International Institute of Social Studies

Erasmus

Impacts of Long Working Hours on the Mental Health of European Above the Age of 50

A Research Paper presented by:

Le Thi Anh Mai

(Vietnam)

in partial fulfilment of the requirements for obtaining the degree of

MASTER OF ARTS IN DEVELOPMENT STUDIES

Major:

Economics of Development

(ECD)

Members of the Examining Committee:

Dr. Matthias Rieger

Dr. Thang Vo (UEH)

The Hague, The Netherlands

December 2020

Disclaimer:

This document represents part of the author's study programme while at the International Institute of Social Studies. The views stated therein are those of the author and not necessarily those of the Institute.

Inquiries:

International Institute of Social Studies
P.O. Box 29776
2502 LT The Hague
The Netherlands

t: +31 70 426 0460
e: info@iss.nl
w: www.iss.nl
fb: http://www.facebook.com/iss.nl
twitter: @issnl

Location:

Kortenaerkade 12
2518 AX The Hague
The Netherlands

Contents

List of Tables

List of Figures

List of Appendices

List of Acronyms

FE	Fixed effects
IV	Instrumental variables
OLS	Ordinary least squares
RE	Random effects
RDD	Regression discontinuity design
RKD	Regression kink design
SHARE	Survey of Health, Ageing and Retirement in Europe

Acknowledgement

First of all, I would like to send my deepest thanks to my beloved parents. Thank you for your unconditional love, for never losing hope and trust in me, for forgiving all of my life mistakes.

I would like to express my sincere gratitude to my supervisors, Dr. Matthias Rieger and Dr. Thang Vo. Thank you, Dr. Thang Vo, for your great inspiration on the research topic, your high expectation and instruction along the way. Thank you, Dr. Matthias Rieger, for your guidance and advice through and through. Thank you very much, professors, without your academic and mental, emotional support, I could not have finished this paper!

Finally, to all of my special friends, who have been coming and staying by my side for all these years, through thick and thin, thank you. Thank you for making me a better and stronger me!

Abstract

Using data from three waves of the Survey of Health, Ageing and Retirement in Europe (SHARE), this paper examines the effects of long working hours on the depression symptoms for a population aged over 50 years in 11 European countries. The fixed effects regressions and instrumental variable approach give evidence of positive effects of work on the mental health outcome while confirms the adverse impact of working longer than the European regulatory recommendation, which is a 48-hours weekly limit. In an attemp of applying a regression kink design, the study also find a gradual negative effect of retirement on the old-age depression.

Relevance to Development Studies

Labour force is one of the most crucial factors in the economic development. Studies of the impact of working hours on health outcome is extremely relevant, especially for the highly-developed continent of Europe, where the population is aging with a rapid rate. The study results could serve as suggestion for better labour legislation, which can improve productivity as well as the overall well-being of European citizens, and encouragement for similar studies in other regions worldwide.

Keywords

Long working hours, mental health, aged population, Europe, SHARE, EURO-D depression scale, panel data, fixed effects, instrumental variables, regression kink design

CHAPTER 1 INTRODUCTION[1]

1.1 Research background

According to the European Union's Working Time Directive (European Parliament, 2003), EU Member States need to guarantee to limit the average working hours of their residents to the maximum of 48 hours per week, including over time. The regime seems to have strong enforcement as the work duration in Europe shows a decline trend over time. In 2018, statistics shows that European works 37.1 hours weekly on average in their full-time jobs (Eurostat, 2020)

On a different note, the age structure of Europe indicates an aging continent with the majority of countries having more than 15 or even 20% of total population aged 65 years or older (World Bank, 2019). Whether such demographic trend negatively impacts the national and regional economies is still under debate, but there are clear efforts in increasing retirement age to sustain the labour force (Finnish Centre for Pensions, 2020). In addition, studies such as of Kim and Fieldman (2000) and Dingemans and Henkens (2019) about working after retirement in Europe shows a positive relationship between life satisfaction and continuous working at the old age. As a result, there is an absolute need for research about working intensity and conditions and more focused labour regulations for the older age group.

