AW
7/21/08
1:00 pm

Gender and regional differences in self-rated health in Europe*

Franco Peracchi
Tor Vergata University and EIEF

Claudio Rossetti
Tor Vergata University

May 2008

Abstract

This paper shows that gender and regional differences in self-rated health in Europe are partly explained by differences in the prevalence of the various conditions. However, a non-negligible part of these differences is due to other causes, which may include differences in reporting own health. We employ the tool of "anchoring vignettes" to understand whether and how women and men living in different regions differently report levels in a number of health components or domains. We find that vignettes help identifying gender and regional differences in response scales. After correcting for these differences, both gender and regional variation in reported health is substantially reduced, although not entirely eliminated. Our results suggest that differences in response style should be taken into account when using self-assessment of health in socio-economic studies. Failing to do so may lead to misleading conclusions.

Keywords: self-rated health, health domains, anchoring vignettes, reporting bias

JEL codes: C35, C81, I12, J14

* We thank Anne Case, Chris Paxson and Arthur van Soest for helpful discussions. Franco Peracchi also thanks the Center for Health and Wellbeing at Princeton University for generous hospitality during summer 2007, supported by the National Institute of Aging grant P30 AG024361.

1 Introduction

Self-rated health (SRH) tends to be worse for women than for men at all ages, although women are less likely to die and do not present higher hospitalization rates than men at ages when pregnancy-related hospitalization is no longer an issue. In Europe, not only gender differences, but also regional differences in SRH are observed. Both men and women living in Mediterranean countries tend to report worse health than those living in Continental and Scandinavian countries, but they are not more likely to be hospitalized or die.

This paradox could have different explanations, not necessarily mutually exclusive. One explanation is that gender and regional differences in SRH could be due to differences in the distribution of chronic conditions, for either biological or behavioural reasons. Suffering from conditions that are painful, but not life threatening, could lead to poorer SRH but need not imply higher hospitalization or mortality rates. Indeed, Case and Paxson (2005) show that the difference in SRH between women and men in the U.S. can almost entirely be explained by differences in the distribution of chronic conditions.

Another explanation is that there are gender and regional differences in the way people report their health status. This may depend on a different perception of health problems, or on a different mapping of true health status into SRH. In fact, since true health status and subjective thresholds may both vary across individuals, it is not possible, using answers to the subjective scale questions alone, to know how much of the individual rating on these scales reflects true objective differences among people and how much it reflects variation across people in their subjective thresholds. Several studies have focused their attention on heterogeneity of health reporting (see for example Sen 2002, Lindeboom and van Doorslaer 2004, Jürges 2008). Jürges (2007) shows that when differences in reporting styles are taken into account, cross-country variation in SRH in Europe are substantially reduced.

In this paper, we decompose gender and regional differences in morbidity into the contribution of differences in the distribution of chronic conditions and the contribution of the impact of such conditions. For this purpose, we compare men and women living in the same European region, as well as people of the same gender living in different regions, after controlling for differences in socio-demographic characteristics and other health measures, such as body mass and grip strength. Controlling for socio-demographic characteristics is important in order to avoid confounding effects in the relation between SRH and chronic conditions. The fact that differences in SRH between men and women living in different regions can partly be explained by differences in the distribution

of chronic conditions does not exclude the possibility that these groups might use systematically different response scales. For this reason, we employ the tool of "anchoring vignettes" to correct self-assessment of health on six components or domains of health. The domains considered here are pain, mobility, sleeping problems, shortness of breath, concentration problems, and depression. Because reported general health can be regarded as a scalar summary that depends on the level in these different domains (Salomon et al. 2003), understanding whether and how men and women living in different regions differently report levels in these domains may provide helpful insight into differences in SRH.

Anchoring vignettes have been developed as a new component of survey instruments that may be used to position self-reported responses on a common, interpersonally comparable scale. Respondents are first asked to evaluate their position on a scale in a given domain. They are then asked to evaluate the vignette on the same scale they used to rate their own position. Because the objective situation of the person described in the vignette is the same for all respondents, anchoring vignettes have the potential to identify individual variation in subjective thresholds. Vignette questions have been applied in works on international comparisons of health (Salomon, Tandon and Murray 2004, King and Wand 2007, D'Uva et al. 2008), political efficacy (King et al. 2004) and work disability (Kapteyn, Smith and van Soest 2007). In all these applications, subjective scales were used and significant differences were found across groups or countries in the subjective outcomes. Anchoring vignettes were employed to assess whether these groups also differed in their subjective thresholds. A validation study of the use of vignettes for correcting subjective response scales is provided by van Soest et al. (2007).

Our analysis is based on Release 2 of the first (2004) wave of the Survey of Health, Ageing and Retirement in Europe (SHARE). This survey is ideal for our purpose because it contains information on subjective measures of health (such as SRH) and more objective measures (such as hospitalization and interviewer-measured grip strength), as well as detailed information on chronic health conditions. Release 2 of the data also includes the use of vignettes in self-administered questionnaires given to a randomly selected subsample of respondents. For our purpose, the survey is better than other comparable surveys, such as the European Community Household Panel (ECHP), because the latter does not provide detailed information about chronic health conditions, contains little information on objective health measures, and does not include vignettes.

Our results indicate that the differences between men's and women's health are only partially explained by differences in the prevalence of the various conditions. A non-negligible part of the

differences depends on unexplained factors, which may possibly include gender differences in reporting own health. Furthermore, most of the regional differences in the fraction reporting poor health is unexplained by the differences in health conditions and limitations, which again may possibly be due to differences in how people report their health. Socio-demographic characteristics turn out to be much less important than chronic conditions in explaining gender and regional differences in SRH. We find that vignettes help identifying differences in how men and women living in different European regions report their health. Specifically, after correcting for response scales, both gender and regional variations in reported health are substantially reduced, although not eliminated. Our results suggest that differences in response styles should be taken into account when using self-assessment of health in socio-economic studies. Failing to do so may lead to misleading conclusions.

The remainder of this paper is organised as follows Section 2 describes the data used for this study. Section 2.3 provides preliminary evidence. Section 3 examines gender and regional differences in the relationship between chronic conditions and SRH. Section 4 examines gender and regional differences in self-assessment of health, using anchoring vignettes to correct for the possibility that different groups might use systematically different response scales. Finally, Section 5 offers some conclusions.

2 Data and descriptive statistics

2.1 Data

The data in this study are from Release 2 of the first (2004) wave of the Survey of Health, Ageing and Retirement in Europe (SHARE), a multidisciplinary and cross-national longitudinal survey on health, socio-economic status, and social and family networks. The target population of SHARE consists of individuals aged 50+ (born in 1954 or earlier), and their spouses/partners regardless of age, living in private households in Europe. Partners may be younger than 50, but must be living at the exact same address as the selected age-eligible respondent.

Eleven countries have contributed data to the 2004 SHARE baseline study. They are a representation of the various regions of Europe, ranging from Scandinavia (Denmark, Sweden) through Central Europe (Austria, France, Belgium, Germany, Netherlands, Switzerland) to the Mediterranean region (Greece, Italy, Spain). The survey has been administered by means of computer assisted personal interviews (CAPI) in the fall of 2004 to probability samples of individuals aged 50+ in the participating countries. For a detailed description, see Börsch-Supan et al. (2005), and

Börsch-Supan and Jürges (2005).

The survey collects information on health variables (SRH, physical functioning, cognitive functioning, health behavior, use of health care facilities, etc.), psychological variables (psychological health, life satisfaction, etc.), economic variables (current work activity, job characteristics, opportunities to work past retirement age, sources and composition of current income, wealth and consumption, housing, education), and social support variables (assistance within families, transfers of income and assets, social networks, volunteer activities, etc.).[1] The second release of SHARE 2004 also includes vignettes on health as self-administered questionnaires in Sweden, Belgium, Spain, France, Germany, Greece, Italy, and the Netherlands.

We restrict attention to men and women aged 50–90 for whom the vignette information is available. We remove all cases with missing data on any of the variables used. Note that, unlike the case of income or wealth, item nonresponse to health questions is negligible. Nonresponse to the SRH question is lower than 1% in all countries except France, where it is slightly higher than 2%. Even nonresponse to single vignette questions is lower than 1% in almost all countries. Nonetheless, the fraction of respondents for whom answer to at least one vignette question is missing is a bit higher (on average 6%, ranging from about 2% in Greece to about 11% in Sweden).

Table 1 shows the composition of our final sample by country and gender. The subsample of respondents to which the vignettes questions were assigned represents about 18% of the full SHARE sample.

2.2 Variables

Our measure of morbidity is based on the European categorization of SRH into 5 categories:[2] 1="Very good", 2="Good", 3="Fair", 4="Bad", 5="Very bad". We use a dichotomization of SRH, namely a binary indicator equal to one if an individual reports herself to be in fair, bad or very bad health, and equal to zero otherwise. ¿From now on we refer to such binary indicator as "poor health".

SHARE also includes self-assessments and vignette questions on a set of health related concepts or domains, namely pain, mobility, sleeping problems, shortness of breath, concentration problems, depression, and work limitations. This set of health domains is sufficiently exhaustive to capture the common meaning of health. On the other hand, health domains provide a parsimonious description

[1] The data may be downloaded by registered users from the SHARE website (http://www.share-project.org).

[2] We also carried out our analysis by using the US categorization of SRH. The results obtained by using the latter do not differ from the results obtained by using the EU categorization. For this reason we decided to report only the results from the EU categorization.

of health avoiding overlap and redundancy (Salomon et al. 2003). Respondents are asked to rate their own health problems in the six domains on an ordered qualitative scale. The five response categories are: (1) None, (2) Mild, (3) Moderate, (4) Severe, (5) Extreme. For parsimony, in the empirical work we merge the categories "Moderate", "Severe" and "Extreme" into a single one.[3] A detailed description of the self-assessment questions for all six domains is reported in Appendix A.

Two sets of covariates are used to model health outcomes. The first set includes indicators for diagnosed chronic conditions and illnesses, interviewer-measured grip strength, and a measure of relative body weight. The second set includes standard socio-demographic characteristics. The self-reported diagnosed conditions[4] considered are heart attack, high blood pressure, high blood cholesterol, stroke, diabetes, chronic lung disease, asthma, arthritis, osteoporosis, ulcer, Parkinson disease, cataracts, hip or femoral fracture, reproductive cancer, and other cancer. Illnesses which may be symptoms of diseases are pain in back, heart trouble, breathlessness, persistent cough, swollen legs, sleeping problems, falling down, fear of falling down, dizziness, stomach problems, incontinence, and other symptoms.

Grip strength is a core physical measure of health that potentially overcomes the measurement issues arising from subjectivity of SRH. Grip strength is also known to be a good predictor of future medical problems (Rantanen et al. 1999). It is measured here as the maximum of up to four measurements made by the interviewer, two on the left hand and two on the right hand. We use an indicator for the respondent's grip strength (normalized for height, weight and sex) being in the bottom quartile. We label such indicator as "low grip strength".

We include a measure of relative body weight to control for the effects of excessive body weight on physical health. Individuals are classified by relative weight based on their body mass index (BMI), computed from self-reported weight and height as weight (in kilograms) divided by the square of height (in meters). We use the evidence-based clinical guidelines for the classification of overweight and obesity in adults, published by the National Heart, Lung and Blood Institute of the National Institutes of Health (NIH) to classify the respondents into four weight classes: underweight (BMI < 18.5), normal weight ($18.5 \leq$ BMI < 25), overweight ($25 \leq$ BMI < 30), and obesity (BMI ≥ 30), (National Heart, Lung, and Blood Institute).

The set of socio-demographic characteristics includes a polynomial in age, the logarithm of per-capita household income, an indicatorsfor living with a spouse or a partner, and indicators

[3] Our main conclusions do not change if these categories are considered separately.

[4] These conditions are self-reports about medical diagnosis. In fact, the exact questions are "Has a doctor ever told you that you had" a certain condition.

for upper secondary and post-secondary completed education based on the international standard classification of education (ISCED). Household income, in Euros and before tax, is adjusted for purchasing power parity and is the sum of a number of income components that are asked separately in the questionnaire. For many observations, one or more of these components are missing. For observations with missing values, the SHARE data provide imputations largely based on the answers to the sequence of unfolding bracket questions asked to initial nonrespondents. We use the first of the five imputations available in SHARE. To adjust for household size, income is divided by the number of household members.

