SHARE ERIC
SURVEY OF HEALTH, AGEING AND RETIREMENT IN EUROPE

SHARE WORKING PAPER SERIES

Can we trust older people's statements on their childhood circumstances? Evidence from SHARELIFE

Enkelejda Havari and Fabrizio Mazzonna

Working Paper Series 05-2011

SHARE-ERIC | Amalienstr. 33 | 80799 Munich | Germany | share-eric.eu

mea | MAX-PLANCK-INSTITUT FÜR SOZIALRECHT UND SOZIALPOLITIK
Munich Center for the Economics of Aging
Horizon 2020
Federal Ministry of Education and Research
DFG
NIH National Institute on Aging

Can we trust older people's statements on their childhood circumstances? Evidence from SHARELIFE*

Enkelejda Havari
Tor Vergata University

Fabrizio Mazzonna
Ca' Foscari University of Venice and MEA†

Draft version
August 2011

Abstract

This study provides evidence about the quality of retrospective assessments of individuals aged 50+ regarding their childhood histories in 3rd wave of the Survey of Health Ageing and Retirement in Europe (SHARE). Early life events are important to social scientists in predicting individuals' outcomes in adulthood. Nevertheless, there is wide skepticism about the ability of old age respondents to recall with good accuracy events which happened decades ago. We assess the internal and external consistency of some measures of childhood health and socio-economic status and find that overall respondents seem to remember well their health status and living conditions between ages 0-15. Thanks to the cross-country dimension of SHARE (13 European countries), we are able to compare individual responses with aggregate data (e.g. GDP per capita) at country level. The results we find should mitigate doubts on retrospective data collection and promote their use for research purposes.

Keywords: retrospective, childhood, health, SHARE, methods.

JEL codes: I10, J10, J14.

*This paper uses data from SHARE release 2.3.1. SHARE data collection in 2004-2007 was primarily funded by the European Commission through its 5th and 6th framework programmes (project numbers QLK6-CT-2001- 00360; RII-CT- 2006-062193; CIT5-CT-2005-028857). Additional funding by the US National Institute on Aging (grant numbers U01 AG09740-13S2; P01 AG005842; P01 AG08291; P30 AG12815; Y1-AG-4553-01; OGHA 04-064; R21 AG025169) as well as by various national sources is gratefully acknowledged (see http://www.share-project.org for a full list of funding institutions).

†Munich Center for the Economics of Ageing (MEA) at the Max-Planck-Institute for Social Law and Social Policy (MPISOC).

1 Introduction

Nowadays, researchers are becoming increasingly interested in assessing causal relationships between childhood circumstances and outcomes in adulthood. This is due to a large literature on child development which documents that early life events are good predictors for health and labor market outcomes in adulthood (e.g. Currie 2009, Case et al. 2002, 2005, 2010). As stressed by Almond and Currie (2010), many empirical works based on life-course data show that childhood characteristics (measured at age 0-5) can predict 20% of the variation in wages at age 33, which is a good result for labor economists. Cunha and Heckman (2007) show that cognitive and non-cognitive skills - that are fundamental for later life outcomes - are the results of the interaction of genetic traits and early environmental factors. Another body of research in developmental epidemiology arrives to similar conclusions by documenting the role of adverse early environment on adult health status (Barker 1998, Gluckman and Hanson 2005).

However, to assess whether childhood circumstances affect directly adult health and socioeconomic status or indirectly through human capital accumulation (e.g. education), one needs to have detailed data on individuals' lives from birth to adulthood. Unfortunately prospective surveys like nationally representative cohort studies (e.g. British cohort studies) that follow individuals for a long time are expensive and available for a limited number of countries (e.g. US, UK).

The alternative is to collect retrospective information where individuals provide subjective assessments on their health status and living conditions in childhood, as well as their experiences in health, education, employment, life satisfaction, etc. As stressed by Schröder (2008) this method is "*faster, less costly and the risk that respondents drop out of the study is much lower than in a prospective survey*".

Nonetheless, this method may suffer from problems of recall bias and coloring. Recall bias arises when individuals do not remember accurately when and how an event took place in the past, whereas coloring (anchoring) is a consequence of projecting the current status (e.g. health) in the past, when answering to questions on childhood or young adulthood.

Despite these potential concerns, recent longitudinal surveys have started to collect retrospective information on individuals life histories (including information on their childhood circumstances), to have a picture of their experiences before the baseline year of the survey. As Smith (2009a) argues, information on pre-baseline health histories is crucial for researchers in order to avoid untestable

assumptions on the initial conditions. This is particularly true for a high-quality longitudinal survey like the Survey of Health Ageing and Retirement in Europe (SHARE) that collects micro-data on health and socioeconomic status of individuals aged 50 or more across 15 European countries. For this reason, the third wave, called SHARELIFE, provides retrospective data on individuals aged 50+ and their spouses, independent of age, regarding their health and working history, childhood circumstances and family background and other general questions on well-being from childhood to adulthood. Since respondents have to recall events that took place 50 years ago or longer, the assessment of data quality in SHARELIFE is of primary interest. So far, there are only a few contributions that look at the accuracy of retrospective assessments in SHARELIFE (data released in November 2010). Garrouste and Paccagnella (2010) provide an overall assessment of recall bias in SHARE concentrating mostly on how coherent are individuals in answering to the same questions at different points in time (2004, 2006, 2008). Yet, there are no works that carefully look at responses on childhood characteristics and circumstances.

Our contribution is therefore to analyze the quality of retrospective information on childhood characteristics in SHARELIFE, which might be of interest for many researchers.

Studies based on longitudinal surveys similar to SHARE (e.g. PSID, HRS), have been mostly dealing with the quality of retrospective childhood health measures such as self-reported childhood health status (SRHS) and chronic diseases (e.g. Haas 2007, Smith 2009a). However, as stressed by Smith (2009a): "*other domains of childhood, in addition to health status, may be critical for later life. These might include child exposure to adverse physical or social environment*". For this reason, we analyze not only the quality of some childhood health measures but we also consider indicators of family socio-economic background (e.g. number of books, accommodation features), as well as episodes of hunger and financial hardship that should be important predictors of later life outcomes.

In evaluating the quality of childhood measures we concentrate on their within wave consistency but we also exploit the cross-country dimension of SHARE. In particular, we use external data (i.e. GDP per capita, average years of schooling and war episodes) at country level to provide some external validity on variables related to childhood SES, hunger and financial hardship episodes.

It is clear that in absence of administrative data to compare with the retrospective assessments on childhood characteristics it is difficult to have a sharp evaluation of their quality[1]. Nevertheless,

[1]Only for Germany there is a project that links survey data to administrative records (see Korbmacher and

our results show a good level of internal and external consistency of most of the self-reported childhood measures, especially when we exploit the cross-country dimension of SHARE.

The remainder of this paper is organized as follows: Section 2 presents a brief review of the major works that assesses the quality of retrospective information; Section 3 describes the data and the methods used for this study; Section 4 presents our results; Section 5 offers some conclusions.

