Job penibility: measurements and policy discussions

Thesis presented by
Arno Baurin

Supervisor
Jean Hindriks (UCL)

Reader
Vincent Vandenberghe (UCL)

Academic year 2017-2018

Master 120 en Sciences économiques, Orientation générale, Finalité spécialisée

Ecole d'économie de Louvain/UCL • Place Montesquieu 3 • 1348 Louvain-la-Neuve
Département des Sciences économiques/UNamur • Rempart de la Vierge 8 • 5000 Namur

Job penibility: measurements and policy discussions

Arno BAURIN[1]

Abstract

The aim of this paper is to discuss the possible recognition of the job penibility[2] in the pension system. This discussion can be investigated in several ways and we have chosen to concentrate on the objectivation of the job penibility. This master thesis is, therefore, a paper that searches for coefficients of job impact on health. We investigate health with two kinds of dependent variables, a subjective one (self-rated health) with the data of the "European Working Conditions Survey, 2015" and an objective one (life expectancy) with the data of the "US National Longitudinal Mortality Study, 80s-90s". This second database contains 1 835 072 observations and 364 possible occupations. We find significant job effects on self-rated health and significant difference of life expectancy across professions (e.g., 57,18 years of life expectancy at 25 for lawyers (men) and only 48,73 years for "food and beverage serving workers" (men)). We continue by discussing, from a philosophical and economic point of view, whether we should compensate, how and who must finance it.

Keywords: retirement, hazardous job, arduous job, penible job, life expectancy by occupations, health inequalities, work environment

This paper uses data provided from the public-use file of the National Longitudinal Mortality Study (hereafter NLMS), the UK Data Archive and includes information from the O*NET 22.2 Database by the U.S. Department of Labor, Employment and Training Administration (USDOL/ETA). O*NET® is a trademark of USDOL/ETA and is used under the CC BY 4.0 license. The NLMS file is provided to persons interested in research by the U.S. Census Bureau. The author has modified all or some of the O*Net information and USDOL/ETA has not approved, endorsed, or tested these modifications. The information and views expressed in this paper are those of the author and do not necessarily reflect the views of the National Longitudinal Mortality Study, the Bureau of the Census, or the project sponsors: the National Heart, Lung, and Blood Institute, the National Cancer Institute, the National Institute on Aging, and the National Center for Health Statistics and of any providing institutions. Responsibility for the information and views expressed herein lies entirely with the author.

[1] E-mail: arno.baurin@hotmail.com

[2] We use the term "penibility" derived from the French word "pénibilité" which has to be understood as "arduous (mental and/or physical) and hazardous".

Acknowledgments

This master thesis is the end of a process of education which began long ago in preschool in Montigny-le-Tilleul. Although all the people, that I have met, have – to some extent – contributed to it; some deserve more recognition and I will, by quoting their names, acknowledge them.

Therefore, I would like to chronologically thank Michèle Dutz, Jean-Louis Choquet, Raoul Francart and Jean-Michel Genot for their helps and spirits during my secondary studies. I would also like to express my honest gratitude to Eddy Corbisier, Alain De Crombrugghe and Paul Wynants for their guidance and helps during my Bachelor at the University of Namur.

This master thesis would not have been possible without the original idea of objectivating job penibility given by Jean Hindriks. I show him my appreciation for the time and expertise he spent with me on this work. I am also grateful to Vincent Vandenberghe for being the reader of my master thesis. Elodie Orth-Soane, Véronique Vanlanduyt, Stéphane Berberat and Philippe Bocquier also deserve a mention for the help provided. Ideas are of no use without data; therefore, the statistical offices and institutions that provided them to me also merit a mention.

Last, but not least, I would like to thank my parents, Katia Cini and Frédéric Baurin, and my girlfriend, Smaranda Molle, for their unconditional support and patience all along the last months and years.

Table of contents

List of Figures

List of Tables

List of abbreviations

KM: Kaplan-Meier

LE: life expectancy

PCA: principal component analysis

SRHS: self-rated health status

Introduction & Literature

"Every worker has the right to working conditions which respects his or her health, safety and dignity".
Charter of Fundamental Rights of the European Union (2000: Art. 31, §1)

"There are relatively few studies on work conditions, such as exposure to physical demands, and health in the economics literature"
Fletcher et al. (2011: 555)

1 Introduction

Our pension system is at risk due to the increasing dependency ratio and it brings the discussion (in Europe) to raise the retirement age. However, people claim that workers in certain professions are not able to continue due to the heavy nature of their occupations and the deterioration of their health they have suffered of. This situation pushes a real demand for governments to legislate exemptions for these particular workers. However, one problem lies in the impossibility of arriving at a clear classification of what these professions are and, unfortunately, the special current schemes that exist (across Europe) do not lead to a straightforward classification (Zaidi & Whitehouse, 2009; Buelen et al., 2014). Interestingly, even at national level, only five of the 28 European countries, that recognize some jobs as penible, have an official definition of its meaning. This shows the absence of an in-depth thinking about this topic (Natali et al., 2016). As Borghans & ter Weel (2012) have argued, creating a list based on self-reported job penibility will not work as everyone will have an incentive to include their own job and, so, an urgent need of research in this particular field is, clearly, required to help governments achieving their aims. For example, the Belgian government said that "all difference […] in retirement should be *objectively* justified" (Accord de gouvernement, 2014: 36; own translation and italics added). The goal of this master thesis is to find how to objectivate such a concept in order to foster the current political discussion. Although this purpose is sufficient *per se*; it should also be known that "inequalities in work environment are one of the major sources of inequalities in health" (Hämmig & Bauer, 2013: 2) and new evidences about this are always desirable.

The starting point of this master thesis was the discovery of the Delwarde's work (2004) in which, using the retirement public sector database for the years between 1998 and 2002, he found that Belgian teachers live one year longer than the rest of the public sector officials. From this moment onwards, we wanted to investigate the possibility to clearly objectivate the debate, which is an underdeveloped subject in the economic literature. In order to better show the relevance of our task, we have done a broad data review and have also inquired into the "epidemiological" literature. This last inquiry was made to understand the roots of the health impacts.

Studying the impact of a job on health is a very challenging task for two reasons. The first lies in the impossibility of defining, in an objective manner, the word "health" (see Section 2.1) and the second is almost the impossibility to find non-biased coefficients (see Section 5). Regarding the first problem, we choose to base our research on two indicators (self-rated health status (SRHS) and life expectancy[3] (LE)) which are widely used in the literature. The second problem lies in the lack of available data to estimate without interference the work impact. The two main nuisances come from the fact that people self-select their jobs and that people can change from jobs; it is, therefore, almost impossible to find out unbiased estimates without experimental data. However, in order to create those data, one should randomize all "kinds" of people between all occupations and, then, wait long enough to see long-lasting health effect with the hope that all other health behaviours would not have changed during this period. It is, therefore, not plausible to find totally unbiased estimates and we will discuss the limits of our paper in Section 5. However, and regarding the current retirement debate, the most an economist could do is to search for the best *possible* unbiased coefficients. Therefore, the purpose of this master thesis is to search for them keeping in mind their limits. To this aim, we will use two databases to investigate SRHS and LE by occupations.

Our first database is the "European Working Conditions Survey, 2015", which is a survey about SRHS and working conditions. This database allows us to enter the subject by investigating relations between "general" health and occupations. The second database is the "National Longitudinal Mortality Survey, 80s-90s" from the United States which is, contrary to the first, a longitudinal database. This will help us to calculate an objective indicator of health, LE, using the Kaplan-Meier (KM) technique. The KM estimator is the most frequently used technique in survival analysis (Jager et al., 2008) and is used to compute survival estimates for follow-up data in which all the observations do not experience the risk (in our case, the death) before the end of the follow-up. It is important to notice that, to the best of our knowledge, it is the first time someone attempts to create LE by occupations at such a detailed level. In fact, we were able to create 91 different LE. After this step, we will use principal components analysis (PCA) to "divide" jobs based on their characteristics with the "O*Net Data". These data detail the work requirements and conditions and allow us to classify occupations. Based on this codification, we will investigate if "hard" jobs have some particular features.

[3] Life expectancy "is defined as how long, on average, [someone at a specific age] can expect to live, if current death rates do not change. However, the actual age-specific death rate of any particular birth cohort cannot be known in advance. If rates are falling, actual life spans will be higher than life expectancy calculated using current death rates" (OECD, 2018).

After these empirical sections, we will investigate from a more philosophical and economic point of view why governments should compensate people for the difficulty of a job that they have (freely) chosen and then, we will conclude. We have to add that this is a very large issue which can be scrutinized by means of four questions; the first one concerns the possibility to measure the penibility of an occupation. The second is to know why governments should compensate people; the third one determines how to compensate them (e.g., early retirement, special working arrangements, etc). The last one is to identify who should pay for this compensation. This master thesis will review in depth only the first question and will investigate briefly the others leaving them open for further research. The rest of the thesis is divided into three chapters: chapter "Introduction & Literature" presents the literature, chapter "Empirical analysis" presents our empirical analysis on SRHS using the "European Working Conditions Survey, 2015", on LE using the "US National Longitudinal Mortality Survey, 80s-90s" and on the roots of the job penibility using the "O*Net Data" and chapter "Policy discussions & Conclusion" discusses whether the public authority should or should not compensate people.

2 Literature

2. 1 What is health?

First and foremost, it is always important to clearly define a subject before entering it. This is the reason why we should explain the meaning of "health", as this whole paper will investigate the relations between it and jobs. This should not only be seen as "an abstract exercise but [as] a way of shaping the world metaphysically and structuring the world politically" (Callahan, 1973: 78). In fact, the way we define "health" has a lot of implications. For example, the fact of recognizing "attention deficit hyperactivity disorder" as a disease or not, is a consequence of our description of "health" (Lanzerath, 2012). Firstly, it should be mentioned that there is not unique (scientific) definition (Jylhä, 2009). However, it could be assumed that its meaning is not arbitrarily chosen; but depends on the "general" cultural understanding of this word (ibid.). In fact, the problem lies in the various interested parties that argue for different definitions from the more theoretical to the more practical (Law & Widdows, 2007). The "official" one is as follows: "Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity" (World Health Organization, 1946). This characterisation has been criticized in various ways, notably for being too restrictive due to the adjective "complete" (Huber et al., 2011) or for its possible abuse of defining some social problems as "sickness" (Callahan, 1973). It must also be observed that this definition includes more than the absence of disease, but requires the presence of well-being. Other definitions exist like the "naturalistic approach", which says that health could be "read off" and that nature creates "norms" and deviations from these norms should be seen as a disease (Lanzerath, 2012). Another is the "conventionalist approach" which states that "health" is more a convention within a society (ibid.). In the end, we could say that health is "a constantly renewed construction, realised by different observers (e.g., individuals or health professionals) and influenced by their cultural context or the time they live in" (Hafen, 2016: 438). This first review of the signification of "health" was important for us to keep in mind that all possible health indicators always bring the problem of subjectivity. Consequently, our selection of indicators (self-rated health and life expectancy) has the same drawback *than all other possible measures*.

2. 2 History of health inequality

Before entering the different literatures, we will sum up the history of health inequalities (Adler & Ostrove, 1999) to put this review literature into context. Although these inequalities were a characteristic of all societies; their presences were, surprisingly, only discovered in the 19th century, due to better availability of data. As medicine was making progress, it was believed that the health gap was decreasing but the "Black report" (in 1980) put an end to this thought as it described the opposite in the United Kingdom. It was also thought that health does not improve (with income) above the poverty threshold and so, medicine literature did not enquire about the social gradient. However, around 1985, the threshold model was challenged by the "Whithehall study" which showed that health, among the British civil servants, improves at each higher step of occupational grade and so the model was contradicted. It was time to make new studies and the "World Health Organization" now defines the social gradient as follows: "[…] Within countries, the evidence shows that in general the lower an individual's socioeconomic position the worse his health. […] This is a global phenomenon, seen in low, middle and high-income countries. The social gradient in health means that health inequalities affect everyone" (WHO, 2017). From 1995 to today, medicine literature has explained the different pathways by which socio-economic status influences health (see Figure 1 for a simplified model) which is done by a lot of variables running from childhood conditions (education,…) and living conditions habits (junk food, sports,…) to job occupation (hard work,…). We will only describe the literature relating to the work impact in this thesis as this is its aim.

2. 3 Health & occupation: Economic literature

On the economic side, Case & Deaton (2003) used the Le Grand & Muurinen (1985) model with three capitals: health (body), human (education) and financial (wealth). This model explains the trade-offs between them and demonstrates the struggle of poor people having less human and financial capitals which leads them to hang on more severely to their health capital. Following this model, they illustrated empirically, with the 16 cross-sectional waves of the "US National Health Interview Survey, 1986-2001", that people with a manual occupation have worse SRHS, controlling for education and income (both having protective effect). They also showed that people with worse health self-selected

them out of the labour market. They used this last evidence, and the stratification of health and income, to describe a flow running from poor health to low income. Following this paper, Choo & Denny (2006), using the "Canadian Community Health Survey, 2000-2001", discovered the same empirical result in Canada. They investigated further to know the part played by life-style decisions (obesity and smoking) and chronic conditions (asthma, allergies, cancer,…) and integrated these three factors into their regression. They found significant coefficients for these factors. However, their job coefficients also remained significant and fell by only 10 %. Contrary to the two previous studies which use cross-sectional data, Fletcher et al. (2011) used panel data ("US Panel Study of Income Dynamics" for the years from 1968 to 2001) which allows them to study the impact of the last 5-year environmental and physical job demands. The result of their model is that an occupation with more physical and environmental demands results in more deterioration in health; although, the consequences are different by demographic groups (age, gender and race). Morefield et al. (2011) also used this dataset (for the years from 1984 to 2007) and a five-year history to detect the probability of change of SRHS from good into bad health. They found that blue-collars have much more probability to enter into bad health but did not however find a significant difference in the other side (from bad to good health). These four studies are complementary in demonstrating that some jobs are more harmful than others. Although this is interesting to know; there is a lack of explanation behind their result. Before turning to the epidemiological literature, which explains more the foundations behind the different coefficients, we will first make a broad data review in order to better grasp the importance of our study.

2. 4 Health & occupation: Data review

On a more statistical side, as explained in the introduction, Delwarde (2004), who explored the retirement public sector database for the years from 1998 to 2002, found that Belgian teachers live one year more than average in the public sector. However, his conclusion is not controlled for heterogeneity in the public sector and he acknowledged himself that "the low mortality of teachers could be caused by the high percentage of high school and university alumni" (Delwarde, 2004: 31; own translation). Luy et al. (2015) estimated LE by vocational class based on a sample carried out between 1984 and 1986 in the West German population. Vocational class were defined according to the KldB-92 classification which does not differentiate between chief and senior positions or between workers and clerks in an occupational group. In their study, a man in "social service and education"

has a LE of 40,6 at 40 years and a man in transport industry has a LE of only 31,9 years. Cambois et al. (2011) analysed the French case and tried to distinguish LE and healthy LE, using the "French Health Survey, 2002-2003", between 7 broad occupations. Their study also demonstrates the difference of LE between occupations with one for a farmer (men) of 30,9 years at 50 years instead of one of 32,2 for the "high qualifications" (men). Geoffroy-Perez (2006) described, also in France, the mortality and causes of death for 53 sectors, using the "Demographic Permanent Sample", which is a sample of 1 % of the French population, for the period between 1968 and 1999. We can summarize her study by saying that some sectors are more at risk for some causes of death, e.g., a 70 % more risk of oesophagus cancer for the sector of "melting and metal works" compared to the rest of the working population. The different mortality risks are not particularly linked; for example, in the "leather and shoes sector" there is a 44 % less chance to have a stomach cancer, but a 69 % more risk of prostate cancer. It must be noted that this study gives a short explanation for every sector to explain the roots of some causes of death. This study goes along the same lines as the one done by Gass & Bopp (1997) who studied the mortality of men in Switzerland between 1979 and 1983 and gave a mortality and cause of death risk compared to the rest of the population. In their study, an accountant, for example, has a 38,4 % less chance of death than the rest of the population. Our last study comes also from Switzerland with Gubéran & Usel (2000), who detailed the percentage of invalidity by profession in Geneva for the period between 1970 and 1990. They detected a huge difference between professions running from 2,9 % in scientists to 40 % in building workers. Although this difference should not be overlooked, it should be warned that it is biased upwards. In fact, the disability assessment in Switzerland depends on the main occupation and the same illness could turn or not into disability regarding professions. However, they also calculated the percentage of death between 45 and 65 years old and found huge differences with low mortality rates for e.g. teachers (9,7 %) and high mortality rates for e.g. car workers (22,8 %).

2. 5 Health & occupation: Epidemiological literature

Finally, we will end our review of the literature with the epidemiological one, which shows a bigger interest between certain job characteristics and the workers' health. In some ways, this goes beyond the aim of this paper which just wants to objectivate the penibility without knowing in depth its roots. Nevertheless, it is important to review this literature in order to understand how job affects

health and to be certain to not have a spurious correlation (i.e. a significant correlation without any plausible explanation behind). We will summarize the findings from four studies (Borg & Kristensen, 2000; Hemström, 2005; Bauer et al., 2009 and Hämmig & Bauer, 2013) which emphasize the link between physical exposure and lack of job control with the poor health of people. Borg & Kristensen (2000), investigating the data from the "Danish National Work Environment Cohort Study, 1990-1995", found that two thirds of the change in SRHS could be explained by differences in five work factors and life style factors (obesity and smoking). Hemström (2005) calculated, with the "Swedish Survey of Living Condition, 1998-1999", that work environment counts for 25 % of health difference by income for men and 29 % for women. Bauer et al. (2009: 26) said that "carrying heavy loads enhanced the risk for less than good SRHS by 42 % compared to employees doing mostly sedentary work" (with the data provided by a representative sample of the "Swiss National Health Survey, 2002"). However, a finding that is somewhat surprising is the presence of a "psychological" reverse gradient, described as the fact that high-paid jobs have more psychological demands. Yet, in their papers, this "psychological gradient" does not result in a poorer health.

To summarize the literature, the health of people results from both daily habits and job characteristics. People at the lower socio-economic level are harmed by jobs that are more dangerous due to some characteristics like poor job-control and high physical demands. These characteristics affect people and statistical literature has shown their impact in various ways. However, despite the existing work, there is no clear criterion to quantify the difficulty of job leaving all other factors constant. Eurofound (2012) uses four dimensions which are earnings, prospects (referring to job trajectories and job security), intrinsic job quality (work quality and intensity plus physical and social environment) and working time quality. Nevertheless, this could only be seen as an indicator of quality and not as an indicator of health impact. Our aim in the following chapter will be to find out if it is possible to extract the health impact of each occupation.

Empirical analysis

"The goal [creating a healthy life expectancy index] is difficult of attainment,
but the stakes are enormous for our nation and for mankind".
Sanders (1964: 1069)

"Men differ, often violently, in their philosophy of life, in their views […] as to the means best adapted to those ends […]
It will be possible by the gradual extension of the field of exact knowledge ultimately to narrow the contentious field
and correspondingly to enlarge the basis upon which useful social action may rest."
Gay (1st President of the NBER, 1926: 5)

In this chapter, we will investigate empirically the relation between health and jobs. First, we will use the "European Working Conditions Survey, 2015" in order to investigate the connections between SRHS and occupations. SRHS is a very relevant indicator of health (see Section 3); but it has the drawback to be a subjective one (as the assessment is done by the respondent himself). The relation between this measure and the WHO definition is a little ambiguous as we do not know if respondents refer totally to it; nonetheless, this measure approaches it very well. This first section is important, as SRHS is widely used, one can be tempted to use it to identify hard work. Therefore, it is necessary to define a methodological framework to analyse it. In the second section, we will use the "US National Longitudinal Mortality Survey, 80s-90s" from the United States to calculate LE using KM technique. We have chosen to calculate LE as it is one "objective" measure of health and is based on a clear fact (death of people) which does not require medical attention to be known. This last section is our major contribution to the literature; it could be seen as a very detailed extension of the work of Delwarde (2004), Cambois et al. (2011) and Luy et al. (2015). However, it is worth noting that this measure does not follow the WHO definition as LE refers more to the absence of disease than to the presence of well-being (if we except the case of suicides). It is also for this deficiency that Sanders (1964) proposed to create a "Healthy Life Expectancy" indicator (done by Sullivan, 1971); but unfortunately, we were unable to do it due to a lack of data. After these sections, we will investigate if there are some relations between some jobs characteristics and our results using the information provided by the "O*Net Data"; which are, basically, very detailed job descriptions.

3 Self-rated health: "European Working Conditions Survey"

In this section, we will explore the association between SRHS and jobs. Although SRHS is a very common variable in the literature, it is always important to remember its relevance. In his paper, Bound (1991) explained that SRHS could be biased for the reason that all respondents do not use the same scale. Some people want to "legitimate" their conditions (e.g., not being in the labor force) or there is a "reward" for being in bad health (e.g., invalidity benefits). Although, these problems exist; "objective" measures are not perfect either. It is difficult to assess all the aspects of a person's health and moreover, it is not obvious how these different facets should be aggregated. The use of a simple question to a General Practitioner (GP) bears also the risk of being biased by the interaction between the GP and the person. Investigating statistical properties, Bound (1991) found that all measures have

some limitations. Idler & Benyamini (1997) reviewed 27 studies from all over the world and assessed that SRHS is a very good indicator. They pointed out several reasons why SRHS could be the best health indicator, notably that "it may be that ``the whole [health] is more than the sum of the parts´´ and while researchers are measuring the parts, respondents have access to the whole" (ibid.: 28). They concluded by stating that "its significant effect is impressive, regardless of whether it operates as an indicator of some unmeasured process or simply as a most effective summary of all the other measures" (ibid.: 31). To assess more this finding, Idler et al. (2000) examined the relevance of SRHS compared to a lot of clinical indicators and found that SRHS always brings a unique information. Recently, Schnittker & Bacak (2014) investigated the variation of powerfulness of this indicator over time. They discovered that SRHS became an increasing relevant indicator. According to their investigations, this is due to the fact that people are nowadays more educated and have better access to health and disease information. To summarize, although this measure is subjective, it is a very good one that can be used to assess the "true" health of people. We will now examine the relation between it and occupations as it is a good manner of entering our subject.

3. 1 Data

We use data from the 6th "European Working Conditions Survey" which interviewed workers from February to December 2015. This survey was done by Ipsos NV, on behalf of the European Foundation for the Improvement of Living and Working Conditions and released to us by the UK Data Archive. The last update of the data available at the time of the writing is the one from 4th October 2017. This wave interviews 43 850 workers, using computer-aided personal interviewing (CAPI) in 35 countries[4] on their working conditions and work-related health. After dropping observations with missing values, we have obtained a group of 32 778 subjects. This large figure is an advantage to this dataset; on the other side, the inconvenient is that the subjects are from different countries.

[4] The countries are: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, United Kingdom, Montenegro, Republic of Macedonia, Serbia, Turkey, Norway, Switzerland and Albania.

The dataset contains 49,65 % of male and 50,35 % of female observations. There are some countries with a bigger or smaller male/female ratio, e.g., 72 % of the observations in Turkey are male; but, generally, the proportion is close to half between the two genders (see Table 1). The mean age is 42,84 and 48,4 % of the observations are between 35 and 55; the age distribution can be observed in the Figure 2. The dataset contains 85,18 % of employee and 14,82 % of self-employee[5] with some differences between countries with, for example, 29,33 % of self-employees in Turkey and only 4,92 % in Denmark. The total percentage of self-employees is the same as the one in reality (14,25 % in 2015 for Europe (28 countries); Eurostat, 2018); however, the difference between countries in the sample does not reflect the real one. The net incomes by month[6] are reported in Table 1; obviously, there are large variations between countries[7].

