Health Outcomes and Socio-economic Status among the Elderly in China: Evidence from the CHARLS Baseline Data

Xiaoyan Lei
CCER, Peking University

Xiaoting Sun
CCER, Peking University

John Strauss
University of Southern California

Yaohui Zhao
CCER, Peking University

Gonghuan Yang
China Center for Disease Control and Prevention

Perry Hu
University of California, Los Angeles

Yisong Hu
ISSS, Peking University

Xiangjun Yin
China Center for Disease Control and Prevention

For presentation at the Population Association of America Annual Meetings,
April 2013, New Orleans

Abstract Using a very rich set of health indicators that include both self-reported measures and biomarkers from the CHARLS baseline data, we document health conditions of the Chinese elderly and examine correlations between these health outcomes and two important indicators of socio-economic status: education and log of per capita expenditure (PCE), our preferred measure of household resources. As expected, we find that Chinese elderly are facing challenges from chronic diseases including hypertension. Overnutrition has become a bigger problem than undernutrition, particularly for women, reflected in a higher rate of overweight compared to underweight. Disability rates are also high, especially for rural women, who also report suffering from more pain. In general, education and PCE tend to be positively correlated with better health outcomes, as it is in other countries. For PCE the relationship is very nonlinear. At low levels of PCE, there exists a positive correlation with better health outcomes, while for higher levels of PCE the relationship flattens out. Unmeasured community influences turn out to be highly important, much more so than one usually finds in other countries. We also find a large degree of under-diagnosis of hypertension, a major health problems that afflicts the aged. This implies that the current health system is not well prepared to address the rapid aging of the Chinese population.

Keywords Health status; Health-SES correlations; Chinese elderly

Introduction

We are concerned in this paper with measuring health outcomes among the elderly in China, and examining the relationships between different dimensions of health status and measures of current socio-economic status (SES). China has undergone a health revolution over the past 50 years, with life expectancy having risen from 46 in the 1950s to just over 74 in 2009 (Wagstaff et al. 2009; World Health Organization 2012). Driving this change, the mortality rates for those under 5 fell dramatically from 225 per 1,000 live births in 1960 to 48 in 1990 and 18 in 2010 (Wagstaff et al. 2009; UNICEF 2012). Most of this decline was due to an increasing control over infectious disease and undernutrition. As a result, infectious diseases have been progressively replaced by chronic diseases as the major source of ill-health and mortality (Hossain 1997; Lopez et al. 2006).

As China has been passing through its health transition, it has also been undergoing a nutrition transition, which has both positive and negative aspects (Popkin et al. 1993, 1995a; Popkin 1999, 2002). Among the principle dimensions of this transition has been a dramatic rise in body mass index (BMI) among adults and a large change in diet towards more 'fatty foods' (Popkin et al. 1995b). For instance Luo (2003), using the China Health and Nutrition Survey (CHNS), documents an increase in overweight adults over 20 years from 1989 to 1997, for women from 11% to 21% and from 6% to 17% for men. [1]At the same time, Luo shows that the fraction of elderly adults who are undernourished (a BMI under 18.5) has fallen, particularly so for those over 60 years old, from 19% to 13% for women and 20% to 12% for men from 1991 to 1997.

Related to these health and nutrition transitions has been China's demographic transition. China's elderly population will increase from under 10% of the total population in 2000 to 30% in 2050 (Kinsella and He 2009). The number of workers per pensioner has already fallen from over 12 in 1980 to 2 in 2005. This sharp demographic transition is likely to place stress on China's health system, which has been focused on disease at younger ages and on infectious, as opposed to chronic, diseases.

[1] Overweight is defined using World Health Organization standard of having a BMI 25 or above.

In this paper, we use the China Health and Retirement Longitudinal Study baseline data (CHARLS) to document health conditions among the elderly (aged 45 and over) in China. We use a very rich set of health indicators that include both self-reported measures and biomarkers. We also examine correlations between these health outcomes and two important indicators of socio-economic status (SES): education and log of per capita expenditure (log PCE), our preferred measure of household resources. While there exists a very large literature that examines the relationships between SES and health measures, little has been done on Chinese data to see whether correlations reported in many other countries are replicated in China, particularly so for the aged. While we cannot infer causality from these estimates, they tell us something important about the degree of health differentials by education and per capita expenditure (PCE).

In general, education and PCE tend to be positively correlated with better health outcomes, as it is in other countries. The PCE association is quite nonlinear, positive at lower levels of PCE and flattening out for higher levels. Unmeasured community influences turn out to be highly important, much more so than one usually finds in other countries. While it is not yet clear which aspects of communities matter and why they matter, we set up an agenda for future research on this topic. We also find a large degree of under-diagnosis of hypertension, a major health problem that afflicts the aged. This implies that the current health system is not well prepared to address the rapid ageing of the Chinese population.

This paper is divided into five sections. The next section briefly introduces the topic. Section "Data and Empirical Specifications" describes the data while our main empirical findings are presented in section "Results". The final section highlights our main conclusions.

Health-SES Correlations

Across most country settings, no matter which measures of SES are used (income, wealth or education), the evidence of this association between health and SES being large and pervasive is abundant (Marmot 1999; Smith 1999; Strauss and Thomas 1998).

Mainly due to the absence of high quality data, far less research has been conducted on the magnitude and underlying reasons for the SES-health gradient in China. China is about to age very rapidly and has at the same time been experiencing high but very unequal rates of economic growth. The extent to which this recent economic growth has improved the overall health of the Chinese population depends in part on understanding in the Chinese context the influence of economic well-being on health. Similarly, the degree to which growing inequalities in levels of economic resources in China are producing similarly large inequalities in health requires better understanding.

Part of the reason for a lack of any substantial research on this topic in the Chinese context is that until recently data in China were simply not up to the task or were largely unavailable to scholars either inside or outside China. Fortunately, this situation is changing. The investigators on this research project have been involved in a key ongoing data collection effort—the Chinese Health and Retirement Longitudinal Study (CHARLS)—that aims to remedy this situation.

This paper documents in detail the nature of the SES-health gradient in China. The analysis spans many salient measures of health status—including general health status, functional measures of disability (ADLs and IADLs), body pain, body mass index (BMI), hypertension, and survival expectations. In addition to standard measures of economic status widely used in other country settings, notably schooling, we examine a measure of income, household consumption, which is arguably a much better index of economic wellbeing in a country at the level of economic development that China now is, especially in more rural regions.

The salient nature of the SES-health gradient is not difficult to document. In most countries, at each age those in higher income or wealth groups are in much better health. These differences are quantitatively large. In the US the fraction reporting excellent health among the highest income quartile is 40 percentage points higher than those in the lowest quartile (Smith 1999). The health-income gradient widens until around age 50 after which it gradually contracts. Similar gradients of this magnitude appear in many other countries so that the US is not unique. While there is a broad consensus about the 'facts,' the interpretation of underlying forces creating a strong SES-health gradient and how their relative importance varies across countries at different levels of economic development are controversial.

A basic question in any country is whether differences in health by SES indicators such as education, income, and wealth largely reflect impacts from SES to health, or vice versa. Medical scientists often conclude that the dominant pathway is that variation in SES produces large health disparities; their main debate is about why lower economic status leads to poor health (Marmot 1999). Important insightful contributions by such scholars has involved investigating the influence of other factors besides access to high-quality health care or deleterious personal behaviors, both of which are believed to offer incomplete explanations. These contributions have emphasized long-term impacts of early childhood or even inter-uterine environmental factors (Barker 1997; Smith 2009), the cumulative effects of prolonged exposures to individual stressful events (Seeman et al. 1997), or reactions to macro-societal factors such as rising levels of income inequality (Wilkenson 1996).

This view that the principal direction of causation flows almost exclusively from SES to health has been challenged (Lleras-Muney 2005; Cutler and Lleras-Muney 2010; Smith 1999; Strauss 1986, 1993; Strauss and Thomas 1995, 1998, 2008; Thomas and Strauss 1997; Thomas 2010). The best evidence is in fact that causation goes in both directions, from SES to health and from health to SES.

Moreover, many omitted variables determine both health measures and SES. A good example, the subject of much recent research, is health when very young, which appears to affect health in older ages, but also SES outcomes such as education and incomes as adults (see in particular, Barker 1994; Gluckman and Hanson 2005; Hoddinott et al. 2008; and Maluccio et al. 2009). The latter two experimental studies document that nutritional supplements given to children before 3 years of age in Guatemala were associated with more completed schooling by the time they were adults, as well as higher adult labor force earnings. Other recent careful nonexperimental studies such as Lleras-Muney (2005) have established that the influence of SES on health status is in part causal.

With these caveats, it is still of much interest to examine the health-SES gradients. As discussed, these correlations will also indicate the degree of socio-economic inequalities in health outcomes and their relationships to inequality in SES, education and resources in particular. Income inequality is now of much policy interest. Health inequality is likely to be of policy interest as well in China, as it is in industrial countries.

Data and Empirical Specifications

We use the CHARLS national baseline data, which was designed after the Health and Retirement Study (HRS) in the US as a broad-purposed social science and health survey of the elderly in China. The baseline survey was conducted between June 2011 and March 2012. It is a nationally representative sample of people aged 45 and over, and their spouses, living in households in China.

The CHARLS baseline data represent some major advantages over existing Chinese data. First, CHARLS is nationally representative, which many other related surveys in China are not. Second, many Chinese datasets are not publicly available, though they may be available to some with special connections. One example is the Survey of the Aged Population, which started as a cross-sectional survey and now has a small panel component. The 2002 Nutrition and Health Survey is a specialized health survey of the entire population patterned after the United States' National Health and Nutrition Examination Survey (NHANES), and is another example. Even if they were made public, both have only very scant economic information (e.g., no consumption, poor income and wealth data) with which to examine SES-health relationships, a problem with other data sets.

The Chinese Longitudinal Healthy Longevity Survey is publicly available, but not nationally representative. It began as a survey of the oldest old (over 80 years), although younger cohorts have been added in recent years. Yet it too is far different from an HRS-style survey, particularly in its lack of economic variables such as consumption and wealth. In addition, there were few biomarkers collected in the earlier waves. Likewise the World Health Organization's Study on Global Ageing and Adult Health (SAGE), a series of global surveys on the aged, including one for China, is rather light on socio-economic information on income, wealth, retirement and family support.

Finally, the China Health and Nutrition Survey (CHNS) may be the best known household survey on China to the outside world, and has been the workhorse for many scholars. However there are important differences from an HRS survey. CHNS is not strictly an ageing survey, and its sample of elderly is small. Income is also not measured well in CHNS and there are no total consumption data (only food consumption). Moreover, CHARLS has richer health data, particularly on biomarkers. Very important for this study, CHARLS has available numeric codes for primary sampling units so that community-fixed effects regressions may be run. CHNS does not release community codes with its data. Moreover CHNS, with only nine provinces, is not nationally representative.

