Ruhr Economic Papers

Andrew M. Jones and Stefanie Schurer

How Does Heterogeneity Shape the Socioeconomic Gradient in Health Satisfaction?

#8

Ruhr Graduate School in Economics
ECON
RWI ESSEN

Ruhr Economic Papers

Published by

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstraße 150, 44801 Bochum, Germany

Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstraße 12, 45117 Essen, Germany

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI Essen)
Hohenzollernstrasse 1/3, 45128 Essen, Germany

Ruhr Economic Papers #8

Responsible Editor: Justus Haucap

ISSN 1864-4872 (online) – ISBN 978-3-86788-000-8

Ruhr Economic Papers

#8

Andrew M. Jones and Stefanie Schurer

How Does Heterogeneity Shape the Socioeconomic Gradient in Health Satisfaction?

Ruhr Graduate School in Economics

ECON

RWI ESSEN

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

ISSN 1864-4872 (online)
ISBN 978-3-86788-000-8

Andrew M. Jones and Stefanie Schurer*

How Does Heterogeneity Shape the Socioeconomic Gradient in Health Satisfaction?

Abstract
Individual heterogeneity plays a key role in explaining variation in self-reported well-being and, in particular, health satisfaction. It is hypothesised that the influence of this heterogeneity varies over levels of health and increases over the life-cycle. These hypotheses are tested with data on health satisfaction from 22 waves of the German Socioeconomic Panel (GSOEP). Nonlinear fixed effects methods that allow for unobserved heterogeneity are not readily available for categorical measures of well-being. One common solution is to revert to conditional fixed effects methods, at the price of a high degree of information loss. Another common solution is to ignore the association between unobserved heterogeneity and socio-economic status by using pooled or random effects models, at the price of potential bias. We use a generalization of the conditional fixed effects logit, that allows for individual-specific reporting bias, heterogeneity in health endowments, and heterogeneity in the impact of income on health satisfaction. Adjusting for unobserved heterogeneity accounts for the relationship between income and very good health, but not between income and poorer health states. The income gradient for older age-groups is more strongly affected by controlling for unobserved heterogeneity: revealing an increasing influence of heterogeneity on health satisfaction over the life-span.

JEL Classification: I12, C23

Keywords: Panel data, generalized conditional fixed effects logit, generalized ordered logit, health, GSOEP

May 2007

---

*Andrew M. Jones, University of York; Stefanie Schurer, Ruhr Graduate School in Economics and RWI Essen. – We would like to thank Nigel Rice and the Centre for Health Economics, York, UK, for their support and valuable contributions. We also would like to thank Jan Brenner, Paul Frijters, Pilar García Gomez, John Haisken-DeNew, Michael Shields, and participants of the Health, Econometrics, and Data Group at the University of York, York, UK, the Brown Bag Seminar at the RWI Essen, Essen, Germany, the First Doctoral Conference of the RGS in Economics, University of Dortmund, Germany. – All correspondence to: Stefanie Schurer, RGS Econ, c/o RWI Essen, Hohenzollernstr. 1-3, 45128 Essen, Germany. Fax: +49 201 81 49-200, email: schurer@rwi-essen.de.

1 Introduction

Individual heterogeneity plays a major role in determining self-reported life and health satisfaction. Evidence from the psychological literature suggests that 50 to 80 percent of the variation in happiness should be attributed to genetics, personality traits, or early childhood experiences (Lykken and Tellegen, 1996). The question is whether the link between life satisfaction and observable factors such as inflation (DiTella et al., 2001, 2003; Helliwell, 2004; Alesina et al., 2004), unemployment (Clark and Oswald, 1994; Clark, 2001; Winkelmann and Winkelmann, 1998), income growth (Easterlin, 1995; Blanchflower and Oswald, 2004) or relative income (Ferrer-i-Carbonell, 2005; Frijters et al., 2004a, 2006; Lelkes, 2006; Boes and Winkelmann, 2006) can be assessed reliably in the presence of confounding latent variables? How reliable are subjective well-being indicators for predicting economic outcomes or the willingness-to-pay for a public good (Oswald, 1997; Layard, 2005; Frey et al., 2006; Van Praag and Ferrer-i-Carbonell, 2005; Kahnemann and Krueger, 2006)?[1]

Similar questions arise in the study of the relationship between socio-economic status and health satisfaction, the latter being a constituent part of overall life satisfaction (Van Praag and Ferrer-i-Carbonell, 2004). Health satisfaction can be understood as an alternative measure to self-assessed health which is often considered to be a reliable predictor of mortality (Mossey and Shapiro, 1982; Idler and Benyamini, 1997; Mackenbach et al., 2002). The literature on the socio-economic gradient of health provides extensive evidence that an association between these variables exists, but both the causal pathways and the factors that explain this association remain controversial. Thus, as for overall well-being, we need to ask how much of the variation of health status is driven by latent factors such as health-related genetics and personality traits and by how much do these correlate with socio-economic status. For example, in the context of health, an interesting candidate for the composition of the unobserved effect is cognitive ability, a factor widely acknowledged by labour economists, that biases wage regressions if left unaccounted for (Card, 1995;

[1] A comprehensive and excellent review of the literature is provided by Clark et al. (2006).

Blackburn and Neumark, 1995).

The link between cognitive ability and socio-economic status has been well documented. For instance, children's IQ predicts adult socio-economic outcome better than parents' background variables (Jencks et al., 1972, 1979) and 20 to 40 percent of total observed variation in education, occupation, and earnings can be linked to genetic differences in cognitive ability (Gottfredson, 2004). In health economics cognitive ability has not yet received similar attention. One exception is Auld and Sidhu (2005) who show that both schooling and cognitive ability are strongly associated with lower health status and that intelligence accounts for about one quarter of the association between schooling and health. There is also evidence that general intelligence is a very good predictor of health outcomes (Gottfredson, 1997, 2002; Gottfredson and Deary, 2004). Children's IQ is also a reliable predictor of survival rates in older age (Whalley and Deary, 2001; Betty and Deary, 2004). Singh-Manoux et al. (2005) conclude that up to 40 percent of the relationship between socio-economic status and self-rated health can be explained by cognitive ability. Several non-exclusive explanations for the association between IQ and objective and subjective health indicators have been advanced. On the one hand, IQ could be considered a good proxy for health, representing birth complications, suboptimal postnatal care or general body integrity measured via the brain's capacity to process information rapidly and accurately. On the other hand, IQ may be an indirect proxy for health behaviour with respect to substance abuse, stress management or adherence to complex treatment regimes.

Moreover, the influence of unobserved heterogeneity on health and well-being may not necessarily remain constant over the life-span of an individual. The idea of dynamic heterogeneity has been expressed by Frijters et al. (2005a) in the context of a mixed proportional hazard model. They assume unobserved heterogeneity to grow over time due to cumulating experiences of random health shocks. In our paper, we hypothesise that the influence of unobserved heterogeneity on health and socio-economic outcomes changes nonlinearly over the life-cycle. We justify this hypothesis by several arguments. First, Scarr and McCartney (1983) argue in a theoretical paper on individual development that genetic factors become more influential in determining choices in older age. This has to do with the fact that once leaving the nurturing environment of family and compulsory schooling, individuals engage in niche-building that correlates with their talents, interests, and personality characteristics. Hence, choices determining socio-economic status (tertiary schooling, professional training, on-the-job training) and health behaviours (diet, smoking, exercise), correlates more strongly with unobserved factors such as intelligence in adulthood. To this extent, we would expect younger age-groups to exhibit less bias when not controlling for unobserved heterogeneity. Furthermore, genetic factors play an even stronger role when the limits of life-span (above 80) are reached (McClearn et al., 1997; Perls, 2002). To be able to surpass a specific age, individuals must be endowed with a genetically determined health status that helps them to suppress age-related diseases and, thus, increase life expectancy. In this respect, we would expect the oldest age-group to exhibit the greatest level of bias when not controlling for unobserved heterogeneity. Second, it has been shown that genetic factors increasingly determine individual differences in cognitive ability across age-groups (20 percent in childhood, 40 percent in adolescence, and 50 to 60 percent in adulthood) (McGue et al., 1993; Plomin, 1986; Plomin and Petrill, 1997; Bouchard, 1998). If cognitive ability proxies physical and mental functioning, then the link between unobserved heterogeneity and health outcomes should increase over the life-time as well. Third, health status generally worsens over the life-cycle due to idiosyncratic health shocks, age-related diseases and the outbreak of particular genetic diseases (Liang et al., 2005).

In light of these arguments, the appropriate econometric model to assess the socio-economic gradient of health satisfaction should capture the presence of unobserved heterogeneity and its correlation with socio-economic factors. Fixed-effects panel data methods, that allow for a correlation between unobserved heterogeneity and the regressors of the model, are not, however, readily available for categorical ordered data due to what is called the "incidental parameter problem" (Neyman and Scott, 1948). To assess self-reported health data researchers have adopted two common approaches. Studies in the happiness literature typically revert to conditional fixed effects methods, originally proposed by Chamberlain (1980). Although this provides consistent parameter estimates, even when the individual effect correlates with the regressors, the approach causes a range of practical inconveniences for the researcher. First, the approach incurs a great loss of information and a reduction in sample size. This is because the dependent variable needs to be collapsed into a binary format. Empirical studies usually use an arbitrary threshold value so that only those individuals can be retained in the sample who surpass this threshold. Second, marginal effects of a variable of interest cannot be easily calculated due to the inherent lack of information on the distribution of the individual heterogeneity (Wooldridge, 2002, p.492). A recent extension of Chamberlain's model, the conditional ordered fixed effects logit (COFEL), proposed by Ferrer-i-Carbonell and Frijters (2004) and applied by Frijters et al. (2004a,b, 2005b), suggests a method to reduce the drastic loss in the number of observations by identifying individual-specific threshold values to collapse the dependent variable into a binary format. However, its original formulation is highly computation intensive, a fact which makes its wider application less attractive. In a simplification of this estimator, one can simply use the individual means as cut-off criterion. This approach has the limitation that it still collapses the dependent variable into a binary format and that one can only assess the probability to report a health status greater than one's own mean. This focus may be of limited practical value if the researcher is interested in testing whether the impact of income on health (well-being) differs for high and low values of health (well-being).

In light of these problems many researchers give up the orthogonality assumption or ignore the presence of an individual unobserved effect to be able to apply the conventional models for categorically ordered data. Pooled ordered, generalized ordered (Terza, 1985) or random effects ordered models and their extensions based on the 'correlated effects' approach of Mundlak (1978) and Chamberlain (1980) are used predominantly in the literature on self-reported health. Generalized ordered probit models, which relax the single index assumption, are applied in the health literature to control for reporting bias in the subjective measures (Kerkhofs and Lindeboom, 1995; Groot, 2000; Sadana et al., 2000; Shmueli, 2002; Van Doorslaer and Jones, 2003; Hernandez-Quevedo et al., 2004;

Contoyannis et al., 2004). A further extension, the random effects ordered probit model with a dynamic specification has been applied by Jones et al. (2006) to assess health-related panel attrition and by Gannon (2005) to assess the effect of disability on labour market status. Boes and Winkelmann (2006) use generalized ordered probit models to identify the asymmetric impact of income on life satisfaction.

In this paper we use 22 waves of the German Socio-Economic Panel to answer four questions concerning the income-health nexus: First, we ask whether the trade-off between parameter bias, that occurs when using ordered or random effects models, and information loss, that occurs when using conditional fixed effects logit models, matters. Second, we ask whether the impact of income on health is asymmetric for transitions out of poor health and transitions into very good health. To this extent we go a step further than Ferrer-i-Carbonell and Frijters (2004) by proposing a simple solution to overcome the loss of information. In what we call the generalized conditional fixed effects logit (GCFEL), we estimate for each possible cut-off value for which the dependent variable can be dichotomized a simple conditional fixed effects logit. This allows us to investigate the effect of income on the upper and lower bounds of the distribution of reported health states while controlling for unobserved heterogeneity. Third, we test the hypothesis that the differences between marginal effects of income across models vary across age-cohorts. We do this to take into account the changing influence of unobserved heterogeneity over the life-cycle. In so far as unobserved factors become more important for the older groups we expect smaller differences in the relative performance of models for the younger age-groups. Finally, to check for robustness, we ask whether differences in marginal effects can be attributed to different sample sizes that result from conditional fixed and random effects models, from the particular choice of the proxy for health or from the computational methods.

We find that adjusting for unobserved heterogeneity accounts for the relationship between income and very good health, but not between income and poorer health states. The income gradient for older age-groups is more strongly affected by controlling for unobserved heterogeneity: revealing an increasing influence of heterogeneity on health satisfaction over the life-span.

The remainder of the paper is structured as follows: Section 2 introduces a general panel data specification for ordered categorical data and the methods to calculate marginal effects in the presence of individual heterogeneity. In section 3 we introduce the dataset, explain the construction of the measure of satisfaction with health, present the relevant descriptive statistics, and display the conditional distributions of the dependent variable. In section 4 we present our results. We focus on a graphical depiction of the marginal effects of income and their confidence intervals and a simulation of predicted probabilities. This section is complemented by a comprehensive sensitivity analysis of our results. Section 5 concludes and discusses the implications of the results.