We have two health variables; the first one asks if people think that their health is affected by their work[8] and the second one asks the interviewees about their SRHS. The first one was not, *a priori*, our interest; but we found interesting to keep it. Overall, 69,6 % of the participants declared that their health is not or is positively affected by their work. However, the answer runs from 51,9 % for the "building and related trade workers" to 83,4 % for the "general and keyboards clerks" (see Table 2). We can say that the fact that work impacts health is not felt in the same way between jobs. Our goal will be to find if this perception is true or false. We will nevertheless remain cautious that people saying that their health is affected by their job could express a worse one due to justification bias (Black et al., 2017). However, although this problem should not be overlooked; it should be known that the relation between objective measures (e.g., mortality) and SRHS does not vary between kinds of jobs (McFadden et al., 2009). We have computed in the 6th column the "percentage of people" that are in bad health in each profession. This is interesting to compare with the third column as they are both done by asking the respondents and so, are both subjective. We see that there is a link between the two answers but it is not very strong (coefficient of correlation: 36,91 %).

[5] Self-employee is defined by the survey as someone who has his own business or is partner in a business (e.g., producers' cooperatives), regardless of the number of employees.

[6] The net income is reported by the interviewees themselves. It is emphasized that this is the income coming from their main jobs.

[7] In order to compare earnings between countries, all answers were converted to euros (by the survey) using the following exchange rates: Bulgaria: 0,5113; Croatia: 0,1321; Czech Republic: 0,03657; Denmark: 0,134043; Hungary: 0,003318; Poland: 0,2507; Romania: 0,2258; Sweden: 0,1067 and United Kingdom: 1,3557.

[8] The question is straightforward: "Does your work affect your health?".

To come back to our main health variable, 26,27 % of the sample think that their health is very good; 51,57 % good; 19,33 % fair; 2,50 % bad and 0,33 % very bad. One can question if this general good SRHS is not related to an "healthy worker effect". In fact, one issue that arises when questioning working people is the absence of people out of work (due to bad health). The Table 3 illustrates it with data from the "Northern Ireland Health Survey, 2010-2011". We can observe that people having worked last week have been, for 72,20 %, in good health over the last 12 months while the economically inactive people have only been, for 42,84 %, in good health over the last 12 months. We can also see in Table 1 that the repartition of SRHS is not the same across countries. This could be due to true difference in health or to different cultural scale when assessing SRHS (Jürges, 2007); for this reason, we will add the country in the control variable during the regression.

The control variable related to studies is the ISCED 2011 code which is a standard tool of education level. Its classification is detailed in 9 levels[9] which we recoded into primary education, secondary education and post-secondary education. Overall, 5,14 % of the dataset only have primary education; 62,08 % have secondary education and 32,78 % have post-secondary education. Our main variable of interest is the occupation of people which is collected by the ISCO code. The ISCO (International Standard Classification of Occupations) code is a standardized code made by the International Labour Organization. The version used is the 2-digit one from 2008 for which our database contains 43 different unique values. The survey asks the main job of interviewees with as much information as possible to recode it into ISCO code. We have to assume that people rest (and have been) in the same occupation as the one they have provided. Although, this is an important assumption; it is a natural consequence of doing it due to a lack of temporal data (e.g., Gubéran & Usel, 2000; Case & Deaton, 2003; Delwarde, 2004 and Luy et al., 2015).

[9] These levels are: early childhood education, primary education, lower secondary education, upper secondary education, post-secondary non-tertiary education, short-cycle tertiary education, bachelor's or equivalent level, master's or equivalent level and doctoral or equivalent level.

3. 2 Logistic regression

Our aim is to estimate the effect of the different independent variables on SRHS, with a special attention for the job. Our regressand is an ordinal variable; in the sense that the health status could be ranked (Cameron & Trivedi, 2010: 525). We have collapsed the five categories (very good/good/fair/bad/very bad) into three (good/fair/bad). SRHS grouping is not an unusual practice in the literature; as for example, Borg & Kristensen (2000), Morefield et al. (2001), Hemström (2005) and Bauer et al. (2009) dichotomized it in only two categories in their respective studies. Manor et al. (2000) analysed the possible drawbacks of this practice and did not find any particular issues.

As the dependent variable is coded as consecutive integers, it should be tempting to analyse it with a linear regression model (OLS). However, this use would violate the assumptions of the latter framework and would lead to incorrect conclusions (Scott Long & Freese, 2014: 309). We will use an ordinal logistic regression model instead of a probabilistic one to be able to interpret the odds ratio. The model estimated by logistic regression will be the following:

$$Pr(SRHS = m \mid \boldsymbol{x}_i) = F(\tau_m - \boldsymbol{x\beta}) - F(\tau_{m-1} - \boldsymbol{x\beta}) \quad (1)$$

Where

$$\boldsymbol{x\beta} = \beta_1 Age + \beta_2 Age^2 + \beta_3 Self-employment + \beta_4 ISCO + \beta_5 ISCED + \beta_6 ln(Income) + \beta_7 Country + \beta_8 Sex$$

$$\tau_m \text{ are the different categories of self-rated health}$$

This regression states that SRHS is a function of the age and age-square of the people, the self-employment (baseline: employee), the job (baseline: "teaching professional"), the level of education (baseline: secondary education), the logarithm of income, the country (baseline: Austria) and the sex of the observations (baseline: man). Normally, the age will impact health in a negative way; because people tend to be in a worse health as the time elapses. We have also put the age at the quadratic form in the regression to account for non-linear effects. The "self-employment" variable is used due to empirical fact that, in Belgium, self-employed people are less in invalidity than employees (3,46 % in 2015 for self-employed against 7,26% for employees; Inami-Riziv, 2015). However, this fact can be due to legal difference in invalidity benefits. The ISCO code is our main variable of interest.

The ISCED code is used to control for childhood socio-economic conditions which are known to have an impact on health and health related behaviour at adult age (Gubéran & Usel, 2000). Furthermore, education indicator is often used as control as it is an objective and stable (not changing in time) regressor (Vandresse, 2015). Normally, SRHS will be better with a higher level of education. The income is also a control variable; for which, we choose the logarithmic form to account for non-linear relation. The gender and countries are our last control variables.

3. 3 Empirical results

The output of the ordered logit regression is reported in Table 4. We can observe that some jobs are not statistically significant and a large part of others seems to impact negatively the health compared to the baseline ("teaching professionals"). The insignificance of some jobs should not necessary be understood as a problem; it could only mean that they are not particularly more difficult that the reference one. The baseline level of education is the secondary education and a higher education impacts, as expected, positively the SRHS. Many countries are statistically significant which means that the SRHS is not related to the same scale across them and/or that people are *really* in better/worse health in this country. This could also be due to the power of countries variables to control for difference in income. We can also notice that the earnings affect positively the SRHS and being self-employed has a negative influence (or, possibly self-employed people continue to work despite their poor health more than employees). Before proceeding to more interpretations (odds ratio), we have to check the consistency of our model. We begin by testing the predictive power of our model (detailed in Table 5) and the difference is almost inexistent. However, we have to remain cautious that "many [model] can approximately mimic any given set of facts; that one [model] can does not mean that it is even close to right" (Summers, 1986: 24). Therefore, we verify the proportional odds assumption which if violated would lead to incorrect interpretations and biased estimates (Williams, 2006 & 2016). The proportional odds assumption states that the coefficients are the same across the logit equations; in fact, the ordered logit computes several logit equations (e.g., bad versus fair and good; bad and fair versus good) and the proportional odds assumption constraints these coefficients to be the same.

We begin our screening by using the command `oparallel` (Buis, 2013) in Stata which indicates that the proportional-odds assumption does not hold. Following this first test, we use the `brant` command (Scott Long & Freese, 2014) which allows to identify the departure for every variable. We see that only some variables, e.g., self-employment, have this problem. We should now consider other types of regression that can be used with these characteristics. We choose to compute a partial proportional odds model which allows the coefficients that "violate the proportional odds assumption to vary across logit equations" (Fullerton, 2009: 313). The principal difference between an ordered logit (i.e. with proportional odds assumption) and the partial proportional odds model is the coefficients allowance to vary across the logit equations. However, it is qualified as "partial" as only the coefficients that need to vary change and the others remain constraint.

The partial proportional odds model is, therefore, less restrictive than the ("overly restrictive") ordered logit; but more parsimonious than a multinomial one (Williams, 2006). We can rewrite (1) like:

$$Pr(SRHS = m \mid \boldsymbol{x}_i) = F(\tau_m - \boldsymbol{x_1\beta_{1m}} - \boldsymbol{x_2\beta_2}) - F(\tau_{m-1} - \boldsymbol{x_1\beta_{1m-1}} - \boldsymbol{x_2\beta_2})$$

Where $\boldsymbol{\beta_1}$ is a vector of coefficients allowed to vary across equations and $\boldsymbol{\beta_2}$ is a vector of fixed coefficients (Fullerton, 2009). We select the parameters allowed to vary by categories regarding three sources of information; the two first come from the subcommand `autofit` (of `gologit2`) at 5 % and at 1 % in Stata and the third comes from the `details` option of the Brant test (which, basically, gives the result of binary regressions). Self-employment is allowed to vary as well as "general and keyboard clerks", "production and specialised services managers", "metal, machinery and related trade workers", ln(Income), primary education and some countries. The final output can be seen in Table 4. We have also predicted the outcome from this model (see Table 5); the difference between the prediction and the reality is also almost insignificant.

We will now turn to the interpretation of our model and (slightly) compare it with the first (biased) one. The logit coefficients cannot be directly interpreted; but different ways exist to transform them into meaningful quantities, one of which is the odds ratio (Scott Long & Freese, 2014: 332). The odds ratio of our models can also be read in Table 4. First of all, the "constant" odds ratio does not vary much between the two models; the fact of being a woman lowers the chance of reporting a higher SRHS by 14,66 % compared to being a man. Self-employees have 25,30 % less

chance to report their health as fair or good and 9,42 % less chance to report it as good. The different jobs impact in various ways the SRHS; some jobs are insignificant and others affect more or less health. For example, being a "driver and mobile plant operator" decreases the odds by 32,17 % of reporting a higher SRHS where the decrease is only 19,87 % for a personal care worker. The effect of income is positive, but decreasing with the categories; an increase of a standard deviation increases the odds of reporting fair or good SRHS by 42,1 % and an increase of a standard deviation increases the odds of reporting good SRHS by 29 %. This seems to indicate that income has a bigger protective effect against bad health than for improving to good health.

To sum up this first section, we started by describing the answers to a question about health impact of jobs and see that the correlation between SRHS and these responses was not very high. Then, we investigated econometrically, the job impact on health which was done by using a partial proportional model. This model gave us some indications about which jobs have a greater impact on SRHS when controlling for income, education, self-employment, age, sex and country. We will now turn to the estimation of LE with the data of the "US National Longitudinal Mortality Survey, 80s-90s" from the United States.

4 Life expectancy: "US National Longitudinal Mortality Survey"

In this section, we will use a longitudinal database to calculate LE by using survival analysis and the KM technique. This introductory step will allow us to show differences in LE by occupations; after having done this first task, we will compare the SRHS by profession with the computed LE. We will then use a Cox model to demonstrate the impact of occupations on LE, leaving other factors constant. As stated before, LE by occupation is our main contribution to the literature.

4. 1 Data

We use the longest follow-up file (11 years) of the 5th version of the "US National Longitudinal Mortality Study" (NLMS – Public Use) which is a mortality study sponsored by the National Heart, Lung, and Blood Institute, the National Cancer Institute, the National Institute on Aging, the National Center for Health Statistics and the U.S. Census Bureau. This record consists of a subset of all cohorts in the full NLMS that were collected throughout 1980s and 1990s. The observations were made to reflect the U.S. non-institutionalized population on April 1, 1990. The dataset contains 1 835 072 observations aged from 0 to 90 years old.

The dataset contains very precise information about jobs; however, we have to regroup them in order to have more observations by "groups". We cluster the 364 different occupations into 91 "group of occupations" with regard to the classification of the "Bureau of Labour Statistics". The record contains a bit more men (with occupations) (54,15 %) than women (45,85 %); on the whole, 91,24 % of the sample survived until the end of the follow-up (160 750 people died). The repartition between the two sexes with regard to the different occupations could be find in Table 6. There are professions that are more important than others, e.g., there are 48 008 "construction trade workers" and only 2 027 "physical scientists". There are also 855 325 people without jobs in the dataset; and for those we know the reason, 98,25 % of them are not working because they are students, retired or unable to work. The age distribution (at the time of the sample) of the whole data can be seen in the Figure 3.

In addition to these variables, the NLMS provides a lot of other indicators like the state where the observation lives, his "race" or the place where he was born. We keep the same variables as in the preceding section (i.e. income, education level and self-employment). We have the income in 1990 dollars reported into several categories[10] which we transform into "continuous" variable by using the midpoint of each categories. The mean income is 35 349,71 $ for people aged 25 and more and the median is 32 499,50 $. As expected, the mean is very different between professions; as for example a "computer & mathematical occupation" has a mean of 54 318,89 $ where a "food and beverage serving worker" has only one of 24 952,68 $. As in the previous section, we regroup the education level into three categories: elementary school, high school and college. Half of the sample, aged 25 and more, are in the category "high school" and 38 % are in the college category. There is also, like for the previous variable, a lot of variations between professions. We also have the SRHS by work which we will compare with our estimation of LE. One main advantage of our database is the large number of observations and the length and precise follow-up (in days). One disadvantage is the absence of migrants (in or out the United States) in this record which can biased our estimation. Another drawback is the mix of different cohorts and so, we do not have the estimation for a particular cohort; but it is not a problem so far.

4. 2 Kaplan-Meier estimations

We now turn to the estimation of life expectancy of the different professions of our database using Kaplan-Meier (KM) estimator. Our data are longitudinal and we have a maximum follow-up of 11 years which means that we lack information, i.e. the death date of a lot of observations. The KM estimator was developed to cope with this problem because it is "impossible to make complete measurements on all members of a random sample" (Kaplan & Meier, 1958: 458). The principal characteristic of the KM technique is the possibility to use it with data affected by right/left censoring/truncation. The right censoring occurs when subjects do not experience the risk before the end of the studies which is often seen as they only have a limited follow-up[11]. This is the case in our

[10] These categories are 0$-4 999$; 5 000$-7 499$; 7 500$-9 999$; 10 000$-12 499$; 12 500$-14 999$; 15 000$-19 999$; 20 000$-24 999$; 25 000$-29 999$; 30 000$-34 999$; 35 000$-39 999$; 40 000$-49 999$; 50 000$-59 999$; 60 000$-74 999$ and more than 75 000$.

[11] For a description of left censoring and right/left truncation, see Kartsonaki (2016).

study where 91,24 % of the observations are still alive at the end of it. This technique is often used[12] and is described as "the most appropriate method to present and/or describe survival characteristics" (Etikan et al., 2017: 2). Another technique to estimate LE is by using life time tables (see Chiang, 1984) which gives similar results (Marubini & Valsecchi, 1994; Kartsonaki, 2016). However, this last technique is cross-sectional and need more data which is the reason why we prefer to use KM.

KM is a nonparametric estimator as it does not make any distributional assumptions and its formula is as follows:

$$\hat{S}(t) = \prod_{t_i < t} \frac{n_i - d_i}{n_i}$$

where n_i stands for the number of people at risk and d_i is the number of people who experience the risk (e.g., death). This is done until *t* which is the end of the follow-up and this is the reason why this estimator is often called the "product-limit" estimator. The mean lifetime, which in our work is the LE, is defined as "the mean of the product limit estimate of the distribution which is equal to the area under the corresponding survivorship function" (Kaplan & Meier, 1958: 467). It is so calculated with the following formula:

$$\widehat{LE} = \int_0^\infty \hat{S}(t)\, dt$$

One of the advantage of this technique is that we also use the information provided by the people that do not experience death; although, we do not know when they will experience it. An implicit assumption of the framework is that people who are censored (i.e. those still alive at the end of the follow-up) are assumed to have the same survival prospects than the no-censored one. This model is also better fitted when the follow-up is as much detailed as possible which is our case with a follow-up of eleven years in days. It is also important to notice that this LE is the one of people having lived during this particular time; it is not *per se* a LE at a particular date.

[12] A research on PubMed with "Kaplan Meier" as keywords give us 9 578 papers for 2017.

As in the previous section, we have to assume that people stay (and have been) in the same occupation all their life. We have to warn that there are some professions with less observations, e.g., "baggage porters, bellhops and concierges", for which the estimator could be biased due to the small number of them; however, we could be confident for a lot of professions due to the high number of observations and the length of follow-up (4018 days). We compute LE at 25 and at 50 (see Table 6) in order to know if some differences arise with time. We think, notably, that some self-selection could create biased estimates. For example, if people having done a penible job choose later to move into a simpler one; their presence/absence in the two estimates should impact the LE. However, the correlation between the two life expectancies is 96,34 % for men and 98,77 % for women. An analysis by scatter plot (see Figure 4 and 5) reveals no outlier for men and one for women (rail transportation workers); however, this outlier has very few observations and so, should be caused by a lack of data. We also have computed LE at 40 and 45 as a robustness exercise and no particular problem has arisen.

To turn to the analysis of our results, it should be, first, noticed that women live longer than men which is a comforting fact. The analysis of the results for men at 50 gives a range between the 10^{th} and the 90^{th} percentile of 5,85 years which is more or less the same as the one at 25. The range interquartile is 3,51 years at 25 and 3,18 years at 50. For the women, the scope is larger with a range of 11,23 years at 25 and of 11,88 years at 50 between the 10^{th} percentile and the 90^{th} percentile. It reduces to 7,37 at 25 and 7,13 at 50 for the interquartile range. We also see that, like in Delwarde (2004), teachers have a high LE. The Figure 6 gives an example of two survival functions for male engineers (LE of 55,99) and male metal and plastic workers (LE of 51,17). The relation between SRHS and life expectancy is not really strong and this is important to notice as SRHS could be a good predictor of how people feel and how healthy they are, it is nevertheless not a good predictor of LE. This is due to the fact that LE requires some computational techniques which cannot be simply replaced by averaging SRHS.

4. 3 Cox Model

Our first step was to compute LE in order to analyse it and to find if there are some differences between occupations. We notice, in the figures, that they exist and a log-rank test of equality of survival functions confirms that they differ, either we stratify by sex or not. A first conclusion could be that, e.g., "cooks and foods preparation workers" is a more penible job than "sales representatives" because they cannot expect to live longer. However as long as we have not controlled yet for the other characteristics; this conclusion would reduce to an *ignoratio elenchi*. As, e.g., in our example, "sales representatives" have a better income than "cooks and food preparations workers", it is "normal" that they have a longer LE. However, "the limitation of Kaplan-Meier estimate is that it cannot be used for multivariate analysis as it only studies the effect of one factor at the time" (Etikan et al., 2017: 2). We should therefore use another model in order to be able to disentangle the different influences of the variables. We will use a Cox model which was developed to extend the KM estimator and "more generally to the incorporation of regression-like arguments into life-table analysis" (Cox, 1972: 187). This model is frequently used and is "perhaps the most popular regression technique for survival analysis" (van Dijk et al., 2008: 705). It is based on the following equation:

$$h(t|\boldsymbol{x}_i) = h_0(t)exp(\boldsymbol{x}_i\boldsymbol{\beta}_x)$$

Before starting to explain our use of it, it is always useful to remember some of its aspects. This model is based on a "baseline hazard" ($h_0(t)$) which is assumed to be the same for everyone and depends on time only. In fact, this baseline does not need to be known as the hazard ratio calculation does not need it. There is also no distributional assumption about the baseline which means that the Cox model is non parametric. The multiplicative part of the equation ($\exp(\mathbf{x}_i\boldsymbol{\beta}_x)$) is used to introduce the covariates that "shift" the baseline. The multiplicative part is assumed to depend on the covariates, but to be constant over time. The advantage of this technique is that we do not make assumption about h_0 which could turn out to be wrong; however, we left efficiency as, if we have known its form, we could have produced better estimate (Cleves et al., 2004). We will use, first, the following model for our data:

$$h(t|\boldsymbol{x}_i) = h_0(t)exp(\beta_1 age_i + \beta_2 age_i^2 + \beta_3 sex_i + \beta_4 educ_i + \beta_5 ln(income)_i + \beta_6 occupation_i + \beta_7 self-emloyed_i) \quad (2)$$

The covariates are the same as in the previous empirical research and we also change our data to make our baseline correspond to someone with high school degree at age 25 working as a "secretary and administrative assistant".

We now turn to the result of our regression (see Table 7) in which we can see that being a woman reduces the baseline hazard by 49 % compared to a man. The elementary education level has a strange effect as it seems to slightly decrease the hazard. For our main variable of interest, we can see that a third of professions are significant (28 at 5 % and 7 at 10 %). Some professions decrease the hazard like being in a "computer and mathematical occupation" (0,71 the hazard of the secretaries) and some increase the hazard like people serving in restaurant (1,23 the hazard). The baseline survival function of our Cox model can be observed in the Figure 7. The fact of being self-employed decreases slightly the baseline (0,96 the "normal" baseline). As being self-employed has an effect, we were interest to test another specification instead of (2), we will continue with:

$$h(t|\boldsymbol{x}_i) = h_0(t)exp(\beta_1 age_i + \beta_2 age_i^2 + \beta_3 sex_i + \beta_4 educ_i + \beta_5 ln(income)_i + \beta_6 occupation_i * self - emloyed_i) \quad (3)$$

This new framing lets us test for a cross coefficient between occupation and self-employed in order to have the impact by occupation distinguishing by work status. The outcome of this framework can be seen in Table 7 too. The coefficients of the other variables do not vary much between the two models. The jobs that were significant in (2) are still significant in (3) either in employed, or in self-employed or in both. The interaction coefficient of self-employed and occupation tends to be lower (which means that they decrease the hazard rate) which is consistent with the estimate of β_7 in (2). We will now test a last model in which we incorporate the SRHS of people in interaction with occupation:

$$h(t|\boldsymbol{x}_i) = h_0(t)exp(\beta_1 age_i + \beta_2 age_i^2 + \beta_3 sex_i + \beta_4 educ_i + \beta_5 ln(income)_i + \beta_6 self - employed + \beta_7 occupation_i * SRHS_i) \quad (4)$$

The last term that we have added is to know the interaction of job and SRHS on the hazard rate. We have dropped the interaction between job and self-employment for computational problem. This could be a problem; however, as we have computed the cross-effect in (3), we have to keep in mind the different effects that we have discovered when we will analyse (4). The advantage of having the cross-effect of occupation and the current health is to take into account the current condition. In the specification (1), we consider all workers as being equal in term of SRHS; however, the baseline could

be shifted differently due to this interaction. The first thing to note is that the different SRHS impact, obviously, differently the baseline with a more detrimental effect for bad health (see Table 8). Because the size of the sample for each "classes" is reduced, our standard errors could be larger for some estimations. We can observe that the shift is not the same between the different occupations; for example, a "nursing, psychiatric and home health aid" in bad health has a coefficient of 3,05 instead of 1,23 (increase of 150 %) in (1) where a "building cleaning and pest control worker" has an increase of 100 % (2,23 instead of 1,14). This last specification can be seen as an illustration of how hard it is to define a penible job when we take more and more controls into account. Although it is always possible to detect jobs more difficult than others; it is really difficult to clearly find "the" coefficient.

The last thing that we should do is to test the proportional hazard assumption. As said before, a Cox model is based on a multiplicative part which does not depend on time and so, this part should be proportional to the baseline. Testing this hypothesis could be computed mathematically (with Stata) or observed in a graph. Normally, "if the proportionality assumption holds, a plot of Schoenfeld residuals against time should exhibit no time trend (a slope of zero)" (Powers & Xie, 2008: 213). The Figure 8 does not indicate a slope and so, the proportional assumption should hold. However, and for the sake of scientific integrity, we have to mention that computing the statistical test (on Stata) does not provide support for this conclusion. Nevertheless, and regarding the graphs plotted in Figure 8, we think that it is only due to the high number of observations as the graphs cannot indicate any slope. We are thus confident in our model and we now turn to a general analysis and discussion of our results from the two databases.