2.3 Descriptive statistics and preliminary evidence

Figure 1 shows the fraction reporting poor health by gender for each of the countries considered.[5] The fraction of women reporting poor health is always higher than the fraction of men, excepted in France and the Netherlands. The gender difference in SHARE is particularly high for Mediterranean countries (Greece, Italy, and Spain) and is much lower for non-Mediterranean countries (Belgium, France, Germany, Netherlands, and Sweden).

Figure 2 shows the fraction reporting poor health by region, gender and age. In Mediterranean countries, the fraction of women reporting poor health is higher than the fraction of men at almost all ages.

Figure 3 shows the histograms of self assessments for the health domains considered here by region and gender. For most health domains, women are more likely to report themselves to have moderate, severe or extreme health problems than men.

Figure 4 shows prevalence rates of some selected conditions by gender, age and region (non-Mediterranean countries in the top panel, Mediterranean countries in the bottom panel). Women are more likely to suffer from painful conditions such as arthritis, rheumatism, or osteoporosis than do men. On the other hand, men are more likely to suffer from life threatening conditions such as heart attack, or stroke.

Table 2 shows descriptive statistics of the variables used here by region and gender. In non-Mediterranean countries, about 36% of men and 39% of women report poor health. In Mediterranean countries, these percentages are about 35% and 54% respectively for men and women. For both men and women, the fraction of people with low hand grip strength is higher in non-

[5] We carried out a similar analysis using data from the ECHP. For both men and women, the fraction reporting poor health in the ECHP is systematically higher than in SHARE. Apart from this, the main conclusions about gender and regional differences in SRH are similar for the two survey. The results from the ECHP are available from the authors upon request.

Mediterranean countries. Average age varies little, from 63 to 65 years. The fraction with secondary and post-secondary completed education is always higher for men than for women, but people living in non-Mediterranean countries are on average more educated and have higher household income than people living in Mediterranean countries.

3 SRH and chronic conditions

In this section we analyze the relationship between the probability of reporting poor health on the one hand, and socio-economic characteristics and health problems and limitations on the other hand. To facilitate comparison with the results of Case and Paxson (2005) for the U.S., we largely follow their approach.

3.1 Model specification and estimation

We model the probability of reporting poor health ($H = 1$) as a linear function of a set of health problems and limitations C and a set of socio-economic characteristics W

$$\Pr\{H = 1|C, W\} = \alpha + \beta' C + \gamma' W. \tag{1}$$

The set of socio-economic characteristics includes age, age squared, the logarithm of per-capita household income, and indicators for educational attainments and for living with a spouse or a partner. The set of health problems and limitations depends on the specification of the model. In the first specification (Model 1), this set includes indicators for the presence of chronic conditions and symptoms, low grip strength and BMI. The second specification (Model 2) replaces the indicators for the presence of chronic conditions, low grip strength and BMI with a set of indicators for reported mild or moderate, and severe or extreme problems in the six health domains. The third specification (Model 3) contains all the regressors included in (Model 1) and (Model 2).

We estimate model (1) pooling data by country and gender, thus constraining coefficients to be the same for men and women living in different regions, and separately for four groups: non-Mediterranean women (NW), non-Mediterranean men (NM), Mediterranean women (MW), and Mediterranean men (MM). In the second case, the two sets of covariates C and W include all the variables in the third specification (Model 3). The OLS estimates of β_{NW}, β_{NM}, β_{MW}, and β_{MM} provide information on regional and gender differences in how health problems and limitations map into health measures. Following Case and Paxson (2005), we use these estimates and the information about the prevalence of the various conditions and limitations to construct measures of "severity"

and "prevalence" effects. Because the model is linear, we can decompose the differences in the probability of reporting poor health between any two groups, j and k ($j, k = NW, NM, MW, MM$), into a number of components. The first component is a "prevalence effect" (or endowments effect), capturing differences in the distributions of conditions and limitations. It is measured by the differences in prevalence rates weighted by a vector β_* of chronic condition's benchmark coefficients

$$\beta_*'(\bar{C}_j - \bar{C}_k).$$

The second component is a "severity effect" (or coefficients effect), due to differences in the impact of conditions and limitations

$$(\beta_j - \beta_*)'\bar{C}_j + (\beta_* - \beta_k)'\bar{C}_k'.$$

The other components are the endowment effects and the coefficient effects of the control variables in W, and a residual term which includes other regional differences (country dummies) and "unexplained" differences (the constant term). Alternative choices of benchmark coefficients are $\beta_* = \beta_j$, $\beta_* = \beta_k$, $\beta_* = (\beta_j + \beta_k)/2$, or β_* equal to the coefficients in the pooled sample of the two groups. To ensure comparison with Case and Paxson (2005), we set $\beta_* = (\beta_j + \beta_k)/2$.

3.2 Pooled data

Table 3 contains the estimated coefficients of the OLS regression for the probability of reporting poor health and our three different specifications using the pooled data.

In the first specification (Model 1), most of the indicators of chronic conditions and symptoms have a positive and statistically significant effect on the probability of reporting poor health. Low grip strength also has a positive and statistically significant coefficient, while the coefficients on the indicators for BMI turn out to be small and not statistically significant. There is a negative gradient in education, as the probability of poor health declines monotonically with educational attainments. The R^2 of this regression is about 30%.

In the second specification (Model 2), the use of indicators for reported problems in the six health domains achieves a similar fit as Model 1. Not surprisingly, pain and mobility problems have the highest impact on the probability of reporting poor health.

In the third specification (Model 3), most chronic conditions are still significant after controlling for problems in the health domains. Considering both sets of variables improves the R^2 from 30 to about 38%. This is interesting because it indicates that the six health domains are not just summaries of the information provided by the chronic conditions.

3.3 Gender differences

Table 4 shows gender differences in the impact of each condition on the probability of reporting poor health. Estimated OLS coefficients for the four groups are reported in Appendix C. In most cases, the differences in the coefficients between groups are not statistically significant. Further, the hypothesis that all the coefficients associated with conditions and limitations are the same for men and women cannot be rejected at conventional levels. This is consistent with the finding of Case and Paxson (2005) for the U.S. of no significant gender differences in how chronic conditions map into SRH.

Although we observe no gender differences in how conditions map into reported poor health, there are important gender differences in the prevalence of conditions. Table 5 shows excess prevalence of each condition and limitations in women relative to men. Women report significantly higher pain and have higher prevalence of painful conditions such as arthritis, rheumatism, osteoporosis, and other non-life-threatening problems such as sleeping problems and depression. Men, on the other hand, are significantly more likely to suffer from heart attack.

Table 6 shows the decomposition of gender differences in the probability of reporting poor health. The first column shows the decomposition of the differences between non-Mediterranean women and non-Mediterranean men. Women are only about 3% more likely to report poor health than men. The second column shows the decomposition of the differences between Mediterranean women and Mediterranean men. The former are about 19% more likely to report poor health than the latter. The difference between men's and women's health is partly explained by differences in the prevalence of the various conditions. Furthermore, estimated prevalence effects are much more important than severity effects. In particular, the latter explain only less than 3% of the differences. This is again consistent with the findings in Case and Paxson (2005). Nonetheless, a non negligible part of the differences is due to other causes, which may include gender differences in reporting own health.

3.4 Regional differences

The preliminary evidence in Section 2.3 showed that while SRH does not differ much by region for men, this is not true for women. In fact, women living in Mediterranean countries report themselves to be in poorer health than women living in non-Mediterranean countries, although the latter have lower life expectancy than the former. In this section we examine the relationship between regional differences in the probability of reporting poor health and regional differences in the prevalence of

health conditions and limitations.

Table 7 shows regional differences in the impact of each condition on the probability of reporting poor health. In most cases, coefficients are not statistically different between groups and the hypothesis that the coefficients on conditions are the same for people living in non-Mediterranean and Mediterranean countries cannot be rejected at conventional levels. This suggests the absence of significant regional differences in how conditions and limitations map into reports of poor health.

On the other hand, Table 8 suggest that there are important regional differences in the prevalence of the various conditions. The table shows excess prevalence of each condition and limitations in women and men living in Mediterranean countries relative to women and men living in non-Mediterranean countries. Women living in Mediterranean countries have significantly higher rates of arthritis and osteoporosis than women living in non-Mediterranean countries. On the other hand, men living in non-Mediterranean countries are more likely to suffer of hearth attack or stroke than men living in Mediterranean countries.

Table 9 shows the decomposition of the regional differences in the probability of reporting poor health. Although very small for men, these differences are sizable for women. Mediterranean women are about 15% more likely to report poor health than non-Mediterranean women, but a large part of this regional difference remains unexplained. Consistently with the findings in Jürges (2007), this is possibly due to differences in how women living in different regions report their own health.

4 Anchoring vignettes

The results obtained thus far do not exclude the possibility that men and women living in different regions use systematically different response scales when reporting their health. In this section we employ the information contained in anchoring vignettes to check whether this is the case and to control for such differences.

Anchoring vignettes have been developed as a new component of survey instruments that may be used to position self-reported responses on a common, interpersonally comparable scale. Specifically, "an anchoring vignette is a description of a concrete level on a given health domain that respondents are asked to evaluate with the same questions and response scales applied to self-assessments on that domain. Vignettes fix the level of ability on a domain, so that variation in categorical responses is attributable to variation in response category cut-points " (Salomon et al. 2003). Because the same hypothetical situation is presented to each respondents, variability in vi-

gnette answers reveals lack of comparability. In practice, the self-assessment is usually asked first, followed by the vignettes randomly ordered. In SHARE, the names on each vignette are changed to match a respondent's gender and country.

4.1 A simple example

The following example illustrates how vignettes help identifying differences in response scale.

Suppose we want to characterize the amount of pain two groups of individuals have. Figure 5 presents the distribution of the density of the true but unobserved continuous level of pain for groups A and B. On average, people in group B have more pain than people in group A. However, people in the two groups use different response scales when asked whether or not they have pain on a three-point scale. The most common terminology for interpersonal incomparability is "differential item functioning" (DIF). The term originated in the educational testing literature, where a test question is said to have DIF if equally able individuals have unequal probabilities of answering the question correctly. In this example, pain is better tolerated by people in group A than by people in group B. The distribution of self-reports in the two groups suggests that people in A have more pain than those in B. This is in fact the opposite of the true distribution. Correcting for the differences in the response scales is essential to compare the actual level of pain in the two groups.

Vignettes can be used for this purpose. The hypothetical individual described in the vignettes is the same and its objective pain level is marked by the dashed line. This is evaluated as "Mild" by group A and as "None" by group B. Since the actual level of pain of the vignette person is the same, the difference in the evaluations by the two groups is likely to be due to DIF. Hence, vignette evaluations help identify differences in response scales. In fact, using the scales in one of the two groups as the benchmark, the distribution of evaluations in the other group can be adjusted by evaluating them on the benchmark scale. The corrected distribution of the evaluations can then be compared since they are now on the same scale.

4.2 Health on six domains and vignettes

Vignettes included in SHARE refer to the six health domains described in Section 2, namely pain, mobility, sleeping problems, shortness of breath, concentration problems, and depression, plus work limitations. We do not use the vignettes for work limitations because strictly speaking work limitations cannot be considered as a health domain. The reason why there are no vignettes for general health is that this is a multi-dimensional concept and therefore cannot be related to just one domain.

In this section, we use anchoring vignettes to correct for the lack of interpersonal comparability in reported health levels on each of the six domains. Although correction of reported health on the six domains does not offer a direct correction of self-rated general health, it may provide helpful insight into differences in how men and women living in different European regions report their own health.

For each of the six domains, three vignette questions were asked in a random order after the self-assessment question (Appendix B reports a detailed description of the vignettes questions). For each vignette situation, respondents were asked to rate health problems of the hypothetical persons on the same five-point ordered scale ranging from "None" to "Extreme" used for the self-assessment question. As for self-assessments, we merge the categories "Moderate, "Severe" and "Extreme" into a single one. The health problems in the three hypothetical situations in each domain may be viewed as ordered from least to most severe.

Using anchoring vignettes to correct for self-assessment requires two key assumptions (King et al. 2004). The first ("response consistency") is the assumption that each individual uses the response categories for a particular survey question in the same way when providing self-assessment and when assessing each of the hypothetical situations in the vignettes. The second ("vignette equivalence") is the assumption that the level of the variable represented in each vignette is perceived by all respondents in the same way and on the same uni-dimensional scale, apart from random measurement error.

