Physical Measures and Comparability of Health among Europeans Ages 50 and Above

Arie Kapteyn, RAND

Tatiana Andreyeva, Pardee RAND Graduate School

Erik Meijer, RAND

May 14, 2006

Acknowledgements

The authors thank the National Institute of Aging for funding under research grant P01 AG022481-01.

Abstract

One of the most intractable problems in international research on health is comparing health measures across countries or cultures. Arguably, physical measures are less influenced by cultural and linguistic differences than self-reports on general health or even on health conditions. The aim of this paper is to characterize some aspects of the health distributions within and across European countries. We model the relation between functional limitations, self-reports and physical measures (like grip strength) in a multiple indicator-multiple cause framework, where in principle all measurement scales may vary across countries, with the exception of grip strength. Using the estimated model we apply an Oaxaca type decomposition to disentangle the effect of differences in demographic composition and the effect of institutions and culture. We find that the observed large differences in health across European countries appear to be the result of different institutional environments and cultures rather than differences in the demographic composition of the nations.

1. Introduction

One of the most intractable problems in international research on health is the comparability (or incomparability) of health measures across countries or cultures. The conventional approach to evaluating health within and across nations relies heavily on using measures of subjective health assessment such as self-reports of health and health conditions. Arguably, these measures are conditioned by cultural or social norms, differences in thresholds for medical diagnosis and access to health care resources so that comparisons of health across different populations may be difficult or impossible with such gauges. In response to this issue, research on modeling comparable health measures has focused on finding an objective measurement tool that would provide consistent evaluation of genuine health across and within nations.

The ability to compare health across countries is a prerequisite for understanding the role of national policies and institutions in influencing behavior. Health plays a substantial role in many economic models, including models, of labor force participation, retirement, or savings decisions. Omitting health for a lack of comparable health measures may produce biased estimates of model parameters if health is correlated with the variables of interest. Although economic models differ greatly in what categorization and specific pecuniary factors they use, the reality is that economic incentives (e.g. disability benefits) are often conditioned on health. In a cross-national study of economic behavior, the use of comparable health measures helps not only to provide unbiased assessment of behavior but also to predict the effects of changes in policies. Based on comparable measures of health, we can evaluate effectiveness of different policy initiatives, assess health interventions across countries, and set priorities for intervention.

The analysis of inequalities in health within and across nations locates another dimension of research that needs comparable health measures. Health inequalities that are generally traced to inequalities in income, education and other socioeconomic categories persist in all countries but there are cross-national differences in their level, rate of change and strength of association (Carlson 1998, Kopp 2000, Kunst et al. 2005, Macinko et al. 2003, Van Doorslaer et al. 1997).

Efforts to develop comparable, composite measures of population health have a long history. Yet, despite many efforts to develop a consistent instrument to measure health, there seems to be no standard that is universally accepted (WHO 2002). The measures developed to date differ methodologically (on, for example, the use of weights for health problems or coverage of health domains) and conceptually (composite measures of individual health vis-à-vis population health). From a conceptual point of view, indices designed to capture the detailed components of individual health require a different set of considerations than more general, population-oriented health status measures. The latter include generalizable data on mortality, the prevalence, incidence and natural history of non-fatal conditions, prevalence-based valuations for the disability weights associated with these conditions (WHO 2002).

One of the best known summary measures of population health is the disability-adjusted life year (DALY) that made its debut in the World Development Report 1993 presented by the World Bank. The DALY measures the gap between a population's actual health and some explicit goal, and is calculated as the present value of the future years of disability-free life that is lost, to all causes, whether from premature mortality or from some degree of disability during a period of time (Erikson et al. 1995, WHO 2002).

Another common summary measure of population health used by the World Health Organization (WHO) is the Disability-Adjusted Life Expectancy (DALE) that measures the expected number of years of life to be lived in full health, or healthy life expectancy. DALE estimates are based on the estimates of severity-weighted disability prevalence developed for the non-fatal component of disease and injury burden (WHO 2002).

There is no single instrument to monitor population health in the US. Measures used by the US government agencies include the Centers for Disease Control and Prevention Health-Related Quality-of-Life 14-Item Measure (CDC HRQOL-14) "Healthy Days Measures"[1] and the Health and Activity Limitation Index (HALex), also known as Years of Healthy Life (YHL) that is based on age-specific mortality rates, activity limitation and self-rated overall health (Erikson 1998, Sondik 2002, Stewart et al. 2005). YHL was used as a summary health measure in Healthy People 2000, the primary prevention initiative in the US. For the Healthy People 2010 program no single summary measures have been identified. About 20 leading health indicators serve as a summary set of nation's health measures (Sondik, 2002).

While summary measures like DALE, YHL or HRQOL are useful for comparisons of overall health across countries and for the measurement of progress of one nation's health, they offer limited power in measuring the current health status of an individual that is essential in economic models. Self-rated health and reports of doctor-diagnosed chronic conditions have been common measures of the current individual

[1] The CDC HRQOL-14 combines three modules: "Healthy Days core module" that evaluates self-reported general health, number of days over the last 30 days in ill physical health, mental health or health-related limited functional ability; "Activity Limitations Module" with 5 questions about physical, mental, or emotional problems or limitations in daily life, and "Healthy Days Symptoms Module" with 5 questions about the number of days in the past 30 days when experiencing specific symptoms. More detail about the CDC HRQOL-14 is available at the CDC website http://www.cdc.gov/hrqol/hrqol14_measure.htm.

health status in most types of modeling. On the one hand, health self-assessment is a good measure of health because the question has high response rates and predictive power for other health measures and mortality. On the other hand, self-reported health evaluations are subject to many biases related to differences in culture, language and institutional environment. In international comparisons of health, it is impossible to separate the observed variation in the subjective health responses into the variation in genuine health and the differences related to cultural or social norms. Furthermore, self-reported measures may be unstable over short periods of time.

Recent innovations in the design and data collection of some household surveys make it possible to construct internationally comparable health measures using more objective and accurate evaluation of health than self-perceived health. These advances include collecting information on physical measures like grip strength and walking speed in cross-national multidisciplinary studies such as the Survey of Health, Ageing and Retirement in Europe (SHARE), as well as preparations for such data collection in other household surveys (the US Health and Retirement Study (HRS) and the English Longitudinal Study of Ageing (ELSA)). Interviewers take physical measures of health like grip strength using the same protocol across all countries. Such assessments are therefore less likely to be subject to the biases affecting self-reports of health, and may overcome the measurement issues of cultural differences in how people evaluate their health.

The primary objective of this paper is to construct internationally comparable measures of health. To address the scaling issues inherent in cross-national comparability of subjective health questions, we develop a model that relies on objective health indicators like grip strength to arrive at comparable scales. The second objective is to compare health inequality within and across countries. A third objective is to decompose the health measures so as to understand the sources of the variation in health. We will use an Oaxaca decomposition approach to isolate the effect of different demographic compositions of populations from the effect of different estimated model parameters, assumed to represent institutions, culture, and genetic endowments.

Section 2 of the paper describes the data, while Section 3 presents our model. The empirical results are discussed in Section 4. There we also present the outcomes of some model simulations and decompositions. Section 5 provides a further interpretation and discussion of the results.

2. The Data

We use information collected in the first (2004) wave of the Survey of Health, Ageing and Retirement in Europe (SHARE), which is a multidisciplinary cross-national longitudinal survey of Continental Europeans over the age of 50 and their spouses. The baseline 2004 SHARE study includes data on 11 countries providing a balanced representation of the different European regions from Scandinavia (Denmark and Sweden) through Central Europe (Austria, France, Germany, Switzerland, Belgium, The Netherlands) to the Mediterranean (Spain, Italy and Greece). We use data from SHARE Release 1 (April 28th, 2005), in which data for Belgium are not available yet.

Designed after the role models of the HRS and ELSA, the 2004 SHARE combines information on health (e.g., self-reported health, physical and cognitive functioning, health behaviors, health care utilization and expenditure), psychological conditions (e.g., psychological health, well-being, life satisfaction), socio-economic status (e.g., work activity, job characteristics, income, wealth and consumption, housing, education), and social support (e.g., social networks, volunteer activities).

The 2004 SHARE Release I sample includes 22,777 respondents from 10 European countries. The survey has been administered as computer assisted personal interviewing (CAPI) in the fall of 2004 to probability samples of individuals of 50 and over in all participating countries. The sampling plan follows a complex probabilistic multistage design to produce estimates representative of the non-institutionalized population aged 50 and above in each country. The study also interviews spouses younger than 50. The response rate varies by country but on average is 57% for households and 86% for individuals within participating households. A detailed description of the SHARE data and methodology has been published elsewhere (Borsch-Supan et al, 2005). The data is available to registered users from the SHARE website (http://www.share-project.org).

We impose several sample restrictions. First, we exclude individuals younger than 50 (759 observations or 3.3% of the original sample). The second exclusion (3317 observations or 14.5% of the original sample) is for the data with missing or incomplete responses on at least one measure used in our estimation. Due to these selections, 19,460 individuals remain eligible for the analysis (9055 men and 10,405 women). We find that individuals in worse health have a higher probability of being dropped from the sample according to our inclusion rule, but this tendency is similar across countries. To obtain estimates representative of the non-institutionalized population over the age of 50, we use individual sample weights when presenting sample statistics.

Measures

SHARE contains extensive modules on physical health combining information on subjective health assessment (based on the US categorization on the five-point scale from “poor” to “excellent” and the European categorization on the five-point scale from “very bad” to “very good”), indicators of doctor-diagnosed chronic conditions (heart disease, high blood cholesterol, hypertension, stroke, diabetes, lung disease, asthma, arthritis/rheumatism, osteoporosis, cancer, ulcer, Parkinson’s disease, cataracts, hip fracture), a battery of functional limitations from more severe limitations with activities of daily living (ADL) to less disabling problems with instrumental activities of daily living (IADL) and mobility limitations. In addition SHARE contains a limited number of physical measures, including self-reported body weight and height, interviewer-measured walking speed (for adults aged 76 and older) and grip strength (for all respondents).