Adequate amount of work has always been a topic of debate over the concerns of economic growth, public health and human development. Even though longer working hours appear to utilize resources, fixed costs and unproductive time, they are also associated with worker fatigue, increase in task errors, absenteeism, occupational injuries and illnesses (Dembe, 2005; Keller, Berryman and Lukes, 2009; Holden *et al.*, 2010; Collewet and Sauermann, 2017). There is empirical evidence that long working hours adversely affect health, physically as well as psychologically. People who work longer hours face higher risks of heart disease, anxiety, insomnia, and depression (Bannai and Tamakoshi, 2014) . Other studies also find links between extended working hours with bad health behaviors such as smoking, alcohol consuming, over eating, lack of exercise. These negative impacts do not only affect the workers but also their families, employers and society (Caruso *et al.*, 2006).

[1] Some parts of this paper including the data used have been used for my assignments in the following course: 3015, 4317, 4391.

1.2 Research question and contribution

This research aims to answer the question: Do long working hours affect the mental health of the European aged population? In other words, we explore the relationship between long working hours and mental health outcome in a specific population of European who are over 50 years old. The research contributes to the literature in three ways. First, the research makes use of a secondary panel data set from The Survey of Health, Ageing and Retirement in Europe (SHARE) (Börsch-Supan, 2020a, 2020b, 2020c) which is an intensive multidisciplinary and cross-national database with the size significantly large compared to the majority of other studies on the same topic. Secondly, the employment of the panel data with fixed effects regressions could control for unobservable changes within-individual over time. In addition, the identification strategies also consist of two approaches to tackle the problem of reverse causality. Besides the instrumental variable (IV) FE approach which is commonly used to infere causal relationship in the absence of randomized controlled experiments, we also attemp a regression kink design (RKD) in order to overcome the problem of weak instrumental variables.

1.3 Scopes and limitation

There are still certain limitations remained in the study. As being conducted in multiple nations, the released database does not provide clear information about educational attainment and income level, which are two important confounders in the relationship between working hours and depression. The self-reported data on the health conditions could be not accurate and could bias the results in unexpectable ways. Finally, the research might not be able to include all the confounders in the regression models, and those obimitted variables could create biases on the impact of long working hours on mental health.

1.4 Structure of the research paper

The remainder of the paper proceeds as follows. Chapter 2 critically summarizes the literature on the subjects of long working hours, the relevant theoretical frameworks and provides an overview of the labour market in Europe. The following section describes the data and the empirical strategies applied to answer the research question. Chapter 4 presents and discusses the descriptive statistics of the data used and the estimates from the regression models. The last chapter concludes the study with the policy implications.

CHAPTER 2 LITERATURE REVIEW

This chapter reviews the key literature on the topic of working hours and health outcomes, and the theoretical framework that the research paper plans to apply. As the subject of study is the European aged population, there is also a summary about the situation of working hours and the labour market in the region. The contribution of the research to literature concludes the chapter.

2.1 The impacts of long working hours

The systematic review of Bannai and Tamakoshi (2014) includes a thorough analysis of 17 articles from 19 studies published in Medline and PsycINFO during seven-year period of 1995-2012 after the initial search from 5,088 articles. The chosen studies are observational with prospective cohort, case-control, and cross-sectional studies; shift-working studies are excluded; the average time (or plus one standard deviation) spent on work exceeds 40 hours/week or 8 hours/day; the studied health outcomes could directly cause death or diseases. The review documents the associations long working hours and circulatory diseases, depressive state, anxiety, sleep condition while cannot conclude on the effects of long working hours on all-cause mortality, diabetes, other mental states, cognitive function and behaviour.

