SHARE WORKING PAPER SERIES

Two measures of lifetime resources for Europe using SHARELIFE

Christoph T. Weiss

Working Paper Series 06-2012

Two measures of lifetime resources for Europe using SHARELIFE

Christoph T. Weiss*

March 15, 2012

Abstract

I use the detailed retrospective information provided by the third wave of the Survey of Health, Ageing and Retirement in Europe (SHARE) to estimate two different measures of lifetime resources. The first is a measure of lifetime earnings which corresponds to the income flowing from the asset value of working at age ten. This asset value is the discounted sum of all wages and other benefits earned during the career from age ten until retirement. The second is a measure of lifetime income at age ten which includes both labour and pension income earned over the life cycle. The measure of lifetime income includes expected pension income until death using cohort and country specific mortality tables. I provide graphical evidence that lifetime earnings are positively associated with years of education. I also report descriptive statistics on annual pension benefits, years of education, years of work and number of jobs during the career for individuals surveyed in SHARE.

*Department of Economics, University of Padua, Via del Santo 33, 35122 Padova, Italy. Contact: christoph.weiss@unipd.it. I thank Giorgio Brunello and Guglielmo Weber for helpful discussion and Agar Brugiavini, Danilo Cavapozzi and Elisabetta Trevisan for valuable comments. Financial support from the Antonveneta Centre for Economic Research (CSEA) is gratefully acknowledged. This paper uses data from SHARELIFE release 1, as of November 24th 2010 or SHARE release 2.5.0, as of May 24th 2011. The SHARE data collection has been primarily funded by the European Commission through the 5th framework programme (project QLK6-CT-2001- 00360 in the thematic programme Quality of Life), through the 6th framework programme (projects SHARE-I3, RII-CT-2006-062193, COMPARE, CIT5-CT-2005-028857, and SHARELIFE, CIT4-CT-2006-028812) and through the 7th framework programme (SHARE-PREP, 211909 and SHARE-LEAP, 227822). Additional funding from the U.S. National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, Y1-AG-4553-01 and OGHA 04-064, IAG BSR06-11, R21 AG025169) as well as from various national sources is gratefully acknowledged (see www.share-project.org for a full list of funding institutions). The usual disclaimer applies.

1 Introduction

Most empirical studies in labour economics rely on short-term measures of income even though the object of interest is almost always a longer-term concept. Researchers use current earnings as a proxy for lifetime earnings because they seldom have access to data that span the entire career of workers. Haider and Solon (2006) show that the association between current and lifetime earnings varies over the life cycle. A regression model that uses current income as a proxy for lifetime income produces inconsistent estimates because of this life cycle bias.[1] This implies that the standard errors-in-variables model incorrectly characterizes the relationship between current and lifetime earnings. This misspecification leads to inconsistent estimates of the model coefficients above and beyond the bias due to classical measurement error. That is, there is a bias even when current income is used as the dependent variable. Measures of lifetime earnings are available only for a few countries and they are usually not comparable across countries.[2]

I use the retrospective information provided by the third wave of the Survey of Health, Ageing and Retirement in Europe (SHARE) to estimate a measure of lifetime earnings at age ten which corresponds to the income flowing from the asset value of working at age ten. This asset value is the discounted sum of all net wages and other benefits earned during the career from age ten until retirement using a discount rate of 2%.[3] Individuals who have ever been self-employed are excluded.[4] I also estimate an alternative measure which includes both labour and pension income earned over the life cycle. I name this wider definition of lifetime resources lifetime income. The measure of lifetime income includes expected pension income until death using cohort and country specific mortality tables.

SHARE is a multidisciplinary and cross-national panel database of micro data on health, socio-economic status and social and family networks of more than 25,000 individuals aged 50 or over. The survey's third wave of data collection, SHARELIFE, collects detailed retrospective life and labour market histories in thirteen countries in 2008-09. The survey's first and second waves were conducted in 2004 and 2006-07, respectively. They are a balanced

[1] Mazmumder (2001) shows that the bias is substantial even when an average of earnings over 5 years instead of earnings in one year is used.

[2] Measure of lifetime earnings have been constructed for males for the U.S. (Haider and Solon, 2006) and Norway (Bhuller et al., 2011), and for both females and males for Sweden (Böhlmark and Lindquist, 2006) and Germany (Brenner, 2010).

[3] Haider and Solon (2006), Böhlmark and Lindquist (2006) and Brenner (2010) also assume a constant real interest rate of 2% to construct a measure of lifetime earnings. Bhuller et al. (2011) use instead an interest rate of 2.3%.

[4] Murphy and Welch (1990) also exclude the self-employed in their analysis of age-earnings profiles.

representation of the various regions in Europe, ranging from Scandinavia (Denmark and Sweden) through Central Europe (Austria, Belgium, France, Germany, Poland, the Czech Republic, the Netherlands and Switzerland) to the Mediterranean (Greece, Italy and Spain).

An alternative to retrospective data are national longitudinal studies (e.g. British cohort studies) where individuals are followed throughout their lives and surveyed at multiple points in time. However, prospective data are much more expensive to collect and are only available for few countries or a short time period. The validation studies by Havari and Mazzonna (2011) and Garrouste and Paccagnella (2011) find that recall bias is not severe in SHARELIFE data, possibly because of the state-of-the-art elicitation techniques used: respondents are helped to locate events along the time line, starting from domains that are more easily remembered, and then asked progressively more details about them.[5]

Recall data on earnings collected in surveys are of course subject to measurement error, often not of the classical type. Bound et al. (2001) and Kapteyn and Ypma (2007) provide reviews of validation studies on the relation between earnings in survey data and administrative data. Several studies find a negative correlation between the true value of earnings and measurement error (e.g. if low earners tend to overreport their earnings in survey data).[6] In the Panel Study of Income Dynamics Validation Study (PSIDVS), Pischke (1995) finds a weak negative correlation for hourly earnings but no correlation for monthly earnings.

Administrative data are often based on different administrative files. They may be prone to measurement error because of the problem of mismatch when linking an observation across datasets. This implies that they do not necessarily contain the true value of earnings for each individual. Kapteyn and Ypma (2007) discuss the problem of measurement error in both administrative and survey data by linking Swedish males who were surveyed in both SHARE and LINDA (a Swedish Longitudinal Individual Data Base). They find that ignoring the possibility of mismatches in administrative data can lead to substantial biases. Meijer et al. (2011) generalize the mixture factor analysis model of Kapteyn and Ypma (2007). They show that even when the probability of a mismatch is small, administrative data can perform very poorly. They argue that survey data are sometimes more reliable despite their higher measurement error.

Administrative data can have other flaws that go beyond mismatch as well. They could

[5] In their study of the effect of childhood environment on economic and social outcomes for Yemenites who immigrated to Israel in 1949, Gould et al. (2011) find evidence that retrospective data on childhood environment from more than 50 years ago is of high quality.

[6] Assuming that administrative data contain the true value of earnings, measurement error is defined as the difference between the measure reported in survey data and the true value.

deviate from the true value of income even when there is no mismatch (e.g. if someone is working in the underground economy). For instance, Haider and Solon (2006) use Social Security earnings histories of participants in the U.S. Health Retirement Study (HRS) for the period 1951-1991 to recover a measure of lifetime earnings. But their earnings data are only available for jobs covered by U.S. Social Security and in some years a large proportion of the sample (sometimes above 50%) are right-censored because of the Social Security taxable limit for that year.

Administrative data are usually collected for a specific purpose (e.g. social security contributions for entitlements to pension benefits) and are therefore limited in the amount of information collected. Overall, the main advantage of survey data such as SHARE over administrative data is that it obtains rich information on family background, social environment, physical and mental health, educational attainment, labor market outcomes and that there is no risk of mismatch across datasets. Moreover, in SHARE, rigorous procedural guidelines and programs ensure that the data is comparable across countries.

The reminder of this paper is organized as follows. In the next section, I present the methodology used to recover the measures of lifetime earnings and lifetime income. Section 3 provides some descriptive statistics across countries on the key variables. The last section concludes.

2 Methodology

The first step is to compute the length of each employment spell. When the years at the beginning and at the end of the spell are identical, I assume that the individual spent an entire year in the job, i.e. working from January 1 to December 31. When the years are different, I assume that they started and stopped working in the same month, e.g. working from March 1984 to March 1996. This implies that someone who reports to have started working in an employment spell in 1984 and stopped in 1984 will be treated equally to someone who started in 1984 but stopped in 1985.

In SHARELIFE, survey participants are asked to report the amount they were paid monthly after taxes each time they started an employment spell. They are also asked the monthly new wage in their current job (if they are still working) and the monthly net wage at the end of the main job in their career (if they have already retired). Whenever the current income is missing, I use the income measure from the imputation module in wave 2 (if the current employment spell started before the interview year of wave 2) or from wave 1 (if the current employment spell started before the interview year of wave 1). The imputation modules in waves 1 and 2 contain a measure of annual net income from employment (or self-employment) in the previous year.[7]

All wages and pension benefits are transformed using PPP exchange rates and CPI measures into 2006 Euro. PPP-adjusted exchange rates and CPI measures are taken from the OECD and national sources.[8] I only keep annual wages that are below the 99th percentile or above the 1st percentile of the empirical wage distribution. I proceed in a similar fashion for pension benefits. The trimming should not create too much harm to the data. One could (or should) of course consider other floors or ceilings.

