INFERRING OCCUPATION ARDUOUSNESS FROM POOR HEALTH BEYOND THE AGE OF 50

Arno Baurin, Sandy Tubeuf and Vincent Vandenberghe

LIDAM Discussion Paper IRES
2022 / 05

UCLouvain

LIDAM
IRES | Louvain Institute of Data Analysis and Modeling in economics and statistics

Inferring Occupation Arduousness from Poor Health Beyond the Age of 50

Arno Baurin[*], Sandy Tubeuf[†] & Vincent Vandenberghe[‡]

March 2, 2022

Abstract

This paper shows that the analyst with no information on occupation arduousness could reasonably infer it from poor health beyond 50. Using retrospective lifetime data from the Survey of Health, Ageing and Retirement in Europe (SHARE), including the respondents' professional career described with ISCO 2-digit, this paper finds a statistically significant link between many occupations and the risk of poor health beyond the age of 50. Next, we quantify the relative contribution of professional occupation to poor health compared to other factors decomposing the variance of health disparities between sources. We find that occupation's arduousness - although a significant predictor of poor health - is less consequential than initial health endowment, demographics or country fixed effects in explaining differences in health at an older age.

Keywords: Health, Work, Occupation Arduousness, Variance Decomposition
JEL Codes: I10, J26, J28

[*]Institute of Economic and Social Research (IRES/LIDAM, UCLouvain) & FRESH-FNRS, arno.baurin@uclouvain.be

[†]Institute of Economic and Social Research (IRES/LIDAM, UCLouvain) & Institute of Health and Society (IRSS, UCLouvain), sandy.tubeuf@uclouvain.be

[‡]Institute of Economic and Social Research (IRES/LIDAM, UCLouvain), vincent.vandenberghe@uclouvain.be

1 Introduction

The rise of average healthy life expectancy calls for an overall increase in the age of retirement. However, many observers suggest that to preserve retirement equality, the age of retirement should be differentiated (see e.g. Vermeer et al., 2016). Same-age older individuals differ a lot in terms of health (Wise, 2017; Vandenberghe, 2021) and also as to their remaining (healthy) life expectancy.[1] There is strong evidence that ill-health at 50 is correlated with a shorter life span/early death, De Nardi et al. (2016) show that lifespan is 3.3 years shorter for those with bad health than for those with good health, while Pijoan-Mas and Ríos-Rull (2014) show the equivalent numbers are 5.6 for men and 4.7 for women at age 50.

Historically, in most retirement systems, a uniform age has been used as a proxy for poor health, the ensuing loss of work capacity and remaining (healthy) life expectancy. But now comes this proposition to adopt a slightly more refined proxying strategy: one that consists of using several retirement ages to better match the distribution of late years health status and work capacity across socio-demographic groups, and in particular across occupations that vary a lot in terms of their arduousness. Many stakeholders, but also economists (Ayuso et al., 2016) call for the abolition of the uniform age of retirement. Paradoxically, the implementation of this idea seems complicated. There is a lack of a direct and consensual measure of arduousness and its actual contribution to the the risk of poor health/short life beyond 50. This is one of the reasons why governments struggle to implement it. "Arduous" jobs are often defined more or less arbitrarily. An exception is perhaps Poland, where daily calories needed to perform the job have been used to evaluate the arduousness of a profession (see e.g. Zaidi and Whitehouse, 2009). Still, some sedentary occupations might not require as much calories as a manual work but could also be detrimental to health.

[1] In Belgium, for people that entered the labour market at the age of 20, there is a 10% risk that they will not reach the uniform age of retirement of 65. At the age of 65, the remaining life expectancy gradient is in the range of 8 to 10 years considering gender$\times$education SES categories.

This paper aims at clarifying the above debate. It should be seen as contribution to the economic literature on retirement age differentiation and demanding occupations (Pestieau and Racionero, 2016; Vermeer et al., 2016; Vandenberghe, 2021). More generally, it contributes to the work-health literature, by exploiting unique, and so far untapped retrospective data on career, together with very detailed data on health beyond the age of 50. Finally, this paper relates to the life course literature as it as it stresses the role of pre-labour market entry determinants of late-years health (Trannoy et al., 2010).

The impact of work on health has long been investigated by the eponymous work-health literature in epidemiology, psychology, and sociology literature; but to a lesser extent in economics (for a review of the literature, see Barnay, 2016; Bassanini and Caroli, 2015). Most research and policy debates underline various negative consequences of work and working conditions such as stress, physical exhaustion, work-related disabilities, overall poor health as well as premature death.

Using data from the US National Health Interview Survey (NHIS), Case and Deaton (2005) showed that manual workers, with low education or low wealth have higher rates of health deterioration and that their physical health will deteriorate more rapidly with age than non-manual workers. However, overall, economists have been cautious at interpreting these correlations as evidence of a causal effect of occupation on health (see Ravesteijn et al., 2013, for a thorough literature review). The main problem is that individuals in poor health could not only face reduced employment opportunities, but could also self-select in specific jobs because of their health status. In order to overcome this issue, Sindelar et al. (2007) have assessed the impact of the first occupation on later health. Their underlying assumption is that health should be homogeneous at the start of the career. However, several studies have shown that health may already differ at entry in the labor market. In particular, unhealthy individuals may self-select into occupations that are less arduous. To cope with this problem, Fletcher and Sindelar (2009) use the father's occupation or the

US State economic conditions at the start of the career to instrument respondent's first occupation. But the exogeneity of both instruments could largely be questioned as they are likely to affect health directly. Fletcher et al. (2011) combined external information on the physical requirements of work and environmental conditions with data from US Panel Study of Income Dynamics to estimate the health impact of exposure to physical and environmental conditions at work. However, they failed to establish a causal effect of occupation on health. More recently, Morefield et al. (2012) and Ravesteijn et al. (2018) have considered dynamic models controlling for health status at different time points. When controlling for lagged health one year before, Ravesteijn et al. (2018) found that at least 60% of the correlation between occupation and health is due to people with worse health self-selecting into specific occupations. In short, arduousness as the impact of an occupation on health is not easy to estimate.

Beyond the link between work/occupation and health, there is another stream of the economic literate that shows that past (in particular early life) personal experience can undermine health and professional pathways cumulatively over the life course (Lindeboom, 2012). There are long-lasting effects of family and social background on health status in adulthood. Three concurrent channels of transmission from one generation to another have been identified (Trannoy et al., 2010): a direct channel where social background influences adult health following a latency period; an indirect channel where social background influences health through its influence on employment and life trajectories; and the third channel is an inter-generational transmission of health a common genetic capital within families. More generally, a large body of literature equally acknowledges the role played by the social determinants of health (Marmot and Wilkinson, 2005). Such selection effects from early life experience must also be taken into account to measure the net effect of work on health.

Our contribution to the above literature is essentially twofold.

- The first one is our demonstration that *career arduousness* – at least when it comes to *ranking* individuals with different careers – could relatively easily be inferred from health data in the European context in the absence of direct and detailed measures of arduousness. There are survey providing direct estimates of average "arduousness" by occupation. An example is the US Occupational Information Network (O*Net)[2], that provides an impressively detailed description of working conditions for almost 1,000 professions, or its less known and less detailed European counterpart the European Working Conditions Survey (EWC).[3] However, those surveys are no panacea for two reasons. First, they require governments to collect a lot of (new) data. Second, they are not able to summarize the career arduousness in one indicator, but only gives a very detailed list of working conditions. Ideally, governments should be able to have a synthetic indicator.

In this paper, we propose a quantification of arduousness that is indirect but more in line with the way the concept is defined by job demands and job quality literature (Bakker and Demerouti, 2007; Chen et al., 2017). A more arduous/demanding occupation or job requires more physical and/or psychological effort or skills and consumes more physiological and/or psychological resources. If occupations are unequally stressful or physically demanding they should contribute to individuals' health gradient. We thus posit that it is via a careful examination of the relationship between health and occupational differences that we should be able to best quantify arduousness. In this paper, we ask whether the health status at 50 and beyond could be used to infer a measure of work arduousness that is as relevant as those provided by surveys like O*Net or EWC. Detailed data about old age health status are readily available. We use those amassed via the 7$^{\text{th}}$ wave of the Survey on Health, Ageing and Retirement in Europe (SHARE) to analyze the link between health

[2]See onetcenter.org.
[3]See eurofound.europa.eu/surveys/european-working-conditions-surveys-ewcs.

at old age[4] and work history of older individuals in 28 European countries, also available in that survey.

What differentiates our approach from most of the job demands or work-health literature is that we are not just interested in analysing the health consequences of the current or most recent occupation, but the occupation(s) representing a *career*. That objective directly derives from the recent availability in SHARE of a detailed account (including in terms of occupation) of people's entire career. As far as we know, quantifying arduousness using information about people's career and their late year health is something new in the economic literature.

- The second key contribution of the paper pertains to the quantification of the *relative contribution* of career arduousness to the risk of poor health beyond 50,[5] thus in comparison with other factors. In other words, we do not only assess the impact of career arduousness on health, but also "how much" career arduousness is responsible for differences in health at old age. For that purpose, we use the natural decomposition of the variance as proposed by Shorrocks (1982). This allows us to quantify the relative importance of career arduousness to health inequality compared to other variables like what epidemiologists call people's health endowment and other *pre-labour-market* entry determinants of late-life health (e.g. education, gender). One of the remarkable features of SHARE is to inform on respondents' initial health endowment before entering the labour market. SHARE collects the health status during teenage along with information on whether parents are dead and if so, their age at death. These variables allow us to control for the latter's direct impact on late-life health but also its potential role on entry-level occupational choice. From an econometric point of view, these represent a source of selection bias. They

[4] SHARE samples respondents from the age of 50.
[5] And by extension, the risk of short life.

must be taken into account to properly measure the net effect of professional occupation on later health.

The first key result of the paper is to show that an indicator of career arduousness could indeed be inferred from health at old age and could serve as a reasonable substitute for a detailed occupational survey. The second question at the core of this paper is the relative contribution of career arduousness to people's health. And we show that career arduousness, although a statistically significant determinant of health in old age, does not emerges as being the main driver of late year health inequalities. The rest of this paper is structured as follow. The methods are explained in Section 2 and data are described in Section 3. Section 4 presents the results and Section 5 concludes.

2 Method

The aim of this paper is to analyze the link between professional occupations and health at older age in order to provide a ranking of job arduousness. Let us consider $Health_{i,j}$, a measure of poor health considered binary and equals to 1 if individual i in country j is in poor health and 0 otherwise. We consider that $Health_{i,j}$ is a function that can be written as follows:

$$Health_{i,j} = \alpha + \beta\, OCC_i + \gamma\, X_i + \delta_j + \epsilon_{i,j} \tag{1}$$

where OCC refers to the occupation, X is a vector of controls and δ is the country fixed effect. (1) will be estimated separately for men and women. The occupation variable can refer to the main, the first or the last job or the time spent in each occupation (coded at the ISCO 2 digit). We will test those different specifications in order to know if our results depend on them. The vector of controls contain age, education, father's occupation and initial health endowment. We will add progressively those different controls to see how much our arduousness coefficients change. We will also provide a coefficient decomposition "à la Gelbach (2016)" that allow us to identify the key factors driving the gradual reduction of the magnitude of our arduousness estimates when enriching our model, in particular the role of the correlation between the level of pre-labour market endowments (education, father's education and initial health) and the likelihood to exert different occupations. Note that all our models are estimated using OLS instead of logit. This has the advantage of generating coefficient that can be decomposed. We have checked that this does not affect our other key results. Second, we provide a variance decomposition in order to analyze the relative importance of some variables (e.g. occupation) versus others variables (e.g. country). We draw from the method pioneered by Shorrocks (1982) and used by Fields (2003) in labor economics and Jusot et al. (2013) in health economics. It consists of combining regression analysis with variance decomposition. The procedure consists of two stages. At stage 1, using equation (1), we predict the health based on the vectors of regressors:

$$
\begin{aligned}
\widehat{Health}^{OCC_{i,j}} &= \widehat{\beta^{OCC}} \times OCC_{i,j} \\
\widehat{Health}^{X^k_{i,j}} &= \widehat{\gamma}^k \times X^k_{i,j} \\
\widehat{Health}^{\delta_{i,j}} &= \widehat{\delta_j}
\end{aligned}
\tag{2}
$$

At stage 2, we use the variance of health as a reference to quantify the contribution of each variable. The decomposition is given by the covariance between each regressor and the health outcome.

$$
\sigma^2\left(\widehat{Health_{i,j}}\right) = \sigma\left(\widehat{Health_{i,j}}, \widehat{Health^{OCC}_{i,j}}\right) + \sigma\left(\widehat{Health_{i,j}}, \widehat{Health^{X^k}_{i,j}}\right) + \sigma\left(\widehat{Health_{i,j}}, \widehat{Health^{\delta}_{i,j}}\right)
\tag{3}
$$

Therefore, the relative importance of a particular variable (or group of variables) is the ratio of its covariance divided by the total model-explained variance.

$$
ratio^{OCC} = \frac{\sigma\left(\widehat{Health_{i,j}}, \widehat{Health^{OCC}_{i,j}}\right)}{\sigma^2\left(\widehat{Health_{i,j}}\right)}
\tag{4}
$$

We now describe our data before moving to the presentation of our results.