2 Econometric Specification

2.1 The Model

We work with an ordered dependent variable that measures health status (HS) as reported by the interviewee[2]. HS^*_{it}, the true, but unobserved health status of an individual is given by:

$$HS^*_{it} = \alpha_i + \beta' X_{it} + u_{it}, \tag{1}$$

where i $= 1, \ldots, N$, t $= 1, \ldots, T_i$, and T_i reflects an unbalanced panel, α_i is an intercept term that varies for individuals and thus represents time-invariant unobserved heterogeneity, X_{it} is a vector of exogenous variables, and u_{it} is an idiosyncratic error term. We assume a standard logistic distribution for $u_{it} \sim \Lambda(0, \frac{\pi^2}{3})$.

We observe the reported health status $HS_{it} = j$ for j $\in \{1, \ldots, J\}$ if the true, underlying health status lies within an interval τ_{j-1} and τ_j:

$$HS_{it} = j \text{ if } \tau_{j-1} < HS^*_{it} \leq \tau_j, \tag{2}$$

[2]We proxy health status by satisfaction with health. From now on health status is synonymous with satisfaction with health.

We allow the individual threshold value τ_j to be not only a function of observable characteristics X_{it} according to Terza (1985), but also to vary across individuals i:

$$\tau_j = \tilde{\tau}_{ij} + \gamma_j' X_{it}, \tag{3}$$

where $\tilde{\tau}_{ij}$ is a scalar that varies across thresholds j and individuals i and γ_j' is a $1 \times k$ row vector of response specific parameters. Plugging equation (3) into (2) and replacing HS_{it}^* by the linear index of equation (1), then $\mathrm{HS}_{it} = j$ if

$$\tilde{\tau}_{ij-1} + \gamma_{j-1}' X_{it} < \alpha_i + \beta' X_{it} + u_{it} \leq \tilde{\tau}_{ij} + \gamma_j' X_{it}. \tag{4}$$

By rearrangement, we obtain:

$$\begin{aligned}
\tilde{\tau}_{ij-1} + \gamma_{j-1}' X_{it} - (\alpha_i + \beta' X_{it}) \;&<\; u_{it} \;\leq\; \tilde{\tau}_{ij} + \gamma_j' X_{it} - (\alpha_i + \beta' X_{it}) \\
-(\alpha_i - \tilde{\tau}_{ij-1}) - (\beta - \gamma_j)' X_{it} \;&<\; u_{it} \;\leq\; -(\alpha_i - \tilde{\tau}_{ij}) - (\beta - \gamma_{j-1})' X_{it} \qquad (5) \\
-\alpha_{ij-1} - \beta_j' X_{it} \;&<\; u_{it} \;\leq\; -\alpha_{ij} - \beta_j' X_{it}, \qquad (6)
\end{aligned}$$

where $\beta_j' = (\beta - \gamma_j)'$ and $\alpha_{ij}' = (\alpha_i - \tilde{\tau}_{ij})'$. This means that neither β and γ_j nor α_i, the individual specific health endowment and $\tilde{\tau}_{ij}$, the reporting bias, can be separately identified. Hence, the probability that an individual will report a specific health status $\mathrm{HS}_{it} = j$ is

$$P_{it,j} = P(HS_{it} = j | \alpha_{ij}, X_{it}) = F(-\alpha_{ij-1} - \beta_{j-1}' X_{it}) - F(-\alpha_{ij} - \beta_j' X_{it}), \tag{7}$$

where F $= \Lambda(.)$ is the logistic distribution function. The sample loglikelihood function to be maximized is:

$$\ln L(\beta) = \sum_{j=1}^{J} \sum_{i=1}^{N} \sum_{t=1}^{T_i} \ln \Big[\Lambda(-\alpha_{ij} - \beta_j' X_{it}) - \Lambda(-\alpha_{ij-1} - \beta_{j-1}' X_{it}) \Big]. \tag{8}$$

Equation (8) presents a generalization of Terza (1985) with individual effects $\tilde{\tau}_{ij}$ in the cut-points τ_i. The threshold values allow for individual specific reporting bias due to cut-point shift. Economically, this could mean that individuals in a higher category of perceived health experience a smaller marginal effect of income on health and that the individual unobserved heterogeneity not only differs across individuals, but depends also on the current health status which an individual reports.

In practice, the generalized model dichotomizes the dependent variable for J-1 threshold values that an individual can surpass and estimates the probability of observing $\mathrm{HS}_{it}^{B,j} = 1$. This can be expressed as:

$$HS_{it}^{B,j} = 1 \text{ if } HS_{it} \leq j \tag{9}$$

$$HS_{it}^{B,j} = 0 \text{ if } HS_{it} > j, \tag{10}$$

where j $\in \{1, \ldots, \mathrm{J}-1\}$ and B stands for "binary variable". Which regression model should be applied depends on the assumption one imposes on the α_{ij} and β_j,

In light of our theoretical discussion on the possible links between health, cognitive ability, and socio-economic status, we allow α_{ij} to be correlated with the regressors X_{it}. In this setting we can use the Chamberlain (1980) conditional fixed effects logit to estimate the β_js. To eliminate the individual fixed effect from the loglikelihood function, this approach takes advantage of a sufficient statistic. We have to find J-1 sufficient statistics η_j for α_{ij}, for which the distribution of the sample, given η_j, does not depend on α_{ij}:

$$f(HS_{it}^{B,j}|X_{it}, \alpha_{ij}, \eta_j) = f(HS_{it}^{B,j}|X_{it}, \eta_j). \tag{11}$$

In the case of the logistic regression, Andersen (1970, 1971) shows that $\sum_{t=1}^{T_i} HS_{it}^{B,j}$ is a sufficient statistic for α_{ij} and that conditional ML estimates are consistent. We use this result for J-1 binary equations. Conditioning on $\sum_{t=1}^{T_i} HS_{it}^{B,j} = \sum_{t=1}^{T_i} d_{it,j}$, i $= 1, \ldots, N_j$, and t $= 1, \ldots, T_i$, where $\mathrm{d}_{itj} = 1$ if $\mathrm{HS}_{it}^{B,j} = 1$ and 0 otherwise, the loglikelihood is in our case:

$$\ln L = \sum_{j=1}^{J-1} \ln L_j[I(HS_{i1} > j), \ldots, I(HS_{i,T_i} > j)| \sum_{t=1}^{T_i} I(HS_{it} > j) = c_j] =$$

$$\sum_{j=1}^{J-1} \sum_{i=1}^{N_j} \ln \frac{\exp((\sum_{t=1}^{T_i} HS_{it}^{B,j} X_{it}')\beta_j)}{\sum_{d \in B_{ij}} \exp((\sum_{t=1}^{T_i} (d_{itj} X_{it}')\beta_j)}, \tag{12}$$

where

$$B_{ij} = \{\mathbf{d}_j = (d_{i1j}, \ldots, d_{iT_ij})|d_{itj} \in \{0, 1\} \text{ and } \sum_{t=1}^{T_i} d_{itj} = \sum_{t=1}^{T_i} = HS_{it}^{B,j} = c_j\}. \tag{13}$$

is the set of all possible sequences of 0s and 1s for which the sum of T_i binary outcomes equals $\sum_{t=1}^{T_i} d_{itj} = c_j$. For those individuals for which $0 < \sum_{t=1}^{T_i} HS_{it}^{B,j} < T_i$ does not hold true do not contribute to the log-likelihood and, therefore, will be dropped from the sample. Hence, sample sizes N_j across the J$-$1 categories will differ, i.e. $\mathrm{N}_{j=1} \neq \ldots \neq \mathrm{N}_{j=J-1}$. In total, there will be $T_i - 1$ alternative sets B_{ij}. Thus, in the denominator of the conditional likelihood function we find the sum of probabilities of each possible sequence of 0s and 1s that is not equal to 0 or T_i.

This formulation is a generalization of Chamberlain (1980), Das and Van Soest (1999), and Ferrer-i-Carbonell and Frijters (2004). All three allow only for dichotomization of one pairing of adjacent categories, and, thus implicitly assume that $\beta_j = \beta$ and $\alpha_{ij} = \alpha_i$.

Application of the Chamberlain (1980) approach requires an arbitrary threshold to dichotomize the variable, e.g. $\sum_{t=1}^{T_i} I(HS_{it} > 3)$. Choosing this arbitrary threshold value incurs a loss of those individuals who never cross, at least once, the threshold j= 3, i.e. for whom $0 < \sum_{t=1}^{T_i} HS_{it}^{B,j=3} < T_i$ does not hold true. Das and Van Soest (1999) combine adjacent categories so that the dependent variable is summarized as a binary variable, and then use Chamberlain's method. They repeat this for all the possible combinations of adjacent categories to get J$-$1 estimates of the parameters of interest. Das and Van Soest (1999) then combine these estimates into one final estimate β_{FIX} which is a linear combination of J$-$1 β_js. The optimal weighting matrix is obtained from a minimum distance approach. Ferrer-i-Carbonell and Frijters (2004) propose an estimator that still collapses the ordered variable into a binary format, but that uses an individual specific threshold value j_i. To find this individual threshold, the authors maximize a weighted sum of J−1 log-likelihood functions, similar to Das and Van Soest (1999), subject to the constraint that the sum of squared weights across all possible threshold values across all individuals must be equal to the number of individuals in the sample. This constraint means that we can use only weights $w_{ij} \in \{0,1\}$. This is tantamount to saying that only one out of J−1 log-likelihood functions for each individual actually contributes to the total log-likelihood, all the others will drop out. The major question of this set-up is to which of the J−1 log-likelihood functions should one assign the weight. Ferrer-i-Carbonell and Frijters (2004) suggest it should be the one for which the analytical expected Hessian is minimized. In practice, the estimator of interest can be derived within a four step procedure. In a first step the ordered dependent variable is dichotomized into $\mathrm{HS}_{it}^{B,j}$ for each possible threshold value j. In a second step Chamberlain's estimator is used choosing one arbitrary threshold value that equally applies to all individuals, let's say j= 3. The predetermined parameter vector $\hat{\beta}_{j=3}$ along with the realizations of $\mathrm{HS}_{it}^{B,j}$ and X_{itj} for all j $\in 1,2,3,4$ is used to calculate analytically for each individual, which surpasses the corresponding threshold j, the expected Hessian. In a third step, the log-likelihood function L_{ij} for which the analytical expected Hessian is minimized receives the weight $\mathrm{w}_{ij} = 1$ and the corresponding threshold value j_i is earmarked. In a last and fourth step the parameter vector of interest is estimated by dichotomizing the ordered variable into HS_{it}^{B,j_i} using the individual-specific threshold j_i and applying the Chamberlain method. As an approximation one can simply use the within-individual mean values of the health status score as the threshold. This would mean that one constructs for each individual $j_i = \bar{HS}_i = \frac{1}{T_i}\sum_{t=1}^{T_i} HS_{it}$ and collapse the data into binary format:

$$\begin{aligned} HS_{it}^{B,j_i=\bar{HS}_i} &= 1 \text{ if } HS_{it} > \bar{HS}_i \\ HS_{it}^{B,j_i=\bar{HS}_i} &= 0 \text{ if } HS_{it} \leq \bar{HS}_i \end{aligned}$$

Using the mean values as cut-off values one loses only those individuals who never change their health status over time[3].

If one still allows for threshold-specific unobserved heterogeneity α_{ij}, but assumes it to be independent of X_{it}, and category specific parameter vectors β_j, we are in the world of a generalized random effects logit (GREL). The individual effect α_{ij} is then integrated out of the loglikelihood by:

$$\ln L(\beta_j) = \prod_i \Big\{ \int_{-\infty}^{\infty} \prod_{t=1}^{T_i} \ln[1-\Lambda(\beta_j^{*\prime} X_{it}+\sqrt{\frac{\rho}{1-\rho}})\alpha_{ij}^*]^{1-HS_{it}^{B,j}} [\Lambda(\beta_j^{\prime} X_{it}+\sqrt{\frac{\rho}{1-\rho}}\alpha_{ij}^*)]^{HS_{it}^{B,j}} \phi(\alpha_{ij}^*) d\alpha_{ij}^* \Big\}, \tag{14}$$

where $\beta_j^* = \frac{\beta_j}{\sigma_u}$ and $\alpha_{ij}^* = \frac{\alpha_{ij}}{\sigma_u}$. If we assume $\alpha_{ij} = \alpha_j$, we are in the world of a generalized ordered logit (GOL). Finally, if we additionally, assume $\beta_j = \beta$ we are in the world of pooled ordered logit (POL).

2.2 Marginal Effects

We calculate the marginal effect of the logarithm of net monthly household equivalent income for each age-specific sub-group, evaluated at the most representative value of the independent variables for each subgroup. For instance, we calculate the marginal effect for a married individual in the age-group 41 to 50, who has two children, who live in the same household, who obtained a ten-year school education and underwent an apprenticeship, and who has the average disposable income of the respective age-group.