One of the most recent meta-analysis of Wong, Chan and Ngan (2019) examines the effect of long working hours on occupational health studied in 243 articles since 1997. The analysis adopts the random effect model to obtain odds ratios; I-square statistics are used as a heterogeneity indicator. After adjusting for publication bias, the study shows that workers who work longer than 50 hours/week or 10 hours/day are likely more likely to suffer from occupational health problems. The effects on two out of five categories of health are statistically significant in the meta-regression, which are physiological health and related health. In the category of physiological health, the effects of long working hours on cardiovascular disease and metabolic syndrome are statistically significant, with the stronger effects are on cardiovascular disease (Odds ratio of 1.539 and 1.100 respectively). The effects are all statistically significant on five symptoms of related health, naming fatigue, injury, poor sleep quality, short sleep duration and sleep disturbance, with short sleep duration is the most severe problem (highest odds ratio of 1.909). The results on sleep condition are quite consistent with that of Bannai and Tamakoshi (2014), although the cut-off point for long working hours are 10 hours weekly different.

Following Caruso *et al.* (2006) 's framework to looking further into the effects that could moderate the association between long working hour and psychosomatic health complaints, Müller et al. (2018) employ a German nationwide representative survey of full-time employees, define long working hours as over 48 hours per week, and use the Index of Psychosomatic Complaints. While being able to confirm the hypothesis that there is a positive relation between long working hours and health complaints, the study does not find valid support for moderating roles of adverse working conditions, including deadline and performance pressure, permanent availability and change in work hours.

Empirical evidence in literature about the relationship between long working hours and productivity is not consistent either. Some studies conclude that working hours is proportional to output - the measure of productivity; however the estimating models in those studies could suffer from biases (Collewet and Sauermann, 2017). The majority of studies find evidence of productivity decreased with longer working time with both macro and micro databases. Using two sample sets of meta-database and one observation per countries, Burger (2015) observes a correlation close to zero for the former set and a negative one for the latter. Many other studies also find similar decreasing returns to hours in specific sectors, such as in manufacturing (Shepard and Clifton, 2000; Shimizu *et al.*, 2004), in construction (Hanna Awad S., Taylor Craig S., and Sullivan Kenneth T., 2005), and in service sector (Collewet and Sauermann, 2017). For example, in the study of Collewet and Sauermann (2017), they study the link between working hours and productivity using data from a call centre in the Netherlands. Some reasons behind the lower productivity could be work fatigue, higher accident and error rate, or self-pacing by workers in longer working days (Hanna, Taylor and Sullivan, 2005; Collewet and Sauermann, 2017).

2.2 Issues in studies of long working hours

The literature on impacts of long working hours is quite extensive, but many of those studies employ cross-sectional data thus are unable to account for time-invariant characteristics, focusing on specific industries or occupations, or do not provide clear definition of long working hours (Angrave, Charlwood and Wooden, 2014; Milner, Smith and LaMontagne, 2015; Kim *et al.*, 2016; Collewet and Sauermann, 2017).

Caruso et al. (2006) realize the difficulties and limits lying in studies of working hours and health impacts as inconsistency in definition of long working time, various factors that can lead to extended working hours and moderate the impacts, the sample size and characteristics. They argue that there are economic, societal as well as individual needs and preferences that generate complete work schedules, not just mere numbers of working hours. Therefore, confounding factors such as socio-economic status, demographic variables, lifestyles, and household structures could also hinder the research results as many studies fail to control for those factors. Those work schedule characteristics combined with long duration of work reduce the time available for other non-work activities and give workers longer exposure and make them more vulnerable to work demand and workplace hazards. According to Caruso and colleagues, long working hours then produces immediate effects on workers' health, such as sleep disturbance, fatigue, stress, neurological, cognitive, and physiological dysfunctions, moderated by factors including the characteristics of individuals, the jobs and working environments. The immediate effects of long working hours can impact not only the health quality of workers, but also their families, employers and communities. Therefore, in order to have effective interventions, Caruso *et al.* propose a thorough conceptual framework that highlights the complex aspects of long working hours. In their report, they also emphasize the need of studies on higher vulnerable group such as pregnant women, socio-economically disadvantaged workers, or older workers. Among the confounding factors that could affect both working hours and health, age seems to be the most observable one.

Another empirical challenge in studying the causal effects of long working hours on health is reverse causality. Working hours and health outcome are likely to be endogenous. Healthy people might have the tendency to self-select into working extra hours, and people who have health problem might not be able to work for long.