2.1 Imputing missing wages and predicting wages at the end of an employment spell

I impute missing wage values using predictive mean matching. This is done separately for females and males. Predictive mean matching method is an imputation method used for continuous variables and is similar to a regression method. It finds the observation whose predicted value are closest to the predicted value of the missing observation but uses the observed value for the imputation.[9] Imputed values are obtained by regressing annual wages on the following list of variables: ISCED education level (3 different levels), birth cohort (3 cohorts), decade of the start of the employment spell (4 different decades), whether the worker is a white collar during the spell, whether he worked part-time during the spell, and country. Approximately 30% of the wage values are imputed. Unsurprisingly, there are more missing values for employment spells that were started in earlier decades. They are also more missing values for females than for males.

In SHARELIFE, individuals are asked the amount of their monthly net pay at the start of each employment spell. They are not asked how much they were paid at the end of the spell except for the main spell in their career (if they have retired) or the current employment spell

[7]See Christelis (2011) for more details on the imputed variables in SHARE.

[8]More details can be found in Trevisan et al. (2011). I do not consider individuals from Poland because of unreliable income data. Trevisan et al. (2011) argue that Poles answering the SHARE questionnaire got confused between new and old Zloty around the devaluation in 1995 and misreported earnings during the hyper inflation of the 80s and 90s.

[9]One can also draw at random from a set of observed values whose predicted values are close to the one of the observation with missing value. See, e.g., Heitjan and Little (1991), Schenker and Taylor (1996) and Horton and Kleinman (2007).

(if they are still working). That is, only the current and the main employment spells have wage measures both at the beginning and the end of the spell. I predict wages at the end of an employment spell using potential labour market experience as a running variable. Potential experience is defined as $A_t - S - IS1$ where A_t denotes age in year t, S years of education and $IS1$ age at school entry. Wage predictions are done separately for females and males. For males, I regress the logarithm of the current wage on potential experience, potential experience squared and a series of characteristics: years of education, an occupation dummy (white-collar job), industry dummies (agriculture, manufacturing, services, public sector, community services) and the interaction of these characteristics with potential experience. For females, I regress the logarithm of the wage at the end of the main job of the career on the same list of covariates.

I also control for characteristics that are constant over the life cycle: country, 3 birth cohorts, an indicator of the number of books at age ten in the place where the individual was living (excluding magazines, newspapers or school books)[10], whether the individual was better (or much better) to others in mathematics at age ten (as opposed to about the same, worse or much worse), whether the individual was better (or much better) to others in the country's language at age ten (as opposed to about the same, worse or much worse), the features of the accommodation at age ten (5 indicators for whether or not the accommodation had a fixed bath, cold running water supply, hot running water supply, inside toilet and central heating), and an indicator of the number of rooms occupied by the household divided by the number of people living in the household at age ten[11]. I estimate the following model by OLS

$$y_c = \beta_1 E_c + \beta_2 E_c^2 + \beta_3 E_c S + \beta_4 E_c X_c + \beta_5 S + \beta_6 X_c + \beta_7 W + U$$

where y_c denotes the logarithm of current wage for males (or the wage at the end of the main job in the career for females), E potential experience, S years of education, X characteristics that are specific to an employment spell (i.e. white-collar job and industry), W characteristics that are constant over the life cycle, and U is a disturbance term. I then use the wage at the beginning of employment spell j (which is typically observed) and the estimated coefficients from the regression on the current spell c to predict the wage at the

[10] The indicator takes value one if people report to have less than 10 books ("less than a shelf") in the household at age ten.

[11] The indicator takes value one if the number of rooms occupied by the household at age ten (including bedrooms but excluding kitchen, bathrooms, and hallways) divided by the number of people living in the household is equal or higher to one, and zero otherwise. That is, I compute whether there are more rooms than people in the household at age ten.

end of employment spell j

$$\hat{y}_{t1j} = y_{t0j} + b_1 \left(E_{t1j} - E_{t0j}\right) + b_2 \left(E_{t1j}^2 - E_{t0j}^2\right) + b_3 \left(E_{t1j}S - E_{t0j}S\right) + b_4 \left(E_{t1j}X_j - E_{t0j}X_j\right)$$

where $\hat{y}_{t1j}$ is the predicted logarithm of wage at the end of the spell, y_{t0j} is the logarithm of the observed (or imputed) wage at the beginning of spell, E_{t1j} and E_{t0j} denote potential experience at the end and the beginning of the spell respectively. Armed with the wages at the beginning and the end of the spell, I compute the annual growth of earnings during an employment spell as follows: $\left(\hat{y}_{t1j} - y_{t0j}\right)/len_j$ where len_j denotes the length of the employment spell. I use this growth rate to generate annual earnings in each employment spell.

To check the accuracy of the procedure, I apply it to the current and main employment spells, for which there is information on wages at both the beginning and the end of the employment spell. Tables 1 and 2 compare the predicted with the actual values and show that the predicted wage values are close to the reported values. For males, the predicted current income is closer to the reported value because the estimated coefficients used for the prediction are taken from a regression on current income. For males, the hypothesis that the predicted mean value of main income is equal to the mean observed value is rejected at the 1% significance level. For the other 3 other predicted values, I cannot reject the hypothesis that the predicted mean values are equal to the observed ones.

2.2 Validation of the wage prediction procedure

To provide some evidence on the validity of the procedure to predict wages, I use data from the German Socio-Economic Panel Study (SOEP). The SOEP is a longitudinal panel dataset of the population in Germany which started in 1984. It obtains information on household composition, occupation, employment, earnings, health and life satisfaction. I use annual data from 1984 to 2008. SOEP data are integrated into the Cross National Equivalent File (CNEF) which contains equivalently defined variables for panel databases from the UK (BHPS), Australia (HILDA), South Korea (KLIPS), the U.S. (PSID), and Canada (SLID). I use the variables in the CNEF file for the validation study.

I define potential labour market experience in SOEP as age - schooling - 6 (the age at school entry in Germany) and estimate a model which is very similar to the one above

$$y_c = \beta_1 E_c + \beta_2 E_c^2 + \beta_3 E_c S + \beta_4 E_c X_c + \beta_5 S + \beta_6 X_c + \beta_7 W + U$$

where y_c denotes the logarithm of individual labor earnings in 2008, E potential experience, S education, X characteristics that are specific to an employment spell (i.e. white-collar job and industry), W characteristics that are constant over the life cycle (birth cohort dummy), and U is a disturbance term. Wages are deflated using the CPI and are in 2005 prices. The analysis is done separately for males and females. For both gender, the sample for the regression consists of all people born between 1945 and 1956 who provide information on individual labor earnings, age, schooling, occupation and industry in 2008. I then use the wage in 1984 (the first year in SOEP) and the estimated coefficients from the regression on the wage in 2008 to predict the wage in each year from 1985 to 2008

$$\hat{y}_t = y_{t0} + b_1 \left(E_t - E_{1984}\right) + b_2 \left(E_t^2 - E_{1984}^2\right) + b_3 \left(E_t S - E_{1984} S\right) + b_4 \left(E_t X_j - E_{1984} X_{1984}\right)$$

where $\hat{y}_t$ is the predicted logarithm of wage in year t and y_{1984} is the logarithm of the wage in 1984. The variables contained in this model are very similar to the one used for the wage predictions using SHARE data where the dependent variable was also the wage in 2008. The main difference is that we are now focusing on one country - Germany - and that we do not include covariates describing early life conditions as they are not available in SOEP.

Table 3 and 4 report the means of the observed wage, the predicted wage, the prediction error, and p-value for a two-sample T-test with unequal variances using SOEP data. The sample used for for each year from 1985 to 2008 consists of all individuals who report information in 1984 (the starting year) and year t. Table 3 suggests that for males the hypothesis that the mean predicted wage and the mean true wage are equal is rejected only for 4 years (2003 to 2006). Table 3 shows that for females the same hypothesis is not rejected in any year. It should be noted that unlike in SHARE data, the standard deviations of the observed wage and predicted wage are of very similar magnitude, both for males and females. More importantly, the prediction procedure also captures the concavity in average wage profiles for the cohorts of interest. This implies that the average wage profiles are rather well predicted for most years during the period 1985-2008 for both males and females in SOEP. This evidence is quite reassuring and suggests that the procedure that uses experience as a running variable to estimate wages at the end of a job is rather accurate.[12]

[12]I repeated the same exercise using data from the Italian Survey of Household Income and Wealth (SHIW). Results for males are very similar to the ones using the German SOEP. They are not reported here but are available from the author upon request.

2.3 Earnings during the career

All monthly wages are multiplied by 12. They are annualized because the time period for an employment spell is expressed in years while the earnings are in months. Of course, in some countries, some individuals are paid 13 or 14 months of salary per year. This is ignored here. For each individual, the discounted sum of all annual incomes is

$$A = r \cdot \sum_{j=1}^{J} Y_j \sum_{k=1}^{K} (1 + gr_j)^{(k-1)} / (1 + r)^{(S_j - (BY+11)+k)}$$

where j refers to job (or employment spell) j, J is the total number of jobs, Y_j are annual earnings at the beginning of employment spell j, k refers to the year in the employment spell, K is the total length of each employment spell (in years), $1 + gr_j$ the annual growth rate of earnings during the employment spell j, S_j the year when the employment spell j started, BY is the year of birth and r is the interest rate. To illustrate, if someone is born in 1940 and starts working in 1950, the first wage in 1950 is not discounted, in the wage in 1951 is discounted with $1 + r$, in 1952 with $(1 + r)^2$ and so forth. While the first wage is reported by the individual, subsequent earnings are predicted as above.