Hence for our purposes, CHARLS represents a major improvement for the study of health-SES relationships of the elderly in China, and it is publicly available.[2] The CHARLS baseline sample was drawn in four stages. County-level units (counties or urban districts) were sampled directly. All county-level units in all provinces except for Tibet were stratified by 8 regions, by whether they were urban districts (*qu*) or rural counties (*xian*), and by county GDP. They were sorted based on this stratification and 150 were randomly chosen proportional to population size.[3] These counties cover 28 out of 30 provinces, other than Tibet.

After the county units were chosen, the National Bureau of Statistics helped us to sample villages and communities within county units using recently updated village level population data. Our sample used administrative villages (*cun*) in rural areas and neighborhoods (*shequ*), which comprise one or more formal resident committees (*juweihui*), in urban areas as primary sampling units (PSUs). We selected three PSUs within each county-level unit, using PPS sampling, for a total of 450 PSUs. Note that rural counties contain both rural villages and urban neighborhoods, and it is also possible for urban districts to contain rural administrative villages.

In each PSU, we constructed our own sampling frame using Google Earth base maps.[4]

[2] See http://charls.ccer.edu.cn.

[3] A random number was drawn for the first county selected and then every Nth county thereafter, where N was determined in order to sample 150 counties.

[4] The original plan was to locate Google Earth base maps based on names of the villages and load the maps to mapper-PC before sending them to the field. However, because the Google Earth maps do not contain the boundaries of the villages, it is extremely difficult to know exactly how large an area to include in the base maps. Thus we decided to send the mapper/lister to the field with a GPS unit to get the boundary first and then extract the Google Earth map based on this boundary. This requires sending picture files back and forth many times between the field staff and the headquarters, with associated quality checks. Considering that some of the villages may not have Google Earth maps available for mapping, we made separate training

Advance teams verified that all buildings in the PSU had been properly identified, and that dwelling units within multi-dwelling buildings had been correctly coded to complete the frame. Once the sampling frame for a PSU was completed and entered into the computer, a CAPI (computer assisted personal interview) program was used to sample households and to conduct the interviews using laptops.

The number of households sampled was greater than the targeted sample size of 24 households per PSU in anticipation of non-response and sampled households' not having any members aged 45 or older.[5] We interviewed all age-eligible sample households in each PSU who were willing to participate in the survey, ultimately interviewing 10,257 households containing 17,587 respondents aged 45 and over and their spouses. The response rate was 80.5% of those households that were chosen to be sampled. This is much better than that of HRS-type surveys in the US and Europe, which now tends to be in the 60% or even 50% range, and compares favorably with other surveys done in Asia.

In this paper, we use data on all respondents 45 year of age and older,[6] some 17,343 respondents. Tables and figures are weighted using individual sample weights, adjusted for non-response.[7] All figures are nonparametric and drawn using LOWESS. Regressions are run unweighted.

In this paper, the data collected on self-reported health outcomes, and on biomarkers are used extensively. Specifically, our health measures include a general health measure; measures of disability (specifically activities of daily living [ADLs] and instrumental activities of daily living [IADLs]) and body pain; body mass index (BMI), underweight and overweight; hypertension and under-diagnosis of hypertension; and survival expectations to age 75 (for those aged 65 and under).[8] A high or low body mass index (BMI) is associated

---

materials for these situations. At the end, of the 450 village-level units, 379 (84.2%) had clear and usable Google Earth maps, 66 (14.7%) of the villages had maps which were illegible and 5 (1%) had no maps at all. In cases where maps were illegible or nonexistent, mappers drew maps using CHARLS-GIS software and every mapper/lister was trained to do this.

[5] We first sampled 80 households in each PSU and from that derived an estimate of age-eligibility and empty dwellings. From these, plus an assumed non-response rate, we re-sampled from the frame, allowing for both age-ineligibility (calculated at the PSU-level), empty dwellings and household non-response).

[6] Spouses who are under 45 years old are dropped from this analysis.

[7] The non-response adjustment is different for self-reported data than for biomarkers because the non-response for biomarkers is larger. See Zhao et al. (2013) for details.

[8] Heights were measured using a lightweight SECA aluminum height board, the SECA 214 portable stadiometer. Weights

with higher risk of subsequent mortality (Waaler 1984). Measures of functioning such as ADLs and IADLs have been found to be important health indicators of the elderly. Pain has a high impact on physical, psychological and social health (Elliott et al. 1999; Smith et al. 2001; Verhaak et al. 1998) and studies show an association between the report of widespread pain and subsequent death from cancer in the medium and long term (e.g. Macfarlane et. al 2001). The survival expectations of a person have been shown to be highly predictive of subsequent mortality (Banks et al. 2009), as have measures of general health. CHARLS also collected venous blood, which are being analyzed for several conditions, but those analyses have not yet been completed, so are not considered in this paper. In our regressions we start in column 1 with dummies for age group and education levels, a linear spline in log of household per capita expenditure, [9] location dummies for rural region and counties. In column 2, we replace rural and county dummies with community dummies. For the age and education dummies, being age 45–49 and having no schooling are the omitted groups respectively. In addition to life-cycle progression, these age dummies will also capture birth year cohort effects. With only a cross-section, we cannot distinguish the two.[10]

Education may proxy for many factors. Since we are controlling for an income measure, education will capture factors over and above income. These will include allocative efficiency effects, which may represent better-educated women having better information and their better understanding of what health inputs to choose to ensure good health (Schultz 1984). Of course education will also be correlated with preferences towards health perhaps in part due to more forward looking behavior (Fuchs 1982). Since past health (which would be endogenous) is correlated with current health and is an omitted variable in our analysis, it

---

were measured using a portable digital scale, the Beaver Tech HTS7270. Blood pressure was taken with a digital meter, the Omron HEM 712c meter.

[9] A linear spline allows different slopes to the left and right of the knot point with the two lines being joined at the knot point. The first coefficient reported is the slope to the left of the knot point and the second coefficient is the change in the slope between the two knot points.

[10] Cohort effects would arise because younger birth cohorts have more schooling and also faced better health conditions when they were babies and in the fetus, compared to older cohorts. There is an accumulation of evidence now that better health conditions when young are associated with better health in older age (for instance Barker 1994; Gluckman and Hanson 2005; Strauss and Thomas 2008, for an economist's perspective).

may be that past health “caused” in part education attainment, so that causation is going in both directions with cross-sectional data.

For household resources, as noted, we use log of household per capita expenditure (log PCE) for the household. This is a better measure of long-run resources than is current income, particularly so in low-income rural settings, where incomes can vary so much year to year because of variation in weather, pests, plant diseases, and so on. Per capita expenditure includes the value of food production which is self-consumed, which ought to be included in income, but may not be in all measures of income. Per capita expenditure also tends to be measured with less error than income (Deaton 1997; Lee 2009). Because income impacts may be highly nonlinear, even when PCE is in logs, we use a linear spline around the median log PCE.

Since location dummies are so important to our story, we discuss them in more detail. In our first specification we include rural and county dummies. In our second specification we use community dummies, using the PSU as our definition of community, which amounts to PSU, or community, fixed effects. As noted in the discussion of the sampling, our PSUs are either administrative villages or urban neighborhoods, and there are 459 of them in the sample.[11] These PSUs are small areas that are likely to be more homogeneous than cities or counties, or certainly than provinces. The idea here is that each community has factors that will affect health outcomes that are not captured by county dummies. These factors will include health care and other prices, inherent healthiness of the area, public health infrastructure and other factors. F-tests for all combinations of dummy variables are reported as well.[12]

Throughout this paper, we use ordinary least squares for continuous dependent variables and the linear probability model (LP) for binary dependent variables. LP model estimates are

[11] In a few cases there was no variation in the dependent variable among observations within the PSU. In these cases we refined the PSU to be a more aggregate area.

[12] We also investigate whether the SES variables vary in their coefficients by age group, distinguishing older (age 60 and above) from younger (age45-59) and by rural-urban residence. In general we do not find significant SES interactions. Where we do, we discuss this in the text.

consistent for estimating average partial effects of the regressors, which is our main interest. Robust standard errors of the regression coefficients are computed, that also allow for clustering at the county or community levels. By using robust standard errors for the linear probability regressions, we ensure that these standard error estimates are consistent (Wooldridge 2002).

Weighted means and standard errors for the variables used in the regressions are shown in Table A (see appendix), for men and women respectively. Mean respondent age is approximately 60 years, while most respondents are between 45 and 64. Women tend to be slightly younger on average because when they are chosen in the sample as a spouse, not a main respondent, they tend to be younger than their husbands. Education tends to be low, especially for women, for whom 39% have no schooling. For men, only 12% have no schooling, though 29% did not complete primary school. On the other hand, the fraction of men that did complete primary and junior high school and above is much higher than for women.

Results

General Health

We first examine self-reported general health. CHARLS followed the HRS example and asked respondents to assess their general health using two different scales: 1) excellent, very good, good, fair, poor, and 2) very good, good, fair, poor, very poor. Here we use the second scale and code it with a dummy variable indicating whether respondents report poor or very poor health. Table 1 displays the distribution of general health by age and sex group. About 23.1% of men and 29.7% of women report that they are in poor or very poor health. The fraction of women reporting poor health is larger than for men, as is common; our other health measures, both self-reports and biomarkers also indicate worse health for women, again commonly observed. For both men and women, the proportion reporting poor or very poor health increases with age. The fraction of respondents reporting fair health is quite high, about 48%. This is one reason why we do not combine fair and poor health as is often done in US studies. Apparently “fair” translates in Chinese to a word which is very commonly answered.

Table 2 presents SES regressions for reported general health for both men and women. The fraction reporting poor or very poor health rises with age for men and plateaus at age 70, while for women it rises and plateaus at age 65.These results could constitute either or both age or birth cohort effects. The fraction in poor or very poor health is negatively correlated with education and the effect is stronger for men than for women. Self-reported poor or very poor health is also negatively correlated with log PCE with a very non-linear relationship for both men and women. Higher PCE is associated with a lower likelihood of reporting poor health, until the median PCE, at which point the slope becomes close to zero.[13] Rural residence matters for women and is more highly associated with a probability of reporting poor or very poor health. The community fixed effects (and county dummies) are strongly jointly significant. It is also interesting to notice that the inequalities of health remains even after we control for community or county dummies and the pattern is consistent throughout the paper.

Disability and Body Pain

To make our study comparable with the mainstream of studies (e.g. Cutler et al. 2006; Schoeni et al. 2005), we define disability as the presence of any impairment in any of the activities of daily living (ADLs) or instrumental activities of daily living (IADLs). CHARLS includes information on 6 ADL measures: dressing, bathing, eating, walking, toileting, urination and defecations and 5 IADL measures: doing housework, shopping, cooking, managing money and taking medicine. Table 3 presents the fraction of disabled elderly by age, sex and urban/rural group. Overall, 18.8% of men and 26.5% of women suffer from disability and as is the case for our other health measures, there is a higher female disadvantage. The proportion rises monotonically with age, which may reflect a combination effect of age and birth cohort effects. The disability rates are smaller compared with the study of 1,348 men and women aged over 60 who were part of the Alumni Health Study in the U.S.