5 Discussion of our results

The main message that can be retrieved from our two studies is that some professions are more harmful than others. We test this assumption with SRHS and LE and the results indicate this tendency after having controlled for education and income which are two important causes of inequalities in health. The insignificance of some occupations could mean that some jobs have the same difficulty level. These findings are hardly surprising as the literature has already found these relations; but, to the best of our knowledge, this study was never done to such a deep level. On one side, we should be pleased of our results; on the other side, we should always remain careful because econometric techniques can also lead to problems when we forget the needed retrospect. This is the reason why we turn to the discussion of the limits of our study.

The first – evident – problem is the impossibility of having an objective health indicator and the validity of our studies can only be believed if one accepts the two selected subjective indicators. Another kind of problem is the assumption that people stay always in the same occupation which is, indeed, the current way of proceeding, but is clearly not the more realistic one. It is, in fact, almost impossible to know the kind of occupational mobility made by people. Kambourov & Manovskii (2008) found that occupational mobility is 15 % a year at a two-digit profession code level (consistent with Moscarini & Thomsson, 2007) in the US; however, the use of occupational code could be misleading as, for example, a change from a health service worker to a health technician is computed as an occupational mobility which is not really relevant in our case. One interesting result from their study is that government workers tend to have a relatively lower level of occupational mobility (6 %), which is better for our estimations.

There are two other particular issues that make "particularly difficult to measure cumulative effects of occupation on health" (Ravesteijn et al., 2013: 49): the first is that people select themselves into jobs and the second is that people can change their job. Following the first – evident – observation, one may argue that coefficients are not "pure" because, e.g., stronger people could choose heavier jobs. For example, Case & Deaton (2003: 16) argued that "selection […] is sufficiently severe to offset the deterioration associated with the work itself" when they noticed that "police and fire workers" have a

health status like non-manual worker. Although, this first example should not be overlooked; the whole argument has to be taken carefully as occupations are not unidimensional and the penibility is also created by the interactions between the worker and the work. To put it clearly, teaching is, clearly, less physically demanding than brickworking; however, this should not imply that turning a brickworker into a teacher would imply less penibility for him. His new placement could be felt very meaningless and, at the end, resulting into a burn-out. Another setting is the case of an employee that accepts a job that is not fitted for him (because, e.g., he needs money) which turns to be harmful for him, whatever the "real" level of penibility of it. To gather our argument, we do not state that the selection bias is neither uninteresting, nor important; but, that this argument has to be thought deeply before considering it. The second problem (job change) could be more problematic and some papers in the literature could not control for it due to cross-sectional data (e.g., Case & Deaton, 2003). A way of dealing with this problem could be found in Fletcher et al. (2011) who controlled for lag health. However, some occupational problems only arise late in life and so, this control does not allow neither to be sure that the observed effect is really the one of the current occupation. Econometrics, although an useful tool in some cases, has the weakness to depend on the availability of data that cannot be totally overcome for this field. However, whatever the technique used (cross-sectional, longitudinal, control for lag health, impact of the first occupation (Fletcher & Sindelar, 2009),…), the tendency remains the same which means that the results are, nonetheless, interesting for policy implications. To sum it up, although it is hard to isolate job effects only; the coefficients should not be seen as correlation only, but as causality too.

In our case, the first study has the drawback of being cross-sectional and so, we cannot control for change in jobs. However, as the categories are broad, a change is less of a problem as people could likelier stay in the same "broad" category. We have the same problem for the LE; but we have to mention that we do not observe in the data a particular different shape of age distribution between jobs. We were also not (totally) able to control for smoking and obesity which are known to be important factors of health status. If occupations are linked with these two behaviours, our coefficients will be biased because the job variable will act as proxy for these two effects; the same should be mentioned for the contrary effect of potential "health investment". However, Choo & Denny (2006) found that adding chronic conditions, obesity and smoking in their control factors downsizes by only 10 % their coefficients. "In sum, it is a very serious challenge to estimate cumulative effects of job conditions on health with non-experimental data" (Ravesteijn et al., 2013: 50) and we do not see how

experimental data could be achieved. As said in the introduction, in order to create those data, it would be needed to randomize all "kinds" of people between all occupations. It should, then, be waited long enough to see long-lasting effect with the hope that all other health behaviours will not change during this time. Although, these conditions are not theoretically impossible; they are barely plausible. The other solution of controlling for all factors (physical/mental strengths, preferences over job, health behaviour,…) is not more plausible neither (as well as using a two-step procedure like the Heckman model).

We also should remain cautious that our calculations could not be extended to other countries without verification. For example, high-rank Nigerian civil servants had the higher risk of cardio-vascular disease which was the contrary of the Whitehall study at the same time (Adler & Ostrove, 1999). This should warn us not to expand too quickly our findings. Indeed, the problem in economics is that we cannot find universal numerical constants (Klant, 1994: 83) and we have to remain constantly aware of it. To conclude, "pure" coefficients of job penibility are not achievable; however, and regarding the current political debate, the most an economist could do is to search for the best unbiased coefficients and, therefore, ours should be seen as the best achievable results when keeping in mind their potential biases. We will now explore the job contents and search for a possible link between work characteristics and our results.

6 Sources of the job penibility

In the two precedent sections, we have spoken about the job penibility without investigating about its possible roots. This should not be seen as something that does not attract our interest; but we wanted firstly to objectivate the penibility dimension before knowing its origins. Furthermore, the sources could be better explained by health researchers than by economists. However, we did not want to close our work without turning – a little bit – to its possible roots. To this aim, we use the O*Net data which are, basically, a very detailed job description with "hundreds of standardized and occupation-specific descriptors on almost 1 000 occupations covering the entire U.S. economy" (O*Net, 2018). The aim of this database is to assess the capacity needed to perform a particular job and, especially, the change in the abilities required to help the US government to maintain an "up-to-date" workforce. This database is sponsored by the US Department of Labor/Employment and Training Administration. The information are collected for a very detailed level of jobs and we use the same grouping as the one used in the "National Longitudinal Mortality Survey" section to be able to compare our findings. The last update at the time of the writing was the one of August, 2017.

We use the information provided by the "Abilities", "Work activities" and "Work context" files which give us 148 variables. These variables are particularities about the job contents and cover them in various ways. We list some examples of questions below to provide an illustration about it:

- What level of coaching and developing others is needed to perform your current job?
- What level of operating vehicles, mechanized devices, or equipment is needed to perform your current job?
- What level of handling and moving objects is needed to perform your current job?
- How often does your current job require you to work in an environment that is <u>not</u> environmentally controlled (like a warehouse <u>without</u> air conditioning)?
- How often does your current job require you to work in an environment that is environmentally controlled (like a warehouse <u>with</u> air conditioning)?
- …

The assessment is done by eight analysts in the case of the "Abilities" file and the mean standard error is 0,165 (maximum: 0,84) for a scale from 0 to 7. The assessment is done by incumbents (for 77 % of the professions) and by occupational experts (for 23 % of the professions) for the "Work activities" and "Work context" files. The mean standard error is 0,447 (max: 1,7; scale 0-7) for the "Work activities" file and 0,252 (max: 1,23; scale 1-5) for the "Work context" file. These incumbents and occupational experts are selected randomly and varies in number between occupations (from 9 to 90).

We choose to compute a principal component analysis to gather our 148 variables into a small number of uncorrelated principal components. The principle of a PCA is to "describe the variation of the variables in a set of multivariate data as *parsimoniously* as possible" (Rabe-Hesketh & Everitt, 2004: 260; italics added) by creating new components based on an uncorrelated linear combination of the original variables. We choose to base our PCA on the correlation matrix to compensate for the different scales of measurements between the different variables as argued in Afifi et al. (2011: 366). We keep the two first components which explain 41,21 % and 22 % of the total variance (the third one only 7,14 %).

The first component is related to the job physical difficulty as it is positively correlated with variables like "exposed to minor burns, cuts, bites, or stings", "spend time standing" or "performing general physical activities" and negatively correlated with "electronic mail", "spending time sitting" or "interacting with computers". The second component is subtler and is related to be in a "middle management" occupation as it is positively correlated with variables like "coaching and developing others", but also to variables like "outdoors under cover". The second component describes occupations involving management, but also presence "on the ground".

The Figure 9 shows all the occupations on the two dimensions. For readability, we create a second one (Figure 10) with less occupations. For example, we can see that "metal workers and plastic workers" is a very physical job while "mathematical science occupations" is not. On the other dimension, we see that "supervisors" score high which is the opposite of "communication equipment workers".

We have now a reading grid to analyse the roots of our coefficients. We have to mention that O*Net data are description of the current work characteristics which are, basically, not the same as the one in our NLMS file with cohort during the 80's and 90's. However, the bias created by using current job descriptions should be low as we do not see an occupation that totally changes over time. For example, "agricultural work" involves new characteristics; however, the core of the activity remains the same over time. We regress our coefficient of the Cox model (2) on the principal components (see Table 9) and find a significant positive coefficient for the first component (job physical difficulty) and a significant negative coefficient for the second component ("middle management"). Therefore, being in a physical job leads to deteriorations in health (after controlling for the different factors in (2)) as opposed to people in "middle management". This conclusion is not surprising as we have read in the medical literature that physical work deteriorates health; however, there was not a link between the "inverse" psychological gradient and a poorer health. We will now stop to investigate empirically the penibility problem and will turn now to discuss on a more philosophical and economic side whether governments should or should not compensate people for their occupations and, then, we will conclude.

Policy discussions & Conclusion

"Econometric models […] should lead all investigators to the same conclusions, independents of their personal whims. The usual experience in the field of economic policy is that there are about as many types of advice as there are advisors (sometimes even more!)"

Klein (1947: 111)

"The achievement of any State in the promotion and protection of health is of value to all"

World Health Organization (1946: 1)

7 Policy discussions

So far, we have reviewed the literature and computed some statistics, which can help politicians to achieve their aims. However, we should not stop our work here and only provide "rough statistics". The work of an economist is far more complex and complete and therefore we will now turn to discuss the reasons why public authorities should or should not recognize the job penibility. This section shall not be understated as economic ideas are very powerful and Keynes even said (1936: 383) that "the world is ruled by little else". This section will enquire about the pros and cons of compensation and then analyse how to indemnify people and who should pay for it. In our view, the previous sections are of little use without this one as there is, until now, no advice on how to use our figures.

Before entering this section, it has to be pointed out that it speaks *in abstracto* about compensations. We are aware that creating new compensations is not the same as withdrawing old ones. The difference is that someone may have chosen a job due to its indemnities and withdrawing them will not be the same as creating them. However, our discussion does not cope with this problem.

7. 1 Arguments for compensation

The first argument is the public support for the compensation of the job penibility. In Belgium, 83 % of the population agree with it (Ogeo Fund, 2016) and in the Netherlands, Vermeer et al. (2016) found that Dutch people are willing to contribute to an early pension for people having physical penible jobs[13]. However, this first argument is questionable for two reasons. The first is people's understanding of the survey questions is not complete and they may not consider all the ramifications of their answers (e.g., they "may not take account the economic constraints of their choices" (Vermeer et al., 2016: 161)). The second reason is the fact that we should compensate "because people care about, […] is only valid if we accept the normative position that society should care about what people care about" (Decancq et al., 2015: 70). A second argument is that everyone should have the possibility

[13] Contrary to the "simple" survey of Ogeo Fund, Vermeer et al. (2016) investigated if this willingness could only be due to self-interest. They found, e.g., that "almost half of the respondents whose job is not similar to that of a construction worker indicate that they are probably or certainly willing to contribute to a retirement scheme for construction workers" (ibid.: 167).

to live a healthy life like all others. To put it in accordance with Fishkin's pluralism of opportunity[14] (2014), we should develop healthy opportunities for every person stuck at a certain bottleneck in order for them to have the ability to enjoy their lives. Fishkin (2014: 195) himself said that there cannot be an agreement about the goals one should pursue in his life, but that "there is a broad agreement […] that […] health is important". It is, in our view, incomprehensible why people should have less-healthy lives and how, e.g., behind a "veil of ignorance" (to use now a Rawlsian framework) people would accept it. Furthermore, one's work is a function of his education attainment which depends a lot on his parents (for which nobody is responsible) (Keane & Roemer, 2009). Not compensating someone for not having achieved some high diploma related to a better job, because of his poor childhood, could be very unfair. Additionally, even if the lack of a high diploma is due to his own fault; how long should a person be punished for his mistake? A job is something you carry out your entire life and there is not a lot of errors that deserves such a long lasting punishment.

The three first arguments are philosophical and, sometimes, economists do not feel at ease when they are dealing with this kind of matter. However, a fourth one exists on which all economist will agree: it is the fact that the penibility of a job being not *fully* compensated is a market imperfection. In fact, in the beginning of economics, Adam Smith (1776: 111) said that: "The whole of the advantages and disadvantages of the different employments of labour and stock, must, in the same neighbourhood, be either perfectly equal, or continually tending to equality". However, this statement is only valid if we assume perfect mobility between jobs and the same bargaining power for both parties. Adam Smith pointed himself that "the ease or hardship, the cleanliness or dirtiness" (ibid.: 112) is one of the five principal circumstances that should be pecuniary compensated and also mentioned that his theory "would be the case in a society […] where every man was perfectly free to choose […] and to change [his occupation]" (ibid.: 111). Empirical studies have indeed found some wage compensation differentials for heavier jobs (Cousineau et al., 1992; see also Ehrenberg & Smith, 2006: 244-246 and Cahuc et al., 2014: 176-178). However, the real question does not only cope with their presences but also with their fullness. We can indeed think of possible wage compensations that indemnify to some extent, but not totally; due, e.g., to market power for the firm or because the worker

[14] Fishkin (2014) showed that it is not possible to equalize everyone's opportunity. His theory is that, in our current civilization, people should pass through "bottleneck" (e.g., University degree) to climb in the society. He advocated that we should help people to *pass through the bottleneck* and that we also should *create paths around the bottleneck* in order that people, stuck at each "level", can enjoy their lives.

does not know the long-lasting effect of his job conditions. In this case of *partial* wage compensation, governments are just creating, in a sense, a more perfect market (which is an aim on which all economists will agree)[15]. A last and more practical argument lies in the need of people's existence in penible job (e.g., policemen). In fact, we cannot eradicate all the sources of this kind of work (e.g., fire for firemen) but society should at least try to smoothen the penibility of these workers.

7. 2 Arguments against compensation

Although the five reasons listed above are sensible, there are also some arguments against compensation. The first argument is that we have to avoid moral hazard from employers (Boulet et al., 2015). In fact, the *real* problem lies in the possible absence of a *full* wage premium and if government creates some compensations; they will distort the incentives for employers to create a better work place (i.e. solving the problem at its roots). In reality, these compensations will act as a public subsidy for these firms and can – even – lead, in the long run, to less and less healthy working conditions. "In fact, the very existence of compensating differentials is central to the study of market incentives for employers to provide adequate levels of safety at the firm's level" (Cousineau et al., 1992: 166). Regarding this argument, the best way legislators could help workers is by forcing companies to pay them more for their job characteristics or to enhance the legal minimum work protection.

The second argument is that people know in advance their job characteristics. In fact, interviews of foundry workers lead to know that they are aware that few of them pass 70 (Buelen et al., 2014). There is a very particular issue now because people are, in some ways, taking risks consciously. However, people do not *necessarily* realize all the long-lasting effects of their job and – even – labour inspectors are not fully aware of them either (ibid.). Moreover, Akerlof & Dickens (1982) gave an example of cognitive dissonance[16] in which Ph.D.s, working in nuclear plant (!), did not wear their radiation exposure badges, which is an evidence of "forgetting" our own job danger. Another counterargument comes when the job hazard is only discovered further in time (e.g., asbestos) with the consequences that workers could not have known it when assessing their job penibility.

[15] It will take a whole thesis to find if the wage differentials are fully or not. Obviously, the question of unobservable characteristics, notably the worker preference over jobs, will pose a problem.

[16] The cognitive dissonance is the fact of manipulating, with long-lasting effect, one's belief about the state of the world. In our case, "people prefer to believe that their work is safe" (Akerlof & Dickens, 1982: 308).

7. 3 How to compensate?

Based on these arguments, we believe that – and *if there is not a full wage compensation* – workers in penible situations need to be compensated[17]. The next questions are: how (and to what extent) should they be compensated and who has to pay for it? As said in the introduction, we will leave these questions for further research; however, we give a (brief) first discussion on these matters. Nowadays, public authorities want to create some compensations for penible work and this is often related to an early retirement (due to a shorter life expectancy). A historical example is the one about the train drivers who could retire earlier due to inhalation of coal vapors (Zaidi & Whitehouse, 2009). From our point of view, this way of compensating is not optimal and could be, easily, related to the Sen's critics about the Rawl's fetishism when he spoke about the "primary goods" (Sen, 1979: 216). In our field, the fetishism is related to the compensation with an extra year of retirement although the person is ill. In a sense, his set of functionings is reduced, and this extra year of retirement could not really compensate him for his illness. The best thing that the government should do is to preserve the set of functionings by protecting people during *all their lifetime*. This policy could be achieved, e.g., by creating special work arrangements[18], possible working time reduction at the end of the career or retraining in order to be able to switch to a less demanding job at a certain age. As Papavlassopulos & Kepper (2011) explained, people value longer life with better health conditions; so there is no reason to expand the time in retirement for a sick person.

Moreover, on a more practical side (– not to speak on a more ethical side), there are two reasons to prefer protection to compensation. The first argument refers to the possibility to extend the career tenure because people determine it with regards to their jobs requirements. Indeed, Neumark & Song (2012) and Frins et al. (2016) have assessed that the "desired" duration of a career is related to job conditions. Therefore, in preventing harm to workers, governments will be able to better delay

[17] It could also be noticed that, in the absence of a full wage compensation, workers in penible job are twice punished: the first during their lifetime when they earn less than they merit and secondly during their retirement time when pension is based on lifetime income.

[18] For example, Belgian nursing workers can choose to have an extra day-off per month at 45, 2 days at 50 and 3 days at 55 (Arrêté royal du 1er Octobre 2008, 2008). In the Netherlands, there are "68 collective agreements that exempt older workers from overtime, irregular working hours, weekend work and team work" (Natali et al., 2016: 16).

retirement. The second reason is the problematic, almost, impossibility of withdrawing early retirement – even when it becomes irrelevant (e.g., the one from train workers). In fact, workers tend to perceive it as a "enshrined right" as the time elapses[19] (Vukorepa, 2017). It is, therefore, better to prevent harm to workers than to create some early retirements that will stay when their original motives have disappeared. We have to notice that it could be easier to remove them in a point system[20]. In this system, workers get "penibility points" for each year of penible job (Schokkaert et al., 2017). In the case of a decreasing penibility, stopping them is easier than stopping an earlier retirement as workers do not lose their past "bonuses" (Boulet et al., 2015). This system is also, in our point of view, a better way to deal with penible job if – nevertheless – governments want to use an early retirement. In fact, this system takes into account the demanding job tenure as you get points for every year in it which gives more equity to the pension system (Devolder & Hindriks, 2018). Moreover, if a worker in a heavy job wants to switch for an easier one; he will not lose his advantage because he keeps his points of penibility in his account. This particularity allows to lengthen the career tenure as workers in demanding jobs can "smooth" their working conditions by switching to a "softer" job which is not possible when the last job determines the retirement age, as pointed out by Schokkaert et al. (2017).

Despite this last remark, we maintain that we are not in favour of early retirement; furthermore, that invalidity benefits already exist in a lot of countries which are, *de facto*, an early retirement for demanding jobs[21]. Their advantages are that they are far more flexible and objective (and personal). To assess more deeply this argument, Pestieau & Racionero (2016) studied the implications of allowing people in a particular job to retire early against to have early retirement conditional on invalidity testing. They argued that the cost of the screening is an important factor to take into account; but they also pointed out that people in bad health in the "safe" occupation could find unfair the early retirement for the one in the "harsh" occupation – even more for those always in good health. At least, the early

[19] We should not totally blame workers for this perception. Nowadays, someone who faces his career choice can choose to be a railway worker due to the early retirement. No one warns him that it is due to the past inhalation of vapors and so, it is "normal" for him to not be in favour of the withdrawal of this advantage as it was an element of his past choice.

[20] A retirement point system is, basically, a framework in which workers get "points" for every working year. These "points per year" can be adjusted for multiple reasons: penibility, family arrangements and so on. At the end of the career, points are "transformed" into retirement benefit based on a point value and some other factors. See Hindriks et al. (2017) for an example with figures and Schokkaert et al. (2017) for a detailed and thorough explanation of the mechanism and its properties.

[21] One can also question if compensating for the job penibility should be the aim of the retirement system or if it is an aim *per se* which should have its proper "tools" (e.g., invalidity benefit).

retirement should always be possible for the people in bad health in the "safe" occupation with, possibly, an automatic early retirement for the "harsh" one. The second question is to what extend do we have to compensate for? In fact, if we indemnify for the job penibility, do we also have to compensate for day-to-day behaviour related to work (e.g., drinking after-work versus doing sports with colleagues)? In our view, we only have to compensate for the job because people should be kept responsible for their behaviour out-of-work.

7. 4 Should the company pay for the compensation?

Our last question refers to the indemnity funding, as whatever form it takes, someone has to pay for it. There are several options that can be followed and we will discuss one in particular which is the indemnity paid by the firm who is responsible for it. This has been shown to be a proper incentive to reduce invalidity in the Netherlands where companies have to pay regarding their history of employee invalidity (De Groot & Koning, 2016). We will argue the pros and cons of this solution by extending some arguments made by Blanchard & Tirole (2003) who detailed, on their side, how to make the companies aware of their unemployment responsibilities.

The argument in favour of imputing the cost to the companies is to make them aware of their duties for the health of their employees. It is the same logic as the "polluter-pays principle"; the firms who "destroy" the health must pay and the ones that preserve it should not. However, this solution is not so clear-cut and we also have to think of the possible repercussions of this policy. First, we can obviously fear that it could create barriers to hire "non-healthy" people due to the bigger likeliness that they will go into invalidity. There is also the possible drawback that an employer fires someone that becomes ill in order to avoid paying for him. We can think of a – even more problematic – behaviour by creating "sub-companies" to transfer demanding activities in them and closing regularly these "sub-companies" to avoid paying invalidity benefits. This is not so hypothetical as shown by Ringleb & Wiggins (1990) for environmental issues.

The possible consequences in the previous paragraph are only about possible abuses of the system, but we should also consider the possible drawbacks when it is play in a fair way. The first should be a greater unemployment as the whole labour cost will be higher. We should also mention that the price product will, normally, go up as the cost for the firm increases and, depending on the product elasticity (and the possible importation of the product), it can impact a second time the employment rate. Another possible problem is the liquidity constraint of the firm; but, it can be managed with delayed payment.

Regarding the previous arguments, it is difficult to be totally in favour of imposing the whole cost to the companies. In addition, we also have to add that all sectors are not the same and it is "normal" than a brick working firm has a higher invalidity rate than a secretary firm. Therefore, imposing the cost regardless of the sector could be misleading as it did not take into account the *true* effort made by the firm. It is for that reason we would prefer – if this solution is used – that the companies only have to pay for the difference between their invalidity rate and the one of their sector. We also think that it could be better to try to make the adherence of firms to the reduction of health impacts by not imposing cost on them. Other solutions are indeed feasible like making them pride of their achievements in reducing them. To this aim, some labels or symbols "healthy working environment" could, for example, be created. Another solution could be an active concertation between workers and firms to find solutions during all their career paths to preserve their health.