2.3 Theoretical framework

As mentioned, hours spent working reduce the amount of time allocated for other activities. Therefore, the Becker's theory of the allocation of time (Becker, 1965) and a further development from it – the Grossman model of human capital (Grossman, 1972, 2000) are highly relevant in the study of long working hours. The theory of Becker (1965) makes a major assumption that household are both producers and consumers. In the utility maximization function, the inputs are both goods (income) and time, and the increase in earning by increasing hours of work also induces a decline in leisure time.

In the human capital and demand for health model of Grossman (1972), health could be view as a form of human capital or investment, as well as a commodity of good health. According to the model, individuals own a certain amount of health stock that depreciates with time and can be used to improve quality of life and work, but also can be increased with good investment. The Grossman model has been extensively applied in studies about costs and benefits of medical treatments and health-related decision making (Khwaja, 2010; Papageorge, 2016; Cronin, Forsstrom and Papageorge, 2017). The possible implication from the Grossman model is that the long working hours could negatively affect mental health by reducing time spent for healthy activities such as entertainment and exercise or have positive influence as extra work raises income that reduce financial constraint and possibly increase fiscal investment in health. That two-way meachnism suggests a reverse casual correlation. In addition, as the stock of health is time-depreciated, age could be the one of themost crucial confounders in the relationship between working hours and mental health that requires extra attention during empirical design.

2.4 Working hours and age in Europe

Statistics of 2019 shows that people in 28 European countries work 37.0 hours weekly on average in their full-time jobs (Eurostat, 2020). According to Eurostat (2020), work duration in Europe shows a decline trend since the issue of European Union's Working Time Directive 2003/88/EC (European Parliament, 2003), which limit the average working hours to the maximum of 48 hours per week, including over time. As the labour market in Europe tends to move towards a healthy "work-life balance" trend, academia also appears to pay little attention on the issues of extended working hours in this region (Burger, 2015). However, the study of Burger (2015) by looking at the trends and causes for extreme working hours in different North American and European countries during the period of 1970-2010 detects a divergence: a few countries such as France and Scandinavian countries maintain a balanced working hour profile while many others shows a convergence with North America's pattern in which ratios of extreme working hours increases over the period. The trend is found to start since the beginning of the 1990s. In fact, the 2003/88/EC Directive has the opt-out option which allows employers to not be obliged to the limit of 48 hours with workers' consent. 18 Member States use this opt-out either regardless of the industry or with some restrictions. For this opt-out option, there is no maximum limit of the number of working hours (Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee, 2017).

On a different note, the European labour market faces a problem of declining proportion of working age as the greying continent has the majority of its countries with 15 or even 20% of total population aged 65 years or older (World Bank, 2019). 65 years is currently the most common retirement age in Europe, but many countries have established attempts to raise this limit to 67 or even higher in order to maintain the labour force and sustain the pension system (Finnish Centre for Pensions, 2020). However, even in the group of older workers who are from 55 to 64 years old - still under retirement age, the employment rate is below 50% for the EU-27 in 2010, and the figure is significant lower for women than men (European Commision, 2012). For older workers, to extend their working life, working environment and job autonomy such as the flexible working hours seems to be key determinants (Mullan, Vargas Llave and Wilkens, 2017). There is also empirical evidence that continuous working at old age or bridge employment after retirement have positive effects on retirement satisfaction as well as life satisfaction (Kim and Feldman, 2000; Dingemans and Henkens, 2019).

2.5 Contribution of the research paper

There is limited research focusing on the impacts of long working hours in the older generation, before and after retirement age; therefore, this research aims to fill in this gap. The research employs a significantly large cross-national database that covers general population instead of workers in specific sectors, thus could draw a more general conclusion for the region's labour force. The longitudinal data also allows for the examination of changes within person over time in the number of working hours and health conditions. Furthermore, this study contributes to the literature by applying an instrumental variables (IV) approach to confirm the impact of long working hours on mental health and overcome the issue of reverse causality. Lastly, anticipating the insufficient strength of our IVs, the study also applies a RKD design to investigating the impact of the sudden change in working hours at the retirement age on the depression of the aging population.