2 Literature on retrospective data

In the past 20 years there has been a growing interest in assessing the quality of retrospective data. This is strictly linked to the widespread use of household survey data that cover different dimensions of individuals and their families' life (e.g. lifetime earnings, health status, health care). The alternative way, embraced by most developed and developing countries (e.g. HRS, SHARE, ELSA, DHS), is to collect individual data at different points in time in a longitudinal framework by asking information on individuals' life histories before the baseline year of the survey. Unfortunately, the desirability to collect retrospective data over a long period of an individual's life is counterbalanced by different critiques on data quality (Bound et al. 2001). That is especially true when individuals are asked to remember events occurred up to 50 years ago or even more, like in SHARELIFE. The presence of recall errors, in fact, may bias the estimates of the long-run effects of childhood events on later-life outcomes.

Several contributions (Dex 1991, Beckett 2001, Smith and Thomas 2003), agree that the recall of an event depends on its attributes (e.g. timing, frequency with which it has occurred, etc.), respondents' characteristics (age, schooling) and the survey practices implemented (e.g. the use of temporal landmarks derived by the most important life events during life). Intuitively, an event occurred decades ago would be recalled with less accuracy with respect to an event that took place last week or last year. However, the more salient an event with more accuracy it will be recalled, even after 50 years ago since it has occurred.

To deal with recall bias the literature has developed in two directions: *i*) find ways to minimize recall errors in a survey (ex-ante approach); *ii*) assess the quality of recall data (ex-post). On the former Bound et al. (2001) and Beckett (2001) provide useful summaries.

To improve the design of retrospective surveys, scientists have tried to exploit how memory

Schröder 2010).

works. Two are the most intuitive points behind such approach. First, events that are more important are easier to remember. Second, memory is organized in a hierarchical format, namely it moves from general structure to generic ones (e.g. Belli 1998).

So far, one practical way has been the use of Life Calendar Methods (LCM). They are based on a bounded recall framework, according to which the presence of outstanding events during one's life - such as marriage date, child birth etc. - can be used as an anchor for recalling other events that have changed more frequently (e.g. patterns of mobility in accommodation). The first attempt was done by Friedman (1988) in a paper and pen version (Börsch-Supan 2008). Nowadays it is possible to opt for modern computerized versions, where respondents can see the placement of the events in a monitor while they answer. Belli et al. (2005) implemented this method within the PSID survey, arguing that graphical devices in general and LCM approaches in particular can improve respondents recall ability.

The other strand of the literature evaluates the quality of retrospective assessments once the data are collected. One way to do it is to validate retrospective information with more objective assessments. Such practice is followed especially by epidemiologists when dealing with retrospective childhood illnesses. Krall et al. (1988) validated self-reports of respondents' childhood diseases, accidents, hospitalizations at age 30, 40 and 50 by using physician assessments collected during their infancy in a longitudinal prospective study. Illnesses were recalled with a high level of accuracy of about 85% at age 50. Similarly, Berney and Blane (1998) revealed that a substantial majority of subjects had recalled simple socio-demographic information, such as a father's occupation and simple residential information, after a period of 50 years with a useful degree of accuracy[2].

Smith (2009a), on the other hand, validates the quality of responses on child diseases (infectious diseases, asthma, allergies etc.) by linking disease prevalence rates in the HRS and PSID with objective prevalence rates from the National Health Examination Survey when the respondents were children. However, with the enrichment of household surveys other methods have been applied. Using the HRS and PSID, Haas (1997) and Haas and Bishop (2000) compare individual's responses over time showing that the retrospective measure of SRCH is reliably reported. A similar study on HRS has been done by Elo (1998) where she finds a high level of consistency between childhood SRHS and indicators for missing school due to health problems.

[2]The study based on 57 observation compare interview data on selected items with archive material achieving a level of agreement between 83% to 100%.

3 Data and methods

3.1 Data

The Survey of Health, Ageing and Retirement in Europe (SHARE) is a multidisciplinary, cross-national bi-annual household panel survey started in 2004. The survey collects data on health, socio-economic status (SES), and social and family networks for nationally representative samples of elderly people in the participating countries. SHARE is designed to be cross-nationally comparable and is harmonized with the U.S. Health and Retirement Study (HRS) and the English Longitudinal Study of Ageing (ELSA). We use data from 13 countries where individuals participate to both the second and the third wave of the survey. These countries represent different European regions, from Scandinavia (Denmark, Sweden) through Central Europe (Austria, Belgium, France, Germany, the Netherlands, Switzerland) and Mediterranean countries (Greece, Italy, Spain) to Eastern European (Poland and Czech Republic). The target population consists of individuals aged 50 and over who speak the official language of each country and do not live abroad or in an institution, plus their spouses or partners irrespective of age.

The third wave of SHARE, called SHARELIFE, has been implemented to collect the retrospective histories of the SHARE respondents in order to obtain information about the respondents' lives before the baseline year of the survey (2004). The survey design of SHARELIFE has been implemented following the above mentioned literature on retrospective data collection in order to improve the respondents' recall ability. First, SHARELIFE orders the different interview modules according to what is usually most important for the respondent and thus remembered most accurately. The interview, in fact, starts with questions about the children, then follow a module about partner and marriage, events that should be easily remembered. Second, the interview is supported with a life grid computerized version of the LCM that serves as the basis for the SHARELIFE interview. "*Life events are recorded into this life grid, where sets of topics such as children, partners, or work are combined with the time dimension and external historical events*" (Schröder 2010). Basically, as the respondent answers, the information appears in the calendar for both the respondent and the interviewer to see, so that the interviewer has an easy way of linking questions to personal events.

As mentioned before, this method was experimented first in the PSID but then it was done by telephone interview, so the respondent could not see the calendar (Belli et al. 2005). The En-

glish Longitudinal Study of Ageing (ELSA), instead, implemented a face-to-face computer assisted interview in 2007 that serves as a basis for the SHARELIFE interview with the benefit of the comparability of these two studies.

3.2 Measures of childhood circumstances

In this paper we investigate the quality of some of retrospective questions in SHARELIFE that contains valuable childhood information. In particular, we focus our attention on three modules of the questionnaire: childhood health, childhood SES and general life.

The health section opens by asking respondents to rate their health from birth up to age 15 in five categories (*excellent*, *very good*, *good*, *fair*, *poor*), following the same format of adult SRHS. Besides, questions on whether the respondent has missed school because of health problems are asked. Finally, to have more details on health status, it is asked whether or not individuals experienced any of the following diseases from birth until age 15 : *infectious diseases, polio, asthma, other respiratory problems, allergies, severe diarrhea, meningitis, chronic ear problems, speech impairment, difficulty seeing, severe epilepsy/seizures, emotional, nervous, or psychiatric problems, fractures, appendicitis, diabetes, heart problems, leukemia, cancer or other not listed.*

The childhood SES section asks respondents to provide information on their living conditions and family characteristics when they were 10 years old. Among many variables of interest we concentrate on the number of rooms and number of people in the house (excluding bathrooms and kitchens), features of accommodation (fixed bath, cold and hot running water supply, inside toilet and central heating), number of books at home (from none to 2 or more bookcases) and the occupation of the main breadwinner (10 categories). They give an idea about the living standard of the family in which the respondent grew up. Furthermore, in the same section we present assessments of respondents' relative position in Math and Language at age 10 with respect to their class-mates.