At the end, the dilemma lies between two views: the internalization of the cost or the mutualisation of it (Freyssinet, 2003) and we think that both should be used. The internalization is needed to make firms aware of their responsibilities but the mutualisation is needed to not put too much pressure on some heavy sectors – even more if they are exposed to world competition. The aim of this section was to review the reasons of a possible compensation, how this compensation should take place and who has to pay for it. We will now proceed to conclude.

8 Conclusion

Nowadays, the increasing life expectancy leads the public authorities to postpone the retirement age. However, some people argue that workers in "penible jobs" should be exempted of it. The problem lies in the lack of studies that search to objectivate the penibility of a job. To remedy this situation and to help the political representatives, we spent the main part of this master thesis in discussing its measurements. At the end, we discussed on a more philosophical and economic side the reasons to compensate people in demanding jobs.

We started our work by reviewing the literature that shows, too partially, that some occupations are more harmful than others. After this introductory step, we turned to our own estimates by first investigating the relations between work and self-rated health status with the data from the "European Working Conditions Survey, 2015" by using a partial proportional model. This first research gave us 15 significant job coefficients at 5 % (on 39 possible jobs). We found, e.g., that working in constructions diminishes by 40 % the odds of reporting a higher SRHS compared to working as a teacher, controlling for age, income, self-employment and education level. After this first empirical study, we deepened into the relations between life expectancy and professions using the data from the "US National Longitudinal Mortality Survey, 80s-90s". We computed LE for 91 professions by using the Kaplan-Meier technique and found differences between occupations. For example, a male secretary has a LE of 52,76 (at 25) while a man serving in a restaurant has only a LE of 48,73. After this first important step, we used a Cox model to disentangle the job impact from the other factors (age, income, self-employment and education level). A third of the professions were significant (28 at 5 % and 7 at 10 %) and, e.g., the risk of death for a "computer and mathematical occupation" is decreased by 29 % compared to a secretary.

These two studies clearly show that occupations are, indeed, very important for how long someone can expect to live. However, it has to be pointed out that our coefficients are not totally biases free. In fact, two particular problems occur when assessing the impact of profession on health

which are, firstly, that people self-select themselves into jobs and, secondly, that people can move from one job to another. Therefore, it is impossible to estimate *totally* unbiased coefficients without experimental data (and it cannot be plausible to create experimental data in this field). Regarding this problem, the best an economist could do, in order to help in the current debate, is to search for the best *possible* unbiased coefficients when keeping in mind the limit of his study. For this reason, we need to remain careful and humble about the interpretation of our results. Nevertheless, our coefficients should not only be seen as correlations; but also as causality as the literature shows that the same tendency remains whatever the technique used for correcting these biases. Then, we use the O*Net data to search for possible roots of the job impacts on health. This database contains, basically, very detailed job descriptions and we used a principal component analysis in order to create two summary indicators. The first one was based on the job physical difficulty and the second one was based on being in a middle management position. The regression of our Cox coefficient on them gave us a negative impact for the physical difficulty which is hardly surprising as the epidemiological literature has already provided some evidences about it.

After this empirical chapter, we discussed from a more philosophical point of view why governments should or should not compensate workers. We argued positively and negatively and, according to us, one important feature to know is the presence or absence of a *full wage differential* as a lot of *raison d'être* of the present debate depends on it. We also discussed the best possible compensation and explained that, from our point of view, the current political tendency of allowing early retirement is not optimal. We instead suggested that the policy aim should be the health preservation as long as possible which is better than allowing early retirement for people whose health is already "destroyed". Moreover, we pointed out that invalidity benefits already exist in a lot of countries which are, *de facto*, an early retirement; but this latter technique has the advantage of referring to the personal situation and is also far more flexible.

After having summed this paper up, we have to detail the next steps to be taken in future research (in addition to the other questions mentioned in the introduction). First, our aim was "only" to extract the job impact without searching in depth for real explanations behind our results. It should be of first importance, in the future, for governments to know the reason of a job difficulty in order

to take preventive measures. This work can only be done by collaboration between economists, demographists and health scientists; but because LE will continue to increase in the future, it is vital to preserve people's health as long as possible. It is an aim *per se*; but also an aim for the stability of our health and retirement system. Second, before creating some compensation, it should be deeply studied about the interactions between the markets and these possible laws. Normally, workers in hard jobs will progressively shy away from them and in the long run these occupations should disappear (Zaidi & Whitehouse, 2009: 9). If governments protect these companies, laws will act as counterforces against the market which could be more detrimental. This aspect should be deeply explored in order to avoid non-desirable second round effects.

We will conclude by pointing that, although, we have found some coefficients and procedures; we do not think that they should be the only indicator to be used. Like Sen (1979: 220) said, when he explained his new capability approach, we "did not argue that others were irrelevant". For example, in Austria, Hungary and Poland, KJoule consumption needed to perform a job is used to assess the penibility (Zaidi & Whitehouse, 2009). Although, there are some limitations to this measure, we think that it can be used in a scorecard for deciding which occupations are penible or not. On the other hand, retirement policy is in the hand of governments which based their actions on the people's demand. If they want that only some jobs are compensated (e.g., only physical jobs in Vermeer et al., 2016); this is also an important point to keep in mind. As Piketty (2014: 2) said: "Democracy will never be supplanted by a republic of experts – and that is a very good thing". In our opinion, our coefficients are, only, a first approach to objectivate the end of the "one-age-fits-all"; but we humbly acknowledge that this is *only* a first step. A lot remains to be done in order to better grasp this complex health problem and it is urgent to deal with it.

9 References

Accord de Gouvernement (2014), Gouvernement Fédéral – Michel 1er, Belgique

Adler, N. & J. Ostrove (1999), "Socioeconomic status and health: what we know and what we don't", *Annals of the New York Academy of Sciences*

Afifi, A.; S. May & A. C. Virginia (2011), *Practical Multivariate Analysis – 5th Edition*, CRC Press, Boca Raton (US), p. 357-377

Akerlof, G. & W. Dickens (1982), "The Economic Consequences of Cognitive Dissonance", *The American Economic Review*, 72 – 3, p. 307-319

Arrêté royal du 1er Octobre 2008 (2008), "Arrêté royal rendant obligatoire la convention collective de travail du 26 octobre 2005, conclue au sein de la Commission paritaire des services de santé, relative à la dispense de prestations de travail dans le cadre de la problématique de fin de carrière et l'octroi de congés supplémentaires au profit de certaines catégories de membres du personnel"

Bauer, G. F.; C. A. Huber, G. J. Jenny, F. Müller & O. Hämmig (2009), "Socioeconomic status, working conditions and self-rated health in Switzerland: explaining the gradient in men and women", *International Journal of Public Health*, 54, p. 23-30

Black, N.; D. W. Johnston & A. Suziedelyte (2017), "Justification bias in self-reported disability: New evidence from panel data", *Journal of Health Economics*, 54, p. 124-134

Blanchard, O. & J. Tirole (2003), *Protection de l'emploi et procédures de licenciement*, La documentation française, Paris (FR)

Borg, V. & T. S. Kristensen (2000), "Social class and self-rated health: can the gradient be explained by differences in life style or work environment?", *Social Science & Medicine*, 51, p. 1019-1030

Borghans, L. & B. Ter Weel (2012), "Zware dilemma's voor zware beroepen", *Demos*, 28 – 10, p. 5-8

Boulet, J.; B. Cantillon; P. Devolder; J. Hindriks; R. Janvier; F. Masai; G. Perl; E. Schokkaert; Y. Stevens & F. Vandenbroucke (2015), "Métiers pénibles, pension à temps partiel – Avis complémentaire de la Commission de réforme des pensions 2020-2040 "

Bound, J. (1991), "Self-reported versus objective measures of health in retirement models", *The journal of human resources*, 26 – 1, p. 106-138

Buelen, W.; A. I. Martinez Garcia, P. Alfonso, D. Theodorakis, C. Menne, E. Nygren, E. Brentnall, E. Somaglia, R. Pond, S. Jänicke, U. Spannow, E. Voss, F. Warneck, K. Szczpanik, M. Sptari, J. Serrano, J. Jensen, T. B. Hansen & M. Giaccone (2014); "Better Understanding of ``Arduous Occupations´´ within the European Pension Debate"

Buis, M. L. (2013), oparallel,: Stata module providing post-estimation command for testing the parallel regression assumption, Boston College Department of Economics, Statistical Software Components, S457720

Cahuc, P.; S. Carcillo & A. Zylberberg (2014), *Labor economics – 2nd edition*, THE MIT Press, Cambridge (UK), p. 151-188

Callahan, D. (1973), "The WHO Definition of `Health´", *The Hasting Center Studies*, 1 – 3, p. 77-87

Cambois, E.; C. Laborde, I. Romieu & J.-M. Robine (2011), "Occupational inequalities in health expectancies in France in the early 2000s: Unequal chances of reaching and living retirement in good health", *Demographic Research*, 25 – 12, p. 408-436

Cameron, A. C. & P. K. Trivedi (2010), *Microeconometrics using Stata - Revised Editon*, Stata press, Texas, p. 525-529

Case, A.C. & A. Deaton (2003), "Broken down by work and sex: how our health declines", *NBER Working paper*, 9821 ©

Charter of Fundamental Rights of the European Union (2000), *Official Journal of the European Communities*

Chiang, C. L. (1984), *The life table and its applications*, Robert Krieger Publishing Company, Malbar (US), p. 153-164

Choo, E. & M. Denny (2006), "Wearing out – The decline of health"

Cleves, M. A.; W. W. Gould & R.G. Gutierrez (2004), *An introduction to survival analysis using Stata- Revised edition*, Stata Press, Texas (US), p. 121-195

Cousineau, J.-M.; R. Lacroix & A.-M. Girard (1992), "Occupational Hazard and Wage Compensating Differentials", *The Review of Economics and Statistics*, 74 – 1, p. 166-169

Cox, D. R. (1972), "Regression Models and Life-Tables", *Journal of The Royal Statistical Society, Series B (Methodological)*, 34 – 2, p. 187-220

De Groot, N. & P. Koning (2016), "Assessing the Effects of Disability Insurance Experience Rating: The Case of the Netherlands", *IZA Discussion Paper*, 9742

Decancq, K.; M. Fleurbaey & E. Schokkaert (2015), *Handbook of Income Distribution*, 2A, Elsevier, Amsterdam (NL), p. 67-140

Delwarde, A. (2004), "Analyse de la mortalité des pensionnés du secteur public"

Devolder, P. & J. Hindriks (2018), "La pension à points: 5 principes pour plus d'équité dans les régimes de pension en Belgique", *IRES Regards Economiques,* 139

Ehrenberg, R. G. & R. S. Smith (2006), *Modern labor economics: theory and public policy – 9th edition*, Pearson Education, Boston (US), p. 239-274

Etikan, I.; s. Abubakar & R. Alkassim (2017), "The Kaplan-Meier Estimate in Survival Analysis", *Biometrics & Biostatistics International Journal*, 5 – 2: 00128, p. 1-5

Eurofound (2012), "Trends in job quality in Europe", *Publications Office of the European Union*

Eurostat (2018), *Employment by sex, age, professional status and full-time/pat-time (1 000)* [on line], available at: ec.europa.eu/eurostat/en/web/products-datasets/-/LFSA_EFTPT (consulted the 30th January, 2018)

Fishkin, J. (2014), *Bottlenecks, A new theory of equal opportunity*, Oxford University Press, Oxford (UK)

Fletcher, J. M. & J. L. Sindelar (2009), "Estimating causal effects of early occupational choice on later health: evidence using the PSID", *NBER Working Paper*, 15256 ©

_____; _____ & S. Yamaguchi (2011), "Cumulative effects of job characteristics on health", *Health economics*, 20, p. 553-570

Freyssinet, J. (2003), Comment in: Blanchard, O. & J. Tirole (2003), *Protection de l'emploi et procédures de licenciement*, La documentation française, Paris (FR), p. 51-55

Frins, W.; J. van Ruysseveldt, K. van Damullerton & S. N. J. van den Bossche (2016), "Older Employees' desired retirement age: a JD-R perspective", *Journal of Managerial Psychology*, 31 – 1, p. 34-49

Fullerton, A. S. (2009), "A conceptual framework for ordered logistic regression models", *Sociological Methods & Research*, 38 – 2, p. 306-347

Gass, R. & M. Bopp (1997), *Berufsspezifische Mortalitätsrisiken der Männer in der Schweiz 1979/83*, Bundesamt für Statistik, Bern (CH)

Gay, E. F. (1926), Foreword in: Thorp, W. L. (1926), *Business annals,* NBER, New York (US), p. 5-8

Geoffroy-Perez, B. (2006), "Analyse de la mortalité et des causes de décès par secteur d'activité de 1968 à 1999 à partir de l'Echantillon démographique permanent"

Gubéran, E. & M. Usel (2000), "Mortalité prématurée et invalidité selon la profession et la classe sociale à Genève"

Hafen, M. (2016), "Of what use (or harm) is a positive health definition?", *Journal Public Health*, 24, p. 437-441

Hämmig, O. & G. F. Bauer (2013), "The social gradient in work and health: a cross-sectional study exploring the relationship between working conditions and health inequalities", *BMC Public Health*, 13

Hemström, Ö. (2005), "Health inequalities by wage income in Sweden: The role of work environment", *Social Science & Medicine*, 61, p. 637-647

Hindriks, J.; P. Devolder, E. Schokkkaert, F. Vandenbroucke (2017), "Réforme des pensions légales: le système de pension à points", *IRES Regards économiques*, 130

Huber, M.; J. A. Knotterus, L. Green, H. Van Der Horst, A. R. Jadad, D. Kromhout, B. Leonard, K. Lorig, M.I. Loureiro, J.W.M. Van der Meer, P. Schnable, R. Smith, C. Van Weel & H. Smid (2011), "How should we define health?", *BMJ*

Idler, E. L. & Y. Benyamini (1997), "Self-rated health and mortality: a review of twenty-seven community studies", *Journal of Health and Social Behavior*, 38 – 1, p. 21-37

_____; L. B. Russell & D. Davis (2000), "Survival, Functional Limitations, and Self-rated Health in the NHANES I Epidemiologic Follow-up Study, 1992", *American Journal of Epidemiology*, 152 – 9, p. 874-883

Inami-Riziv (2015), "Facteurs explicatifs de l'augmentation du nombre d'invalides – Régime général et régime des indépendants – Période 2006-2015"

Jager, K. J.; P. C. van Dijk, C. Zoccali & F. W. Dekker (2008), "The analysis of survival data: the Kaplan-Meier method", *Kidney International*, 74, p. 560-565

Jürges, H. (2007), "True health vs response styles: exploring cross-country differences in self-reported health", *Health economics*, 16, p. 163-178

Jylhä, M. (2009), "What is self-rated health and why does it predict mortality? Towards a unified conceptual model", *Social Science & Medicine*, 69, p. 307-316

Kambourov, G. & I. Manovskii (2008), "Rising occupational and industry mobility in the United States: 1968-97", *International Economic Review*, 49 – 1, p. 41-79

Kaplan, M. & P. Meier (1958), "Nonparametric estimation from incomplete observations", *Journal of the American Statistical Association*, 53 – 282, p. 457-481

Kartsonaki, C. (2016), "Survival analysis", *Diagnostic Histopathology,* 22 – 7, p. 263-270

Keane, M. P. & J. E. Roemer (2009), "Assessing policies to equalize opportunity using an equilibrium model of educational and occupational choices", *Journal of Public Economics*, 93 (7-8), p. 879–898

Keynes, J. M. (1936), *The general theory of employment, interest and money* (Printed version of 1949), MacMilland, London (UK), p. 372-385

Klant, J. J. (1994), *The nature of economic thought*, Edward Elgar Publishing Limited, Cornwall (UK), p. 83-84

Klein, L. R. (1947), "The use of Econometric Models as a Guide to Economic Policy", *Econometrica*, 15 – 2, p. 111-151

Lanzerath, D. (2012), "Health and disease as normative concepts", *Bioethica Forum*, 5 – 2, p. 59-66

Law, I. & H. Widdows (2007), "Conceptualising health: insights from the capability approach"

Le Grand, J. & J. M. Muurinen (1985), "The economic analysis of inequalities in health", *Social Science and Medicine*, 20, p. 1029-1035

Luy, M.; C. Wegner-Siegmundt, A. Wiedemann & J. Spijker (2015), "Life expectancy by education, income and occupation in Germany: Estimations using the Longitudinal Survival Method", *Comparative Population Studies*, 40 – 4, p. 400-436

Manor, O.; S. Matthews & C. Power (2000), "Dichotomous or categorical response? Analysing self-rated health and lifetime social class", *International Journal of Epidemiology*, 29, p. 149-157

Marubini, E. & M. G. Valsecchi (1994), *Analysing Survival Data from Clinical Trials and Observational Studies*, John Wiley & Sons, Chichester (UK), p. 54-56

McFadden, E.; R. Luben; S. Bingham; N. Wareham; A.-L. Kinmonth & K.-T. Khaw (2009), "Does the association between self-rated health and mortality vary by social class?", *Social Science & Medicine*, 68, p. 275-280

Morefield, G. B.; D. C. Ribar & C. J. Ruhm (2011), "Occupational Status and health transitions", *NBER Working Paper Series*, 16794 ©

Moscarini, G. & K. Thomsson (2007), "Occupational and job mobility in the US", *Scandinavian Journal of Economics*, 109 – 4, p. 807-836

Natali, D.; S. Spasova & B. Vanhercke (2016), "Retirement regimes for workers in arduous or hazardous jobs in Europe – A study of national policies"

Neumark, D. & J. Song (2012), "Barriers to Later Retirement: Increases in the Full Retirement Age, Age Discrimination and the Physical Challenges of Work", *Michigan Retirement Research Center Working Paper*, 265

O*Net (2018), *About O*Net* [on line], available at: onetcenter.org/overview.html (consulted the 17th March, 2018)

OECD (2018), *Life expectancy at birth* [on line], available at: data.oecd.org/healthstat/life-expectancy-at-birth.htmonetcenter.org/overview.html (consulted the 7th May, 2018)

Ogeo Fund (2016), "Les Belges et leur pension – Enquête 2016 commanditée par Ogeo Fund et menée par Ipsos"

Papavlassopulos, N. & D. Keppler (2011), "Life expectancy as an objective factor of a subjective well-being", *Social Indicator Research*, 104, p. 475-505

Pestieau, P. & M. Racionero (2016), "Harsh occupations, health status and social security", *Journal of Economic*, 117, p. 239-257

Piketty, T. (2014), *Capital in the Twenty-First Century*, The Belknap press of Harvard university press, Cambridge (UK), p. 1-35

Powers, D. A. & Y. Xie (2008), *Statistical Methods for Categorical Data Analysis - 2nd edition*, Emerald Editions, Bingley (UK), p. 167-242

Rabe-Hesketh; S. & B. Everitt (2004), *A Handbook of Statistical Analyses using Stata – 3rd edition*, CRC Press, Boca Raton (US), p. 258-268

Ravesteijn, B.; H. van Kippersluis & van Doorslaer, E. (2013), "Long and healthy careers?", *Netspar Panel Papers*, 36

Ringleb, A. H. & S. N. Wiggins (1990), "Liability and Large-Scale, Long-Term Hazards", *Journal of Political Economy*, 98 – 3, p. 574-595

Sanders, B. S (1964), "Measuring community health levels", *American journal of public health and the nation's health*, 54 – 7, p. 1063-1070

Schnittker, J. & V. Bacak (2014), "The increasing predictive validity of Self-Rated Health", *PLoS ONE9*, 1

Scott Long, J. & J. Freese (2014), *Regression Models for Categorical Dependant Variables Using Stata - 3rd edition*, Stata Press, Texas (US), p. 309-381

Schokkaert, E.; P. Devolder, J. Hindriks & F. Vandenbroucke (2017), "Towards an equitable and sustainable points system, A proposal for pension reform in Belgium, *KU Leuven Department of economics Discussion Paper Series*, 573241

Sen, A. (1979), "Equality of What?", *The Tanner Lecture on Human Values*, p. 198-220

Smith, A. (1776), *An Inquiry into the Nature and Causes of the Wealth of Nations* (Printed Version of 1950), Edwin Cannan, p. 111-161

Sullivan, D. F. (1971), "A Single Index of Mortality and Morbidity", *HSMHA Health Reports*, 86 – 4, p. 347-354

Summers, L. H. (1986), "Some Sceptical Observation on Real Business Cycle Theory", *Federal Reserve Bank of Minneapolis Quarterly Review*, 10, p. 23-37

Van dijk, C. P.; K. J. Jager, A. H. Zwinderman, C. Zoccali & F. W. Dekker (2008), "The analysis of survival data in nephrology: basic concepts and methods of Cox regression", *Kidney International,* 74, p. 705-709

Vandresse, M. (2015), "Les espérances de vie/de santé selon le statut socioéconomique, Revue de la littérature"

Vermeer, N.; M. Mastrogiacomo & A. Van Soest (2016), "Demanding Occupations and the Retirement Age", *Labour Economics*, 43, p. 159-170

Vukorepa, I. (2017), "Retirement policy changes for workers in arduous and hazardous jobs: comparative overview and lessons for Croatia"

WHO (2017), *Social determinants of health* [on line], available at: who.int/social_determinants/thecommission/finalreport/key_concepts/en/ (consulted the 24th December, 2017)

Williams, R. (2006), "Generalized ordered logit/partial proportional odds models for ordinal dependent variables", *The Stata Journal*, 6 – 1, p.58-82

_____ (2016), "Understanding and interpreting generalized ordered logit models,", *The Journal of Mathematical Sociology*, 40 – 1, p. 7-20

World Health Organization (1946), "Constitution of the World Health Organization"

Zaidi, A. & E. Whitehouse (2009), "Should Pension Systems Recognise ``Hazardous and Arduous Work´´?", *OECD Social, Employment and Migration Working Papers*, 91

Data:

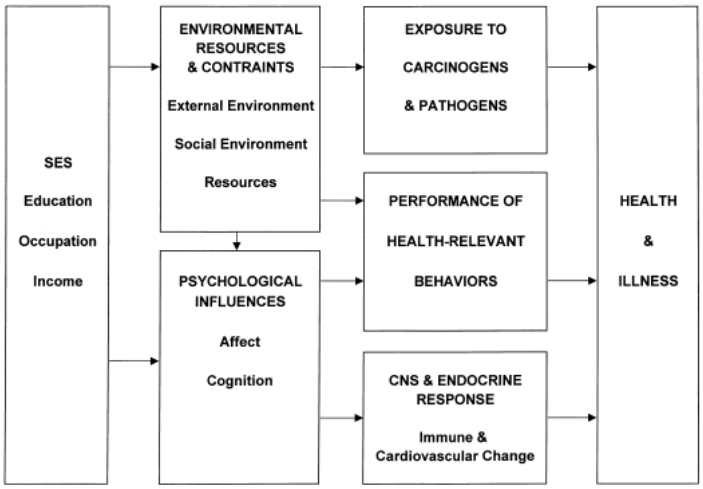
Department of Health, Northern Ireland. (2016). Northern Ireland Health Survey, 2010-2011. [data collection]. 2nd Edition. UK Data Service. SN: 7258, doi.org/10.5255/UKDA-SN-7258-2

European Foundation for the Improvement of Living and Working Conditions, European Working Conditions Survey, 2015 [computer file].4th Edition. Colchester, Essex: UK Data Archive [distributor], June 2017. SN: 8098, dx.doi.org/10.5255/UKDA-SN-8098-4

National Longitudinal Mortality Study (NLMS) Public-use file, United States Department of Commerce, Economics and Statistics Administration, U.S. Census Bureau, October 2015

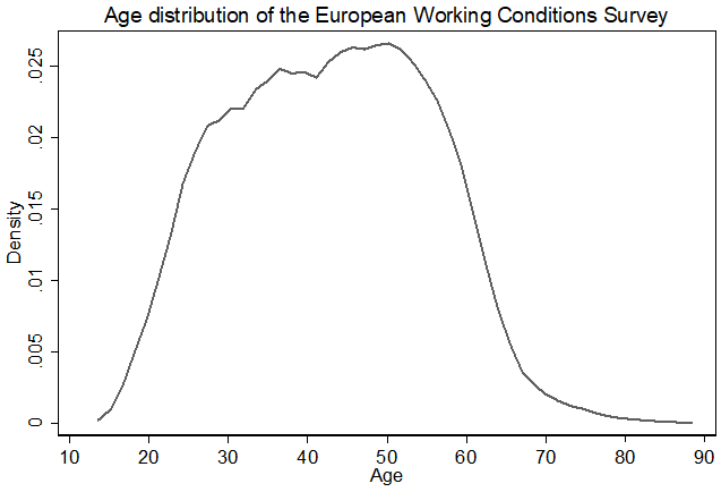
10 Annexes: List of figures and List of tables

Figure 1: Mode of the pathways by which socio-economic status (SES) influences health



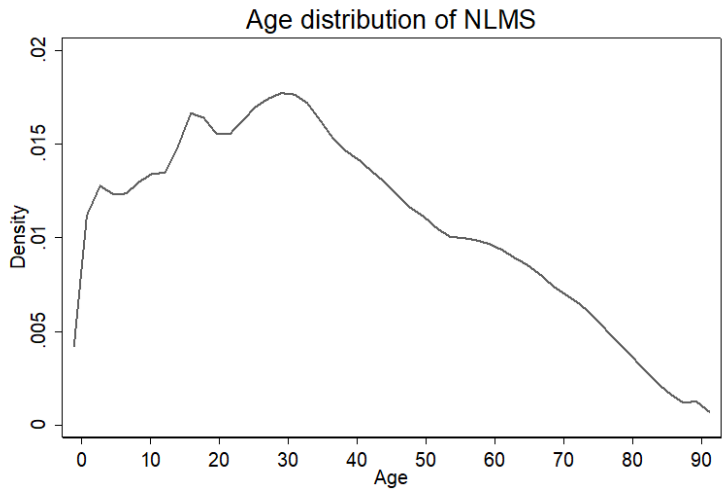
Notice: the Figure explains the different pathways by which socio-economic status affect health. For example, our socio-economic status (SES) influences the environmental resources and constraints, which have an effect our exposure to carcinogens and pathogens. This is a simplified model provided for illustration since the model should normally be completed by feedback loops and interaction effects (Adler & Ostrove, 1999). The Figure comes from Adler & Ostrove (1999: 12).