2.6 Summarization of reviewed papers

Table 2.1 contains the list of the key literature reviewed in this chapter. The summary includes data description, the studied sector(s), the theoretical model applied, and the empirical design. Not only are the outcomes of those studies valuable to our research, but their definitions of long working hours also serve as guidance to seek for benchmarks of the work duration that can have significant impacts on mental health.

Table 2.1
Critical papers on the study of long working hours' impacts

Study	Geographical focus	Data period	Observations	Long working hours definition	Sector	Theoretical model	Design	Outcome
Angrave *et al.* (2014)	Australia, UK	1992 (UK)/2002 (Australia) - 2011	Initially: 7682 households (Australia) - 5538 (UK)	Moderately long: 40-49 hours/week Very long: ≥ 50 hours/week	various	none	FE logistic and Poisson	Working long hours reduces quitting chance, increase relapsing chance and cigarette consumption. Effects are stronger for very long working hours.
Bannai and Tamakoshi (2014)	Japan, UK, the Netherlands, Australia	1984 - 2005	19 studies in 17 articles	≥ 40 hrs/week (8hrs/day)	various	none	Systematic review	Long working hours is associated with circulatory diseases, depressive state, anxiety, sleep condition.
Burger (2015)	EU, United States, Canada	1970 - 2010	104 observations	≥ 50 hrs/week	various	none	Pooled OLS	There is divergence in the evolution of working hours in Europe. Most Western European countries show increasing ratios of extreme jobs. Extreme jobs do not necessarily lead to higher productivity.
Caruso *et al.* (2006)	The US	NA	NA	NA	NA	VA	Literature review, conference discussion	The paper proposes a conceptual framework to study the links between long working hours and health.
Collewet and Sauermann	The Netherlands	2008 - 2010	332 agents - 33,123 working days	none	Service	Cobb-Doughlas function	FE	Average handling call time increases as working time increases, the possible mechanism could be due to work fatigue.

(2017)								
Müller *et al.* (2018)	Germany	2015	13,452 observations	> 48 hours/week	various	none	Hypothesis testing, stepwise regression	Long working hours and health has negative association. Hypotheses about moderating effects of adverse working conditions on this association.
Wong *et al.* (2019)	Japan, Korea, China, US, UK, EU and New Zealand	1998 - 2018 (published time)	243 articles	> 50 hrs/week (10 hrs/day)	various	none	Meta-analysis, RE, odds ratio	Long working hours is strongly associated with cardiovascular disease and short sleep duration. The effects are also statistically significant for metabolic syndrome, fatigue, injury, poor sleep quality, and sleep disturbance.

CHAPTER 3 RESEARCH METHODOLOGY

This chapter introduces the dataset used in general, and describes the variables used, including information available in the database and data extracted from other sources. The second main section of the chapter explains the method and regression models applied in order to solve the limitations of the reviews and answer the research question.

3.1 Data

3.1.1 Data source

The dataset used is a part of The Survey of Health, Ageing and Retirement in Europe (SHARE) (Börsch-Supan, 2020a, 2020b, 2020c). The survey has been conducted since 2004 until recently, consisting of eight waves. The main participants of SHARE are individuals from 28 European countries and Israel, who are 50 years old or older and their family members living in the same household. The data collection methods vary from in-person interviews to web surveys, with in-depth questions about socio-economic status, family and social networks, employment, and the physical and mental health of the main participants.