Grip strength is a core physical measure of health that potentially enables cross-national comparability of health estimates and avoids some of the endogeneity problems inherent in more subjective health measures like self-rated health. It also helps to overcome the measurement issues related to biases that arise from subjectivity of self-reported health and health conditions due to cultural differences across and within countries, differential physician contacts or cross-national differences in the criteria for thresholds of medical diagnosis. At the same time, predictive validity of grip strength for assessing health was established in studies that found grip strength be to a better predictor of future medical problems than self-reported health (Christensen et al., 2000; Rantanen et al. 1999, 2001 Al Snih et al. 2002).

Table 1 summarizes the distribution of our analysis sample across countries and gender for selected health indicators. We report four measures of subjective health assessment in SHARE: any limitation with ADL, IADL, mobility limitations, and self-reports of fair or poor general health. The distribution of the data on self-reported health is particularly illustrative of the large cross-country differences embedded in self-reports. For example, the percentage of men who rated their health as poor or fair is more than four-fold larger in Germany than in Sweden, whereas approximately the same proportion of men in both countries report having some chronic health condition (~70%). Another example is the male population of Denmark, whose life expectancy is on average one year less than of French men, but who are 25% less likely than the French to rate their health as poor/fair.

Table 2 presents the mean value of grip strength measurements across countries and gender. The cross-national variation in grip strength is much smaller than the observed differences in self-reports of health. The difference between the highest and the lowest average national measurements is about 25% for both men and women. In all countries, the average grip strength of women is around two-thirds of the average level for men.

SHARE asked respondents to report whether they had any difficulties doing various activities because of a health or physical problem in the last month before the interview (difficulties expected to last less than three months were excluded). We have selected 25 indicators to measure health and functional ability in SHARE, including reports of limitations with 10 activities of mobility, arm function and fine motor function, 6 ADLs, 7 IADLs, self-reported health, and grip strength. Table 3 describes all indicators, and presents their structure in modeling health (we will return to this structure later in the text).

We use a set of standard socio-demographic covariates in modeling physical health and functional ability. These include a third degree age polynomial, educational achievement (secondary and tertiary education, primary or no education was the reference group), household size, household income (an inverse hyperbolic sine[2] of annual household income before taxation), and marital status (married and living together with a spouse or registered partnership). We also include a measure of relative body weight to account for the well-documented effects of excessive body weight or obesity on physical health and functioning. Individuals are classified by relative weight based on their body mass index (BMI), calculated 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 stratify the study respondents into five weight classes: underweight (BMI<18.5), normal weight (BMI: 18.5-24.9), overweight (BMI: 25.0-29.9), moderate obesity (BMI: 30.0-34.9), and severe obesity (BMI: >=35.0) (National Heart, Lung, and Blood Institute). The sample size for extreme obesity (BMI: >=40.0), another weight class in the NIH guidelines, is too small to enable meaningful analysis. We divide the obesity group into moderate and severe obesity because there are differential health effects by degree of obesity. Severe obesity is associated with more chronic health problems than moderate

[2] The inverse hyperbolic sine function *h(.)* is defined as follows: $h(x) = \ln[x + \sqrt{(x^2 + 1)}]$, for any *x*. Note that *h(x)=h(-x)*. For positive values of *x* the inverse hyperbolic sine is similar to a logarithm, but it has the advantage of also being defined for values less or equal to zero.

obesity, and the onset is at earlier ages (Field et al, Hillier and Pedula, Must et al). Table 4 presents socio-demographic characteristics and body weight distribution of the sample.

3. The Model

Our model is of the Multiple Indicators Multiple Causes type. We postulate a limited number of latent health variables (in the empirical work the number of latent variables will chosen to be equal to two). The health variables are not directly observable, but instead we observe the *indicators* ADLs, IADLs, mobility indicators, self reported health, and grip strength. In turn the latent health variables are assumed to be causally related to a number of observable variables, like education, age and other demographics as well as unobserved factors. In the modeling we make a distinction between grip strength and other indicators for reasons that will become apparent below.

In writing down the model we follow the familiar LISREL notation

$$y_{cn} = \Lambda_{yc}\eta_{cn} + \varepsilon_{ycn} \qquad (1)$$

$$g_{cn} = \Lambda_{g}\eta_{cn} + D_c h_{cn} + \varepsilon_{gcn} \qquad (2)$$

The first equation links the vector of indicators *y* (with subscripts *c* and *n* to indicate country and individual respectively) to the latent health vector η_{cn} and a vector of i.i.d. errors ε_{ycn}. Notice that the matrix of parameters Λ_{yc} has a subscript *c* indicating that parameters may be different across countries. The second equation explains grip strength as a function of the latent health vector η_{cn} and a vector of i.i.d. errors ε_{gcn}; the vector h_{cn} contains polynomials in height and weight to allow for the fact that grip strength will be affected by an individual's weight and height. The vector D_c may vary across countries. Notice that Λ_g does not have a country subscript signifying that the relation between health and grip strength is assumed to be the same across countries, once the effect of height and weight is accounted for.

The vector η_{cn} depends on a vector of demographics x_{cn} as follows:

$$\eta_{cn} = \Gamma_c x_{cn} + \zeta_{cn} \qquad (3)$$

The dimension of η_{cn} is not determined a priori, but the result of preliminary exploratory factor analysis. It turns out that two dimensions provide an adequate description of the data, as will be shown in the next section. Table 3, provides the assumed structure of the matrix Λ_{yc}. An X indicates a non-zero entry, whereas all other entries are set equal to zero. The elements of the row vector Λ_g are set equal to minus one without loss of generality. This partly fixes the scale of the various parameters in the model. Appendix A provides an analysis of the identifiability of the parameters in the model[3]. The conclusion of that analysis is that one more normalizing assumption needs to be made to fix the scale of all parameters. We choose "difficulty with dressing, including putting on shoes and socks" as a second indicator (next to grip strength) with equal coefficients across countries.

4. Results

Preliminary factor analysis suggests that two factors do a reasonable job of providing a description of the correlation structure of the 25 indicators of functional limitations and self-perceived health. This was generally found for all countries in the by-country analysis. When we pool the data for all countries we find that a larger number of

[3] This appendix will be added to the next version of the paper.

factors should be retained (See Table B.1 in Appendix B), but for comparability of the pooled results and the results by country we also retain two factors for the pooled sample. The factor analysis results are similar for men and women in all countries.

We therefore retain two dimensions of health, which we define as physical health and functional ability. The indicators have been grouped into the physical health or functional ability category based on each indicator's adequacy for the definition of a physically demanding functional task. Table 3 gives the structure of the indicators by the group of physical health and the group of functional ability.

Table 5 presents the estimation results of the structural parameters from the pooled model (i.e. equation (3)). Higher values of the structural estimates indicate more impaired functional ability and worse physical health. A positive sign means that the corresponding variable is associated with a lower level of physical health status or functional ability. Many of the variables are significant, particularly for women, yet the magnitudes of some coefficients are modest. We observe that the variables predicting physical health generally are associated with higher estimated coefficients in terms of both statistical significance and the size of the coefficients.

We find that our estimates square well with the findings in the literature on the role of socio-demographics and obesity for health. For example, we note that people with low education are more likely to have limited functional ability and poor health, whereas household income plays a significant role only for physical health. Married men and women are more likely to be in better health, yet the size of the household seems to statistically matter only for the functional ability of women (perhaps indicating that women with disabling conditions are likely to live in homes with their children). We observe that across the weight groups the largest negative effects for physical health and functional ability appear for underweight older individuals. Among other weight groups, severe obesity and moderate obesity are correlated significantly with adverse physical health of all individuals and functional limitations among women.

The estimates of age effects are not straightforward to interpret because we used the polynomial structure to model the age effects on health. Figures 1 to 4 in Appendix B depict the age polynomials for both health measures in the pooled sample. As one would expect, the age polynomials are monotonically increasing in physical health for men and women. They are almost constant for functional limitations before a rise at the age of about 70 for women and 65 for men. Appendix B also provides graphs of age polynomials for selected countries.

The next tabulation (Table 6) presents the parameters of the vector D_c in equation (2). These estimates cannot be interpreted individually because of the polynomial and interaction structure that we used.

Table 7 shows the estimates of the country dummies from the pooled model by gender. We interpret the dummies as measures of average health, after controlling for demographics. Switzerland is taken as the reference country. Compared to Switzerland, physical health and functional ability in the other countries seem to be worse after controlling for demographics. For males Spain, Germany and Austria are estimated to perform particularly poorly on the physical health scale; Austrian males also do poorly on the functional ability scale together with males in France and Denmark. For females, Spain is once again the country with the highest (worst) score for physical health, followed by France and Greece; Like their male counterparts, Danish females fare poorly on the functional ability scale, now followed by Spain and Italy. The correlation between the country estimates for men and women is high (the correlation coefficient is 0.71 for physical health and 0.66 for functional ability).

We have conducted a test to compare the pooled model (where parameters are fixed for all countries) to a country-specific model (where parameters are varied by country except for those in Λ_g in the grip strength equation (2)). It turns out that the pooled model is rejected against the model with country parameters. From now on, we therefore focus our attention on the analysis and interpretation of the results from the country-specific models that follow.

Using the estimated models one can simulate the distribution of health in the various countries. The simulated means of the health indices, based on the by country models are presented in Table 8. Similar to the analysis of the pooled model (Table 7), Switzerland serves as the reference country, so that countries with negative values of the health indices are estimated to have populations in better physical health or functional ability than the Swiss. The physical health indices vary substantially more by country than the functional ability measures. We also observed this when we compared countries based on the estimated country dummies in the pooled model (Table 7).

We observe substantial variability among the means of the health variables. The mean health indices for men and women are highly correlated across countries: the correlation coefficient is 0.93 for physical health and 0.68 for functional ability. The within gender correlation between the two health instruments is relatively large for women (0.63) but less strong for men (0.32). The ranking of countries is different from what is found when using the pooled model. Regarding male physical health, Spain, Italy, and Greece are at the bottom, with Denmark scoring best. Male functional ability appears to be best in The Netherlands, with France, Greece, and Italy scoring worst. Female physical health is best in Austria, while Spain, Italy, and Greece score worst. Female functional ability is best in The Netherlands, with France, Italy, and Spain at the bottom

Table 9 presents the dispersion of the simulated health indices by country. The within country variation of the health indices is clearly a lot less than the across variation of means. The variation in the health indices of men appears to be somewhat larger than the variation in the women's indices.