[13] Note the coefficients of log PCE for those above the median reported are marginal effects in addition to the coefficients for those below the median.

and who had been participating in a longitudinal study of risk factors for physical disability since 1986 (Murtagh and Hubert, 2004). Over the lifetime of this study (1986–1999), they find the disability rates of 52% for women and 37% for men. However, our results are similar to the research focused on the elderly population (ages sixty-five and older) using Phase 1 of the National Health Interview Disability Supplement of 1994 and 1995 (NHIS-D) in the U.S., in which 9.5% report ADL disability and 22.7% IADL disability (Cutler et al. 2006). Rural respondents are more likely to be disabled and subject to much larger female-male differences. For example, 11.1% of men and 18.2 of women aged 45-49 in rural areas are disabled and it is until 55 and over for men and 60 and over for women, that people in urban areas reach the equivalent proportions.

Another related dimension is measures of body pain. Pain has a high impact on physical, psychological and social health (Elliott et al. 1999; Smith et al. 2001; Verhaak et al. 1998) and studies show an association between the report of widespread pain and subsequent death from cancer in the medium and long term (e.g. Macfarlane et. al 2001). CHARLS respondents were asked about whether they were often troubled with any body pain and if so the level of pain (mild, moderate, or severe). On the whole, 24.9% of men and 36.5% of women suffer from body pain (shown in Table 4). Women are more likely to report body pain than men. Rural women are the most vulnerable group and urban men the group least likely to report any pain. Even the youngest rural women group, 45-49, have a reported pain rate higher than that of the oldest male group of age 70 and above (35.4% vs. 32.4%). Women are not only more prone to feel body pain but also feel pain of higher severity. Table 5 shows the fraction of three levels of pain (mild, moderate and severe) for each age and sex group. 27.0% of women vs. 17.6% of men report moderate or severe body pain. For men, as age goes up, the fraction of each level of pain rises and that reporting moderate or severe pain rises more sharply than mild pain. For women, the fraction of moderate or severe pain increases even more dramatically than for men and the fraction of mild pain goes down as age increases, which indicates that for women, the level of pain exacerbates with age.

Regressions for any ADL or IADL disability are shown in Table 6. The disability rate is positively correlated with age. There is an obvious negative gradient in education and log PCE in all specifications. The log PCE effect is again nonlinear for both men and women, being negatively correlated with reported pain for log PCE below the median and flattening out to near zero for those above the median. Rural residence matters only for women, positively. The community fixed effects (and county dummies) are strongly jointly significant, as is true for our other regressions.

Table 7 reports the analysis of moderate or severe pain for men and women. For men there is not a strong age gradient, unlike for women. For women, the age effect flattens out about age 65, after a slight tick up at ages 60-64. For education, only higher levels of education, at junior high school and above, is associated with reporting less moderate or severe pain, although the education variables are only jointly significant for men. The same nonlinear relationship we saw for log PCE in the previous results are repeated here. Higher log PCE is strongly negatively associated with moderate or severe pain, for both men and women, at levels of log PCE below the median, and then the curve flattens. Both men and women in rural areas report more serious pain than their urban counterparts, though the interaction is much larger for older age groups.[14] The community fixed effects (and county dummies) continue to show strong joint significance.

BMI

BMI is measured as weight (in kg) divided by height squared (in meters). Extreme values of BMI may be related to hypertension, diabetes, and in general to higher adult mortality (Waaler 1984). Across countries, the BMI distribution is shifted to the right for countries with higher incomes. Figure 1, which is reproduced from Strauss and Thomas (2008), demonstrates this, showing nonparametric relationships between BMI and years of schooling, for men and women aged 25–70 from six countries, ranging in GDP from Bangladesh to the United States.[15] China as of 1991 is included among these six countries

[14] The interaction results are available upon request.

[15] The sources are the Matlab Health and Socioeconomic Survey (1996); the China Health and Nutrition Survey (1991); the Indonesia Family Life Survey (2000); the South African Demographic Health Survey (1998); the Mexican Family Life Survey (2002); and the NHANES3 (National Health and Nutrition Examination Survey III) (1988–1994).

and is closest in its BMI distribution to Indonesia. Note that for men, except in the US, BMI rises with more education. For women the story is quite different. Again, the US excepting and also for low levels of schooling for women, BMI rises with education, but at higher levels, it falls. Bangladesh is an exception; BMI rises with female schooling, probably because women are still so close to levels of undernutrition. In the other developing countries, including China, the inverted U-shape relation is apparent. It may be that at higher levels of female schooling, women recognize the health benefits of reducing their BMI. Why this is not the case for men is a key question for future research.

Table 8 shows mean BMI by age and sex group, as well as the fraction underweight, overweight, and obesity in each group. Overall, 38.1% of women are overweight compared to 28.1% of men. The fact that a higher fraction of women are overweight compared to men is a common result found in many other countries (see Strauss and Thomas 2008). BMI tends to fall with age, although again one has to be careful because with a cross-section it is not possible to distinguish age from birth cohort effects. The proportions that are overweight are higher than the elderly in Indonesia (see Witoelar, Strauss and Sikoki, 2012). Women are not only more prone to be overweight but also obese. 8.6% of women suffer from obesity compared to 3.4% of men. Note too that underweight is still a problem, and particularly so for the very elderly, those 75 years and over, for whom approximately 17% are underweight. These proportions are just under the proportions that are overweight for these older ages, unlike for younger ages, for whom the proportions overweight are far larger than for underweight.

The regressions examine the extreme values of BMI: underweight and overweight. Table 9 gives the analysis of underweight for men and women, respectively. Men over 60 and women over 55 are more likely to be underweight and the likelihood tends to increase with age/birth cohort. Education has no association with underweight for women, but does for men, particularly schooling at junior high school or above. For both genders, log PCE is negatively correlated with underweight for both genders for those whose log PCE is below the median, but the relationship flattens and is insignificant for those whose log PCE is above the median.

Rural residence matters only for women and rural women are more likely to be underweight. Rural and province dummies are significant for both men and women (see Strauss et al. 2010), as are the community dummies in the community fixed effect models.

Regressions for overweight are presented in Table 10. For women, the chance of overweight falls with age/cohort while for men, the age/cohort effect presents a U-pattern. Education has no association with being overweight for women. However, men with junior high school and above show a lower probability of overweight and the education dummies are jointly significant. The effect of log PCE is positively correlated with overweight for people whose log PCE is under the median level. For men, the relationship flattens out and is not significant for those above the median. For women, however, the relationship becomes significantly negative at levels of log PCE above the median.[16] For both genders, those living in rural areas is negatively correlated with being overweight. The community fixed effects are jointly significant.

Hypertension and Its Under-diagnosis

Along with BMI, blood pressure is an indicator of risk of coronary heart disease. Respondents were measured three times for blood pressure in the survey. We drop the first reading because respondents may have been more nervous the first time and take the mean of last two systolic and diastolic measurements separately. We then form a variable for being hypertensive if the mean systolic is 140 or greater or the mean diastolic is 90 or greater. These are the conventional cutoffs for high blood pressure, or hypertension. Also persons who report that they have been diagnosed with hypertension by a doctor are classified as hypertensive, including those who take medications for hypertension.[17] Since we combine biomarker measurements and self-reported hypertension, our sample here exclude respondents that didn't participate in the physical examination, but in the descriptive tables sample weights include an adjustment for non-response. As shown in Table 11, 41.0% of men

[16] The results with level coefficients on the spline are available upon request.

[17] Because of the way the questionnaire is designed, those who report taking medicines are a subset of those who report a positive doctor diagnosis.

and 45.2% of women in China are hypertensive. The hypertension rate increases with age for both genders, but especially so for women over age 64. For women older than 75, the fraction being hypertensive reaches up to 70%. Women are more likely to be hypertensive than men starting at age 55 and urban respondents tend to have a higher hypertension rate.

Table 12 displays the regressions predicting being hypertensive for men and women respectively. Hypertension is strongly increasing with age for both genders. Interestingly, the rural residence coefficient is significantly negative for both men and women, apparently older rural Chinese are less likely to suffer from hypertension than the urban elderly. SES gradients exist only for men and then only for schooling. Men finishing primary school and above are more likely to have hypertension. As always, the community fixed effects are jointly significant.

An important policy issue that emerges from the health transition that China has been going through is that chronic diseases tend to be under-diagnosed, at least during this transition period. Other countries that are undergoing the health transition seem to be experiencing a similar phenomenon (e.g. Witoelar, Strauss and Sikoki, 2012, and Parker, Teruel and Rubalcava, 2010, for Indonesia and Mexico respectively). Hypertension turns out to be a good example of the degree of under-diagnosis of disease among the elderly in China. We calculate the proportion of hypertensive people who report not being diagnosed, shown in Table 11. Among those who have hypertension, 42.8% of men and 40.6% of women are under-diagnosed. This seems quite large, although estimates for Indonesia are much higher, 74% for men and 62% for women (Witoelar, Strauss and Sikoki, 2012) and estimates for Mexico from the Mexican Family Life Survey are also high (see Parker, Teruel and Rubalcava, 2010). One interpretation is that the health system in China is not yet set up to focus on chronic conditions of the elderly, perhaps because the emphasis is on infectious disease and on children and mothers. Additional research will be required to examine this issue more properly. The under-diagnosis rate of the youngest and the oldest groups is higher and the middle group is lower. Men in rural areas are most likely to be under-diagnosed (45.4%).

In addition to undiagnosed disease, another key health issue is good adherence to treatment when the disease is diagnosed (Goldman and Smith 2002). Table 11 also shows that 78.4% of men and 82.2% of women, who have hypertension by our definition, and report having been diagnosed, are taking medications. Thus, conditional on being diagnosed a preponderance of respondents is taking medications. However, those who are undiagnosed are not.

In Table 13, using the sample of men and women who have measured or self-reported diagnosis of hypertension, we regress a dummy of not being diagnosed on the same set of covariates used in the other regressions for urban and rural subsamples respectively.[18] Age matters except for men in urban areas. Rural men over 55, rural women aged 55-69, and urban women aged 50-69 are less likely to be under-diagnosed. Education is significantly associated with under-diagnosis only in rural areas and only for men. For rural men, having junior high school or more education is negatively related to under-diagnosis and the effect of being able to read and write is also negative and weakly significant (at the 10% level). For rural women, education of primary school and above leads to less under-diagnosis in the case of community fixed-effects included. However, if we look at the joint significance of education dummies, it's only significant for rural men. The effect of log PCE is only significant at 10% for men and 5% for women, and negatively correlated with being under-diagnosed, in urban areas for both genders. Yet again these relationships are nonlinear for both men and women, flattening out for those with log PCE above the median. The community fixed effects are jointly significant except for the urban women subsample. Hence there are important community differences in the degree of under-diagnosis of hypertension. The SES differentials are not so large, but some do exist, particularly with respect to log PCE. This is a health policy area where improvements seem possible, although more research is needed to fill in more details.