Figure 2: Age distribution of the "European Working Conditions Survey, 2015"



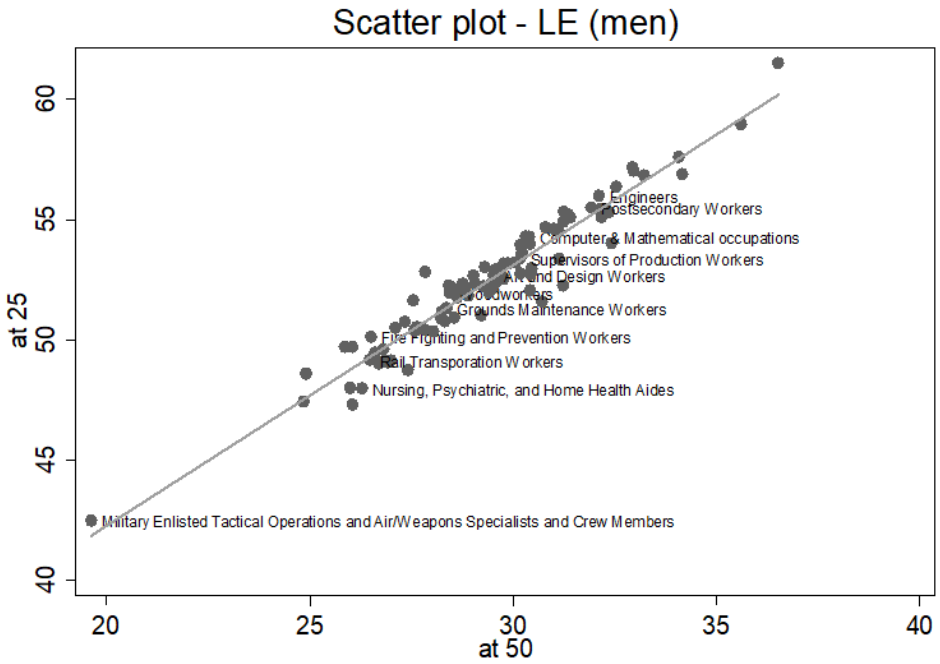
Notice: the Figure presents the age distribution of the "European Working Conditions Survey, 2015". The mean age is 42,84 and 48,4 % of the observations are between 35 and 55.

Figure 3: Age distribution of the "US National Longitudinal Mortality Survey, 80s-90s"



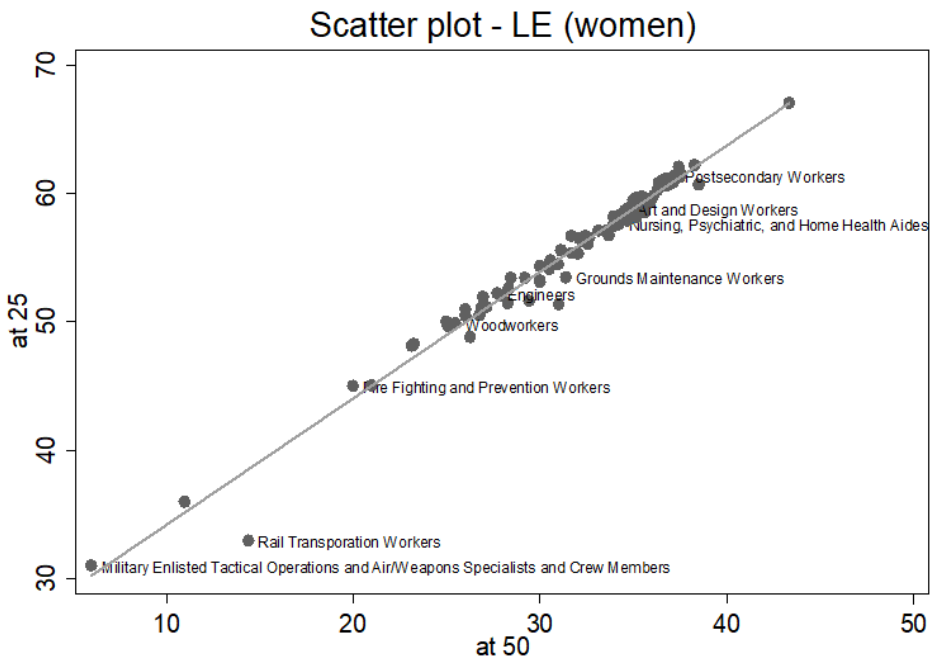
Notice: the Figure presents the age distribution of the "US National Longitudinal Mortality Survey, 80s-90s". The mean age is 35,19 and the median age is 33.

Figure 4: Scatter plot – LE for men at 25 and at 50



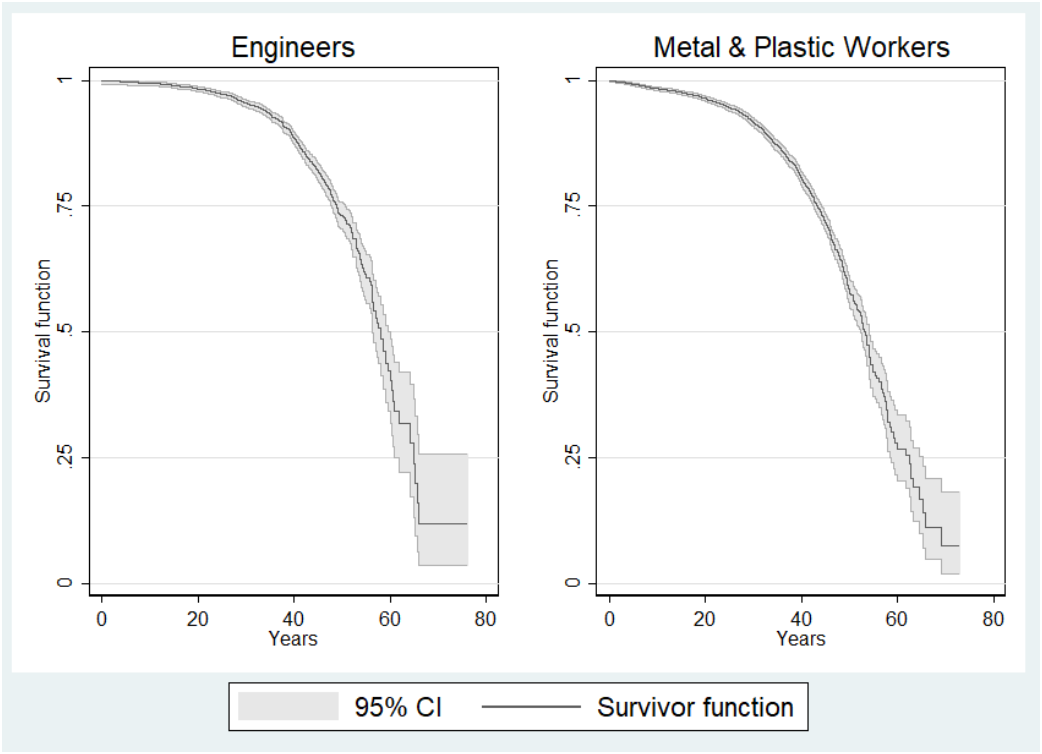
Notice: the Figure presents our estimate of life expectancy for men at 25 and at 50. These estimates were obtained by using Kaplan-Meier technique on the "US National Longitudinal Mortality Survey, 80s-90s". The correlation between the estimate at 25 and at 50 is very strong (coefficient of correlation: 96,34 %).

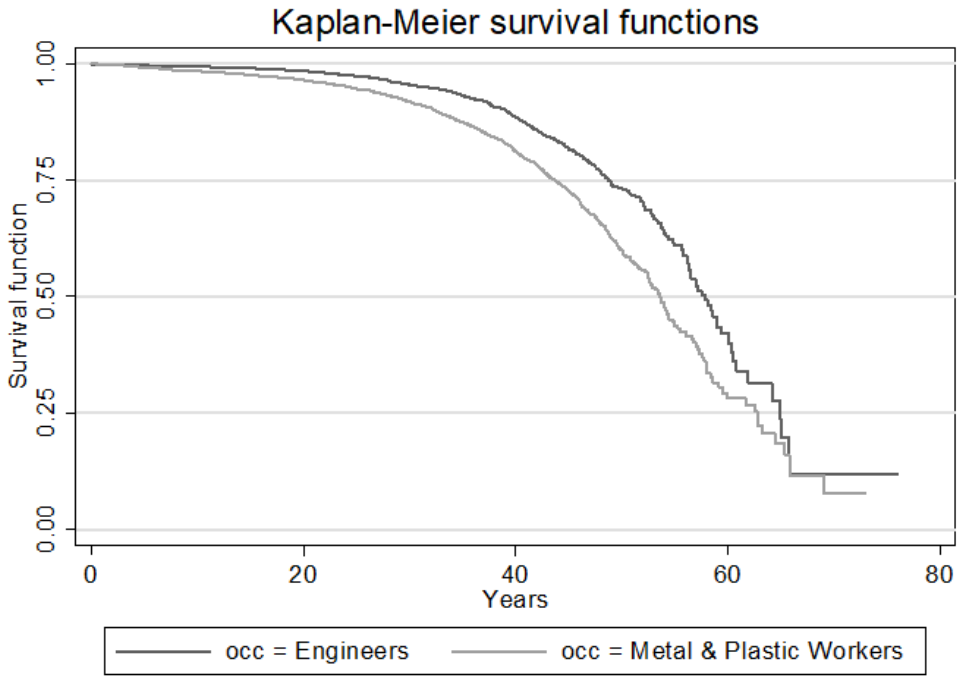
Figure 5: Scatter plot – LE for women at 25 and at 50



Notice: the Figure presents our estimate of life expectancy for women at 25 and at 50. These estimates were obtained by using Kaplan-Meier technique on the "US National Longitudinal Mortality Survey, 80s-90s". The correlation between the estimates at 25 and at 50 is very strong (coefficient of correlation: 98,77 %). The only outlier (Rail Transportation Workers) should be due to a lack of data as there are very few observations for this profession.

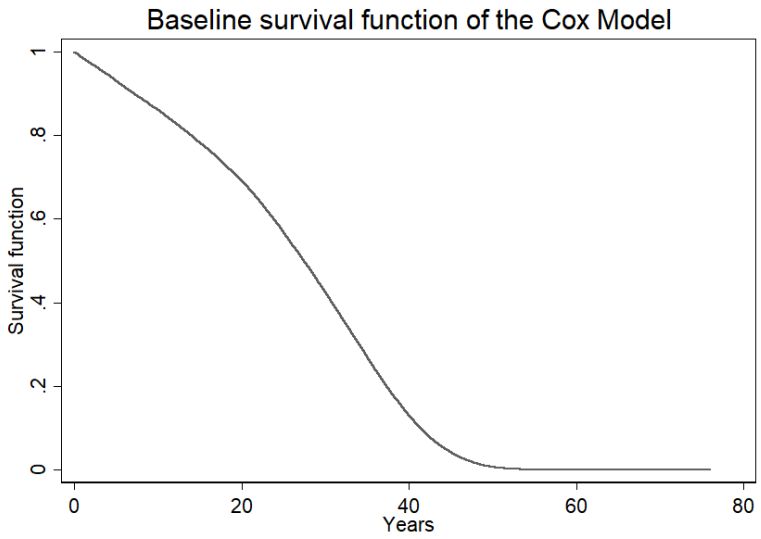
Figure 6: Survival functions for Engineers and Metal & Plastic Workers (men)





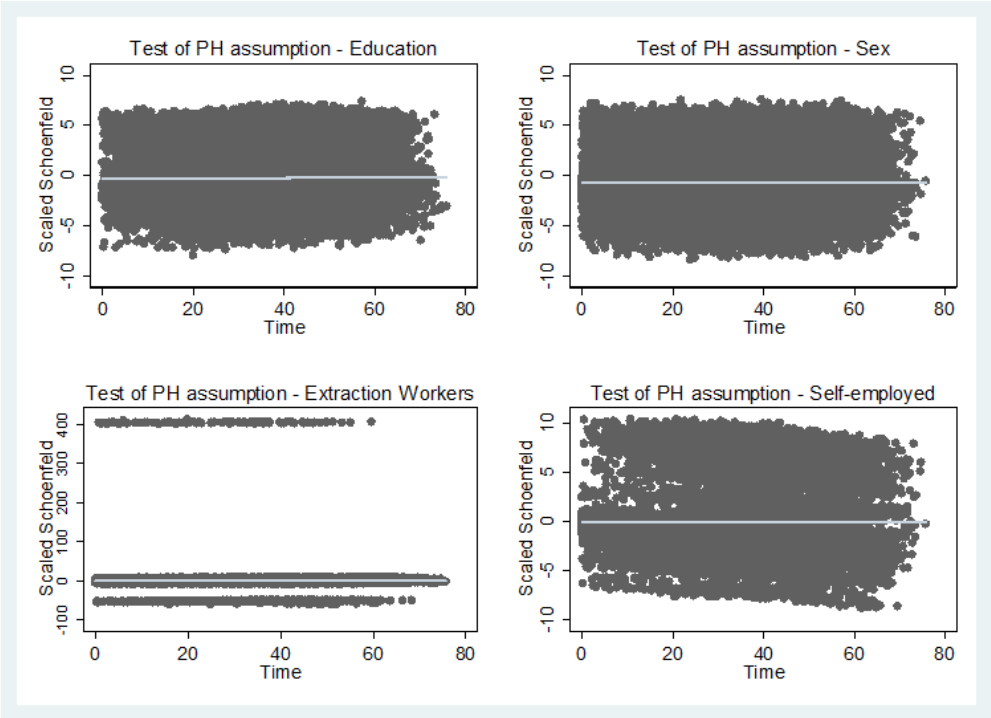
Notice: the Figure shows the survival functions for Engineers and Metal & Plastic Workers (male) from the "US National Longitudinal Mortality Study, 80s-90s". We can see that they differ which is reflected in the life expectancy estimates of 55,99 for Engineers and of 51,17 for Metal & Plastic Workers (at 25).

Figure 7: Baseline survival function of the Cox model



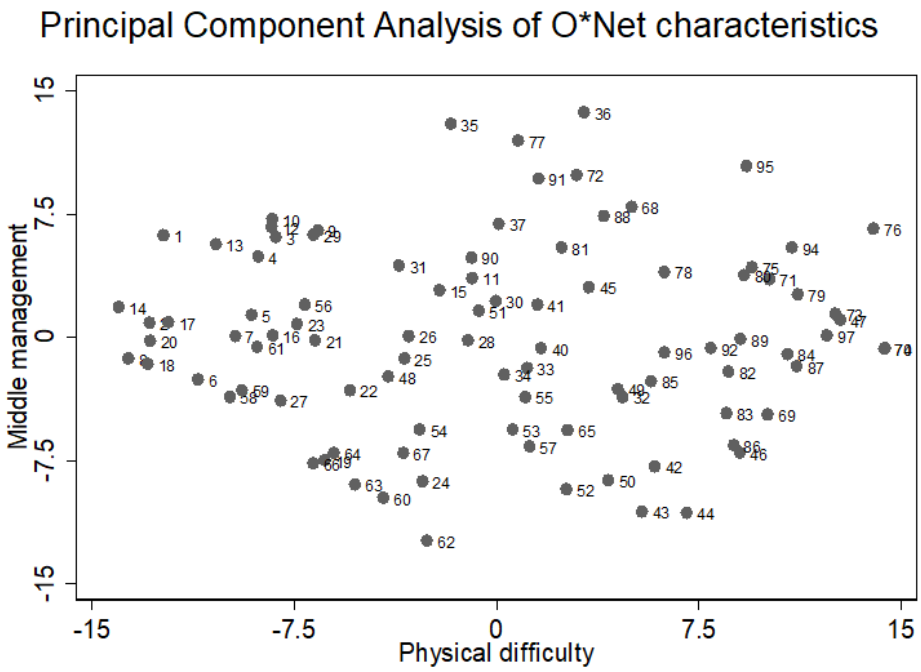
Notice: the Figure shows the baseline survival function of the Cox model (2) using the "US National Longitudinal Mortality Study, 80s-90s" database.

Figure 8: Test of the proportional hazard (PH) assumption



Notice: the Figure presents the test of the proportional hazard assumption of our Cox model (using the data from the "US National Longitudinal Mortality Study, 80s-90s"). We can see that there is no slope in the different plots which is a confirmation that the assumption holds.

Figure 9: Principal Component Analysis of O*Net characteristics



Notice: see Figure 10

Legend:

1: Top Executives

2: Advertising, marketing, promotions, publics relations and sales managers

3: Operations specialities managers

4: Others management occupations

5: Business & Financial operations occupations

6: Financial specialists

7: Computer & Mathematical occupations

8: Mathematical Science Occupations

9: Architects, Surveys and Cartographers

10: Engineers

11: Drafters, Engineering Technicians, and Mapping Technicians

12: Life Scientists

13: Physical Scientists

14: Social Scientists and Related Workers

15: Life, Physical, and Social Science Technicians

16: Counsellors, Social Workers, and Other Community and Social Service Specialists

17: Religious Workers

18: Lawyers, Judges, and Related Workers

19: Legal Support Workers

20: Postsecondary Teachers

21: Preschool, Primary, Secondary, and Special Education School Teachers

22: Other Teachers and Instructors

23: Librarians, Curators, and Archivists

24: Other Education, Training, and Library Occupations

25: Art and Design Workers

26: Entertainers and Performers, Sports and Related Workers

27: Media and Communication Workers

28: Media and Communication Equipment Workers

29: Health Diagnosing and Treating Practitioners

30: Health Technologists and Technicians

31: Other Healthcare Practitioners and Technical Occupations

32: Nursing, Psychiatric, and Home Health Aides

33: Occupational Therapy and Physical Therapist Assistants and Aides

34: Other Healthcare Support Occupations

35: Supervisors of Protective Service Workers

36: Fire Fighting and Prevention Workers

37: Law Enforcement Workers

40: Other Protective Service Workers

41: Supervisors of Food Preparation and Serving Workers

42: Cooks and Food Preparation Workers

43: Food and Beverage Serving Workers

44: Other Food Preparation and Serving Related Workers

45: Supervisors of Building and Grounds Cleaning and Maintenance Workers

46: Building Cleaning and Pest Control Workers

47: Grounds Maintenance Workers

48: Supervisors of Personal Care and Service Workers

49: Animal Care and Service Workers

50: Entertainment Attendants and Related Workers

51: Funeral Service Workers

52: Personal Appearance Workers

53: Baggage Porters, Bellhops, and Concierges

54: Tour and Travel Guides

55: Other Personal Care and Service Workers

56: Supervisors of Sales Workers

57: Retail Sales Workers

58: Sales Representatives, Services

59: Sales Representatives, Wholesale and Manufacturing

60: Other Sales and Related Workers

61: Supervisors of Office and Administrative Support Workers

62: Communications Equipment Operators

63: Financial Clerks

64: Information and Record Clerks

65: Material Recording, Scheduling, Dispatching, and Distributing Workers

66: Secretaries and Administrative Assistants

67: Other Office and Administrative Support Workers

68: Supervisors of Farming, Fishing, and Forestry Workers

69: Agricultural Workers

70: Fishing and Hunting Workers

71: Forest, Conservation, and Logging Workers

72: Supervisors of Construction and Extraction Workers

73: Construction Trades Workers

74: Helpers, Construction Trades

75: Other Construction and Related Workers

76: Extraction Workers

77: Supervisors of Installation, Maintenance, and Repair Workers

78: Electrical and Electronic Equipment Mechanics, Installers, and Repairers

79: Vehicle and Mobile Equipment Mechanics, Installers, and Repairers

80: Other Installation, Maintenance, and Repair Occupations

81: Supervisors of Production Workers

82: Assemblers and Fabricators

83: Food Processing Workers

84: Metal Workers and Plastic Workers

85: Printing Workers

86: Textile, Apparel, and Furnishings Workers

87: Woodworkers

88: Plant and System Operators

89: Other Production Occupations

90: Supervisors of Transportation and Material Moving Workers

91: Air Transportation Workers

92: Motor Vehicle Operators

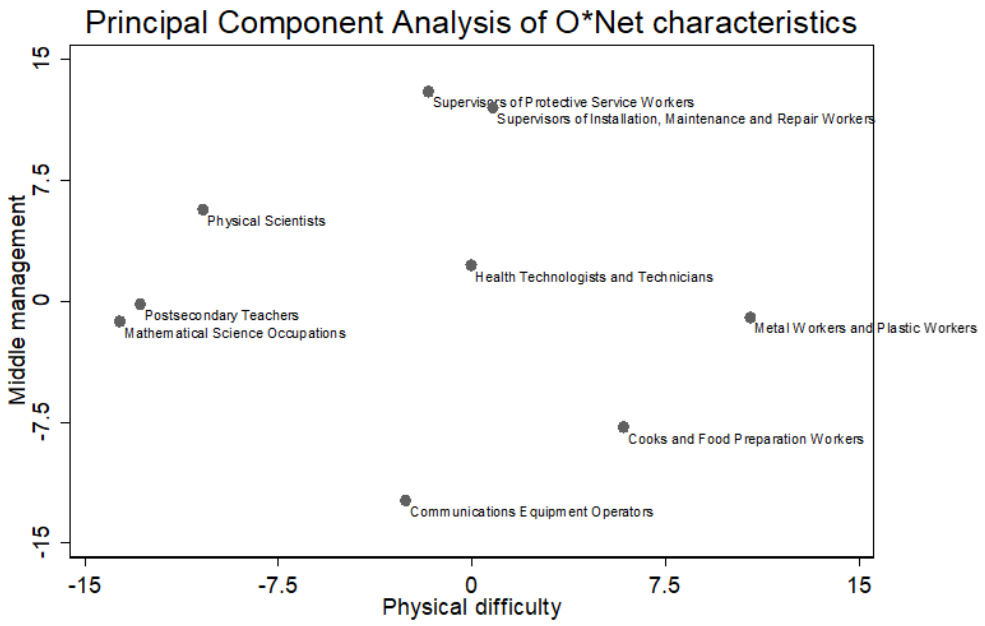
94: Rail Transportation Workers

95: Water Transportation Workers

96: Other Transportation Workers

97: Material Moving Workers

Figure 10: Principal Component Analysis of O*Net characteristics – 2nd Figure



Notice: this Figure presents some professions on a scatter plot scaled on our two components. The first component is the physical difficulty and the second component is related to be in a "middle management" occupation. The PCA was based on the "O*Net" database which is, basically, a very detailed job description. Some jobs are more physical than others (e.g., metal or food workers versus mathematical occupations) and, on the other dimension, supervisors functions score obviously high.