To create a measure of lifetime income, I then add to this discounted sum of all annual incomes until the interview year the discounted flow all expected incomes and pension benefits until age 110 (the time of death for all individuals). Each annual income and pension benefit received after the interview year of SHARELIFE is multiplied by the survival probability within elementary age interval, i.e. 1 - the probability of death between age_t and age_{t+1}. This is because I am considering income and pension benefits that will be received in the future. I hence need to take into account the possibility that individuals may die after the interview. Of course, all individuals in our sample have survived up to the interview year. The probability of death qx_m varies across country, gender, year of birth and age. The data on life expectancy are based the Human Mortality Database (Department of Demography at the University of California, Berkeley and Max Planck Institute for Demographic Research). I make some modifications to the raw data. As Austria does not have data before 1947, all individuals born before 1947 are assigned the probability of death of individuals born in the year 1947. Data from Belgium replace Germany (German data start in 1991) and the Czech Republic (no data).

2.4 Lifetime resources for those who have retired

Some people are still working at the time of their interview, others have already retired. For those who have retired before the interview year of SHARELIFE, the rest of their lifetime income consists of pension benefits and expected pension benefits. This income flow is given by

$$B_{ret} = r \cdot pension \cdot \sum_{l=1}^{L} 1/(1+r)^{(RET-(BY+11)+l)}$$

where *pension* refers to the annual pension benefits that are currently received, l refers to the year in the current pension spell, L is the length of the current pension spell (i.e. the difference between the interview year and the retirement year), RET is the retirement year, BY is the year of birth and r is the interest rate. I assume that pension benefits do not increase or decrease during retirement. Similarly, expected pension benefits (from interview year until age 110) are computed as

$$D_{ret} = r \cdot pension \cdot \sum_{m=1}^{M} (1 - qx_m)/(1+r)^{(INT-(BY+11)+m)}$$

where m refers to each year spent in the expected pension spell, M is the expected length of the pension spell (i.e. the difference between age 110 and the age at the time of the interview), qx_m is the probability of death within elementary age interval $[BY+m, BY+m+1)$, INT is the interview year, BY is the year of birth and r is the interest rate. I assume that current pension benefits are expected to be continuously received until death.

I use information from SHARELIFE on monthly benefits after tax from social security or pensions, i.e. the sum of all pensions (public, occupational or private). I multiply the monthly benefits by 12 to obtain annual measures. When the sum of pension benefits is missing or below/above the trimming thresholds (1st and 99th percentiles), I use information from wave 2. I compute the sum of all annual pension benefits reported in wave 2. I include public old age pension, public early or pre-retirement pension, public disability insurance, public unemployment benefit or insurance, public survivor pension from partner, war pension, private (occupational) old age pension, private (occupational) early retirement pension, private (occupational) disability insurance, private (occupational) survivor pension from partner's job, public old age supplementary pension, secondary public disability insurance pension, secondary public survivor pension from spouse/partner, occupational old age pension from a second job, occupational old age pension from a third job, and private (occupational) disability insurance. I only keep values below the 99th percentile of above the 1st percentile. I then use this measure to replace the missing values of pension benefits that could not be recovered using SHARELIFE. When the sum of pension benefits is still missing, I use information from wave 1. I include public old age pension, public early or pre-retirement pension, public disability insurance, public unemployment benefit or insurance, public survivor pension from partner, public invalidity or incapacity pension, war pension, private (occupational) old age pension, private (occupational) early retirement pension, private (occupational) disability insurance, and private (occupational) survivor pension from partner's job. Similarly, I only keep values below the 99th percentile or above the 1st percentile. I then use this measure to replace the missing values of pension benefits that could not be recovered up to this point. The remaining missing values for pension benefits (approx. 5% of the sample) are imputed using predictive mean matching. I regress pension on ISCED education level (3 different levels), birth cohort (3 cohorts), decade of the retirement year (4 different decades), and country.

2.5 Lifetime resources for those who are still working

For the individuals who are still working, I do not have information on the exact amount of pension benefits that they will receive after retirement. But I know from previous SHARE waves their expected retirement age and replacement rate. The rest of their lifetime income consists of expected income until expected pension age and expected pension benefits from expected pension age until death. To compute a measure of lifetime income that includes all working episodes over the life cycle for all individuals, I create a new artificial employment spell that should correspond to the last employment spell until retirement. Obviously, for those who have already retired, the length of this artificial employment spell is equal to zero. For those who are still working, the length of the employment spell is the difference between the age at which they expect to collect pension benefits and their current age. If these two ages are equal, I assume that they retire immediately and start collecting pension benefits.

In this future employment spell, I assume that all individuals who are still working at the time of the interview in SHARELIFE will continue working until their expected retirement age. That is, they will not stop working before retirement age and will never be unemployed until retirement. I also assume that individuals who are still working but have passed the retirement age will immediately stop working and retire. I predict the wage at the end of this spell in a similar fashion to the way I predict the wage at the end of each employment spell. I then compute the growth of income from interview year of until retirement year. When individuals do not report at what age they will start collecting pension, I use information on statutory retirement age in their country.[13] Sometimes the statutory retirement also varies across gender within a country.

I compute the discounted sum of expected incomes until expected pension age for each individual as

$$C_{work} = r \cdot Y_{curr} \cdot \sum_{t=1}^{T} (1+gr)(1-qx_t)/(1+r)^{(INT-(BY+11)+t)}$$

where Y_{curr} refers to current earnings, t is each year spent in the current employment spell until expected pension age, T is the expected length of the artificial employment spell (i.e. the difference between the expected pension age and the interview year), $1+gr$ the annual growth rate of income during the employment spell, qx_t is the probability of death within elementary age interval $[BY+t, BY+t+1)$, INT is the interview year, BY is the year of birth and r is the interest rate. Similarly, expected pension benefits (from expected pension age until age 110) are given by

$$D_{work} = r \cdot reprate \cdot Y_{curr} \cdot \sum_{v=1}^{V} (1-qx_v)/(1+r)^{(PY-(BY+11)+v)}$$

where $reparte$ refers to the replacement rate (or percentage of salary received as pension), Y_{curr} refers to current earnings, t refers to each year spent in the expected pension spell, V is the expected length of the retirement spell (i.e. the difference between age 110 and the expected pension age), qx_v is the probability of death within elementary age interval $[BY+v, BY+v+1)$, PY is the expected retirement year (the year in which the individual will start receiving pension benefits), BY is the year of birth and r is the interest rate. In the formula above, I use current wage and not the predicted wage at expected retirement age because individuals are asked what is the percentage of current wage that will be received as pension (and not the percentage of expected wage at retirement age).

The expected percentage of salary received as pensions are reported by the individuals who are working in wave 2 of SHARE. I take the sum of all expected percentages for each type of pension and consider those in the range between 50% and 100%. When values from wave 2 are missing, I use data from wave 1. I fill the remaining missing values at the individual level by computing the median replacement rate within each country, gender and 3 birth cohorts. The replacement rate multiplied by the current income should be a good approximation to the pension benefits. As an alternative measure for the replacement rate, I could compute

[13]Information on statutory retirement age is mainly taken from Angelini et al. (2009).

the ratio of pension benefits over the main wage in the career. I obtain this information from the individuals who have already retired. I could then use the median of this ratio by country and apply it to the individuals who are still working. But I prefer to use information provided by individuals who are working and report their own expected percentage of salary received as pensions. I believe that this is superior to using information reported by individuals who have retired in different years and at different ages.

2.6 Adding widow's, widower's or surviving civil partner's pension

If two individuals are leaving together and one of the partner dies, the surviving partner can receive a survivor pension for the rest of his life (unless he remarries or cohabits with another person). The pension is usually payable regardless of other income. I set the amount of the expected future survivor pension to 60% of the pension of the partner. If both partners are retired, they immediately start receiving expected future survivor pension benefits. If the individual is retired but his wife is still working, he receives the expected future survivor pension benefits only after the partner retires (i.e. from the expected pension year of the partner). If both are working, they have to wait until the last one retires to receive expected future survivor pension benefits from partner. I match couples using the variable on household identifier.

Expected survivor pension from partner for an individual who is retired and whose partner is also retired (from interview year until age 110) is given by

$$E_{ret} = 0.6 \cdot r \cdot pension_{part} \cdot \sum_{m=1}^{M} (1 - qx_m)\, qx_{(m,part)} / (1+r)^{(INT-(BY+11)+m)}$$

where $pension_{part}$ refers to the annual pension benefit currently received by the partner, m refers to each year in the expected pension spell, M is the expected length of the pension spell (i.e. the difference between age 110 and the age at the time of the interview), qx_m is the probability of death within elementary age interval $[BY + m, BY + m + 1)$, $qx_{(m,part)}$ is the partner's probability of death of the within elementary age interval $[BY_{part} + m, BY_{part} + m + 1)$, INT is the interview year, BY is the year of birth and r is the interest rate.

Expected survivor pension from partner for an individual who is retired and whose partner is still working is given by

$$E_{ret} = 0.6 \cdot r \cdot reprate_{part} \cdot Y_{(curr,part)} \cdot \sum_{m=1}^{M} (1 - qx_m)\, qx_{(m,part)}/(1+r)^{(PY_{part}-(BY+11)+m)}$$

where $reparte_{part}$ refers to the replacement rate (or percentage of salary received as pension) of the partner, $Y_{(curr,part)}$ refers to the the partner's current annual income (which is expected to be continuously received by the partner until her pension age), m refers to the year in the expected pension spell, M is the expected length of the partner's pension spell (i.e. the difference between age 110 and the age at which the partner is expected to retire), qx_m is the probability of death within elementary age interval $[BY + m, BY + m + 1)$, $qx_{(m,part)}$ is the partner's probability of death within elementary age interval $[BY_{part} + m, BY_{part} + m + 1)$, PY_{part} is the partner's expected retirement year (the year in which the partner will start receiving pension benefits), BY is the year of birth and r is the interest rate. That is, $reprate_{part} \cdot Y_{(curr,part)}$ corresponds to the expected pension benefits of the partner.