As the study plans to employ FE regression models, following individuals over the same time span, obtaining the most balanced dataset possible is the first concern in the data processing steps. Therefore, the sample only takes data from three consecutive waves 4, 5, and 6 (conducted in 2011/2012, 2013 and 2015 respectively) within 11 countries that participate in all three waves – Austria, Germany, Sweden, Spain, France, Italy, Denmark, Switzerland, Belgium, Czech Republic, and Slovenia. Information about working hours and work status are the most important for studies about relationship between working hours and health; hence, the sample excludes observations that miss data on all the questions about the number of working hours, current job status, paid work done in the last four weeks, any paid work done ever, working sectors and the ending time of last jobs. It is likely that those surveyed participants have not worked for a long time or not worked at all in their entire life, thus not relevant in the study that seeks to explore the change within individuals in work and health. The data loss of 1,924 observations is as modest as 3% of the sample size and should not create further selection bias. There could already be a downward bias due to this selection as those who have more severe mental health problems over time could refuse to participate in the survey at the baseline or in the later waves, which lowers the external validity of the findings. This concern is raised in several studies using SHARE and other alike database (Radon, Goldberg and Becklake, 2002; Spierenburg *et al.*, 2015; Loerbroks, Karrasch and Lunau, 2017) . The final analytical sample comprises 132,321 observations from 64,098 individuals, which is unfortunately still an unbalanced dataset as there are dropping out (due to natural death or other reasons) and new participants coming in every new wave. Relatedly the number of observations also varies over different model specifications.

3.1.2 Variable description

(i) Dependent variable

The key dependent variable of the research models is a depression score based on 12 aspects of the standardized EURO-D depression scale. The scale is developed under the Concerted Action Programme of the European Commission to assess late-life depression. The 12 EURO-D items include symptoms about feeling depressed, pessimism, suicidality, guilt, sleep quality, interest, irritability, appetite, fatigue, concentration, enjoyment, and tearfulness (M. J. Prince *et al.*, 1999; M. Prince *et al.*, 1999; Guerra *et al.*, 2015). The mental health section of SHARE contains at least one question for each symptom. In order to compute the EURO-D score, we create a set of dummy variables that are coded 1 for the confirmation of the symptoms and 0 for no symptom reported. With all symptoms assigned an equal weight of 1, the EURO-D depression score is the sum of all dummy values and ranges from 0 to 12 points.

The higher the depression score, the more severe mental health problem the individual has. An EURO-D score greater than 3 could be classified as clinically significant depression or depressive disorder which should be therapeutically intervened (M. J. Prince *et al.*, 1999; M. Prince *et al.*, 1999; Guerra *et al.*, 2015). Therefore, a dummy variable indicating depression disorder is created with EURO-D score greater than 3 taking the value of 1 and 0 otherwise. To confirm the estimated effects of working hours on mental health, we replace this indicator of significant depression as the dependent variable in some estimations.

(ii) Explanatory variables

The primary explanatory variable is the number of hours the individuals work in their main job. There are 92,706 observations where their working hours are missing, which is 70% of the studied sample. By studying the survey questions that elaborate the employment history of the participants, we realize that for people who have their questions about employment status answered as currently unemployed (either retired, unemployed, permanently sick or disabled, homemaker, or for other reasons), they would not be asked about the working hours in their main job (and the first side job). Therefore, we make an assumption to replace missing values as 0 hours, which is possibly a strong assumption as people could still have a heavy amount of work without being in the official labour force. However, this assumption is not only necessary to preserve the adequate number of observations. More importantly, it could also be helpful in estimations of the effects of working hours changed for individuals who have their employed status changed from employed to retired over the study period.

To allow for possible non-linearities in the relationship between working hours and mental health, the working hours are categorized into 8 groups: not working at all (0 hours), under 10 hours, from 10 to 19 hours, from 20 to 29 hours, from 30 to 34 hours, from 35 to 39 hours, from 40 to 47 hours, from 48 to 54 hours, and over 55 hours. The categorization mostly serves the data exploration purpose and does not enter regression models.

In order to distinguish the durations of long and extremely long working hours as well as to confirm the effects of long working hours on mental health, we assign two other bivariate explanatory variables: one indicates the weekly work duration of at least 48 hours and the other indicates the minimum working time of 55 hours per week. The study decides on these two cut-off points based on the Directive of EU which limits the working time at 48 hours, and those studies of Afonso *et al.*, Milner *et al.*, Mueller *et al.*, and Virtanen *et al.* (European Parliament, 2003; Milner, Smith and LaMontagne, 2015; Afonso, Fonseca and Pires, 2017; Müller, Tisch and Wöhrmann, 2018; Virtanen *et al.*, 2018). Virtanen et al. takes the value of 55 hours as long working hours, and this study decides to set it as extremely long working hours.

