	−0.143∗	−0.339∗∗∗	−0.199∗∗∗	−0.983∗∗∗
	(0.079)	(0.057)	(0.030)	(0.101)
Leaving inheritance	0.379∗∗∗	0.312∗∗∗	0.144∗∗∗	0.531∗∗∗
	(0.052)	(0.031)	(0.015)	(0.054)
Married	0.443∗∗∗	0.163∗∗∗	0.055∗∗∗	0.267∗∗∗
	(0.079)	(0.033)	(0.016)	(0.059)
Widowed	0.022	−0.008	−0.006	−0.033
	(0.069)	(0.043)	(0.021)	(0.079)
Divorced	−0.073	−0.094	−0.045∗	−0.113
	(0.080)	(0.055)	(0.024)	(0.094)
Separated	−0.125	−0.030	−0.022	−0.016
	(0.087)	(0.069)	(0.032)	(0.114)
Registered partner	0.512∗∗∗	0.184∗∗∗	0.072 ∗ ∗	0.322∗∗∗
	(0.093)	(0.052)	(0.026)	(0.092)
N. of children	0.018∗	−0.007	−0.006	−0.004
	(0.009)	(0.007)	(0.004)	(0.013)
N. of grandchildren	0.003	0.000	−0.000	−0.002
	(0.003)	(0.002)	(0.001)	(0.005)
Hrooms	0.034 ∗ ∗	0.038∗∗∗	0.015∗∗∗	0.062∗∗∗
	(0.013)	(0.010)	(0.005)	(0.017)
Big city	0.145 ∗ ∗	0.068 ∗ ∗	0.030	0.136 ∗ ∗
	(0.053)	(0.031)	(0.017)	(0.058)
Suburbs	0.046	0.068 ∗ ∗	0.032 ∗ ∗	0.132 ∗ ∗
	(0.071)	(0.029)	(0.014)	(0.055)
Large town	0.138 ∗ ∗	0.097 ∗ ∗	0.049 ∗ ∗	0.175 ∗ ∗
	(0.063)	(0.042)	(0.020)	(0.074)
Small town	0.151∗	0.126∗	0.063∗	0.225∗
	(0.073)	(0.061)	(0.030)	(0.107)
Austria	0.479∗∗∗	0.281∗∗∗	0.109∗∗∗	0.363∗∗∗
	(0.034)	(0.026)	(0.012)	(0.046)
Belgium	0.202∗∗∗	0.024∗∗∗	−0.021∗∗∗	−0.077∗∗∗
	(0.011)	(0.007)	(0.003)	(0.012)
Czech Rep.	−0.553∗∗∗	−0.608∗∗∗	−0.312∗∗∗	−1.130∗∗∗
	(0.042)	(0.034)	(0.016)	(0.058)
Switzerland	0.965∗∗∗	0.515∗∗∗	0.219∗∗∗	0.866∗∗∗
	(0.031)	(0.017)	(0.008)	(0.028)
Spain	0.223∗∗∗	−0.031	−0.032∗	−0.140 ∗ ∗
	(0.044)	(0.032)	(0.015)	(0.059)
Germany	0.260∗∗∗	0.180∗∗∗	0.072∗∗∗	0.195∗∗∗
	(0.017)	(0.010)	(0.005)	(0.017)
Greece	−0.105 ∗ ∗	−0.310∗∗∗	−0.214∗∗∗	−0.647∗∗∗
	(0.041)	(0.027)	(0.012)	(0.045)
Denmark	0.746∗∗∗	0.286∗∗∗	0.114∗∗∗	0.385∗∗∗
	(0.057)	(0.034)	(0.015)	(0.059)
Italy	0.010	−0.251∗∗∗	−0.181∗∗∗	−0.599∗∗∗
	(0.035)	(0.028)	(0.013)	(0.050)

(Continued on next page)

(Continued from previous page)

Netherlands	0.563***	0.593***	0.252***	0.894***
	(0.011)	(0.008)	(0.004)	(0.013)
Poland	−0.545***	−0.256***	−0.142***	−0.494***
	(0.044)	(0.021)	(0.010)	(0.038)
Sweden	0.569***	0.051	0.005	0.025
	(0.064)	(0.040)	(0.018)	(0.068)
Constant	4.898***	5.725***	2.317***	−2.824***
	(0.288)	(0.197)	(0.086)	(0.339)
Observations	30519.000	29593.000	30624.000	29593.000
R-squared	0.154	0.247	0.236	0.242
AIC	116016.800	81147.490	40658.840	117308.600
BIC	116116.800	81247.030	40758.790	117408.100
Log-Likelihood	−57996.000	−40561.740	−20317.000	−58642.000

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Reference Categories: Age class:50-54; Marital Status:Single; Urban area:Rural; Country:France.