3 Data

This paper uses the 7^{th} wave of the Survey on Health, Ageing and Retirement in Europe (SHARE). This wave was conducted across 28 European countries and Israel in 2017. The 7^{th} wave contains several "retrospective" modules, whose aims are to provide detailed data about the respondent's history.[6] Extensive information is provided about, among others, childhood health and job history. The original sample contains 77 263 observations.

Our main variables of interest are the health at old age, the job history and the childhood conditions. The health variable used is the self-rated health. It is the answer to the question "How would you rate your health?" on a 5-items scale: Excellent, Very Good, Good, Fair and Poor. This variable is known for years to be a reliable indicator of health (see e.g. Bound, 1991; Idler and Benyamini, 1997), physical as well as mental health (Han and Jylha, 2006). It is frequently reassessed with the same conclusion about its validity (see e.g. Schnittker and Bacak, 2014).[7] Following Etilé and Milcent (2006), we dichotomized self-reported health into "Good, Fair and Poor" versus "Excellent and Very good". Table 1 details the health variable before having been dichotomized.

Descriptive statistics of the sample can be visualized in Table 2; for example, the average health of Food preparations assistants is worse than the average health of Teachers. In the 7^{th} wave of SHARE, respondents were asked to retrace their complete job history by providing the starting/ending year of each of their jobs that lasted more than 6 months. Job titles are reported at ISCO-4 digits, which we collapsed into ISCO-2 digits for data issues. We transformed the job history into several variables: the first job, the main job, the last job and the time spent in each

[6] The 3^{rd} wave also contains a retrospective module. However, the occupation variable is only detailed at ISCO-1 digit. Due to this limitation, we only use the 7^{th} wave.

[7] It is also a good predictor of more elaborate health indices that can be computed using SHARE numerous subjective and objective health items (Vandenberghe, 2020).

ISCO code. We define the main job as the one having spanned the longest. We withdrew observations having as main job: “Commissioned armed officers”, “Non-commissioned armed forces officers”, “Armed forces occupations, other ranks” and “Street and related sales and services workers” due to the small number of observations from those categories.

One of the strengths of this paper is that SHARE data allows us to control for the initial health endowment. We are not only able to control for the health status of the individual at age 15, but we can also account for the inherited health endowment. The health status at age 15 is reported retrospectively by the respondent in five items that we group as a binary variable as we do with the health outcome at older age. We then proxy inherited health endowment by the death status of the parents. To do so, we consider whether parents are currently alive, and if they have died we consider whether they “prematurely” died (i.e. they died younger than the median age at death in the considered country) or not. This variable can be considered as a proxy of the “genetic” background of the respondent under the assumption of intergenerational transmission of health (Trannoy et al., 2010). We also control for the father’s occupation as a proxy for socioeconomic condition in childhood.

SHARE data has some limitations. First, they do not include a repeated assessment of health during the job history. This a limitation if one is interested in exploring the dynamic relationship between the evolution of health and occupational choices (Ravesteijn et al., 2018). However, in this paper, we care more about the long-term impact of occupation. Second, the participant’s history is reported retrospectively and a long time since it happened (i.e. a retiree in 2017 provides his work history since 1970 if he started working at 20). This can lead to memory biases. To reduce this problem, the survey uses a “Life History Calendar” approach to help the respondent to report accurately his history.

4 Results

4.1 Coefficients of the impact of occupation on health

This section details our analysis of the impact of occupation on health. Remember that the goal of this paper is to provide a ranking of occupation arduousness. We have two main concerns. The first is to know if the specification of our variable "occupation" can change our ranking. The second is to know if the inclusion of variables of control can do it. We start our study by analysing the first point. First, we consider equation (1), where OCC refers to the *main job* and X to age. We have chosen "Business and associate professionals" as the reference occupation.[8] The results are displayed numerically in Table 3 (male) and 4 (female) and graphically in Figure 1. The magnitudes of the coefficients show a clear gradient with presumably less arduous professions, like Teaching professionals being negatively associated with poor health, and more arduous ones like Refuse and other elementary workers or Food preparation assistants, showing the opposite. A sample of 16 occupation over 38 have a significant (at 5 %) coefficient for male and 22 for female. Being older is associated with a higher probability of being in bad health.

However, those coefficients could be biased due to a mixture of measurement error and selection problems regularly highlighted by people doing occupational cohort studies. Measurement error could be driven by the main job dummy being poor proxy of the actual exposure to arduousness, in particular its duration. Also individuals move from job to job in a non random way (e.g. the choice of a less arduous main job could be a response to the health deterioration caused by a first job that was particularly arduous). We will discuss endogeneity issues more thoroughly below (see Section 4.2). In short, to assess the magnitude of these potential biases, we re-estimate (1) with alternative specifications of our occupation variable OCC. Figure 2 (male) and Figure 3 (female) show the correlation between our previous results (main job) and the results, where OCC refers to the *time spent* working in each

[8]The reason being that it is the most frequent occupation in our sample.

ISCO code. There is a clear link between the two and the correlation coefficients are 97.52 % for male and 95.07 % for female. We have also tested the robustness of our ranking by computing (1) with *first* or *last job* as the independent variable. As could be seen in Figure 4 (male) and Figure 5 (female), all coefficients follow the same trend. The correlation between those of the main job and those of the first job is 81.99 % for male and 84.80% for female; the one between the main job and the last job is 94.91 % for male and 96.73 % for female. Still, this does not mean that our coefficients are free from potential biases. However, the goal of this paper is to focus on ranking occupations and this shows that our ranking is stable, whatever the occupation variable used.

We now analyze if adding new variables of control change the ranking of occupations arduousness. We have recomputed (1) and added education in Model 2 along with initial health endowment and father's occupation in Model 3. The results are displayed in columns 2 and 3 of Table 3 (male) and of Table 4 (female). The reference level for education is post-secondary, non-tertiary education ($ISCED\,4$). The education variable has the expected sign. A higher level of education increased the likelihood of being in good health. The overall magnitude of occupation coefficients fall by around 25 % for male and 45 % for female, but remain significant in explaining health status. Turning now to the last regression, we see that being in poor health during teenage increases the probability of reporting a poor health beyond 50. Similarly having a parent who died prematurely is associated with a higher probability of being in poor health in old age comparatively to having parents still alive. The father's occupation are not significant, except for the woman who had a father who was a Senior manager and professionals, which decreases the probability of being in poor health (compared to Office clerks, service workers and sales workers). The coefficients associated with professional occupations remain significant with only a slight decrease compared to the second specification. The coefficients of the different occupations across the regressions can be visualized in Figure 6 (male) and 7 (female). The reduction we find here is lower than the one found by Ravesteijn et al. (2018) (-60%). This difference could be explained by the fact that Ravesteijn et al.

(2018) controlled for one-year lagged health while we control for initial health endowment during the adolescence. Both control strategies have their pros and cons. Controlling for health a year before withdraw a potential bias; however it cannot control for the long-lasting effect of one's occupation on health. The reverse applies when we control for health during the adolescence.

As an extension to the regression analysis above, we provide a decomposition "à la Gelbach (2016)". Gelbach reminds us of what are the two key determinants of the evolution of a coefficient of interest (here a series of coefficients $\widehat{\beta^{OCC}}$ capturing the link between an occupation and health beyond 50) when adding to the initial model a series of controls X_k (education, initial health endowment), see Figure 6 and 7. Gelbach refers to baseline $[b]$ vs. full $[f]$ model (i.e. the one with education and health endowment controls in our case). He shows that

$$\begin{gathered}\widehat{\beta_b^{OCC}} = \widehat{\beta_f^{OCC}} + \sum_k \delta^{X_k} \\ \text{where } \delta^{X_k} = \widehat{\rho^{X_k}}\ \widehat{\gamma^{X_k}}\end{gathered} \tag{5}$$

and where $\widehat{\gamma^{X_k}}$ represents the impact of variable X_k on health beyond 50 and $\widehat{\rho^{X_k}}$ the link between X_k and occupation OCC. In technical terms, $\widehat{\rho^{X_k}}$ is the outcome of the OLS regression of X_k on OCC. Note also that as we are dealing with occupation dummies, everything in terms ρ_k^X is to be interpreted as deviation from what happens with the reference occupation. The point is that the two determinants of the reduction of the magnitude of our arduousness coefficients (25% for males and 45% for females) visible in Figures 6 and 7 are: *i)* the propensity of our pre-labour endowment variables (education, initial health endowment) to impact health beyond the age of 50 ($\widehat{\rho^{X_k}}$) and, also, *ii)* the propensity of people exerting different occupations to vary in terms of the level of these pre-labour entry endowment variables ($\widehat{\rho^{X_k}}$).

The full list of δ^{X_k} delivered by the Gelbach decomposition are reported in Table

5 (male) and in Table 6 (female). There is one by control variable × the number of occupations. To illustrate the outcome of the decomposition of the δ's suggested by (5), consider education (EDU) as control, and two a priori contrasted occupations: "Food preparation assistants" and "Chief Executives, senior officials and legislators". Remember that our reference profession is "Business and associate professionals". Table 5 shows that δ^{EDUC} is .01158 for (male) "Food preparation assistants" and -.01528 for "Chief Executives, senior officials and legislators". We know that these numbers combine *i)* the (negative) contribution of education to the risk of being in poor health[9] and *ii)* the propensity of these occupations to be associated with (low/high) educational attainment (relative to the reference occupation). As the vector of γ's is the **same** for the two occupations, the different δ's we have can only come difference in terms of ρ's . As we would expect, the implicit ρ is positive for "Chief Executives, senior officials and legislators" (i.e. they have a higher educational attainment than the reference occupation), whereas the implicit ρ is negative for "Food preparation assistants" (i.e. they have a lower educational attainment than the reference occupation). See Appendix A for the detailed decomposition of δ^{EDUC} as cross product of ρ's and δ's.

4.2 Endogeneity concerns

Establishing an unbiased estimate of the impact of occupation on health is challenging. In addition to measurement errors discussed above, there are several endogeneity concerns. First, health at entry in the labor market could affect the choice of the first job. Second, the deterioration of health during the career (health shocks) could lead people to abandon more arduous occupations (this is known as the simultaneity/reverse causality bias). Third, other unaccounted factors (e.g. unobserved heterogeneity in terms of risk preferences) can affect both health and occupational choice. The literature always struggles to cope with either of these problems. It is challenging to find a plausible exogenous variation of occupations - even if they are

[9]The reported γ's for education in Table 3 and 4 confirm that education's contribution is negative.

simply dichotomized between blue and white collar.

In this paper, we are not able to address fully the unobserved heterogeneity problem, but we believe we have a good chance of addressing the two other sources of endogeneity. First, we control for health at 15 and both parents' longevity. Note that we control for health *during the adolescence*, hence before the choice of an occupation and the *entry in the labor market*. Moreover, parents' longevity and vital status permit us to proxy the inherited part of people's initial health endowment and we also control for father's occupatioon to proxy the socioeconomic childhood conditions. As to the simultaneity problem, it is important to stress that we consider the impact of past occupation on health beyond 50. This means that there is a potentially important lag between the moment of exposure to a certain degree of job arduousness and the moment health is assessed. By construction, this eliminates to a large extent the risk of simultaneity bias. Moreover we use the job arduousness of the *main* job (defined as the longest spell). If we make the reasonable assumption that the other (shorter) job spells are more likely to correspond to responses to health shocks, this also contributes to limiting the risk of simultaneity/reverse causality. Note finally that we have estimated our regression (1) with alternative variables for OCC: the first, main, last occupation or the time spent in each job. The correlation between all this estimates are large (more than 80 %). This means that job mobility could bias the coefficients, but that it is unlikely to affect the overall ranking of arduousness. This is the principal aim of this paper and as such, the high correlation between our estimates means that we can be confident in our results.

4.3 Variance Decomposition

The first objective of this paper was to build an indicator of occupation arduousness. In addition to this work, it is important to ask the relative importance of working conditions versus other factors (e.g. education) to determine poor health. It is of primary importance for policy maker to know which factors affect health status the most in order to formulate effective preventive policies. This section analyzes this point by considering (1), where OCC refers to the main job and X to age, education, father's education and initial health endowment. The method used is described in Section 2 and the results are reported in Table 7 (male) and Table 8 (female).