4.3 The statistical model

Our statistical model is a simple adaption of the approach proposed by King et al. (2004) for correcting interpersonal incomparability of self-assessed variables. Their approach is based on a parametric ordered probit model for the self-assessments where, under the assumption of response consistency, the individual specific thresholds depend on the same parameters as in the ordered probit model for the responses to the vignettes.

Specifically, consider one of the six health domains described above. The self-reports on that domain are assumed to be driven by an underlying latent index on a continuous scale

$$Y^* = \mu + U, \tag{2}$$

where μ is a linear function of observed variables and U is a regression error. We specify μ as

$$\mu = \alpha_0 + \alpha_1 F + \alpha_2 M + \alpha_3 F \cdot M + \beta' C + \gamma' W \tag{3}$$

where $\alpha = (\alpha_0, \alpha_1, \alpha_2, \alpha_3)$, β and γ are vectors of unknown parameters, F is an indicators for being a female, M is an indicator for living in Mediterranean countries, C is a set of indicators for the presence of chronic conditions, and W is a set of other controls. We further assume that U is distributed independently of F, M, C and W as $\mathcal{N}(0, \omega^2)$. Instead of the latent index Y^* we observed a categorical variable Y taking value $l = 1, \ldots, L$ whenever $\xi_{l-1} < Y^* \leq \xi_l$, where the thresholds $\xi_0, \ldots, \xi_L$ are given by

$$\begin{aligned} &\xi_0 = -\infty \\ &\xi_1 = \eta_1 + \delta_1' X \\ &\xi_l = \xi_{l-1} + \exp(\eta_l + \delta_l' X), \quad l = 2, \ldots, L-1, \\ &\xi_L = \infty. \end{aligned} \tag{4}$$

The variables contained in X may include F, M, or some of the variables contained in C or W. The nonlinearities in (4) are introduced to guarantee that $\xi_1 < \xi_2 < \cdots < \xi_{L-1}$. A test of homogeneity in response scales (that is, no DIF) is a test of the hypothesis that $\delta_1 = \cdots = \delta_{L-1} = 0$.

Restrictions on the model parameters are needed to ensure identifiability. First of all, the latent index must be assigned a location and a scale. Here, we fix the location by setting $\eta_1 = 0$ and the scale by setting $\omega = 1$. It is easy to verify that, if the variables in F, M, C and W are the same as those in X, and we use only the self-assessments, then the parameter vectors α, β and γ in (2) cannot be separately identified from the parameter vectors $\eta = (\eta_2, \ldots, \eta_{L-1})$ and $\delta = (\delta_1, \ldots, \delta_{L-1})$ in (4). In this case, identification only depends on the nonlinearities in the model for the thresholds. Since identification by functional form is undesirable, strong identification requires using at least one vignette.

Responses to each of the three vignettes are also modeled using an ordered probit model

$$Z_m^* = \psi_m + V_m, \qquad m = 1, 2, 3,$$

where, under the vignette equivalence assumption, the vector $\psi = (\psi_1, \psi_2, \psi_3)$ is the same for all respondents and the V_m are assumed to be independently and identically distributed as $\mathcal{N}(0, \sigma^2)$ independently of F, M, C, W, X and U. The scale parameter σ^2 measures how well vignettes are understood. What is observed is a categorical variable Z_m, which takes the value $l = 1, \ldots, L$ whenever $\xi_{l-1} < Z_m^* \leq \xi_l$. Under the assumption of response consistency, the thresholds in the self-assessment and the vignette components of the model are the same, which ensures identifiability of the entire parameter vector $\theta = (\alpha, \beta, \gamma, \eta, \delta, \psi, \sigma)$.

The basic model presented in this section may be generalized by relaxing distributional assumptions and by introducing time-invariant individual effects (Rossetti 2008). In the next section we

consider a fully parametric version of the model that retains normality of the errors in the latent model (2) but controls for unobserved individual heterogeneity by introducing a set of random individual effects with a distribution which is known up to a finite set of parameters.

4.4 Model estimation

A simple way of controlling for unobserved heterogeneity in the model of the previous section is to treat the intercept η_1 in (4) as a random variable with a distribution of a known shape. Conditional on a specific value $\eta_1 = e$ of the individual effect, the likelihood contribution from the self-assessment component for the ith individual in the sample is

$$\mathcal{L}_i^s(\theta \mid e) = \prod_{l=1}^{L} [\Phi(\xi_{il} - \mu_i) - \Phi(\xi_{i,l-1} - \mu_i)]^{1\{Y_i = l\}},$$

where $1\{\cdot\}$ is the indicator function and $\Phi(\cdot)$ is the standard normal distribution function, while the likelihood contribution from the vignette component is

$$\mathcal{L}_i^v(\theta \mid e) = \prod_{m=1}^{3} \prod_{l=1}^{L} \left[\Phi\left(\frac{\xi_{il} - \psi_m}{\sigma} \right) - \Phi\left(\frac{\xi_{i,l-1} - \psi_m}{\sigma} \right) \right]^{1\{Z_{im} = l\}}.$$

Because the likelihood from the self-assessment and the vignette components share the parameter vectors η and δ, they must be maximized jointly. Conditional on $\eta_1 = e$, the overall likelihood contribution for the ith individual is the product of four univariate normal probabilities (one for the self-assessment component and three for the vignette component). Since the individual effect is not observed, the unconditional likelihood contribution for the ith individual can be computed by taking expectations with respect to η_1. Given a random sample of n individuals, and assuming that the individual effects are distributed independently of F, M, C, W, X and U as $\mathcal{N}(0, \varphi^2)$, a ML estimator of θ is obtained by maximizing the sample likelihood

$$\mathcal{L}(\theta) = \prod_{i=1}^{n} w_i \int \mathcal{L}_i^s(\theta \mid e)\, \mathcal{L}_i^v(\theta \mid e)\, \frac{1}{\varphi} \phi\left(\frac{e}{\varphi} \right) de, \tag{5}$$

where w_i is the survey weight for the ith individual and $\phi(\cdot)$ is the standard normal density. The integral in (5) can be approximated numerically using Gauss-Hermite quadrature. We estimate separate models for each of the six health domains.[6] The maximization routine is written in MATA, the matrix programming language of STATA, and is based on the Newton-Raphson algorithm, with numerical first and second derivatives.

[6] We also estimated a model with common thresholds for all six domains, but such model is rejected against the model with different response scales for each of the six domains.

We estimate the model after pooling data by gender and region, thus constraining the slope coefficients to be the same for men and women living in different regions. Nonetheless, the hypothesis that all the coefficients associated with conditions and limitations are the same for men and women living in different regions cannot be rejected at conventional levels.[7] The vector C_i includes indicators for chronic conditions, low grip strength and BMI. The vector W_i includes a set of socio-economic characteristics (age, age squared, the logarithm of household income, and indicators for educational attainments and for living with a spouse or a partner). The vector X_i in the threshold equation includes the same variables contained in C_i and W_i, the indicators for being a female F and for living in Mediterranean countries M, and their interaction.

Table 10 reports the estimated coefficients of both the ordered probit model with constant thresholds (the baseline) and the ordered probit model with individual specific thresholds for the three domains which have the highest impact on SRH, namely pain, mobility and concentration.[8] For parsimony, only the coefficients of the gender and regional dummies and their interactions are reported. Complete parameter estimates of the vignettes model are reported in Appendix D. A likelihood ratio (LR) test rejects the hypothesis of no DIF (constant thresholds) in favor of the model with individual specific thresholds for all health domains. The hypothesis that $\varphi = 0$ (no random individual effects in the response scale) is also rejected by a likelihood ratio (LR) test. The results of the ordered probit model with constant thresholds appear in the first numerical column of Table 10. After controlling for chronic conditions, grip strength, BMI and socio-economic characteristics, people living in Mediterranean countries report significantly lower health problems in each domain (the coefficient of the indicator for living in Mediterranean countries is negative and significant for each domain). Female respondents report significantly higher pain.

The second numerical column of Table 10 presents the parameter estimates using the vignettes to correct for differences in thresholds among respondents. First of all, the estimates of the actual values of the three vignettes for each domain turn out to be ordered in exactly the way we expected (from least to most health problems in each domain). This also provides some evidence that each concept being measured is likely to be unidimensional. For most health domains the estimated coefficient of the dummy for living in Mediterranean countries substantially reduce in magnitude compared to the model with constant thresholds. Furthermore, for pain such dummy is no longer significant. For concentration problems, the interaction between the dummy for female and the

[7] Given the small sample size, a model with full heterogeneity in the parameters (both in the latent index and in the thresholds) could not be estimated due to problems of convergence.

[8] Results for the other health domains are available from the authors upon request.

dummy for living in Mediterranean countries is no longer significant. Finally, for pain the female dummy reduces in magnitude compared to the model with constant thresholds. The explanation for these differences in the estimated coefficients between the model with individual specific thresholds and the model with constant thresholds is given by the estimates of threshold parameters. In fact, significant shifts in the thresholds are observed both by gender and region for all considered domains. This indicates that there are both gender and regional differences in response scales.

4.5 Decomposition of gender and regional differences

Analogously to the decomposition exercise computed for SRH in the first part of this paper, we now decompose gender and regional differences in the level of health problems in each domain. Because the latent model (3) is linear in such level, we can decompose the differences between any two groups, j and k, into a number of components. The first component is a "prevalence effect", capturing differences in the distributions of conditions

$$\beta'(\bar{C}_j - \bar{C}_k).$$

The "severity effect" is zero under the assumption that coefficients are the same for men and women living in different regions. The second component is the "endowment effect" of the socio-economic characteristics W

$$\gamma'(\bar{W}_j - \bar{W}_k).$$

The last component is a residual term which includes differences in the health measure that cannot be explained neither by differences in the distributions of conditions, nor by differences in the distributions of the socio-economic characteristics. Specifically, the unexplained difference between non-Mediterranean women and non-Mediterranean men is α_1. The unexplained difference between women and men living in Mediterranean countries is $\alpha_1 + \alpha_3$. The unexplained difference between Mediterranean women and non-Mediterranean women is $\alpha_2+\alpha_3$. Finally, the unexplained difference between Mediterranean men and non-Mediterranean men is α_2.

Table 11 shows the decomposition of gender differences in the level of health problems for selected domains. The decomposition is reported for both the ordered probit model with constant thresholds and the ordered probit model with individual specific thresholds. The top panel of Table 11 shows the decomposition of the differences between non-Mediterranean women and non-Mediterranean men. Non-Mediterranean women have a higher level of health problems than non-Mediterranean men. Unexplained differences in pain are reduced from about 71% to about

64% when correcting for differences in response scales. Unexplained differences in mobility and concentration problems are instead increased when correcting for differences in response scales. The bottom panel of Table 11 shows the decomposition of the differences between Mediterranean women and Mediterranean men. Mediterranean women have much higher level of health problems than Mediterranean men. Unexplained differences in pain, mobility and concentration are all reduced when correcting for differences in response scales.

Table 12 shows the decomposition of the regional differences in the level of health problems for selected domains. All regional differences are substantially reduced when correcting for differences in response scales.

5 Conclusions

In this paper we looked at gender and regional differences in SHR using data from Release 2 of the first (2004) wave of SHARE. Our results indicate that the difference between men's and women's health is partly explained by differences in the prevalence of the various conditions. However, a non negligible part of the difference is due to "other causes", which may possibly include gender differences in reporting own health. Furthermore, most of the regional differences in the fraction reporting poor health is unexplained by differences in health conditions and limitations or by socio-demographic characteristics. Again, this may reflect differences in how people report their health.

We employ the tool of "anchoring vignettes" for correcting response scales in the self-assessment of health on six domains: pain, mobility, sleeping problems, shortness of breath, concentration problems, and depression. Understanding whether and how women and men living in different regions differently report levels in these domains can give us helpful insight into differences in SRH. We find that vignettes help identifying both gender and regional differences in how respondents report their health. In particular, the fraction of gender differences in the level of health which cannot explained by chronic conditions nor by socio-economic characteristics is substantially reduced after correcting for differences in response scales. Furthermore, after correcting for such differences, regional differences in the level of health are substantially reduced, although not entirely eliminated. Our results suggest that differences in response styles should be taken into account when using self-assessment of health in socio-economic studies. Failing to do so may lead to misleading conclusions.