The last module that we consider is a general life section that provides relevant information on hunger and financial hardship episodes. We decided to focus on these questions because they might be particularly relevant for research on the long-term consequences of childhood circumstances on adult outcomes. Information on these episodes are asked in following two questions: i) *Looking back on your life was there a period in which you suffered from hunger (financial hardship)? If yes, when did it start and when did it stop?* Knowing year and country of birth it is possible to have a

general view on the period in which hunger or financial hardship problems have started.

4 Results

In this section we present our results. They are organized as follows: in paragraph 4.1 retrospective information on childhood health is exploited by looking at both childhood SRHS and incidence of diseases in the SHARELIFE sample; Section 4.2 shows some descriptive statistics on respondents socio-economic conditions at age 10; finally in Section 4.3 we assess the reliability of self-reported episodes of hunger and financial hardship by using yearly data on GDP per capita for each country and controlling for World War II.

4.1 Childhood health

There are some studies that analyze the relationship between child and adult health by using subjective retrospective measures (SRHS). For instance, Smith (2009b) based on the PSID finds a strong negative impact of poor childhood SRHS on US adults outcomes (health, earnings and labor supply). If we plot a naive correspondence between good health today (after age 50) and health in childhood (age 0-15) using SHARE data we obtain the relationship in Figure 1. At age 60 individuals that were in good health during childhood report good health as adults 15% more often than those whose health in childhood was bad. With our data we cannot cover the whole pathway from childhood to adulthood but we can say that discrepancies between the two groups are higher at age 50 and then shrink as one becomes older.

Figure 1 is informative but it does not reveal the whole story because of the retrospective nature of the data. Thus, before analyzing causal relationships between childhood and adult health, it is important to evaluate the accuracy of retrospective records about health conditions during infancy. As largely explained in Section 2, different solutions have been proposed in the literature to validate such assessments: *i*) using external medical records; *ii*) testing the consistency of responses on the same questions over time (Haas 2007) or *iii*) comparing individual responses to aggregate data on prevalence disease rates for a given country and year (Smith 2009a).

None of these methods can be implemented here because SHARELIFE data are not linked to medical records, we have only one wave of retrospective questions and for the moment we lack data on disease prevalence rates in Europe. Nevertheless, the third method could be feasible in the

near future. The approach we follow, instead, is to analyze the determinants of childhood SRHS using information within SHARELIFE (e.g. selected disease responses and health indicators) and exploiting the cross-country dimension in SHARE.

As in Smith (2009a) and Elo (1998), we start investigating the quality of childhood SRHS by looking at how coherent individuals are in reporting information on selected diseases and frequency with which they went to school when in bad health. Table 1 shows logit estimates where the dependent variable is childhood SRHS being *excellent* or *very good* and controls the large set of self reported illness in childhood presented before. This is widely used in the literature to check whether individuals are coherent in answering to similar questions about health. In the first column we control for country fixed effects (baseline specification) while in the second we add controls for a quadratic polynomial in age and a dummy for female. From Table 1 we notice that suffering from illnesses like asthma, respiratory problems, heart trouble, diarrhea, etc. reduces the probability of reporting good health about childhood period. The more severe the disease is greater is the effect. All coefficients have negative sign and are statistically significant at 1% level with the exception of short term diseases like infectious diseases and fractures. The same story can be told about Table 2. Missing school for more than one month or being in hospital three times a year reduces the probability of reporting good health during childhood but in the last case the effect is bigger. Although these two tables are not informative about recall bias and coloring, they suggest that there is a good internal consistency between self-reported information on child SRHS, reported diseases and missing school.

The following analysis concentrates on disease rates within SHARELIFE, with the aim of evaluating the presence of recall bias problem. Table 3 shows how many respondents (%) declare to have suffered from a specific disease in their childhood (age 0-15). Results refer to different cohorts. It is noticeable how disease rates decrease sharply with age. For instance, only 76% of individuals aged 70+ report to have had infectious diseases (mumps, chicken pox, etc.) compared to 82% of younger respondents.

There can be two interpretations for this pattern. On one hand, it can be attributable to recall bias, as for older individuals is more difficult to recall selected diseases than is for the younger ones. On the other hand, lower rates for the oldest cohort can be explained by the lack of good technology fifty years ago in diagnosing diseases and by the absence of disease prevention in most of the countries. For this reason, in Table 4 we exploit differences in prevalence rates by geographical

areas, to see whether we find a similar pattern across countries. Countries are grouped according to their geographical location that at the time reflected also different stages of the European economic development. For this reason we report in the last row the average GDP per capita by region referring to the period 1926–1956 (GDP per capita is expressed in thousands of 1990 international Geary-Khamis dollars, Maddison 2010). We find a large variation of response rates across countries with low percentages for Mediterranean and Eastern Europe and higher ones for Central Europe and Scandinavia. Individuals born in poorer countries (Italy, Spain, Greece, Czech. Republic, Poland) characterized by a limited access to medical services show lower response rates for most diseases (e.g. infectious diseases, asthma, etc.). Thus, the second interpretation seems to be plausible for our case.

A further check on the recall bias is done in Table 5. We regress the number of chronic diseases (by summing up responses on each disease) on a proxy of memory capacity, which summarizes the outcomes of two cognitive ability tests from wave 2[3]. In the baseline specification we include also a cubic polynomial in age, a dummy for gender and country fixed effect. In the second column we add an indicator for education level (a dummy for an individual having at least a high school degree) and in the last column we control for indicators of childhood SES (e.g. accommodation features, number of books in the house, father's occupation, number of rooms per capita).

It is evident from the fist column that memory capacity is not important in predicting the number of reported chronic diseases, whereas age matters and this is more consistent with the "lack of diagnosis" interpretation rather than recall bias. Education - a significant predictor in the second column - is no longer significant when we include proxies for childhood SES. Similar conclusions arise for the age controls. Standing to our results, recall problems do not matter much for self-reported health diseases, whereas family socio-economic conditions better explain childhood health, being closely linked to health care knowledge and disease prevention in the area of residence of the respondent.