The marginal effects for the (ordinal) pooled ordered and generalized ordered logit are programmed according to Cameron and Trivedi (2005, p. 520) and Boes and Winkelmann (2006, p. 12), respectively. For calculating the marginal effects for the (binary) conditional fixed and random effects we need to make an additional assumption on the distribution of the unobserved effect α_{ij}. One common solution is to set its estimate $\hat{\alpha}_{ij} = 0$. In this case the marginal effect is calculated by fitting a pooled logit density function multiplied

[3]Many thanks to Paul Frijters for valuable comments, the GAUSS syntax and for making us aware of this simplification that can be implemented easily in STATA. We estimated our models by both the full computation method in GAUSS and its simplification in STATA. Since results are very similar, we conducted the entire analysis with the simplified approach.

by the coefficient of interest:

$$ME_j = \hat{\beta}_j \Lambda(X_{it}\hat{\beta}_j + \hat{\alpha}_{ij})(1 - \Lambda(X_{it}\hat{\beta}_j + \hat{\alpha}_{ij})) = \hat{\beta}_j \Lambda(X_{it}\hat{\beta}_j)(1 - \Lambda(X_{it}\hat{\beta}_j)), \tag{15}$$

where j stands for the particular threshold value.

Another solution is to obtain $\hat{\alpha}_{ij}$ from the generalized conditional fixed effects logit. Analytically, the individual fixed effect is the mean difference between the latent health status and the prediction of the dependent variable, evaluated at the mean value of all regressors. We approximate this by:

$$\hat{\alpha}_{ij} \simeq \Lambda^{-1}(\bar{HS}_i^{B,j}) - \bar{X}_i \hat{\beta}_j, \tag{16}$$

with $\bar{HS}_i^{B,j}$ being the sample average of the observed health status variable that resulted from dichotomizing the originally ordinal variable at threshold j, with $\bar{X}_i$ being the sample average of each exogenous variable in the model, and with Λ^{-1} being the inverse of the logistic function. In light of the hypothesis, that there should be an increasing variability in the individual fixed effect over the life-span of an individual, we calculate the average of the individual fixed effect for each age-group k:

$$\bar{\hat{\alpha}}_k = \frac{1}{N_k} \sum_{i=1}^{N} \hat{\alpha}_{ij} D_k, \tag{17}$$

where D_k is an indicator variable that takes the value 1 if the individual belongs to the particular age-group k, and 0 otherwise. The total number of individuals in each age-group is indicated by N_k[4].

[4] A third possibility is to calculate 'pseudo-marginal effects' proposed and implemented by Ferrer-i-Carbonell and Frijters (2004); Frijters et al. (2004a,b, 2005b). The pseudo marginal effect makes the marginal effect of a variable of interest proportional to the effect of changes in observed health on latent health. It represents the effect of an increase of 1 in a variable with coefficient β as an effect of $\hat{\mu}\beta$ on latent health. μ estimates the response of observed health satisfaction levels to estimated changes in latent health satisfaction, i.e.: $\hat{\mu} = \frac{{}_{t}\ {}_{S_t}(HS_{it+1} - HS_{it})}{{}_{t}\ {}_{S_t}(\bar{z}_{t+1} - \bar{z}_t)\hat{\beta}_z}$, where $\bar{z}_t$ is the average of any variable of interest and S_t is the number of individuals observed in all time periods. This marginal effect relies on the crucial assumption that a change in latent health is linear to a change in observed health.

Last, to make generalized random effects logit (GREL) marginal effects comparable to those of pooled ordered logit, we have to multiply the GREL coefficients β_j^* with $(1-\hat{\rho})$ (Arulampalan, 1999), where $\hat{\rho} = \frac{\sigma^2_{\alpha_j}}{1+\sigma^2_{\alpha_j}}$.

3 Data

We use the full set of 22 available waves of the German Socio-Economic Panel (GSOEP) running from 1984 to 2005. The data was set up with the help of Panel Whiz, a data retrieval programme written by John Haisken-DeNew[5]. The GSOEP provides three subjective measures to proxy health status, namely "satisfaction with health" (SWH), "self-assessed health" (SAH), and "worry about health" (WAH). The measure SWH ranges from 0 for the lowest to 10 for the highest level of health satisfaction. The second measure, the one most commonly used in the literature on health determination, ranges from bad to very good health on a five point scale. The last, WAH, is a three point measure ranging from a lot of worry to no worries about one's own health status. However, SAH is available only for the years 1994 to 2005, and WAH only for the years 1999 to 2005. To take advantage of the largest possible time series for each individual, we follow Frijters et al. (2005b) in interpreting satisfaction with health as an adequate proxy for SAH. However, to make the measure more comparable to the SAH measure that has been widely used in the literature, and to make our results more manageable, we collapse this measure into a five point scale. For mapping SWH into SAH we use the cumulative distribution functions of the two measures, the cross tabulations of each category and the correlation structure between each sub-category of the two measures. Comparing the cumulative distribution functions[6] for the two variables allows crudely to compare the sum of individuals who report values of SWH up to let's say 2 with the sum of individuals who report a sub-category of SAH, let's say 1.

Looking at the cross-tabulations allows to compare for each sub-category the percent-

[5]Panel Whiz can be accessed via www.panelwhiz.eu. We use plug-ins written by John Haisken-DeNew, Markus Hahn, Mathias Sinning, and Ingo Geishecker.

[6]The graph of the cumulative distribution function can be provided upon request by the authors.

age of individuals who report, let's say SWH$=0$ who also report SAH$=1$. One can calculate the percentage of overlap for each possible sub-category combination of SWH and SAH. In total, there are 55 pairwise comparisons. These percentages of overlap can be expressed from the perspective of an individual who reports a sub-category of SWH or SAH. A value of SWH is then mapped into a value of SAH for those pairwise comparisons for which the percentage is the greatest.

Finally, one can look at the sign and the magnitude of the correlations between the 55 possible pairs of comparisons. The strongest positive correlation between two sub-categories of the two variables is then used as a guideline for mapping. For example, if the correlation between SWH$=2$ and SAH$=1$ is stronger (and positive) than any other correlation between SWH$=2$ and SAH>1, then SWH will be recoded to 1. Details of these latter two methods can be found in the Appendix: Coding of Health Satisfaction.

Mapping the lower and the higher values of health satisfaction (value of SWH from 0 to 6 and from 8 to 10) is unambiguous across the three different methods of mapping. Ambiguity arises due to the question into which category of SAH SWH$=7$ should be merged. We choose the following recoding:

- SWH$=0,1,2$ into SAH$=1$ (bad health)
- SWH$=3,4$ into SAH$=2$ (poor health)
- SWH$=5,6$ into SAH$=3$ (satisfactory health)
- SWH$=7,8$ into SAH$=4$ (good health)
- SWH$=9,10$ into SAH$=5$ (very good health)

Thus, we proxy health status (HS) by satisfaction with health collapsed into a five point scale.

As independent variables we include six age categories: from 16 to 30, 31 to 40, 41 to 50, 51 to 60, 61 to 70 and 71 and older. We include these age-groups to test the hypothesis that the influence of time-invariant unobserved heterogeneity should become more influential for older individuals in determining health. Our main variable of interest, socio-economic status, is proxied by household equivalent income[7]. We construct this variable from net monthly household income and adjust for the household size by dividing this variable by the square root of the number of household members[8]. We interact the income variable with the age-group dummies to obtain a vector of six different income coefficients. We also include a set of four dummy variables representing educational and professional training. "Secondary schooling" includes those individuals who have less or completed nine years of education and who have no professional education. This category also includes those who have any "other degree", an educational status which refers usually to immigrants. "Intermediate degree" includes those individuals who have at least nine years of schooling and an apprenticeship. "Upper degree" includes those who have thirteen years of schooling and chose an apprenticeship as professional education or those who have ten years of schooling and an advanced professional technical school degree. "University degree" includes all those who obtained a degree from university or a polytechnic.

In addition, we control for immigrant and marital status, the number of household members, geographical location (East versus West Germany) and time effects. Household equivalent income is log-transformed. Last, we separate our sample into men and women to account for the gender specific difference in the relationship between socio-economic status and health. A full description of the variables, their descriptive statistics, and sample sizes can be found in the Data Appendix.

Below, we present the distribution of SWH split by various demographic, socio-economic, and household characteristics. In addition, we display the evolution of health

[7]In the happiness literature we find a large emphasis on the importance of relative income in determining life satisfaction. However, we choose a measure of absolute income to test its effect on health satisfaction. Recent evidence has shown that relative income has no effect on self-reported health (Jones and Wildman, 2007; Gravelle and Sutton, 2003, 2006; Lindley and Lorgelly, 2003; Miller and Paxson, 2006). Some of the arguments in favour of no relative income effect are that relative income may have a delayed effect only after 15 years (Subramanian and Kawachi, 2004) or due to omitted variable bias caused by not including area or year effects (Gravelle and Sutton, 2006).

[8]This is a simplification recently applied in OECD studies to adjust household income for needs. It is an approximation of the OECD modified equivalence scale (Hagenaars et al., 1994), which assigns to the household head a value of 1, each additional member of the household a value of 0.5, and each child below the age of 16 a value of 0.3. We tested both methods. Since the results do not differ, we opted for the simplified version.

status over time. All graphs are constructed separately for men (left column) and women (right column). Figures 1(a) and 1(b) show the evolution of the five categories of SWH over time, reported for every second year from 1984 to 2004. It is clear that over the 21 years reported here the number of individuals reporting very good health, SWH= 5, declines significantly, but in a non-linear manner for both women and men. The number of individuals, who report bad health, however, remains relatively stable over the years, showing only a slight tendency to decline.

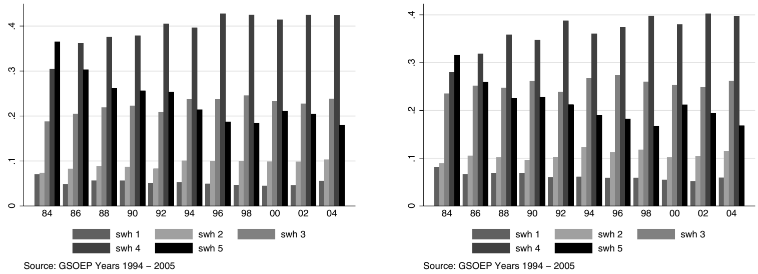
Figures 2(a) and 2(b) depict the differences in reporting each category of SWH over the age-groups for women and men. These figures illustrate the decline (increase) in number of individuals reporting very good health (bad health) the older the age-group they belong to.

Figures 3(a) and 3(b) depict the change in reporting various health states over the range of income quintiles (increasing from left to right). For both women and men we notice an interesting phenomenon. Whereas, as expected, the number of individuals reporting good health increases the higher the income quintile, we do not find the same trend for very good health. The number of individuals reporting very good health initially drops for the first three income quintiles, and then slightly increases for the two highest income quintiles.

A similar phenomenon occurs when looking at education groups as depicted in Figures 4(a) and 4(b). The higher the educational background of men and women (from left to right), the greater the number of individuals who report good health. The same does not hold true for the highest possible health status when looking at men. Nevertheless, the number of individuals reporting bad health declines the higher the income or the educational background for women and men.

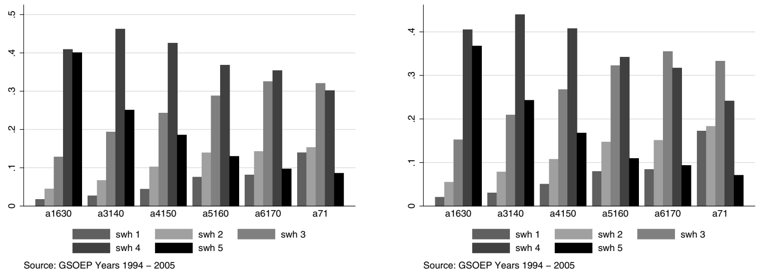
4 Results

The next two sub-sections show the estimation results for men and women for various models. These models comprise, on the one hand, the pooled ordered logit (POL) and gen-



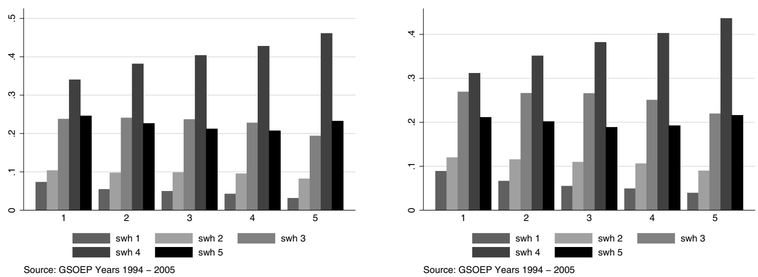
(a) Men (b) Women

Figure 1: SWH by years



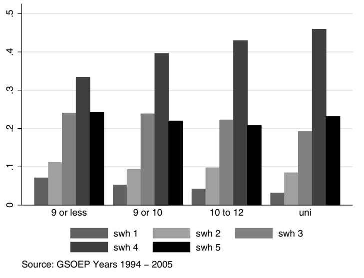
(a) Men (b) Women

Figure 2: SWH by age-groups

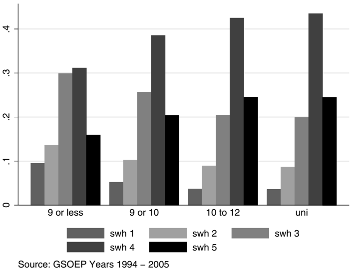


(a) Men (b) Women

Figure 3: SWH by income quintiles



(a) Men



(b) Women

Figure 4: SWH by education level

eralized ordered logit (GOL), which do not model time-invariant, individual- and health status-specific unobserved heterogeneity explicitly and on the other hand, the generalized random effects logit (GREL) and the generalized conditional fixed effects logit (GCFEL). The latter both assume the presence of health-status specific time-invariant heterogeneity, but differ in their assumption on the relationship between this heterogeneity and the regressors of the model. Looking at very good self-reported health, we dichotomize the ordered categorical variable into a binary format using the threshold value j= 4 for the GOL, GREL, and GCFEL. We contrast theses results with those obtained from the conditional ordered fixed effects logit (COFEL) that uses the individual means as cut-off values. Last, we simulate the predicted probabilities of reporting very good health over a range of household incomes. Finally, we report also the changes in the probabilities for other health states, namely the probability to exit bad or bad and poor health, or to enter good and very good health for one age-group. Ultimately, we test the robustness of our results with respect to the calculation of the marginal effects, the proxy for health status chosen, and the sample definition.