We have conducted a set of simulation exercises to look at how the variability in the health instruments change once we remove the observed heterogeneity in socio-demographic variables. We estimated six scenarios where we assigned everyone in the sample to have the median household income of one specific country (for example, Austria, Switzerland or Greece), the same age (50, 60 or 70-years old), and education (primary, secondary or tertiary education). Figures B.1.1 to B.2.3 in Appendix B show results from the base estimation and the simulated scenarios. We present simulations for two polar cases where we assign everybody to have socio-demographic characteristics that are associated with better health (tertiary education, younger age (50) and high income, i.e. the median household income in Switzerland) or poorer health (primary education, older age (70), and the median income in Greece).

Given that the variation in physical health across countries appears to be larger than within countries, one would not expect this exercise to have dramatic effects on the ranking of countries. Yet, there are some interesting changes taking place if we move from the base case to the two polar cases. The lowest ranked countries do not change position if we artificially change the demographics. For both male and female physical health the lowest ranking countries are Spain, Italy, Greece, and France no matter if we consider the base case, the case with a favorable demographic composition or the case with an unfavorable demographic composition. Among the top-ranked countries we do see some movement in relative rank depending on whether we consider favorable or unfavorable demographics. For instance, for both male and female physical health, Germany is ranked as the healthiest country if we assign favorable demographics, but as number 5 if we assign unfavorable demographics. The Netherlands on the other hand is ranked number 5 in average physical health for both males and females with favorable demographics and number 2 and number 3 respectively for male and female physical health with unfavorable demographics.

The rankings of countries with respect to functional ability show a fair amount of sensitivity with respect to the demographic scenario we impose. For instance for male functional ability, Greece is at the bottom if we assign favorable demographics and moves to second place if we assign all countries unfavorable demographics. On the other hand, a country like France scores at or near the bottom for both males and females under any scenario. The Netherlands scores at the top of the ranking, except for females when we assign favorable demographics; in that case The Netherlands scores in the middle.

As a final approach to comparing health across countries and disentangling the effect of demographic composition from the estimated parameters in the by country models we apply an Oaxaca decomposition. We interpret the effect of the estimated parameters as the effect of institutional differences and possibly genetic differences. These may include cross-national differences along at least five dimensions such as: 1) environment, climate, and geography; 2) culture and lifestyle; 3) health care system, 4) social capital and crime, and 5) genes. For brevity we will just refer to these factors as institutional differences.

We have experimented with each country to serve as a basis in comparisons, yet the results were robust to the choice of the reference country. We found the effect of the population's composition to be small compared to the effects of the institutional differences. The results presented in Table 10, take Switzerland as the reference country. In line with outcomes of the exercise above, we observe that the variation in the socio-demographic composition of the SHARE countries explains only a minor part of the variability in the physical health indices. For functional ability the situation is different; the effects of demographics is somewhat smaller than the effect of institutions, but the order of magnitudes of the effects are similar.

5. Interpretation and Conclusions

We constructed and estimated a model to compare health across countries using physical measures of health, ADLs, IADLS, mobility measures, and self reported health. Important for our approach to the international comparisons is that we assume that the relation between certain measurements (in particular grip strength) and the underlying health dimensions is the same in all countries. We find that our data can be reasonably described by two dimensions of health. The findings are indicative of the critical role that institutions, environment, cultural and social norms play in the health of Europeans ages 50 and above. The observed large differences in the average health indices across European countries appear to be the result of different institutional environments and perhaps genetic disposition, more than than of differences in the demographic composition of the nations.

Previous work on developing a cross-national composite index of health has primarily looked at mortality, morbidity experiences, and health self-reports across countries with some adjustment for quality of life. Due to data paucity these indices have often excluded factors that might affect health such as lifestyle patterns, work-related stress, employment, and household structure. Another problematic part of health comparability research has been an absence of a benchmark for health or some universally accepted standard that all results could be compared to and the instruments validated against. As a result, the validity of the proposed health indices often had to be tested using sensitivity analysis and qualitative procedures.

Our approach of modeling health as a vector of latent variables and assuming that some measures can be reasonably compared across countries goes part way towards addressing these issues. We have furthermore validated our health constructs using several methods. A correlation analysis of the health indices and unadjusted data on functional limitations, assessment of general health, and health conditions has been conducted by gender for each country. We observe that the physical health measure correlates relatively well (with the correlation coefficient of about 0.2-0.3) with the measures of functional limitations and disability such ADL-disability, IADL-disability, and limited mobility functioning (disability defined as reporting at least one limitation in each case). The correlation between the physical health variable and self-assessed health is also about 0.2-0.3. We however do not find a strong correlation between our health variables and health conditions. This is consistent with the fact that in the data reports of doctor-diagnosed health conditions are hardly correlated with most measures of functional limitations and self-rated health. The exceptions are arthritis and heart disease with stronger effects for women. In contrast to doctor-diagnosed chronic illnesses, depression is highly related to both self-assessed health and functional limitations[4].

We checked for the validity of our measures using their stratification by socio-demographic group, and compared cross-country differences in the means of the health indices by education and income. We found health inequalities across nations to exceed socio-demographic differences in health within countries– the result that we have also observed in the Oaxaca decomposition analysis and simulations.

Finally, we have compared our estimates with the mortality data from various global sources like the World Health Organization (WHO). We find little correlation between our data and mortality statistics from the WHO.

Clearly, our model invites further improvements and proposes directions for the future research avenue. On the data side, another cross-validation technique could be a comparison of our estimates with the anchoring vignette data that have been collected in subsamples in most SHARE countries . We could consider a wider array of health indices including self-reports of chronic health conditions, and compare how sensitive the estimates are to changes in the model specification.

[4] Depression was assessed using the EURO-D scale that summarized respondents' experiences with any of the 12 feelings, including sadness or depression, pessimism, suicidality, guilt, sleep trouble, interest, concentration, appetite, irritability, fatigue, enjoyment, and tearfulness. The threshold for depression was scoring 4 and above on the EURO-D depression scale.

References

References

Al-Snih, S., K. Markides, L. Ray, G. Ostir, and J. Goodwin (2002), "Handgrip Strength and Mortality in Older Mexican Americans," *Journal of the American Geriatric Society* 50(7): 1250-6.

Bardage C, Pluijm S, Pedersen N, Deeg D, Jylha M, Noale M, et al. (2005) Self-rated health among older adults: a cross-national comparison. *European Journal of Ageing*, 2:149-158.

Borsch-Supan A, Brugiavini A, Jurges 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. Manheim: MEA.

Borsch-Supan A, Jurges H (eds) (2005). *Health, ageing and retirement in Europe - methodology*. Manheim: MEA (forthcoming).

Christensen, K., M. McGue, A.I. Yashin, I.A. Iachine, N.V. Holm, J.W. Vaupel (2000), "Genetic and Environmental Influences on Functional Abilities Among Danish Twins Aged 75 Years and Older," *Journal of Gerontology: Medical Sciences* 55A(8), M446-M452.

Carlson P, (1998). Self-perceived health in East and West Europe: another European health divide. *Social Science and Medicine*. 46:1355-66.

Erickson P, (1998). Evaluation of a population-based measure of quality of life: the Health and Activity Limitation Index (HALex). *Quality of Life Research*. 7(2):101-14.

Erickson P, Wilson R, Shannon I, (1995) *Years of healthy life*. (CDC/NCHS, Healthy people, Statistical notes no. 7). US Department of Health and Human Services, National Center for Health Statistics, Hyattsville, MD.

Field A, Coakley E, Must A, et al. (2001). Impact of overweight on the risk of developing common chronic diseases during a 10-year period. *Arch Intern Med*.;161(13):1581-6.

Hillier T, Pedula K. (2001) Characteristics of an adult population with newly diagnosed type 2 diabetes: the relation of obesity and age of onset. *Diabetes Care*, 24(9):1522-7.

Kopp MS, Skrabski A, Szedmak S, (2000). Self-rated health and social transitions. In: Nilsson P, Orth-Gomer K (eds) Self-rated health in a European perspective. The Swedish Council for Planning and Coordination of Research, Ord & Form AB, Uppsala, pp 85-102.

Kunst AE, Bos V, Lahelma E, Bartley M, Lissau I, Regidor E, Mielck A, et al. (2005) "Trends in socioeconomic inequalities in self-assessed health in 10 European countries." *International Journal of Epidemiology*, 34:295-305.

Macinko JA, Shi L, Starfield B, Wulu JT, (2003). Income inequality and health: a critical review of the literature. *Medical Care Research and Review*. 60(4):407-52.

Mathers C, Sadana R, Salomon J, Murray C, Lopez A. (2000) Estimates of DALE for 191 countries: methods and results. Global Program on Evidence for Health Policy Working Paper No. 16. World Health Organization.

Must A, Spadano J, Coakley E, Field A, Colditz G, Dietz W. (1999) The disease burden associated with overweight and obesity. *JAMA* 282(16):1523-9.

National Center for Health Statistics (2000) Developing an index of health: approaches, current developmental activities, and implications. (A report to the Senate Committee on Appropriations by the National Center for Health Statistics.) National Center for Health Statistics, Center for Disease Control and Prevention, Washington, DC.

National Heart, Lung, and Blood Institute (1998) *Clinical Guidelines on the Identification, Evaluation and Treatment of Overweight and Obesity in Adults: The Evidence Report*. Rockville, MD: National Heart, Lung, and Blood Institute.

Rantanen, T., J.M. Guralnik, D. Foley, K. Masaki, S. Leveille, J.D. Curb, and L. White (1999), "Midlife Hand Grip Strength as a Predictor of Old Age Disability, *Journal of the American Medical Association 281*, 6, 558-560.

Rantanen, T., T. Harris; S.G. Leveille, M. Visser; D. Foley, K. Masaki, J.M. Guralnik (2001)," Muscle Strength and Body Mass Index as Long-Term Predictors of Mortality in Initially Healthy Men, *Journal of Gerontology, Biological Sciences and Medical Sciences* 55(3): M168-73.

Sondik E. (2002). "Summary measures of population health: applications and issues in the United States". Chapter 2.4 in *Summary measures of population health: concepts, ethics, measurement and applications*. eds. Murray, C., Salomon, J., Mathers, C. and A. Lopez. World Health Organization, Geneva.