[18] We also do regressions for men and women separately without dividing the sample by urban and rural like the analyses of other health measures, which are available unpon request. The reason for using the urban and rural subsamples is to see whether education and log PCE work differently by gender or by region.

Survival Expectations

In CHARLS, as in the HRS, we ask respondents about the chances that they expect to live to a particular age. Respondents answer the question on a five point scale, from 1 "almost impossible", to 5 "almost certain".[19] Experience with HRS and other ageing surveys have shown that answers to this question are highly correlated with survival to subsequent waves (for example see Banks et al. 2009). The future age about which each respondent is asked depends on their current age. Older respondents are asked about survival to older ages. This raises an issue in that answers across respondents who were asked about different ages may not therefore correspond. Here we take respondents under age 65, all of whom were asked about survival to age 75, so that this issue does not arise. We construct our dependent variable to indicate whether the respondent thinks it is not very likely, or almost impossible, to reach age 75, the two lowest scores.

Table 14 gives the distribution of the chances to reach 75 for each age and sex group. Women of all ages are more pessimistic than men. At the other end of the distribution, men are more likely than women to be almost certain that they will reach age 75.

Regressions are reported in Table 15. Men of older cohorts feel more confident to survive to 75 years old, which may be a cohort effect or an age effect; though, age doesn't matter for women. Education is associated with a higher level of confidence of reaching age 75 for both men and women. For men, education of junior high school and above leads to better expectations. For women, there's a distinct education gradient in survival expectations, starting with "can read or write" (some primary school). Expectations also improve with higher log PCE for both men and women, but yet again, flatten out for those above median log PCE. The location effects are strongly jointly significant although the rural dummy is not significant. Presumably these SES gradients result from the fact that real health variables which may be related to likely future mortality have SES gradients, as we have seen.

---

[19] We do not ask probabilities directly since our pretest experience, and experience in other low- income countries indicated a real difficulty for respondents to understand probabilities.

Conclusions

This paper has presented estimates of the health-SES gradient in China using multiple measures of health and of SES. China has undergone a significant health and nutrition transition such that under-nutrition is much less of a problem for the elderly than it had been in the past and overnutrition has become much more of an issue. In China, where the CHARLS baseline was fielded, health conditions of the elderly, such as better general health, less disability, less body pain, and positive survival expectations are all correlated positively with education for both genders while being overweight, hypertension and being under-diagnosed is correlated only for men; with better education being associated with better health outcomes except for hypertension. Evidence for correlation between being underweight and education is not found in this paper. The correlations of health measures with income, as measured by per capita expenditure, are strongly significant for all health outcomes for both genders with one exceptions, hypertension. In virtually all of our cases, log PCE is positively correlated with better health measures, for respondents with log PCE less than the sample median, and is flat for those above the median.

One of the most important findings in this analysis is the apparent importance of community factors. What exactly lies behind this is not yet clear and needs to be the subject of future research. From economic theory there are a number of factors that should be part of the story. Prices of health inputs and of other commodities is surely one such factor, as should be the availability and quality of health care services. Public health infrastructure should be another such factor, as should the inherent healthiness of a community due to factors like water, sanitation and air quality. Different and changing food or diet preferences are also no doubt related to these findings. Given the strength of the relationships, however, it may well be that there are other community influences that are important, including factors that related to social interaction and stress, that are particularly important in China. However at this point, all of these hypotheses represent speculation.

The other important finding in this research is the large-scale under-diagnosis of hypertension, which is correlated negatively with education in rural areas and with PCE in urban areas for women. It is most strongly correlated with community location for all the people except for urban women. This represents a major health system gap and one which is probably more serious for other, less prevalent, chronic conditions of the elderly. This problem is certainly not unique to China and seems to exist in other countries that are still in the midst of the health transition from infectious to chronic diseases. Health systems in such health transition countries apparently take time to re-orient their systems to diagnose and treat chronic diseases of the ageing and aged. This is an important step that the health systems in China will need to work out in the future.

References

Banks, J., Marmot, M., Oldfield, Z., & Smith, J. P. (2006). Disease and disadvantage in the United States and in England. Journal of the American Medical Association, 295(17), 2037–2045.

Banks, J., Muriel, A., &Smith, J. P. (2009). Disease prevalence, incidence and determinants of mortality in the United States and England, manuscript, Department of Economics, University College London.

Barker, D. (1994). Mothers, babies and health in later life. London: BMJ Publishing Group.

Barker, D. (1997). Maternal nutrition, fetal nutrition and diseases in later life. Nutrition, 13(9), 807–813.

Cutler, D., Landrum, M., & Stewart, K. (2006). How do the better educated do it? Socioeconomic status and the ability to cope with underlying impairment (No. w12040). National Bureau of Economic Research.

Cutler, D., & Lleras-Muney, A. (2010). Understanding differences in health behaviors by Education. Journal of Health Economics, 29(1), 1–28.

Das, J., Quy-Toan, D., Friedman, J., McKenzie, D., & Scott, K. (2007). Mental health and poverty in developing countries. Social Science and Medicine, 65(3), 467–480.

Deaton, A. (1997). The analysis of household surveys: A microeconometric approach. Baltimore: Johns Hopkins University Press.

Elliott, A., Smith, B., Penny, K., Smith, W., & Chambers, W. (1999). The epidemiology of chronic pain in the community. Lancet, 354(9186), 1248.

Friedman, J., & Thomas, D. (2008). Psychological health, before, during and after an economic crisis: results from Indonesia 1993–2000. World Bank Economic Review, 23(1), 57–76.

Fuchs, V. (1982). Time preference and health: an exploratory study. In V. Fuchs (Ed.), Economic aspects of health. Chicago: National Bureau of Economic Research.

Gluckman, P., & Hanson, M. (2005). The fetal matrix: Evolution, development and disease. Cambridge.

Goldman, D., & Smith, J. P. (2002). Can patient self-management help explain the SES health gradient? Proceedings of the National Academy of Sciences, 99(16), 10929–11093.

Hoddinott, J., Maluccio, J., Behrman, J., Flores, R., & Martorell, R. (2008). Effect of a nutrition intervention during early childhood on economic productivity in Guatemalan adults. Lancet, 371, 411–416.

Hossain, S. (1997). Tackling health transition in China, Policy Research Working Paper No. 1813, Washington DC: World Bank.

Kenkel, D., Lillard, D., & Liu, F. (2009). An analysis of life-course smoking behavior in China. Health Economics, 18, S147–S156.

Kinsella, K., & He, W. (2009). An aging world: 2008, US Census Bureau, International Population Reports, PS95/09-1. Washington DC: US Government Printing Office.

Lance, P., Akin, J., Dow, W., & Loh, C. (2004). Is cigarette smoking in poorer Nations

sensitive to price? Evidence from Russia and China. Journal of Health Economics, 23, 173–189.
Lee, N. (2009). Measurement error and its impact on estimates of income and consumption dynamics, mimeo. Department of Economics: Chinese University of Hong Kong.
Lee, J., & Smith, J. P. (2008). Explanations for education gradients in depression—The case of Korea, mimeo, RAND Corporation, Santa Monica, CA.
Lleras-Muney, A. (2005). The relationship between education and adult mortality in the US. Review of Economic Studies, 72(1), 189–221.
Lopez, A., Mathers, C., Ezzati, M., Jamison, D., & Murray, C. (2006). Measuring the global burden of disease and risk factors, 1990–2001. In A. Lopez, C. Mathers, M. Ezzati, D. Jamison, & C. Murray (Eds.), Global burden of disease and risk factors. Oxford: Oxford University Press.
Luo, Z. (2003). Socioeconomic determinants of body mass index of adults Chinese in 1990s, presented at the Northeast Consortium Development Conference (NEUDC), October 2003, mimeo, Department of Epidemiology, Michigan State University.
Macfarlane, G. J., McBeth, J., Silman, A. J., & Crombie, I. K. (2001). Widespread body pain and mortality: prospective population based studyCommentary: An interesting finding, but what does it mean?. Bmj, 323(7314), 662.
Maluccio, J., Hoddinott, J., Behrman, J., Martorell, R., Quisumbing, A., & Stein, A. (2009). The impact of improving nutrition during early childhood on education among Guatemalan adults. Economic Journal, 119(537), 734–763.
Marmot, M. G. (1999). Multi-level approaches to understanding social determinants. In L. Berkman & I. Kawachi (Eds.), Social epidemiology (pp. 349–367). Oxford: Oxford University Press.
McArdle, J., Fisher, G., & Kadlec, K. (2007). Latent variable analysis of age trends in tests of cognitive ability in the Health and Retirement Survey, 1992–2004. Psychology and Aging, 22(3), 525–545.
McArdle, J., Smith, J. P., & Willis, R. (2009). Cognition and economic outcomes in the Health and Retirement Survey, manuscript, RAND Corporation, Santa Monica, CA.
Murtagh, K., & Hubert, H. (2004). Gender differences in physical disability among an elderly cohort. American Journal of Public Health, 94(8), 1406.
Parker, S., Teruel, G., & Rubalcava L. (2010). Perceptions and knowledge of underlying health conditions in Mexico, paper presented at the Population Association of America Annual Meetings, Dallas, Texas, 2010.
Patel, V., & Kleinman, A. (2003). Poverty and common mental disorders in developing countries. Bulletin of the World Health Organization, 81(8), 609–615.
Popkin, B. (1999). Urbanization, lifestyle changes and the nutrition transition. World Development, 27 (11), 1905–1916.
Popkin, B., Ge, K., Zhai, F., Guo, X., Ma, H., & Zohoori, N. (1993). The nutrition transition in China: a cross-sectional analysis. European Journal of Clinical Nutrition, 47, 333–346.
Popkin, B. (2002). The shift in stages of the nutrition transition in the developing world differs from past experiences. Public Health Nutrition, 5(1A), 205–214.

Popkin, B., Paeratakul, S., Zhai, F., & Ge, K. (1995a). A review of dietary and environmental correlates of obesity with emphasis on developing countries. Obesity Research, 3(S2), S145–S153.

Popkin, B., Paeratakul, S., Zhai, F., & Ge, K. (1995b). Dietary and environmental correlates of obesity in a population study of China. Obesity Research, 3(S2), S135–S143.

Schoeni, R. , Martin, L., Andreski, P., & Freedman, V. (2005). Persistent and growing socioeconomic disparities in disability among the elderly: 1982-2002. American Journal of Public Health, 95(11), 2065.