We report only the marginal effects and their confidence intervals for our variable of interest: the logarithm of net monthly household equivalent income for each age-specific sub-group and we present these graphically as box-plots. Age-groups are identified by ‘Age 1630’, for those aged between 16 and 30, up to ‘Age 71’, for those who are 71 years and older. We test the hypothesis of an increasing correlation between the individual fixed effect with socio-economic status due to an increase in magnitude of the individual fixed effect over time. Therefore, we expect the parameter bias to be greater for older age-groups; the difference in the parameter coefficients between the POL/GOL to the GREL/GCFEL should be no smaller for the older age-groups than for the younger ones.

4.1 Box-Plots of Marginal Effects

In Figure 5 we show the box-plots of the marginal effects of log-income on the probability to report very good health and their confidence interval for each age-specific sub-group (5(a) to 5(f)). The dot in the figure represents the magnitude of the marginal effect, while the vertical, capped lines represent the upper and the lower bound of a 95 percent confidence interval. The marginal effects for men (women) are indicated by a M (W) on the horizontal axis. From left to right we see in each chart the box-plot resulting from the pooled ordered logit (POL), the generalized ordered logit (GOL), the generalized random effects logit (GREL) model, and the generalized conditional fixed effects logit (GCFEL) model. Separated from these models by a vertical line, we report the marginal effects and its confidence interval for the conditional ordered fixed effects logit (COFEL) on the far right. This marginal effect is not directly comparable with the other marginal effects, because it represents the marginal effect of income on the probability of reporting a health status higher than one's own individual mean. This probability is labelled on the righthand vertical axis. We report this marginal effect to illustrate the differences in conclusions that result from a choice of a threshold value to dichotomize the dependent variable that is driven by statistical rather than by economic considerations.

Consider the marginal effects of income on the probability to report very good health (SWH= 5) for men in Figures 5(a) to 5(f). The POL yields the greatest marginal effect for all age-groups. The effect is largest for age-group 51 to 60 (Figure 5(d)): all things being equal, a doubling of log income per month will increase the probability to report very good health by about six-fold[9]. The effect is the smallest for the youngest (0.015, Figure 5(a)) and the oldest age-group (< 0.03, Figure 5(f)).

The more we adjust for unobserved heterogeneity and the fewer restrictions we impose on the model, the smaller the effect of a marginal increase of income on the probability to report very good health. The generalized ordered logit (GOL) model, which relaxes the single index assumption, yields a marginal effect of about 1 to 2 percentage points less than the POL model. The GREL model, which assumes the individual unobserved effect to be independent from income, suggests that a doubling of log income doubles and even triples the probability to report excellent health for the three middle-aged groups (41 to 50, 51 to 60, and 61 to 70, see Figures 5(c), 5(d), 5(e)). For the oldest and the second youngest age-group (71 and older, and 31 to 40, Figures 5(f) and 5(b)), this effect is less than 1 percentage point for an additional 1 percent increase in log income. For the youngest age-group the effect is even negative (Figure 5(a)). When relaxing all assumptions imposed on the relationship between unobserved heterogeneity and income in the generalized conditional fixed effects logit (GCFEL) model, the marginal effect is close to zero. It is negative for the youngest age-group (-0.005)[10].

This last result highlights one important finding in favour of our generalized approach: once the individual fixed effect is controlled for, the influence of observable socio-economic status on very good self-reported health disappears. This decrease is the most extreme for the middle-aged and older groups (41 to 50, 51 to 60, and 61 to 70), but not for the very old ones (71 and older). For these groups the difference between the marginal effects of income between the POL and GCFEL model varies between 4 to 6 percentage points. The effect is the least extreme for the youngest age-group (16 to 30). This result strengthens our hypothesis that the influence of the individual fixed effect plays a greater role in mediating the relationship between income and health status in older age.

Last, it has to be mentioned that using the conditional ordered fixed effects (COFEL) model may yield misleading results when one is interested in the upper bounds of the

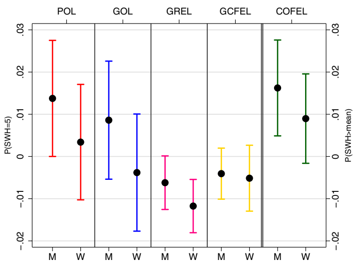
[9] A one percent increase in log-income increases the probability to report very good health by 0.06.

[10] Confidence intervals are not missing. In some cases they are so small that they do not show on the graph

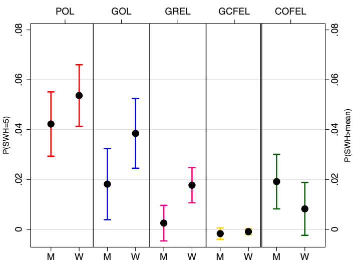
health distribution. The marginal effect of the COFEL model measures exclusively the impact of a change in income on any change in health status beyond one's own mean. For women a very similar picture emerges as documented in the same in the same figure under W (W=Women), except for the fact that the marginal effect across all possible models is by 1 to 2 percent points smaller than for men.

In Figure 6 we compare the marginal effects for age-group 51 to 60 for the remaining possible changes in health states (thresholds j= 1, j= 2, and j= 3). We choose to illustrate the changes for the middle age-group, because for this group the marginal impact of an increase in household income on the probability to report very good health has turned out to be the strongest. The box-plots in the left (right) column of Figure 6 represent the marginal effects for men (women). The models in Figure 6(a) can be interpreted as the effect of a percentage change in income on the probability to exit a bad health status (P(SWH> 1)). Figure 6(b) depicts the effect of an increase on income on the probability to exit bad and poor health states (P(SWH> 2)). Last, Figure 6(c) shows the impact of an increase in income on the probability to enter a good or very good health state (P(SWH> 3)). Similarly, as for the models investigating the probability to report very good health, the generalized ordered logit model, an extension of the POL, yields the highest impact of income on any change in health status for both men and women alike. For instance, a doubling of log monthly income for men (women), 15-folds the probability to report good or very good health (13-fold). In contrast, the effects of an increase in income on the probability of exiting bad (and poor) health states are much weaker. A doubling of log monthly income increases the probability to exit the lowest possible health status by four times only (eight times for the two lowest health states). This suggests, in general, that income has a parabolic effect on health: extra income supports individuals to reach higher reported levels of health more effectively than to exit lower reported health states, but levels off again for the highest health status.

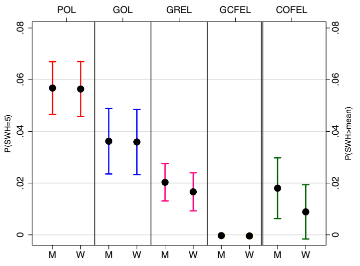
Interestingly, once controlling for unobserved heterogeneity with the generalized conditional fixed effects (GCFEL) model, the impact of income on health is relatively high for exiting the lower health states (0.07 for both men and women), but equal to zero for



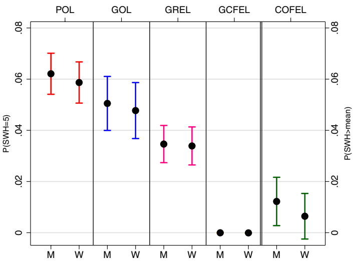
(a) Age-group 16 to 30



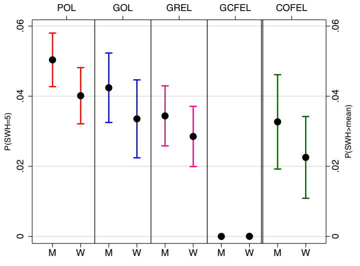
(b) Age-group 31 to 40



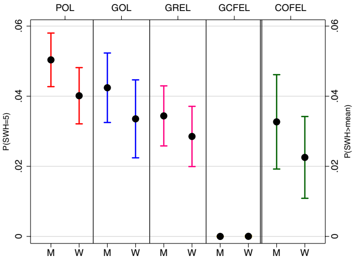
(c) Age-group 41 to 50



(d) Age-group 51 to 60

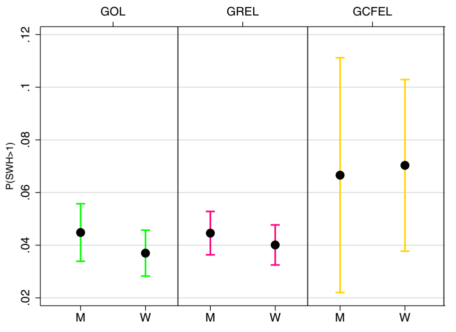


(e) Age-group 61 to 70

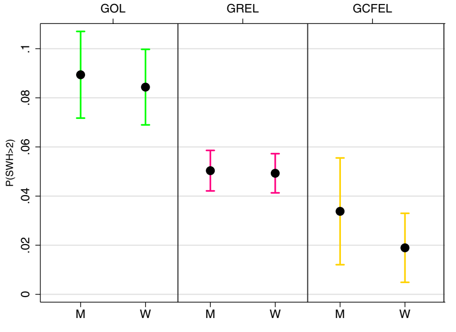


(f) Age-group 71

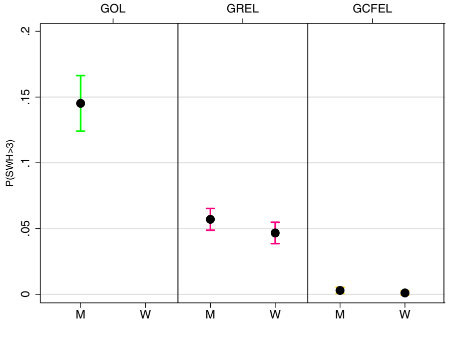
Figure 5: Marginal effects of income for men and women ($\alpha_{i,j=4} = 0$ for all age-groups)



(a) Prob to exit bad health (men and women)



(b) Prob to exit bad or poor health (men and women)



(c) Prob to enter good or very good health (men and women)

Figure 6: Marginal effects of income for men and women in age-group 51 to 60

the probability to report good or very good health. The marginal impact of a 100 percent increase of log income 4 to 7-fold the probability to exit lower health states. This reveals an important point: the individual fixed effect seems to drive the impact of income on entering higher health states, whereas it doesn't for exiting lower health states.

For this middle-aged group the confidence intervals of the marginal effects are quite large for changes in the lowest health states in the GCFEL. According to Table 1, we have only a fifth of the original sample of middle-aged individuals that report at least once SWH= 1 (1259 out of originally 5428 individuals). These smaller sample sizes could be an explanation for greater imprecision, i.e. in terms of the size of confidence intervals in the marginal effects. This finding illustrates that the generalized version of the conditional fixed can only be used if the original sample size is large enough.

4.2 Predicted Probabilities for Reporting Very Good Health

In Figure 7 we present the predicted probabilities to report very good health for the pooled ordered logit (POL), generalized ordered logit (GOL), generalized random effects logit (GREL), and the generalized conditional fixed effects logit (GCFEL) model for both men (left column) and women (right column). We graph the change in reporting probabilities over a range of potential net monthly household incomes from 500 to $5,500$ Euro for men and women separately[11].