The study tackles the problem of outliers in data of working hours by utilizing the information on the working hours in the side job. 414 observations that have the total weekly working hours on both jobs (their main job and the first side job if there is any) exceeding 168 hours (the total hours of seven days) are dropped from the sample. The total working hours of 168 hours are far over the 1.5 times of inter quartile range and obviously impossible, that could be due to data recording errors or some extreme cases of mental problems which make it difficult to validate other information to have meaningful regression and interpretation. Even though the working hours are still unrealistically high in a certain number of respondents (such as 160 hours), we do not make further observation removal to avoid overmanipulating the data. Thanks to the large size of the sample, those outliers should not significantly affect the data estimations.

(iii) Instrumental variables (IVs)

The study employs two IVs in the second part of the estimations, which is explained further in the next subsection 3.2 of the paper. The first IV is the information about the industry the individuals work in and already available in the survey. The second one indicates the average working hours in each of 11 studied countries, and is obtained from the labour force statistics provided by Organisation for Economic Co-operation and Development (OECD) (OECD.Stat, 2020). The statistics specifically measures the average usual weekly hours worked on the main job for the whole labour force in both gender and from the age of 15 including both dependent employment and self-employed work, either full-time or part-time (**Error! Reference source not found.**).

(iv) Control variables

The main socio-demographic potential confounders are age, gender, marital status, education attainment, financial situation. They are the most common cofounders in social studies We calculate the age based on the year of birth. 2,285 observations of the database are removed from the analytical sample because either their birth year information is missing, or the computed age is as unrealistic as 2000 years old.

As there are discrepancies in the respective educational systems, the recorded data on educational levels is ambiguous and it is impossible to obtain the information about the highest educational degree. For education attainment variable, we use the number of years in full-time education. The question is not repeated if already answered in the previous wave; therefore, information needs to be filled up with the individuals re-joined the survey in the later waves.

We use two binary variables for marital status – the Married variable is for those who are married or in an official partnership and live together with their spouse/partner, and the Widowed variable for widowers. Marriage and partnership are expected to affect mental health positively, in contrast to being widowed.

Unable to derive numerical values of individual and household's income from the recorded database, we decide to proxy the financial status with two dummy variables, one indicates whether the individual is the only income contributor in the household and the other is a subjective evaluation if the household is able to make ends meet. We expect the Only-contributor variable to have positive sign and the Make-ends-meet variable to have negative sign in our further regressions as the former is an indicator of financial concern or constraint while the latter signals a positive attitude about their household financial status.

Our study controls for work characteristics with variables indicating current job situation, the industry the individuals work in and the job sector, which is either public, private or self-employed. When industry information is not used as an IV, it is included in estimations as covariates differentiating white collar and blue-collar jobs. We categorize those who work in Agriculture, hunting, forestry, fishing/ Mining and quarrying/ Manufacturing/Electricity, gas, and water supply/ Construction/Wholesale and retail trade, repair of motor vehicles, motorcycles, personal and household goods blue collar workers. White collar workers are decided as those who are employed in finance, real estate, public administration and defence, education, health, social work, or other social service activities. White-collar jobs are expected to give rise to depression syptoms while blue-collar works could be negatively associated with depression.

We also employ a set of control variables for health risk behaviours, including smoking, drinking and exercising. For smoking, the variable reveals the individual does any kind of smoking at the time of interview. Drinking behaviours are demonstrated with two different questions – whether the respondent drinks weekly and whether the number of alcoholic drinks exceed 5 at least once a week. Another unhealthy habit is inactiveness. Those replies of hardly ever, or never doing any sports or vigorous activities are marked as those who do not exercise; and respondents who hardly ever, or never perform any activities that require a moderate level of energy are classified as inactive. We suspect that those risk factors are positively correlated with mental health problems (negative signs).