Schultz, T. P. (1984). Studying the impact of household economic and community variables on child mortality. In W. H. Mosely, & L. Chen (Eds.), Child survival: Strategies for research, population and development review, 10(Supplement):215–236.

Seeman, T., Burt Singer, J., Rowe, R. H., & McEwen, B. (1997). Price of adaption-allostatic load and its health consequences. Archives of Internal Medicine, 157(19), 2259–2268.

Smith, B. H., Elliott, A. M., Chambers, W. A., Smith, W. C., Hannaford, P. C., & Penny, K. (2001). The impact of chronic pain in the community. Family practice, 18(3), 292-299.

Smith, J. P. (1999). Healthy bodies and thick wallets: the dual relation between health and economic status. Journal of Economic Perspectives, 13(2), 145–167.

Smith, J. P. (2009). The impact of childhood health on adult labor market outcomes. Review of Economics and Statistics, 91(3), 478–489.

Strauss, J. (1986). Does better nutrition raise farm productivity? Journal of Political Economy, 94(2), 297–320.

Strauss, J. (1993). The impact of improved nutrition on labor productivity and human resource development: An economic perspective. In P. Pinstrup-Andersen (Ed.), The political economy of food and nutrition policies. Baltimore: Johns Hopkins.

Strauss, J., & Thomas, D. (1995). Human resources: Empirical modeling of household and family decisions. In J. R. Behrman & T. N. Srinivasan (Eds.), Handbook of development economics (Vol. 3A). Amsterdam: North Holland.

Strauss, J., & Thomas, D. (1998). Health, nutrition and economic development. Journal of Economic Literature, 36(3), 766–817.

Strauss, J., & Thomas, D. (2008). Health over the life course. In T. P. Schultz & J. Strauss (Eds.), Handbook of development economics (Vol. 4). Amsterdam: North Holland.

Strauss, J., Lei, X., Park, A., Shen, Y., Smith, J. P., Yang, Z., et al. (2010). Health outcomes and socioeconomic status among the elderly in China: Evidence from the CHARLS pilot, Journal of Population Ageing, 3(3):111-142.

Thomas, D. (2010). Health and socio-economic status: The importance of causal pathways. In J. Y. Lin & B. Pleskovic (Eds.), Annual World Bank Conference on Development Economics, Global 2009 (pp. 355–384). Washington DC: World Bank.

Thomas, D., & Strauss, J. (1997). Health and wages: evidence for men and women in urban Brazil. Journal of Econometrics, 77(1), 159–186.

Thomas, D., Strauss, J., & Henriques, M.-H. (1991). How does mother's education affect child height? Journal of Human Resources, 26(2), 183–211.

UNICEF. (2012). China statistics, www.unicef.org/infobycountry/china_statistics.html.

Verhaak, P. F., Kerssens, J. J., Dekker, J., Sorbi, M. J., & Bensing, J. (1998). Prevalence of chronic benign pain disorder among adults. Pain, 77, 231-239.

Waaler, H. (1984). Height, weight and mortality: the Norwegian experience. Acta Medica Scaninavica, 215(S679)), 1–56.

Wagstaff, A., Yip, W., Lindelow, M., & Hsiao, W. (2009). China's health system and its reform: a review of recent studies. Health Economics, 18, S7–S23.

Wilkenson, R. G. (1996). Unhealthy societies: The afflictions of inequality. London: Routledge.

Witoelar, F., Strauss, J., & Sikoki, B. (2012). Socioeconomic success and health in later life: Evidence from the Indonesia family life survey, in M. Majmundar and J.P. Smith (eds.), *Aging in Asia: Findings from New and Emerging Data Initiatives*, Committee on Policy Research and Data Needs to Meet the Challenge of Aging in Asia, National Research Council. Washington D.C.: National Academies Press.

Wooldridge, J. (2002). Econometric analysis of cross-section and panel data. Cambridge: MIT Press.

World Health Organization. (2012). China, WHO Country Health Information Profiles.

Zhao, Y., Strauss, J., Yang, G., Giles, J., Hu, P., Hu, Y., Lei, X., Park, A., Smith, J. P. & Wang, Y. (2013). China health and retirement longitudinal study – 2011-2012 national baseline user's guide, China Center for Economic Research, Peking University.

Table 1. Self-reported health, by age and sex

	Men						Women					
	Very good	**good**	**fair**	**poor**	**very poor**	**N**	**Very good**	**good**	**fair**	**poor**	**very poor**	**N**
45-49	11.4 (1.4)	24.4 (1.8)	50.7 (2.4)	11.2 (1.1)	2.2 (0.4)	1553	7.1 (0.8)	20.3 (1.7)	51.9 (1.8)	18.4 (1.6)	2.3 (0.4)	1893
50-54	11.1 (1.7)	23.9 (1.9)	48.0 (2.4)	14.9 (1.2)	2.1 (0.4)	1250	7.6 (1.7)	15.6 (1.4)	52.0 (2.2)	21.3 (1.5)	3.5 (0.5)	1310
55-59	10.1 (1.8)	21.9 (1.4)	45.7 (1.7)	19.0 (1.3)	3.4 (0.5)	1736	6.5 (1.6)	18.6 (2.5)	46.7 (2.0)	24.0 (1.4)	4.2 (0.5)	1802
60-64	6.6 (0.7)	17.4 (1.3)	48.5 (1.7)	22.9 (1.7)	4.6 (0.7)	1439	4.7 (0.7)	14.9 (1.2)	48.3 (1.7)	25.6 (1.3)	6.4 (0.8)	1448
65-69	5.9 (0.9)	16.7 (1.4)	52.9 (2.5)	21.0 (1.7)	3.5 (0.6)	950	3.8 (0.8)	10.0 (1.2)	46.2 (2.2)	33.6 (2.4)	6.4 (0.9)	892
70-74	6.3 (1.4)	12.8 (1.7)	48.8 (2.7)	25.3 (2.1)	6.8 (1.2)	713	4.9 (1.2)	15.0 (1.7)	44.7 (2.4)	28.5 (2.1)	6.9 (1.2)	655
75+	3.2 (0.8)	14.7 (1.7)	44.6 (2.7)	29.2 (2.7)	8.3 (1.5)	730	4.6 (0.9)	14.8 (1.5)	39.7 (2.2)	31.5 (2.2)	9.3 (1.3)	821
Total (45+)	8.5 (0.6)	19.9 (0.8)	48.4 (0.9)	19.2 (0.6)	3.9 (0.3)	8371	5.9 (0.5)	16.5 (0.8)	47.9 (0.9)	24.7 (0.8)	5.0 (0.4)	8821

Weighted at individual level with household and response adjustment. Sample are respondents not younger than 45.Standard errors in parentheses.

Table 2. Regressions for self-reported health poor or very poor

	Men		Women	
	(1)	(2)	(1)	(2)
Aged 50-54	0.037**	0.040***	0.031*	0.025
	(0.016)	(0.014)	(0.016)	(0.015)
Aged 55-59	0.054***	0.058***	0.065***	0.066***
	(0.015)	(0.015)	(0.015)	(0.014)
Aged 60-64	0.080***	0.081***	0.076***	0.077***
	(0.016)	(0.015)	(0.016)	(0.015)
Aged 65-69	0.079***	0.078***	0.138***	0.132***
	(0.017)	(0.019)	(0.019)	(0.018)
Aged 70-74	0.154***	0.153***	0.124***	0.123***
	(0.020)	(0.022)	(0.021)	(0.021)
Aged 75 and over	0.140***	0.138***	0.135***	0.135***
	(0.020)	(0.020)	(0.021)	(0.023)
Can read and write	-0.043**	-0.045**	0.005	0.005
	(0.017)	(0.018)	(0.014)	(0.015)
Finished primary	-0.065***	-0.065***	-0.026*	-0.024
	(0.016)	(0.018)	(0.015)	(0.015)
Junior high and above	-0.101***	-0.101***	-0.064***	-0.055***
	(0.017)	(0.018)	(0.015)	(0.015)
logPCE (< median)	-0.052***	-0.058***	-0.041***	-0.040***
	(0.008)	(0.009)	(0.008)	(0.010)
logPCE (> median, marginal)	0.070***	0.080***	0.049***	0.050***
	(0.011)	(0.013)	(0.011)	(0.014)
Rural	0.019		0.037***	
	(0.013)		(0.014)	
County Dummies	YES	NO	YES	NO
Community Dummies	NO	YES	NO	YES
F-test for all age dummies	14.11***	13.55***	14.16***	14.34***
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)
F-test for all education dummies	13.53***	10.95***	7.43***	6.37***
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)
F-test for all logPCE splines	22.55***	20.25***	12.82***	7.81***
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)
F-test for all location dummies	2.73***	1.94***	3.73***	2.55***
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	8291	8291	8728	8728

Standard error in parentheses, all clustered at community level. * p<.1 ** p<.05 *** p<.01. logPCE (>median, marginal) represents the change in the slope from the interval for logPCE blow the median, see footnote 7. Sample are respondents not younger than 45.

Table 3. Disability (any ADL or IADL), by age, sex, and urban/rural

	Rural				Urban				Total			
	Men		Women		Men		Women		Men		Women	
	%	N	%	N	%	N	%	N	%	N	%	N
45-49	11.1 (1.3)	905	18.2 (1.5)	1089	6.4 (1.2)	647	8.1 (1.4)	802	8.7 (0.9)	1552	12.8 (1.1)	1891
50-54	13.6 (1.4)	754	23.7 (1.7)	764	7.6 (1.6)	493	11.3 (1.6)	544	10.7 (1.1)	1247	17.3 (1.3)	1308
55-59	19.7 (1.5)	1055	27.9 (1.8)	1069	12.5 (1.8)	677	15.3 (1.7)	728	16.2 (1.3)	1732	21.4 (1.4)	1797
60-64	19.6 (1.6)	882	34.5 (2.0)	868	12.8 (1.7)	556	19.5 (2.1)	579	16.4 (1.2)	1438	27.4 (1.6)	1447
65-69	27.2 (1.9)	610	41.0 (2.5)	557	17.9 (2.9)	341	28.0 (4.8)	335	23.2 (1.8)	951	35.0 (2.5)	892
70-74	33.8 (2.5)	444	47.7 (2.9)	366	22.5 (3.9)	268	25.2 (3.1)	292	28.3 (2.3)	712	35.6 (2.3)	658
75+	50.2 (3.2)	429	66.3 (2.5)	466	44.1 (4.9)	298	53.4 (3.4)	351	47.1 (2.9)	727	60.1 (2.1)	817
Total (45+)	22.2 (1.0)	5079	33.6 (1.2)	5179	15.2 (0.9)	3280	19.6 (1.1)	3631	18.8 (0.7)	8359	26.5 (0.9)	8810

Weighted at individual level with household and response adjustment. Sample are respondents not younger than 45.Standard errors in parentheses.