The table is split in four columns. In column "All", we use all the countries. In column "high", "middle", "low", we use a subset of countries, classified depending on their GDP per capita,[10] to see if the decomposition differs between them. First of all, we see that occupation is a minor contributor of poor health (7.07 % for male and 4.06 % for female). The two most important variables are countries and initial health endowment. Decomposing by group of countries, we see that occupations is more important in richer countries. Nevertheless, it is still a relative marginal contributor compared to initial health endowment. This underlines that the debate of job arduousness is important but cannot ignore the importance of initial health endowment when designing policies.

4.4 Robustness

As robustness checks, we have considered two different tests. First, we have re-estimated our model considering a truncated sample where we withdrew individuals aged above age 70. This new sample restricted the analysis to people at working age or close to working age. Second, we have re-run our model using another health indicator than self-assessed health. The new health indicator is based on a principal component analysis of a set of health indicators, namely self-rated health, long-term illness, limitation in daily activities, number of limitations and limited with instrumental activities of daily living. Results are displayed in Table 9 and 10 (male) and in Table 11 and 12 (female). The new estimated coefficients are very correlated with our previous results and the professional occupation remains a relatively minor determinant to poor health.

[10] We classified as follow. High: Austria, Germany, Sweden, The Netherlands, France, Denmark, Switzerland, Belgium, Ireland, Luxembourg, Finland, Israel. Middle: Spain, Italy, Czech Republic, Portugal, Cyprus, Malta, Slovenia, Estonia. Low: Poland, Hungary, Croatia, Lithuania, Bulgaria, Latvia, Romania, Slovakia, Greece.

5 Conclusion

The first objective of this paper was to assess the possibility to rank the arduousness of occupations by using micro data on health beyond the age of 50 instead of "direct" measure of arduousness provided bv survey describing working condition by occupation. The results show that it is indeed possible to infer it from health and retirement surveys like SHARE that comprise job-history modules. We have shown that the ranking is robust to different specifications of the occupation variable as well as the inclusion of different controls, in particular pre-labour-market-entry variables (education, father's occupation and initial health endowment). The second significant finding of this paper is that whilst occupation arduousness is a significant contributor to poor health at later age, it is still (quantitatively) a minor determinant. Initial health endowment (proxied here by teenage health and the longevity of the respondent's parents) and country fixed effects (that presumably capture GDP per capita differences) explain a larger part of health differences *ceteris paribus.*

In policy terms, our findings provide justification to policy makers wishing to differentiate pension policy based on people's career. Occupations vary in terms of arduousness and, furthermore it is possible to rank them ex ante using readily available micro-data. At the same time, the research underlines the importance for late-years health of other early life, pre-labour determinants than occupation arduousness. And these comprise people's initial health endowment.

This result calls for further research, but in policy terms, it tentatively suggests that compensating individuals for poor health (and by extension longevity differences) calls for more than retirement age differentiation. It probably requires very early-stage interventions targeting childhood health inequalities, via public health and related policies.

A Full Gelbach decomposition: an illustration for two contrasted occupations

We illustrate here the decomposition of the Gelbach-estimated contribution of education δ^{EDU} to lowering the magnitude of the correlation between occupation and the risk of being in poor health. We consider two a priori contrasted main occupation "Food preparation assistants" and "Chief Executives, senior officials and legislators", bearing in mind the reference profession is "Business and associate professionals".

Table 5 (results for male respondents) shows that δ^{EDUC} is .01158 for "Food preparation assistants" and -.01528 for "Chief Executives, senior officials and legislators". As our education variable consists of a series of ISCED dummies, for each considered occupation, $\delta^{EDU} = \sum_l \delta^{ISCED_l}, l = 0, 1, 2, 3, 5, 6\ (ref. = 4)$. And following (5) each δ^{ISCED_l} is the cross-product of the γ^{ISCED_l}'s (reported in Tables 3 and 4) by ISCED specific ρ's. More precisely:

- "Food preparation assistants"

$$\underbrace{.01158}_{\delta^{EDU}} = \underbrace{.23764 \times .022}_{\rho^{ISCED\,0} \times \gamma^{ISCED\,0}} + \underbrace{-.01932 \times .026}_{\rho^{ISCED\,1} \times \gamma^{ISCED\,1}} + \underbrace{-.07331 \times .029}_{\rho^{ISCED\,2} \times \gamma^{ISCED\,2}} + \underbrace{.12980 \times .009}_{\rho^{ISCED\,3} \times \gamma^{ISCED\,3}} + \underbrace{-.16606 \times -.046}_{\rho^{ISCED\,5} \times \gamma^{ISCED\,5}} + \underbrace{-0.00234 \times -.113}_{\rho^{ISCED\,6} \times \gamma^{ISCED\,6}}$$

  Note in role of the role of $\rho^{ISCED\,3} = -.12980$ (i.e. a higher propensity to be $ISCED\,3$) and $\rho^{ISCED\,5} = -.16606$ (i.e. a lower propensity to be $ISCED\,5$) in contributing to $\delta > 0$.

- "Chief Executives, senior officials and legislators"

$$\underbrace{-.01528}_{\delta^{EDU}} = \underbrace{.00008 \times .022}_{\rho^{ISCED\,0} \times \gamma^{ISCED\,0}} + \underbrace{-.00644 \times .026}_{\rho^{ISCED\,1} \times \gamma^{ISCED\,1}} + \underbrace{-.05863 \times .029}_{\rho^{ISCED\,2} \times \gamma^{ISCED\,2}} + \underbrace{-.13735 \times .009}_{\rho^{ISCED\,3} \times \gamma^{ISCED\,3}} + \underbrace{.21522 \times -.046}_{\rho^{ISCED\,5} \times \gamma^{ISCED\,5}} + \underbrace{.02032 \times -.113}_{\rho^{ISCED\,6} \times \gamma^{ISCED\,6}}$$

Acknowledgement

This project has received financial support from the Belgian French-speaking Community (convention ARC n°18/23-088 on "Sustainable, adequate and safe pensions: financial architecture, social justice and governance"). Arno Baurin acknowledges financial support from the Fonds de la Recherche Scientifique - FNRS (FRESH Scholarship). We would like to thank the audience at at the LEGOS Seminar (Online, 2021), Belgian Summer Doctoral Workshop (Online, 2021), 9th ECINEQ Meeting (Online, 2021), EALE Conference (Online, 2021) and the 43th Journées des Economistes de la Santé Français (Online, 2021) for their helpful feedback. We are grateful to Jonathan Sicsic (discussant) for his valuable comments.

This paper uses data from SHARE Waves 1, 2, 3, 4, 5, 6, 7 and 8 (DOIs: 10.6103/SHARE.w1.710, 10.6103/SHARE.w2.710, 10.6103/SHARE.w3.710, 10.6103/SHARE.w4.710, 10.6103/SHARE.w5.710, 10.6103/SHARE.w6.710, 10.6103/SHARE.w7.711, 10.6103/SHARE.w8cabeta.001), see Börsch-Supan et al. (2013) for methodological details. The SHARE data collection has been funded by the European Commission through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARE-LIFE: CIT4-CT-2006-028812), FP7 (SHARE-PREP: GA N°211909, SHARE-LEAP: GA N°227822, SHARE M4: GA N°261982, DASISH: GA N°283646) and Horizon 2020 (SHARE-DEV3: GA N°676536, SHARE-COHESION: GA N°870628, SERISS: GA N°654221, SSHOC: GA N°823782) and by DG Employment, Social Affairs & Inclusion. Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01_AG09740-13S2, P01_AG005842, P01_AG08291, P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064, HHSN271201300071C) and from various national funding sources is gratefully acknowledged (see www.share-project.org).

Declarations of interest

None

Data and code availability

All code necessary for replication of the results are freely available at https://github.com/sirarno93/Inferring-Occupation-Arduousness-from-Poor-Health-Beyond-the-Age-of-50. The data should be download from http://www.share-project.org/. The authors remain available for all requests concerning the data and their treatments.

References

Ayuso, M., Bravo, J., Holzmann, R., 2016. Addressing Longevity Heterogeneity in Pension Scheme Design and Reform. IZA Discussion Papers 10378. Institute of Labor Economics (IZA).

Bakker, A.B., Demerouti, E., 2007. The job demands-resources model: State of the art. Journal of managerial psychology .

Barnay, T., 2016. Health, work and working conditions: a review of the European economic literature. The European Journal of Health Economics 17, 693–709.

Bassanini, A., Caroli, E., 2015. Is work bad for health? the role of constraint versus choice. Annals of Economics and Statistics , 13–37.

Bound, J., 1991. Self-reported versus objective measures of health in retirement models. Journal of Human Resources 26, 106–138.

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., Schaan, B., Stuck, S., Zuber, Sabrina, o.b.o.t.S.C.C.T., 2013. Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). International Journal of Epidemiology 42, 992–1001.

Case, A., Deaton, A.S., 2005. Broken down by work and sex: How our health declines, in: Analyses in the Economics of Aging. University of Chicago Press, pp. 185–212.

Chen, Y., Li, S., Xia, Q., He, C., 2017. The relationship between job demands and employees' counterproductive work behaviors: The mediating effect of psychological detachment and job anxiety. Frontiers in Psychology 8.

De Nardi, M., French, E., Jones, J.B., 2016. Savings after retirement: A survey. Annual Review of Economics 8, 177–204.

Etilé, F., Milcent, C., 2006. Income-related reporting heterogeneity in self-assessed health: evidence from france. Health economics 15, 965–981.

Fields, G., 2003. Accounting for income inequality and its change: A new method, with application to the distribution of earnings in the united states. Articles Chapters 22.

Fletcher, J.M., Sindelar, J.L., 2009. Estimating causal effects of early occupational choice on later health: Evidence using the psid. Technical Report. National Bureau of Economic Research.

Fletcher, J.M., Sindelar, J.L., Yamaguchi, S., 2011. Cumulative effects of job characteristics on health. Health economics 20, 553–570.

Gelbach, J.B., 2016. When do covariates matter? And which ones, and how much? Journal of Labor Economics 34, 509–543.

Han, B., Jylha, M., 2006. Improvement in depressive symptoms and changes in self-rated health among community-dwelling disabled older adults. Aging and Mental Health 10, 599–605.

Idler, E.L., Benyamini, Y., 1997. Self-rated health and mortality: a review of twenty-seven community studies. Journal of health and social behavior , 21–37.

Jusot, F., Tubeuf, S., Trannoy, A., 2013. Circumstances and efforts: how important is their correlation for the measurement of inequality of opportunity in health? Health economics 22, 1470–1495.

Lindeboom, M., 2012. Health and work of older workers, in: The Elgar Companion to Health Economics, Second Edition. Edward Elgar Publishing.

Marmot, M., Wilkinson, R., 2005. Social determinants of health. OUP Oxford.

Morefield, B., Ribar, D.C., Ruhm, C.J., 2012. Occupational status and health transitions. The BE Journal of Economic Analysis & Policy 11.

Pestieau, P., Racionero, M., 2016. Harsh occupations, life expectancy and social security. Economic Modelling 58, 194–202.

Pijoan-Mas, J., Ríos-Rull, J.V., 2014. Heterogeneity in expected longevities. Demography 51, 2075–2102.

Ravesteijn, B., Kippersluis, H.v., Doorslaer, E.v., 2018. The wear and tear on health: What is the role of occupation? Health economics 27, e69–e86.

Ravesteijn, B., van Kippersluis, H., van Doorslaer, E., 2013. Long and healthy careers? the relationship between occupation and health and its implications for the statutory retirement age. Netspar Panel Papers .

Schnittker, J., Bacak, V., 2014. The increasing predictive validity of self-rated health. PloS one 9, e84933.

Shorrocks, A.F., 1982. Inequality decomposition by factor components. Econometrica: Journal of the Econometric Society , 193–211.

Sindelar, J.L., Fletcher, J., Falba, T., Keenan, P., Gallo, W.T., 2007. Impact of first occupation on health at older ages. Technical Report. National Bureau of Economic Research.

Trannoy, A., Tubeuf, S., Jusot, F., Devaux, M., 2010. Inequality of opportunities in health in France: a first pass. Health economics 19, 921–938.

Vandenberghe, V., 2020. Health, cognition and work capacity beyond the age of 50: International evidence on the extensive and intensive margin of work. International Labour Review .

Vandenberghe, V., 2021. Differentiating retirement age to compensate for health differences. IZA Journal of Labor Policy 11, –.

Vermeer, N., Mastrogiacomo, M., Van Soest, A., 2016. Demanding occupations and the retirement age. Labour Economics 43, 159–170.

Wise, D.A., 2017. Social Security Programs and Retirement around the World: The Capacity to Work at Older Ages. Number 22 in NBER Books, National Bureau of Economic Research, Inc.