Expected survivor pension from partner for an individual who is still working and whose partner has already retired is given by

$$E_{work} = 0.6 \cdot r \cdot pension_{part} \cdot \sum_{m=1}^{M} (1 - qx_m)\, qx_{(m,part)}/(1+r)^{(PY-(BY+11)+m)}$$

where $pension_{part}$ refers to the annual pension benefit currently received by the partner, m refers to the year in the expected pension spell, M is the expected length of the individual's pension spell (i.e. the difference between age 110 and the age at which the individual is expected to retire), qx_m is the probability of death within elementary age interval $[BY + m, BY + m + 1)$, $qx_{(m,part)}$ is the partner's probability of death within elementary age interval $[BY_{part} + m, BY_{part} + m + 1)$, PY is the expected retirement year (the year in which the individual will start receiving pension benefits), BY is the year of birth and r is the interest rate. That is, $reprate_{part} \cdot Y_{curr,part}$ corresponds to the expected pension benefits of the partner.

Expected survivor pension from partner for an individual who is still working and whose partner is also still working is given by

$$E_{work} = 0.6 \cdot r \cdot reprate_{part} \cdot Y_{(curr,part)} \cdot \sum_{m=1}^{M} (1 - qx_m) qx_{(m,part)}/(1+r)^{(\max(PY,PY_{part})-(BY+11)+m)}$$

where $reparte_{part}$ refers to the replacement rate (or percentage of salary received as pension) of the partner, $Y_{(curr,part)}$ refers to the the partner's current annual income (which is expected to be continuously received by the partner until her pension age), m refers to the year in the expected pension spell, M is the expected length of the individual's pension spell (i.e. the difference between age 110 and the age at which the individual is expected to receive partner's pension benefits), qx_m is the probability of death within elementary age interval $[BY + m, BY + m + 1)$, $qx_{(m,part)}$ is the partner's probability of death within elementary age interval $[BY_{part} + m, BY_{part} + m + 1)$, PY is the expected retirement year (the year in which the individual will start receiving pension benefits), PY_{part} is the partner's expected retirement year (the year in which the partner will start receiving pension benefits), BY is the year of birth and r is the interest rate. That is, $reprate_{part} \cdot Y_{(curr,part)}$ corresponds to the expected pension benefits of the partner. The year in which the individual is expected to receive partner's pension benefits depends on whether the partner retires before or after the individual, i.e. they have to wait until the last one has retired in order to receive partner's pension benefits. This explains why the expression $\max(PY, PY_{part})$ is in the denominator of the equation above.

2.7 Sums of lifetime resources

For the individuals who have retired at the time of the interview, their lifetime income is given by

$$NPV_{10} = A + B_{ret} + D_{ret} + E_{ret}$$

where NPV_{10} refers to the net present value at age 10 of all earnings, pension benefits and expected pension benefits earned until death. Their measure of lifetime earnings is simply A, the first component of the sum above. For the individuals who are still working at the time of the interview, their lifetime income is given by

$$NPV_{10} = A + C_{work} + D_{work} + E_{work}$$

where NPV_{10} refers to the net present value at age 10 of all earnings, expected earnings and expected pension benefits earned until death. Their measure of lifetime earnings is equal to the sum of A and C_{work}, the first two components in the sum above. I only retain in the final dataset those individuals with an estimated lifetime income between the 1st percentile and the 99th percentile of the distribution of lifetime income.

3 Some descriptive statistics

Some descriptive statistic (mean, standard deviation, median and interquartile range) at the country level for the measure of lifetime earnings net of pensions are reported for males in Table 5. Swiss men earn the most during their career on average. The lowest average lifetime earnings are found in Greece and the Czech Republic. Table 6 shows the same set of descriptive statistics for females. Average lifetime earnings are lower for women than for men in all countries. Swedish women are on top of their league as they earn a bit more than Swiss women on average. The Czech Republic and Spain have the lowest average lifetime earnings for females. Figures 1 and 2 are box plots of lifetime earnings and give some more information on the distribution of lifetime earnings in each country. Tables 7 and 8 provide descriptive statistics for the measure of lifetime income which includes pension and expected pension benefits. The ranking across countries is very similar to the previous measure, both for males and females. The correlation coefficient between the two measures of lifetime resources is very high, at 0.95, and is almost constant across gender or country.

The median age for males in the sample is 66 (ranging from a median age of 64 in Denmark, France and East Germany to 68 in Italy), while it is 64 for females (ranging from 60 in Greece to 67 in Austria). 69.3% of men have already retired (ranging from 56.1% in Denmark to 85.1% in Austria), while the same statistic is 71.9% for women (ranging from 59.7% in Denmark up to 88.4% in Austria). Median retirement age for men is at 61 (ranging from age 58 in Italy to age 65 in Sweden and Switzerland), whereas median retirement age for women is at 60 (ranging from age 57 in Austria, the Czech Republic and Greece to age 65 in Denmark and Sweden). Average annual pension benefits are lower for women than for men in all countries. Table 9 suggests that average annual pension benefits for Swiss men, at 23,204 Euro 2006, are almost three times as high as what is received in the Czech Republic. Average annual pension benefits for Czech women, at 6,717 Euro 2006, are a bit more than half of what is received by Swiss women, as shown in Table 10.

People who are still working will retire at a later age than people who have already retired. The median age at which men expect to retire is 65 (ranging from age 60 in France to 65 in Germany, Greece, Italy, the Netherlands, Spain, Sweden and Switzerland). Women who are still working expect to retire at a median age of 63 (ranging from age 60 in Austria, Belgium, the Czech Republic, France, Greece and Italy to 65 in Germany, the Netherlands, Spain and Sweden). People who are still working also expect to receive higher pension benefits than those who have already retired. Table 11 reports expected annual pension benefits for men and shows that Swiss men expect to receive more than 3 times what it is expected by men in the Czech Republic. As shown in Table 12, Czech women expect to receive as pension benefits 40% of what will be received by women in Denmark and Switzerland.

The average years of education for men in the sample is at 11.29, as shown in Table 13. German men are the most educated as they went to school for more than 13 years, while men in Italy and Spain only attended school for less than 9 years on average. The average educational attainment is lower for females, at 10.85 years. Table 14 shows that women in Austria, Italy and Spain spent less than 9 years in school on average, while Danish and German women went to school for more than 12 years.

Figure 3 shows kernel density estimates of lifetime earnings net of pension benefits for four education categories: 0 to 8, 9 to 11, 12 to 14, and 15 to 25 years of education. The top panel is for males and the bottom panel for females. Each education category contains approximately one quarter of the sample. The kernel density estimates suggest that there is a positive association between education and lifetime earnings: there is a clear difference between people in the first and the fourth categories. However, people in the second and the third categories have similar distributions of lifetime earnings. There is a potential problem of composition for the sample of each education category: older cohorts or some countries are likely to be more represented in the lowest educational categories. I hence regress the logarithm of lifetime earnings net of pension benefits on country and cohorts dummies. This allows me to recover a measure of log lifetime earnings net of country and cohort effects. Figure 4 shows the mean of log lifetime earnings net of country and cohort effects for each year of education.[14] The 95% confidence interval based on the corresponding standard errors is smaller for males than for females. The pattern in the association between education and lifetime earnings is less stable for females than for males but both patterns confirm the positive association presented in Figure 3.

Table 15 reports that men have worked 34.14 years on average (from less than 32 years in Denmark, France and Italy up to almost 38 years in the Czech Republic). Note that years of work for those who are still working also include the expected number of years of work until retirement. Females work less than males in all countries. In Table 16, they have worked (or will work until retirement) 24.22 years on average (less than 19 years in the Netherlands and Spain, more than 35 years in the Czech Republic). Table 17 shows that there is some variation in job mobility across countries. The average number of jobs held by males during the career is 2.92 but it goes from almost an unique job for Greek workers to more than 4 jobs in Denmark. There is not much variation in the average number of jobs across gender

[14] For these calculations, individuals with more than 21 years of education (less than 1% of the sample) have been assigned 21 years of education.

within a country. As reported in Table 18, the average number of jobs for women in the sample is 2.72, ranging from less than 2 jobs for Greek, Italian and Spanish women to more than 4 jobs for Danish women.

4 Conclusion

Most empirical studies in labour economics rely on short-term measures of income even though the object of interest is almost always a longer-term concept. Researchers use current earnings as a proxy for lifetime earnings because they seldom have access to data that span the entire career of workers. I use the retrospective information provided by the Survey of Health, Ageing and Retirement in Europe (SHARE) to estimate a measure of lifetime earnings at age ten which corresponds to the income flowing from the asset value of working at age ten. This asset value is the discounted sum of all wages and other benefits earned during the career from age ten until retirement using a discount rate of 2%. I also estimate an alternative measure which includes both labour and pension income earned over the life cycle. I name this wider definition of lifetime resources lifetime income. The measure of lifetime income includes expected pension income until death using cohort and country specific mortality tables. I also report some descriptive statistics at the country level on the two measures of lifetime resources and on pension benefits, years of education, years of work and number of jobs during the career.