Stewart S, Woodward RM, Cutler DM, (2005). A proposed method for monitoring US population health: linking symptoms, impairments, chronic conditions, and health ratings. NBER Working Paper 11358.

US Department of Health and Human Services (2000) *Healthy people 2010*. Public Health Service, Washington, DC.

World Bank (1993) *World development report: investing in health*. Oxford University Press, New York.

World Health Organization (2000). *The World Health Report 2000 Health Systems: Improving Performance*. World Health Organization: Geneva.

World Health Organization (2002). *Summary Measures of Population Health. Concepts, Ethics, Measurement and Applications*. World Health Organization: Geneva.

van Doorslaer E, Wagstaff A, Bleichrodt H, et al (1997). Income-related inequalities in health: some international comparisons. *Journal of Health Economics* 16:93-112.

Table 1

Health Measures by SHARE Country

	Sample Size	ADL (at least one limitation), %	IADL (at least one limitation), %	Mobility (at least one limitation), %	Fair/poor self-reported health, %
			Men		
Austria	723	6.4 (24.5)	9.7 (29.6)	41.6 (49.3)	24.6** (43.1)
Germany	1263	6.7 (25.0)	9.5 (29.3)	45.9** (49.8)	35.5** (47.8)
Sweden	1316	5.8* (23.4)	8.2 (27.5)	33.3** (47.1)	8.3** (27.6)
The Netherlands	1256	6.1 (23.9)	9.8 (29.7)	31.0** (46.3)	24.9** (43.3)
Spain	887	6.9 (25.5)	13.9** (34.6)	39.9 (49.0)	32.1 (46.7)
Italy	1019	7.3 (26.0)	6.8** (25.3)	38.9 (48.8)	30.4 (46.0)
France	661	11.3** (31.7)	10.4 (30.5)	35.0** (47.7)	29.5 (45.6)
Denmark	720	9.3 (29.1)	10.7 (30.9)	33.5** (47.2)	23.7** (42.5)
Greece	797	4.3** (20.2)	8.9 (28.6)	43.3 (49.6)	23.9** (42.7)
Switzerland	413	4.2** (20.0)	4.3** (20.3)	27.9** (44.9)	13.1** (33.8)
Total	*9055*	*7.5 (26.4)*	*9.6 (29.4)*	*39.6 (48.9)*	*29.9 (45.8)*
			Women		
Austria	897	8.3 (27.6)	19.4 (39.6)	56.7 (49.6)	25.4** (43.6)
Germany	1400	9.3 (28.9)	14.8** (35.6)	59.2 (49.1)	39.2 (48.8)
Sweden	1419	8.8 (28.3)	17.4 (37.9)	51.8** (49.9)	11.7** (32.1)
The Netherlands	1367	7.1** (25.7)	17.4 (37.9)	48.3** (50.0)	25.9** (43.8)
Spain	1139	10.8 (31.1)	24.6** (43.1)	60.9* (48.8)	45.7** (49.8)
Italy	1207	10.4 (30.5)	15.9 (36.6)	56.1 (49.6)	43.9** (49.6)
France	826	10.6 (30.8)	19.8 (39.8)	57.5 (49.5)	32.9** (47.0)
Denmark	802	9.1 (28.8)	19.8 (39.9)	48.5** (50.0)	24.3** (42.9)
Greece	905	7.9 (26.9)	21.7** (41.2)	63.3** (48.2)	34.3 (47.5)
Switzerland	443	7.5 (26.4)	10.6** (30.8)	44.0** (49.7)	17.1** (37.7)
Total	*10405*	*9.7 (29.6)*	*17.8 (38.3)*	*57.3 (49.5)*	*37.0 (48.3)*

Note: Data are presented as percentages. Means adjusted for sample weights. Standard deviation is in parentheses. * Significantly different from overall SHARE mean at $p<0.05$. ** Significantly different from overall SHARE mean at $p<0.01$.

Table 2

Physical Measure of Health by SHARE Country

	Grip strength, kg **Men**	**Grip strength, kg** **Women**
Austria	45.8** (9.91)	28.9** (7.73)
Germany	46.0** (10.92)	28.3** (7.77)
Sweden	45.0** (9.96)	26.4** (7.28)
The Netherlands	45.7** (10.38)	27.8** (7.50)
Spain	37.7** (10.43)	22.7** (7.55)
Italy	39.8** (11.20)	23.1** (7.27)
France	42.8 (10.50)	25.7 (7.21)
Denmark	46.7** (10.46)	26.9** (7.32)
Greece	41.4** (10.49)	25.0** (6.72)
Switzerland	44.4** (9.39)	27.3** (7.06)
Total	*42.8* *(11.13)*	*25.8* *(7.78)*

Note: Data are presented as the maximum measurement of two grip strength measurements in kilograms. Means adjusted for sample weights. Standard deviation is in parentheses.

* Significantly different from overall SHARE mean at p<0.05.
** Significantly different from overall SHARE mean at p<0.01.

Table 3

Indicators for Modeling Health indices

Description	**Physical health measure**	**Functional ability measure**
ADLs		
Dressing, including putting on shoes and socks		X
Walking across a room		X
Bathing or showering		X
Eating, such as cutting up food		X
Getting in or out of bed		X
Using the toilet, including getting up or down		X
IADLs		
Using a map to figure out how to get around in a strange place		X
Preparing a hot meal		X
Shopping for groceries		X
Making telephone calls		X
Taking medications		X
Doing work around the house or garden	X	X
Managing money, such as paying bills and keeping track of expenses		X
Mobility, arm and fine motor function limitations		
Walking 100 meters	X	
Sitting for about two hours	X	
Getting up from a chair after sitting for long periods	X	
Climbing several flights of stairs without resting	X	
Climbing one flight of stairs without resting	X	
Stooping, kneeling, or crouching	X	
Reaching or extending arms above shoulder level	X	
Pulling or pushing large objects like a living room chair	X	
Lifting or carrying weights over 10 pounds/5 kilos, like a heavy bag of groceries	X	
Picking up a small coin from a table		X
Self-reported general health		
Five-point scale of the US categorization from “poor” to “excellent”	X	X
Grip strength		
Maximum measurement of both hands or one hand	X	X

Note: “X” means using an indicator in the estimation for a particular measure.

Table 4

Socio-demographic characteristics of the SHARE Sample

	Men **N=9055**	**Women** **N=10405**
Age, years	63.9 (9.7)	65.6 (10.5)
Educational achievement, %		
No or primary education	43.4 (49.6)	56.1 (49.6)
Secondary education	35.0 (47.7)	30.1 (45.9)
Tertiary education	21.6 (41.1)	13.8 (34.5)
Household size, no of people	2.3 (1.0)	2.0 (1.0)
Annual household income before taxes PPP-adjusted, Euros	51,864 (93,155)	44,417 (102,759)
Married or registered partnership, %	76.8 (42.2)	55.9 (49.6)
Relative body weight, %		
Underweight	0.3 (5.6)	1.8 (13.2)
Normal weight	32.8 (46.9)	43.8 (49.6)
Overweight	50.4 (50.0)	36.4 (48.1)
Moderate obesity	13.5 (34.2)	13.6 (34.2)
Severe obesity	3.0 (17.1)	4.4 (20.6)

Note: Means adjusted for sample individual weights. Standard deviation is in parentheses.

Table 5

Structural Estimates for the Pooled Estimation

	Age	Secondary education	Tertiary education	House-hold size	Marital status	Household income	Under weight	Overweight	Obese	Severely obese	Age second polynomial	Age third polynomial
						Men						
Functional ability												
Estimate	0.753	-0.030	-0.045	0.0039	-0.029	-0.0003	0.7584	-0.0065	0.0093	0.0587	-0.1515	0.0100
St.er.	(0.546)	(0.0148)	(0.0176)	(0.0063)	(0.0163)	(0.0038)	(0.1974)	(0.0111)	(0.0164)	(0.0346)	(0.0836)	(0.0043)
Physical health												
Estimate	-9.4124	-0.6090	-0.9909	0.0219	-0.2869	-0.1190	5.9763	0.1266	1.2189	2.3674	1.2872	-0.0466
St.er.	(4.337)	(0.1101)	(0.1218)	(0.0519)	(0.1246)	(0.0320)	(0.7885)	(0.0924)	(0.1401)	(0.2747)	(0.6362)	(0.0307)
						Women						
Functional ability												
Estimate	5.6366	-0.0624	-0.0762	0.0337	-0.0878	0.0019	0.2358	0.0306	0.1058	0.2320	-0.9334	0.0516
St.er.	(0.918)	(0.0193)	(0.0234)	(0.0097)	(0.0202)	(0.0048)	(0.0615)	(0.0165)	(0.0245)	(0.0426)	(0.1407)	(0.0072)
Physical health												
Estimate	-6.9734	-0.5489	-0.7458	0.0461	-0.2509	-0.0494	0.8501	0.6274	1.7371	3.0914	0.9811	-0.0359
St.er.	(3.189)	(0.0791)	(0.0964)	(0.0392)	(0.0785)	(0.0203)	(0.2448)	(0.0705)	(0.1029)	(0.1703)	(0.4673)	(0.0225)

Note: Estimates for country dummies are not reported.