Zaidi, A., Whitehouse, E., 2009. Should pension systems recognise "Hazardous and Arduous Work"? .

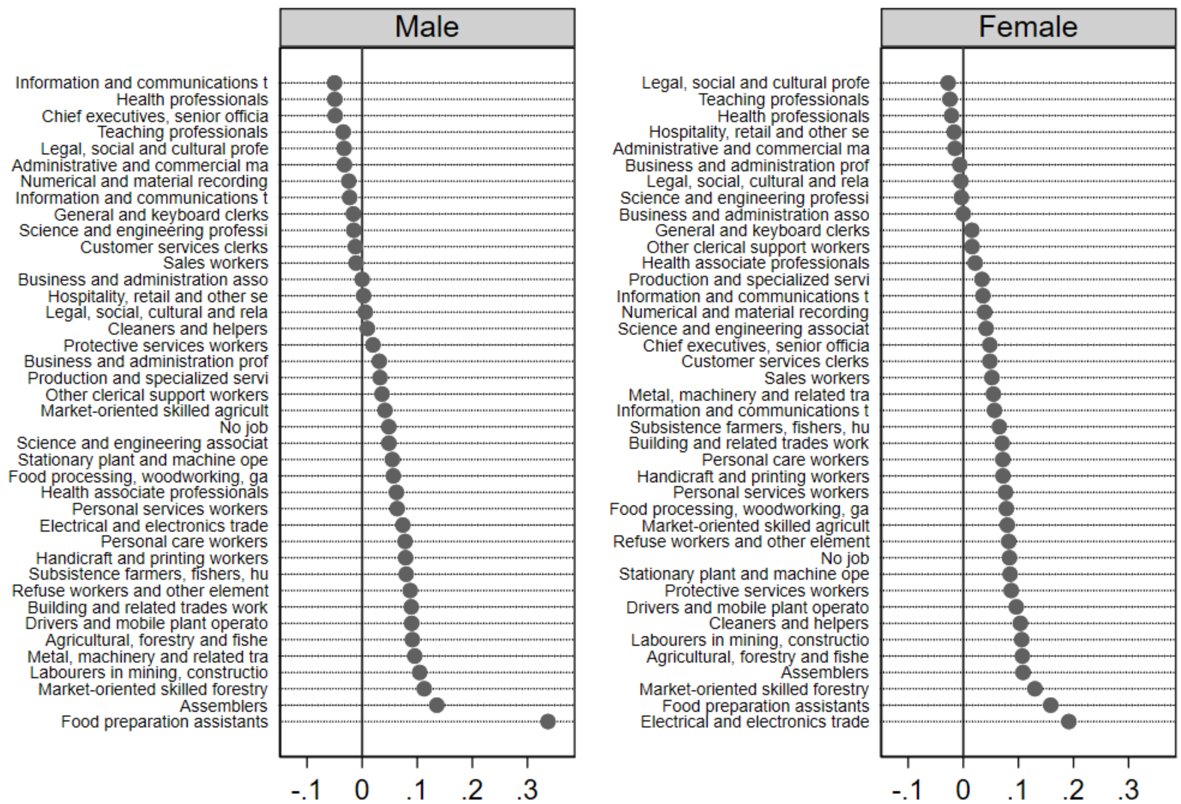

Figure 1: Regression coefficients by sex between jobs and poor health

The figure shows the coefficient of the regression (1) with main job as the occupation variable controlling for age and country fixed effects. Data are from the 7$^{\text{th}}$ wave of the SHARE survey after dropping incomplete data for the full specification.

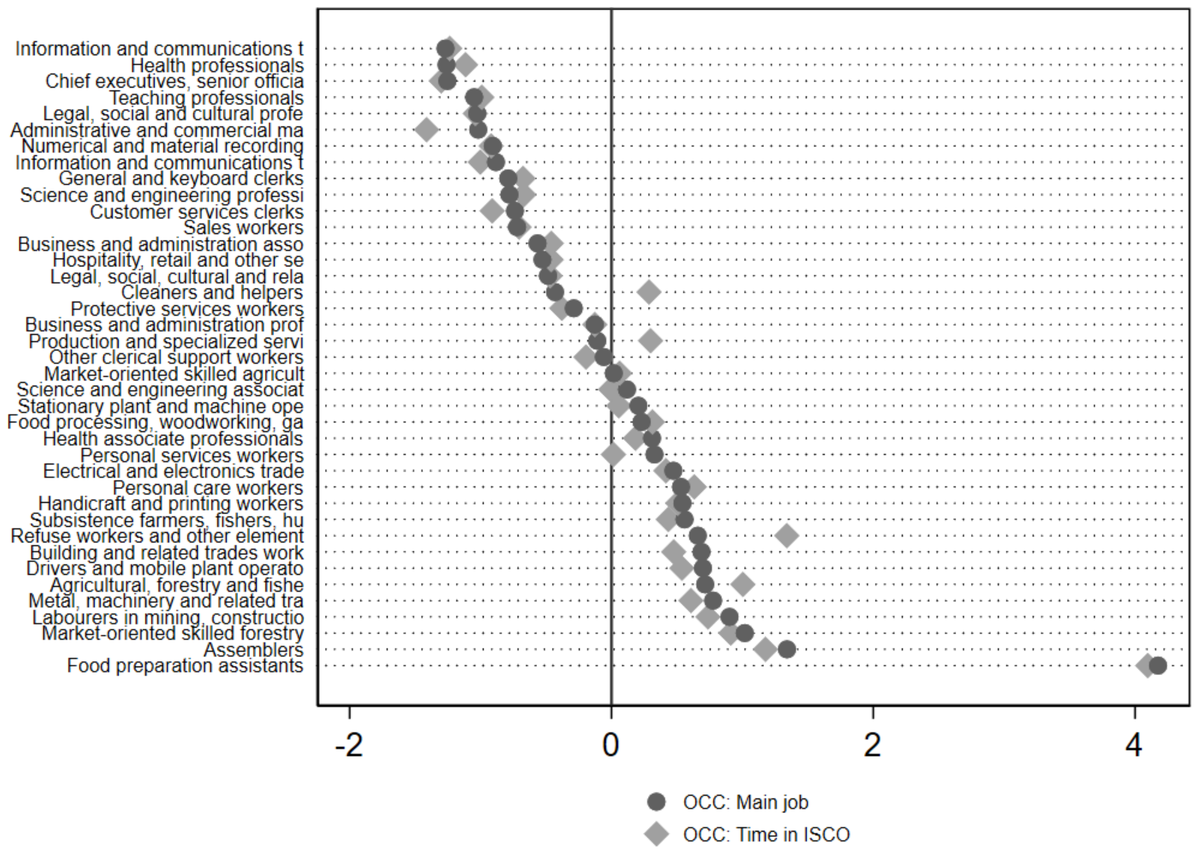

Figure 2: Correlation between coefficients (standardized) with main job and time in each ISCO code as independent variable - Male

The figure shows the correlation of the coefficients of the regression (1) with main job or time in each ISCO code as the occupation variable controlling for age and country fixed effects. Data are from the 7$^{\text{th}}$ wave of the SHARE survey after dropping incomplete data for the full specification.

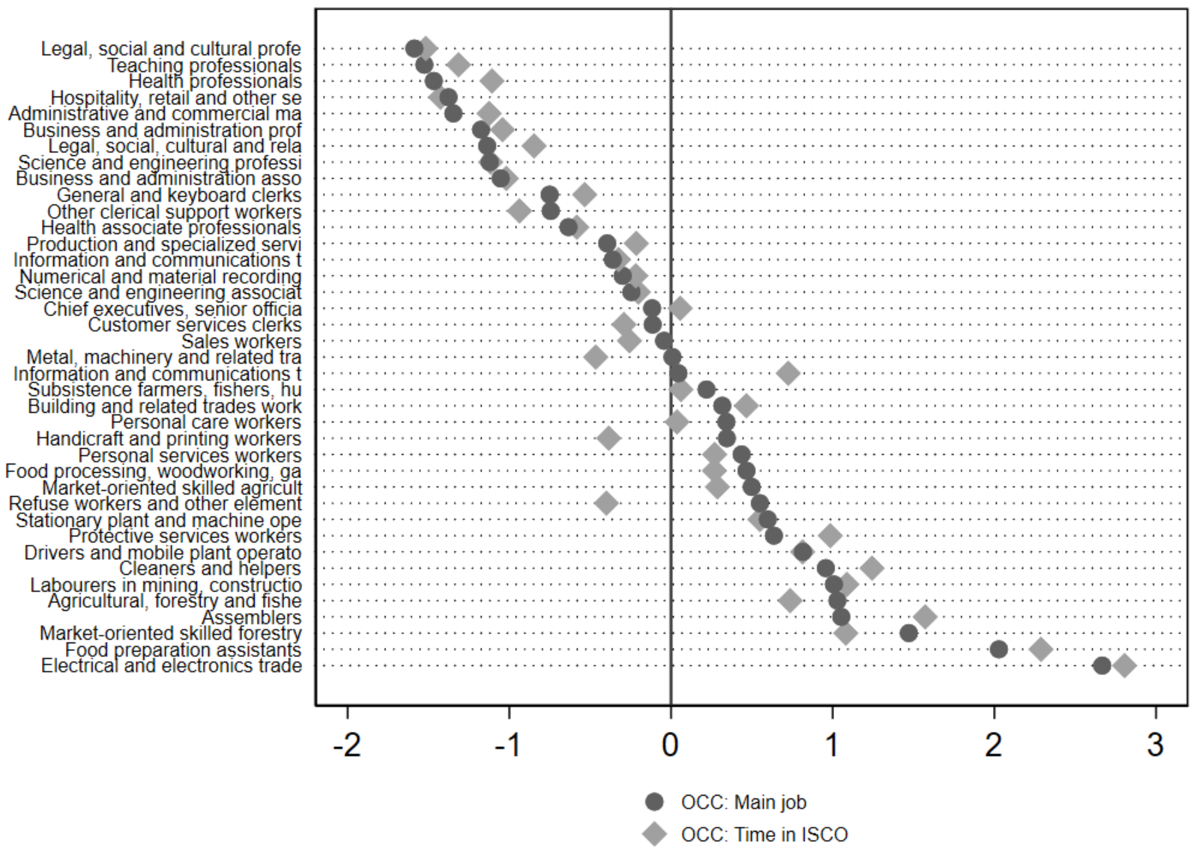

Figure 3: Correlation between coefficients (standardized) with main job and time in each ISCO code as independent variable - Female

The figure shows the correlation of the coefficients of the regression (1) with main job or time in each ISCO code as the occupation variable controlling for age and country fixed effects. Data are from the 7$^{\text{th}}$ wave of the SHARE survey after dropping incomplete data for the full specification.

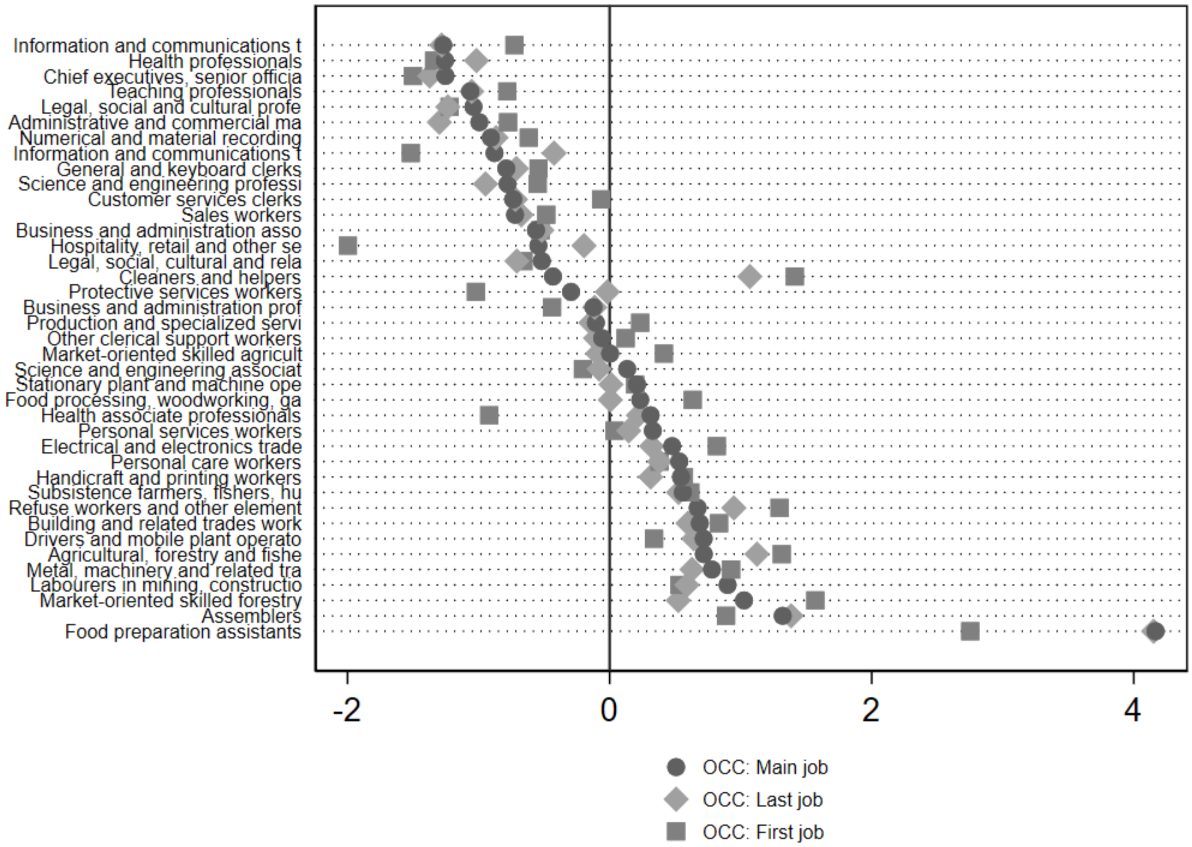

Figure 4: Correlation between coefficients (standardized) with first, main and last job as independent variable - Male

The figure shows the correlation of the coefficients of the regression (1) with main job, first job or last job as the occupation variable controlling for age and country fixed effects. Data are from the 7$^{\text{th}}$ wave of the SHARE survey after dropping incomplete data for the full specification.