Recall data coming from a survey such as SHARE is likely to be affected by measurement error. Kapteyn and Ypma (2007) have linked the Swedish subsample of SHARE with administrative data from LINDA (a Swedish Longitudinal Individual Data Base). Böhlmark and Lindquist (2006) have created a measure of lifetime income for men and women in Sweden using data from LINDA. The possibility to link the measure of Böhlmark and Lindquist (2006) with a Swedish subsample of SHARE could provide some evidence on the validity of the measure of lifetime earnings net of pension benefits developed in this paper. More recently, the SHARE-RV project has linked administrative data from the German Pension Fund (Deutsche Rentenversicherung) with the German subsample of SHARE using the social security number. The combination of administrative data collected for a specific purpose (such as social security contributions for entitlements to pension benefits) with rich information from survey data about many different aspects of respondents' working histories could provide some further possibilities to validate the quality of the recall data collected in SHARE.

Figure 1: Box plots of lifetime earnings net of pension benefits, males

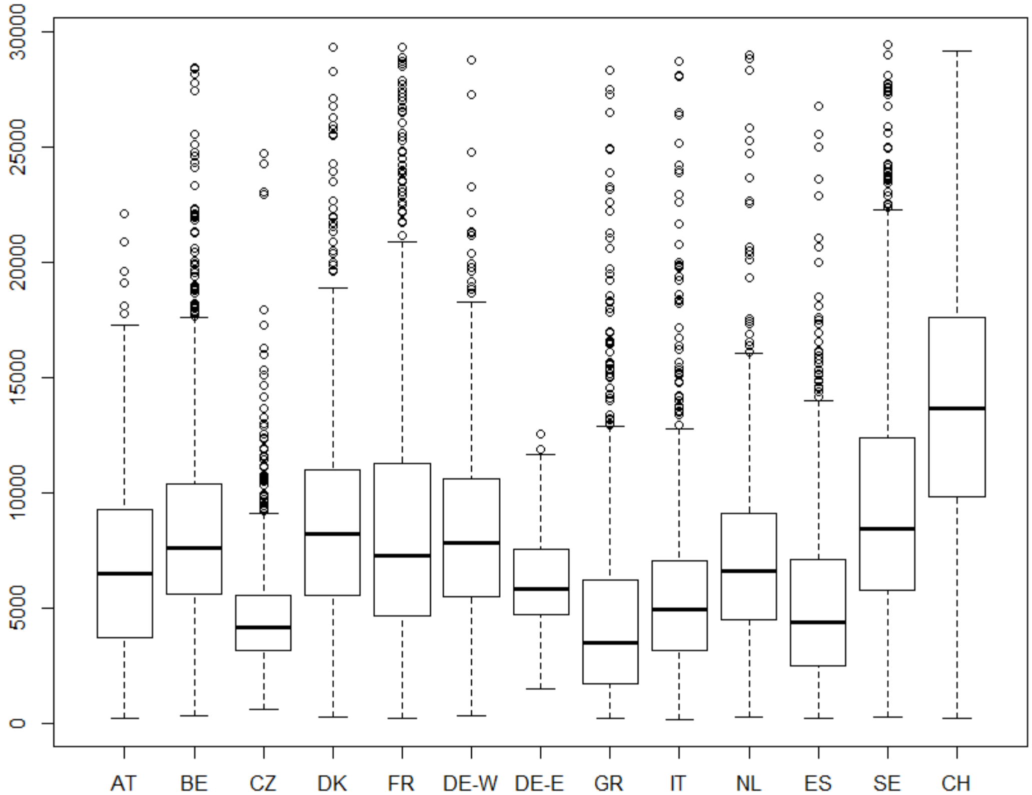

Figure 2: Box plots of lifetime earnings net of pension benefits, females

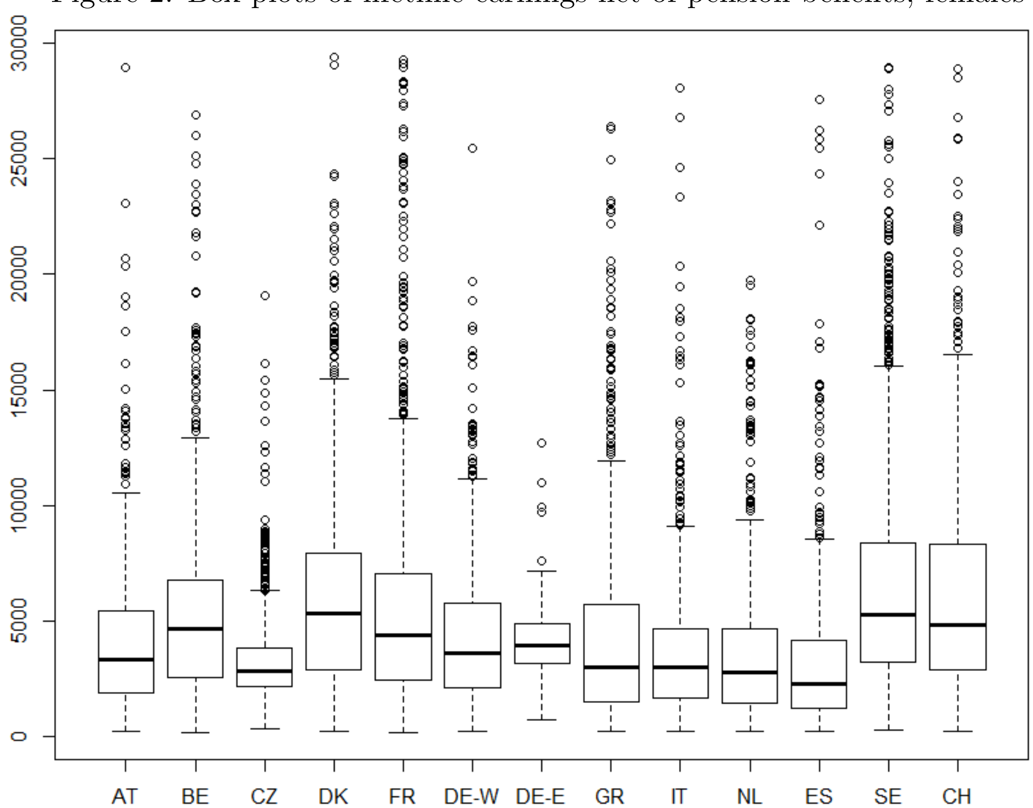

Figure 3: Kernel density estimates of lifetime earnings net of pensions, by education level

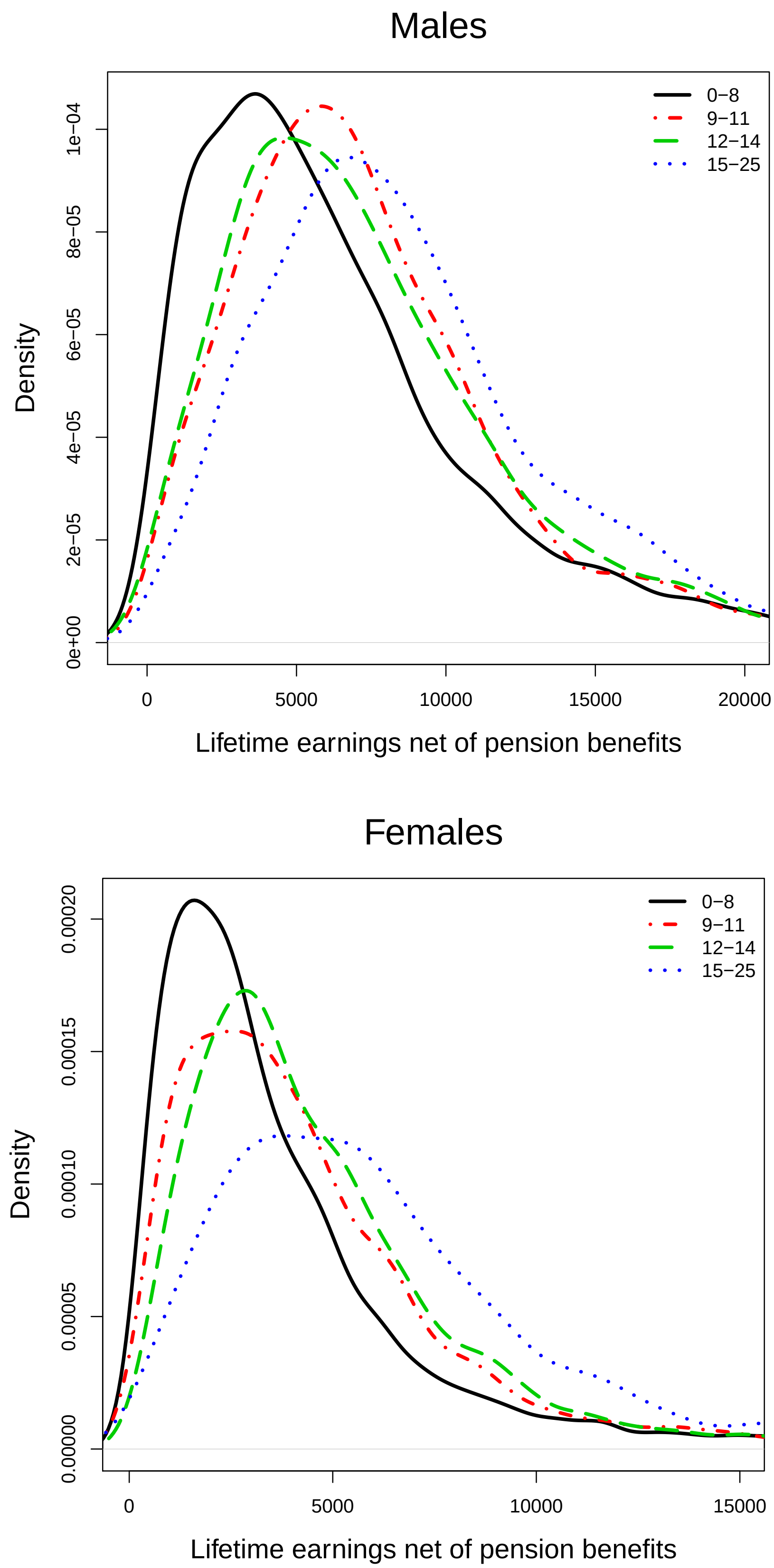

Figure 4: Average log lifetime earnings net of country and cohort effects, by year of education