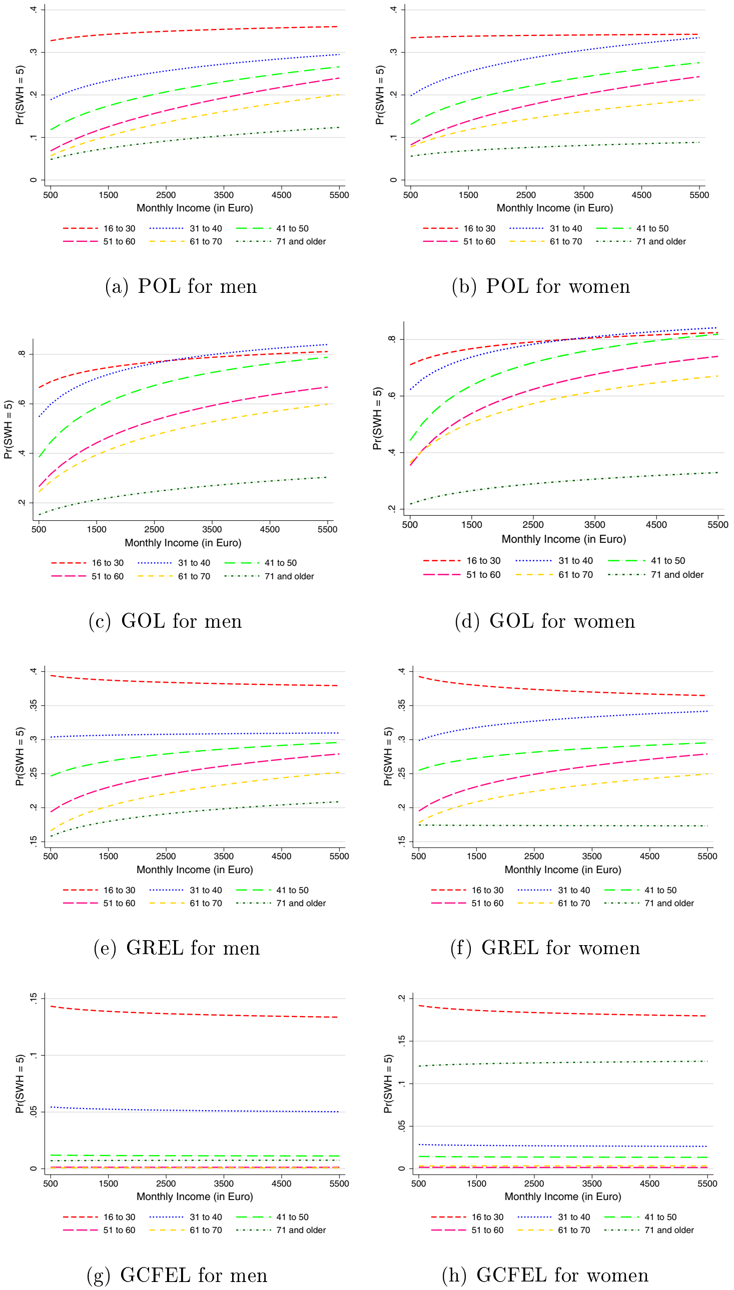
The predicted probabilities for both men and women make it clear that the overall effect of income on health is relatively low. At maximum, giving an extra $5,000$ Euro per month to a poor person raises the probability to report very good health by 0.3 in the POL (Figures 7(a) and 7(b)), and by 0.4 in the GOL (Figures 7(c) and 7(d)). This is the case for those aged between 51 and 60, who show the steepest income gradient. The four graphs show that after controlling for individual heterogeneity (Figures 7(g) for men and Figure 7(h) for women) the effect of income on the probability to report very good health

[11] To be able to graph the various conditional fixed effects models, we approximate the coefficients for time-invariant variables, such as the coefficient for age-group and education level, with the coefficients obtained from the generalized ordered logit.

Table 1: Number of individuals for men and women resulting from different models

Thresholds	Models	Age 16 to 30	Age 31 to 40	Age 41 to 50	Age 51 to 60	Age 61 to 70	Age 71 and older
MEN							
SWH > 1	GOL	7508	6724	6532	5428	3952	2143
	GREL	7508	6724	6532	5428	3952	2143
	GCFEL	640	928	1153	1259	1045	646
SWH > 2	GOL	7343	6512	6242	5163	3808	2071
	GREL	7343	6512	6242	5163	3808	2071
	GCFEL	1738	2291	2658	2533	1982	1065
SWH > 3	GOL	7343	6512	6242	5163	3808	2071
	GREL	7343	6512	6242	5163	3808	2071
	GCFEL	3400	3917	3972	3347	2483	1242
SWH > 4	GOL	7343	6512	6242	5163	3808	2071
	GREL	7343	6512	6242	5163	3808	2071
	GCFEL	4191	3839	3155	2176	1383	609
	POL	7508	6724	6532	5428	3952	2143
	GOL	7508	6724	6532	5428	3952	2143
SWH > mean	COFEL	5321	5385	5207	4344	3252	1676
SAH > 4	GOL	4910	4877	4531	3936	3194	1664
	GREL	4910	4877	4531	3936	3194	1664
	GCFEL	2788	3075	2634	2211	1717	834
WOMEN							
SWH > 1	GOL	7416	6735	6317	5048	3992	2942
	GREL	7416	6735	6317	5048	3992	2942
	GCFEL	745	1089	1319	1370	1263	1089
SWH > 2	GOL	7416	6735	6317	5048	3992	2942
	GREL	7416	6735	6317	5048	3992	2942
	GCFEL	2043	2640	2900	2604	2227	1637
SWH > 3	GOL	7416	6735	6317	5048	3992	2942
	GREL	7416	6735	6317	5048	3992	2942
	GCFEL	3748	4257	4138	3342	2621	1669
SWH > 4	GOL	7416	6735	6317	5048	3992	2942
	GREL	7416	6735	6317	5048	3992	2942
	GCFEL	4216	4042	3132	2005	1326	777
	POL	7416	6735	6317	5048	3992	2942
	GOL	7416	6735	6317	5048	3992	2942
SWH > mean	COFEL	5460	5672	5370	4315	3472	2369
SAH > 4	POL	4995	5042	4747	3874	3195	2339
	GREL	4910	4877	4531	3936	3194	1664
	GCFEL	2746	3193	2747	2202	1718	1219

The models are abbreviated as follows: generalized ordered logit (GOL), pooled ordered logit (POL), generalised conditional fixed effects logit (GCFEL), generalised random effects logit (GREL), and conditional ordered fixed effects logit (COFEL).



(a) POL for men

(b) POL for women

(c) GOL for men

(d) GOL for women

(e) GREL for men

(f) GREL for women

(g) GCFEL for men

(h) GCFEL for women

Figure 7: Probability of reporting very good health over income for men and women

fades significantly. The effect disappears completely for the two youngest age-groups in both models (Age 16 to 30 and Age 31 to 40), and similarly for the age-group of 41 to 50 year-old.

However, all models coincide in their ranking of overall probabilities in reporting very good health for the six age-groups. In all models the youngest age-group has the highest probability, whereas the oldest age-group has the lowest probability. Only for women the GCFEL results in a slightly different ranking, in which the oldest age-group faces the second highest probability to report very hood health.

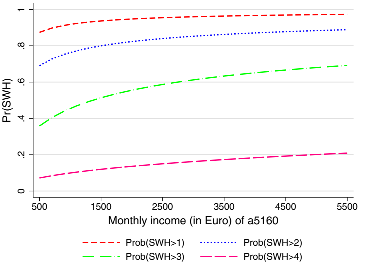
After sweeping out the individual fixed effect, the probabilities to report very good health are significantly lower for all sub-groups. For the older subgroups for both men and women this probability approaches zero. This observation is in line with our hypothesis that the older age-groups should be affected disproportionally from controlling for unobserved heterogeneity.

Last, Figure 8 displays the predicted probabilities of reporting all four health states for age-group 51 to 60 that result from the GOL (Figures 8(a), 8(b)), the GREL (Figures 8(c), 8(d)), and the GCFEL (Figures 8(e), 8(f)). Most interestingly these various probabilities reinforce our findings from Figure 6(a) to 6(c). The individual fixed effect drives the association between health and income for changes into good and very good health, but does not drive it to exit bad or poor health. We find the same phenomenon for all age-groups, except for the youngest one[12].

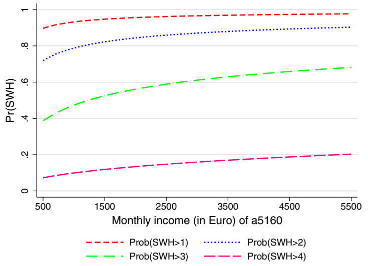
4.3 Sensitivity Analysis

In this section we test whether our results are robust to changes in the dependent variable, in the calculation of the marginal effects or to the definition of the sample size. In particular, we ask whether we can generalize our results for another proxy of health status, namely self-assessed health (SAH). Then we investigate whether the low magnitude of the marginal effect of income on health in the GCFEL model is caused by setting the

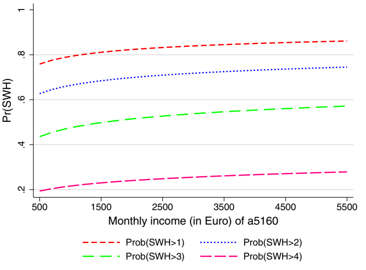
[12]The graphs showing the predicted probabilities of reporting all levels of health for all age-groups will be provided upon request.



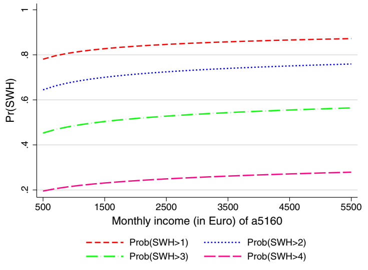
(a) GOL for men



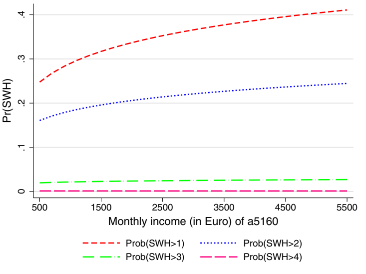
(b) GOL for women



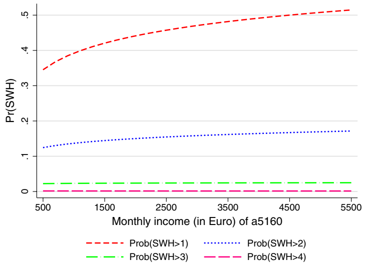
(c) GREL for men



(d) GREL for women



(e) GCFEL for men



(f) GCFEL for women

Figure 8: Predicted probabilities of four possible health states for age-group 51 to 60 for men and women

individual fixed effect equal to zero. Last, we test whether the different results for the GCFEL model are driven by the smaller sample size required by the conditioning method rather than by controlling for the individual fixed effect.

In Figure 10 and 11 we display all marginal effects resulting from these three different types of sensitivity analysis for the three youngest and the three oldest age-groups. For the pooled and the generalized ordered logit (POL and GOL, respectively) the graphs show the marginal effects for the full sample (FS)[13], the smaller sub-sample which uses only observations that are used in the GCFEL (SS), and the regressions that apply self-assessed health as dependent variable (SAH). In addition, we report for the generalized random effects logit (GREL) and the generalized conditional fixed effects logit (GCFEL) the alternative marginal effects, which uses an approximation of the estimate of the individual fixed effect (ALT).

Since we have already shown that results for men and women are very similar we conduct the sensitivity analysis only for men.

4.3.1 Alternative Measure of Health Status

One may ask whether our polarizing results between the GOL and POL and the GREL and GCFEL may be a particular feature of the satisfaction with health proxy for heath status rather than a general trait of health determination models. One way to test this assertion it to replicate our analysis using self-assessed health (SAH) as dependent variable. As shown in Table 3 and 4 in the Data Appendix, the average value of this variable in the male and female sample is one unit smaller than SWH and it is also only available from 1994 onwards. The latter fact causes the sample shrink to one third from the original sample (see Table 1, compare row SAH>4 with SWH>4). The loss of observations is the greater the younger the age-group.

Looking at Figures 10 and 11 we see the difference between the original marginal effects (FS) and the estimated marginal effects for changes in the probability to report very good health when using SAH (SAH) turn out to be mixed. In the GREL this latter marginal

[13]These are the original marginal effects reported in Figure 5.

effect is either the same or slightly bigger than the one under the original model. For the three younger age-groups in the POL and GOL the former holds true (see Figure 10), but for the older age-groups the marginal effects are smaller by greater proportions than in the original model (see Figure 11). In sharp contrast to these results, again, stands the marginal effects resulting from the GCFEL. In this model the marginal effects are much smaller than the original ones by a magnitude ranging between $-.13$ and $.30$ percentage points for the four younger age-groups.

One explanation for this surprising finding could be that the sample sizes for each sub-group are particularly small. With respect to the four youngest age-groups, we lose between 500 to 1000 individuals (Table 1 for age-group 1530 to 5160).

In light of these insights, we conclude that we can generalize our original results with respect to the health proxy chosen for the younger age-groups in the pooled, generalized and random effects models and for the older age-groups in the generalized conditional fixed effects model.

4.3.2 Alternative Marginal Effect Calculation

We may be criticized for making the arbitrary assumption of $\alpha_{ij} = 0$ when calculating the marginal effects for each age-group and that this assumption drives the low magnitude of the effects in the GREL and the GCFEL. One method to test this hypothesis is to calculate an approximation of the individual, threshold-specific fixed effect (from here onwards referred to as IFE) and to plug the mean value of it for each age-group into the formula of the marginal effect. We use the approximation as proposed in equation (16).

Understanding the dimensions of the approximate distribution of the IFE has also the advantage to test its dependence on the covariates of the health determination model. We can test directly, therefore, the hypothesis of no association, which is a necessary condition to apply pooled (ordered) or random effects logit models.