In the end we seek for coloring in childhood SRHS. As thoroughly discussed, this measure could be subject to coloring because health status in adulthood and childhood are contemporaneously reported in SHARELIFE. Since we cannot check the consistency of this question across waves we

[3]This two tests consist of a verbal registration and recall of a list of 10 items. Each respondent hears the list only once. The test is carried out immediately after the encoding phase (immediate recall), and then again after the fluency and numeracy questions (delayed recall).

adopt a different strategy that exploits the availability of the adult SRHS in wave 2 and 3. Basically, we start assuming that the relationship between childhood SRHS and adult SRHS measured in wave 3 is the same as the relationship between childhood SRHS and adult SRHS measured in wave 2. This assumption although untestable, encloses the idea that the relationship between "true" child health status and "true" adult health status after (on average) 50 or 52 years should be almost the same (wave 2 was released in 2006 and SHARELIFE in 2008). Thus, if we observe a stronger relationship between childhood SRHS and adult SRHS in wave 3 than child health and adult SRHS in wave 2, we can suspect the presence of coloring, probably due to simultaneous reporting of child and adult health status in wave 3. This is done by using the Wald test for the difference in the coefficients of childhood SRHS across two regression models. We opt for a cross-model comparison. In the first equation we regress adult SRHS reported in wave 3 (W3) on childhood SRHS and in the second we use as an outcome adult SRHS in wave 2 (W2). The null hypothesis states that the difference in the coefficients of childhood SRHS is equal to zero. We select the same sample and use calibrated cross-sectional weights provided by SHARE to control for presence of attrition between waves. For more accuracy we perform the test separately for men and women controlling for country fixed effects and a quadratic polynomial in age (results are shown in Table 6). In both cases we reject the null hypothesis that the difference in the coefficients is equal to 0, so we conclude that there is some coloring. Moreover, women seem to be more subject to coloring than men.

4.2 Childhood SES

As mentioned in section 3.2, the childhood SES section offers details about living conditions and school performance of the respondents when they were 10 years old. Such information can be used to evaluate the long lasting effects of childhood background on adult and old age outcomes.

As an example of the predicting power of such variables, Figure 2 shows the relationship between the four childhood SES proxies (number of rooms per capita, number of books, number of facilities and breadwinner's occupation) and the old-age per capita income reported in the second wave of the survey (2006). The blue bars show the average per capita income for those who report a value for the first 3 proxies below the country median, while the red ones above that value. In the case of the breadwinner's main occupation, the blue bar shows those with low skill occupation (mainly blue-collar and elementary occupation). The figure clearly shows that individuals from better off families - with higher value of the four proxies - have on average higher income when older.

In order to measure the quality of these variables we use external data and exploit the cross-country dimension of SHARE. Table 7 reports summary statistics by country for three of these indicators (excluding number of books at home), with standard errors in parentheses, and shows the average GDP per capita for the time interval 1926–1956 expressed in thousands of 1990 international Geary-Khamis dollars in the last column (Maddison 2010). The first two columns report the average number of rooms per capita and facilities in the accommodation (in a range of 0 to 4). The third one evidences the proportion of individuals whose breadwinner had a low-skill occupation. There is a very large variation in all indicators between and within countries. Mediterranean countries and Poland as one can expect have lower values for number of rooms per capita and facilities within the house than other countries. Looking at low skill occupation, 86% of the Italian respondents report to have had a parent with a low job position. Scandinavian countries and Switzerland, instead, have the biggest fraction of respondents that grew up in better off households. The validity of such results is confirmed by the historical data on GDP per capita for the reference period (1926-56) reported in the last column. To higher levels of GDP per capita are associated higher average values of the three indicators of living standards.

In the same way, Table 8 compares average years of schooling in 1960 (for individuals aged 25 and over) and the proportion of respondents who report to have had at least a case of books in their accommodation.

As before countries like Italy and Poland, with the lowest values in number of books at home, are associated lower levels of education (years of schooling). This is confirmed by the very high correlation coefficient - 0.69 - between these two indicators. Finally, we look at the self-reported school performances when the respondent was 10. Figure 3 shows the relationship between school performances in math and language and the average years of education reported in wave 2. For each subject there are three categories: i) better than others (blue bars), ii) same as others (red bars), iii) worse than others (green bars). As expected, respondents that report to have performed better than their classmates in math and language at age 10, have on average higher educational attainment. It is worth to remark the differences by gender. For instance, differences in years of education by math level are less marked for females than are for males. This may be due to the lower educational attainment of females from the older cohorts, who very often were obliged to drop out from school earlier (no matter their ability) because of cultural norms.

4.3 General Life Questions

In this final section, we analyze two aspects of general life section in SHARELIFE: hunger and financial hardship episodes. We concentrate on these two variables because striking childhood events matter for future outcomes. Figure 4 evidences the relationship between hunger and financial hardship episodes from birth to age 15 and adult SRHS in wave 3. The blue bars show the proportion of respondents who report to have excellent or very good health in old age among respondents who do not report hunger of financial hardship episodes in childhood (up to the age of 15), whereas the red bars show the same proportion but among those that do report hunger or financial hardship episode. As expected in the group of people that experienced hunger or financial hardship during childhood we have a very small proportion in very good health condition when old. Although not causal, this strong relationship confirms the importance of childhood events in predicting long-term health outcomes.

As we know malnutrition episodes can be an effect of exogenous shocks like famines, wars, natural disasters (earthquakes). To verify the validity of subjective assessments about hunger episodes it is important to assess when such episodes happened and check whether they are consistent with historical events. In Figure 2 we report the histogram of the years in which hunger episodes have started by country, as reported by respondents. As expected, most of the spikes are in the interval 1939-45 for countries that have been involved greatly in World War II (Netherlands, Poland, Germany, etc.), or in a Civil War (1936-1939 for Spain) and the trend is smoother for those countries that were not active during the war (Switzerland, Denmark).

A more formal test of consistency of these episodes with external data on GDP per capita is shown in Table 9. In this table, we use hunger and financial hardship episodes - reported by the individuals - as an outcome in a logit specification. The idea is to verify whether the reported episodes of hunger and financial hardship are consistent with the macro-economic environment, proxied by the average GDP growth rate (at country level), when the respondent was less than 10 or 15 years old. Each regression includes as baseline controls: country fixed effect, a quadratic polynomial in age and the four proxies for family socio-economic status (rooms per capita, number of facilities, number of books at home and breadwinner's occupational level).

The first panel reports the results for hunger and the second for financial hardship. Columns denote different specifications and outcomes. In particular, the first two columns report the logit

estimates on reported hunger (financial hardship) episodes up to age 10, while the other two hunger episodes up to age 15. Starting from the first column of the hunger estimates, there is a noticeable correspondence between hunger episodes and average GDP growth rate. We estimate that a one percent increase in GDP per capita over that interval decreases the probability of reporting hunger episodes to about .72%. This effect, however, may be driven by the WWII that - as seen before - is the period where most of the hunger episodes are concentrated. For this reason the second column includes a dummy variable that is equal to 1 when the respondent was in a country involved in the WWII in the age interval 0-10. Although the inclusion of the dummy variable is positive and significant it only slightly decreases the coefficient of the GDP. In particular, the cohort involved in the WWII shows a 4% increase in the probability of reporting hunger episodes. In order to verify that the control for the average GDP growth is not capturing some country trend effect, we consider a larger interval for hunger episodes (0-15) and we compare the coefficients on two different GDP growth interval (column 3): the old interval (0-10) and the same interval of the dependent variable (0-15). If the average GDP growth rate does capture the effect of macroeconomic conditions and not of other country trends we should observe a significantly larger coefficient in the GDP growth rate interval (0-15). The results confirm this hypothesis. The coefficient in the third column, in fact, is significantly smaller, less then half the coefficient in the fourth column.