Table 1: Summary statistics of the "European Working Conditions Survey, 2015"

Country	Obs.	Gender		Self-rated health status			Age	Type of employment		Income	Education (ISCED 11)		
		Male	Female	Good	Fair	Bad		Employee	Self		Primary	Secondary	Post-secondary
Austria	822	350	472	642	152	28	42.31	717	105	1'557.25	10	653	159
		43%	57%	78%	18%	3%	(12.41)	87%	13%	(884.66)	1%	79%	19%
Belgium	2'016	989	1027	1610	352	54	42.48	1825	191	1'783.20	87	1057	872
		49%	51%	80%	17%	3%	(11.60)	91%	9%	(787.62)	4%	52%	43%
Bulgaria	834	369	465	684	132	18	44.8	727	107	337.48	6	602	226
		44%	56%	82%	16%	2%	(12.22)	87%	13%	(233.09)	1%	72%	27%
Croatia	659	304	355	521	117	21	43.87	557	102	542.48	7	503	149
		46%	54%	79%	18%	3%	(12.20)	85%	15%	(297.93)	1%	76%	23%
Cyprus	844	428	416	788	50	6	36.45	699	145	1'128.00	19	395	430
		51%	49%	93%	6%	1%	(12.24)	83%	17%	(624.18)	2%	47%	51%
Czeech Republic	582	262	320	480	94	8	43.28	509	73	712.01	2	489	91
		45%	55%	82%	16%	1%	(12.14)	87%	13%	(463.84)	0%	84%	16%
Denmark	853	446	407	711	124	18	44.76	811	42	2'785.21	4	426	423
		52%	48%	83%	15%	2%	(13.06)	95%	5%	(1865.20)	0%	50%	50%
Estonia	651	233	418	378	253	20	45.44	595	56	755.12	0	420	231
		36%	64%	58%	39%	3%	(13.06)	91%	9%	(484.36)	0%	65%	35%
Finland	898	441	457	698	183	17	46.33	742	156	2'238.22	24	412	462
		49%	51%	78%	20%	2%	(12.18)	83%	17%	(1252.01)	3%	46%	51%
France	1'372	652	720	1102	238	32	42.84	1272	100	1'846.58	40	747	585
		48%	52%	80%	17%	2%	(11.44)	93%	7%	(1395.94)	3%	54%	43%
Germany	1'694	847	847	1253	393	48	44.85	1531	163	1'598.87	18	1439	237
		50%	50%	74%	23%	3%	(13.14)	90%	10%	(1036.24)	1%	85%	14%
Greece	604	340	264	549	47	8	41.47	434	170	798.03	42	326	236
		56%	44%	91%	8%	1%	(11.66)	72%	28%	(455.99)	7%	54%	39%
Hungary	462	212	250	351	95	16	45.29	415	47	413.38	1	352	109
		46%	54%	76%	21%	3%	(11.66)	90%	10%	(326.88)	0%	76%	24%
Ireland	820	414	406	744	68	8	42.85	693	127	2'066.92	36	387	397
		50%	50%	91%	8%	1%	(12.39)	85%	15%	(1877.17)	4%	47%	48%
Italy	761	384	377	513	211	37	46.01	588	173	1'379.13	32	591	138
		50%	50%	67%	28%	5%	(11.48)	77%	23%	(1257.32)	4%	78%	18%
Latvia	824	348	476	459	330	35	45.46	722	102	490.18	5	518	301
		42%	58%	56%	40%	4%	(13.14)	88%	12%	(341.18)	1%	63%	37%
Lithuania	903	358	545	537	343	23	45.06	797	106	698.80	0	529	374
		40%	60%	59%	38%	3%	(12.63)	88%	12%	(591.61)	0%	59%	41%
Luxembourg	712	349	363	534	154	24	41.05	646	66	2'875.24	45	413	254
		49%	51%	75%	22%	3%	(10.00)	91%	9%	(1547.02)	6%	58%	36%
Malta	793	462	331	624	160	9	41.56	723	70	1'167.70	41	466	286
		58%	42%	79%	20%	1%	(12.91)	91%	9%	(624.00)	5%	59%	36%
Netherlands	832	417	415	675	131	26	43.84	725	107	1'805.36	38	443	351
		50%	50%	81%	16%	3%	(13.57)	87%	13%	(1979.76)	5%	53%	42%
Poland	688	310	378	518	147	23	41.38	619	69	539.09	5	534	149
		45%	55%	75%	21%	3%	(12.89)	90%	10%	(620.13)	1%	78%	22%
Portugal	670	256	414	440	193	37	46.97	500	170	710.37	233	331	106
		38%	62%	66%	29%	6%	(13.44)	75%	25%	(439.50)	35%	49%	16%
Romania	715	359	356	523	175	17	42.06	617	98	296.93	13	563	139
		50%	50%	73%	24%	2%	(11.27)	86%	14%	(170.34)	2%	79%	19%

Slovakia	691	290	401	482	183	26	43.98	624	67	618.70	2	558	131
		42%	58%	70%	26%	4%	(11.59)	90%	10%	(283.95)	0%	81%	19%
Slovenia	1'143	536	607	849	238	56	43.07	1023	120	926.73	1	770	372
		47%	53%	74%	21%	5%	(11.04)	90%	10%	(488.96)	0%	67%	33%
Spain	2'343	1183	1160	1823	446	74	41.5	2054	289	1'140.82	277	1473	593
		50%	50%	78%	19%	3%	(11.19)	88%	12%	(1019.95)	12%	63%	25%
Sweden	868	433	435	701	139	28	45.01	808	60	2'507.18	4	470	394
		50%	50%	81%	16%	3%	(12.81)	93%	7%	(1822.92)	0%	54%	45%
UK	1'282	686	596	1048	207	27	42.9	1124	158	2'588.84	19	701	562
		54%	46%	82%	16%	2%	(12.98)	88%	12%	(8613.66)	1%	55%	44%
Montenegro	789	438	351	590	168	31	43.68	569	220	385.20	13	569	207
		56%	44%	75%	21%	4%	(12.89)	72%	28%	(221.96)	2%	72%	26%
FYROM	767	429	338	615	124	28	41.9	572	195	254.39	170	379	218
		56%	44%	80%	16%	4%	(12.30)	75%	25%	(193.77)	22%	49%	28%
Serbia	727	371	356	528	155	44	45.63	491	236	307.10	27	507	193
		51%	49%	73%	21%	6%	(13.98)	68%	32%	(781.15)	4%	70%	27%
Turkey	1'667	1203	464	1469	157	41	36.68	1178	489	521.08	434	792	441
		72%	28%	88%	9%	2%	(11.95)	71%	29%	(299.49)	26%	48%	26%
Norway	950	438	512	759	172	19	41.99	882	68	2'964.26	16	461	473
		46%	54%	80%	18%	2%	(13.64)	93%	7%	(2010.37)	2%	49%	50%
Switzerland	692	353	339	607	74	11	41.83	638	54	4'589.03	13	549	130
		51%	49%	88%	11%	2%	(12.79)	92%	8%	(4974.18)	2%	79%	19%
Albania	850	384	466	711	131	8	39.77	493	357	249.66	2	523	325
		45%	55%	84%	15%	1%	(12.98)	58%	42%	(143.38)	0%	62%	38%
Total	32'778	16'274	16'504	25'516	6'336	926	42.84	27'920	4'858	1'355.67	1'686	20'348	10'744
		49.65%	50.35%	77.84%	19.33%	2.83%	(12.57)	85.18%	14.82%	(2295.90)	5.14%	62.08%	32.78%

Notice: this Table presents the summary of the statistics of the "European Working Conditions Survey, 2015". The first statistic is the sex which is usually closed to the half for each gender. The second statistic is the self-rated health status for which we have some variations between countries. These variations could be related to "true" differences in health or to different scales due to cultural factors (Jürges, 2007). The mean age and mean income are computed for each country with their standard deviation. The type of employment is also provided and there are some differences across countries (which does not reflect the one in reality). The last statistics relates to the education level for which the majority of the sample is at the secondary level.

Table 2: Subjective impact on health by job – "European Working Conditions Survey, 2015"

Does your work affect your health?

	No and mainly positively	Yes, mainly negatively	No opinion and refusal	Total	% Bad health
Building and related trades workers, excluding electricians	689 51.9%	592 44.6%	47 3.5%	1'328	2.74%
Stationary plant and machine operators	440 55.2%	337 42.3%	20 2.5%	797	3.02%
Metal, machinery and related trades workers	591 56.4%	427 40.7%	30 2.9%	1'048	1.94%
Labourers in mining, construction, manufacturing and transport	422 57.1%	296 40.1%	21 2.8%	739	3.81%
Drivers and mobile plant operators	781 57.6%	530 39.1%	44 3.2%	1'355	2.86%
Market-oriented skilled forestry, fishery and hunting workers	38 60.3%	24 38.1%	1 1.6%	63	7.31%
Commissioned armed forces officers	23 62.2%	14 37.8%	- 0.0%	37	4.17%
Assemblers	113 59.8%	68 36.0%	8 4.2%	189	1.35%
Protective services workers	438 62.6%	246 35.1%	16 2.3%	700	2.94%
Agricultural, forestry and fishery labourers	240 60.3%	139 34.9%	19 4.8%	398	6.11%
Health associate professionals	463 64.1%	246 34.1%	13 1.8%	722	1.62%
Handicraft and printing workers	120 64.2%	63 33.7%	4 2.1%	187	3.34%
Market-oriented skilled agricultural workers	568 63.6%	301 33.7%	24 2.7%	893	4.72%
Health professionals	654 64.2%	339 33.3%	26 2.6%	1'019	1.81%
Refuse workers and other elementary workers	373 64.1%	194 33.3%	15 2.6%	582	6.47%
Cleaners and helpers	987 65.3%	502 33.2%	22 1.5%	1'511	6.11%
Science and engineering associate professionals	493 65.6%	249 33.1%	10 1.3%	752	1.50%
Electrical and electronic trades workers	300 65.4%	145 31.6%	14 3.1%	459	2.27%
Food processing, wood working, garment and other craft and related trades workers	542 65.7%	258 31.3%	25 3.0%	825	3.52%
Personal care workers	884 66.3%	412 30.9%	37 2.8%	1'333	3.54%
Subsistence farmers, fishers, hunters and gatherers	266 68.2%	116 29.7%	8 2.1%	390	7.48%
Production and specialised services managers	420 68.6%	171 27.9%	21 3.4%	612	3.03%
Food preparation assistants	169 71.3%	62 26.2%	6 2.5%	237	3.90%
Other clerical support workers	284 74.3%	91 23.8%	7 1.8%	382	2.91%
Personal service workers	1'573 74.0%	494 23.2%	58 2.7%	2'125	2.79%
Teaching professionals	1'585 74.8%	469 22.1%	64 3.0%	2'118	1.96%
Legal, social and cultural professionals	564 75.5%	164 22.0%	19 2.5%	747	1.95%
Customer services clerks	486 74.9%	143 22.0%	20 3.1%	649	2.68%
Science and engineering professionals	558 76.2%	159 21.7%	15 2.0%	732	0.63%

Hospitality, retail and other services managers	327 76.9%	91 21.4%	7 1.6%	425	3.41%
Administrative and commercial managers	399 77.2%	109 21.1%	9 1.7%	517	0.92%
Information and communications technicians	159 77.9%	43 21.1%	2 1.0%	204	4.70%
Business and administration professionals	711 77.7%	188 20.5%	16 1.7%	915	1.71%
Business and administration associate professionals	1'259 77.3%	328 20.1%	41 2.5%	1'628	2.63%
Numerical and material recording clerks	622 77.5%	158 19.7%	23 2.9%	803	2.59%
Legal, social, cultural and related associate professionals	341 78.2%	86 19.7%	9 2.1%	436	1.92%
Sales workers	2'401 78.6%	582 19.0%	73 2.4%	3'056	2.55%
Information and communications technology professionals	309 79.2%	73 18.7%	8 2.1%	390	0.62%
Chief executives, senior officials and legislators	295 78.9%	69 18.4%	10 2.7%	374	1.27%
Street and related sales and service workers	36 78.3%	7 15.2%	3 6.5%	46	2.78%
General and keyboard clerks	880 83.4%	149 14.1%	26 2.5%	1'055	1.37%
Total	22'803 69.6%	9'134 27.9%	841 2.6%	32'778	2.86%

Notice: this Table presents the answer to the question "Does your work impact your health?" of the "European Working Conditions Survey, 2015". The answer could be understood as an indicator of self-reported job penibility. We can see that the answers vary across professions. For example, 78,6 % of the "sales workers" think that their health is not or positively affected by their work and only 55,2 % of the "stationary plant and machine operators" think in the same way. We have added in the 6th column the percentage of people that are in bad health[22]. The correlation between this statistic and the third column is not very strong (36,91 %) which is an indicator that using a self-reported job penibility indicator is not an efficient way to determine which jobs can be considered as penible.

Table 3: Self-rated health for the last 12 months by current employment status

Current Employment status	Health in last 12 months			Total
	Good	Fairly good	Not good	
Worked last week	**72.20%**	21.24%	6.56%	100%
Away from last week	64.56%	18.99%	16.46%	100%
Looking for work	67.29%	30.84%	1.87%	100%
Not looking - sick	9.78%	25.00%	65.22%	100%
Economically inactive	**42.84%**	33.04%	24.12%	100%
Total	56.62%	27.10%	16.28%	100%

Notice: this Table shows, with the "Northern Ireland Health Survey, 2010-2011", that one problem of interviewing workers is the overrepresentation of people in good health. In this database, 72,20 % of the people having worked last week declare having been in good health over the last 12 months while they are only 42,84 % to state the same among the economically inactive people.

[22] In order to take into account the difference of age in the sample between the different professions; we have summed the percentage of people in bad health aged between 30-39, 40-49 and 50-60 and then divided it by 3. There is almost no difference between this splitting and 30-44, 45-60 which was done for a robustness test.

Table 4: Result of logistic regression – "European Working Conditions Survey, 2015"

$$\boldsymbol{x\beta} = \beta_1 Age + \beta_2 Age^2 + \beta_3 Self-employment + \beta_4 ISCO + \beta_5 ISCED + \beta_6 ln(Income) + \beta_7 Country + \beta_8 Sex$$

	Models			Odds ratio		
	Ordered Logit	Partial Proportional Model		Ordered Logit	Partial Proportional Model	
		All/Bad	Fair		All/Bad	Fair
Self-employed	-0.116 ***	-0.292 ***	-0.099 **	0.890 ***	0.747 ***	0.906 **
	(0.045)	(0.087)	(0.045)	(0.040)	(0.065)	(0.041)
Agricultural, forestry and fishery labourers	-0.510 ***	-0.504 ***		0.600 ***	0.604 ***	
	(0.134)	(0.134)		(0.081)	(0.081)	
Building and related trades workers, excluding electricians	-0.503 ***	-0.504 ***		0.605 ***	0.604 ***	
	(0.098)	(0.098)		(0.059)	(0.059)	
Cleaners and helpers	-0.434 ***	-0.440 ***		0.648 ***	0.644 ***	
	(0.090)	(0.090)		(0.058)	(0.058)	
Food processing, wood working, garment and other craft and related trades workers	-0.277 ***	-0.276 ***		0.758 ***	0.759 ***	
	(0.108)	(0.108)		(0.082)	(0.082)	
Drivers and mobile plant operators	-0.390 ***	-0.388 ***		0.677 ***	0.678 ***	
	(0.097)	(0.097)		(0.066)	(0.066)	
Labourers in mining, construction, manufacturing and transport	-0.608 ***	-0.611 ***		0.516 ***	0.543 ***	
	(0.110)	(0.111)		(0.060)	(0.060)	
Market-oriented skilled agricultural workers	-0.388 ***	-0.366 ***		0.679 ***	0.693 ***	
	(0.108)	(0.108)		(0.074)	(0.075)	
Market-oriented skilled forestry, fishery and hunting workers	-0.675 ***	-0.673 **		0.509 ***	0.510 **	
	(0.290)	(0.290)		(0.148)	(0.148)	
Metal, machinery and related trades workers	-0.379 ***	0.129	-0.405 ***	0.685 ***	1.138	0.667 ***
	(0.104)	(0.246)	(0.105)	(0.071)	(0.280)	(0.070)
Refuse workers and other elementary workers	-0.559 ***	-0.564 ***		0.572 ***	0.569 ***	
	(0.119)	(0.119)		(0.068)	(0.068)	
Subsistence farmers, fishers, hunters and gatherers	-0.595 ***	-0.568 ***		0.551 ***	0.566 ***	
	(0.142)	(0.141)		(0.078)	(0.080)	
Stationary plant and machine operators	-0.402 ***	-0.405 ***		0.669 ***	0.667 ***	
	(0.109)	(0.109)		(0.073)	(0.073)	
Assemblers	-0.425 **	-0.429 **		0.654 **	0.651 **	
	(0.186)	(0.186)		(0.121)	(0.121)	
Personal care workers	-0.215 **	-0.222 **		0.806 **	0.801 **	
	(0.093)	(0.094)		(0.075)	(0.075)	
Street and related sales and service workers	-0.661 **	-0.656 **		0.516 **	0.519 **	
	(0.335)	(0.334)		(0.173)	(0.173)	
Chief executives, senior officials and legislators	0.305 *	0.312 *		1.356 *	1.366 *	
	(0.162)	(0.163)		(0.220)	(0.222)	
Food preparation assistants	-0.321 *	-0.330 *		0.725 *	0.719 *	
	(0.171)	(0.172)		(0.124)	(0.123)	
Protective services workers	-0.219 *	-0.218 *		0.804 *	0.804 *	
	(0.117)	(0.117)		(0.094)	(0.094)	
Administrative and commercial managers	0.095	0.107		1.100	1.113	
	(0.142)	(0.142)		(0.156)	(0.158)	
Business and administration associate professionals	0.094	0.099		1.099	1.104	
	(0.093)	(0.093)		(0.102)	(0.103)	
Business and administration professionals	0.083	0.091		1.086	1.096	
	(0.110)	(0.111)		(0.120)	(0.121)	
Commissioned armed forces officers	0.531	0.533		1.700	1.703	
	(0.615)	(0.615)		(1.046)	(1.048)	
Customer services clerks	-0.158	-0.155		0.854	0.857	
	(0.123)	(0.123)		(0.105)	(0.106)	
Electrical and electronic trades workers	-0.192	-0.186		0.826	0.830	
	(0.140)	(0.140)		(0.115)	(0.116)	
General and keyboard clerks	0.002	0.575 *	-0.012	1.002	1.777 *	0.989
	(0.104)	(0.300)	(0.105)	(0.104)	(0.534)	(0.103)
Handicraft and printing workers	-0.090	-0.096		0.914	0.908	
	(0.194)	(0.194)		(0.177)	(0.176)	
Health professionals	0.048	0.047		1.049	1.048	
	(0.105)	(0.105)		(0.110)	(0.111)	

Health associate professionals	0.115		0.113			1.122		1.120				
	(0.120)		(0.120)			(0.134)		(0.134)				
Hospitality, retail and other services managers	0.112		0.123			1.118		1.131				
	(0.148)		(0.148)			(0.166)		(0.168)				
Information and communications technicians	- 0.263		- 0.254			0.769		0.776				
	(0.202)		(0.202)			(0.155)		(0.157)				
Information and communications technology professionals	- 0.064		- 0.060			0.938		0.942				
	(0.161)		(0.161)			(0.151)		(0.152)				
Legal, social and cultural professionals	0.031		0.031			1.031		1.031				
	(0.117)		(0.118)			(0.121)		(0.121)				
Legal, social, cultural and related associate professionals	- 0.020		- 0.024			0.980		0.977				
	(0.147)		(0.148)			(0.144)		(0.144)				
Numerical and material recording clerks	0.061		0.066			1.063		1.068				
	(0.117)		(0.117)			(0.125)		(0.125)				
Other clerical support workers	- 0.116		- 0.117			0.890		0.890				
	(0.143)		(0.143)			(0.127)		(0.127)				
Personal service workers	- 0.081		- 0.083			0.922		0.920				
	(0.090)		(0.090)			(0.083)		(0.083)				
Production and specialised services managers	0.012		- 0.522	**	0.048	1.012		0.593	**	1.049		
	(0.125)		(0.237)		(0.126)	(0.126)		(0.141)		(0.132)		
Sales workers	- 0.082		- 0.084			0.921		0.920				
	(0.083)		(0.084)			(0.077)		(0.077)				
Science and engineering associate professionals	- 0.003		0.006			0.997		1.006				
	(0.119)		(0.119)			(0.119)		(0.120)				
Science and engineering professionals	0.129		0.134			1.138		1.143				
	(0.126)		(0.126)			(0.143)		(0.144)				
Female	- 0.159	***	- 0.159	***		0.853	***	0.853	***			
	(0.035)		(0.035)			(0.030)		(0.030)				
Age	- 0.083	***	- 0.085	***		0.354	***	0.342	***			
	(0.007)		(0.008)			(0.089)		(0.088)				
Age^2	0.0004	***	0.0004	***		1.5170	***	1.5669	***			
	(0.0001)		(0.0001)			(0.091)		(0.091)				
Ln(Income)	0.271	***	0.356	***	0.258 ***	1.307	***	1.421	***	1.290	***	
	(0.024)		(0.037)		(0.242)	(0.031)		(0.052)		(0.261)		
Primary education	- 0.452	***	- 0.849	***	-0.375 ***	0.636	***	0.428	***	0.687	***	
	(0.064)		(0.107)		(0.065)	(0.041)		(0.046)		(0.045)		
Post-Secondary Education	0.193	***	0.200	***		1.213	***	1.221	***			
	(0.041)		(0.041)			(0.050)		(0.050)				
cut 1 - constant (bad)	- 4.858		4.318	***								
	(0.247)		(0.315)									
cut 2 - constant (fair)	- 2.390		2.546	***								
	(0.246)		(0.247)									
Pseudo R^2	10.12%		10.59%									
AIC	36'309.01		36'148.91									
BIC	37'006.00		36'971.86									

***: statistically significant at 1%, **: at 5%, *: at 10%

Notice: this Table presents our coefficients and odds ratios from our two models. The first estimated model was an ordered logit, but the proportional odds assumption did not hold and we turned to the "partial proportional model" in which the coefficients needed to vary change across the logit equations. The coefficients of our models can be better interpreted when transformed into odds ratios (in the right-hand side of our Table). The first column of the partial proportional model reports the coefficients of being above bad health versus being in bad health for the coefficients allowed to vary and reports the unique coefficient for the non-varying one. The odds ratio are reported for an increased deviation for age, age^2 and ln(Income) and the ones of countries are omitted. Our baseline is a male Austrian employee as a "teaching professional". We see that large part of jobs are significant and impact negatively the SRHS. The fact that some jobs are insignificant should not be seen as a problem as it could only mean that these jobs are not particularly more penible than the reference one. We also see that being self-employed has a negative effect. Finally, the income has a better effect of protecting against bad health than for improving to good health.