References

Börsch-Supan A., Brugiavini 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: Mannheim Research Institute for the Economics of Aging.

Börsch-Supan A., Jürges H. (2005). *The Survey of Health, Aging, and Retirement in Europe Methodology*, Mannheim: Mannheim Research Institute for the Economics of Aging.

Case A., Paxson C. (2005). "Sex differences in morbidity and mortality", *Demography*, 42: 189–214.

D'Uva T.B., Van Doorslaer E., Lindeboom M., O'Donnell O. (2008). "Does reporting heterogeneity bias the measurement of health disparities?", *Health Economics*, 17: 351–375.

Jürges H. (2007). "True health vs response styles: exploring cross-country differences in self-reported health", *Health Economics*, 16: 163–178.

Jürges H. (2008). "Self-assessed health, reference levels, and mortality", *Applied Economics*, 40: 569–582.

Kapteyn,A., Smith J., van Soest A. (2007). "Vignettes and self-reports of work disability in the United States and the Netherlands", *American Economic Review*, 97: 461–473.

Kerkhofs M., Lindeboom M. (1995). "Subjective health measures and state department reporting errors", *Health Economics*, 4: 221–235.

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

King G., Wand J. (2007). "Comparing incomparable survey responses: evaluating and selecting anchoring vignettes", *Political Analysis*, 15: 46–66.

Lindeboom M., van Doorslaer E. (2004). "Cut-point shift and index shift in self- reported health", *Journal of Health Economics*, 23: 1083–1099.

Rantanen T., Guralnik J.M., Foley D., Masaki K., Leveille S.G., Curb J.D., et al. (1999), "Midlife hand grip strength as a predictor of old age disability", *Journal of the American Medical Association*, 21: 558–560.

Rossetti C. (2008). "Non-parametric estimation of ordered probit model with anchoring vignettes", mimeo.

Salomon J.A., Mathers C.D., Chatterji S., Sadana R., Üstün T.B., Murray C.J.L. (2003). "Quantifying individual levels of health: definitions, concepts and measurement issues", in Murray C.J.L. and Evans D.B. (eds.), *Health Systems Performance Assessment: Debates, Methods and Empiricism*, Geneva, World Health Organization.

Salomon J.A., Tandon A., Murray C.J.L. (2004). "Comparability of self rated health: cross sectional multi-country survey using anchoring vignettes", *British Medical Journal*, 328: 258–260.

Sen A. (2002). "Health: perception versus observation", *British Medical Journal*, 324: 860–861.

van Soest A., Delaney L., Harmon C., Kapteyn A., Smith J.P. (2007). "Validating the use of vignettes for subjective threshold scales", mimeo.

Table 1: Final sample by country and gender.

Country	Men			Women			Total		
	Full	Vign.	%	Full	Vign.	%	Full	Vign.	%
Germany	1,263	175	13.9	1,391	220	15.8	2,654	395	14.9
Sweden	1,313	144	11.0	1,410	154	10.9	2,723	298	10.9
Netherlands	1,256	223	17.8	1,365	221	16.2	2,621	444	16.9
Spain	870	179	20.6	1,113	212	19.0	1,983	391	19.7
Italy	1,002	153	15.3	1,183	191	16.1	2,185	344	15.7
France	1,159	318	27.4	1,412	390	27.6	2,571	708	27.5
Greece	1,102	294	26.7	1,181	275	23.3	2,283	569	24.9
Belgium	1,621	209	12.9	1,758	249	14.2	3,379	458	13.6
Total	9,586	1,695	17.7	10,813	1,912	17.7	20,399	3,607	17.7

Table 2: Descriptive statistics.

	Non-Mediterranean				Mediterranean			
	Men		Women		Men		Women	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Poor health	0.360	0.480	0.387	0.487	0.352	0.478	0.541	0.499
Heart attack	0.155	0.362	0.087	0.283	0.120	0.325	0.095	0.294
High blood pressure	0.305	0.461	0.320	0.467	0.284	0.451	0.402	0.491
High blood cholesterol	0.221	0.415	0.181	0.386	0.245	0.430	0.210	0.407
Stroke	0.045	0.207	0.020	0.141	0.012	0.107	0.026	0.160
Diabetes	0.080	0.272	0.100	0.301	0.114	0.319	0.103	0.304
Chronic lung disease	0.051	0.220	0.047	0.212	0.057	0.232	0.084	0.278
Asthma	0.043	0.204	0.030	0.171	0.046	0.211	0.036	0.187
Arthritis	0.139	0.346	0.211	0.408	0.176	0.381	0.379	0.485
Osteoporosis	0.015	0.122	0.090	0.286	0.010	0.099	0.150	0.357
Ulcer	0.073	0.261	0.037	0.189	0.081	0.274	0.042	0.200
Parkinson disease	0.007	0.085	0.002	0.044	0.002	0.039	0.002	0.041
Cataracts	0.051	0.220	0.060	0.238	0.063	0.243	0.100	0.300
Hip or femoral fracture	0.019	0.135	0.013	0.112	0.003	0.050	0.015	0.123
Reproductive cancer	0.017	0.131	0.047	0.212	0.005	0.067	0.022	0.146
Other cancer	0.036	0.185	0.027	0.162	0.051	0.220	0.024	0.153
Pain in back	0.500	0.500	0.562	0.496	0.426	0.495	0.622	0.485
Heart trouble	0.097	0.297	0.063	0.243	0.047	0.212	0.079	0.270
Breathlessness	0.110	0.313	0.086	0.280	0.081	0.273	0.106	0.308
Persistent cough	0.044	0.206	0.046	0.210	0.037	0.188	0.067	0.251
Swollen legs	0.040	0.197	0.157	0.364	0.076	0.266	0.228	0.420
Sleeping problems	0.143	0.350	0.224	0.417	0.087	0.283	0.258	0.438
Falling down	0.017	0.129	0.031	0.174	0.026	0.160	0.071	0.256
Fear of falling down	0.032	0.177	0.104	0.306	0.034	0.183	0.129	0.335
Dizziness	0.047	0.212	0.069	0.254	0.070	0.256	0.136	0.343
Stomach problems	0.124	0.330	0.151	0.359	0.110	0.314	0.203	0.403
Incontinence	0.021	0.143	0.038	0.191	0.020	0.141	0.088	0.283
Other symptoms	0.045	0.208	0.031	0.175	0.042	0.200	0.056	0.229
Low grip strength	0.274	0.446	0.161	0.368	0.349	0.477	0.330	0.471
Underweight	0.000	0.020	0.011	0.105	0.004	0.064	0.013	0.114
Overweight	0.513	0.500	0.374	0.484	0.510	0.500	0.408	0.492
Obese	0.155	0.362	0.173	0.379	0.193	0.395	0.196	0.397
Pain: mild	0.347	0.476	0.381	0.486	0.366	0.482	0.353	0.478
Pain: mod/sev/extr	0.295	0.456	0.391	0.488	0.226	0.419	0.400	0.490
Sleeping problems: mild	0.240	0.427	0.309	0.462	0.278	0.448	0.275	0.447
Sleeping problems: mod/sev/extr	0.267	0.442	0.341	0.474	0.172	0.378	0.369	0.483
Mobility problems: mild	0.235	0.424	0.268	0.443	0.177	0.382	0.201	0.401
Mobility problems: mod/sev/extr	0.218	0.413	0.233	0.423	0.124	0.330	0.252	0.435
Concentration problems: mild	0.359	0.480	0.369	0.483	0.297	0.457	0.308	0.462
Concentration problems: mod/sev/extr	0.202	0.401	0.231	0.422	0.149	0.356	0.310	0.463
Shortness of breath: mild	0.199	0.399	0.233	0.423	0.144	0.351	0.147	0.354
Shortness of breath: mod/sev/extr	0.155	0.362	0.142	0.349	0.078	0.268	0.118	0.323
Depression: mild	0.252	0.435	0.316	0.465	0.240	0.427	0.290	0.454
Depression: mod/sev/extr	0.175	0.380	0.218	0.413	0.119	0.324	0.323	0.468
Age	63.6	9.4	64.7	10.0	63.7	8.8	64.9	10.2
Living with spouse or partner	0.796	0.403	0.599	0.490	0.782	0.413	0.540	0.499
Secondary education	0.471	0.499	0.399	0.490	0.217	0.413	0.130	0.337
Post-secondary education	0.250	0.433	0.183	0.387	0.091	0.288	0.060	0.237
Log HH income	9.73	0.99	9.69	1.02	9.10	0.99	8.84	1.24
Observations	1,069		1,234		626		678	

Notes: weighted results

Table 3: Estimated coefficients of the OLS regression for poor health. (* significant at 5%; ** significant at 1%).

	Model 1	Model 2	Model 3
Heart attack	0.178 **	.	0.158 **
High blood pressure	0.070 **	.	0.064 **
High blood cholesterol	0.023	.	0.024
Stroke	0.150 **	.	0.108 **
Diabetes	0.200 **	.	0.146 **
Chronic lung disease	0.018	.	0.006
Asthma	0.028	.	0.006
Arthritis	0.157 **	.	0.105 **
Osteoporosis	0.177 **	.	0.143 **
Ulcer	0.073 *	.	0.061 *
Parkinson disease	0.379 **	.	0.287 *
Cataracts	-0.061 *	.	-0.039
Hip or femoral fracture	-0.155 *	.	-0.117 *
Reproductive cancer	0.115 **	.	0.069
Other cancer	0.303 **	.	0.258 **
Pain in back	0.074 **	.	0.004
Heart trouble	0.081 **	.	0.060 *
Breathlessness	0.219 **	.	0.178 **
Persistent cough	0.081 *	.	0.077 *
Swollen legs	0.011	.	-0.018
Sleeping problems	0.068 **	.	0.032
Falling down	-0.029	.	-0.036
Fear of falling down	0.104 **	.	0.057 *
Dizziness	0.056 *	.	0.020
Stomach problems	0.047 *	.	0.023
Incontinence	-0.065	.	-0.087 *
Other symptoms	0.085 *	.	0.050
Low grip strength	0.086 **	.	0.046 **
Underweight	-0.087	.	-0.116
Overweight	-0.011	.	-0.020
Obese	-0.017	.	-0.044 *
Pain: Mild	.	0.121 **	0.106 **
Pain: Mod/sev/extr	.	0.300 **	0.228 **
Sleeping problems: Mild	.	0.008	0.003
Sleeping problems: Mod/sev/extr	.	0.041 *	0.024
Mobility problems: Mild	.	0.121 **	0.097 **
Mobility problems: Mod/sev/extr	.	0.208 **	0.154 **
Concentration problems: Mild	.	-0.042 *	-0.043 **
Concentration problems: Mod/sev/extr	.	0.051 *	0.037
Shortness of breath: Mild	.	0.049 **	0.028
Shortness of breath: Mod/sev/extr	.	0.104 **	0.006
Depression: Mild	.	0.007	0.007
Depression: Mod/sev/extr	.	0.054 *	0.044 *
Age - 55	0.005 **	0.011 **	0.006 **
(Age - 55) squared /100	-0.005	-0.016 *	-0.014 *
Secondary education	-0.053 **	-0.039 *	-0.045 **
Post-secondary education	-0.097 **	-0.072 **	-0.077 **
Living with spouse or partner	-0.030	0.008	-0.010
Log HH income - log(12763)	-0.034 **	-0.035 **	-0.035 **
Constant	0.230 **	0.095 **	0.112 **
Observations	3,607	3,607	3,607
R^2	0.318	0.310	0.388

Notes: weighted results, country dummies omitted

Table 4: Gender differences in the impact of conditions on the probability of reporting poor health. Model 3 (* significant at 2%).