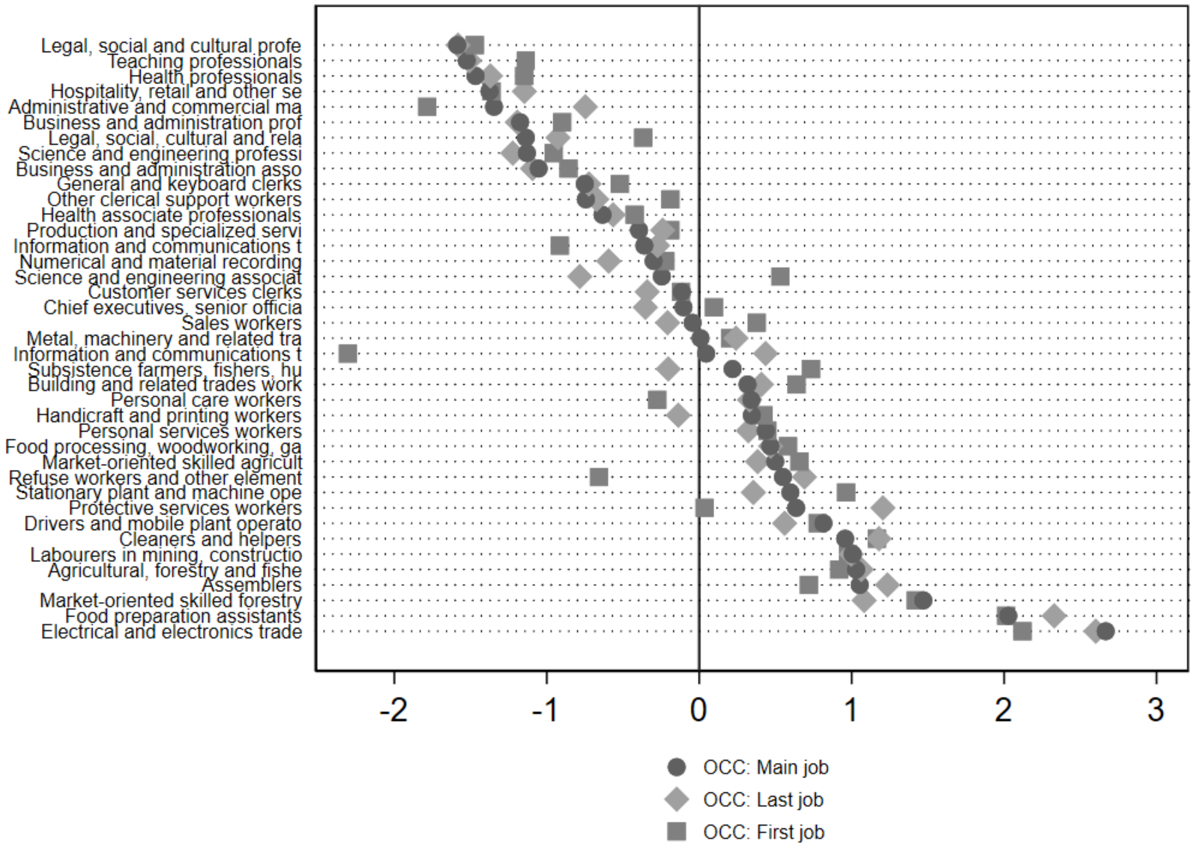

Figure 5: Correlation between coefficients (standardized) with first, main and last job as independent variable - Female

The figure shows the correlation of the coefficients of the regression (1) with main job, first job or last job as the occupation variable controlling for age and country fixed effects. Data are from the 7th wave of the SHARE survey after dropping incomplete data for the full specification.

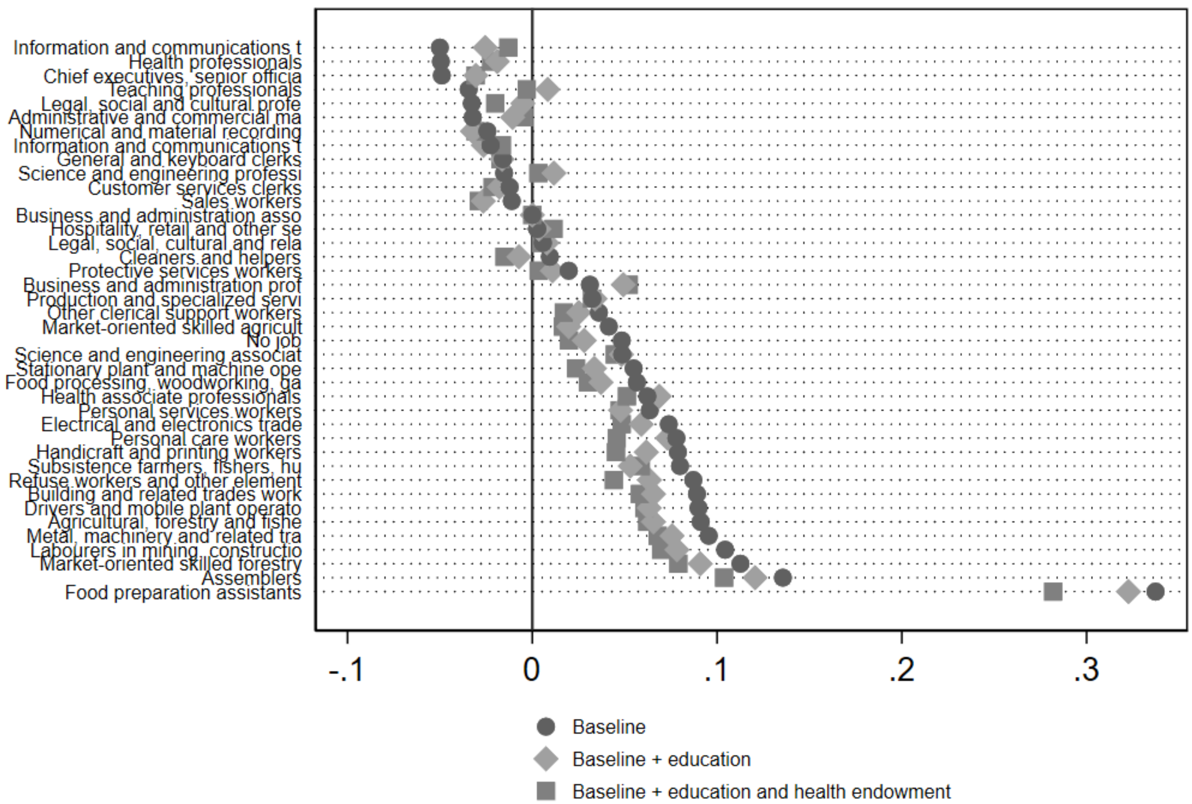

Figure 6: Reduction of the coefficients across the different regressions - Male

The figure shows the correlation of the coefficients of the regression (1) with main job as the occupation variable controlling for 1) age and country fixed effects 2) age and education and country fixed effects 3) age, education, childhood variables and country fixed effects. Data are from the 7$^{\text{th}}$ wave of the SHARE survey after dropping incomplete data for the full specification.

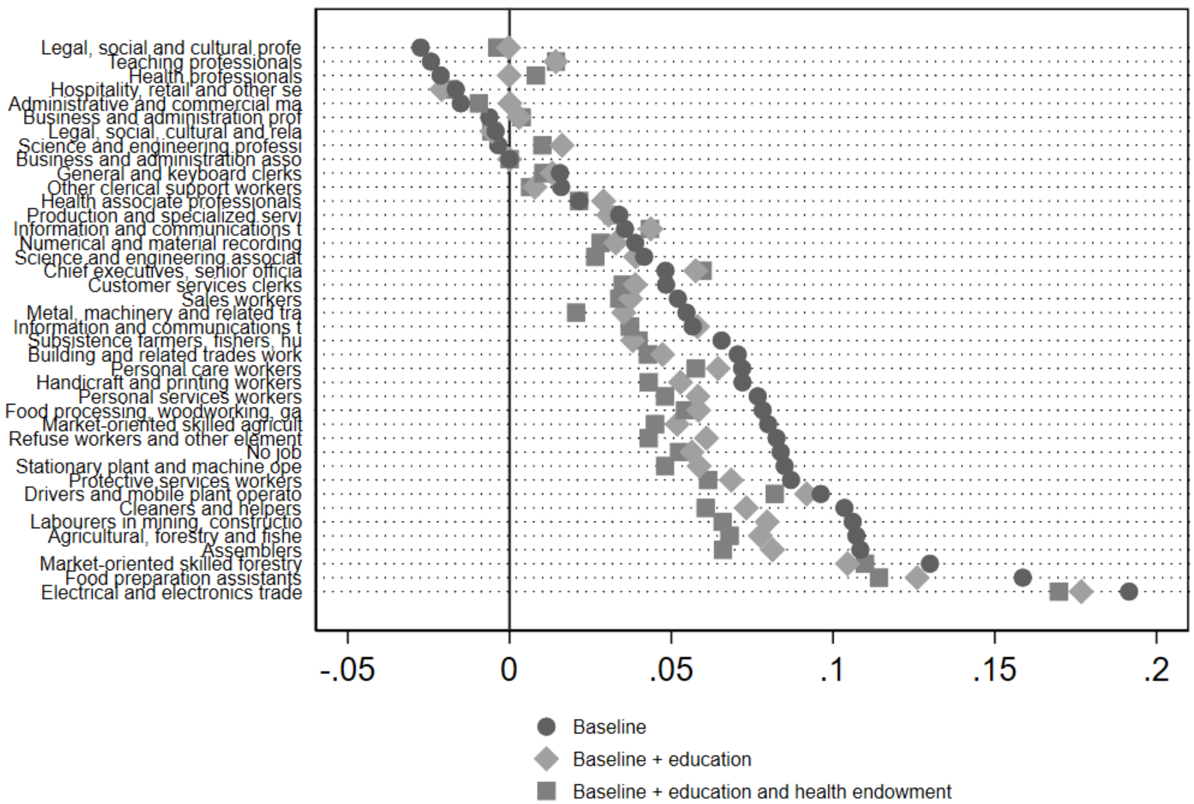

Figure 7: Reduction of the coefficients across the different regressions - Female

The figure shows the correlation of the coefficients of the regression (1) with main job as the occupation variable controlling for 1) age and country fixed effects 2) age and education and country fixed effects 3) age, education, childhood variables and country fixed effects. Data are from the 7$^{\text{th}}$ wave of the SHARE survey after dropping incomplete data for the full specification.