Similar considerations arise when we consider financial hardship episodes, except that the effect of GDP growth rates and WWII on the probability of reporting financial hardship episodes is significantly smaller then in the case of hunger.

5 Conclusions

The importance of childhood circumstances in determining individuals' future health and economic status is well documented in the literature (Currie 2009, Case et al 2010), so our work should help researchers interested in the causal relationships between childhood events and later life outcomes, in presence of retrospective data.

Except for the case of the childhood self rated health status where we have evidence of some coloring, our results seem to indicate that self-reported childhood information collected in SHARE-LIFE shows a good level of internal and external consistency.

Our findings are consistent with the literature on response errors and the quality of recall data.

Simple socio-demographic information is relatively accurate when supported by modern survey techniques (e.g. life grid), as they may minimize recall bias (e.g. Berney and Blane 2003). At the same time, we show that the recall of specific diseases during childhood is more affected by lower access to medical services around fifty years ago - hence less knowledge about preventive care - than by recall bias. This implies that cross-country and cross-cohort comparisons need to take into account differences in terms of socio-economic conditions during childhood, which are influenced by differences in public policy interventions (e.g health care).

Based on the evidence presented in this paper, we expect that similar results should overcome part of skepticism about the retrospective data usage and encourage their collection in a longitudinal study. However, more research is needed before giving a definitive judgment on the reliability of retrospective information on childhood circumstances collected at older ages. In particular, it would be helpful to have repeated subjective assessments of the same childhood information over time in order to verify whether life experience events and the aging process have an influence on the recall of this information.

References

Almond D. and Currie J. "Human Capital Development before Age Five", NBER Working Papers 15827, National Bureau of Economic Research.

Barker D. (1998) "In utero programming of chronic diseases". *Clinical Science*, 95: 115–128.

Beckett M. et al. (2001) "The Quality of Retrospective Data: An Examination of Long-Term Recall in a Developing Country". *Journal of Human Resources*, 36(3): 593–625.

Belli R. (1998). "The structure of autobiographical memory and the history calendar: Potential improvements in the quality of retrospective reports in surveys" . *Memory*, 6(4): 383–407.

Berney L. and Blane D. (1998) "Collecting retrospective data: Accuracy of recall after 50 years judged against historical records". *Social Science and Medicine*, 45(10): 1519–1525.

Berney L. and Blane D. (2003) "The lifegrid method of collecting retrospective information from people at older ages". *Research Policy and Planning*, 21(2): 13–22.

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

Bound J., Brown C. and Mathiowetz N. (2001) "Measurement error in survey data". *Handbook of Econometrics*, 5(59): 3707–3833. Edited by J. Heckman and E. Leamer, Elsevier Science.

Börsch-Supan, A. and Schröder M. (2008) "*Retrospective data collection in Europe*" . Mannheim Research Institute for the Economics of Aging (MEA).

Case A., Lubotsky D. and Paxon C. (2002) "Economic Status and Health in Childhood: The Origins of the Gradient". *American Economic Review*, 92(5): 1308–1334.

Case A., Fertig A. and Paxson C. (2005) "The Lasting Impact of Childhood Health and Circumstance". *Journal of Health Economics* 24(2): 365–389.

Case A. and Paxson C. (2009) "Early Life Health and Cognitive Function in Old Age". *American Economic Review Papers and Proceedings*, 99(2): 104–109.

Case A. (2010) "Causes and consequences of early life health". NBER Working Papers, Working Paper 15637.

Cunha F. and Heckman J. (2007) "The Technology of Skill Formation". *American Economic Review*, 97(2): 31–47.

Currie J. (2009) "Healthy, Wealthy, and Wise? Socioeconmic Status, Poor Health in Childhood, and Human Capital Devleopment". *Journal of Economic Literature*, 47(1): 87–122.

Currie J. et al. (2010) "Child Health and Young Adult Outcomes". *Journal of Human Resources*, 45(3): 517–548.

Dex S. (1991) "The Reliability of Recall Data: A Literature Review". *ESRC Research Centre of Micro-social Change*, WP 11. Colchester: University of Essex.

Elo I., (1998) "Childhood Conditions and Adult Health: Evidence from the Health and Retirement Study", Working Paper Series No. 98-03. Population Aging Research Center, University of Pennsylvania.

Elo I and Preston S. (1992) "Effects of Early-Life Conditions on Adult Mortality: A Review". *Population Index*, 58(2): 186–212.

Garrouste C. and Paccagnella O. (2010) "Data Quality: Three Examples of Consistency Across SHARE and SHARELIFE", in *SHARELIFE Methodology*. Mannheim Research Institute for the Economics of Aging (MEA).

Haas S. (2007) "The Long-Term Effects of Poor Childhood Health: An Assessment and Application of Retrospective Reports". *Demography*, 44(1): 113–135.

Haas S. and Bishop N. (2010) "What Do Retrospective Subjective Reports of Childhood Health Capture? Evidence from the WLS and the PSID". *Research on Aging* 32(6): 698–714.

Gluckman P. and Hanson M. (2005) "The developmental origins of adult disease". *Maternal and Child Nutrition*, 1(3): 130–141.

Korbmacher and Schröder (2010) "Non-response when Linking Survey Data with Administrative Data". Paper presented at the 20th International Workshop on Household Nonresponse, Nuremberg, Germany.

Krall E. et al. (1988) "Recall of childhood diseases". *Journal Of Clinical Epidemiology*, 41(11): 1059–1064.

Schröder M. (2011) " Retrospective Data Collection in the Survey of Health, Ageing and Retirement in Europe. SHARELIFE Methodology". Mannheim Research Institute for the Economics of Aging (MEA).

Smith J. Thomas D. (2004) "Remembrances of Things Past: Test-Retest Reliability of Retrospective Migration Histories". *Labor and Demography* 0403026, EconWPA.

Smith J. (2009a) "Reconstructing childhood health histories". *Demography*. 46(2): 387–403.

Smith J. (2009b) "The Impact of Childhood Health on Adult Labor Market Outcomes". *The Review of Economics and Statistics*, 91(3): 478–489.

Tables and Figures

Table 1: Logit estimates child self-reported health status (excellent, very good).