Table 5: Comparison of the data and the prediction of the models

SRHS	Data	Prediction	
		Ordered Logit	Partial Proportional
Bad	2.83%	2.82%	2.82%
Fair	19.33%	19.33%	19.33%
Good	77.84%	77.86%	77.85%

Notice: this Table presents the prediction of our models with regard to the real data ("European Working Conditions Survey, 2015"). We can notice that the difference is almost insignificant; however, we have to remain cautious that it is not because a model can replicate the data that it is necessarily right.

Table 6: Summary statistics and life expectancies of the "US National Longitudinal Mortality Survey"

				Life expectancy at (men)				SRHS	Life expectancy at (women)				SRHS
	Men	Women	Total	N	25	N	50	men	N	25	N	50	women
Supervisors of Personal Care and Service Workers	161	281	442	161	61.52	87	36.52	6.98%	281	52.59	133	28.31	10.97%
	36.43%	63.57%	0.05%		(1.28)		(1.28)			(2.05)		(2.03)	
Social Scientists and Related Workers	1'475	1'353	2'828	1'475	58.97	728	35.62	1.70%	1'353	61.30	601	37.48	4.50%
	52.16%	47.84%	0.29%		(2.63)		(2.70)			(1.34)		(1.31)	
Health Diagnosing and treating Practitioners	7'235	16'478	23'713	7'234	57.61	4'130	34.08	0.89%	16'476	61.29	7'394	37.13	4.51%
	30.51%	69.49%	2.42%		(0.75)		(0.73)			(1.32)		(1.34)	
Lawyers, Judges and Related Workers	4'362	1'126	5'488	4'362	57.18	2'312	32.93	2.27%	1'126	62.20	407	38.28	3.11%
	79.48%	20.52%	0.56%		(0.74)		(0.74)			(2.02)		(2.04)	
Supervisors of Office and Administrative Support Workers	1'726	2'671	4'397	1'726	57.03	1'004	32.96	4.28%	2'671	60.86	1'338	36.36	5.07%
	39.25%	60.75%	0.45%		(1.01)		(0.99)			(2.21)		(2.23)	
Supervisors of Building and Grounds Cleaning and Maintenance Workers	964	373	1'337	964	56.90	488	34.17	3.16%	373	56.50	241	32.07	11.13%
	72.10%	27.90%	0.14%		(2.94)		(3.02)			(0.97)		(0.90)	
Life Scientists	1'245	486	1'731	1'245	56.84	539	33.23	3.82%	486	59.69	165	35.19	3.21%
	71.92%	28.08%	0.18%		(1.83)		(1.85)			(1.20)		(1.16)	
Supervisors of Farming, Fishing, and Forestry Workers	358	37	395	358	56.36	205	32.54	12.87%	37	50.00	17	25.00	0.00%
	90.63%	9.37%	0.04%		(1.77)		(1.68)			(0.00)		(0.00)	
Engineers	12'343	933	13'276	12'341	55.99	6'556	32.12	2.60%	932	52.22	246	27.72	5.67%
	92.97%	7.03%	1.36%		(0.85)		(0.87)			(1.11)		(1.10)	
Postsecondary Teachers	11'241	24'480	35'721	11'239	55.50	6'065	31.92	4.09%	24'480	61.42	12'123	37.23	4.05%
	31.47%	68.53%	3.65%		(0.65)		(0.66)			(0.98)		(0.99)	
Entertainers and Performers, Sports and Related Workers	1'826	893	2'719	1'826	55.46	558	32.16	6.54%	893	58.15	315	35.19	3.43%
	67.16%	32.84%	0.28%		(1.58)		(1.60)			(1.51)		(1.45)	
Religious Workers	2'512	620	3'132	2'512	55.34	1'626	31.25	5.49%	620	61.61	402	37.52	7.21%
	80.20%	19.80%	0.32%		(0.73)		(0.70)			(1.48)		(1.40)	
Architects, Surveys and Cartographers	959	138	1'097	959	55.33	487	32.36	0.78%	138	50.47	39	26.00	0.00%
	87.42%	12.58%	0.11%		(1.41)		(1.33)			(0.53)		(0.00)	
Media and Communication Workers	2'784	2'700	5'484	2'782	55.23	1'245	31.36	5.46%	2'700	59.71	1'102	35.61	4.42%
	50.77%	49.23%	0.56%		(1.01)		(1.01)			(1.30)		(1.31)	
Sales Representatives, Services	5'226	2'518	7'744	5'226	55.10	2'837	31.41	3.09%	2'518	56.70	955	32.47	3.78%
	67.48%	32.52%	0.79%		(0.73)		(0.73)			(2.56)		(2.60)	
Other Teachers and Instructors	1'482	3'168	4'650	1'480	55.10	634	32.18	2.66%	3'167	59.52	1'325	35.00	5.74%
	31.87%	68.13%	0.47%		(1.57)		(1.58)			(.86)		(0.85)	
Financial specialists	7'295	6'804	14'099	7'293	55.06	3'485	31.33	3.63%	6'803	60.71	2'765	38.50	4.93%
	51.74%	48.26%	1.44%		(0.67)		(0.67)			(1.81)		(1.83)	
Others Management Occupations	45'614	22'446	68'060	45'606	54.90	28'114	31.23	3.50%	22'445	59.62	12'152	35.72	4.63%
	67.02%	32.98%	6.95%		(0.29)		(0.29)			(0.77)		(0.78)	
Life, Physical, and Social Science Technicians	1'695	933	2'628	1'694	54.69	685	30.80	6.56%	932	53.45	307	29.22	3.32%
	64.50%	35.50%	0.27%		(1.37)		(1.38)			(1.21)		(1.19)	
Operations specialities managers	18'498	7'546	26'044	18'496	54.60	11'876	31.00	5.36%	7'546	59.50	4'124	35.98	5.96%
	71.03%	28.97%	2.66%		(0.27)		(0.25)			(.78)		(0.77)	
Other education, Training and Library Occupations	319	4'206	4'525	318	54.51	52	31.14	9.35%	4'205	56.26	2'128	32.27	8.17%
	7.05%	92.95%	0.46%		(2.38)		(2.19)			(0.89)		(0.89)	
Computer & Mathematical occupations	5'364	2'403	7'767	5'362	54.29	1'876	30.40	2.83%	2'402	62.09	736	37.42	4.07%
	69.06%	30.94%	0.79%		(1.09)		(1.11)			(1.20)		(1.20)	
Baggage Porters, Bellhops, and Concierges	241	30	271	241	54.28	76	30.31	17.78%	30	50.00	11	25.00	
	88.93%	11.07%	0.03%		(2.02)		(1.93)			-		(0.00)	
Legal Support Workers	309	1'557	1'866	309	54.01	120	32.42	4.62%	1'557	56.27	649	32.47	5.38%
	16.56%	83.44%	0.19%		(2.82)		(2.77)			(1.34)		(1.35)	
Sales Representatives, Wholesale and Manufacturing	10'960	3'705	14'665	10'953	54.00	5'385	30.41	3.43%	3'704	59.43	1'216	35.91	5.87%
	74.74%	25.26%	1.50%		(0.60)		(0.61)			(1.43)		(1.45)	
Supervisors of Installation, Maintenance and Repair Workers	1'148	122	1'270	1'148	53.95	730	30.18	6.39%	122	54.36	68	30.00	6.67%
	90.39%	9.61%	0.13%		(1.89)		(1.90)			(2.93)		(2.92)	
Counsellors, Social Workers, and Other Community and Social Service Specialists	2'141	4'864	7'005	2'140	53.63	1'087	30.22	5.69%	4'863	56.57	2'319	32.66	8.92%
	30.56%	69.44%	0.71%		(1.34)		(1.35)			(0.75)		(0.75)	
Supervisors of Production Workers	10'195	1'767	11'962	10'193	53.40	5'795	30.18	5.61%	1'766	61.15	969	36.75	7.82%
	85.23%	14.77%	1.22%		(0.78)		(0.80)			(4.09)		(4.14)	
Tour and Travel Guides	232	614	846	232	53.38	79	31.13	0.00%	614	59.62	217	35.34	8.12%
	27.42%	72.58%	0.09%		(2.53)		(2.38)			(1.89)		(1.88)	
Other Sales and Related Workers	5'264	6'008	11'272	5'253	53.22	2'440	30.06	8.39%	6'006	60.22	3'148	36.30	8.10%
	46.70%	53.30%	1.15%		(0.55)		(0.49)			(0.85)		(0.85)	
Supervisor of Sales Workers	17'219	9'645	26'864	17'214	53.18	9'209	29.88	5.22%	9'643	59.20	4'820	35.24	6.28%
	64.10%	35.90%	2.74%		(0.40)		(0.40)			(0.80)		(0.81)	
Media and Communication Equipment Workers	944	496	1'440	944	53.18	396	29.78	8.57%	496	55.57	143	31.15	2.67%
	65.56%	34.44%	0.15%		(1.14)		(1.05)			(2.08)		(2.07)	
Business & Financial operations occupations	5'380	7'278	12'658	5'379	53.02	2'925	29.31	3.99%	7'276	67.07	3'163	43.33	3.67%
	42.50%	57.50%	1.29%		(0.57)		(0.56)			(0.98)		(0.99)	
Assemblers and Fabricators	5'841	6'454	12'295	5'831	52.98	2'081	30.46	10.79%	6'452	57.08	2'881	33.50	13.44%
	47.51%	52.49%	1.25%		(1.63)		(1.72)			(1.03)		(1.05)	
Supervisors of Construction and Extraction Workers	3'405	73	3'478	3'405	52.93	1'901	29.58	4.98%	73	45.03	33	21.00	4.44%
	97.90%	2.10%	0.35%		(1.09)		(1.11)			(1.31)		(0.95)	
Top Executives	114	64	178	114	52.84	96	27.84	13.10%	64	53.09	46	30.00	0.00%
	64.04%	35.96%	0.02%		(1.96)		(1.96)			(2.71)		(2.11)	
Vehicle and Mobile Equipment Mechanics, Installers, and Repairers	14'293	146	14'439	14'279	52.83	5'514	29.69	10.08%	146	51.63	55	29.43	0.00%
	98.99%	1.01%	1.47%		(0.85)		(0.88)			(1.90)		(1.17)	
Secretaries and Administrative Assistants	320	31'865	32'185	319	52.76	174	30.19	7.51%	31'856	60.26	13'896	36.28	5.27%
	0.99%	99.01%	3.29%		(1.73)		(1.35)			(0.92)		(0.94)	
Other Office and Administrative Support Workers	5'457	23'119	28'576	5'452	52.74	1'782	30.44	7.99%	23'108	59.88	8'670	36.11	7.75%
	19.10%	80.90%	2.92%		(0.82)		(0.84)			(0.81)		(0.82)	
Librarians, Curators and Archivists	307	1'624	1'931	307	52.68	157	29.02	8.83%	1'624	61.01	1'036	36.54	5.18%
	15.90%	84.10%	0.20%		(1.64)		(1.50)			(61.01)		(0.92)	
Art and Design Workers	2'415	2'985	5'400	2'414	52.68	1'099	29.53	4.52%	2'985	58.87	1'257	34.71	4.39%
	44.72%	55.28%	0.55%		(1.08)		(1.09)			(0.99)		(0.99)	
Textile, Apparel, and Furnishings Workers	3'339	11'363	14'702	3'333	52.53	1'428	29.77	14.41%	11'363	59.19	6'068	35.59	14.94%
	22.71%	77.29%	1.50%		(1.02)		(1.03)			(0.67)		(0.68)	

Retail Sales Workers	14'640 29.45%	35'065 70.55%	49'705 5.07%	14'611	52.38 (0.51)	4'695	29.59 (0.51)	7.64%	35'034	60.59 (0.59)	11'004	36.80 (0.60)	11.39%
Fishing and Hunting Workers	686 92.33%	57 7.67%	743 0.08%	684	52.36 (1.50)	279	29.06 (1.45)	6.91%	57	48.12 (2.21)	30	23.12 (2.21)	
Agricultural Workers	10'249 71.57%	4'071 28.43%	14'320 1.46%	10'210	52.34 (0.55)	3'203	29.47 (0.54)	11.37%	4'067	58.35 (0.86)	1'767	34.34 (0.86)	12.36%
Law Enforcement Workers	5'082 88.37%	669 11.63%	5'751 0.59%	5'081	52.33 (1.28)	1'919	28.77 (1.31)	5.55%	669	51.44 (1.40)	226	28.30 (1.27)	5.11%
Physical Scientists	1'638 80.81%	389 19.19%	2'027 0.21%	1'638	52.25 (1.16)	812	31.24 (1.13)	3.42%	389	48.81 (1.06)	116	26.28 (0.96)	3.37%
Electrical and Electronic Equipment Mechanics, Installers, and Repairers	4'797 91.39%	452 8.61%	5'249 0.54%	4'795	52.25 (0.78)	2'142	28.43 (0.79)	4.91%	452	49.66 (0.81)	187	25.09 (0.76)	1.80%
Other Production Occupations	22'554 60.96%	14'441 39.04%	36'995 3.78%	22'532	52.24 (0.64)	8'843	29.39 (0.67)	9.23%	14'435	58.26 (0.89)	6'612	34.73 (1.91)	12.64%
Motor Vehicle Operators	26'073 89.71%	2'990 10.29%	29'063 2.97%	26'054	52.24 (0.53)	12'230	29.27 (0.54)	8.72%	2'990	55.38 (0.75)	1'399	31.70 (0.74)	10.03%
Construction Trades Workers	46'800 97.48%	1'208 2.52%	48'008 4.90%	46'732	52.12 (0.43)	17'981	29.46 (0.45)	8.81%	1'206	51.16 (1.02)	387	27.14 (0.99)	8.85%
Advertising, marketing, promotions, public relations and sales managers	3'425 77.26%	1'008 22.74%	4'433 0.45%	3'425	52.10 (0.78)	1'705	28.56 (0.75)	1.50%	1'008	54.80 (1.39)	357	30.56 (1.39)	4.02%
Drafters, Engineering Technicians, and Mapping Technicians	6'898 80.80%	1'639 19.20%	8'537 0.87%	6'894	52.10 (0.60)	2'603	28.55 (0.60)	5.70%	1'638	58.23 (1.61)	530	33.95 (1.61)	6.86%
Health Technologists and Technicians	2'511 19.12%	10'625 80.88%	13'136 1.34%	2'508	52.10 (1.11)	912	28.80 (1.13)	5.16%	10'625	58.66 (1.68)	4'102	34.56 (1.71)	6.62%
Animal Care and Service Workers	376 39.54%	575 60.46%	951 0.10%	374	52.07 (3.01)	108	30.41 (3.00)	6.25%	572	54.46 (1.83)	152	30.48 (1.80)	8.51%
Food Processing Workers	2'869 68.80%	1'301 31.20%	4'170 0.43%	2'862	51.99 (1.27)	1'178	29.41 (1.30)	9.60%	1'301	54.10 (1.05)	566	30.49 (1.01)	13.04%
Woodworkers	1'934 87.67%	272 12.33%	2'206 0.23%	1'930	51.95 (1.01)	803	28.44 (1.00)	9.96%	272	49.90 (1.85)	97	25.46 (1.78)	6.73%
Material Recording, Scheduling, Dispatching, and Distributing Workers	18'916 67.25%	9'213 32.75%	28'129 2.87%	18'872	51.84 (0.48)	6'515	28.91 (0.49)	6.80%	9'210	60.88 (1.27)	3'640	37.15 (1.30)	9.31%
Other Installation, Maintenance, and Repair Occupations	15'985 96.89%	513 3.11%	16'498 1.68%	15'975	51.76 (0.40)	7'527	28.63 (0.39)	8.16%	512	54.50 (1.56)	222	30.97 (1.50)	7.46%
Supervisors of Protective Service Workers	987 88.13%	133 11.88%	1'120 0.11%	987	51.63 (1.27)	685	27.55 (1.25)	3.14%	133	56.68 (1.80)	65	31.68 (1.80)	3.51%
Personal Appearance Workers	1'171 18.41%	5'191 81.59%	6'362 0.65%	1'171	51.59 (1.20)	699	30.71 (1.08)	11.00%	5'190	59.78 (0.91)	2'056	35.44 (0.91)	7.69%
Grounds Maintenance Workers	6'191 93.25%	448 6.75%	6'639 0.68%	6'168	51.31 (0.61)	1'905	28.37 (0.58)	13.00%	446	53.46 (1.99)	134	31.39 (1.81)	9.78%
Metal Workers and Plastic Workers	17'332 88.55%	2'241 11.45%	19'573 2.00%	17'317	51.17 (0.55)	7'504	28.25 (0.56)	8.62%	2'240	57.89 (1.31)	974	34.68 (1.32)	11.40%
Other Transportation Workers	2'988 92.94%	227 7.06%	3'215 0.33%	2'983	51.03 (1.16)	617	29.22 (1.12)	10.81%	227	51.07 (1.62)	52	26.86 (1.44)	19.44%
Other Food Preparation and Serving Related Workers	6'062 47.01%	6'834 52.99%	12'896 1.32%	6'035	50.92 (1.14)	869	28.56 (1.17)	9.44%	6'829	57.58 (0.67)	3'013	34.01 (0.66)	10.81%
Information and Record Clerks	5'096 17.38%	24'217 82.62%	29'313 2.99%	5'091	50.87 (0.73)	1'842	28.24 (0.73)	7.71%	24'209	58.75 (0.65)	9'043	34.77 (0.66)	7.86%
Cooks and Food Preparation Workers	8'092 46.95%	9'142 53.05%	17'234 1.76%	8'056	50.78 (1.28)	1'385	28.32 (1.35)	13.45%	9'133	57.45 (0.75)	4'507	33.88 (0.76)	14.44%
Plant and System Operators	2'046 97.43%	54 2.57%	2'100 0.21%	2'045	50.74 (0.98)	1'143	27.34 (0.98)	3.60%	54	53.45 (4.10)	23	28.45 (4.10)	0.00%
Financial Clerks	2'473 9.50%	23'553 90.50%	26'026 2.66%	2'470	50.53 (0.83)	1'047	27.64 (0.79)	9.56%	23'547	60.87 (0.69)	10'654	36.88 (0.70)	6.02%
Printing Workers	2'594 70.87%	1'066 29.13%	3'660 0.37%	2'590	50.51 (0.78)	1'100	27.11 (0.76)	9.62%	1'065	52.06 (0.82)	411	27.92 (0.78)	12.64%
Material Moving Workers	23'864 79.62%	6'107 20.38%	29'971 3.06%	23'817	50.43 (0.52)	7'866	27.82 (0.54)	10.14%	6'106	56.07 (1.55)	2'400	32.57 (1.60)	11.69%
Other Protective Service Workers	5'042 78.99%	1'341 21.01%	6'383 0.65%	5'038	50.40 (0.60)	2'578	27.58 (0.57)	11.72%	1'340	57.10 (1.08)	550	33.12 (1.04)	9.06%
Supervisors of Food Preparation and Serving Workers	719 36.53%	1'249 63.47%	1'968 0.20%	719	50.37 (1.62)	212	27.61 (1.57)	6.62%	1'248	56.31 (1.79)	364	32.53 (1.79)	11.04%
Building Cleaning and Pest Control Workers	15'343 47.71%	16'816 52.29%	32'159 3.28%	15'315	50.36 (0.36)	7'239	28.04 (0.34)	13.03%	16'815	56.73 (0.50)	9'545	33.69 (0.50)	15.75%
Fire Fighting and Prevention Workers	1'796 97.61%	44 2.39%	1'840 0.19%	1'796	50.12 (2.02)	744	26.51 (2.08)	3.11%	44	45.00 -	11	20.00 (0.00)	
Other Personal Care and Service Workers	867 6.01%	13'556 93.99%	14'423 1.47%	863	49.71 (1.80)	159	26.06 (1.78)	12.22%	13'548	58.04 (0.69)	3'952	33.97 (0.69)	11.80%
Supervisors of Transportation and Material Moving Workers	391 86.50%	61 13.50%	452 0.05%	391	49.71 (1.12)	187	25.86 (1.04)	7.56%	61	51.95 (2.70)	26	26.95 (2.70)	9.70%
Other Construction and Related Workers	1'624 97.07%	49 2.93%	1'673 0.17%	1'623	49.61 (0.88)	742	26.81 (0.83)	6.30%	49	48.26 (1.62)	28	23.26 (1.62)	16.67%
Forest, Conservation and Logging Workers	1'063 96.11%	43 3.89%	1'106 0.11%	1'062	49.47 (1.25)	381	26.61 (1.23)	13.27%	43	51.00 (0.00)	15	26.00 (0.00)	
Rail Transportation Workers	1'427 97.41%	38 2.59%	1'465 0.15%	1'427	49.17 (1.13)	773	26.48 (1.08)	4.43%	38	32.96 (4.67)	12	14.40 (3.80)	
Extraction Workers	2'308 97.47%	60 2.53%	2'368 0.24%	2'304	49.14 (0.92)	680	27.01 (0.91)	9.60%	60	51.37 (3.20)	18	31.00 (0.00)	0.00%
Water Transportation Workers	501 98.62%	7 1.38%	508 0.05%	501	49.12 (1.93)	246	26.85 (1.84)	6.20%	7	36.00 (0.00)	3	11.00 (0.00)	
Helpers, Construction Trades	722 96.78%	24 3.22%	746 0.08%	720	49.07 (1.78)	118	26.93 (1.73)	11.96%	24	45.00 (0.00)	6	20.00 (0.00)	40.00%
Entertainment Attendants and Related Workers	1'514 60.22%	1'000 39.78%	2'514 0.26%	1'511	49.01 (1.27)	344	26.70 (1.21)	0.95%	999	58.53 (3.02)	212	35.61 (3.09)	9.05%
Food and Beverage Serving Workers	4'183 20.11%	16'614 79.89%	20'797 2.12%	4'166	48.73 (0.97)	858	27.41 (0.99)	6.86%	16'591	58.55 (1.58)	3'318	35.18 (1.63)	12.53%
Mathematical Science Occupations	1'264 64.69%	690 35.31%	1'954 0.20%	1'263	48.60 (0.72)	632	24.91 (0.52)	4.52%	690	50.54 (1.12)	286	26.76 (0.97)	3.35%

Air Transportation Workers	960 94.21%	59 5.79%	1'019 0.10%	960	48.01 (0.93)	526	25.99 (0.67)	2.51%	59	53.25 (1.71)	14	30.00 (0.00)	0.00%
Nursing, Psychiatric, and Home Health Aides	1'257 9.88%	11'464 90.12%	12'721 1.30%	1'256	47.98 (1.41)	366	26.29 (1.41)	9.45%	11'459	57.62 (0.99)	5'097	34.23 (1.01)	14.45%
Communications Equipment Operators	205 8.98%	2'079 91.02%	2'284 0.23%	204	47.43 (2.70)	56	24.86 (2.52)	20.04%	2'079	55.33 (1.04)	929	32.05 (1.02)	9.71%
Other Healthcare Support Occupations	517 11.66%	3'916 88.34%	4'433 0.45%	517	47.30 (2.14)	134	26.05 (2.11)	11.53%	3'912	55.53 (0.76)	1'191	31.14 (0.76)	8.47%
Military Enlisted Tactical Operations and Air/Weapons Specialists and Crew Members	280 86.42%	44 13.58%	324 0.03%	279	42.48 (1.78)	71	19.63 (1.54)	5.56%	44	31.00 (0.00)	2	6.00 (0.00)	
Total	530'520 54.15%	449'227 45.85%	979'747	709'295	50.63 (0.04)	340'969	27.63 (0.04)	6.60%	791'251	56.24 (0.04)	401'510	32.50 (0.04)	7.79%

Notice: this Table presents the life expectancy and self-rated health status of the "US National Longitudinal Mortality Survey, 80s-90s". The LE is reported for men and for women by professions at 25 and at 50 (using the Kaplan-Meier method). There are indeed some differences between professions; for example, a financial specialist (male) has a LE of 55,06 at 25 and a material moving worker has a LE of 50,43 at the same age. A standardized SRHS is reported in order to take into account the difference in age between the different occupation samples. This figure is created by summing the percentage of "bad" and "fair " people in the slices 30-39, 40-49 and 50-60 and then divided it by 3. As in the previous part, there are large differences between occupations[23]. There is no a clear relation between SRHS and LE which is due to the fact that LE requires some computational technique that cannot be simply replaced by averaging SRHS.