Table 1: Health statistics

	Observations	Percent
Excellent	3058	6.11 %
Very Good	7745	15.49 %
Good	18212	36.42 %
Fair	14690	29.37 %
Poor	6307	12.61 %
N	50 012	

Table 2: Descriptive statistics

Main job	Obs	Prop Male	Mean age	Mean health	Mean education
Chief executives, senior officials and legislators	679	.69	67.86	.74	3.94
Administrative and commercials managers	568	.52	66.51	.68	4.08
Production and specialized services managers	678	.59	69.02	.77	3.39
Hospitality, retail and other services managers	321	.51	66.74	.74	3.26
Science and engineering professionals	1566	.67	68.25	.74	4.17
Health professionals	1496	.21	65.61	.70	4.16
Teaching professionals	2960	.28	66.93	.70	4.64
Business and administration professionals	831	.35	66.14	.75	3.78
Information and communications technology professionals	356	.54	64.32	.67	4.03
Legal, social and cultural professionals	1105	.35	66.21	.70	4.26
Science and engineering associate professionals	1890	.71	67.59	.79	3.29
Health associate professionals	727	.23	66.20	.76	3.64
Business and administration associate professionals	2556	.36	66.71	.73	3.30
Legal, social, cultural and related associate professionals	458	.38	65.14	.73	3.33
Information and communications technicians	192	.70	66.07	.74	3.47
General and keyboard clerks	2350	.21	66.16	.71	3.18
Customer services clerks	652	.26	66.58	.76	2.98
Numerical and material recording clerks	1177	.30	66.99	.78	3.03
Other clerical support workers	790	.33	65.97	.76	2.96
Personal services workers	1934	.25	65.70	.80	2.57
Sales workers	2518	.24	65.82	.77	2.63
Personal care workers	1375	.06	64.87	.77	2.99
Protective services workers	573	.78	66.21	.75	2.88

Market-oriented skilled agricultural workers	1268	.46	70.87	.84	1.96
Market-oriented skilled forestry, fishery and hunting workers	284	.68	67.93	.85	2.35
Subsistence farmers, fishers, hunters and gatherers	782	.46	67.83	.90	2.20
Building and related trades workers (excluding electricians)	2120	.90	66.56	.82	2.35
Metal, machinery and related trades workers	2281	.90	66.81	.83	2.68
Handicraft and printing workers	712	.45	68.43	.83	2.48
Electrical and electronics trades workers	823	.90	66.19	.81	2.96
Food processing, woodworking, garment and other craft and related trades workers	2219	.31	67.19	.83	2.42
Stationary plant and machine operators	1727	.45	67.17	.84	2.29
Assemblers	341	.52	66.24	.85	2.49
Drivers and mobile plant operators	2093	.91	67.00	.85	2.54
Cleaners and helpers	1342	.04	66.72	.84	1.82
Agricultural, forestry and fishery labourers	391	.36	70.85	.88	1.86
Labourers in mining, construction, manufacturing and transport	1523	.56	66.58	.84	2.20
Food preparation assistants	257	.04	66.92	.91	2.035
Refuse workers and other elementary workers	365	.53	66.91	.85	2.46
No job	3732	.11	70.56	.82	1.65
N	50 012				

Comments: The data are from the SHARE Wave 7. Mean age (resp. health) refers to the mean age (resp. health) at the time of the interview. Health equals to 1 means that the individual is in less than very good health. Mean education is the mean of the ISCED code (0 to 6).

Table 3: Regression coefficients - Male

	(1) Model 1	(2) Model 2	(3) Model 3
Age	0.008***	0.008***	0.006***
	(0.000)	(0.000)	(0.000)
Chief executives, senior officials and legislators	-0.049*	-0.031	-0.031
	(0.027)	(0.027)	(0.027)
Administrative and commercial managers	-0.032	-0.011	-0.005
	(0.031)	(0.031)	(0.030)
Production and specialized services managers	0.032	0.034	0.032
	(0.028)	(0.028)	(0.027)
Hospitality, retail and other services managers	0.003	0.004	0.011
	(0.039)	(0.039)	(0.038)
Science and engineering professionals	-0.015	0.012	0.003
	(0.021)	(0.022)	(0.021)
Health professionals	-0.049*	-0.019	-0.022
	(0.030)	(0.030)	(0.030)
Teaching professionals	-0.034	0.008	-0.003
	(0.022)	(0.023)	(0.022)
Business and administration professionals	0.031	0.049	0.052*
	(0.032)	(0.032)	(0.031)
Information and communications technology professionals	-0.050	-0.025	-0.013
	(0.036)	(0.036)	(0.036)
Legal, social and cultural professionals	-0.033	-0.005	-0.020
	(0.029)	(0.029)	(0.028)
Science and engineering associate professionals	0.049**	0.048**	0.045**
	(0.020)	(0.020)	(0.020)
Health associate professionals	0.062*	0.069*	0.051

	(0.038)	(0.038)	(0.037)
Legal, social, cultural and related associate professionals	0.006	0.008	0.005
	(0.038)	(0.038)	(0.037)
Information and communications technicians	-0.023	-0.026	-0.017
	(0.043)	(0.043)	(0.042)
General and keyboard clerks	-0.016	-0.016	-0.017
	(0.026)	(0.026)	(0.026)
Customer services clerks	-0.012	-0.018	-0.021
	(0.037)	(0.037)	(0.036)
Numerical and material recording clerks	-0.024	-0.032	-0.031
	(0.030)	(0.030)	(0.029)
Other clerical support workers	0.036	0.025	0.017
	(0.034)	(0.034)	(0.033)
Personal services workers	0.064**	0.048*	0.047*
	(0.027)	(0.027)	(0.026)
Sales workers	-0.011	-0.027	-0.029
	(0.025)	(0.025)	(0.025)
Personal care workers	0.078	0.073	0.046
	(0.056)	(0.056)	(0.055)
Protective services workers	0.020	0.011	0.003
	(0.028)	(0.028)	(0.027)
Market-oriented skilled agricultural workers	0.041*	0.020	0.017
	(0.025)	(0.026)	(0.025)
Market-oriented skilled forestry, fishery and hunting workers	0.113***	0.091**	0.079**
	(0.038)	(0.038)	(0.037)
Subsistence farmers, fishers, hunters and gatherers	0.080***	0.053*	0.059**
	(0.029)	(0.029)	(0.029)
Building and related trades workers (excluding electricians)	0.089***	0.066***	0.058***
	(0.019)	(0.019)	(0.019)

Metal, machinery and related trades workers	0.095***	0.076***	0.068***
	(0.019)	(0.019)	(0.019)
Handicraft and printing workers	0.079***	0.062**	0.045
	(0.031)	(0.031)	(0.030)
Electrical and electronics trades workers	0.074***	0.059***	0.048**
	(0.023)	(0.023)	(0.023)
Food processing, woodworking, garment and other craft and related trades workers	0.057**	0.037	0.030
	(0.024)	(0.024)	(0.024)
Stationary plant and machine operators	0.055**	0.034	0.024
	(0.023)	(0.023)	(0.023)
Assemblers	0.136***	0.121***	0.104***
	(0.039)	(0.039)	(0.038)
Drivers and mobile plant operators	0.090***	0.063***	0.061***
	(0.019)	(0.020)	(0.019)
Cleaners and helpers	0.010	-0.007	-0.015
	(0.075)	(0.075)	(0.074)
Agricultural, forestry and fishery labourers	0.091**	0.066	0.062
	(0.044)	(0.044)	(0.043)
Labourers in mining, construction, manufacturing and transport	0.104***	0.078***	0.070***
	(0.023)	(0.023)	(0.022)
Food preparation assistants	0.337**	0.323**	0.282*
	(0.151)	(0.151)	(0.148)
Refuse workers and other elementary workers	0.087**	0.063*	0.044
	(0.038)	(0.038)	(0.038)
No job	0.048*	0.028	0.020
	(0.030)	(0.030)	(0.029)
ISCED 0		0.029	0.022
		(0.024)	(0.023)
ISCED 1		0.038**	0.026

	(0.018)	(0.018)
ISCED 2	0.034**	0.029*
	(0.017)	(0.017)
ISCED 3	0.011	0.009
	(0.015)	(0.015)
ISCED 5	-0.057***	-0.046***
	(0.016)	(0.016)
ISCED 6	-0.124***	-0.113***
	(0.034)	(0.034)
Health childhood		0.162***
		(0.007)
Father: Premature dead		0.088***
		(0.012)
Father: Normal dead		0.072***
		(0.012)
Mother: Premature dead		0.025***
		(0.010)
Mother: Normal dead		0.017*
		(0.010)
Father profession:		
Senior managers and professionals		-0.003
		(0.013)
Technicians and associate professionals and armed forces		-0.028*
		(0.015)
Skilled agricultural and fishery workers		-0.003
		(0.013)
Craftsmen and skilled workers		0.006
		(0.012)
Elementary occupations and unskilled workers		0.006
		(0.013)

Unknown			-0.002
			(0.014)
_cons	0.135***	0.171***	0.150***
	(0.033)	(0.036)	(0.037)
N	15 221	15 221	15 221
R^2	11.30 %	12.15 %	15.77 %

The table shows the result of regression (1). Model 1 includes age in the control variables, Model 2 adds education and Model 3 the childhood circumstances. Our coefficients are reduced when we include more control variables. The job coefficients show a clear gradient with low-arduous occupations, like Teachers, and high arduous occupations, like Refuse workers. Standard errors are in parentheses; *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$.

Table 4: Regression coefficients - Female

	(1) Model 1	(2) Model 2	(3) Model 3
Age	0.008***	0.008***	0.006***
	(0.000)	(0.000)	(0.000)
Chief executives, senior officials and legislators	0.048	0.058*	0.060*
	(0.032)	(0.032)	(0.032)
Administrative and commercial managers	-0.015	0.000	-0.009
	(0.029)	(0.029)	(0.029)
Production and specialized services managers	0.034	0.031	0.032
	(0.028)	(0.028)	(0.027)
Hospitality, retail and other services managers	-0.017	-0.021	-0.019
	(0.036)	(0.036)	(0.035)
Science and engineering professionals	-0.004	0.016	0.010
	(0.022)	(0.022)	(0.022)
Health professionals	-0.021	-0.000	0.008
	(0.017)	(0.017)	(0.017)
Teaching professionals	-0.024*	0.014	0.014
	(0.015)	(0.015)	(0.015)
Business and administration professionals	-0.006	0.003	0.004
	(0.022)	(0.022)	(0.022)
Information and communications technology professionals	0.036	0.044	0.043
	(0.037)	(0.037)	(0.036)
Legal, social and cultural professionals	-0.027	-0.000	-0.004
	(0.020)	(0.020)	(0.020)
Science and engineering associate professionals	0.042*	0.039*	0.026
	(0.022)	(0.022)	(0.022)
Health associate professionals	0.022	0.029	0.021

	(0.021)	(0.021)	(0.021)
Legal, social, cultural and related associate professionals	-0.004	-0.005	-0.006
	(0.029)	(0.029)	(0.028)
Information and communications technicians	0.057	0.058	0.037
	(0.059)	(0.058)	(0.057)
General and keyboard clerks	0.016	0.013	0.010
	(0.015)	(0.015)	(0.015)
Customer services clerks	0.048**	0.039*	0.035
	(0.023)	(0.023)	(0.023)
Numerical and material recording clerks	0.039**	0.033*	0.028
	(0.019)	(0.019)	(0.019)
Other clerical support workers	0.016	0.008	0.006
	(0.022)	(0.022)	(0.022)
Personal services workers	0.077***	0.058***	0.048***
	(0.016)	(0.016)	(0.016)
Sales workers	0.052***	0.037**	0.034**
	(0.015)	(0.015)	(0.015)
Personal care workers	0.072***	0.064***	0.058***
	(0.017)	(0.016)	(0.016)
Protective services workers	0.087**	0.069*	0.061
	(0.040)	(0.040)	(0.039)
Market-oriented skilled agricultural workers	0.080***	0.052**	0.045**
	(0.021)	(0.021)	(0.021)
Market-oriented skilled forestry, fishery and hunting workers	0.130***	0.104**	0.110**
	(0.046)	(0.046)	(0.045)
Subsistence farmers, fishers, hunters and gatherers	0.066***	0.038	0.040*
	(0.025)	(0.025)	(0.024)
Building and related trades workers (excluding electricians)	0.071**	0.047	0.043
	(0.032)	(0.032)	(0.031)

Metal, machinery and related trades workers	0.055*	0.035	0.021
	(0.032)	(0.032)	(0.031)
Handicraft and printing workers	0.072***	0.053**	0.043*
	(0.026)	(0.026)	(0.026)
Electrical and electronics trades workers	0.192***	0.177***	0.170***
	(0.049)	(0.049)	(0.048)
Food processing, woodworking, garment and other craft and related trades workers	0.078***	0.058***	0.054***
	(0.016)	(0.016)	(0.016)
Stationary plant and machine operators	0.085***	0.059***	0.048***
	(0.018)	(0.019)	(0.018)
Assemblers	0.108***	0.081**	0.066*
	(0.038)	(0.038)	(0.037)
Drivers and mobile plant operators	0.096***	0.092***	0.082**
	(0.034)	(0.034)	(0.034)
Cleaners and helpers	0.104***	0.073***	0.061***
	(0.017)	(0.017)	(0.017)
Agricultural, forestry and fishery labourers	0.107***	0.078**	0.068**
	(0.032)	(0.032)	(0.031)
Labourers in mining, construction, manufacturing and transport	0.106***	0.080***	0.066***
	(0.021)	(0.021)	(0.020)
Food preparation assistants	0.159***	0.126***	0.114***
	(0.031)	(0.031)	(0.030)
Refuse workers and other elementary workers	0.083**	0.061*	0.043
	(0.036)	(0.036)	(0.036)
No job	0.084***	0.056***	0.052***
	(0.015)	(0.015)	(0.015)
No education		0.058***	0.053***
		(0.019)	(0.019)
ISCED 1		0.057***	0.043***