Figure 9 displays the distribution of the IFE conditional on the four possible health states. FE1 stands for the IFE calculated from a model in which individuals surpass, at least once, the threshold j= 1, FE2 stands for the IFE calculated from the model in

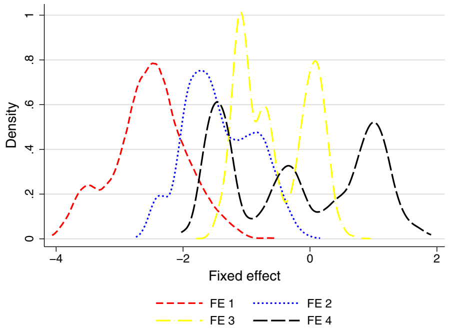


Figure 9: Distributions of individual fixed effects across thresholds j= 1 to 4

which individuals surpass j= 2, and so on. Depending on the health status changes an individual makes, the IFE is bounded by −4 and +2. For health status changes in the upper end of the distribution (FE3, FE4), setting the IFE equal to zero is a reasonable assumption to make. For changes into good and very good health, for instance, one of the two most frequent values of the bimodal distribution is zero. However, for the health changes at the two lower end of the distribution (FE1, FE2), this assumption is difficult to defend. The most frequent value of the IFE in the sample of changes out of bad or poor health centers approximately around −3 and −2, respectively. Neither distribution includes zero as a possible value for the IFE.

Moreover, we regress the threshold-specific IFE on the initial conditions of socio-economic and other control variables to test its correlation with our right-hand side variables. Table 2 reports the results for four regressions. Models (1) to (4) differ with respect to the health status changes made by individuals. Across models (1) and (3) the IFE is the smallest (or the largest in absolute value) for the oldest age-groups, 71 and older. Since according to equation (5) the IFE is a difference between the individual-specific health endowment and individual- and threshold specific reporting bias, we cannot interpret these age-group coefficients unambiguously. In the case of the older age-groups it is intuitive to argue that they dispose of a worse health endowment due to age-related diseases (Liang et al., 2005). In the case for younger age-groups, who also dispose of a very small IFE it is more likely that the reporting bias dominates. Younger individuals are expected, given their age, to have a relative good health endowment. One example is age-group 40 to 51 in Models (3) and (4). We would expect individuals of this group to report a lower level of health given their objective health status. A similar argument would hold for the highly educated from the samples of low health states (Models (1) and (2)). For members of this group the IFE is smaller than for any of the groups with a lower level of education. The latent health endowment of this group is, a priori, expected to be high. We find a reverse direction of the reporting bias for East Germans, who have a higher IFE than West Germans. Unless East Germans do not have a better health endowment[14], we expect East Germans to over-report their true health level. More important for our analysis is the finding that, at least for Models (2) to (4), the IFE significantly and positively correlates with household equivalent income and that the correlation is the strongest for the oldest age-group in models (1) to (3). This supports the hypothesis that the influence of the IFE on individual choices grows for older age-groups (Scarr and McCartney, 1983). Thus, the bias resulting from wrong model assumptions about the links of unobserved heterogeneity with right-hand side variables would be larger for older age-groups. In sum, these results suggest that the IFE, though nonlinearly, varies substantially across age-groups and strongly correlates with both proxies of socio-economic status. In conjunction with the finding of a non-normally distributed IFE (Figure (16)), we suggest that the generalized conditional fixed effects logit is the preferred model.

Finally, we test whether calculating of the marginal effects with the empirical approximation of the IFE has implications for our conclusions. A close look at the younger age-groups in Figure 10 and 11 shows that the alternatively calculated marginal effects (ALT) for the GREL and the GCFEL yields no changes vis-a-vis the original marginal effect (FS). Even though we observe a change in marginal effects for the older age-groups, this difference is limited to 1 to 2 percentage points. These differences for the older age-groups are expected, as we have shown in Table 2 that the assumption of a zero IFE is

[14] A better health endowment among East Germans may have resulted from healthier life-styles during the socialist political regime.

the least tenable for the older age-groups.

We conclude from this sensitivity analysis that the low magnitude of the marginal effects in the GREL and GCFEL models do not result from the arbitrary assumption that $\hat{\alpha}_{ij} = 0$. It may be used with no concerns for the younger age-groups.

4.3.3 Smaller Sub-samples

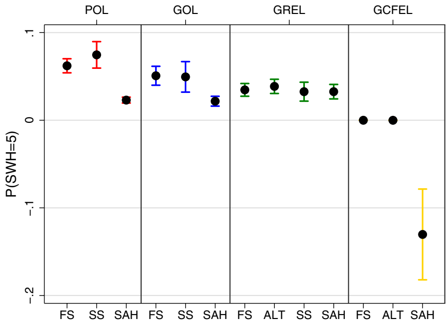
Finally, we may hypothesise that the small marginal effects of the GCFEL are the result of the small sample sizes which the method yields rather than of the control for unobserved heterogeneity. The idea here is that the sample is a highly self-selected group of individuals who, on the one hand, exhibit sufficient variation in self-reported health and on the other hand report, at least once, a very good health status. Table 1 indicates that these self-selected samples in the GCFEL for j= 4 are about one-third (younger groups) to two-thirds (older groups) smaller than in the full sample. In the light of this sample selection argument, the small marginal effects identified in the GCFEL for the highest possible health status could be driven by different mean values of the independent variables and unobserved heterogeneity. One way to test this hypothesis is to repeat our analysis for the pooled ordered logit (POL), the generalized ordered logit (GOL), and the generalized random effects logit (GREL) by using the sample of the GCFEL that results for changes into the highest possible health status. If the marginal effects for the POL, GOL and the GREL are then similar to those of the GCFEL, or at least much smaller, we would interpret this finding in favour of the sample selection hypothesis. Figures 10 and 11 show the small sample marginal effects (SS).

What we find striking is that for all age-groups the smaller sub-samples yield a larger marginal effect (SS) than for the original marginal effect that uses the full sample (FS). This finding holds true for all models, except for the POL in the case of age-group 41 to 50. In particular, the difference sums up to a maximum of 4 percentage points for all age-groups in the GOL and to 1 to 2 percentage points in the POL and GREL models.

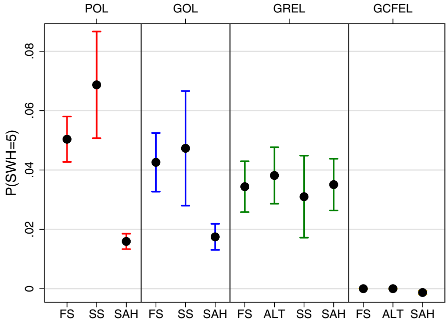
Thus, we do not find evidence of the hypothesis that the small marginal effects that result from the GCFEL model are driven by the selected, smaller sample sizes.

Table 2: Regression of threshold-specific fixed effects on initial values of regressors

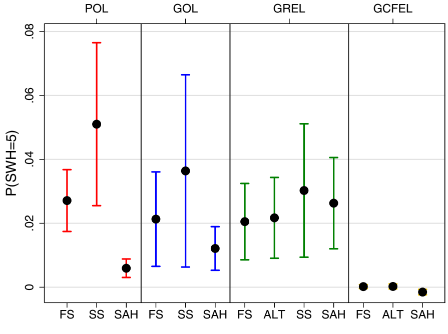
	FE (j=1)	FE (j=2)	FE (j=3)	FE (j=4)
	(1)	(2)	(3)	(4)
East German	.211***	.268***	.252***	.686***
	(.019)	(.015)	(.013)	(.026)
Immigrant	-.002	.013	.028*	.038
	(.023)	(.019)	(.016)	(.029)
9 & 10 yrs & prof training	-.149***	-.025	.078***	.272***
	(.021)	(.018)	(.015)	(.028)
10 to 13 yrs & prof training	-.402***	-.089***	.044**	.251***
	(.028)	(.023)	(.019)	(.038)
13 yrs & university degree	-.163***	-.102***	.008	.141***
	(.029)	(.023)	(.019)	(.036)
Age-group 31 to 40	.115	-.368	-.525***	-1.490***
	(.315)	(.236)	(.175)	(.319)
Age-group 41 to 50	.909***	.107	-1.074***	-2.342***
	(.296)	(.225)	(.172)	(.324)
Age-group 51 to 60	.271	.208	-.331*	-1.853***
	(.291)	(.226)	(.172)	(.360)
Age-group 61 to 70	-1.904***	-.764***	-.437**	-.829*
	(.311)	(.257)	(.219)	(.457)
Age-group 71 & older	-5.254***	-3.009***	-2.634***	-1.477**
	(.339)	(.322)	(.306)	(.692)
Income 16 to 30	.058*	.323***	.352***	.829***
	(.032)	(.024)	(.015)	(.025)
Income 31 to 40	.0008	.364***	.437***	1.070***
	(.036)	(.027)	(.021)	(.041)
Income 41 to 50	-.139***	.282***	.514***	1.158***
	(.031)	(.024)	(.020)	(.040)
Income 51 to 60	-.052*	.278***	.423***	1.110***
	(.029)	(.024)	(.020)	(.045)
Income 61 to 70	.197***	.376***	.420***	.950***
	(.033)	(.029)	(.028)	(.059)
Income 71 & older	.708***	.702***	.738***	1.025***
	(.039)	(.042)	(.042)	(.097)
Unemployed	.490***	.523***	.484***	.606***
	(.030)	(.027)	(.023)	(.047)
Separate	.160***	.151***	-.013	.142***
	(.028)	(.024)	(.021)	(.047)
Single	.182***	.236***	.149***	.346***
	(.025)	(.019)	(.015)	(.028)
HH size	.017***	.005	.003	.023***
	(.006)	(.005)	(.004)	(.007)
Const.	-2.554***	-3.618***	-3.184***	-6.531***
	(.218)	(.161)	(.106)	(.176)
Obs.	2813	6147	9585	8423
R^2	.583	.303	.25	.352
F statistic	195.485	133.14	159.484	228.083



(a) 16 to 30

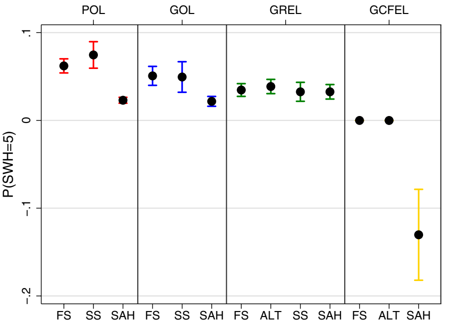


(b) 31 to 40

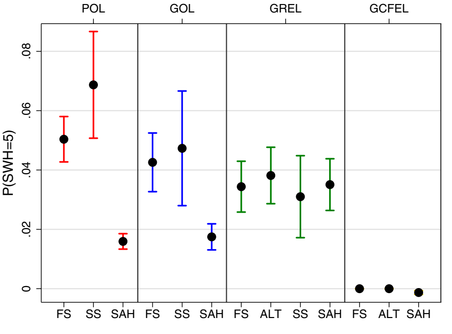


(c) 41 to 50

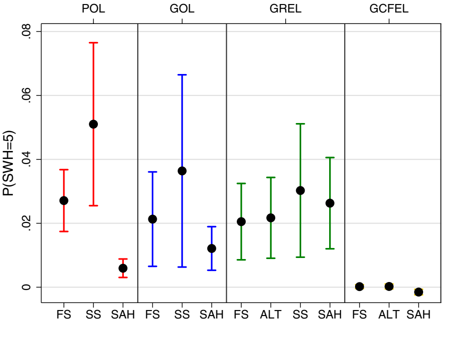
Figure 10: Sensitivity analysis, men only and youngest three age-groups



(a) 51 to 60

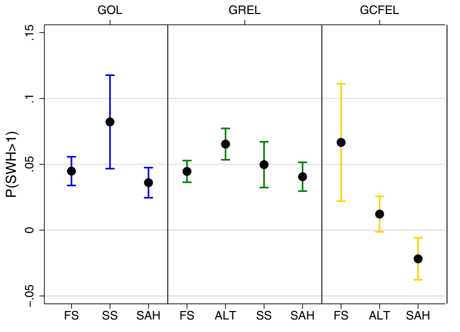


(b) 61 to 70

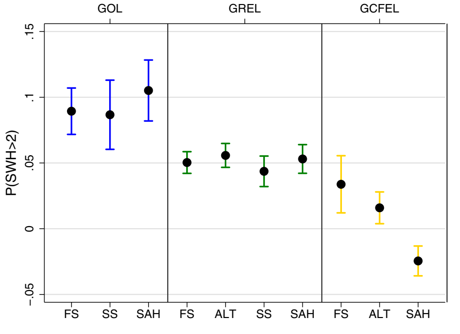


(c) 71 and older

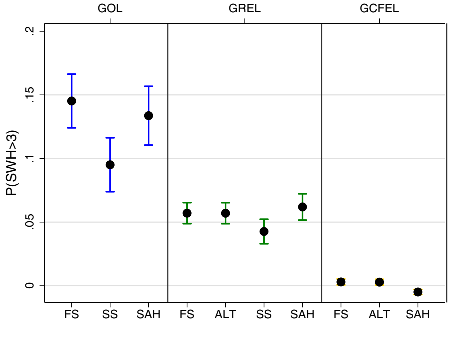
Figure 11: Sensitivity analysis, men only and oldest three age-groups



(a) Probability to exit bad health



(b) Probability to exit bad or poor health



(c) Probability to enter good or very good health

Figure 12: Sensitivity analysis, men only and for age-group 51 to 60

However, this does not hold true when looking at the marginal changes in the probabilities to exit bad or poor health or entering good or very good health. We illustrate these differences in Figures 12(a), 12(b) and 12(c) for age-group 51 to 60. In these three cases the marginal effects of income on the respective probabilities (SS) for both the GOL and the GREL model are smaller than those from the full sample (FS) by a magnitude ranging between 1 to 5 percentage points. Despite these smaller magnitudes, these marginal effects still don't resemble those yielded by the GCFEL model. These are highly negative for the two probabilities of exiting the lower tail of the health distribution.