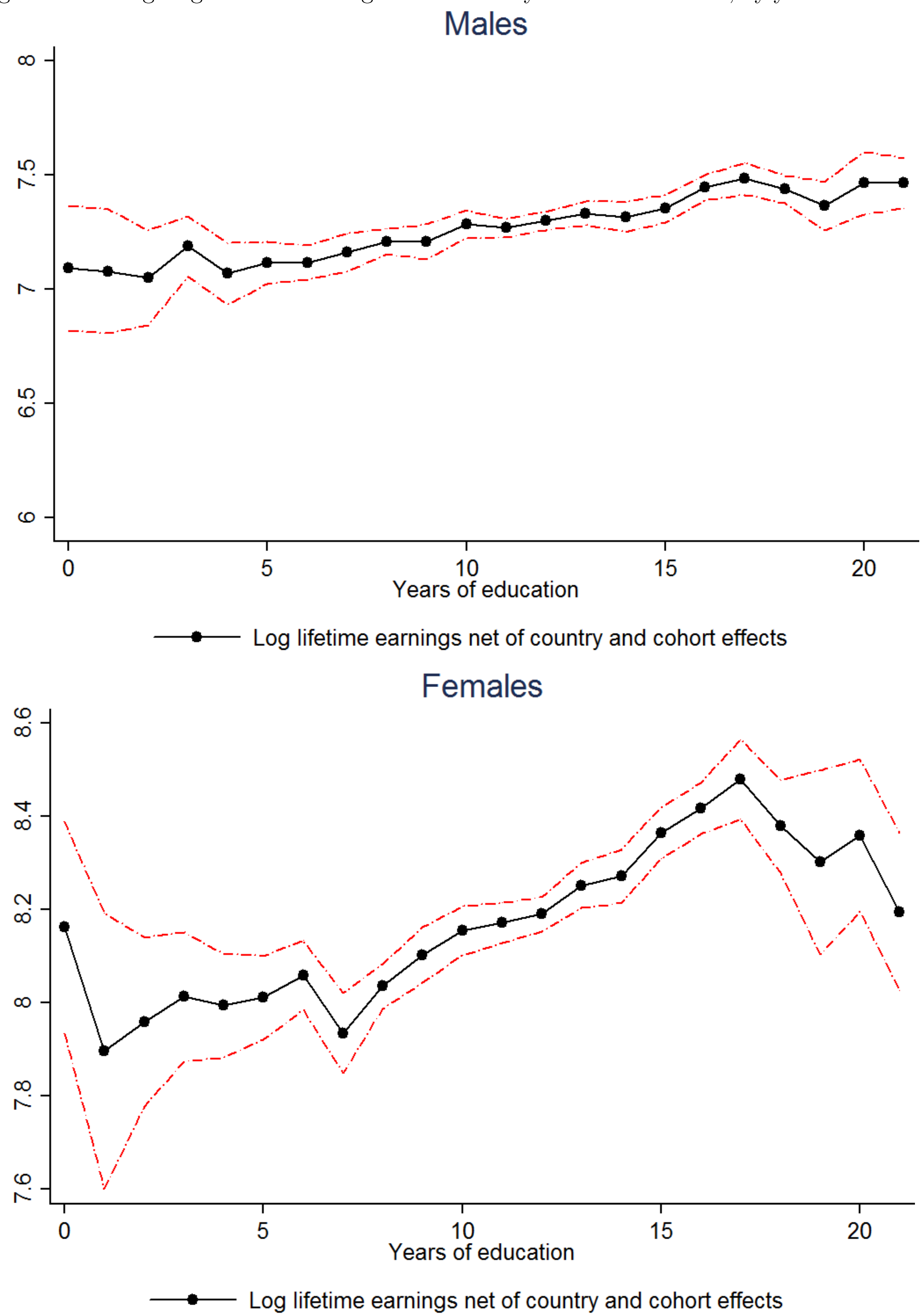

Table 1: Prediction error for current and main wage, males

Variable	Sample size	Mean	Std. Dev.
Log current income	2,298	9.9263	0.4774
Predicted log current income	2,298	9.9244	0.8281
Prediction error	2,298	0.0019	0.7727
Log main income	4,698	9.8067	0.7097
Predicted log main income	4,698	9.7444	1.0375
Prediction error	4,698	0.0623	1.1106

Table 2: Prediction error for current and main wage, females

Variable	Sample size	Mean	Std. Dev.
Log current income	2,582	9.5580	0.5216
Predicted log current income	2,582	9.5167	0.7430
Prediction error	2,582	0.0413	0.6537
Log main income	4,895	9.1484	0.8826
Predicted log main income	4,895	9.1072	1.0230
Prediction error	4,895	0.0412	0.9831

Table 3: Observed wage, predicted wage and prediction error, males

Year	Sample size	Observed wage Mean	Observed wage Std. Dev.	Predicted wage Mean	Predicted wage Std. Dev.	Prediction error Mean	Prediction error Std. Dev.	p-value
1985	984	5.700	0.459	5.675	0.522	0.024	0.432	0.270
1986	877	5.736	0.431	5.713	0.521	0.023	0.463	0.321
1987	838	5.772	0.444	5.760	0.512	0.011	0.471	0.630
1988	772	5.814	0.376	5.789	0.472	0.024	0.417	0.630
1989	784	5.811	0.447	5.782	0.533	0.029	0.515	0.241
1990	756	5.826	0.441	5.809	0.531	0.017	0.528	0.497
1991	729	5.836	0.422	5.820	0.546	0.016	0.532	0.534
1992	693	5.840	0.416	5.844	0.537	-0.004	0.514	0.866
1993	14	5.769	0.382	5.885	0.277	-0.116	0.443	0.366
1994	17	5.875	0.335	5.976	0.354	-0.101	0.365	0.401
1995	585	5.904	0.475	5.879	0.562	0.025	0.552	0.409
1996	556	5.915	0.477	5.904	0.568	0.011	0.568	0.724
1997	528	5.935	0.447	5.929	0.575	0.006	0.573	0.848
1998	480	5.935	0.482	5.948	0.628	-0.013	0.624	0.727
1999	447	5.957	0.470	5.954	0.626	0.002	0.617	0.951
2000	441	5.973	0.512	5.931	0.616	0.041	0.641	0.279
2001	401	5.963	0.500	5.934	0.596	0.028	0.642	0.467
2002	374	5.995	0.511	5.913	0.635	0.082	0.643	0.052
2003	344	6.027	0.530	5.899	0.638	0.128	0.649	0.004
2004	319	6.034	0.457	5.883	0.643	0.151	0.658	0.001
2005	295	6.016	0.489	5.895	0.624	0.121	0.653	0.009
2006	260	5.980	0.523	5.848	0.622	0.132	0.666	0.009
2007	249	5.925	0.653	5.855	0.645	0.070	0.801	0.229
2008	217	5.921	0.695	5.853	0.636	0.068	0.803	0.286

Note: German SOEP 1985-2008. The p-value refers to the hypothesis that the observed wage and the predicted wage have the same mean.

Table 4: Prediction error using the German SOEP, females

	Sample	Observed wage		Predicted wage		Prediction error		
Year	size	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	p-value
1985	569	5.077	0.678	5.062	0.708	0.015	0.533	0.718
1986	477	5.178	0.625	5.139	0.701	0.040	0.509	0.356
1987	438	5.233	0.600	5.178	0.694	0.055	0.564	0.212
1988	400	5.211	0.685	5.212	0.709	-0.001	0.647	0.979
1989	391	5.198	0.683	5.254	0.755	-0.056	0.709	0.280
1990	366	5.228	0.698	5.302	0.721	-0.074	0.723	0.159
1991	359	5.215	0.698	5.307	0.754	-0.092	0.760	0.090
1992	345	5.250	0.657	5.313	0.760	-0.063	0.744	0.241
1993	9	5.309	0.452	5.340	0.591	-0.031	0.512	0.902
1994	10	5.341	0.398	5.369	0.553	-0.027	0.511	0.901
1995	286	5.363	0.679	5.404	0.806	-0.042	0.851	0.502
1996	268	5.363	0.735	5.394	0.814	-0.031	0.863	0.644
1997	253	5.377	0.746	5.412	0.824	-0.035	0.931	0.617
1998	251	5.368	0.742	5.444	0.761	-0.076	0.875	0.258
1999	238	5.341	0.734	5.428	0.761	-0.088	0.886	0.202
2000	232	5.365	0.663	5.455	0.768	-0.090	0.894	0.176
2001	220	5.399	0.691	5.487	0.725	-0.088	0.844	0.195
2002	207	5.383	0.720	5.509	0.702	-0.127	0.867	0.071
2003	191	5.397	0.691	5.525	0.699	-0.127	0.827	0.074
2004	176	5.416	0.655	5.533	0.689	-0.117	0.843	0.103
2005	156	5.345	0.827	5.483	0.713	-0.139	0.972	0.113
2006	137	5.338	0.715	5.502	0.717	-0.164	0.862	0.059
2007	136	5.302	0.792	5.456	0.733	-0.154	0.961	0.098
2008	117	5.260	0.900	5.452	0.749	-0.191	1.043	0.078

Note: German SOEP 1985-2008. The p-value refers to the hypothesis that the observed wage and the predicted wage have the same mean.