	Non-Medit.	Medit.
Heart attack[a)]	-0.037	-0.052
High blood pressure	-0.046	0.138 *
High blood cholesterol	0.115 *	0.030
Stroke	-0.039	-0.184
Diabetes	0.025	-0.092
Chronic lung disease	0.102	-0.190
Asthma	-0.036	-0.068
Arthritis	0.046	0.054
Osteoporosis	0.059	0.043
Ulcer	0.197 *	-0.206
Parkinson disease	-0.106	0.277
Cataracts	-0.185 *	0.061
Hip or femoral fracture	-0.112	0.269
Reproductive cancer	-0.128	-0.102
Other cancer	0.236 *	-0.032
Pain in back	-0.064	-0.093
Heart trouble	0.005	-0.257
Breathlessness	0.018	0.114
Persistent cough	0.082	-0.169
Swollen legs	-0.109	0.071
Sleeping problems	-0.004	-0.073
Falling down	-0.157	0.016
Fear of falling down	-0.013	-0.023
Dizziness	0.104	-0.020
Stomach problems	-0.054	0.168
Incontinence	0.216	-0.057
Other symptoms	0.093	-0.118
Low grip strength	-0.104	-0.042
Underweight	-0.258	-0.221
Overweight	0.074	-0.018
Obese	0.192 *	-0.101
Pain: Mild	0.037	0.025
Pain: Mod/sev/extr	-0.019	0.081
Sleeping problems: Mild	-0.017	0.015
Sleeping problems: Mod/sev/extr	0.024	-0.049
Mobility problems: Mild	0.005	-0.040
Mobility problems: Mod/sev/extr	0.060	-0.184
Concentration problems: Mild	-0.034	0.042
Concentration problems: Mod/sev/extr	-0.069	0.265 *
Shortness of breath: Mild	0.006	-0.052
Shortness of breath: Mod/sev/extr	-0.109	0.042
Depression: Mild	0.002	0.077
Depression: Mod/sev/extr	-0.004	-0.016
All conditions[b)]	0.907	0.972

Notes:
a) significance from t-tests of the hypothesis that the coefficients of each condition on poor health are identical for men and women
b) F-tests of the hypothesis that all the coefficients of chronic conditions on poor health are identical for men and women

Table 5: Excess prevalence of conditions in women (* significant at 2%).

	Non-Medit.[a)]	Medit.[b)]
Heart attack	-0.068 *	-0.036
High blood pressure	-0.003	0.116 *
High blood cholesterol	-0.039	-0.038
Stroke	-0.029 *	0.018
Diabetes	0.004	-0.019
Chronic lung disease	-0.008	0.016
Asthma	-0.013	-0.008
Arthritis	0.072 *	0.189 *
Osteoporosis	0.078 *	0.138 *
Ulcer	-0.039 *	-0.032 *
Parkinson disease	-0.005	0.001
Cataracts	0.004	0.018
Hip or femoral fracture	-0.007	0.014 *
Reproductive cancer	0.032 *	0.019 *
Other cancer	-0.006	-0.029 *
Pain in back	0.055 *	0.163 *
Heart trouble	-0.045 *	0.016
Breathlessness	-0.029	0.025
Persistent cough	-0.002	0.024
Swollen legs	0.112 *	0.137 *
Sleeping problems	0.091 *	0.179 *
Falling down	0.011	0.043 *
Fear of falling down	0.060 *	0.086 *
Dizziness	0.020	0.055 *
Stomach problems	0.028	0.077 *
Incontinence	0.014	0.057 *
Other symptoms	-0.012	0.005
Low grip strength	-0.132 *	-0.043
Underweight	0.012 *	0.008
Overweight	-0.135 *	-0.077 *
Obese	-0.001	-0.008
Pain: Mild	0.044	-0.013
Pain: Mod/sev/extr	0.074 *	0.151 *
Sleeping problems: Mild	0.068 *	-0.003
Sleeping problems: Mod/sev/extr	0.085 *	0.202 *
Mobility problems: Mild	0.026	0.034
Mobility problems: Mod/sev/extr	-0.012	0.097 *
Concentration problems: Mild	0.021	0.017
Concentration problems: Mod/sev/extr	-0.006	0.126 *
Shortness of breath: Mild	0.041	0.011
Shortness of breath: Mod/sev/extr	-0.025	0.029
Depression: Mild	0.067 *	0.046
Depression: Mod/sev/extr	0.024	0.182 *

Notes: Excess prevalence coefficients are the coefficients on an indicator that the respondent is female
a) in the sample of non-Mediterr. countries,
b) in the sample of Mediterr. countries,
in OLS regression for each condition, which also includes a set of control variables W

Table 6: Gender differences. Decomposition of the probability of poor health.

	Non-Mediterranean	Mediterranean
Men	0.360	0.352
Women	0.387	0.541
Difference (women - men)	0.027	0.189
Decomposition of the difference (%)		
Prevalence effect	131.6	62.0
Severity effect	30.0	2.6
Socio-dem. char.: Endowments effect	29.4	10.6
Socio-dem. char.: Coefficients effect	30.2	-41.8
Residual difference	-121.2	66.6

Table 7: Regional differences in the impact of conditions on the probability of reporting poor health. Model 3 (* significant at 2%).

	Women	Men
Heart attack[a)]	-0.074	-0.059
High blood pressure	0.080	-0.104
High blood cholesterol	0.002	0.088
Stroke	-0.000	0.145
Diabetes	-0.108	0.009
Chronic lung disease	-0.149	0.143
Asthma	0.074	0.106
Arthritis	-0.007	-0.015
Osteoporosis	-0.006	0.010
Ulcer	-0.189	0.214 *
Parkinson disease	-0.208	-0.591
Cataracts	0.148	-0.098
Hip or femoral fracture	0.380 *	-0.000
Reproductive cancer	0.212	0.186
Other cancer	0.018	0.286 *
Pain in back	-0.037	-0.008
Heart trouble	-0.098	0.165
Breathlessness	-0.072	-0.168
Persistent cough	-0.107	0.145
Swollen legs	0.187 *	0.008
Sleeping problems	0.066	0.135
Falling down	-0.106	-0.279
Fear of falling down	0.094	0.104
Dizziness	-0.086	0.038
Stomach problems	0.014	-0.208 *
Incontinence	0.002	0.275
Other symptoms	-0.187	0.024
Low grip strength	0.066	0.005
Underweight	-0.251	-0.289
Overweight	0.013	0.105
Obese	-0.096	0.197 *
Pain: Mild	0.060	0.071
Pain: Mod/sev/extr	-0.008	-0.107
Sleeping problems: Mild	-0.038	-0.070
Sleeping problems: Mod/sev/extr	-0.121	-0.049
Mobility problems: Mild	-0.017	0.028
Mobility problems: Mod/sev/extr	-0.089	0.155
Concentration problems: Mild	0.089	0.013
Concentration problems: Mod/sev/extr	0.052	-0.281 *
Shortness of breath: Mild	-0.067	-0.009
Shortness of breath: Mod/sev/extr	0.111	-0.040
Depression: Mild	0.111	0.037
Depression: Mod/sev/extr	-0.004	0.007
All conditions[b)]	1.034	1.392

Notes:
a) significance from t-tests of the hypothesis that the coefficients of each condition on poor health are identical for people living in Medit. and non-Medit. countries
b) F-tests of the hypothesis that all the coefficients of chronic conditions on poor health are identical for people living in Medit. and non-Medit. countries

Table 8: Excess prevalence of conditions in Mediterranean countries (* significant at 2%).

	Women[a)]	Men[b)]
Heart attack	-0.006	-0.006
High blood pressure	0.015	-0.042
High blood cholesterol	0.034	0.021
Stroke	0.013	-0.054 *
Diabetes	-0.043	0.007
Chronic lung disease	0.063 *	-0.006
Asthma	-0.001	0.008
Arthritis	0.260 *	0.129 *
Osteoporosis	0.115 *	-0.008
Ulcer	0.017	0.014
Parkinson disease	0.000	-0.008
Cataracts	0.026	0.004
Hip or femoral fracture	-0.015	-0.019
Reproductive cancer	-0.029	-0.011
Other cancer	0.002	0.014
Pain in back	0.052	-0.115 *
Heart trouble	0.005	-0.059 *
Breathlessness	0.016	-0.015
Persistent cough	0.022	-0.025
Swollen legs	0.071 *	0.056 *
Sleeping problems	0.044	-0.048
Falling down	0.059 *	0.002
Fear of falling down	0.003	-0.014
Dizziness	0.077 *	0.016
Stomach problems	0.095 *	0.003
Incontinence	0.078 *	-0.011
Other symptoms	-0.002	-0.027
Low grip strength	0.185 *	0.030
Underweight	0.022 *	0.009 *
Overweight	-0.070	-0.016
Obese	-0.066 *	-0.038
Pain: Mild	0.100 *	0.121 *
Pain: Mod/sev/extr	-0.066	-0.099 *
Sleeping problems: Mild	-0.008	0.047
Sleeping problems: Mod/sev/extr	0.101 *	-0.040
Mobility problems: Mild	-0.115 *	-0.115 *
Mobility problems: Mod/sev/extr	-0.047	-0.209 *
Concentration problems: Mild	0.049	0.001
Concentration problems: Mod/sev/extr	0.005	-0.117 *
Shortness of breath: Mild	-0.075 *	-0.048
Shortness of breath: Mod/sev/extr	-0.037	-0.151 *
Depression: Mild	0.078 *	-0.018
Depression: Mod/sev/extr	0.053	-0.114 *

Notes: Excess prevalence coefficients are the coefficients
on an indicator that the respondent lives in Medit. countries
a) in the sample of women,
b) in the sample of men,
in OLS regression for each condition, which also includes
a set of control variables W

Table 9: Regional differences. Decomposition of the probability of poor health.

	Women	Men
Non-Med.	0.387	0.360
Medit.	0.541	0.352
Difference (Medit. - non-Med.)	0.154	-0.008
Decomposition of the difference (%)		
Prevalence effect	29.3	762.5
Severity effect	24.9	-832.2
Socio-dem. char.: Endowments effect	31.0	-666.5
Socio-dem. char.: Coefficients effect	-13.0	-666.4
Residual difference	27.8	1502.6

Table 10: Ordered probit model with constant thresholds, and ordered probit model with individual specific thresholds for selected health domains (* significant at 5%; ** significant at 1%).

		Pain		Mobility problems		Concentration problems	
Equation	Variable	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds
SA	medit	-0.150 *	-0.115	-0.442 **	-0.322 **	-0.286 **	-0.147 *
	female	0.280 **	0.237 **	0.022	0.071	-0.001	0.007
	medit*female	-0.100	-0.179	0.141	0.071	0.273 **	0.128
	Constant	-0.320 **	-0.507 **	-0.899 **	-1.042 **	-0.091	-0.154
Thres. 1	medit	.	-0.048	.	0.136 **	.	0.190 **
	female	.	-0.142 **	.	0.022	.	-0.013
	medit*female	.	0.064	.	-0.029	.	-0.127 *
	Constant	0.000	0.000	0.000	0.000	0.000	0.000
Thres. 2	medit	.	0.153 **	.	-0.052	.	-0.118 **
	female	.	0.142 **	.	0.078	.	0.049
	medit*female	.	-0.229 **	.	-0.131	.	-0.049
	Constant	1.200 **	-0.060	0.832 **	-0.505 **	1.039 **	-0.022
Vign. 1	Constant	.	0.494 **	.	0.668 **	.	0.549 **
Vign. 2	Constant	.	1.602 **	.	1.235 **	.	1.282 **
Vign. 3	Constant	.	2.086 **	.	1.400 **	.	1.902 **
ln ω	Constant	.	0.000	.	0.000	.	0.000
ln σ	Constant	.	-0.359 **	.	-0.670 **	.	-0.352 **
ln φ	Constant	.	-0.889 **	.	-0.866 **	.	-0.798 **
Obs.			3,607		3,607		3,607
LR test			735.9		927.5		886.2
p(LR test)			0.0		0.0		0.0
LR test $\varphi = 0$			299.1		517.1		488.0
p(LR test $\varphi = 0$)			0.0		0.0		0.0

Notes: weighted results, only the coefficients of the gender and regional dummies are reported for parsimony. LR test is a Likelihood ratio test of the hypothesis of no DIF (constant thresholds)

Table 11: Gender differences. Decomposition of health level in selected domains.

Non-Mediterranean countries						
	Pain		Mobility problems		Concentration problems	
	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds
Men	0.510	0.246	-0.109	-0.467	0.159	0.029
Women	0.906	0.613	0.009	-0.271	0.269	0.144
Difference (women - men)	0.396	0.367	0.118	0.196	0.110	0.115
Decomposition of the difference (%)						
Prevalence effect	23.4	21.0	39.0	18.9	26.9	4.2
Socio-dem. char.	5.8	14.4	42.0	44.8	73.5	89.4
Residual difference	70.9	64.6	19.0	36.3	-0.5	6.4
Mediterranean countries						
	Pain		Mobility problems		Concentration problems	
	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds
Men	0.322	0.157	-0.571	-0.733	-0.120	-0.069
Women	0.926	0.611	-0.072	-0.273	0.417	0.285
Difference (women - men)	0.604	0.454	0.499	0.459	0.537	0.354
Decomposition of the difference (%)						
Prevalence effect	65.9	74.5	54.8	48.8	32.1	26.9
Socio-dem. char.	4.2	12.8	12.3	20.4	17.2	34.9
Residual difference	29.9	12.7	32.8	30.8	50.7	38.3

Table 12: Regional differences. Decomposition of health level in selected domains.