	(0.015)	(0.015)
ISCED 2	0.043***	0.035***
	(0.014)	(0.014)
ISCED 3	0.011	0.007
	(0.013)	(0.012)
ISCED 5	-0.044***	-0.033***
	(0.014)	(0.013)
ISCED 6	-0.181***	-0.159***
	(0.038)	(0.037)
Health childhood		0.156***
		(0.006)
Father: Premature dead		0.067***
		(0.010)
Father: Normal dead		0.061***
		(0.010)
Mother: Premature dead		0.054***
		(0.008)
Mother: Normal dead		0.034***
		(0.008)
Father profession:		
Senior managers and professionals		-0.040***
		(0.011)
Technicians and associate professionals and armed forces		-0.014
		(0.012)
Skilled agricultural and fishery workers		-0.003
		(0.011)
Craftsmen and skilled workers		-0.009
		(0.010)
Elementary occupations and unskilled workers		0.001
		(0.010)

Unknown			-0.010
			(0.011)
_cons	0.121***	0.169***	0.175***
	(0.025)	(0.028)	(0.029)
N	20 614	20 614	20 614
pseudo R^2	13.10 %	13.58 %	17.39 %

The table shows the result of regression (1). Model 1 includes age in the control variables, Model 2 adds education and Model 3 the childhood circumstances. Our coefficients are reduced when we include more control variables. The job coefficients show a clear gradient with low-arduous occupations, like Teachers, and high arduous occupations, like Refuse workers. Standard errors are in parentheses; *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$.

Table 5: Gelbach decomposition - Male

	$\delta_{Education}$	$\delta_{Father's\,occupation}$	$\delta_{Init.Health}$
Chief executives, senior officials and legislators	-0.01528	0.00051	-0.00352
Administrative and commercial managers	-0.01772	-0.00039	-0.00947
Production and specialized services managers	-0.00159	-0.00055	-0.00227
Hospitality, retail and other services managers	-0.00092	0.00108	-0.00897
Science and engineering professionals	-0.02217	-0.00031	0.00394
Health professionals	-0.02601	-0.00048	-0.00076
Teaching professionals	-0.03568	-0.00021	0.00449
Business and administration professionals	-0.01460	-0.00078	-0.00569
Information and communications technology professionals	-0.01990	-0.00157	-0.01560
Legal, social and cultural professionals	-0.02284	-0.00044	0.01058
Science and engineering associate professionals	0.00021	0.00023	0.00357
Health associate professionals	-0.00555	-0.00067	0.01725
Legal, social, cultural and related associate professionals	-0.00182	-0.00070	0.00348
Information and communications technicians	0.00300	0.00095	-0.01005
General and keyboard clerks	0.00059	0.00083	0.00013
Customer services clerks	0.00438	0.00103	0.00378
Numerical and material recording clerks	0.00654	0.00031	-0.00047
Other clerical support workers	0.00869	0.00195	0.00837
Personal services workers	0.01248	0.00204	0.00176
Sales workers	0.01215	-0.00003	0.00575
Personal care workers	0.00405	-0.00123	0.02948
Protective services workers	0.00709	0.00124	0.00798
Market-oriented skilled agricultural workers	0.01682	-0.00031	0.00841
Market-oriented skilled forestry, fishery and hunting workers	0.01715	0.00092	0.01556
Subsistence farmers, fishers, hunters and gatherers	0.02165	-0.00034	-0.00011

Building and related trades workers (excluding electricians)	0.01873	0.00289	0.00929
Metal, machinery and related trades workers	0.01585	0.00239	0.00941
Handicraft and printing workers	0.01396	0.00357	0.01627
Electrical and electronics trades workers	0.01207	0.00254	0.01088
Food processing, woodworking, garment and other craft and related trades workers	0.01545	0.00249	0.00855
Stationary plant and machine operators	0.01693	0.00271	0.01151
Assemblers	0.01217	0.00287	0.01675
Drivers and mobile plant operators	0.02137	0.00238	0.00557
Cleaners and helpers	0.01296	0.00204	0.00956
Agricultural, forestry and fishery labourers	0.01936	0.00252	0.00703
Labourers in mining, construction, manufacturing and transport	0.02057	0.00203	0.01208
Food preparation assistants	0.01158	-0.00228	0.04611
Refuse workers and other elementary workers	0.01879	0.00169	0.02252
No job	0.01587	0.00137	0.01140

The table shows the Gelbach (2016) decomposition. The $\Delta_{Education}$ is the part due to education of the change between the occupation coefficients of column 1 and 3 in Table 3.

Table 6: Gelbach decomposition - Female

	$\delta_{Education}$	$\delta_{Father's\,occupation}$	$\delta_{Init.Health}$
Chief executives, senior officials and legislators	-0.00717	- 0.00855	0.00426
Administrative and commercial managers	-0.01181	-0.00255	0.00874
Production and specialized services managers	0.00268	-0.00492	0.00439
Hospitality, retail and other services managers	0.00348	- 0.00055	-0.00102
Science and engineering professionals	-0.01535	-0.00524	0.00693
Health professionals	-0.01634	-0.00284	-0.01017
Teaching professionals	-0.02983	-0.00381	-0.00501
Business and administration professionals	-0.00728	-0.00212	-0.00056
Information and communications technology professionals	-0.00619	-0.00212	0.00059
Legal, social and cultural professionals	-0.02105	-0.00467	0.00208
Science and engineering associate professionals	0.00212	-0.00111	0.01409
Health associate professionals	-0.00585	-0.00129	0.00723
Legal, social, cultural and related associate professionals	0.00110	-0.00109	0.00128
Information and communications technicians	-0.00073	0.00285	0.01726
General and keyboard clerks	0.00183	0.00021	0.00305
Customer services clerks	0.00756	-0.00001	0.00580
Numerical and material recording clerks	0.00466	0.00024	0.00573
Other clerical support workers	0.00636	0.00023	0.00294
Personal services workers	0.01446	0.00254	0.01166
Sales workers	0.01133	0.00178	0.00498
Personal care workers	0.00593	0.00089	0.00750
Protective services workers	0.01511	0.00354	0.00697
Market-oriented skilled agricultural workers	0.02278	0.00312	0.00906
Market-oriented skilled forestry, fishery and hunting workers	0.02023	0.00153	-0.00164
Subsistence farmers, fishers, hunters and gatherers	0.02218	0.00391	-0.00041

Building and related trades workers (excluding electricians)	0.01868	0.00245	0.00663
Metal, machinery and related trades workers	0.01535	0.00156	0.01719
Handicraft and printing workers	0.01521	0.00193	0.01188
Electrical and electronics trades workers	0.01169	-0.00270	0.01271
Food processing, woodworking, garment and other craft and related trades workers	0.01546	0.00192	0.00659
Stationary plant and machine operators	0.02092	0.00298	0.01304
Assemblers	0.02155	0.00309	0.01784
Drivers and mobile plant operators	0.00349	0.00273	0.00802
Cleaners and helpers	0.02416	0.00353	0.01518
Agricultural, forestry and fishery labourers	0.02368	0.00235	0.01315
Labourers in mining, construction, manufacturing and transport	0.02090	0.00280	0.01649
Food preparation assistants	0.02581	0.00208	0.01654
Refuse workers and other elementary workers	0.01733	0.00117	0.02106
No job	0.02219	0.00123	0.00796

The table shows the Gelbach (2016) decomposition. The $\Delta_{Education}$ is the part due to education of the change between the occupation coefficients of column 1 and 3 in Table 4.

Table 7: Variance decomposition - Male

	Part of the variance			
	All	High	Middle	Low
Education	5.56%	8.18%	2.81%	3.66%
	(1.056)	(2.275)	(1.498)	(1.341)
Occupations	7.07%	14.76%	6.57%	6.10%
	(1.717)	(4.083)	(4.523)	(3.871)
Age	15.94%	13.76%	19.37%	16.31%
	(1.490)	(2.605)	(2.846)	(2.288)
Country	35.18%	21.67%	30.95%	44.97%
	(1.874)	(3.280)	(3.382)	(4.039)
Childhood conditions	35.80%	40.81%	39.88%	28.32%
	(1.900)	(4.359)	(3.871)	(3.312)
Father's occupation	0.45%	0.83%	0.42%	0.65%
	(0.575)	(1.136)	(1.037)	(1.294)
Difference Occ High-Middle:	8.05**	(3.811)		
Difference Occ High-Low:	8.45**	(3.790)		
Difference Occ Middle-Low:	0.40	(3.029)		

The table shows the part of the variance explained by the different variables. The most important variables are childhood conditions and country.

Table 8: Variance decomposition - Female

	Part of the variance			
	All	High	Middle	Low
Education	6.63%	9.16%	3.11%	5.48 %
	(1.081)	(1.887)	(1.394)	(1.264)
Occupations	4.06%	10.05%	5.10%	3.02%
	(1.264)	(2.818)	(2.464)	(2.766)
Age	18.82%	17.28%	23.37%	18.82%
	(1.384)	(2.369)	(2.659)	(1.766)
Country	32.35%	25.70%	19.58%	38.36%
	(1.624)	(2.655)	(2.161)	(3.057)
Childhood conditions	36.79%	37.33%	47.65%	31.54%
	(1.793)	(3.280)	(2.882)	(3.063)
Father's occupation	1.36%	0.49%	1.18%	2.78%
	(0.590)	(0.812)	(0.886)	(1.266)
Difference Occ High-Middle:	4.95**	(2.266)		
Difference Occ High-Low:	7.03***	(2.289)		
Difference Occ Middle-Low:	2.08	(2.091)		

The table shows the part of the variance explained by the different variables. The most important variables are childhood conditions and country.

Table 9: Robustness test - Male

	(1)	(2)	(3)
	All obs	< 70	ACP
Age	0.006***	0.007***	0.010***
	(0.000)	(0.000)	(0.000)
Chief executives, senior officials and legislators	-0.031	-0.034	-0.020
	(0.027)	(0.036)	(0.032)
Administrative and commercial managers	-0.005	-0.000	0.004
	(0.030)	(0.041)	(0.036)
Production and specialized services managers	0.032	0.048	0.048
	(0.027)	(0.039)	(0.032)
Hospitality, retail and other services managers	0.011	-0.004	-0.031
	(0.038)	(0.050)	(0.046)
Science and engineering professionals	0.003	0.005	-0.020
	(0.021)	(0.029)	(0.025)
Health professionals	-0.022	-0.044	0.023
	(0.030)	(0.039)	(0.035)
Teaching professionals	-0.003	-0.002	0.009
	(0.022)	(0.031)	(0.027)
Business and administration professionals	0.052*	0.037	-0.012
	(0.031)	(0.043)	(0.037)
Information and communications technology professionals	-0.013	-0.008	-0.010
	(0.036)	(0.044)	(0.042)
Legal, social and cultural professionals	-0.020	-0.014	0.011
	(0.028)	(0.038)	(0.034)
Science and engineering associate professionals	0.045**	0.050*	0.022
	(0.020)	(0.027)	(0.024)
Health associate professionals	0.051	0.018	0.037

	(0.037)	(0.048)	(0.044)
Legal, social, cultural and related associate professionals	0.005	0.026	0.014
	(0.037)	(0.046)	(0.044)
Information and communications technicians	-0.017	-0.036	-0.017
	(0.042)	0.052	(0.050)
General and keyboard clerks	-0.017	-0.028	-0.027
	(0.026)	(0.034)	(0.031)
Customer services clerks	-0.021	0.026	0.026
	(0.036)	(0.046)	(0.043)
Numerical and material recording clerks	-0.031	-0.013	0.027
	(0.029)	(0.039)	(0.035)
Other clerical support workers	0.017	0.047	-0.047
	(0.033)	(0.044)	(0.039)
Personal services workers	0.047*	0.074**	0.018
	(0.026)	(0.033)	(0.031)
Sales workers	-0.029	-0.035	-0.034
	(0.025)	(0.033)	(0.029)
Personal care workers	0.046	0.049	0.028
	(0.055)	(0.065)	(0.065)
Protective services workers	0.003	0.028	-0.025
	(0.027)	(0.035)	(0.032)
Market-oriented skilled agricultural workers	0.017	0.036	0.049
	(0.025)	(0.037)	(0.030)
Market-oriented skilled forestry, fishery and hunting workers	0.079**	0.097**	0.078*
	(0.037)	(0.049)	(0.044)
Subsistence farmers, fishers, hunters and gatherers	0.059**	0.062	0.077**
	(0.029)	(0.039)	(0.035)
Building and related trades workers (excluding electricians)	0.058***	0.067***	0.065***
	(0.019)	(0.025)	0.023