Table 6

Structural Estimates for the Pooled Estimation

	Height	Weight	Weight squared	Height squared	Height*Weight
			Men		
Estimate	-124.996	1.7927	0.0815	36.6640	-2.3272
St.er.	(21.2198)	(0.9792)	(0.0167)	(6.8741)	(0.6127)
			Women		
Estimate	17.3189	-0.1635	0.0233	-8.1364	-0.4188
St.er.	(15.8054)	(0.5885)	(0.0084)	(5.1496)	(0.3850)

Table 7

Country Estimates in Health Models
Pooled Sample

	Men		Women	
	Physical health	**Functional ability**	**Physical health**	**Functional ability**
Austria	1.419	0.098	0.913	0.099
Germany	1.779	0.084	1.079	0.108
Sweden	0.256	0.031	0.484	0.088
The Netherlands	0.975	0.053	0.631	0.082
Spain	1.946	0.068	1.765	0.130
Italy	1.099	0.054	1.506	0.142
France	1.301	0.105	1.069	0.098
Denmark	0.986	0.128	0.712	0.159
Greece	1.248	0.015	1.367	0.032
Switzerland	0	0	0	0

Table 8

Average Health Variables with Intercepts
Estimates by Country

	Men		Women	
	Physical health	**Functional ability**	**Physical health**	**Functional ability**
Austria	-1.955	0.063	-2.076	-0.030
Germany	-2.178	0.001	-1.492	0.016
Sweden	-1.479	0.023	0.354	0.063
The Netherlands	-1.859	-0.173	-1.188	-0.100
Spain	7.917	0.082	4.764	0.107
Italy	5.051	0.122	4.266	0.113
France	1.477	0.544	1.680	0.161
Denmark	-3.050	0.057	-0.005	-0.057
Greece	3.153	0.143	2.100	-0.056
Switzerland	0	0	0	0

Table 9

Cross-Country Dispersion of Health Variables
Estimates by Country

	Men		Women	
	Physical health	**Functional ability**	**Physical health**	**Functional ability**
Austria	0.071	0.043	0.108	0.029
Germany	0.129	0.000	0.103	0.010
Sweden	0.107	0.012	0.088	0.024
The Netherlands	0.077	0.065	0.098	0.015
Spain	0.158	0.024	0.122	0.022
Italy	0.204	0.035	0.127	0.044
France	0.142	0.042	0.097	0.039
Denmark	0.167	0.008	0.118	0.034
Greece	0.161	0.043	0.087	0.019
Switzerland	0.181	0.042	0.166	0.012

Table 10

Oaxaca Decomposition of Health indices Across SHARE Countries

	Physical Health		**Functional Ability**	
	Institutional Differences (parameter estimates)	*Difference in population's composition*	*Institutional Differences (parameter estimates)*	*Difference in population's composition*
		Men		
Austria	-2.074	-0.005	0.071	-0.001
Germany	-2.080	0.012	-0.005	0.007
Sweden	-1.798	0.272	-0.023	0.049
The Netherlands	-1.862	0.099	-0.201	0.021
Spain	7.307	0.567	-0.027	0.100
Italy	4.913	0.299	0.064	0.034
France	1.190	0.192	0.475	0.055
Denmark	-3.126	-0.098	0.055	0.011
Greece	2.666	0.467	0.087	0.028
Switzerland	Reference country			
		Women		
Austria	-2.371	0.272	-0.055	0.020
Germany	-1.725	0.325	0.012	0.014
Sweden	0.268	0.146	0.068	-0.001
The Netherlands	-1.310	0.010	-0.101	0.010
Spain	4.331	0.697	0.048	0.043
Italy	4.089	0.213	0.092	0.036
France	1.472	0.233	0.146	0.006
Denmark	-0.067	0.021	-0.056	-0.012
Greece	1.723	0.383	-0.096	0.027
Switzerland	Reference country			

Note: We present the results from the decomposition analysis with Switzerland as a reference country. The results are however not sensitive to the choice of a reference country.