We, therefore conclude that the large differences in the marginal effects of income on health status are mainly driven by the individual, threshold-specific effect and not by a particular sample selection.

5 Conclusions

We investigate whether the assumptions imposed on the relationship between unobserved heterogeneity and socio-economic status in assessing self-assessed health makes any practical difference. To illustrate our point we estimate a model of health satisfaction using 22 waves of the German Socio-Economic Panel and collapsing the eleven-point measure into a five-point scale analogous to self-assessed health.

We propose a generalized conditional fixed effects logit (GCFEL) model that allows a heterogeneous impact of income on health while controlling for unobserved heterogeneity. By imposing various restrictions on this general formulation, we compare the relative performance of pooled and generalized ordered logits (POL and GOL), generalized random effects logit (GREL), and the conditional ordered fixed effects logit (COFEL) with the GCFEL. The main variable of interest is the log of net monthly household equivalent income which we interact with six age-group specific dummies. Estimating the marginal effect of income on the probability to report very good health for various age-groups reflects the hypothesis that the unobserved individual heterogeneity that proxies personality traits, intelligence or genetic endowment, has a growing importance in determining health outcomes in older age. Therefore, we expect that controlling for unobserved heterogeneity may reveal greater differences vis-à-vis models that do not for an older sub-sample.

We find that not controlling for unobserved heterogeneity (POL and GOL) yields a marginal effect that is significantly greater than those yielded by models that consider its presence in the data (GREL and GCFEL). When allowing for correlation between the individual fixed effect and household income the marginal effect approaches zero for all age-groups and this effect is the same for men and women alike. Under the assumption of no correlation between the individual fixed effects and income (GREL), the magnitude of the marginal effect lies half-way between the POL, GOL and the GCFEL. For the middle-aged groups a doubling of the log of disposable income leads to a relatively large increase of 4 to 6 times in the probability to report very good health. Accordingly, the marginal effect yielded by the the GOL and the GREL are smaller than this upper bound, but strictly greater than zero.

The differences in the magnitude of the marginal effects across the pooled and the fixed effects models are the strongest for the older age-groups. This finding supports our hypothesis that controlling for the potential correlation between the fixed effect and income is the more important the older an individual is. In older age the fixed effect correlates stronger with socio-economic status. To this extent the expected bias of unobserved heterogeneity, when not accounted for, is greater for these groups.

Taking our analysis further to assessing other health states we find two interesting phenomena: First, the model that does not account for unobserved heterogeneity (GOL) yields the highest marginal effect. However, the magnitude of this effect is heterogeneous across health states: doubling the log of monthly income will five-fold the probability of exiting a bad health status, whereas it nearly 15-folds the probability to report higher levels of health status. Our results are in line with Boes and Winkelmann (2006) who find varying marginal effects of income on the different levels of general life satisfaction. To this extent we are also able to show that the Ferrer-i-Carbonell and Frijters (2004) proposal, that looks at changes across health changes beyond one's individual mean, is too a narrow focus given the heterogeneity of income effects on health.

Second, controlling for unobserved heterogeneity and, therefore, allowing for its potential correlation with income, the marginal effect is no longer equal to zero for exiting the lower health states. This suggests that the individual fixed effect does mediate the relationship between income and changes in higher health outcomes, but not between income and changes in lower health outcomes.

Last, our sensitivity checks ensure that our results are neither driven by the particular choice of proxy for heath status, nor by the arbitrary assumption of setting the individual fixed effect equal to zero when computing the marginal effects, nor by the self-selected samples incurred by the conditional fixed effects logit approach. By approximating the individual fixed effect we learn that the arbitrary assumptions on normality and orthogonality of the individual fixed effect made in the random effects logit are not tenable. The GCFEL is the preferred model. Nevertheless, the GCFEL suffers from the potential drawback that we are forced to discard many observations for assessing the probability of reporting the upper and lower bounds of health satisfaction. This becomes evident when we use the alternative proxy for health status, self-assessed health, in which we have data only for a panel of 12 years. Given the constant growth of panel datasets, we believe sample sizes will be of smaller concern to researchers in the future.

References

Alesina, A., R. DiTella, and R. MacCulloch (2004). Inequality and happiness: Are European and Americans different? *Journal of Political Economy 88*, 2009–2042.

Andersen, E. (1970). Asymptotic properties of conditional maximum likelihood estimators. *Journal of the Royal Statistical Society. Series B 32*, 283–301.

Andersen, E. (1971). Asymptotic properties of conditional likelihood ratio tests. *Journal of the American Statistical Society 66*, 630–633.

Arulampalan, W. (1999). Practitioner's corner a note on estimated coefficients in a random effects probit models. *Oxford Bulletin of Economics and Statistics 61*(4), 0305–9049.

Auld, M. and N. Sidhu (2005). Schooling, cognitive ability and health. *Health Economics 14*(10), 1019–1034.

Betty, G. and I. Deary (2004). Early life intelligence and adult health. *British Medical Journal 329*, 585–586.

Blackburn, M. and D. Neumark (1995). Are OLS estimates of the return to schooling biased downward? Another look. *Review of Economics and Statistics 77*, 217–229.

Blanchflower, D. and A. Oswald (2004). Well-being over time in Britain and the USA. *Journal of Political Economy 88*, 1359–1386.

Boes, S. and R. Winkelmann (2006). Ordered response models. *Allgemeines Statistisches Archiv 90*, 165–179.

Bouchard, T. J. (1998). Genetic and environmental influences on adult intelligence and special mental abilities. *Human Biology 70*, 257–279.

Cameron, A. and P. Trivedi (2005). *Microeconometrics: Methods and Applications.* New York: Cambridge University Press.

Card, D. (1995). Earnings, ability, and schooling revisited. In S. Polachek (Ed.), *Research in Labour Economics*, Volume 14, pp. 23–48. JAI Press.

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *Review of Economics and Statistics 47*(1), 225–238.

Clark, A. and A. Oswald (1994). Unhappiness and unemployment. *The Economic Journal 104*, 648–659.

Clark, A. (2001). What really matters in a job? Hedonic measurement using quit data. *Labour Economics 8*, 223–242.

Contoyannis, P., A. Jones, and N. Rice (2004). The dynamics of health in the British Household Panel Survey. *Journal of Applied Econometrics 19*, 473–503.

Das, M. and A. Van Soest (1999). A panel data model for subjective information on household income growth. *Journal of Economic Behavior and Organization 40*, 409–426.

DiTella, R., R. McCulloch, and A. Oswald (2001). Preferences over inflation and unemployment. *American Economic Review 91*, 335–341.

DiTella, R., R. McCulloch, and A. Oswald (2003). The macroeconomics of happiness. *Review of Economics and Statistics 85*, 809–827.

Easterlin, R. (1995). Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior and Organization 27*, 35–47.

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

Ferrer-i-Carbonell, A. (2005). Income and wellbeing: An empirical analysis of the comparison income effect. *Journal of Political Economy 89*, 997–1019.

Frey, B., S. Luechinger, and A. Stutzer (2006). The life satisfaction approach to the value of public goods: The case of terrorism. *University of Zurich Working Paper 184*.

Frijters, P., I. Geishecker, M. Shields, and J. Haisken-DeNew (2006). Can the large swings in Russian life satisfaction be explained by ups and downs in real incomes? *Scandinavian Journal of Economics 108*(3), 433–458.

Frijters, P., J. P. Haisken-DeNew, and M. A. Shields (2004a). Changes in the pattern and determinants of life satisfaction in Germany following reunification. *Journal of Human Resources 39*, 649–674.

Frijters, P., J. P. Haisken-DeNew, and M. A. Shields (2004b). Money does matter! Evidence from increasing real incomes and life satisfaction in East Germany following reunification. *American Economic Review 94*, 730–740.

Frijters, P., J. Haisken-DeNew, and M. Shields (2005b). The causal effect of income on health: Evidence from German reunification. *Journal of Health Economics 24*, 997–1017.

Frijters, P., J. Haisken-DeNew, and M. Shields (2005a). Socio-economic status, health shocks, life satisfaction and mortality: Evidence from an increasing mixed proportional hazard model. *IZA Discussion Paper 1488*.

Gannon, B. (2005). The influence of economic incentives on reported disability status. *mimeo*.

Gottfredson, L. and I. Deary (2004). Intelligence predict health and longvity, but why? *Current Directions in Psychological Research 13*(1), 1–4.

Gottfredson, L. (1997). Why G matters: The complexity of everyday life. *Intelligence 24*, 79–132.

Gottfredson, L. (2002). G: Highly general and highly practical. In R. Sternberg and

E. L. Grigorenko (Eds.), *The General Factor of Intelligence*, pp. 331–380. Mahwah, NJ: Erlbaum.

Gottfredson, L. (2004). Intelligence: Is it the epidemiologists' elusive fundamental cause of social class inequalities in health? *Journal of Personality and Social Psychology Copyright 86*(1), 174–199.

Gravelle, H. and M. Sutton (2003). Income related inequalities in self-assessed health in Britain: 1979-1995. *Journal of Epidemiology and Community Health 57*, 125–129.

Gravelle, H. and M. Sutton (2006). Income, relative income, and self-reported health in Britain: 1979-2000. *Centre for Health Economics Discussion Paper No. 2006/06.*

Groot, W. (2000). Adaptation and scale of reference bias in self-assessments of quality of life. *Journal of Health Economics 19*, 403–420.

Hagenaars, A., K. de Vos, and M. Zaidi (1994). *Poverty Statistics in the Late 1980s: Research Based on Micro-data.* Luxembourg: Office for Official Publications of the European Communities.

Helliwell, J. (2004). How's life? Combining individual and national variables to explain seubjective wellbeing. *Economic Modelling 20*, 331–360.

Hernandez-Quevedo, C., A. Jones, and N. Rice (2004). Reporting bias and heterogeneity in self-assessed health. Evidence from the British Household Panel Survey. *HEDG WP University of York 05/04.*

Idler, E. and Y. Benyamini (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior 38*, 21–37.

Jencks, C., S. Bartlett, M. Corcoran, J. Crouse, D. Eaglesfield, G. Jackson, K. McClelland, P. Mueser, M. Olneck, J. Schwartz, S. Ward, and J. Williams (1979). *Who gets ahead? The determinants of economic success in America.* New York: Basic Books.

Jencks, C., M. Smith, H. Acland, and M. Bane (1972). *Inequality: A reassessment of the effect of family and schooling in America.* New York: Basic Books.

Jones, A., X. Koolman, and N. Rice (2006). Health-related non-response in the British Household Panel Survey and European Community Household Panel: Using inverse-probability-weighted estimators in non-linear models. *Journal of the Royal Statistical Society. Series A 169*(3), 1–27.

Jones, A. and J. Wildman (2007). Health, income, and relative deprivation: Evidence from the BHPS. *mimeo*.

Kahnemann, D. and A. Krueger (2006). Developments in the measurement of subjective wellbeing. *Journal of Political Economy 22*, 3–24.

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

Layard, R. (2005). *Happiness. Lessons from a New Science.* London: Allan Lane.

Lelkes, O. (2006). Testing freedom. happiness, religion, and economic transition. *Journal of Economic Behavior and Organization 59*, 173–194.

Liang, J., B. Shaw, N. Krause, J. Bennett, E. Kobayashi, T. Fukaya, and Y. Sugihara (2005). How does self-assessed health change with age? A study of older adults in Japan. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences 60*, 224–232.

Lindley, J. and P. Lorgelly (2003). The relative income hypothesis: Does it exist over time? Evidence from the BHPS. *University of East Anglia, mimeo*.

Lykken, D. and A. Tellegen (1996). Happiness is a stochastic phenomenon. *Psychological Science 7*, 186–189.

Mackenbach, J., J. Simon, C. Looman, and I. Joung (2002). Self-assessed health and mortality: Could psychological factors explain the association? *International Journal of Epidemiology 31*(6), 1162–1168.

McClearn, G., B. Johansson, S. Berg, N. Pedersen, F. Ahern, S. Petrill, and R. Plomin (1997). Substantial genetic influence on cognitive abilities in twins 80 or more years old. *Science 276*, 1560–1563.

McGue, M., J. T.J. Bouchard, W. Iacono, and D. Lykken (1993). Behavioral genetics of cognitive ability: A life-span perspective. In R. Plomin and E. McClearn (Eds.), *Nature, Nurture, and Psychology*, pp. 59–76. Washington, DC: American Psychological Association.

Miller, D. and C. Paxson (2006). Relative income, race, and mortality. *Journal of Health Economics 25*, 979–1003.

Mossey, J. and E. Shapiro (1982). Self-rated health: a predictor of mortality among the elderly. *American Journal of Public Health 72*, 800–808.

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica 46*(1), 69–85.