Metal, machinery and related trades workers	0.068***	0.090***	0.022
	(0.019)	(0.025)	(0.022)
Handicraft and printing workers	0.045	0.043	0.062*
	(0.030)	(0.040)	(0.036)
Electrical and electronics trades workers	0.048**	0.071**	0.079***
	(0.023)	(0.030)	(0.027)
Food processing, woodworking, garment and other craft and related trades workers	0.030	0.041	0.028
	(0.024)	(0.032)	(0.028)
Stationary plant and machine operators	0.024	0.043	0.054**
	(0.023)	(0.030)	(0.027)
Assemblers	0.104***	0.141***	0.117***
	(0.038)	(0.049)	(0.045)
Drivers and mobile plant operators	0.061***	0.078***	0.042*
	(0.019)	(0.026)	(0.023)
Cleaners and helpers	-0.015	0.030	0.129
	(0.074)	(0.088)	(0.088)
Agricultural, forestry and fishery labourers	0.062	0.088	0.021
	(0.043)	(0.064)	(0.051)
Labourers in mining, construction, manufacturing and transport	0.070***	0.085***	0.078***
	(0.022)	(0.029)	(0.027)
Food preparation assistants	0.282*	0.401*	0.256
	(0.148)	(0.238)	(0.176)
Refuse workers and other elementary workers	0.044	0.067	0.060
	(0.038)	(0.048)	(0.045)
No job	0.020	0.023	0.091***
	(0.029)	(0.040)	0.035
ISCED 0	0.022	0.013	0.118***
	(0.023)	(0.034)	(0.028)
ISCED 1	0.026	0.029	0.077***

	(0.018)	(0.025)	(0.021)
ISCED 2	0.029*	0.018	0.057***
	(0.017)	(0.022)	(0.020)
ISCED 3	0.009	-0.003	0.020
	(0.015)	(0.019)	(0.018)
ISCED 5	-0.046***	-0.060***	-0.011
	(0.016)	(0.021)	(0.019)
ISCED 6	-0.113***	-0.160***	-0.088**
	(0.034)	(0.047)	(0.040)
Health childhood	0.162***	0.191***	0.115***
	(0.007)	(0.009)	(0.008)
Father: Premature dead	0.088***	0.077***	0.046***
	(0.012)	(0.014)	(0.015)
Father: Normal dead	0.072***	0.066***	0.019
	(0.012)	(0.014)	(0.015)
Mother: Premature dead	0.025***	0.023**	0.016
	(0.010)	(0.011)	(0.011)
Mother: Normal dead	0.017*	0.008	0.006
	(0.010)	(0.012)	(0.012)
Father profession:			
Senior managers and professionals	-0.003	0.008	-0.025
	(0.013)	(0.018)	(0.016)
Technicians and associate professionals and armed forces	-0.028*	-0.011	-0.019
	(0.015)	(0.020)	(0.018)
Skilled agricultural and fishery workers	-0.003	-0.003	-0.012
	(0.013)	(0.018)	(0.015)
Craftsmen and skilled workers	0.006	0.010	-0.028**
	(0.012)	(0.016)	(0.014)
Elementary occupations and unskilled workers	0.006	0.016	0.003
	(0.013)	(0.017)	(0.015)

Unknown	-0.002	0.006	-0.046***
	(0.014)	(0.018)	(0.016)
_cons	0.150***	0.077	-0.288***
	(0.037)	(0.064)	(0.044)
N	15 221	9 807	15 221
R^2	15.77 %	16.44 %	11.71 %

The table shows the robustness of our results. The first column refers to our previous result, the second to an estimation based on a sample with only the observations younger than 70 and the third column uses another indicator of health. This indicator is obtained from a PCA on several questions related to health. Standard errors are in parentheses; *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$.

Table 10: Variance decomposition (Male) - Robustness

	Part of the variance		
	All	< 70	ACP
Education	5.56%	5.83%	5.23%
	(1.056)	(1.153)	(0.912)
Occupations	7.07%	9.20%	6.37%
	(1.717)	(2.461)	(2.137)
Age	15.94%	6.76%	33.70%
	(1.490)	(1.339)	(2.476)
Country	35.18%	40.05%	34.96%
	(1.874)	(2.379)	(2.193)
Childhood conditions	35.80%	37.89%	18.85%
	(1.900)	(2.375)	(2.122)
Father's occupation	0.45%	0.27%	0.89%
	(0.575)	(0.730)	(0.709)

The table shows the robustness of our results. The first column refers to our previous result, the second to an estimation based on a sample with only the observations younger than 70 and the third column uses another indicator of health. This indicator is obtained from a PCA on several questions related to health.

Table 11: Robustness test - Female

	(1)	(2)	(3)
	All obs	< 70	ACP
Age	0.006***	0.007***	0.011***
	(0.000)	(0.001)	(0.000)
Chief executives, senior officials and legislators	0.060*	0.077*	-0.031
	(0.032)	(0.042)	(0.039)
Administrative and commercial managers	-0.009	-0.008	-0.041
	(0.029)	(0.036)	(0.035)
Production and specialized services managers	0.032	0.039	-0.013
	(0.027)	(0.038)	(0.034)
Hospitality, retail and other services managers	-0.019	-0.045	0.020
	(0.035)	(0.047)	(0.043)
Science and engineering professionals	0.010	-0.008	0.006
	(0.022)	(0.031)	(0.027)
Health professionals	0.008	0.007	0.029
	(0.017)	(0.022)	(0.021)
Teaching professionals	0.014	0.009	0.024
	(0.015)	(0.020)	(0.019)
Business and administration professionals	0.004	-0.014	-0.003
	(0.022)	(0.029)	(0.027)
Information and communications technology professionals	0.043	0.064	-0.010
	(0.036)	(0.047)	(0.045)
Legal, social and cultural professionals	-0.004	0.006	0.000
	(0.020)	(0.026)	(0.035)
Science and engineering associate professionals	0.026	0.050*	0.020
	(0.022)	(0.029)	(0.027)
Health associate professionals	0.021	0.014	-0.006

	(0.021)	(0.027)	(0.026)
Legal, social, cultural and related associate professionals	-0.006	-0.011	0.001
	(0.028)	(0.036)	(0.024)
Information and communications technicians	0.037	-0.009	-0.062
	(0.057)	(0.075)	(0.071)
General and keyboard clerks	0.010	0.002	-0.013
	(0.015)	(0.020)	(0.018)
Customer services clerks	0.035	0.055*	0.015
	(0.023)	(0.031)	(0.028)
Numerical and material recording clerks	0.028	0.029	0.017
	(0.019)	(0.025)	(0.023)
Other clerical support workers	0.006	-0.003	0.009
	(0.022)	(0.028)	(0.027)
Personal services workers	0.048***	0.050**	0.045**
	(0.016)	(0.021)	(0.020)
Sales workers	0.034**	0.022	0.030
	(0.015)	(0.020)	0.019
Personal care workers	0.058***	0.057***	0.041**
	(0.016)	(0.021)	(0.020)
Protective services workers	0.061	0.062	0.009
	(0.039)	(0.054)	(0.048)
Market-oriented skilled agricultural workers	0.045**	0.047	0.060**
	(0.021)	(0.031)	(0.025)
Market-oriented skilled forestry, fishery and hunting workers	0.110**	0.130*	0.075
	(0.045)	(0.069)	(0.056)
Subsistence farmers, fishers, hunters and gatherers	0.040*	0.037	0.128***
	(0.024)	(0.033)	(0.030)
Building and related trades workers (excluding electricians)	0.043	0.042	0.070*
	(0.031)	(0.041)	(0.039)

Metal, machinery and related trades workers	0.021	0.039	0.024
	(0.031)	(0.041)	(0.039)
Handicraft and printing workers	0.043*	0.059*	0.062**
	(0.026)	(0.036)	(0.032)
Electrical and electronics trades workers	0.170***	0.198***	0.159***
	(0.048)	(0.061)	(0.059)
Food processing, woodworking, garment and other craft and related trades workers	0.054***	0.061***	0.062***
	(0.016)	(0.022)	(0.020)
Stationary plant and machine operators	0.048***	0.049**	0.063***
	(0.018)	(0.025)	(0.023)
Assemblers	0.066*	0.048	0.103**
	(0.037)	(0.047)	(0.046)
Drivers and mobile plant operators	0.082**	0.100**	0.085**
	(0.034)	(0.044)	(0.042)
Cleaners and helpers	0.061***	0.073***	0.079***
	(0.017)	(0.022)	(0.021)
Agricultural, forestry and fishery labourers	0.068**	0.121***	0.086**
	(0.031)	(0.049)	(0.039)
Labourers in mining, construction, manufacturing and transport	0.066***	0.063**	0.034
	(0.020)	(0.027)	(0.025)
Food preparation assistants	0.114***	0.139***	0.105***
	(0.030)	(0.042)	(0.038)
Refuse workers and other elementary workers	0.043	0.038	0.076*
	(0.036)	(0.048)	(0.044)
No job	0.052***	0.046**	0.062***
	(0.015)	(0.021)	(0.018)
ISCED 0	0.053***	0.084***	0.133***
	(0.019)	(0.030)	(0.023)
ISCED 1	0.043***	0.092***	0.084***

	(0.015)	(0.021)	(0.018)
ISCED 2	0.035***	0.075***	0.071***
	(0.014)	(0.019)	(0.017)
ISCED 3	0.007	0.033**	0.030*
	(0.012)	(0.017)	(0.016)
ISCED 5	-0.033***	-0.019	-0.011
	(0.013)	(0.018)	(0.016)
ISCED 6	-0.159***	-0.217***	-0.027
	(0.037)	(0.050)	(0.046)
Health childhood	0.156***	0.186***	0.121***
	(0.006)	(0.008)	(0.007)
Father: Premature dead	0.067***	0.057***	0.031***
	(0.010)	(0.011)	(0.012)
Father: Normal dead	0.061***	0.045***	0.008
	(0.010)	(0.011)	(0.012)
Mother: Premature dead	0.054***	0.045***	0.028***
	(0.008)	(0.009)	(0.010)
Mother: Normal dead	0.034***	0.020**	-0.004
	(0.008)	(0.010)	(0.010)
Father profession:			
Senior managers and professionals	-0.040***	-0.041***	-0.041***
	(0.011)	(0.015)	(0.014)
Technicians and associate professionals and armed forces	-0.014	-0.019	-0.000
	(0.012)	(0.017)	(0.015)
Skilled agricultural and fishery workers	-0.003	-0.009	-0.030**
	(0.011)	(0.015)	(0.013)
Craftsmen and skilled workers	-0.009	-0.019	-0.023*
	(0.010)	(0.013)	(0.012)
Elementary occupations and unskilled workers	0.001	-0.002	-0.010
	(0.010)	(0.014)	(0.013)

Unknown	-0.010	-0.010	-0.013
	(0.011)	(0.015)	(0.014)
_cons	0.175***	0.060	-0.374***
	(0.029)	(0.049)	(0.036)
N	20 614	13 374	20 614
pseudo R^2	17.39 %	17.37 %	14.28 %

The table shows the robustness of our results. The first column refers to our previous result, the second to an estimation based on a sample with only the observations younger than 70 and the third column uses another indicator of health. This indicator is obtained from a PCA on several questions related to health. Standard errors are in parentheses; *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$.

Table 12: Variance decomposition (Female) - Robustness

	Part of the variance		
	All	< 70	ACP
Education	5.48 %	7.62%	6.98 %
	(1.264)	(1.171)	(1.158)
Occupations	3.02%	4.79 %	5.03%
	(2.766)	(1.418)	(1.364)
Age	18.82%	7.95 %	42.70%
	(1.766)	(1.407)	(2.171)
Country	38.36%	40.38 %	27.50%
	(3.057)	(1.909)	(1.438)
Childhood conditions	31.54%	38.06 %	17.54 %
	(3.063)	(1.952)	(1.599)
Father's occupation	2.78%	1.20 %	0.25%
	(1.266)	(0.674)	(0.417)

The table shows the robustness of our results. The first column refers to our previous result, the second to an estimation based on a sample with only the observations younger than 70 and the third column uses another indicator of health. This indicator is obtained from a PCA on several questions related to health.

INSTITUT DE RECHERCHE
ÉCONOMIQUES ET SOCIALES

Place Montesquieu 3

1348 Louvain-la-Neuve

ISSN 1379-244X D/2022/3082/05

UCLouvain

IRES | Louvain Institute of Data Analysis and Modeling in economics and statistics