Appendix B

Age Polynomials in the Pooled Model

Figure 1

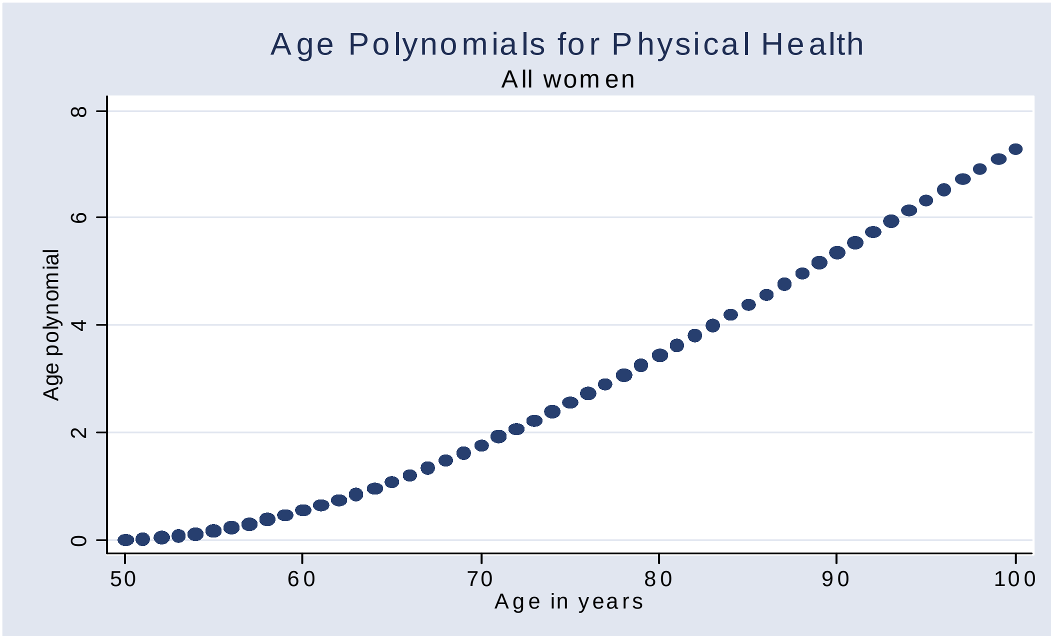


Figure 2

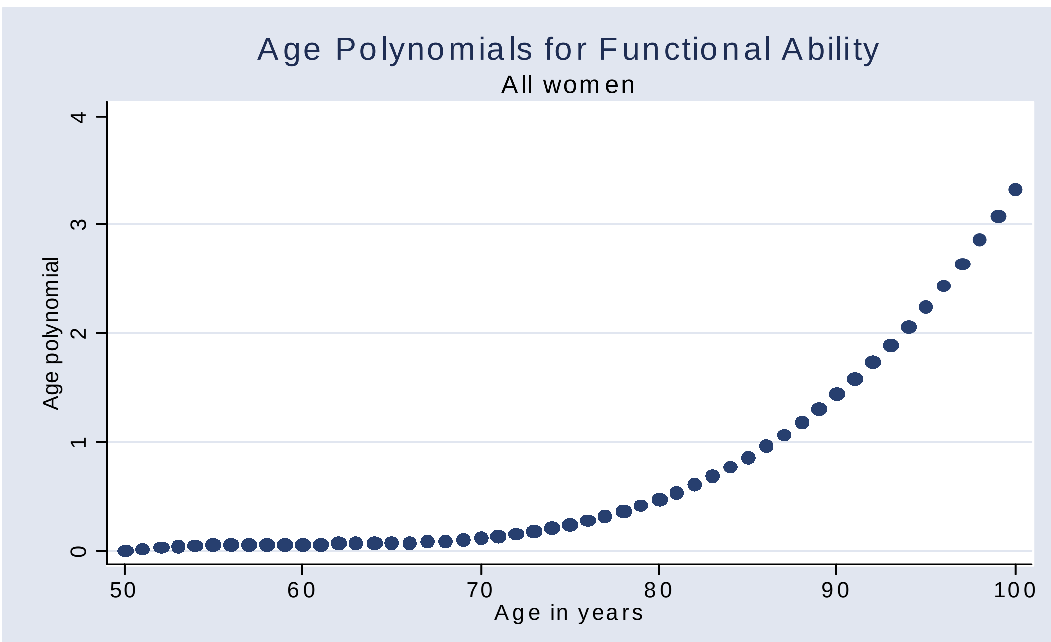


Figure 3

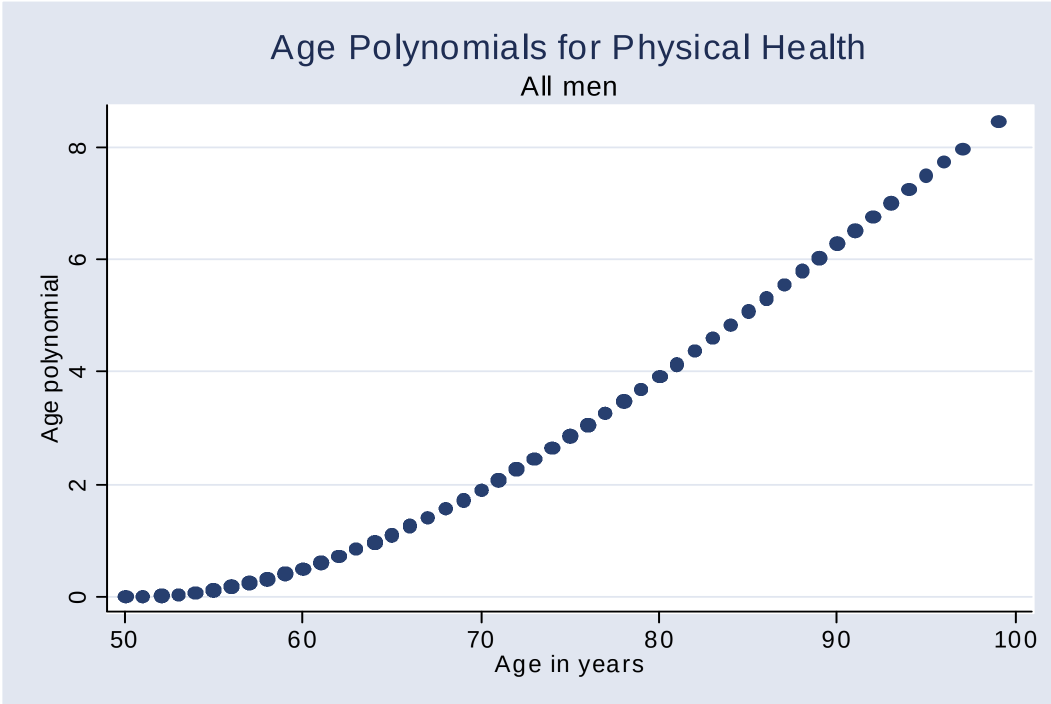
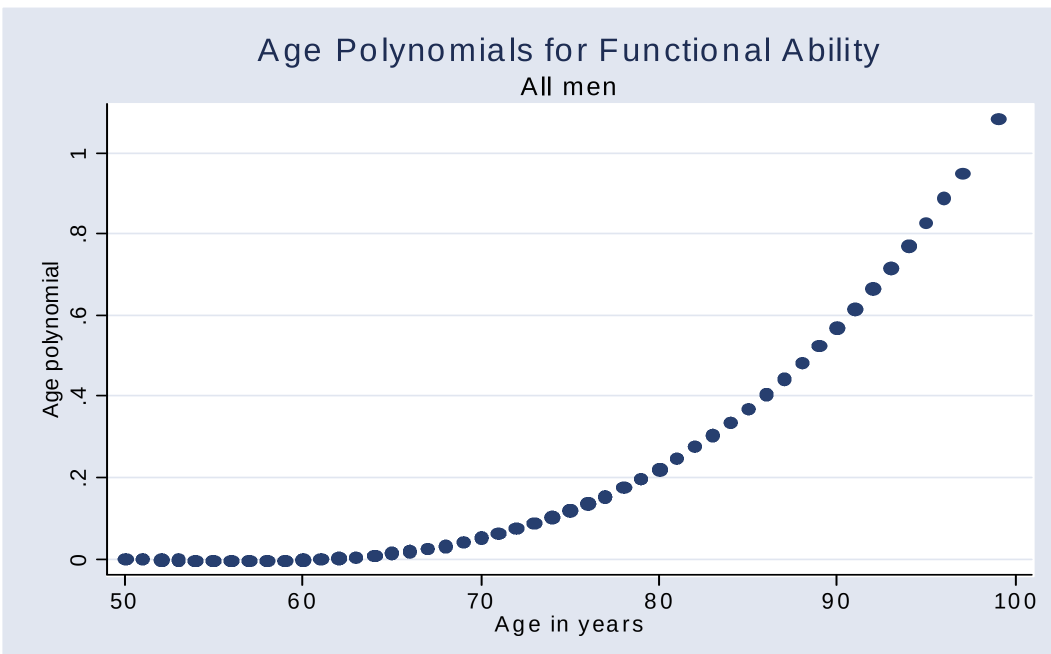
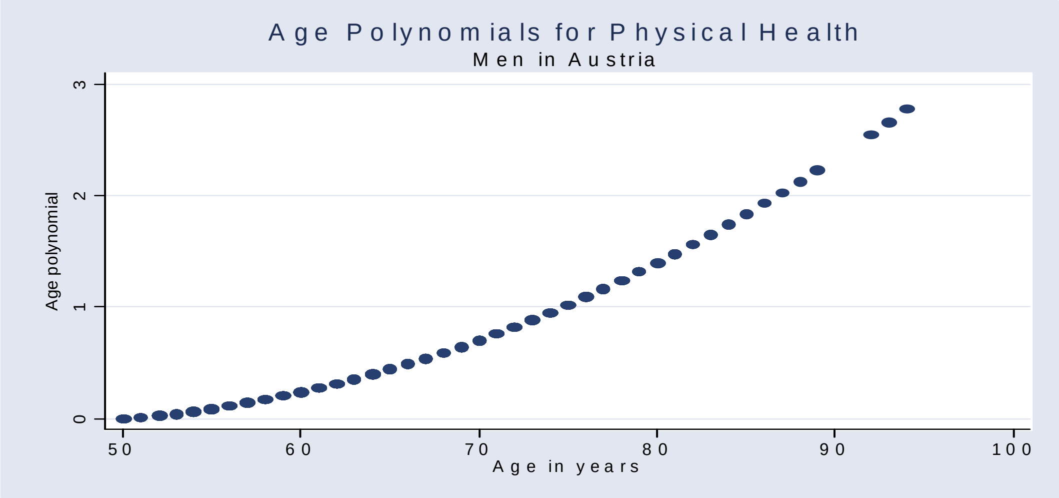
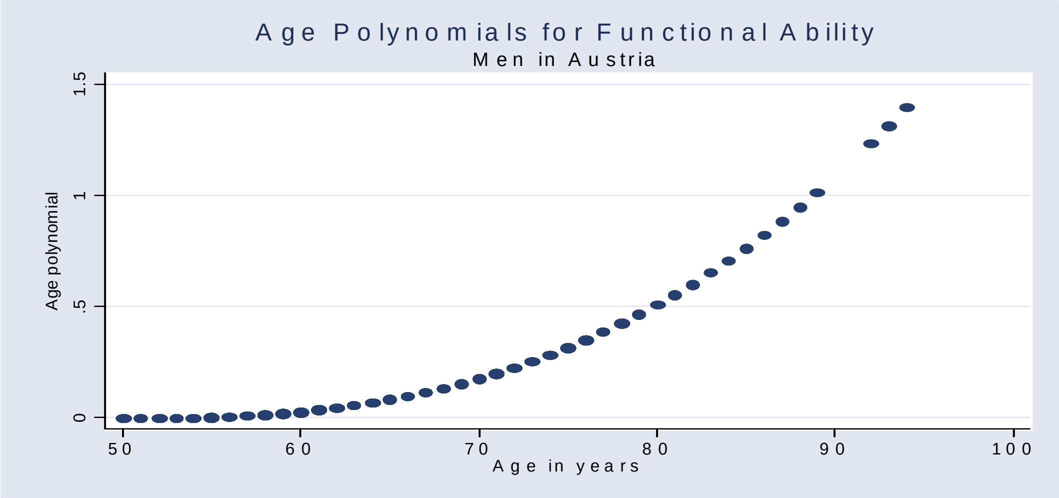