Women						
	Pain		Mobility problems		Concentration problems	
	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds
non-Med.	0.906	0.613	0.009	-0.271	0.269	0.144
Medit.	0.926	0.611	-0.072	-0.273	0.417	0.285
Difference (Medit. - non-Med.)	0.020	-0.003	-0.081	-0.002	0.148	0.141
Decomposition of the difference (%)						
Prevalence effect	1366.1	-9196.9	-275.2	-8512.6	58.6	42.6
Socio-dem. char.	-36.5	-2226.5	5.5	-2306.1	50.7	71.0
Residual difference	-1229.7	11523.5	369.7	10918.7	-9.3	-13.6
Men						
	Pain		Mobility problems		Concentration problems	
	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds
non-Med.	0.510	0.246	-0.109	-0.467	0.159	0.029
Medit.	0.322	0.157	-0.571	-0.733	-0.120	-0.069
Difference (Medit. - non-Med.)	-0.188	-0.090	-0.462	-0.266	-0.279	-0.098
Decomposition of the difference (%)						
Prevalence effect	15.0	29.6	0.8	-3.3	20.2	30.8
Socio-dem. char.	5.4	-57.6	3.5	-17.6	-22.7	-80.7
Residual difference	79.6	128.0	95.7	120.9	102.6	149.9

Figure 1: Fraction reporting poor health by country and gender.

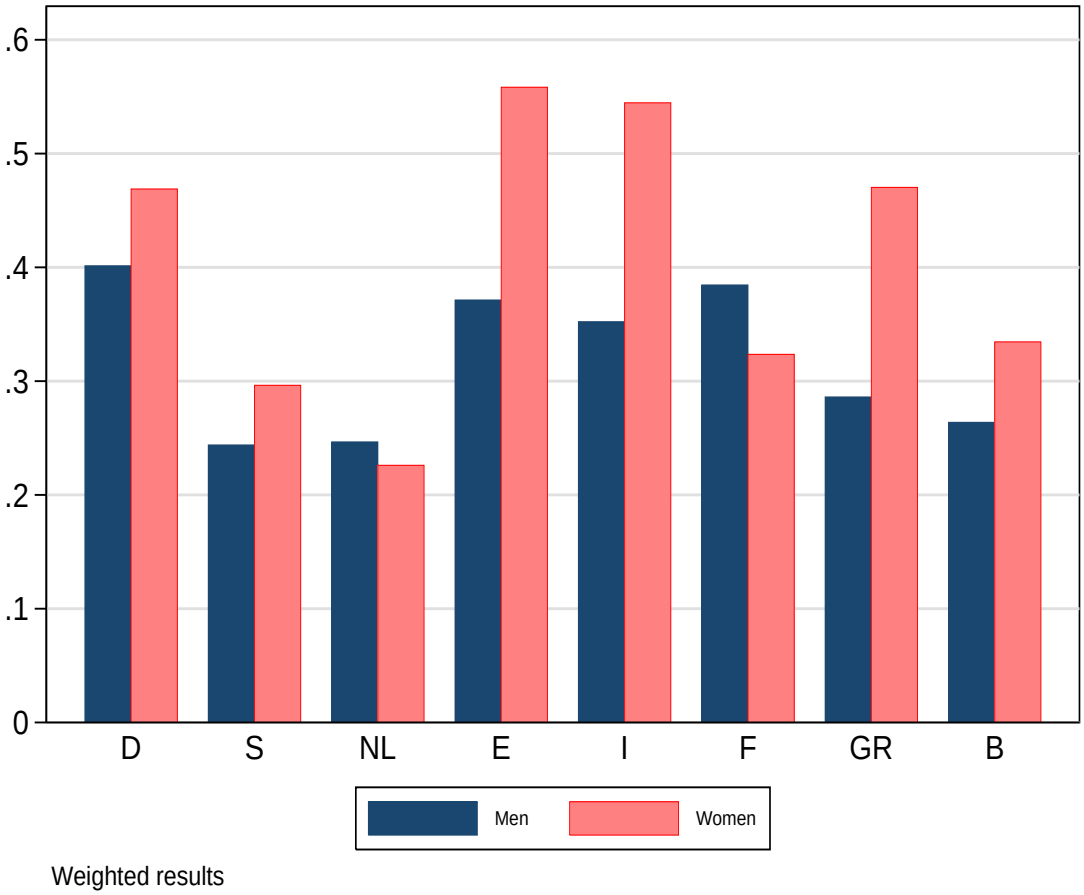
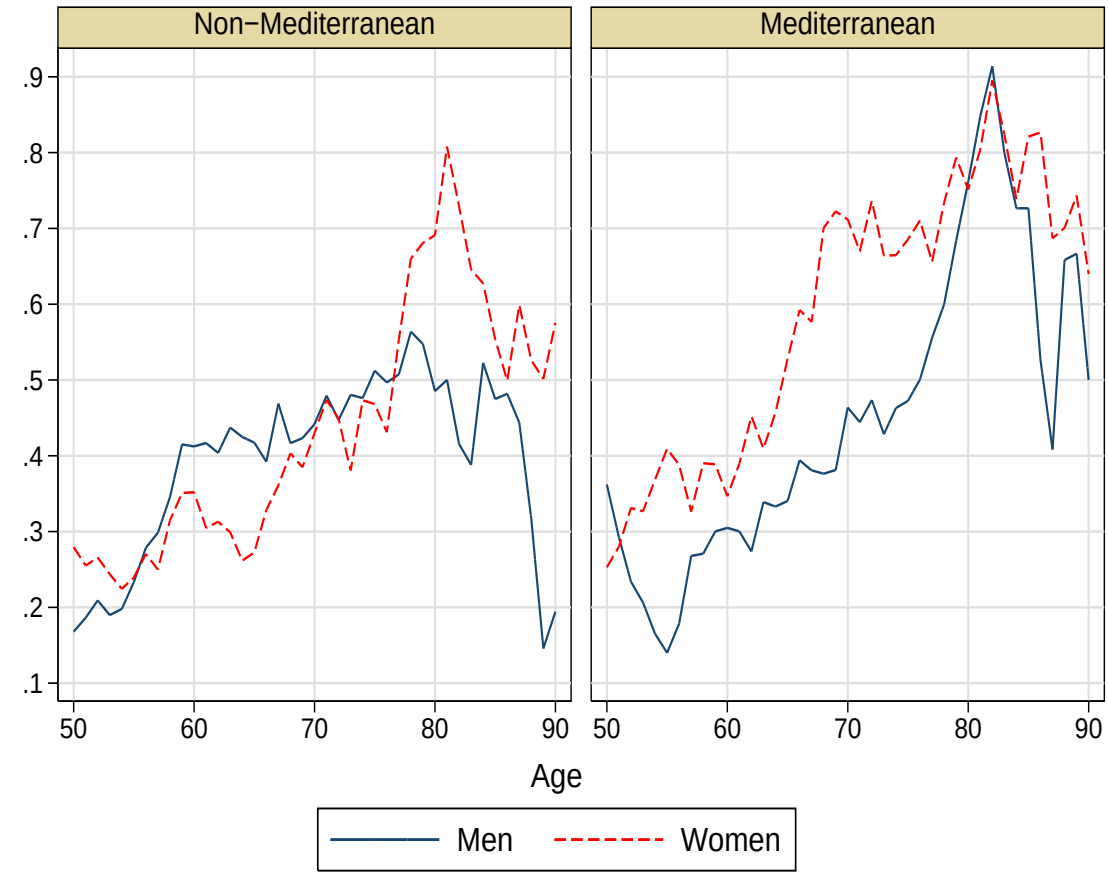
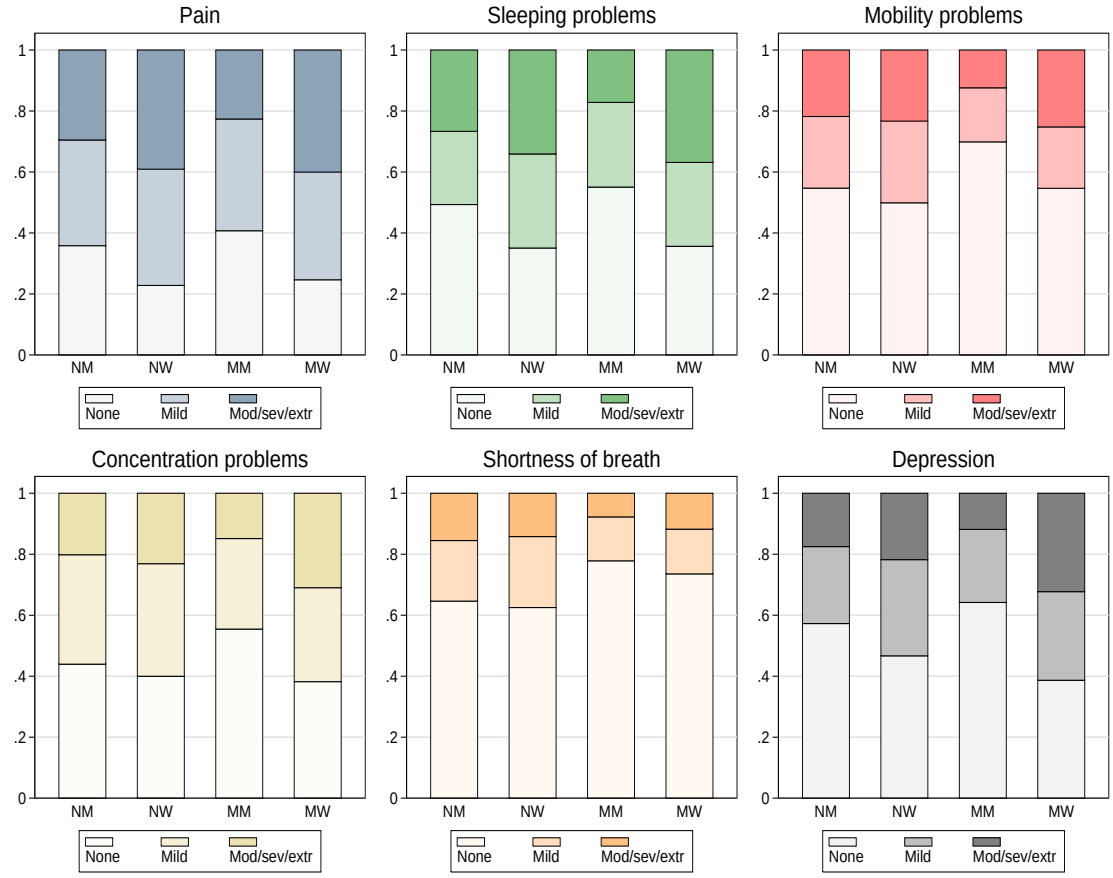


Figure 2: Fraction reporting poor health by age, region and gender.