Table 4. Any body pain, by age, sex, and rural/urban

	Rural				Urban				Total			
	Men		Women		Men		Women		Men		Women	
	%	N	%	N	%	N	%	N	%	N	%	N
45-49	22.9	903	35.4	1088	14.7	647	29.6	804	18.6	1550	32.3	1892
	(1.8)		(2.0)		(2.0)		(2.9)		(1.5)		(1.8)	
50-54	28.3	754	41.2	763	18.9	495	33.3	544	23.7	1249	37.1	1307
	(2.0)		(2.3)		(2.6)		(4.5)		(1.8)		(2.5)	
55-59	31.3	1055	42.4	1071	20.5	678	27.1	730	26.0	1733	34.5	1801
	(1.9)		(2.1)		(4.1)		(2.5)		(2.1)		(1.8)	
60-64	31.7	883	46.9	865	22.6	555	31.6	578	27.5	1438	39.6	1443
	(1.8)		(2.0)		(3.3)		(3.6)		(1.8)		(2.0)	
65-69	33.6	610	49.2	556	18.2	341	36.4	335	27.0	951	43.3	891
	(2.4)		(2.7)		(2.8)		(4.6)		(2.0)		(2.5)	
70-74	32.3	444	47.7	365	18.8	270	28.0	290	25.7	714	37.1	655
	(2.7)		(3.0)		(3.6)		(3.3)		(2.3)		(2.4)	
75+	32.4	428	42.5	464	27.2	296	31.7	347	29.7	724	37.3	811
	(2.8)		(2.9)		(4.7)		(3.3)		(2.7)		(2.2)	
Total	29.8	5077	42.5	5172	19.7	3282	30.6	3628	24.9	8359	36.5	8800
(45+)	(1.2)		(1.4)		(1.2)		(1.5)		(0.9)		(1.0)	

Weighted at individual level with household and response adjustment. Sample are respondents not younger than 45.Standard errors in parentheses.

Table 5. Severity of body pain, by age and sex

	Men				Women			
	Mild	**Moderate**	**Severe**	**N**	**Mild**	**Moderate**	**Severe**	**N**
45-49	5.5	5.9	7.0	1548	11.4	10.8	10.0	1891
	(0.7)	(0.7)	(0.8)		(1.8)	(1.0)	(1.0)	
50-54	7.0	8.9	7.6	1246	10.0	16.0	10.8	1303
	(0.9)	(1.0)	(0.9)		(1.0)	(2.7)	(1.0)	
55-59	7.8	9.0	9.2	1733	8.0	12.3	14.1	1800
	(1.2)	(0.9)	(1.2)		(0.7)	(0.9)	(1.1)	
60-64	6.7	11.0	9.6	1436	7.8	16.6	15.0	1440
	(0.8)	(1.7)	(0.9)		(0.8)	(1.8)	(1.2)	
65-69	7.7	9.1	10.2	951	13.5	14.6	14.8	888
	(1.1)	(1.1)	(1.2)		(2.6)	(1.4)	(1.6)	
70-74	6.9	8.4	10.0	712	9.8	12.9	14.3	654
	(1.2)	(1.2)	(1.6)		(1.3)	(1.6)	(1.5)	
75+	9.4	9.9	10.1	722	5.9	15.2	16.1	811
	(2.6)	(1.4)	(1.3)		(1.0)	(1.5)	(1.6)	
Total	7.1	8.7	8.9	8348	9.4	13.8	13.2	8787
(45+)	(0.4)	(0.5)	(0.5)		(0.5)	(0.6)	(0.7)	

Weighted at individual level with household and response adjustment. Sample are respondents not younger than 45.Standard errors in parentheses.

Table 6. Regressions for disability (any ADL or IADL)

	Men		Women	
	(1)	(2)	(1)	(2)
Aged 50-54	0.009	0.014	0.040***	0.039***
	(0.014)	(0.011)	(0.015)	(0.013)
Aged 55-59	0.037***	0.038***	0.059***	0.064***
	(0.013)	(0.011)	(0.014)	(0.014)
Aged 60-64	0.033**	0.036***	0.099***	0.097***
	(0.014)	(0.013)	(0.015)	(0.013)
Aged 65-69	0.109***	0.113***	0.158***	0.157***
	(0.016)	(0.016)	(0.018)	(0.019)
Aged 70-74	0.146***	0.150***	0.188***	0.185***
	(0.018)	(0.020)	(0.020)	(0.022)
Aged 75 and over	0.259***	0.265***	0.344***	0.346***
	(0.018)	(0.022)	(0.020)	(0.021)
Can read and write	-0.059***	-0.059***	-0.039***	-0.038**
	(0.015)	(0.020)	(0.013)	(0.016)
Finished primary	-0.109***	-0.105***	-0.101***	-0.100***
	(0.015)	(0.020)	(0.014)	(0.016)
Junior high and above	-0.137***	-0.133***	-0.128***	-0.123***
	(0.015)	(0.020)	(0.015)	(0.017)
logPCE (< median)	-0.057***	-0.064***	-0.028***	-0.027***
	(0.007)	(0.009)	(0.008)	(0.009)
logPCE (> median, marginal)	0.075***	0.087***	0.037***	0.039***
	(0.010)	(0.013)	(0.011)	(0.012)
Rural	0.012		0.036***	
	(0.012)		(0.013)	
County Dummies	YES	NO	YES	NO
Community Dummies	NO	YES	NO	YES
F-test for all age dummies	46.13***	35.79***	63.46***	50.53***
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)
F-test for all education dummies	31.36***	19.46***	29.98***	19.47***
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)
F-test for all logPCE splines	32.03***	27.62***	6.76***	4.92***
(p-value)	(0.000)	(0.000)	(0.001)	(0.008)
F-test for all location dummies	3.68***	2.47***	4.67***	3.08***
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	8284	8284	8725	8725

Standard error in parentheses, all clustered at community level. * p<.1 ** p<.05 *** p<.01. logPCE (>median, marginal) represents the change in the slope from the interval for logPCE blow the median, see footnote 7. Sample are respondents not younger than 45.

Table 7. Regressions for moderate or severe body pain

	Men		Women	
	(1)	(2)	(1)	(2)
Aged 50-54	0.033**	0.034**	0.038**	0.033*
	(0.015)	(0.014)	(0.016)	(0.017)
Aged 55-59	0.019	0.021	0.052***	0.050***
	(0.014)	(0.014)	(0.015)	(0.015)
Aged 60-64	0.020	0.022	0.075***	0.068***
	(0.015)	(0.014)	(0.016)	(0.016)
Aged 65-69	0.020	0.019	0.060***	0.055***
	(0.016)	(0.017)	(0.018)	(0.020)
Aged 70-74	-0.004	-0.009	0.049**	0.045**
	(0.018)	(0.017)	(0.021)	(0.020)
Aged 75 and over	0.012	0.013	0.060***	0.056***
	(0.019)	(0.019)	(0.020)	(0.021)
Can read and write	0.008	0.004	0.009	0.008
	(0.016)	(0.018)	(0.014)	(0.015)
Finished primary	-0.008	-0.011	-0.015	-0.015
	(0.015)	(0.017)	(0.015)	(0.015)
Junior high and above	-0.059***	-0.058***	-0.031**	-0.031**
	(0.015)	(0.017)	(0.015)	(0.014)
logPCE (< median)	-0.040***	-0.040***	-0.036***	-0.037***
	(0.007)	(0.008)	(0.008)	(0.009)
logPCE (> median, marginal)	0.056***	0.055***	0.031***	0.032***
	(0.011)	(0.011)	(0.011)	(0.012)
Rural	0.035***		0.039***	
	(0.012)		(0.014)	
County Dummies	YES	NO	YES	NO
Community Dummies	NO	YES	NO	YES
F-test for all age dummies	1.28	1.57	4.20***	3.92***
(p-value)	(0.262)	(0.157)	(0.000)	(0.001)
F-test for all education dummies	11.90***	10.47***	2.27*	2.46*
(p-value)	(0.000)	(0.000)	(0.078)	(0.063)
F-test for all logPCE splines	15.82***	12.28***	12.45***	9.84***
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)
F-test for all location dummies	4.91***	3.20***	6.20***	3.68***
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	8268	8268	8697	8697

Standard error in parentheses, all clustered at community level. * p<.1 ** p<.05 *** p<.01. logPCE (>median, marginal) represents the change in the slope from the interval for logPCE blow the median, see footnote 7. Sample are respondents not younger than 45.

Table 8. BMI, by age and sex

	Men					Women				
	BMI Mean	**%BMI <18.5**	**%BMI >=25.0**	**%BMI >=30.0**	**N**	**BMI Mean**	**%BMI <18.5**	**%BMI >=25.0**	**%BMI >=30.0**	**N**
45-49	24.2	3.2	35.3	5.4	1074	25.0	3.5	42.2	8.7	1490
	(0.3)	(0.9)	(2.5)	(1.1)		(0.5)	(1.1)	(2.6)	(2.3)	
50-54	23.5	3.7	29.8	3.4	907	24.6	3.0	40.6	6.3	1019
	(0.2)	(0.7)	(2.8)	(0.7)		(0.2)	(0.6)	(2.3)	(1.0)	
55-59	23.1	3.6	24.5	3.0	1321	25.6	5.4	42.2	14.6	1447
	(0.2)	(0.6)	(2.4)	(0.7)		(1.2)	(0.8)	(4.4)	(5.9)	
60-64	23.2	6.8	28.5	3.6	1126	24.2	6.0	37.5	7.5	1187
	(0.1)	(0.9)	(1.8)	(0.8)		(0.2)	(0.9)	(1.9)	(1.2)	
65-69	23.2	7.0	30.2	2.5	797	23.9	8.6	36.3	5.4	745
	(0.4)	(1.0)	(4.6)	(0.6)		(0.2)	(1.2)	(3.3)	(1.0)	
70-74	22.6	12.2	24.2	2.3	569	23.8	10.2	31.2	6.8	521
	(0.3)	(2.3)	(3.9)	(0.9)		(0.3)	(1.5)	(2.8)	(1.4)	
75+	21.8	17.3	19.5	1.8	533	22.5	17.0	25.0	5.2	588
	(0.3)	(2.1)	(3.0)	(0.8)		(0.3)	(2.0)	(3.1)	(1.5)	
Total	23.2	6.5	28.1	3.4	6327	24.5	6.6	38.1	8.6	6997
(45+)	(0.1)	(0.5)	(1.1)	(0.4)		(0.4)	(0.5)	(1.5)	(1.7)	

Sample are respondents not younger than 45. Weighted at biomarker individual level with household and response adjustment. Standard errors in parentheses.