	(1)	(2)
Infectious diseases	-0.011	-0.011
Polio	-2.010 ***	-2.052 ***
Asthma	-1.475 ***	-1.503 ***
Respiratory problems	-0.973 ***	-0.977 ***
Allergies	-0.829 ***	-0.848 ***
Diarrhea	-0.903 ***	-0.885 ***
Meningitis	-0.719 ***	-0.726 ***
Ear problem	-0.776 ***	-0.774 ***
Speech impairment	-0.959 ***	-1.011 ***
Difficulties seeing	-0.745 ***	-0.755 ***
Headaches or migraines	-0.654 ***	-0.629 ***
Epilepsy or seizures	-1.647 ***	-1.700 ***
Depression	-0.697 ***	-0.664 ***
Fractures	0.099	0.047
Appendicite	-0.290 ***	-0.281 ***
Diabetes	-2.085 ***	-2.139 ***
Heart problem	-1.301 ***	-1.313 ***
Cancer	-4.439 ***	-4.350 ***
Demographics		X
Country FE	X	X
N	25185	25183
Pseudo-R^2	0.076	0.079

Logit estimates for childhood self-reported health: 1 "excellent, very good" and 0 "good, fair, poor". Demographic controls include a quadratic polynomial in age and a dummy for female. We add country fixed effects (Italy is the reference country). Standard errors are robust to heteroskedasticity. Significance levels: (*) p-values between 10 and 5 percent; (**) p-values between 5 and 1 percent; (***) p-values less than 1 percent.

Table 2: Logit estimates child self-reported health status.

	(1)	(2)
Miss school > 1 month	-0.464 ***	-0.478 ***
Miss school > 1 month (bed)	-0.852 ***	-0.837 ***
In hospital 3 times a year	-0.593 ***	-0.613 ***
Demographics		X
Country FE	X	X
N	25082	25080
Pseudo-R^2	0.071	0.074

Logit estimates for child self-reported health: 1 "excellent, very good" and 0 "good, fair, poor". Demographic controls include a quadratic polynomial in age and a dummy for female. We add country fixed effects (Italy is the reference country). Standard errors are robust to heteroskedasticity. Significance levels: (*) p-values between 10 and 5 percent; (**) p-values between 5 and 1 percent; (***) p-values less than 1 percent.

Table 3: Child diseases (%), by age.

Age in 2008	50-60		61-70		71 +	
Year of birth	1948-1958		1938-1947		1918-1937	
Age 15	1963-1973		1953-1962		1933-1952	
Child Diseases	Mean	SE	Mean	SE	Mean	SE
Infectious diseases	83.23	(0.40)	82.17	(0.41)	76.95	(0.47)
Meningitis	0.96	(0.10)	0.92	(0.10)	0.49	(0.08)
Asthma	2.04	(0.15)	1.90	(0.15)	1.76	(0.15)
Diarrhea	0.93	(0.10)	1.18	(0.12)	0.82	(0.10)
Respiratory problems	3.36	(0.20)	2.87	(0.18)	2.87	(0.19)
Allergies	4.17	(0.21)	2.87	(0.18)	2.16	(0.16)
Speech impairment	0.84	(0.10)	0.48	(0.07)	0.30	(0.06)
Heart trouble	0.65	(0.09)	0.78	(0.09)	0.65	(0.09)
Ear problem	2.93	(0.18)	2.41	(0.16)	2.11	(0.16)
Headaches or migraines	4.19	(0.21)	4.23	(0.21)	3.56	(0.20)
Fractures	9.82	(0.32)	7.38	(0.28)	4.98	(0.24)
Appendicite	9.46	(0.31)	9.73	(0.31)	6.65	(0.28)
Depression	1.24	(0.12)	1.06	(0.11)	0.56	(0.08)
Diabetes	0.11	(0.04)	0.08	(0.03)	0.05	(0.02)
Cancer	0.03	(0.02)	0.02	(0.02)	0.02	(0.02)
Epilepsy or seizures	0.64	(0.08)	0.38	(0.07)	0.17	(0.05)
Leukemia	0.01	(0.01)	0.07	(0.03)	0.00	(0.00)
N	8881		8914		8240	

Table shows disease rates (%) in SHARELIFE, by age group. Calculations are based on individuals answers (whether they have experienced or not a certain disease from birth until age 15). We report the mean and the standard error for each cohort and disease (in %). Year of birth and year of their 15th birthday are shown on top of the table to highlight the period in which they were children.

Table 4: Child diseases (%), by geographic area.

	Central Europe	Eastern Europe	Mediterranean	Scandinavian
Infectious diseases	83.60	66.52	76.74	93.93
Meningitis	0.72	1.22	0.52	1.12
Asthma	2.01	0.55	1.14	2.89
Diarrhea	1.61	0.71	0.49	1.19
Respiratory problems	3.09	2.44	2.70	1.84
Allergies	2.99	0.79	2.26	5.70
Speech impairment	0.62	0.41	0.24	1.22
Heart trouble	1.36	1.36	0.28	0.68
Ear problem	2.77	2.21	1.75	5.02
Headaches or migraines	5.27	2.82	2.84	5.95
Fractures	8.39	7.46	4.86	9.35
Appendicite	11.44	4.38	7.43	6.69
Depression	1.33	0.64	0.55	1.93
Diabetes	0.06	0.02	0.06	0.02
Cancer	0.01	0.00	0.10	0.02
Epilepsy or seizures	0.42	0.19	0.26	1.14
Leukemia	0.06	0.00	0.00	0.00
GDP per capita	5.111	2.729	2.671	5.614
N	11.128	3.716	7.274	3.917

Table reports disease rates (%) in SHARE, by geographic area: Mediterranean (Greece, Italy, Spain,), Scandinavia (Denmark, Sweden), Central Europe (Austria, Belgium, Germany, France, Netherlands, Switzerland), Eastern Europe (Czech. Republic, Poland) We proxy individuals place of birth by their actual residence. Calculations are based on individuals answers (whether they have experienced or not a certain disease from birth until age 15).

Table 5: OLS estimates number of chronic diseases during childhood (self-reported).

Age	-0.003***	-0.003**	-0.002
	(0.001)	(0.001)	(0.001)
Age^2	-0.001	-0.001	-0.004
	(0.020)	(0.020)	(0.020)
Age^3	-0.000	-0.000	-0.000
	(0.001)	(0.001)	(0.001)
Memory	0.000	-0.001	-0.003
	(0.002)	(0.002)	(0.002)
Education		0.037***	0.013
		(0.014)	(0.014)
Other controls:			
Female, country f.e.	X	X	X
childhood SES proxies			X
N	20290	20290	20290

OLS estimates for number of chronic diseases in childhood. Such variable is constructed by summing up all the experienced diseases. We control for a cubic polynomial in age, a dummy for female and country fixed effects (Italy is the reference country). Memory is a proxy for cognitive ability and is constructed using the outcomes from two cognitive tests: recall first and recall delayed. We use years of schooling as a proxy for education level and consider variables such as number of books in the house when 10, number of rooms per capita and a dummy for breadwinner's low skill occupation when 10 as proxies for family socio-economic status. Standard errors are robust to heteroskedasticity. Significance levels: (*) p-values between 10 and 5 percent; (**) p-values between 5 and 1 percent; (***) p-values less than 1 percent.

Table 6: OLS estimates of adult self-reported health by wave and gender.