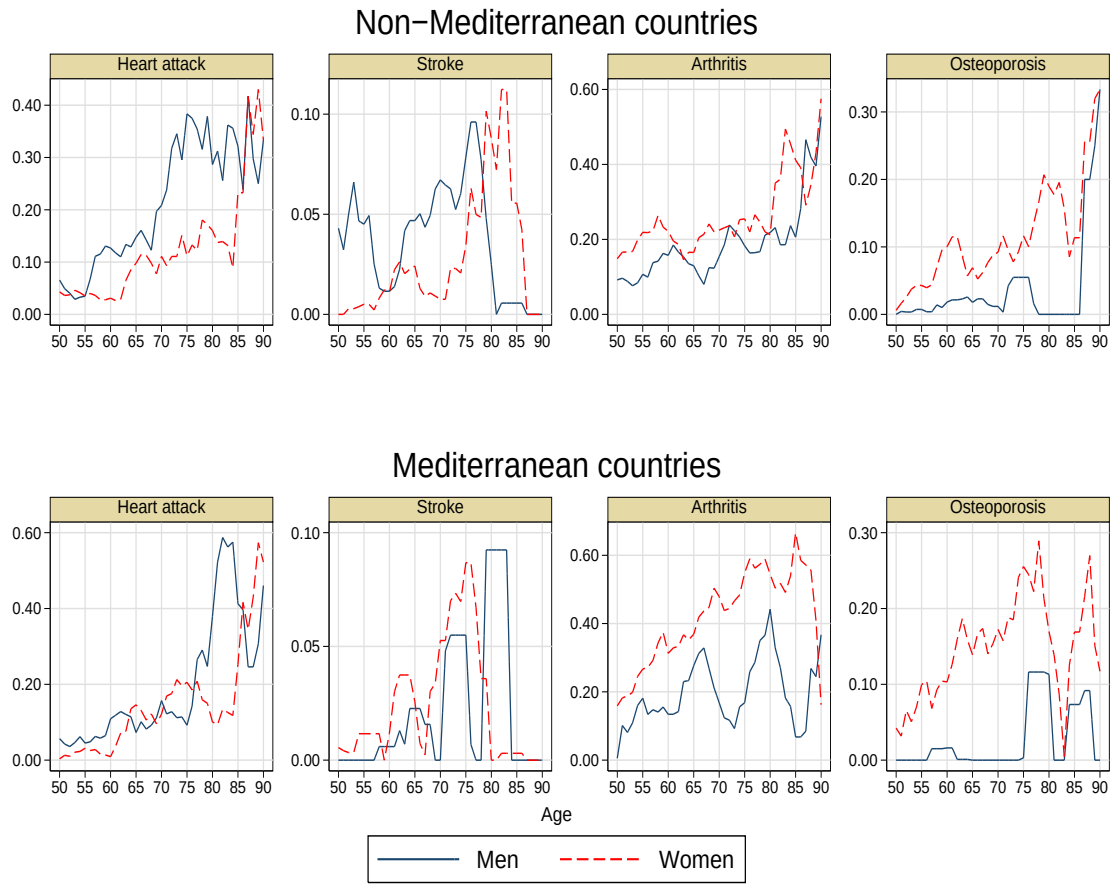
Weighted results

Figure 3: Histograms of self assessments for health domains by region and gender.



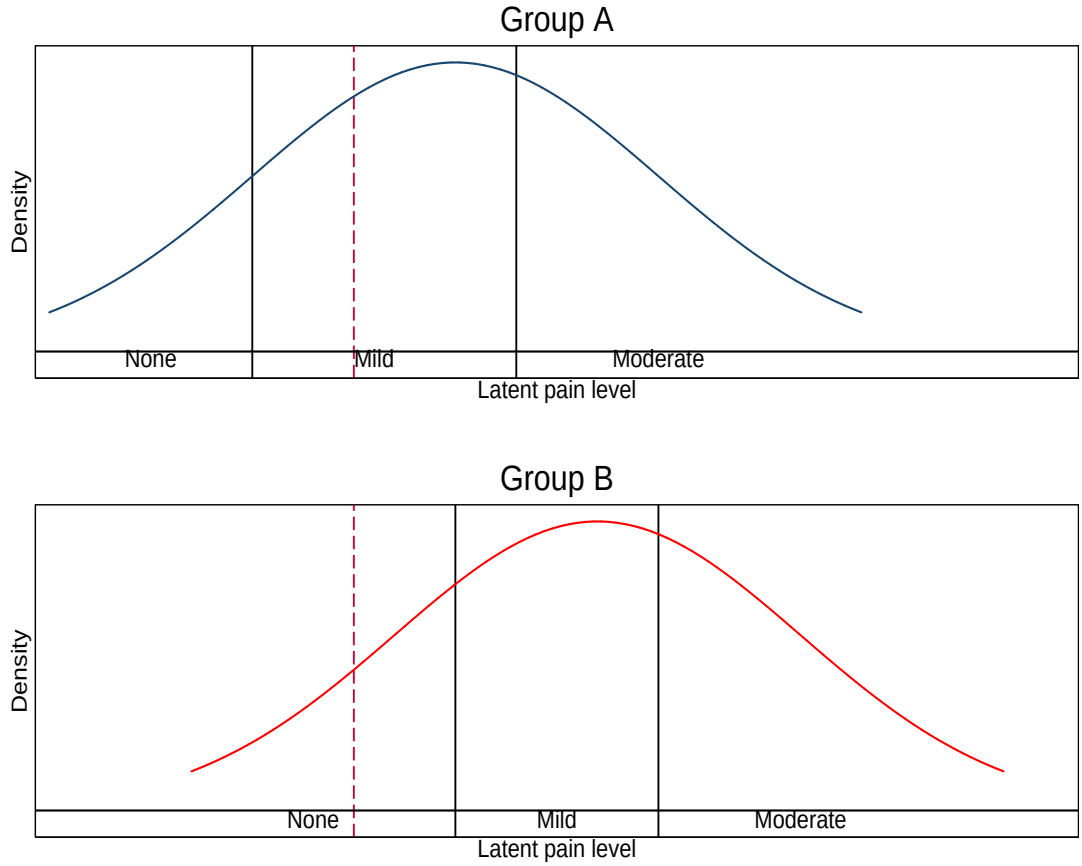
Weighted results

Figure 4: Prevalence rates of some selected conditions by gender and region.



Weighted results

Figure 5: Example of comparison of self-assessed pain in two groups in case of differences in response scales.



APPENDICES

A Description of self-assessments on health domains

Self-assessment questions for each health domain are the following.

Pain

"Overall in the last 30 days, how much of bodily aches or pains did you have?".

Sleeping problems

"In the last 30 days, how much difficulty did you have with sleeping such as falling asleep, waking up frequently during the night or waking up too early in the morning?".

Mobility

"Overall in the last 30 days, how much of a problem did you have with moving around?".

Concentration problems

"Overall in the last 30 days how much difficulty did you have with concentrating or remembering things?".

Shortness of breath

"In the last 30 days, how much of a problem did you have because of shortness of breath?".

Depression

"Overall in the last 30 days, how much of a problem did you have with feeling sad, low, or depressed?".

B Description of vignette hypothetical situations

The three vignette hypothetical situations for each health domain are the following.

Pain

1. *"Paul/Karen has a headache once a month that is relieved after taking a pill. During the headache he/she can carry on with his/her day-to-day affairs."*
2. *"Henri/Maria has pain that radiates down his/her right arm and wrist during his/her day at work. This is slightly relieved in the evenings when he/she is no longer working on his/her computer."*
3. *"Charles/Alice has pain in his/her knees, elbows, wrists and fingers, and the pain is present almost all the time. Although medication helps, he/she feels uncomfortable when moving around, holding and lifting things."*

Sleeping problems

1. *"Charles/Alice falls asleep easily at night, but two nights a week he/she wakes up in the middle of the night and cannot go back to sleep for the rest of the night."*
2. *"Paul/Karen wakes up almost once every hour during the night. When he/she wakes up in the night, it takes around 15 minutes for his/her to go back to sleep. In the morning he/she does not feel well-rested."*
3. *"Henri/Maria takes about two hours every night to fall asleep. He/she wakes up once or twice a night feeling panicked and takes more than one hour to fall asleep again."*

Mobility

1. *"Rob/Eve is able to walk distances of up to 200 metres without any problems but feels tired after walking one kilometre or climbing more than one flight of stairs. He/she has no problems with day-to-day activities, such as carrying food from the market."*
2. *"Kevin/Lisa does not exercise. He/she cannot climb stairs or do other physical activities because he/she is obese. He/she is able to carry the groceries and do some light household work."*
3. *"Tom/Sue has a lot of swelling in his/her legs due to his/her health condition. He/she has to make an effort to walk around his/her home as his/her legs feel heavy."*

Concentration problems

1. *"Kevin/Lisa can concentrate while watching TV, reading a magazine or playing a game of cards or chess. Once a week he/she forgets where his/her keys or glasses are, but finds them within five minutes."*
2. *"Tom/Sue is keen to learn new recipes but finds that he/she often makes mistakes and has to reread several times before he/she is able to do them properly."*
3. *"Rob/Eve cannot concentrate for more than 15 minutes and has difficulty paying attention to what is being said to his/her. Whenever he/she starts a task, he/she never manages to finish it and often forgets what he/she was doing. He/she is able to learn the names of people he/she meets."*

Shortness of breath

1. *"Mark/Karen has no problems with walking slowly. He/she gets out of breath easily when climbing uphill for 20 meters or a flight of stairs."*
2. *"Paul/Karen suffers from respiratory infections about once every year. He/she is short of breath 3 or 4 times a week and had to be admitted in hospital twice in the past month with a bad cough that required treatment with antibiotics."*
3. *"Henri/Maria has been a heavy smoker for 30 years and wakes up with a cough every morning. He/she gets short of breath even while resting and does not leave the house anymore. He/she often needs to be put on oxygen."*

Depression

1. *"Paul/Karen enjoys his/her work and social activities and is generally satisfied with his/her life. He/she gets depressed every 3 weeks for a day or two and loses interest in what he/she usually enjoys but is able to carry on with his/her day-to-day activities."*

2. *"Henri/Maria feels nervous and anxious. He/she worries and thinks negatively about the future, but feels better in the company of people or when doing something that really interests his/her. When he/she is alone he/she tends to feel useless and empty."*

3. *"Mark/Anna feels depressed most of the time. He/she weeps frequently and feels hopeless about the future. He/she feels that he/she has become a burden on others and that he/she would be better dead."*

C Parameter estimates of the OLS regression for poor health by region and gender

This appendix presents parameter estimates of the OLS regression for poor health separately for non-Mediterranean women, non-Mediterranean men, Mediterranean women, and Mediterranean men (* significant at 5%; ** significant at 1%).

	Mediterranean countries		Non-Mediterranean countries	
	Men	Women	Men	Women
Heart attack	0.197 **	0.160 **	0.139 *	0.086
High blood pressure	0.086 **	0.040	-0.018	0.120 **
High blood cholesterol	-0.056	0.059 *	0.032	0.062
Stroke	0.093	0.054	0.238	0.054
Diabetes	0.147 **	0.172 **	0.156 **	0.064
Chronic lung disease	-0.052	0.050	0.091	-0.100
Asthma	0.010	-0.027	0.116	0.048
Arthritis	0.068	0.113 **	0.053	0.107 **
Osteoporosis	0.086	0.145 **	0.096	0.139 **
Ulcer	-0.043	0.154 *	0.170 **	-0.036
Parkinson disease	0.392 **	0.286	-0.198	0.078
Cataracts	0.066	-0.118 *	-0.032	0.029
Hip or femoral fracture	-0.116	-0.228 *	-0.116	0.153
Reproductive cancer	0.150	0.022	0.336	0.234 *
Other cancer	0.088	0.324 **	0.374 **	0.342 **
Pain in back	0.054 *	-0.011	0.045	-0.048
Heart trouble	0.065	0.070	0.230 *	-0.027
Breathlessness	0.217 **	0.235 **	0.049	0.163 *
Persistent cough	0.013	0.094	0.157	-0.012
Swollen legs	0.017	-0.092 **	0.024	0.095 *
Sleeping problems	0.012	0.008	0.148 *	0.074
Falling down	0.140	-0.017	-0.139	-0.123
Fear of falling down	0.018	0.005	0.122	0.099
Dizziness	-0.030	0.075	0.009	-0.011
Stomach problems	0.085 *	0.031	-0.123 *	0.045
Incontinence	-0.248 **	-0.032	0.027	-0.030
Other symptoms	0.014	0.107	0.038	-0.080
Low grip strength	0.089 **	-0.014	0.094 *	0.052
Underweight	0.235	-0.023	-0.053	-0.274
Overweight	-0.080 **	-0.006	0.026	0.007
Obese	-0.161 **	0.031	0.036	-0.064
Pain: Mild	0.063 *	0.099 **	0.134 **	0.159 **
Pain: Mod/sev/extr	0.246 **	0.228 **	0.139 *	0.220 **
Sleeping problems: Mild	0.022	0.005	-0.047	-0.032
Sleeping problems: Mod/sev/extr	0.030	0.054	-0.019	-0.068
Mobility problems: Mild	0.099 **	0.104 **	0.127 *	0.087
Mobility problems: Mod/sev/extr	0.131 **	0.191 **	0.286 **	0.102
Concentration problems: Mild	-0.040	-0.074 **	-0.027	0.015
Concentration problems: Mod/sev/extr	0.121 **	0.052	-0.161 **	0.104 *
Shortness of breath: Mild	0.048	0.054	0.039	-0.013
Shortness of breath: Mod/sev/extr	0.035	-0.074	-0.005	0.037
Depression: Mild	-0.018	-0.016	0.018	0.095 *
Depression: Mod/sev/extr	0.057	0.053	0.064	0.048
Age - 55	0.011 **	0.001	0.007	0.009 *
(Age - 55) squared /100	-0.057 **	0.014	-0.008	-0.025
Secondary education	-0.045	-0.047	-0.066	-0.066
Post-secondary education	-0.051	-0.107 **	-0.050	0.021
Living with spouse or partner	0.022	0.001	0.042	-0.065
Log HH income - log(12763)	-0.043 **	-0.025 *	-0.039 *	-0.033 *
Constant	0.110 *	0.127 **	-0.009	0.119 *
Observations	1,069	1,234	626	678
R^2	0.426	0.445	0.373	0.439

Notes: weighted results, country dummies omitted

D Parameter estimates of an ordered probit model with individual specific thresholds for each health domain

This appendix presents all parameter estimates of the ordered probit model with constant thresholds, and ordered probit model with individual specific thresholds for selected health domains, including parameters in Table 10 (* significant at 5%; ** significant at 1%).