Table 9. Regressions for underweight

	Men		Women	
	(1)	(2)	(1)	(2)
Aged 50-54	0.006	0.001	0.000	-0.000
	(0.011)	(0.008)	(0.010)	(0.008)
Aged 55-59	0.002	0.002	0.027***	0.027***
	(0.011)	(0.008)	(0.010)	(0.009)
Aged 60-64	0.026**	0.026***	0.025**	0.023**
	(0.011)	(0.009)	(0.010)	(0.009)
Aged 65-69	0.042***	0.042***	0.050***	0.049***
	(0.012)	(0.012)	(0.012)	(0.012)
Aged 70-74	0.076***	0.073***	0.087***	0.088***
	(0.014)	(0.015)	(0.013)	(0.016)
Aged 75 and over	0.138***	0.137***	0.157***	0.159***
	(0.014)	(0.020)	(0.014)	(0.018)
Can read and write	-0.017	-0.014	0.009	0.010
	(0.012)	(0.014)	(0.009)	(0.010)
Finished primary	-0.014	-0.013	-0.011	-0.012
	(0.011)	(0.013)	(0.009)	(0.008)
Junior high and above	-0.012	-0.006	0.001	0.001
	(0.012)	(0.013)	(0.010)	(0.008)
logPCE (< median)	-0.023***	-0.027***	-0.017***	-0.015**
	(0.005)	(0.006)	(0.005)	(0.006)
logPCE (> median, marginal)	0.021**	0.026***	0.017**	0.014*
	(0.008)	(0.009)	(0.007)	(0.008)
Rural	0.014		0.041***	
	(0.009)		(0.009)	
County Dummies	YES	NO	YES	NO
Community Dummies	NO	YES	NO	YES
F-test for all age dummies	22.34***	12.26***	29.37***	16.47***
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)
F-test for all education dummies	0.75	0.60	1.24	1.87
(p-value)	(0.522)	(0.617)	(0.293)	(0.135)
F-test for all logPCE splines	10.61***	11.20***	5.80***	3.67**
(p-value)	(0.000)	(0.000)	(0.003)	(0.027)
F-test for all location dummies	2.55***	1.94***	2.47***	2.02***
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	6294	6294	6961	6961

Standard error in parentheses, all clustered at community level. * p<.1 ** p<.05 *** p<.01. logPCE (>median, marginal) represents the change in the slope from the interval for logPCE blow the median, see footnote 7. Sample are respondents not younger than 45.

Table 10. Regressions for overweight

	Men		Women	
	(1)	(2)	(1)	(2)
Aged 50-54	-0.067***	-0.062***	0.011	0.014
	(0.019)	(0.020)	(0.019)	(0.020)
Aged 55-59	-0.067***	-0.062***	-0.045**	-0.040**
	(0.018)	(0.018)	(0.018)	(0.019)
Aged 60-64	-0.036*	-0.030	-0.037*	-0.031
	(0.019)	(0.021)	(0.019)	(0.021)
Aged 65-69	-0.075***	-0.072***	-0.069***	-0.063***
	(0.020)	(0.021)	(0.022)	(0.023)
Aged 70-74	-0.106***	-0.102***	-0.132***	-0.131***
	(0.023)	(0.024)	(0.025)	(0.027)
Aged 75 and over	-0.142***	-0.142***	-0.188***	-0.191***
	(0.024)	(0.025)	(0.025)	(0.025)
Can read and write	0.000	-0.001	-0.001	-0.005
	(0.019)	(0.019)	(0.017)	(0.017)
Finished primary	0.019	0.023	-0.003	0.002
	(0.019)	(0.018)	(0.018)	(0.019)
Junior high and above	0.053***	0.051**	-0.028	-0.024
	(0.019)	(0.020)	(0.018)	(0.020)
logPCE (< median)	0.046***	0.042***	0.035***	0.035***
	(0.009)	(0.009)	(0.009)	(0.010)
logPCE (> median, marginal)	-0.055***	-0.052***	-0.057***	-0.057***
	(0.014)	(0.014)	(0.013)	(0.014)
Rural	-0.090***		-0.096***	
	(0.016)		(0.016)	
County Dummies	YES	NO	YES	NO
Community Dummies	NO	YES	NO	YES
F-test for all age dummies	7.74***	7.65***	14.03***	14.01***
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)
F-test for all education dummies	4.59***	3.69**	0.94	0.65
(p-value)	(0.003)	(0.012)	(0.422)	(0.583)
F-test for all logPCE splines	12.60***	9.78***	8.97***	8.26***
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)
F-test for all location dummies	3.50***	2.44***	4.36***	2.78***
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	6294	6294	6961	6961

Standard error in parentheses, all clustered at community level. * p<.1 ** p<.05 *** p<.01. logPCE (>median, marginal) represents the change in the slope from the interval for logPCE blow the median, see footnote 7. Sample are respondents not younger than 45.

Table 11. Hypertension, under-diagnosis of hypertension and percentage taking medication or treatment, by age, sex and urban/rural

	Rural		Urban		Total	
	Men	**Women**	**Men**	**Women**	**Men**	**Women**
	%	**%**	**%**	**%**	**%**	**%**
Hypertension						
45-49	24.4	24.4	30.8	32.0	27.6	28.6
50-54	31.2	28.1	33.8	37.1	32.4	32.5
55-59	33.5	35.1	45.1	52.7	39.4	44.5
60-64	38.3	44.7	51.9	53.0	44.3	48.5
65-69	43.9	56.2	52.4	61.3	47.8	58.4
70-74	52.0	59.9	55.3	61.2	53.6	60.6
75+	54.3	69.7	64.2	70.4	59.0	70.1
Total (45+)	37.2	41.4	45.1	48.9	41.0	45.2
Under-diagnosis of hypertension						
45-49	53.7	45.8	45.2	50.7	49.0	48.8
50-54	53.7	38.2	41.5	40.7	47.8	39.6
55-59	42.4	32.8	55.4	46.1	50.0	41.2
60-64	36.9	33.0	29.3	37.1	33.0	35.1
65-69	46.5	36.6	28.3	24.5	37.2	31.0
70-74	43.0	44.7	36.7	38.0	40.0	41.1
75+	46.2	45.8	38.4	43.7	42.1	44.7
Total (45+)	45.4	39.3	40.5	41.6	42.8	40.6
Percentage taking medication or treatment						
45-49	57.9	72.2	62.0	88.2	60.3	81.6
50-54	74.7	75.9	74.5	80.9	74.6	78.7
55-59	70.3	76.3	87.7	87.9	79.3	83.0
60-64	77.4	81.7	80.2	78.0	78.9	79.9
65-69	78.4	79.9	89.8	82.5	85.0	81.3
70-74	70.2	79.2	85.9	89.3	78.2	84.9
75+	83.5	80.6	90.0	89.8	87.1	85.3
Total (45+)	73.8	78.4	82.1	85.4	78.4	82.2

Sample are respondents not younger than 45 and exclude respondents that didn't participate in the physical examination.

Weighted at biomarker individual level with household and response adjustment.

Table 12. Regressions for hypertension

	Men		Women	
	(1)	(2)	(1)	(2)
Aged 50-54	0.042*	0.044**	0.070***	0.067***
	(0.021)	(0.021)	(0.019)	(0.018)
Aged 55-59	0.108***	0.099***	0.141***	0.142***
	(0.020)	(0.020)	(0.018)	(0.018)
Aged 60-64	0.170***	0.164***	0.224***	0.222***
	(0.021)	(0.022)	(0.019)	(0.020)
Aged 65-69	0.183***	0.168***	0.339***	0.335***
	(0.023)	(0.024)	(0.022)	(0.021)
Aged 70-74	0.304***	0.296***	0.355***	0.361***
	(0.026)	(0.025)	(0.025)	(0.024)
Aged 75 and over	0.321***	0.304***	0.444***	0.436***
	(0.027)	(0.029)	(0.025)	(0.024)
Can read and write	0.024	0.017	-0.007	0.000
	(0.022)	(0.021)	(0.017)	(0.017)
Finished primary	0.066***	0.060***	-0.005	-0.002
	(0.021)	(0.022)	(0.018)	(0.019)
Junior high and above	0.041*	0.025	-0.012	-0.011
	(0.022)	(0.021)	(0.018)	(0.020)
logPCE (< median)	0.019*	0.016	0.003	-0.002
	(0.010)	(0.011)	(0.009)	(0.010)
logPCE (> median, marginal)	-0.019	-0.019	-0.005	-0.001
	(0.015)	(0.016)	(0.013)	(0.014)
Rural	-0.083***		-0.044***	
	(0.018)		(0.016)	
County Dummies	YES	NO	YES	NO
Community Dummies	NO	YES	NO	YES
F-test for all age dummies	39.96***	37.51***	85.66***	83.30***
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)
F-test for all education dummies	3.76**	3.03**	0.15	0.13
(p-value)	(0.010)	(0.030)	(0.931)	(0.945)
F-test for all logPCE splines	1.96	1.00	0.07	0.20
(p-value)	(0.141)	(0.368)	(0.933)	(0.816)
F-test for all location dummies	3.06***	2.21***	3.01***	2.17***
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	6473	6473	7153	7153

Standard error in parentheses, all clustered at community level. * p<.1 ** p<.05 *** p<.01. logPCE (>median, marginal) represents the change in the slope from the interval for logPCE blow the median, see footnote 7. Sample are respondents not younger than 45.