Table 5: Lifetime earnings net of pension benefits, males

Country	Sample size	Mean	Std. Dev.	Median	IQR
Austria	323	6,999.22	4,068.25	6,489.23	5,602.92
Belgium	1,162	8,330.13	4,441.80	7,604.16	4,797.01
Czech Republic	777	4,809.82	2,838.38	4,138.16	2,382.99
Denmark	917	8,649.61	4,798.24	8,204.63	5,400.35
France	944	8,580.73	5,755.23	7,277.79	6,574.27
West Germany	766	8,301.15	4,242.93	7,820.20	5,104.27
East Germany	80	6,204.33	2,180.79	5,838.61	2,839.80
Greece	853	4,780.89	4,548.10	3,467.05	4,455.00
Italy	992	5,673.64	4,106.90	4,926.23	3,901.38
Netherlands	964	7,114.22	3,979.84	6,601.83	4,627.70
Spain	690	5,368.12	4,168.23	4,362.19	4,603.44
Sweden	779	9,755.31	5,731.41	8,416.35	6,677.36
Switzerland	528	13,567.20	5,918.12	13,666.85	7,811.06
Full sample	9,775	7,529.02	5,080.33	6,519.48	5,999.66

Table 6: Lifetime earnings net of pension benefits, females

Country	Sample size	Mean	Std. Dev.	Median	IQR
Austria	424	4,310.39	3,692.06	3,302.20	3,543.39
Belgium	1,264	5,205.26	3,646.84	4,635.88	4,174.09
Czech Republic	1,062	3,263.82	1,900.54	2,845.18	1,666.84
Denmark	1,133	5,938.91	4,156.85	5,298.51	5,067.39
France	1,165	5,597.02	4,901.15	4,380.26	4,564.90
West Germany	850	4,286.03	3,118.99	3,621.11	3,667.70
East Germany	81	4,268.32	2,126.28	3,954.93	1,770.28
Greece	680	4,529.30	4,539.59	3,014.11	4,208.75
Italy	868	3,679.60	3,230.25	3,002.35	2,979.97
Netherlands	1,106	3,477.65	2,919.39	2,766.99	3,192.68
Spain	622	3,386.37	3,728.25	2,246.66	2,951.78
Sweden	999	6,644.18	5,015.61	5,269.72	5,151.28
Switzerland	675	6,238.86	4,934.88	4,834.59	5,503.39
Full sample	10,929	4,775.06	4,053.13	3,710.21	4,033.25

Table 7: Lifetime income including pension benefits, males

Country	Sample size	Mean	Std. Dev.	Median	IQR
Austria	323	9,434.09	4,665.40	9,121.77	6,498.79
Belgium	1,162	10,774.89	5,028.94	10,291.14	5,908.20
Czech Republic	777	6,033.68	3,284.44	5,442.09	2,760.34
Denmark	917	10,921.88	5,521.68	10,678.96	6,912.74
France	944	11,534.21	6,570.94	10,377.84	7,961.43
West Germany	766	10,816.40	5,182.56	10,262.15	6,244.38
East Germany	80	8,445.94	3,309.26	8,177.26	3,387.01
Greece	853	7,102.01	5,389.93	5,759.79	5,974.63
Italy	992	7,548.73	4,747.02	6,998.13	5,245.92
Netherlands	964	10,191.94	5,092.24	9,744.44	6,496.12
Spain	690	7,446.36	5,016.27	6,281.90	6,074.99
Sweden	779	12,004.19	6,302.80	10,991.97	7,476.37
Switzerland	528	17,191.42	7,412.59	17,107.64	10,094.25
Full sample	9,775	9,929.23	5,974.06	8,951.56	7,450.10

Table 8: Lifetime income including pension benefits, females

Country	Sample size	Mean	Std. Dev.	Median	IQR
Austria	424	6,252.84	4,600.46	5,195.98	4,585.63
Belgium	1,264	7,148.04	4,305.51	6,578.95	5,119.69
Czech Republic	1,062	4,261.03	2,111.75	3,901.95	1,983.03
Denmark	1,133	7,748.02	4,729.29	7,283.52	6,044.35
France	1,165	7,678.06	5,807.94	6,400.83	5,992.64
West Germany	850	5,791.22	3,838.90	4,972.10	4,419.68
East Germany	81	5,741.39	2,437.50	5,582.17	2,698.81
Greece	680	6,822.41	5,442.84	5,474.06	5,586.92
Italy	868	5,213.35	3,758.35	4,449.61	3,845.99
Netherlands	1,106	5,419.65	3,817.19	4,620.90	4,408.15
Spain	622	4,952.19	4,329.53	3,527.45	4,090.18
Sweden	999	8,296.24	5,340.27	7,128.97	5,793.46
Switzerland	675	8,251.72	5,851.15	6,595.93	6,775.95
Full sample	10,929	6,534.84	4,740.83	5,374.64	5,133.31

Table 9: Annual pension benefits for males who are retired

Country	Sample size	Mean	Std. Dev.	Median	IQR
Austria	275	17,287.56	8,400.53	16,785.72	8,864.15
Belgium	862	16,468.11	7,616.30	14,881.38	7,521.75
Czech Republic	529	8,030.30	4,209.90	7,217.69	2,480.75
Denmark	514	12,343.53	6,655.33	10,632.32	6,366.06
France	678	20,808.04	12,762.40	17,885.44	14,589.89
West Germany	556	17,551.25	9,724.14	14,972.43	9,601.16
East Germany	57	13,354.24	10,373.01	10,930.21	6,669.91
Greece	565	16,555.78	9,684.88	14,395.23	10,766.57
Italy	811	13,463.89	7,863.44	11,885.48	8,873.26
Netherlands	646	19,390.84	11,372.68	17,046.91	13,377.59
Spain	482	12,777.52	7,689.11	10,986.36	9,263.73
Sweden	490	17,218.95	9,146.51	14,795.51	9,275.92
Switzerland	311	23,203.96	13,531.32	21,790.19	20,110.05
Full sample	6,776	16,054.00	9,987.83	13,642.35	10,699.86

Table 10: Annual pension benefits for females who are retired

Country	Sample size	Mean	Std. Dev.	Median	IQR
Austria	375	12,729.32	8,827.13	10,849.36	8,003.73
Belgium	955	12,699.87	7,626.23	11,766.64	7,314.53
Czech Republic	823	6,716.56	3,989.39	5,998.11	2,330.45
Denmark	676	10,813.22	6,661.15	9,484.30	4,683.71
France	841	13,788.85	10,455.07	11,890.80	11,202.10
West Germany	616	9,276.87	7,140.91	7,367.13	6,981.93
East Germany	52	9,601.14	5,179.09	8,745.62	6,010.83
Greece	462	13,349.61	9,885.86	10,127.56	8,553.16
Italy	702	8,881.57	6,365.13	7,009.92	6,010.32
Netherlands	800	12,345.77	9,052.41	9,350.51	7,185.82
Spain	476	8,953.67	7,138.93	6,911.82	5,293.09
Sweden	649	11,498.88	5,442.77	10,231.79	5,616.46
Switzerland	438	13,883.37	9,802.96	10,927.49	7,867.96
Full sample	7,865	11,142.30	8,106.60	9,094.93	7,358.06

Table 11: Expected annual pension benefits for males who are still working

Country	Sample size	Mean	Std. Dev.	Median	IQR
Austria	41	15,912.76	7,379.96	13,458.87	8,075.32
Belgium	255	15,568.12	6,782.90	14,358.45	7,060.47
Czech Republic	205	7,839.34	4,675.98	6,601.58	2,981.36
Denmark	336	17,122.29	6,989.54	16,324.94	7,689.05
France	227	17,876.21	9,462.34	15,850.33	12,162.92
West Germany	174	18,750.99	10,894.70	16,540.31	12,039.35
East Germany	19	13,044.10	7,614.93	10,769.15	7,897.38
Greece	231	14,330.86	8,100.98	12,597.03	8,016.83
Italy	132	12,867.96	6,331.15	11,349.47	4,831.41
Netherlands	277	21,663.31	9,882.41	20,654.06	9,653.53
Spain	172	15,150.35	7,428.44	13,824.67	6,943.19
Sweden	233	15,811.90	7,368.45	14,400.50	6,882.59
Switzerland	167	27,924.04	11,302.25	26,321.74	14,340.06
Full sample	2,469	16,814.07	9,366.25	14,812.14	10,178.84

Table 12: Expected annual pension benefits for females who are still working

Country	Sample size	Mean	Std. Dev.	Median	IQR
Austria	42	10,706.89	4,573.39	10,292.59	6,280.80
Belgium	281	11,183.93	4,731.50	10,930.77	6,693.41
Czech Republic	210	5,790.82	2,568.27	5,323.85	2,555.45
Denmark	426	14,255.25	4,933.90	13,876.20	5,702.85
France	308	11,931.90	8,586.14	10,016.53	7,691.26
West Germany	208	9,433.06	6,170.52	8,341.95	7,009.22
East Germany	24	10,229.36	4,915.26	9,830.30	5,950.65
Greece	203	10,228.80	5,129.79	9,690.02	5,496.09
Italy	146	9,751.51	4,472.44	9,774.67	5,882.18
Netherlands	264	12,348.17	7,195.20	11,225.03	9,679.76
Spain	128	11,738.32	6,308.47	10,292.93	6,912.33
Sweden	322	12,093.37	5,037.37	11,700.41	5,506.07
Switzerland	202	14,035.23	9,485.62	11,844.78	12,009.29
Full sample	2,764	11,488.99	6,494.47	10,732.00	7,518.97