	W2 (Male)	W3 (Male)	W2 (Female)	W3 (Female)
childhood SRHS	0.156***	0.186***	0.135***	0.181***
	(0.017)	(0.018)	(0.013)	(0.014)
N	10212	10212	12460	12460
R^2	0.060	0.079	0.099	0.113
Wald test:				
childhood SRHS[W2]-		3.73		11.33
childhood SRHS[W3]=0		(0.0536)		(0.0008)

Logit estimates of adult SRHS: 1 "excellent, very good" and 0 "good, fair, poor"; childhood SRHS 1 "excellent, very good" and 0 "good, fair, poor", by gender and participation in wave 2 and 3. We control for a quadratic polynomial in age. Standard errors are robust to heteroskedasticity. Significance levels: (*) p-values between 10 and 5 percent; (**) p-values between 5 and 1 percent; (***) p-values less than 1 percent.

Table 7: Childhood background proxies by country and average GDP per capita in thousand 1990 international Geary-Khamis dollars.

Country	Rooms p.c.	N. facilities	Low skill occ.	GDP
Austria	0.69 (0.45)	1.49 (1.56)	0.83 (0.38)	3.682
Belgium	0.99 (0.48)	1.80 (1.70)	0.76 (0.43)	5.062
Czech Rep.	0.56 (0.27)	2.30 (1.56)	0.78 (0.41)	3.121
Denmark	0.91 (0.41)	3.00 (1.93)	0.79 (0.41)	5.804
France	0.82 (0.44)	2.13 (1.78)	0.73 (0.44)	4.557
Germany	0.78 (0.40)	2.12 (1.57)	0.75 (0.43)	4.477
Greece	0.54 (0.23)	1.32 (1.40)	0.88 (0.32)	2.174
Italy	0.56 (0.36)	1.43 (1.57)	0.86 (0.35)	3.426
Netherlands	0.80 (0.36)	2.37 (1.13)	0.74 (0.44)	5.441
Poland	0.38 (0.25)	0.62 (1.29)	0.91 (0.29)	2.338
Spain	0.62 (0.41)	1.25 (1.45)	0.88 (0.32)	2.414
Sweden	0.77 (0.41)	3.19 (2.00)	0.75 (0.43)	5.424
Switzerland	0.88 (0.39)	3.14 (1.66)	0.74 (0.44)	7.449

Table 8: Comparison between average year of schooling in 1960 (population 25+) and the proportion of respondents reporting at least one bookcase of books in their childhood accommodation.

Country	Years of schooling	One bookcase
Austria	6.71	.15
Belgium	7.46	.27
Czech Rep.	9.39	.34
Denmark	8.95	.45
France	5.78	.25
Germany	9.48	.29
Greece	4.64	.11
Italy	4.56	.09
Netherlands	5.27	.33
Poland	6.74	.14
Spain	3.64	.14
Sweden	7.65	.37
Switzerland	7.30	.41
Correlation		.69

Table 9: Logit estimates of hunger and financial hardship episodes (marginal effects).

	Hunger			
Before age $\leq$	10	10	15	15
ΔGDP 0-10	-0.725***	-0.587***	-0.399***	
	(0.122)	(0.102)	(0.117)	
WWII 0-10		0.045***		
		(0.011)		
ΔGDP 0-15				-0.864***
				(0.198)
Controls:	quadratic in age, sex, country f.e. and childhood SES			
N	20081	20081	20081	20081
Pseudo R^2	0.132	0.140	0.140	0.142

	Financial Hardship			
Before age $\leq$	10	10	15	15
ΔGDP 0-10	-0.116**	-0.078*	-0.067	
	(0.045)	(0.043)	(0.055)	
WWII 0-10		0.009***		
		(0.004)		
ΔGDP 0-15				-0.195**
				(0.084)
Controls:	quadratic in age, sex, country f.e. and child SES			
N	19943	19943	19943	19943
Pseudo R^2	0.098	0.104	0.100	0.102

Logit estimates for having suffered hunger (financial hardship) between age 0-10 (age 0-15). In each case we control for country average GDP growth rates for years in which individuals were between age 0-10 (age 0-15). Data on GDP per capita are taken from Maddison tables (2010). We include a dummy for World War II (individuals being 10 or 15 years old between 1939–1945). We also control for a quadratic polynomial in age, a dummy for female, country fixed effects and child socio-economic status proxies (number of books, room per capita, etc).

Figure 1: Childhood and adult SRHS.

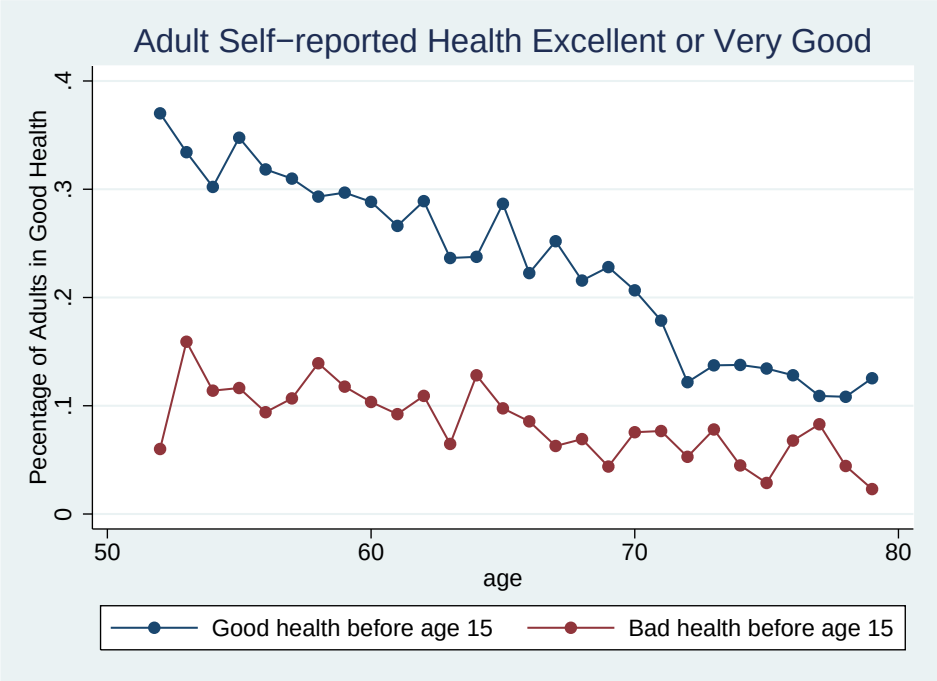

Figure 2: Childhood SES proxies and adult per capita income.