[23] As in the first part (with the "European Working Conditions Survey, 2015"), we have done a robustness test with 30-44 and 45-60; and there is quite no difference.

Table 7: Cox model results of (2) and (3)

$$h(t|\boldsymbol{x}_i) = h_0(t)exp(\beta_1 age_i + \beta_2 age_i^2 + \beta_3 sex_i + \beta_4 educ_i + \beta_5 ln(income)_i + \beta_6 occupation_i + \beta_7 sef - emloyed_i) \quad (2)$$

$$h(t|\boldsymbol{x}_i) = h_0(t)exp(\beta_1 age_i + \beta_2 age_i^2 + \beta_3 sex_i + \beta_4 educ_i + \beta_5 ln(income)_i + \beta_6 occupation_i * self - emloyed_i) \quad (3)$$

	Model(2)		Model(3)	
Self-employed	0.957 (0.014)	***		
Communications Equipment Operators	1.414 (0.148)	***	1.459 (0.153)	***
Rail Transportation Workers	1.388 (0.140)	***	1.413 (0.144)	***
Self-employed:			*1.734 (1.736)*	
Entertainment Attendants and Related Workers	1.346 (0.141)	***	1.373 (0146)	***
Self-employed:			*1.108 (0.785)*	
Nursing, Psychiatric, and Home Health Aides	1.229 (0.068)	***	1.238 (0.070)	***
Self-employed:			*1.200 (0.254)*	
Food and Beverage Serving Workers	1.226 (0.070)	***	1.267 (0.073)	***
Self-employed:			*1.544 (0.263)*	***
Material Moving Workers	1.198 (0.053)	***	1.213 (0.056)	***
Self-employed:			*1.205 (0.162)*	
Cooks and Food Preparation Workers	1.194 (0.061)	***	1.240 (0.065)	***
Self-employed:			*0.717 (0.140)*	*
Other Protective Service Workers	1.192 (0.065)	***	1.206 (0.067)	***
Self-employed:			*0.867 (0.329)*	
Metal Workers and Plastic Workers	1.145 (0.054)	***	1.169 (0.057)	***
Self-employed:			*0.949 (0.127)*	
Building Cleaning and Pest Control Workers	1.136 (0.047)	***	1.150 (0.049)	***
Self-employed:			*1.073 (0.121)*	
Postsecondary Teachers	0.881 (0.042)	***	0.885 (0.044)	***
Self-employed:			*1.099 (0.172)*	
Engineers	0.864 (0.049)	***	0.886 (0.052)	**
Self-employed:			*0.724 (0.118)*	**
Operations specialities managers	0.873 (0.038)	***	1.007 (0.061)	
Self-employed:			*0.826 (0.038)*	***
Physical Scientists	0.716 (0.098)	***	0.684 (0.102)	***
Self-employed:			*0.910 (0.305)*	
Computer & Mathematical occupations	0.713 (0.067)	***	0.729 (0.072)	***
Self-employed:			*0.569 (0.181)*	*
Water Transportation Workers	1.370 (0.217)	**	1.422 (0.234)	**
Self-employed:			*0.814 (0.576)*	
Plant and System Operators	1.209 (0.105)	**	1.226 (0.108)	**
Self-employed:			*0.987 (0.699)*	
Law Enforcement Workers	1.157 (0.086)	**	1.191 (0.090)	**
Self-employed:			*0.964 (0.323)*	
Health Technologists and Technicians	1.127 (0.070)	**	1.154 (0.073)	**
Self-employed:			*0.959 (0.268)*	
Business & Financial operations occupations	1.116 (0.062)	**	1.122 (0.068)	*
Self-employed:			*1.093 (0.119)*	
Other Office and Administrative Support Workers	1.109 (0.053)	**	1.117 (0.055)	**
Self-employed:			*1.178 (0.212)*	
Material Recording, Scheduling, Dispatching, and Distributing Workers	1.101 (0.050)	**	1.108 (0.052)	**
Self-employed:			*1.305 (0.251)*	
Information and Record Clerks	1.096 (0.052)	**	1.104 (0.052)	**
Self-employed:			*1.149 (0.182)*	
Personal Appearance Workers	0.863 (0.063)	**	0.942 (0.104)	
Self-employed:			*0.815 (0.072)*	
Religious Workers	0.841 (0.064)	**	0.867 (0.068)	*
Self-employed:			*0.661 (0.178)*	
Supervisors of Building and Grounds Cleaning and Maintenance Workers	0.761 (0.104)	**	0.868 (0.120)	
Self-employed:			*0.214 (0.151)*	**
Architects, Surveys and Cartographers	0.647 (0.117)	**	0.632 (0.160)	**
Self-employed:			*0.651 (0.170)*	**
Supervisors of Farming, Fishing, and Forestry Workers	0.602 (0.144)	**	0.669 (0.169)	
Self-employed:			0.413 (0.292)	
Military Enlisted Tactical Operations and Air/Weapons Specialists and Crew Members	1.649 (0.480)	*	1.677 (0.488)	*
Other Construction and Related Workers	1.185 (0.121)	*	1.267 (0.137)	**
Self-employed:			*0.852 (0.259)*	
Extraction Workers	1.185 (0.124)	*	1.213 (0.132)	*
Self-employed:			*0.828 (0.340)*	
Other Food Preparation and Serving Related Workers	1.116 (0.064)	*	1.116 (0.066)	*
Self-employed:			*1.291 (0.268)*	
Other Production Occupations	1.085 (0.047)	*	1.114 (0.049)	**
Self-employed:			*0.796 (0.112)*	*
Construction Trades Workers	1.075 (0.045)	*	1.125 (0.050)	***
Self-employed:			*0.943 (0.055)*	
Librarians, Curators and Archivists	0.814 (0.101)	*	0.804 (0.102)	*
Self-employed:			*1.726 (0.865)*	
Media and Communication Equipment Workers	0.825 (0.126)		1.027 (0.180)	
Self-employed:			*0.437 (0.147)*	***
Textile, Apparel, and Furnishings Workers	0.961 (0.050)		1.012 (0.055)	
Self-employed:			*0.749 (0.087)*	***
Advertising, marketing, promotions, public relations and sales managers	0.980 (0.085)		0.953 (0.086)	
Self-employed:			*1.870 (0.568)*	**
Supervisors of Construction and Extraction Workers	0.926 (0.077)		0.977 (0.091)	
Self-employed:			*0.712 (0.122)*	**
Motor Vehicle Operators	1.068 (0.046)		1.062 (0.048)	
Self-employed:			*1.174 (0.087)*	**
Entertainers and Performers, Sports and Related Workers	0.967 (0.101)		1.122 (0.135)	
Self-employed:			*0.715 (0.140)*	*
Other Installation, Maintenance, and Repair Occupations	1.059 (0.051)		1.090 (0.056)	*
Self-employed:			*0.922 (0.090)*	
Agricultural Workers	0.924 (0.049)		0.974 (0.056)	
Self-employed:			*0.714 (0.135)*	
Secretaries and Administrative Assistants *Self-employed:*			*1.157 (0.195)*	
Media and Communication Workers	1.032 (0.077)		0.958 (0.087)	
Self-employed:			*1.198 (0.140)*	
Drafters, Engineering Technicians, and Mapping Technicians	0.954 (0.066)		0.987 (0.070)	
Self-employed:			*0.564 (0.200)*	
Top Executives	1.191 (0.238)		1.233 (0.251)	
Self-employed:			*0.759 (0.759)*	

Counsellors, Social Workers, and Other Community and Social Service Specialists	1.048	1.062
	(0.076)	(0.079)
Self-employed:		*0.998*
		(0.303)
Lawyers, Judges and Related Workers	0.912	0.902
	(0.068)	(0.096)
Self-employed:		*0.881*
		(0.085)
Legal Support Workers	1.092	1.068
	(0.155)	(0.165)
Self-employed:		*1.348*
		(0.479)
Other Teachers and Instructors	1.007	0.896
	(0.083)	(0.099)
Self-employed:		*1.093*
		(0.126)
Other education, Training and Library Occupations	0.989	0.980
	(0.096)	(0.099)
Self-employed:		*1.594*
		(0.534)
Art and Design Workers	1.079	1.092
	(0.083)	(0.108)
Self-employed:		*1.036*
		(0.113)
Health Diagnosing and treating Practitioners	0.932	0.926
	(0.048)	(0.054)
Self-employed:		*0.951*
		(0.070)
Other Healthcare Support Occupations	1.167	1.171
	(0.114)	(0.116)
Self-employed:		*1.106*
		(1.107)
Supervisors of Protective Service Workers	1.006	1.040
	(0.134)	(0.141)
Self-employed:		*0.649*
		(0.460)
Fire Fighting and Prevention Workers	0.934	0.930
	(0.123)	(0.125)
Self-employed:		*1.532*
		(0.886)
Supervisors of Food Preparation and Serving Workers	1.108	1.130
	(0.143)	(0.149)
Self-employed:		*0.694*
		(0.491)
Grounds Maintenance Workers	1.030	1.082
	(0.064)	(0.072)
Self-employed:		*0.872*
		(0.110)
Supervisors of Personal Care and Service Workers	0.804	0.748
	(0.209)	(0.266)
Self-employed:		*0.965*
		(0.395)
Animal Care and Service Workers	1.110	1.186
	(0.194)	(0.246)
Self-employed:		*0.770*
		(0.292)
Baggage Porters, Bellhops, and Concierges	0.546	0.556
	(0.207)	(0.211)
Tour and Travel Guides	0.806	0.939
	(0.174)	(0.203)
Other Personal Care and Service Workers	1.077	1.090
	(0.064)	(0.075)
Self-employed:		*1.035*
		(0.097)
Supervisor of Sales Workers	1.014	0.999
	(0.046)	(0.055)
Self-employed:		*1.003*
		(0.052)
Retail Sales Workers	1.031	1.033
	(0.044)	(0.046)
Self-employed:		*1.065*
		(0.089)
Sales Representatives, Services	0.986	0.942
	(0.064)	(0.075)
Self-employed:		*1.068*
		(0.098)
Sales Representatives, Wholesale and Manufacturing	1.069	1.062
	(0.056)	(0.060)
Self-employed:		*1.118*
		(0.105)
Other Sales and Related Workers	1.012	1.054
	(0.056)	(0.073)
Self-employed:		*0.950*
		(0.069)
Supervisors of Office and Administrative Support Workers	0.883	0.896
	(0.080)	(0.083)
Self-employed:		*0.839*
		(0.485)
Financial Clerks	1.059	1.070
	(0.051)	(0.055)
Self-employed:		*1.117*
		(0.110)
Fishing and Hunting Workers	0.867	1.018
	(0.152)	(0.387)
Self-employed:		*0.804*
		(0.164)
Forest, Conservation and Logging Workers	1.010	1.027
	(0.142)	(0.189)
Self-employed:		*1.014*
		(0.215)
Helpers, Construction Trades	0.913	0.959
	(0.212)	(0.223)
Supervisors of Installation, Maintenance and Repair Workers	0.922	0.952
	(0.119)	(0.125)
Self-employed:		*0.742*
		(0.525)
Electrical and Electronic Equipment Mechanics, Installers, and Repairers	0.965	0.986
	(0.075)	(0.081)
Self-employed:		*0.896*
		(0.173)
Vehicle and Mobile Equipment Mechanics, Installers, and Repairers	1.002	1.015
	(0.053)	(0.059)
Self-employed:		*0.977*
		(0.085)
Supervisors of Production Workers	1.004	1.013
	(0.053)	(0.055)
Self-employed:		*0.947*
		(0.247)
Assemblers and Fabricators	1.092	1.100
	(0.063)	(0.065)
Self-employed:		*1.259*
		(0.309)
Food Processing Workers	0.984	1.024
	(0.079)	(0.085)
Self-employed:		*0.951*
		(0.227)
Printing Workers	1.093	1.107
	(0.091)	(0.096)
Self-employed:		*1.103*
		(0.335)
Woodworkers	1.022	1.125
	(0.096)	(0.119)
Self-employed:		*0.809*
		(0.151)
Supervisors of Transportation and Material Moving Workers	0.966	1.018
	(0.219)	(0.231)
Air Transportation Workers	0.992	0.967
	(0.157)	(0.163)
Self-employed:		*1.433*
		(0.643)
Other Transportation Workers	1.065	1.052
	(0.096)	(0.100)
Self-employed:		*1.323*
		(0.335)
Female	0.507 ***	0.507 ***
	(0.007)	(0.007)
Age	1.028 ***	1.028 ***
	(0.004)	(0.004)
Age^2	0.999 ***	0.999 ***
	(0.0001)	(0.0001)
Ln(Income)	0.780 ***	0.779 ***
	(0.005)	(0.005)
Elementary education	0.963 ***	0.961 ***
	(0.014)	(0.014)
College education	0.826 ***	0.826 ***
	(0.010)	(0.010)

***: statistically significant at 1%, **: at 5%, *: at 10%

Notice: this Table presents our estimates of the Cox model (2) and (3) on the data of the "US National Longitudinal Mortality Survey, 80s-90s". Both models control for the education, age, sex and income. Lot of jobs are significant (28 at 5 % and 7 at 10 %) and impact the baseline hazard differently. For example, serving in a restaurant increases the hazard by 22,6 % while being a librarian decreases it by 18,6 %. As being self-employed is significant, we inquire in regression (3) into the interaction coefficient of profession with being or not self-employed. A lot of our coefficients remains significant in this new framing (either in self-employed, employed or both) and the interaction coefficient with self-employed tends to be lower which is consistent with the coefficient in (2).

Table 8: Cox model results of (4)

$$h(t|\boldsymbol{x}_i) = h_0(t)exp(\beta_1 age_i + \beta_2 age_i^2 + \beta_3 sex_i + \beta_4 educ_i + \beta_5 ln(income)_i + \beta_6 occupation_i * SRHS_i + \beta_7 self - employed) \quad (4)$$

Self-employed	0.933	(0.034)	*
Others Management Occupations			
Bad	*3.427*	*(0.864)*	***
Fair	*1.904*	*(0.364)*	***
Very Good	*0.657*	*(0.110)*	***
Exellent	*0.644*	*(0.109)*	***
Operations specialities managers			
Bad	*2.177*	*(0.597)*	***
Fair	*1.470*	*(0.286)*	**
Good	*0.629*	*(0.119)*	***
Exellent	*0.555*	*(0.112)*	***
Postsecondary Teachers			
Bad	*4.191*	*(1.819)*	***
Fair	*1.953*	*(0.485)*	***
Very Good	*0.650*	*(0.127)*	**
Exellent	*0.481*	*(0.098)*	***
Financial specialists			
Fair	*1.830*	*(0.559)*	**
Exellent	*0.486*	*(0.130)*	***
Engineers			
Fair	*2.074*	*(0.757)*	**
Very Good	*0.498*	*(0.128)*	***
Exellent	*0.425*	*(0.116)*	***
Lawyers, Judges and Related Workers			
Fair	*3.066*	*(1.076)*	***
Very Good	*0.526*	*(0.184)*	*
Exellent	*0.481*	*(0.156)*	**
Health Diagnosing and treating Practitioners			
Fair	*2.316*	*(0.620)*	***
Very Good	*0.656*	*(0.140)*	**
Exellent	*0.419*	*(0.096)*	***
Media and Communication Workers			
Exellent	*0.391*	*(0.150)*	***
Nursing, Psychiatric, and Home Health Aides			
Bad	*3.053*	*(1.065)*	***
Very Good	*0.595*	*(0.168)*	*
Supervisor of Sales Workers			
Fair	*1.694*	*(0.378)*	**
Very Good	*0.703*	*(0.133)*	*
Exellent	*0.490*	*(0.105)*	***
Other Sales and Related Workers			
Bad	*4.839*	*(1.965)*	***
Fair	*2.108*	*(0.531)*	***
Very Good	*0.491*	*(0.136)*	***
Retail Sales Workers			
Bad	*3.847*	*(1.068)*	***
Fair	*1.463*	*(0.310)*	*
Very Good	*0.687*	*(0.131)*	**
Exellent	*0.651*	*(0.137)*	**
Vehicle and Mobile Equipment Mechanics, Installers, and Repairers			
Bad	*3.490*	*(1.647)*	***
Very Good	*0.522*	*(0.138)*	***
Exellent	*0.597*	*(0.170)*	*
Other Construction and Related Workers			
Bad	*5.446*	*(3.936)*	***
Fair	*2.522*	*(1.318)*	*
Good	*2.371*	*(0.675)*	***
Sales Representatives, Services			
Fair	*2.519*	*(0.880)*	***
Very Good	*0.597*	*(0.170)*	*
Exellent	*0.558*	*(0.166)*	**
Food Processing Workers			
Fair	*2.852*	*(1.042)*	***
Very Good	*0.351*	*(0.183)*	**

Metal Workers and Plastic Workers			
Bad	*3.548*	*(1.442)*	***
Woodworkers			
Fair	*4.391*	*(2.071)*	***
Motor Vehicle Operators			
Excellent	*0.660*	*(0.139)*	**
Other Transportation Workers			
Good	*2.269*	*(0.739)*	***
Material Moving Workers			
Bad	*3.861*	*(1.253)*	***
Fair	*1.701*	*(0.389)*	**
Other Installation, Maintenance, and Repair Occupations			
Bad	*3.694*	*(1.350)*	***
Fair	*1.784*	*(0.457)*	**
Excellent	*0.621*	*(0.157)*	*
Agricultural Workers			
Bad	*2.560*	*(0.983)*	***
Fair	*1.861*	*(0.510)*	**
Very Good	*0.474*	*(0.149)*	**
Other Office and Administrative Support Workers			
Bad	*5.122*	*(1.717)*	***
Fair	*2.429*	*(0.517)*	***
Excellent	*0.594*	*(0.195)*	**
Secretaries and Administrative Assistants			
Fair	*1.910*	*(0.553)*	**
Very Good	*0.602*	*(0.134)*	**
Excellent	*0.617*	*(0.147)*	**
Other Protective Service Workers			
Bad	*2.455*	*(0.997)*	**
Personal Appearance Workers			
Bad	*2.576*	*(1.123)*	**
Very Good	*0.405*	*(0.176)*	**
Excellent	*0.114*	*(0.115)*	**
Textile, Apparel, and Furnishings Workers			
Very Good	*0.537*	*(0.160)*	**
Excellent	*0.303*	*(0.158)*	**
Supervisors of Protective Service Workers			
Good	*2.041*	*(0.686)*	**
Financial Clerks			
Fair	*1.790*	*(0.507)*	**
Very Good	*0.600*	*(0.136)*	**
Excellent	*0.620*	*(0.160)*	*
Assemblers and Fabricators			
Bad	*3.093*	*(1.344)*	***
Very Good	*0.614*	*(0178)*	*
Excellent	*0.515*	*(0.180)*	*
Information and Record Clerks			
Fair	*1.995*	*(0.467)*	***
Very Good	*0.709*	*(0.145)*	*
Excellent	*0.644*	*(0.149)*	*
Building Cleaning and Pest Control Workers			
Bad	*2.233*	*(0.564)*	***
Fair	*1.417*	*(0.267)*	*
Supervisors of Office and Administrative Support Workers			
Very Good	*0.487*	*(0.198)*	*
Excellent	*0.266*	*(0.158)*	**
Business & Financial operations occupations			
Bad	*4.660*	*(2.196)*	***
Fair	*1.871*	*(0.628)*	*
Excellent	*0.657*	*(0.159*	*
Social Scientists and Related Workers			
Fair	*3.932*	*(2.840)*	*
Excellent	*0.309*	*(0.161)*	**
Other Teachers and Instructors			
Very Good	*0.509*	*(0.195)*	*
Computer & Mathematical occupations			
Excellent	*0.599*	*(0.174)*	*
Life Scientists			
Very Good	*0.171*	*(0.173)*	*
Legal Support Workers			
Very Good	*0.309*	*(0.223)*	*
Other education, Training and Library Occupations			
Very Good	*0.476*	*(0.207)*	*
Health Technologists and Technicians			
Fair	*1.831*	*(0.667)*	*
Very Good	*0.617*	*(0.171)*	*
Excellent	*0.520*	*(0.174)*	**
Other Healthcare Support Occupations			
Good	*1.754*	*(0.588)*	*

Law Enforcement Workers			
Fair	*2.097*	*(0.851)*	*
Good	*1.720*	*(0.512)*	*
Cooks and Food Preparation Workers			
Fair	*1.507*	*(0.365)*	*
Excellent	*0.545*	*(0.177)*	*
Other Food Preparation and Serving Related Workers			
Fair	*1.964*	*(0.616)*	**
Other Personal Care and Service Workers			
Excellent	*0.528*	*(0.202)*	*
Material Recording, Scheduling, Dispatching, and Distributing Workers			
Very Good	*0.680*	*(0.140)*	*
Excellent	*0.663*	*(0.154)*	*
Supervisors of Installation, Maintenance and Repair Workers			
Excellent	*0.185*	*(0.187)*	*
Electrical and Electronic Equipment Mechanics, Installers, and Repairers			
Excellent	*0.485*	*(0.197)*	*
Female	0.606	(0.019)	***
Age	1.087	(0.010)	***
Age²	0.999	(0.0001)	***
Ln(Income)	0.837	(0.014)	***
Elementary education	0.659	(0.034)	***
College education	0.857	(0.025)	***

***: statistically significant at 1%, **: at 5%, *: at 10%

Non-significant coefficients omitted

Notice: this Table presents our last Cox model using the data of the "US National Longitudinal Mortality Survey, 80s-90s". The last term was added to take into account for the interaction between the profession and the SRHS. SRHS has an impact as being in bad health increases the coefficients. The increase between our coefficients in (2) and in (4) is not the same across professions which is an indication that some profession have more impact depending on the "starting" level of health.

Table 9: Regression of the Cox model (2) on the principal components

	Coefficients	
Phsyical difficulty	0.007	***
	(0.002)	
Middle management	- 0.006	**
	(0.003)	
constant	1.010	***
	(0.016)	
R^2	15.81%	

***: statistically significant at 1%, **: at 5%

Notice: this Table presents our regression of the Cox model (2) (based on the "US National Longitudinal Mortality Survey, 80s-90s") on the principal component (based on the "O*Net" data). The coefficient of physical difficulty is positive which means that being in a physical work impacts negatively the health. However, being in a middle management does not seem to impact it negatively.