Table 13: Years of education, males

Country	Sample size	Mean	Std. Dev.	Median	Min	Max
Austria	323	9.00	4.74	8	1	25
Belgium	1,162	12.17	3.83	12	0	25
Czech Republic	777	12.34	3.28	12	1	24
Denmark	917	12.71	3.67	11	4	21
France	944	11.89	4.16	12	0	25
West Germany	766	13.50	3.31	13	5	25
East Germany	80	13.73	3.06	13	8	21
Greece	853	10.12	4.34	11	0	24
Italy	992	8.53	4.48	8	0	25
Netherlands	964	11.64	3.83	11	1	25
Spain	690	8.54	4.77	8	0	25
Sweden	779	11.34	4.08	11	0	25
Switzerland	528	12.13	4.80	12	1	25
Full sample	9,775	11.29	4.36	11	0	25

Table 14: Years of education, females

Country	Sample size	Mean	Std. Dev.	Median	Min	Max
Austria	424	8.68	4.07	8	1	25
Belgium	1,264	11.63	3.34	12	1	25
Czech Republic	1062	11.28	2.81	12	0	25
Denmark	1,133	12.19	4.00	12	4	18
France	1,165	11.19	3.87	11	0	25
West Germany	850	12.23	3.05	12	1	25
East Germany	81	12.20	2.69	12	6	18
Greece	680	9.79	4.34	10	0	22
Italy	868	8.24	4.24	8	0	25
Netherlands	1,106	10.97	3.30	10	0	24
Spain	622	8.18	4.35	8	0	25
Sweden	999	11.53	3.90	11	1	25
Switzerland	675	10.96	4.16	11	1	25
Full sample	10,929	10.85	3.96	11	0	25

Table 15: Years of work, males

Country	Sample size	Mean	Std. Dev.	Median	Min	Max
Austria	323	36.35	9.78	39	1	59
Belgium	1,162	33.89	10.02	36	1	53
Czech Republic	777	37.88	7.47	40	1	56
Denmark	917	31.73	11.21	35	1	62
France	944	31.84	10.89	35	1	59
West Germany	766	35.36	9.91	38	1	57
East Germany	80	33.74	8.07	35	13	58
Greece	853	33.83	11.47	36	1	70
Italy	992	31.51	11.48	35	1	68
Netherlands	964	35.02	10.56	38	2	56
Spain	690	35.69	12.78	39	1	71
Sweden	779	35.71	10.87	38	1	63
Switzerland	528	33.88	11.62	38	1	77
Full sample	9,775	34.14	10.89	37	1	77

Table 16: Years of work, females

Country	Sample size	Mean	Std. Dev.	Median	Min	Max
Austria	424	22.08	13.29	23.5	1	51
Belgium	1,264	22.12	13.07	24	1	49
Czech Republic	1,062	33.55	7.67	35	1	64
Denmark	1,133	26.40	11.32	29	1	57
France	1,165	24.28	12.76	27	1	55
West Germany	850	24.76	12.72	27	1	63
East Germany	81	30.33	11.13	32	2	67
Greece	680	22.47	12.02	23	1	59
Italy	868	20.33	12.67	20	1	51
Netherlands	1,106	18.57	12.08	17	1	58
Spain	622	18.51	13.60	14	1	54
Sweden	999	31.33	10.26	33	1	72
Switzerland	675	20.33	12.95	21	1	59
Full sample	10,929	24.22	12.86	26	1	72

Table 17: Number of jobs during the career, males

Country	Sample size	Mean	Std. Dev.	Median	Min	Max
Austria	323	2.70	1.61	2	1	13
Belgium	1,162	2.70	1.67	2	1	11
Czech Republic	777	2.26	1.58	2	1	11
Denmark	917	4.40	2.67	4	1	18
France	944	3.08	1.90	3	1	14
West Germany	766	2.76	1.75	2	1	10
East Germany	80	3.52	1.83	3	1	9
Greece	853	1.35	0.72	1	1	10
Italy	992	2.45	1.49	2	1	11
Netherlands	964	3.32	1.98	3	1	14
Spain	690	2.45	1.71	2	1	13
Sweden	779	3.86	2.41	3	1	17
Switzerland	528	3.68	2.21	3	1	14
Full sample	9,775	2.92	2.03	2	1	18

Table 18: Number of jobs during the career, males

Country	Sample size	Mean	Std. Dev.	Median	Min	Max
Austria	424	2.53	1.62	2	1	13
Belgium	1,264	2.17	1.41	2	1	11
Czech Republic	1,062	2.39	1.58	2	1	11
Denmark	1,133	4.22	2.70	4	1	19
France	1,165	2.62	1.71	2	1	16
West Germany	850	2.76	1.69	2	1	11
East Germany	81	3.27	2.12	3	1	11
Greece	680	1.29	0.68	1	1	6
Italy	868	1.83	1.14	1	1	10
Netherlands	1,106	2.95	1.79	3	1	10
Spain	622	1.73	1.05	1	1	7
Sweden	999	3.73	2.37	3	1	20
Switzerland	675	3.62	2.30	3	1	16
Full sample	10,929	2.72	1.98	2	1	20

References

Angelini, Viola, Agar Brugiavini and Guglielmo Weber (2009). "Ageing and Unused Capacity in Europe: Is There an Early Retirement Trap?" *Economic Policy*, 24, 463-508.

Bhuller, Manudeep, Magne Mogstad and Kjell G. Salvanes (2011). "Life-cycle Bias and the Returns to Schooling in Current and Lifetime Earnings." IZA Discussion Paper No. 5788.

Böhlmark, Anders and Matthew J. Lindquist (2006). "Life-Cycle Variations in the Association between Current and Lifetime Income: Replication and Extension for Sweden." *Journal of Labor Economics*, 24(4), 879-896.

Bound, John, Charles Brown and Nancy Mathiowetz (2001). "Measurement Error in Survey Data." In *Handbook of Econometrics, Vol. 5.* Heckman, J.J. and E. Leamer. (eds.), North-Holland, Amsterdam, 3705-3843.

Brenner, Jan (2010). "Life-Cycle Variations in the Association between Current and Lifetime Earnings: Evidence for German Natives and Guest Workers." *Labour Economics*, 17(2), 392-406.

Christelis, Dimitris (2011). "Imputation of Missing Data in Waves 1 and 2 of SHARE." SHARE Working Paper 01/2011.

Garrouste, Christelle and Omar Paccagnella (2011). "Data Quality: Three Examples of Consistency Across SHARE and SHARELIFE Data." In *Retrospective Data Collection in the Survey of Health, Ageing and Retirement in Europe. SHARELIFE Methodology.* Schröder, M. (ed.), Mannheim Research Institute for the Economics of Ageing (MEA), Mannheim, 62-72.

Gould, Eric D., Victor Lavy and M. Daniele Paserman (2011). "Sixty Years after the Magic Carpet Ride: The Long-Run Effect of the Early Childhood Environment on Social and Economic Outcomes." *Review of Economic Studies*, 78(3), 938-973.

Haider, Steven and Gary Solon (2006). "Life cycle Variation in the Association between Current and Lifetime Earnings."*American Economic Review*, 96(4), 1308-1320.

Havari, Enkelejda and Fabrizio Mazzonna (2011). "Can We trust older people's statements on their childhood circumstances? Evidence from SHARELIFE." SHARE Working Paper 05/2011.

Heitjan, Daniel F. and Roderick J.A. Little (1991). "Multiple Imputation for the Fatal Accident Reporting System." *Journal of the Royal Statistical Society C*, 40(1), 13-39.

Horton, Nicholas J. and Ken P. Kleinman (2007). "Much Ado About Nothing: A Comparison of Missing Data Methods and Software to Fit Incomplete Data Regression Models." *American Statistician*, 61(1), 79-90.

Kapteyn, Arie and Yelmer Y. Ypma (2007). "Measurement Error and Missclassification: A Comparison of Survey and Administrative Data." *Journal of Labor Economics*, 25(3), 513-551.

Mazumder, Bhashkar (2001). "The Mis-Measurement of Permanent Earnings: New evidence from Social Security Earnings Data." Federal Reserve Bank of Chicago Working Paper 2001-24.

Meijer, Erik, Susann Rohwedder and Tom Wansbeek (2011). "Measurement Error in Earnings Data: Using a Mixture Model Approach to Combine Survey and Register Data." *Journal of Business & Economics Statistics*, forthcoming.

Murphy, Kevin M. and Finis Welch (1990). "Empirical Age-Earnings Profiles." *Journal of Labor Economics*, 8(2), 202-229.

Pischke, Jörn-Steffen (1995). "Measurement Error and Earnings Dynamics: Some Estimates From the PSID Validation Study." *Journal of Business & Economics Statistics*, 13(3), 305-314.

Schenker, Nathaniel and Jeremy M.G. Taylor (1996). "Partially Parametric Techniques for Multiple Imputation." *Computational Statistics & Data Analysis*, 22(4), 425-446.

Trevisan, Elisabetta, Giacomo Pasini and Roberta Rainato (2011). "Cross-Country Comparison of Monetary Values from SHARELIFE." SHARE Working Paper 02/2011.