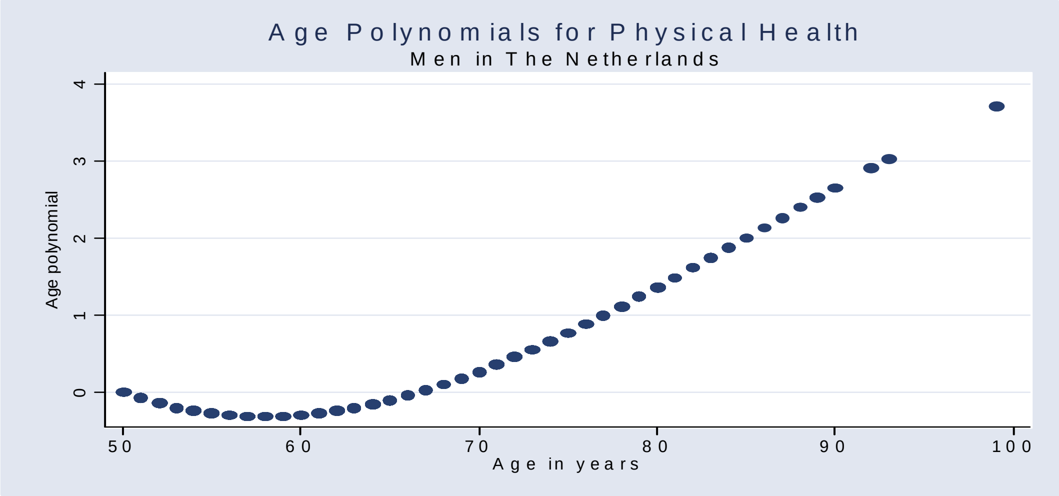
Figure 4

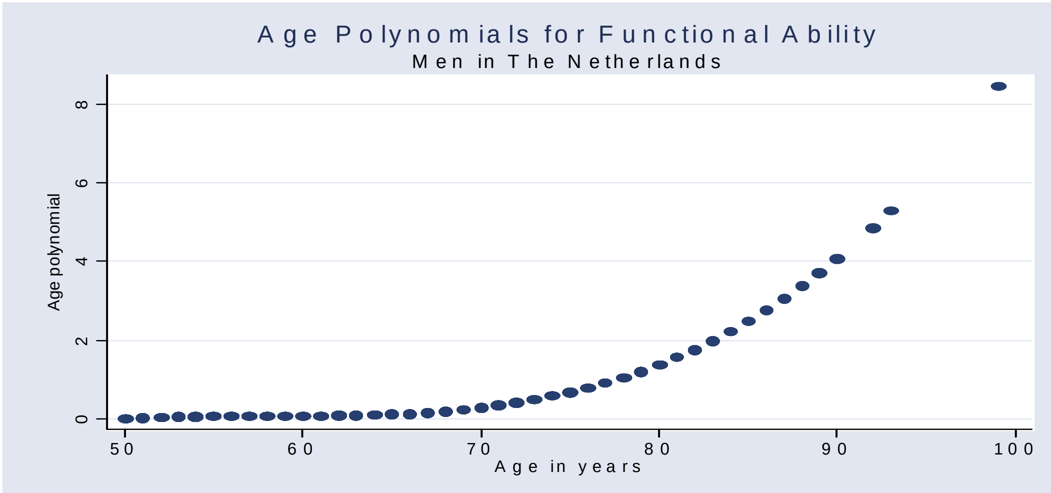


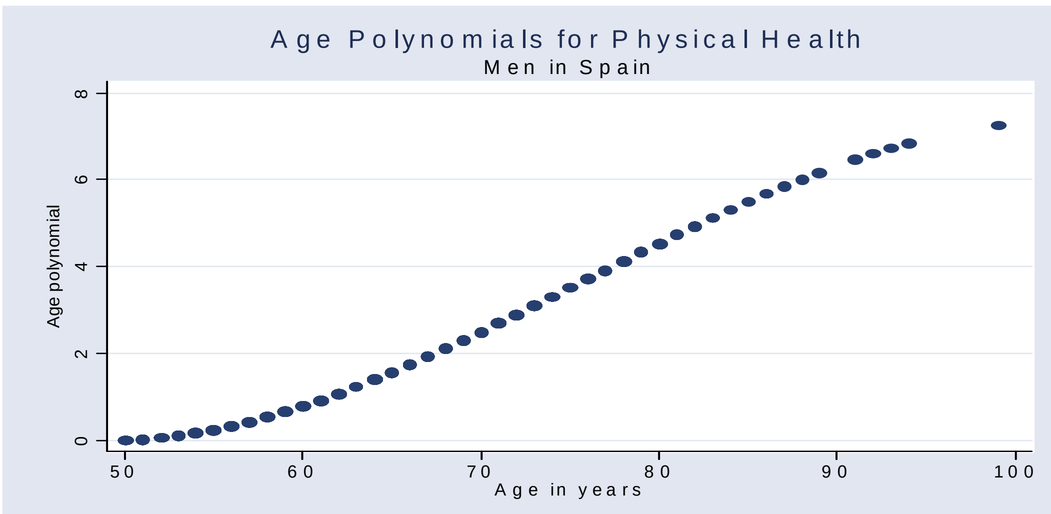
Age Polynomials in Selected Countries

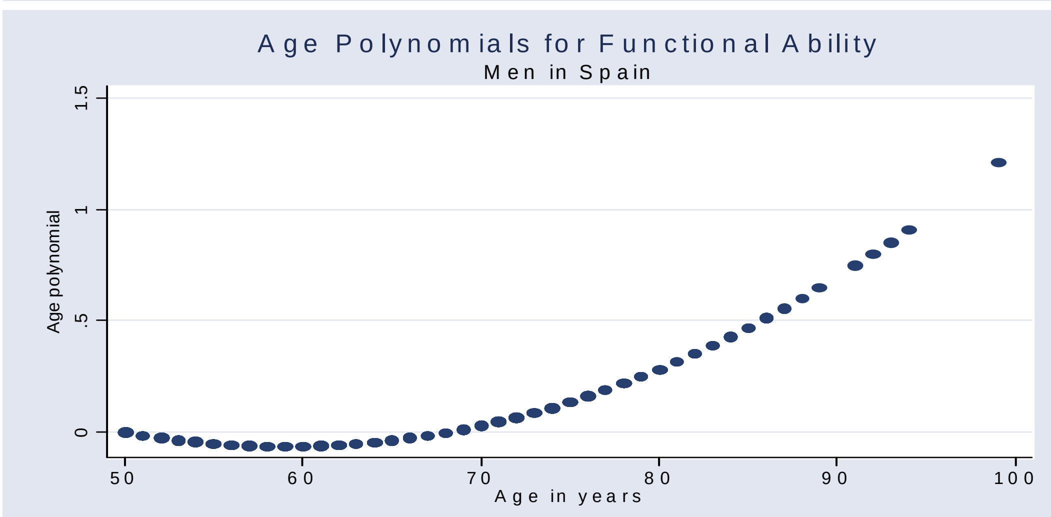


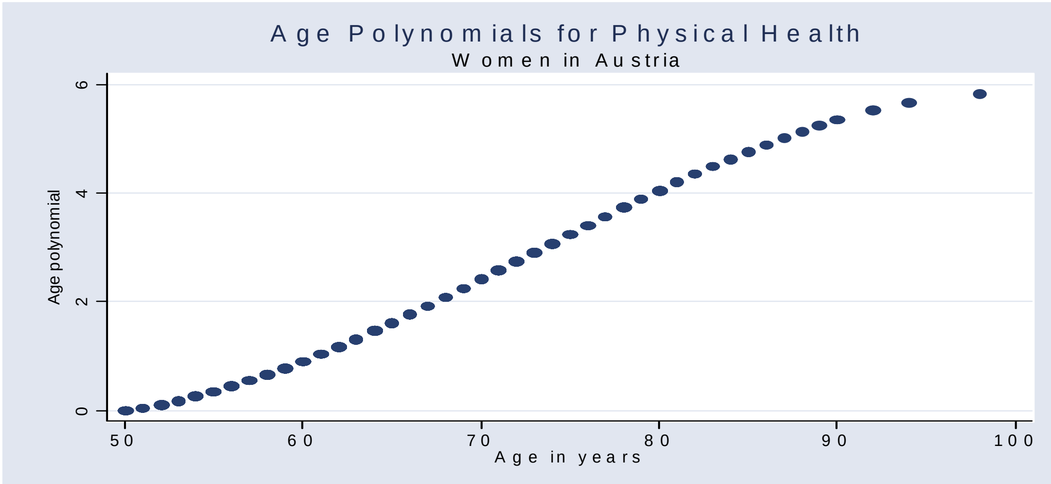


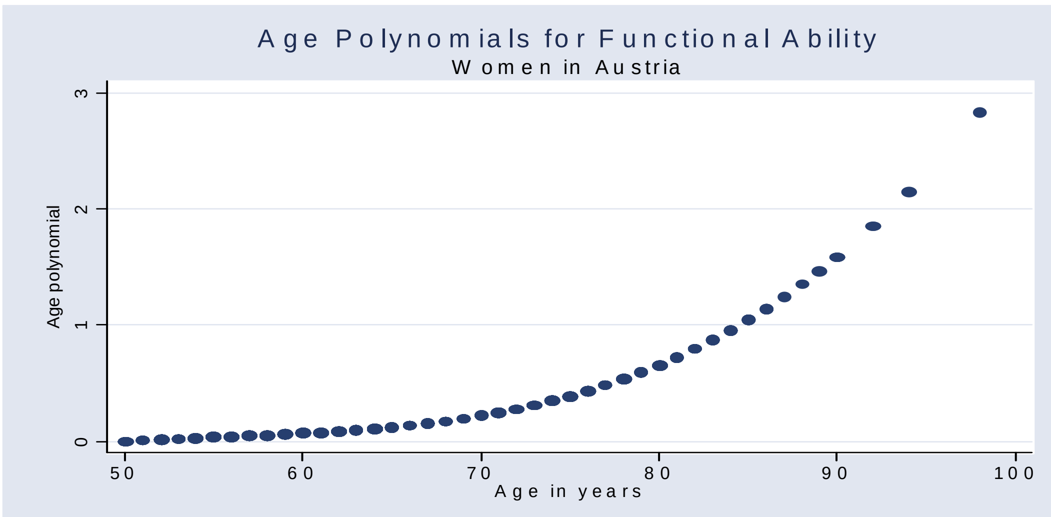


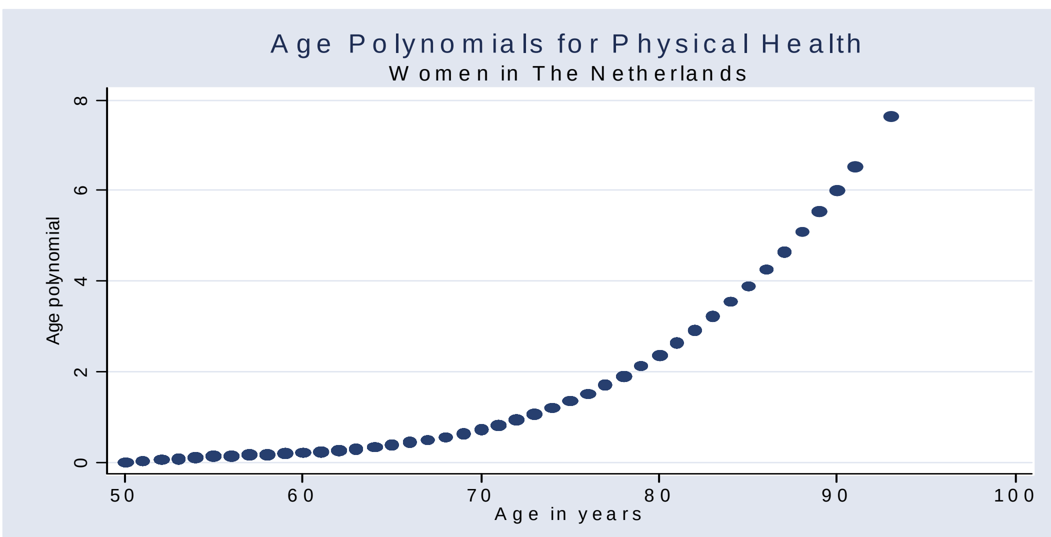

Age Polynomials for Functional Ability
Men in The Netherlands
Age polynomial
0
2
4
6
8
50
60
70
80
90
100
Age in years

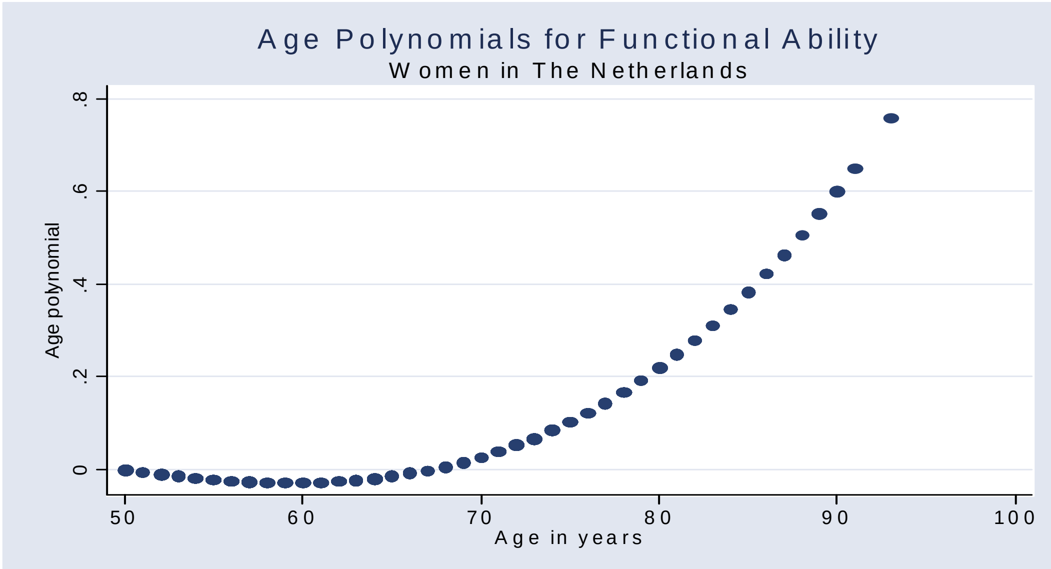

Age Polynomials for Physical Health
Men in Spain
Age polynomial
0
2
4
6
8
50
60
70
80
90
100
Age in years