Neyman, J. and E. Scott (1948). Consistent estimates based on partially consistent observations. *Econometrica 16*, 1–32.

Oswald, A. (1997). Happiness and economic performance. *The Economic Journal 107*, 1815–1831.

Perls, T. (2002). Genetic and environmental influences on exceptional longevity and the AGE nomogram. *Annals of the New York Academy of Sciences 959*, 1–13.

Plomin, R. and S. A. Petrill (1997). Genetics and intelligence: What's new? *Intelligence 24*, 53–77.

Plomin, R. (1986). *Development, Genetics, and Psychology.* Hillsdale, NJ: Erlbaum.

Sadana, R., C. Mathers, A. Lopez, C. Murray, and K. Iburg (2000). Comparative analysis of more than 50 household surveys on health status. *World Health Organisation. GPE Discussion Paper 15* (EIP/GPE/EBD).

Scarr, S. and K. McCartney (1983). How people make their own environment: A theory of genotype # environment effects. *Child Development* (2), 424–435.

Shmueli, A. (2002). Reporting heterogeneity in the measurement of health and health-related quality of life. *Pharmacoeconomics 20* (6), 405–412.

Singh-Manoux, A., J. Ferrie, J. Lynch, and M. Marmot (2005). The role of cognitive ability (intelligence) in explaining the association between soecioeconomic position and health: Evidence from the Whitehall II prospective cohort study. *American Journal of Epidemiology 161* (9), 831–839.

Subramanian, S. and I. Kawachi (2004). Income inequality and health: What have we learnt so far? *Epidemiological Reviews 26*, 78–91.

Terza, J. (1985). Ordered probit: A generalization. *Communications in Statistics - A Theory and Methods 14*, 1–11.

Van Doorslaer, E. and A. Jones (2003). Inequalities in self-reported health: Validation of a new approach to measurement. *Journal of Health Economics 22*, 61–87.

Van Praag, B. and A. Ferrer-i-Carbonell (2004). *Happiness Quantified: A Satisfaction Calculus Approach*. Oxford: Oxford University Press.

Van Praag, B. and A. Ferrer-i-Carbonell (2005). Using happiness surveys to value intangibles: The case of airport noise. *The Economic Journal 115*, 224–246.

Whalley, L. and I. Deary (2001). Longitudinal cohort study of childhood IQ and survival up to age 76. *British Medical Journal 322*, 1–5.

Winkelmann, L. and R. Winkelmann (1998). Why are the unemployed so unhappy? Evidence from panel data. *Economica 65*, 1–15.

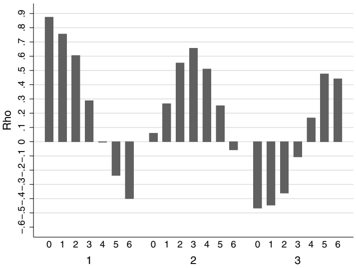
Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: The MIT Press.

Appendix: Coding of Health Satisfaction

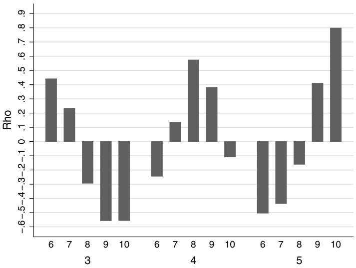
In Figure 13 we present the cross-correlations between the various sub-categories of satisfaction with health (SWH) coded from 0 to 10 and the various sub-categories of self-assessed health (SAH) coded from 1 to 5. The horizontal axis scales the value of the corresponding sub-category of the both measures. In Figure 13(a) we displayed values of SWH from 0 to 6 and of SAH from 1 to 3. In Figure 13(b) the horizontal axis displays values of SWH from 6 to 10 and of SAH from 3 to 5. The bars tell us the strength and the direction of the association between the various values of each sub-category. These associations have been calculated by the authors according to the concept of tetrachord correlations for binary variables in the presence of latent variables. We subsumed each sub-category of SWH under a value of SAH for which the tetrachord correlation was positive and maximized across all possible combinations of the respective sub-category of SWH with the remaining sub-categories of SAH. For instance, we compared the correlation between SWH$=3$ and SAH$=1$ with all possible correlations between SWH$=3$ and SAH$=2,3,4,5$.

In Figure 13(c) we crudely graph the same tetrachord correlations in three-dimensional space. On the vertical axis we graph ρ, the degree of correlation between two possible combinations of sub-categories of SWH and SAH. On the two horizontal axes we graph satisfaction with health (left axis) and self-assessed health (right axis). Small values for both SWH and SAH run from left to right of each axis. A closer look allows to notice that the correlations (high vertical lines) occur where high (low) values of SWH coincide with high (low) values of SAH.

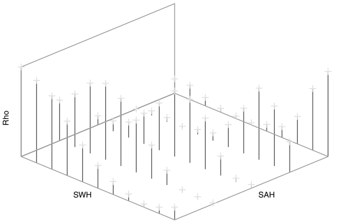
The subsequent five Figures 14(a) to 14(e) show the percentage of individuals who report, let's say SWH$=1$, who also report SAH$=1$. The light grey bars represent the perspective of someone reporting SAH, and the dark grey bar represents the perspective of someone reporting SWH.



(a) SAH= 1 to 3 & SWH= 0 to 6

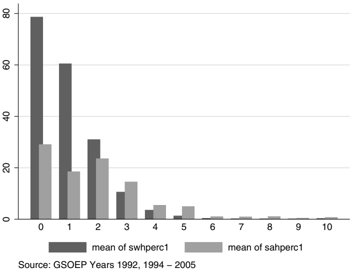


(b) SAH= 3 to 5 & SWH= 6 to 10

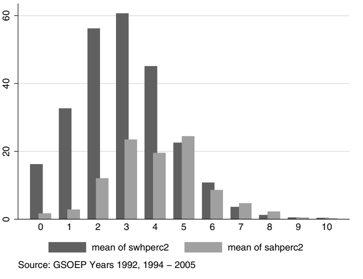


(c) Cross-correlations for SAH and SWH

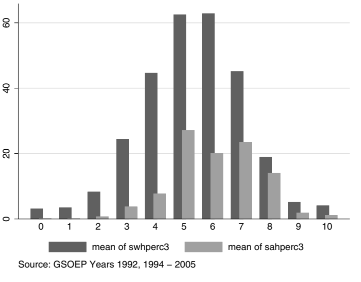
Figure 13: Cross-correlation of subcategories of SWH and SAH



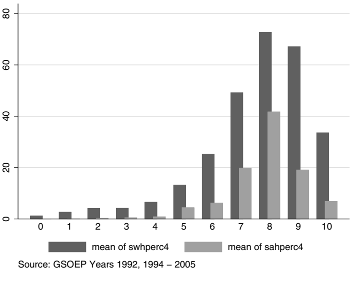
(a) SAH= 1



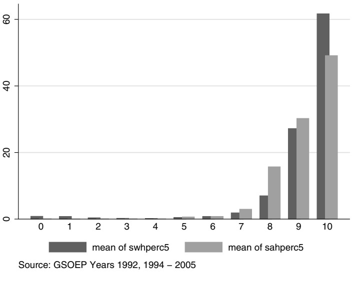
(b) SAH= 2



(c) SAH= 3



(d) SAH= 4



(e) SAH= 5

Figure 14: Proportion of individuals reporting any value of SWH if SAH= 1 to 5

Appendix: Data

In Tables 3 and 4 we report the descriptive statistics for the measures of subjective health status, SWH, SAH, and WAH, and the set of independent variables for men and women, respectively. In the bottom two rows of each table we report the number of individuals and person-year observations for each age-group. Standard deviations are given in parentheses.

Sample sizes, both in total and for the age-specific subgroups, are approximately the same for both men and women. In total, we have more than $20,000$ individuals in each group. For the age-specific subgroups, the youngest group aged 16 to 30 is largest in size with nearly $7,500$ individuals, and the oldest age-group is the smallest with less than $3,000$ individuals. On average, individuals remain in the panel for seven years.

All three measures of subjective health are decreasing in age, whereas the smallest variation in the mean value is reported for WAH. This is a common phenomenon encountered for measures with small units on a Likert scale. Across all age-group individuals report a value of SAH$\approx$ 2.7 and WAH$\approx$ 2.1, and SWH$\approx$ 3.6). Average household equivalent income ranges from approximately $1,200$ Euro per month for the youngest age-group to approximately $1,350$ Euro per month for the oldest age-group. Individuals from the age-group 51 to 60 have the highest average household income, which illustrates growing income over the life-cycle before retirement, when income falls again. The unemployment data reflects the long-term situation in the German labour market, in so far as these are the averages taken over the past 22 years. Individuals above the age of 50 face the highest risk of being unemployed ($\approx$ 12 percent), followed by the youngest age-group ($\approx$ 8 percent).

Table 3: Descriptive statistics of selected variables for men

	ALL	16 to 30	31 to 40	41 to 50	51 to 60	61 to 70	71
Variable							
SAH	2.54	2.02	2.30	2.55	2.82	2.97	3.22
	0.94	0.80	0.82	0.87	0.93	0.91	0.95
SWH	3.66	4.13	3.84	3.61	3.34	3.24	3.04
	1.08	0.92	0.96	1.04	1.10	1.07	1.16
WAH	2.19	2.47	2.33	2.21	2.07	1.96	1.83
	0.69	0.65	0.65	0.65	0.68	0.66	0.66
Secondary	0.18	0.19	0.16	0.19	0.22	0.17	0.13
	0.38	0.39	0.36	0.39	0.41	0.38	0.34
Intermediary	0.47	0.54	0.46	0.42	0.42	0.47	0.53
	0.50	0.50	0.50	0.49	0.49	0.50	0.50
Upper	0.15	0.12	0.18	0.16	0.14	0.14	0.17
	0.36	0.32	0.38	0.36	0.35	0.35	0.38
University	0.20	0.16	0.21	0.23	0.22	0.22	0.17
	0.40	0.36	0.41	0.42	0.42	0.41	0.38
Income	1348.72	1198.60	1271.09	1423.04	1514.56	1439.67	1345.01
	966.36	668.98	666.97	1039.92	1085.82	1433.39	992.81
Unemployed	0.07	0.07	0.06	0.07	0.12	0.03	0.00
	0.25	0.26	0.24	0.25	0.32	0.17	0.03
Separate	0.08	0.02	0.07	0.09	0.09	0.11	0.21
	0.27	0.13	0.25	0.29	0.28	0.31	0.41
Single	0.25	0.76	0.21	0.07	0.04	0.02	0.01
	0.43	0.43	0.40	0.25	0.19	0.15	0.11
Number HH members	3.02	3.33	3.27	3.41	2.78	2.22	1.99
	1.37	1.51	1.34	1.36	1.25	0.85	0.68
Years in panel	7.09	4.67	7.44	7.58	7.90	8.58	8.64
	5.43	3.30	5.35	5.67	5.74	6.08	6.29
N*T	148417	35882	31150	28935	24604	17744	10097
N	20171	7365	6570	6343	5259	3854	2081

Table 4: Descriptive statistics of selected variables for women

	ALL	16 to 30	31 to 40	41 to 50	51 to 60	61 to 70	71
Variable							
SAH	2.65	2.14	2.36	2.62	2.91	3.05	3.41
	0.97	0.82	0.84	0.89	0.92	0.90	0.94
SWH	3.55	4.04	3.79	3.53	3.25	3.18	2.86
	1.11	0.96	0.99	1.05	1.09	1.07	1.17
WAH	2.11	2.37	2.29	2.17	2.01	1.87	1.71
	0.69	0.67	0.64	0.66	0.66	0.65	0.67
Secondary	0.28	0.20	0.19	0.25	0.34	0.39	0.47
	0.45	0.40	0.39	0.43	0.47	0.49	0.50
Intermediary	0.47	0.50	0.48	0.47	0.46	0.46	0.42
	0.50	0.50	0.50	0.50	0.50	0.50	0.49
Upper	0.11	0.15	0.14	0.10	0.07	0.07	0.06
	0.31	0.36	0.35	0.30	0.26	0.25	0.24
University	0.14	0.15	0.19	0.18	0.14	0.09	0.05
	0.35	0.35	0.39	0.39	0.35	0.29	0.22
Income	1278.96	1136.78	1234.33	1430.13	1452.00	1277.71	1152.33
	863.11	637.13	728.02	898.20	1142.87	948.41	814.72
Unemployed	0.06	0.07	0.08	0.07	0.11	0.01	0.00
	0.24	0.26	0.27	0.26	0.31	0.11	0.03
Separate	0.17	0.03	0.10	0.13	0.17	0.29	0.61
	0.38	0.16	0.30	0.34	0.37	0.45	0.49
Single	0.19	0.61	0.12	0.04	0.03	0.04	0.06
	0.39	0.49	0.33	0.20	0.16	0.19	0.24
Number HH members	2.89	3.22	3.53	3.34	2.50	1.97	1.65
	1.39	1.49	1.27	1.28	1.11	0.84	0.93
Years in panel	7.16	4.66	7.34	7.78	8.03	8.36	8.58
	5.45	3.28	5.25	5.71	5.84	5.96	6.21
N*T	156821	36190	32588	29648	23689	19247	15458
N	20950	7438	6779	6385	5084	4009	2951