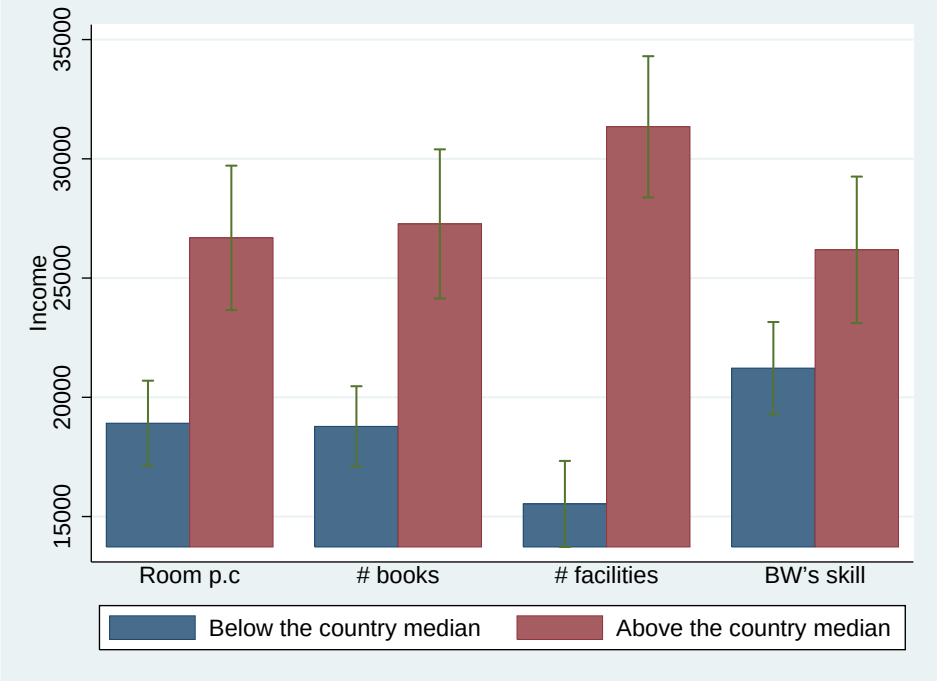

Figure 3: School performance at 10 and years of education.

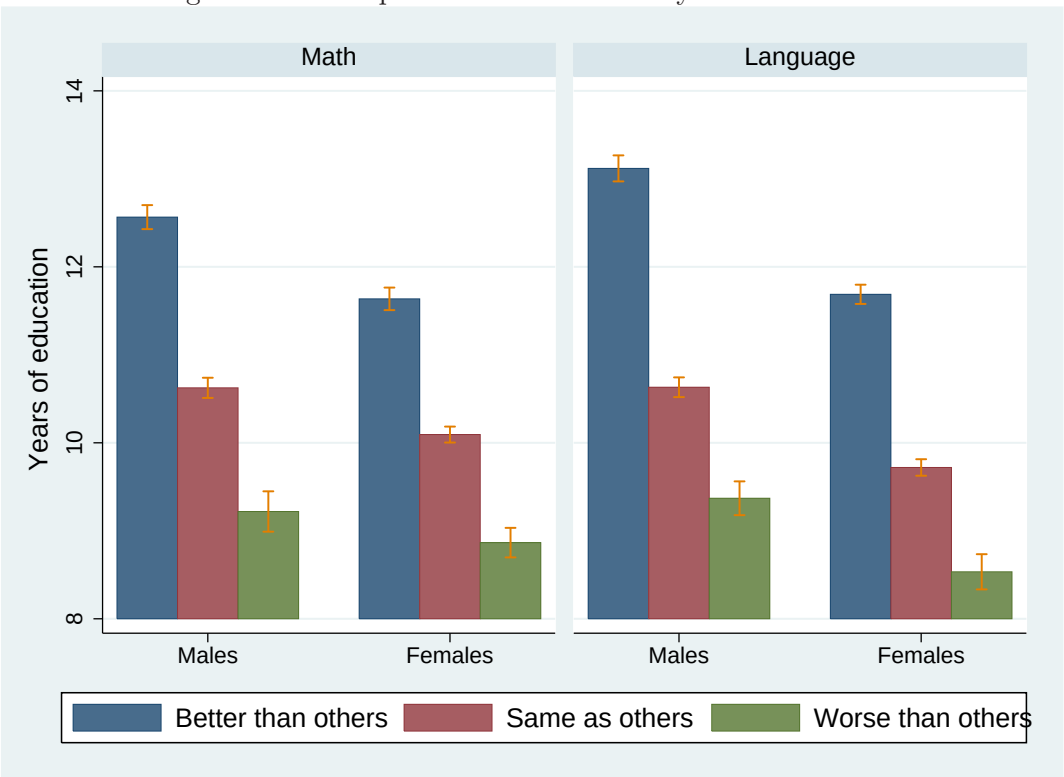

Figure 4: Childhood hunger and financial hardship and adult health.

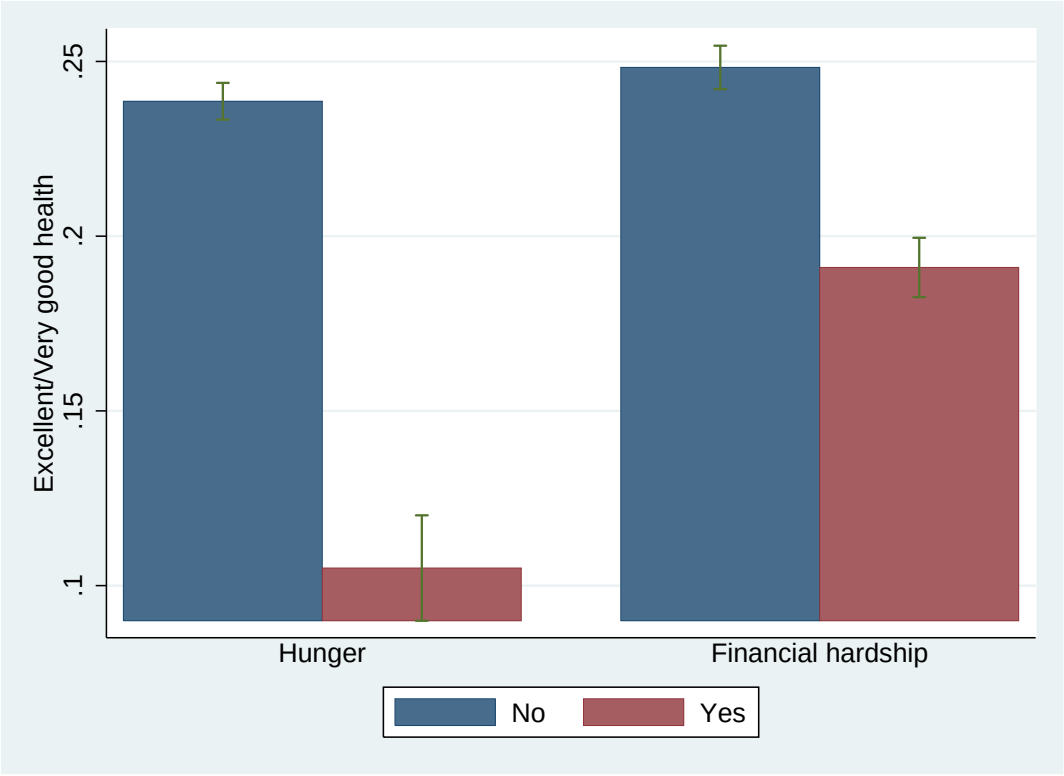

Figure 5: Hunger episodes by countries.