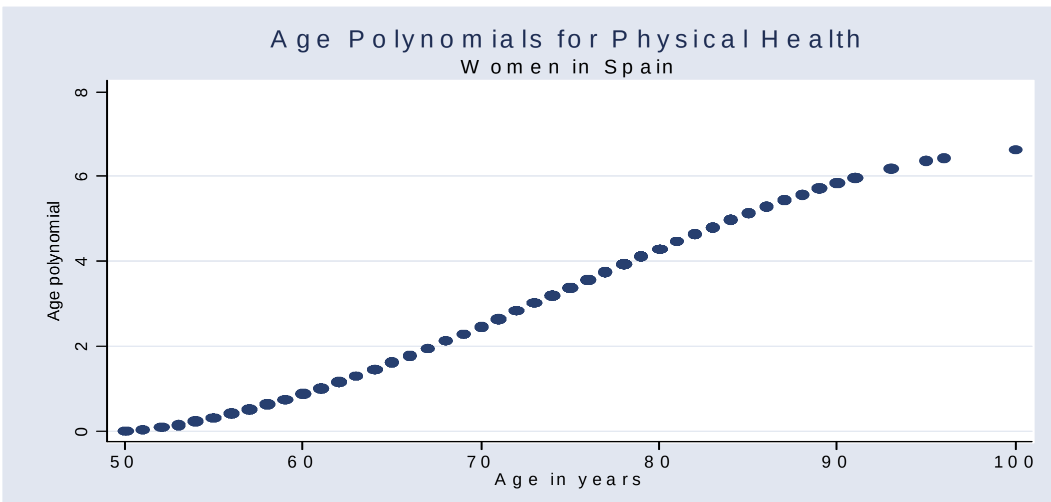

Age Polynomials for Functional Ability
Men in Spain
Age polynomial
0
.5
1
1.5
50
60
70
80
90
100
Age in years


Age Polynomials for Physical Health
Women in Austria
Age polynomial
0
2
4
6
50
60
70
80
90
100
Age in years


Age Polynomials for Functional Ability
Women in Austria
Age polynomial
0
1
2
3
50
60
70
80
90
100
Age in years


Age Polynomials for Physical Health
Women in The Netherlands
Age polynomial
0
2
4
6
8
50
60
70
80
90
100
Age in years


Age Polynomials for Functional Ability
Women in The Netherlands
Age polynomial
0
.2
.4
.6
.8
50
60
70
80
90
100
Age in years


Age Polynomials for Physical Health
Women in Spain
Age polynomial
0
2
4
6
8
50
60
70
80
90
100
Age in years

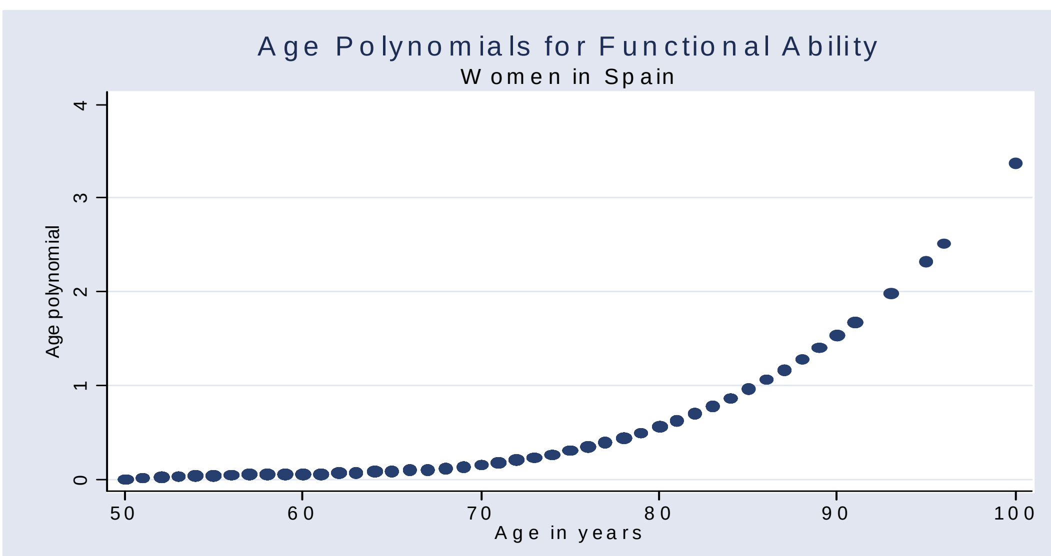

Age Polynomials for Functional Ability
Women in Spain
Age polynomial
0
1
2
3
4
50
60
70
80
90
100
Age in years

Table B.1

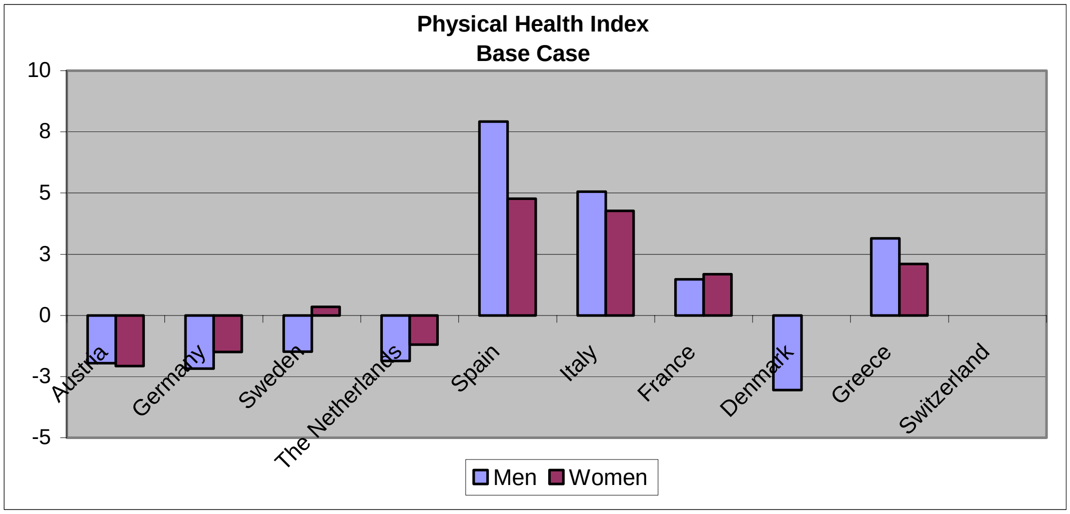
Principal-Component Factors for All Indicators in Factor Analysis

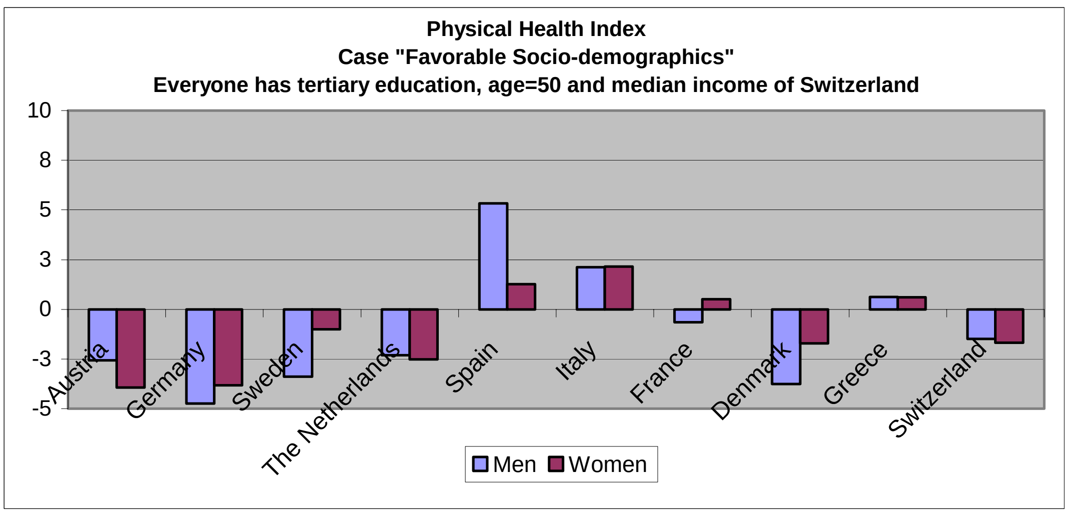
Pooled Sample

	Eigenvalue (unrotated) Factors with eigenvalues above 1 retained	Variance (orthogonal varimax rotation) 4 factors retained	Variance (orthogonal varimax rotation) 3 factors retained	Variance (orthogonal varimax rotation) 2 factors retained
		Men		
Factor 1	6.651	3.440	3.777	4.467
Factor 2	1.961	2.940	3.477	4.144
Factor 3	1.299	2.736	2.656	
Factor 4	1.113	1.907		
		Women		
Factor 1	6.731	4.214	4.332	4.582
Factor 2	2.259	3.018	3.203	4.408
Factor 3	1.349	2.818	2.805	
Factor 4	1.022	1.313		

Note: Factor analysis of a correlation matrix using the method of principal-component factors with rotation.

Figure B.1.1-B.1.3. Simulations for Physical Health Index





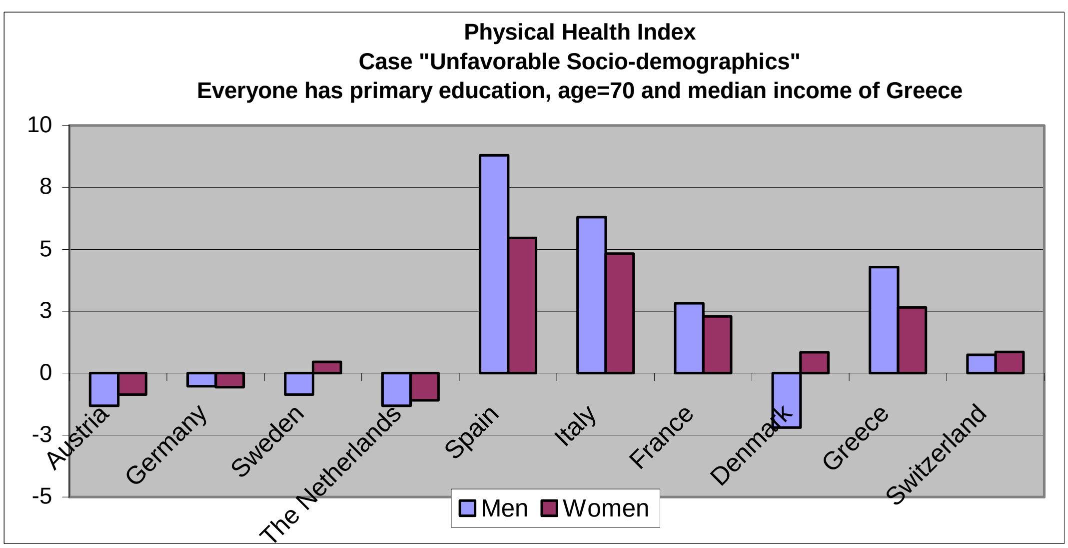


Figure B.2.1-B.2.3. Simulations for Functional Ability Index