Table 13. Regressions for underdiagnose of hypertension (by sex and urban/rural)

	Rural				**Urban**			
	Men		Women		Men		Women	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Aged 50-54	-0.034	-0.035	-0.055	-0.048	0.026	0.023	-0.155***	-0.157***
	(0.054)	(0.052)	(0.048)	(0.046)	(0.063)	(0.069)	(0.057)	(0.052)
Aged 55-59	-0.180***	-0.178***	-0.102**	-0.108***	0.033	0.029	-0.130**	-0.149**
	(0.050)	(0.050)	(0.043)	(0.039)	(0.055)	(0.067)	(0.052)	(0.058)
Aged 60-64	-0.242***	-0.254***	-0.106**	-0.114**	-0.055	-0.046	-0.109**	-0.134**
	(0.050)	(0.053)	(0.044)	(0.048)	(0.057)	(0.063)	(0.053)	(0.056)
Aged 65-69	-0.158***	-0.144***	-0.114**	-0.129**	0.028	0.007	-0.193***	-0.203***
	(0.052)	(0.049)	(0.046)	(0.051)	(0.062)	(0.069)	(0.057)	(0.052)
Aged 70-74	-0.193***	-0.169***	-0.051	-0.052	-0.063	-0.047	-0.053	-0.067
	(0.056)	(0.055)	(0.051)	(0.053)	(0.065)	(0.062)	(0.061)	(0.059)
Aged 75 and over	-0.195***	-0.199***	-0.056	-0.050	0.050	0.056	-0.075	-0.101*
	(0.060)	(0.064)	(0.050)	(0.051)	(0.066)	(0.072)	(0.062)	(0.055)
Can read and write	-0.077*	-0.085**	-0.038	-0.040	-0.101	-0.090	-0.041	-0.039
	(0.042)	(0.043)	(0.033)	(0.036)	(0.072)	(0.076)	(0.045)	(0.054)
Finished primary	-0.063	-0.066	-0.054	-0.066*	0.000	0.004	-0.004	-0.020
	(0.041)	(0.040)	(0.036)	(0.040)	(0.069)	(0.078)	(0.044)	(0.048)
Junior high and above	-0.135***	-0.129***	-0.066	-0.077*	-0.012	-0.004	0.006	-0.013
	(0.043)	(0.041)	(0.044)	(0.044)	(0.069)	(0.080)	(0.044)	(0.049)
logPCE (< median)	0.024	0.020	-0.017	-0.016	-0.063**	-0.056**	-0.065**	-0.077***
	(0.020)	(0.022)	(0.018)	(0.018)	(0.028)	(0.026)	(0.025)	(0.023)

logPCE (> median, marginal)	-0.061*	-0.054*	0.028	0.032	0.080**	0.068	0.078**	0.099***
	(0.033)	(0.032)	(0.027)	(0.026)	(0.040)	(0.044)	(0.036)	(0.031)
County Dummies	YES	NO	YES	NO	YES	NO	YES	NO
Community Dummies	NO	YES	NO	YES	NO	YES	NO	YES
F-test for all age dummies	5.51***	4.93***	1.65	2.34**	1.18	0.71	2.76**	3.21***
(p-value)	(0.000)	(0.000)	(0.131)	(0.033)	(0.317)	(0.646)	(0.011)	(0.006)
F-test for all education dummies	3.49**	3.37**	1.21	1.48	1.59	1.72	0.44	0.20
(p-value)	(0.015)	(0.020)	(0.305)	(0.222)	(0.191)	(0.166)	(0.726)	(0.897)
F-test for all logPCE splines	1.84	1.70	0.57	0.76	2.53*	2.55*	3.24**	5.68***
(p-value)	(0.159)	(0.186)	(0.568)	(0.470)	(0.081)	(0.083)	(0.039)	(0.005)
F-test for all location dummies	2.12***	1.65***	2.55***	1.91***	1.74***	1.46***	1.29**	1.13
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.003)	(0.033)	(0.178)
Observations	1513	1513	1777	1777	1019	1019	1240	1240

Standard error in parentheses, all clustered at community level. * p<.1 ** p<.05 *** p<.01. logPCE (>median, marginal) represents the change in the slope from the interval for logPCE blow the median, see footnote 7. Sample are respondents not younger than 45.

Table 14. Possibility of reaching 75, by age and sex

	Men						Women					
	Almost impossible	**Not very likely**	**Maybe**	**Very likely**	**Almost certain**	**N**	**Almost impossible**	**Not very likely**	**Maybe**	**Very likely**	**Almost certain**	**N**
45-49	4.9	10.1	24.9	14.0	19.5	1218	4.0	13.1	28.6	16.5	14.0	1535
	(0.7)	(1.1)	(1.7)	(1.2)	(1.8)		(0.5)	(1.1)	(1.7)	(1.7)	(1.4)	
50-54	3.7	11.0	31.7	15.1	18.1	991	4.7	14.5	27.9	13.5	15.8	1044
	(0.7)	(1.1)	(2.9)	(1.3)	(1.6)		(0.7)	(1.2)	(1.7)	(1.3)	(1.9)	
55-59	3.9	9.5	29.7	15.2	19.8	1364	4.4	15.6	31.3	12.6	14.5	1437
	(0.5)	(0.8)	(1.4)	(1.2)	(2.5)		(0.6)	(1.2)	(1.8)	(1.0)	(1.5)	
60-64	4.6	12.8	30.5	16.0	18.2	1193	5.4	13.7	30.8	13.6	14.0	1145
	(0.5)	(1.1)	(1.5)	(1.2)	(1.4)		(0.7)	(1.1)	(1.7)	(1.3)	(1.3)	
Total	4.3	10.7	28.9	15.0	19.0	4766	4.6	14.2	29.7	14.2	14.5	5161
(45-65)	(0.4)	(0.6)	(1.0)	(0.8)	(1.1)		(0.4)	(0.7)	(1.0)	(0.8)	(1.0)	

Sample are respondents between 45 and 65. Weighted at individual level with household and response adjustment. Standard errors in parentheses.

Table 15. Regressions for not very likely or almost impossible to reach 75

	Men		Women	
	(1)	(2)	(1)	(2)
Aged 50-54	-0.000	-0.004	0.020	0.015
	(0.017)	(0.018)	(0.017)	(0.018)
Aged 55-59	-0.042***	-0.043***	0.011	0.009
	(0.016)	(0.016)	(0.016)	(0.016)
Aged 60-64	-0.039**	-0.038**	-0.015	-0.010
	(0.017)	(0.018)	(0.018)	(0.017)
Can read and write	0.007	0.013	-0.032*	-0.039**
	(0.025)	(0.031)	(0.018)	(0.020)
Finished primary	-0.032	-0.024	-0.060***	-0.072***
	(0.024)	(0.030)	(0.019)	(0.019)
Junior high and above	-0.096***	-0.085***	-0.077***	-0.077***
	(0.024)	(0.030)	(0.019)	(0.022)
logPCE (< median)	-0.045***	-0.043***	-0.071***	-0.069***
	(0.011)	(0.013)	(0.011)	(0.013)
logPCE (> median, marginal)	0.041**	0.041**	0.082***	0.083***
	(0.016)	(0.017)	(0.016)	(0.018)
Rural	0.028*		0.019	
	(0.017)		(0.018)	
County Dummies	YES	NO	YES	NO
Community Dummies	NO	YES	NO	YES
F-test for all age dummies	3.65**	3.32**	1.39	0.90
(p-value)	(0.012)	(0.020)	(0.245)	(0.441)
F-test for all education dummies	14.01***	12.06***	6.24***	5.68***
(p-value)	(0.000)	(0.000)	(0.000)	(0.001)
F-test for all logPCE splines	9.48***	5.89***	20.75***	14.62***
(p-value)	(0.000)	(0.003)	(0.000)	(0.000)
F-test for all location dummies	3.87***	2.58***	4.49***	2.62***
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	4725	4725	5117	5117

Standard error in parentheses, all clustered at community level. * p<.1 ** p<.05 *** p<.01. logPCE (>median, marginal) represents the change in the slope from the interval for logPCE blow the median, see footnote 7. Sample are respondents not younger than 45.

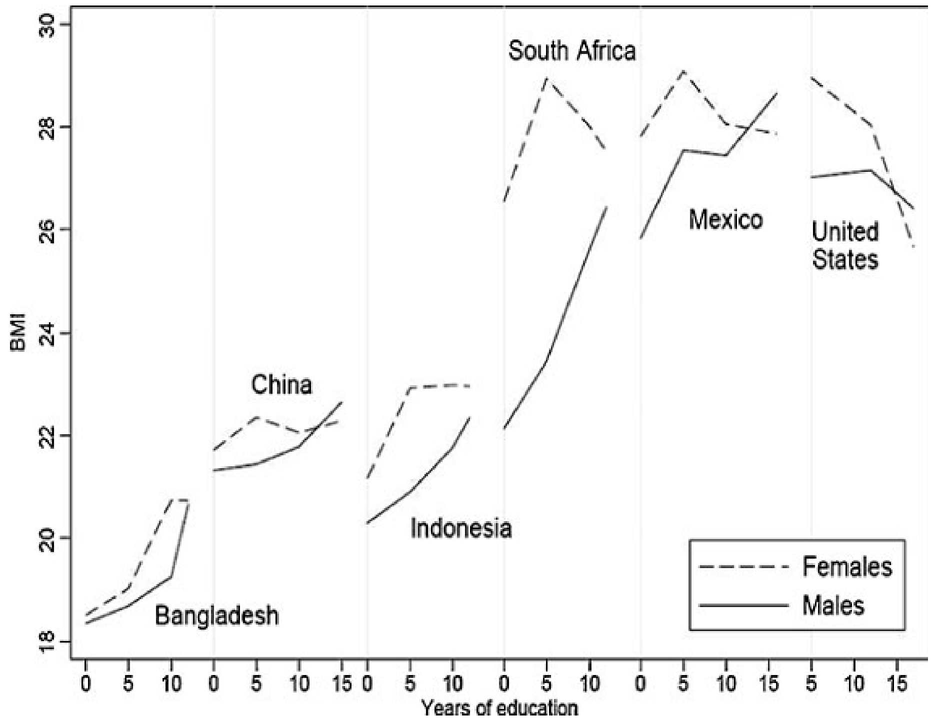

Figure 1 Relationship between BMI and education in selected countries. Reproduced from Strauss and Thomas (2008)

Appendix

Table A. Means and standard deviations of variables, by sex

	Men			Women		
	Mean	**SE**	**N**	**Mean**	**SE**	**N**
Demographics						
Age	59.8	0.165	8426	59.64	0.226	8890
Aged 45-54	0.35	0.007	8426	0.37	0.01	8890
Aged 55-64	0.36	0.007	8426	0.34	0.007	8890
Aged 65-74	0.19	0.006	8426	0.17	0.005	8890
Aged 75+	0.1	0.005	8426	0.12	0.005	8890
Illiterate	0.12	0.007	8415	0.39	0.013	8872
Can read or write	0.17	0.008	8415	0.16	0.007	8872
Finished primary	0.25	0.008	8415	0.17	0.007	8872
Junior high or above	0.46	0.016	8415	0.27	0.013	8872
logPCE	9.07	0.046	8320	9.11	0.045	8765
Rural	0.52	0.027	8320	0.5	0.026	8765
Health Outcomes						
Self-reported health poor or very poor	0.22	0.007	8372	0.27	0.009	8822
Disability (Any ADL or IADL)	0.19	0.007	8360	0.27	0.009	8811
Any body pain	0.25	0.009	8360	0.36	0.01	8801
Moderate or severe body pain	0.18	0.007	8349	0.27	0.009	8788
BMI	23.24	0.101	6328	24.49	0.402	6998
Underweight	0.06	0.005	6328	0.07	0.005	6998
Overweight	0.28	0.011	6328	0.38	0.015	6998
Hypertension	0.41	0.013	6501	0.45	0.015	7190
Under-diagnosis of hypertension	0.43	0.02	2542	0.41	0.023	3025
"Not very likely" or "almost impossible" to reach 75	0.19	0.01	4766	0.24	0.011	5161

Sample are respondents not younger than 45. All weighted at individual level with household and response adjustment expect BMI, underweight, overweight, hypertension, and under-diagnosis of hypertension. BMI, underweight, overweight, hypertension, and under-diagnosis of hypertension are weighted by biomarker individual level with household and response adjustment.