		Pain		Mobility problems		Concentration problems	
Equation	Variable	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds
SA	Heart attack	0.244 **	0.210 *	0.091	0.011	0.067	-0.024
	High blood pressure	0.062	0.048	0.002	-0.022	0.055	0.033
	High blood cholesterol	0.035	-0.034	-0.068	-0.076	0.124 *	0.113 *
	Stroke	0.361 **	0.325 *	0.367 **	0.102	0.404 **	0.150
	Diabetes	0.360 **	0.497 **	0.514 **	0.623 **	0.259 **	0.336 **
	Chronic lung disease	0.133	0.264 *	0.017	0.100	0.043	0.163
	Asthma	0.371 **	0.341 *	0.091	0.165	0.005	0.106
	Arthritis	0.547 **	0.513 **	0.123 *	0.113	0.068	0.025
	Osteoporosis	0.293 **	0.176	0.228 **	0.160	0.070	-0.024
	Ulcer	0.109	0.172	0.116	0.033	0.157	0.164
	Parkinson disease	0.868 *	0.490	0.994 **	0.579	0.435	0.355
	Cataracts	-0.232 **	-0.260 **	-0.054	0.000	-0.267 **	-0.156
	Hip or femoral fracture	-0.029	-0.148	-0.169	-0.284	-0.289	-0.559 **
	Reproductive cancer	0.236	0.265	0.252	0.284	0.153	0.220
	Other cancer	0.461 **	0.475 **	0.399 **	0.284 *	-0.157	-0.148
	Pain in back	0.622 **	0.690 **	0.507 **	0.564 **	0.229 **	0.293 **
	Heart trouble	0.215 *	0.231 *	0.304 **	0.402 **	0.090	0.100
	Breathlessness	0.169 *	0.169	0.285 **	0.317 **	0.391 **	0.368 **
	Persistent cough	0.167	0.073	0.003	-0.308 *	0.097	-0.183
	Swollen legs	0.122	0.132	0.273 **	0.253 **	0.111	0.143
	Sleeping problems	0.197 **	0.236 **	0.119 *	0.021	0.208 **	0.134 *
	Falling down	0.090	0.061	0.060	0.237	-0.006	0.027
	Fear of falling down	0.292 **	0.026	0.448 **	0.203 *	0.124	-0.066
	Dizziness	0.236 **	0.197 *	0.324 **	0.247 **	0.187 *	0.069
	Stomach problems	0.245 **	0.140	0.047	0.072	0.077	0.097
	Incontinence	0.195	0.018	0.062	-0.107	0.144	-0.106
	Other symptoms	0.178	0.033	0.357 **	0.152	0.535 **	0.379 **
	Low grip strength	0.228 **	0.188 **	0.410 **	0.432 **	-0.024	0.033
	Underweight	0.460	0.278	-0.220	-0.198	-0.110	-0.286
	Overweight	0.100 *	0.104 *	0.159 **	0.105	0.027	0.086
	Obese	0.274 **	0.333 **	0.343 **	0.218 **	-0.200 **	-0.139 *
	Age - 55	-0.004	0.001	0.007	0.008	0.012 *	0.004
	(Age - 55) squared /100	0.043 *	0.040	0.044 *	0.049 *	0.036	0.068 **
	Secondary education	0.021	-0.033	0.080	-0.033	-0.105 *	-0.075
	Post-secondary education	-0.076	-0.250 **	-0.042	-0.293 **	-0.195 **	-0.285 **
	Living with spouse or partner	-0.058	-0.096	-0.155 **	-0.208 **	-0.175 **	-0.253 **
	Log HH income - log(12763)	0.021	-0.009	0.004	0.013	-0.007	-0.026
	medit	-0.150 *	-0.115	-0.442 **	-0.322 **	-0.286 **	-0.147 *
	female	0.280 **	0.237 **	0.022	0.071	-0.001	0.007
	medit*female	-0.100	-0.179	0.141	0.071	0.273 **	0.128
	Constant	-0.320 **	-0.507 **	-0.899 **	-1.042 **	-0.091	-0.154

		Pain		Mobility problems		Concentration problems	
Equation	Variable	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds
Thres. 1	Heart attack	.	-0.045	.	-0.090	.	-0.110
	High blood pressure	.	-0.070	.	-0.048	.	-0.013
	High blood cholesterol	.	-0.075	.	0.011	.	-0.004
	Stroke	.	0.129	.	-0.305 **	.	-0.319 **
	Diabetes	.	0.304 **	.	0.131 **	.	0.047
	Chronic lung disease	.	0.116	.	0.097	.	0.190 *
	Asthma	.	-0.195	.	0.047	.	0.052
	Arthritis	.	-0.162 **	.	-0.011	.	-0.091 *
	Osteoporosis	.	-0.179 *	.	-0.039	.	-0.082
	Ulcer	.	0.098	.	-0.191 *	.	-0.004
	Parkinson disease	.	-0.699	.	-0.601	.	-0.221
	Cataracts	.	-0.034	.	0.082	.	0.099
	Hip or femoral fracture	.	-0.356	.	-0.171	.	-0.440 **
	Reproductive cancer	.	-0.091	.	-0.027	.	0.088
	Other cancer	.	-0.045	.	-0.219 *	.	0.019
	Pain in back	.	0.019	.	0.035	.	0.031
	Heart trouble	.	-0.012	.	0.053	.	0.028
	Breathlessness	.	0.076	.	0.051	.	0.009
	Persistent cough	.	-0.264 **	.	-0.364 **	.	-0.409 **
	Swollen legs	.	0.043	.	-0.074	.	0.027
	Sleeping problems	.	0.057	.	-0.096 *	.	-0.066
	Falling down	.	-0.061	.	0.214 *	.	0.014
	Fear of falling down	.	-0.399 **	.	-0.354 **	.	-0.237 **
	Dizziness	.	-0.163 *	.	-0.042	.	-0.137 *
	Stomach problems	.	-0.121 *	.	0.065	.	0.008
	Incontinence	.	-0.388 **	.	-0.120	.	-0.189 *
	Other symptoms	.	-0.116	.	-0.340 **	.	-0.178 *
	Low grip strength	.	0.013	.	0.019	.	0.095 *
	Underweight	.	-0.148	.	0.078	.	-0.381
	Overweight	.	-0.011	.	-0.065	.	0.050
	Obese	.	0.132 **	.	-0.145 **	.	0.097 *
	Age - 55	.	0.013 **	.	0.001	.	-0.013 **
	(Age - 55) squared /100	.	-0.044 *	.	0.005	.	0.053 **
	Secondary education	.	-0.106 *	.	-0.140 **	.	0.045
	Post-secondary education	.	-0.283 **	.	-0.325 **	.	-0.101 *
	Living with spouse or partner	.	-0.081 *	.	-0.070 *	.	-0.104 **
	Log HH income - log(12763)	.	-0.040 *	.	0.006	.	-0.018
	medit	.	-0.048	.	0.136 **	.	0.190 **
	female	.	-0.142 **	.	0.022	.	-0.013
	medit*female	.	0.064	.	-0.029	.	-0.127 *
	Constant	0.000	0.000	0.000	0.000	0.000	0.000

		Pain		Mobility problems		Concentration problems	
Equation	Variable	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds	Constant thresholds	Ind. spec. thresholds
Thres. 2	Heart attack	.	0.001	.	0.043	.	0.008
	High blood pressure	.	0.074 *	.	0.083 *	.	0.001
	High blood cholesterol	.	0.005	.	-0.028	.	-0.034
	Stroke	.	-0.231 *	.	0.114	.	0.063
	Diabetes	.	-0.224 **	.	-0.068	.	0.059
	Chronic lung disease	.	-0.022	.	-0.024	.	-0.127
	Asthma	.	0.154 *	.	0.019	.	0.066
	Arthritis	.	0.115 **	.	-0.004	.	0.089 *
	Osteoporosis	.	0.037	.	-0.076	.	-0.037
	Ulcer	.	-0.048	.	0.285 **	.	0.031
	Parkinson disease	.	0.269	.	0.222	.	0.114
	Cataracts	.	0.032	.	-0.076	.	0.077
	Hip or femoral fracture	.	0.274 *	.	0.184	.	0.402 **
	Reproductive cancer	.	0.131	.	0.106	.	-0.078
	Other cancer	.	0.024	.	0.248 **	.	-0.010
	Pain in back	.	0.055	.	0.008	.	0.043
	Heart trouble	.	0.017	.	0.027	.	-0.055
	Breathlessness	.	-0.080	.	-0.081	.	-0.087
	Persistent cough	.	0.160 *	.	0.160	.	0.204 **
	Swollen legs	.	-0.042	.	0.153 **	.	0.005
	Sleeping problems	.	-0.055	.	-0.032	.	-0.048
	Falling down	.	0.055	.	-0.125	.	0.025
	Fear of falling down	.	0.115	.	0.212 **	.	0.092
	Dizziness	.	0.125 *	.	-0.136	.	-0.004
	Stomach problems	.	-0.002	.	-0.075	.	0.009
	Incontinence	.	0.189 *	.	-0.113	.	-0.090
	Other symptoms	.	-0.056	.	0.334 **	.	-0.018
	Low grip strength	.	-0.097 **	.	-0.077	.	-0.099 **
	Underweight	.	-0.063	.	-0.208	.	0.330 *
	Overweight	.	0.012	.	0.036	.	0.010
	Obese	.	-0.150 **	.	0.050	.	-0.044
	Age - 55	.	-0.009 **	.	0.000	.	0.012 **
	(Age - 55) squared /100	.	0.051 **	.	-0.002	.	-0.055 **
	Secondary education	.	0.073 *	.	0.088 *	.	-0.021
	Post-secondary education	.	0.184 **	.	0.243 **	.	0.055
	Living with spouse or partner	.	0.067 *	.	0.074	.	0.073 *
	Log HH income - log(12763)	.	0.021	.	-0.005	.	0.000
	medit	.	0.153 **	.	-0.052	.	-0.118 **
	female	.	0.142 **	.	0.078	.	0.049
	medit*female	.	-0.229 **	.	-0.131	.	-0.049
	Constant	1.200 **	-0.060	0.832 **	-0.505 **	1.039 **	-0.022
Vign. 1	Constant	.	0.494 **	.	0.668 **	.	0.549 **
Vign. 2	Constant	.	1.602 **	.	1.235 **	.	1.282 **
Vign. 3	Constant	.	2.086 **	.	1.400 **	.	1.902 **
ln ω	Constant	.	0.000	.	0.000	.	0.000
ln σ	Constant	.	-0.359 **	.	-0.670 **	.	-0.352 **
ln φ	Constant	.	-0.889 **	.	-0.866 **	.	-0.798 **
Obs.			3,607		3,607		3,607
LR test			735.9		927.5		886.2
p(LR test)			0.0		0.0		0.0
LR test $\varphi = 0$			299.1		517.1		488.0
p(LR test $\varphi = 0$)			0.0		0.0		0.0

Notes: weighted results