Physical health conditions that possibly affect the state of mental health is captured in our models in three dummy variables. The Long-term illness variable distinguishes individuals who have chronic or long-term physical health problem. The Poor health variable is created from a self-perceived general health question, in which the respondents rank their health from poor to excellent. The last variable in the set indicates the physical constraints, as the respondents answer the question whether they experience any limit in activities during the last six months before the survey. We also expect these variables to aversely influence mental health, especially in the older generation (positive signs).

The study also follows Artacoz *et al.* (2016) to divide the countries into regions, with the following typologies: Continental (Austria, Belgium, Germany, Switzerland, and France), Eastern European (Czech Republic, and Slovenia), Southern European (Spain and Italy) and Nordic countries (Denmark and Sweden).

As described above, Table 3.1 summarizes the full list of variables used in the analysis and the expected signs of their coefficients in estimations based on both the literature (Caruso *et al.*, 2006) and intuition. Besides the dependent and explanatory variables, a variety of control variables, there is also a set of instrumental variables that comprises data from an external source.

Table 3.1
Summary of variables used with measurement methods and their expected signs

Variable	Type	Measurement method	Expected sign
EURO-D score	Dependent variable Continuous	The sum of symptoms reported from 12 symptoms of the EURO-D scale.	NA
Significant depression	Dependent variable Dummy	Coded 1 if the responded reports at least 3 symptoms of the EURO-D scale, 0 otherwise.	NA
Number of working hours (hours)	Explanatory variable Continuous	Number of working hours in the main job	(+)
Working hour categories	Explanatory variable Categorical	The total number of working hours divided in 8 categories: not working at all (0 hours), under 10 hours, from 10 to 19 hours, from 20 to 29 hours, from 30 to 34 hours, from 35 to 39 hours, from 40 to 47 hours, from 48 to 54 hours, and over 55 hours.	(+/−)
Long working hours	Explanatory variable Dummy	Coded 1 if the total number of working hours is at least 48 hours, 0 otherwise.	(+)

Extremely long working hours	Explanatory variable Dummy	Coded 1 if the total number of working hours is at least 55 hours, 0 otherwise.	(+)
Industry	Instrumental variable Categorical	Coded 1 to 14 for different labour fields.	NA
Average hours	Instrumental variable Continuous	The average weekly hours individuals work on the main job.	NA
Age (years old)	Covariate Continuous	The year of birth subtracted from the wave year (2011, 2013, or 2015)	(+/−)
Male	Covariate Dummy	Coded 1 if the respondent is male, 0 otherwise.	(+/−)
Country typology	Covariate Categorical	Coded 1 for continental European countries, 2 for Eastern Europe, 3 for Southern Europe, and 4 for Nordic countries.	(+/−)
Years of education (years)	Covariate Continuous	The number of years in full-time education.	(−)
Married	Covariate Dummy	Coded 1 if the respondent is married or in a registered partnership and living with their partner, 0 otherwise.	(−)
Widowed	Covariate Dummy	Coded 1 if the respondent is widowed, 0 otherwise.	(+)
Only income contributor	Covariate Dummy	Coded 1 if there is no other income contributor in the household, 0 otherwise.	(+)
Able to make ends meet	Covariate Dummy	Coded 1 if the respondent thinks their household is able to make ends meet easily or fairly easily, 0 otherwise.	(−)
Blue collar worker	Covariate Dummy	Coded 1 if the respondent works/worked in agriculture, mining, fishing, manufacture, energy supply, construction or vehicle repair, 0 otherwise.	(−)
White collar worker	Covariate Dummy	Coded 1 if the respondent works/worked in finance, real estate, public administration and defence, education, health, social work, or other social service activities, 0 otherwise.	(+)

Smoke at present	Covariate Dummy	Coded 1 if the respondent does any type of smoking at the time of survey, 0 otherwise.	(+)
Drink weekly	Covariate Dummy	Coded 1 if the respondent drinks at least once per week, 0 otherwise.	(+)
Heavy drinker	Covariate Dummy	Coded 1 if the respondent has more than 5 drinks at least once a week, 0 otherwise.	(+)
No exercise	Covariate Dummy	Coded 1 if the respondent hardly ever, or never does any sports or activities that are vigorous, 0 otherwise.	(+)
Inactive	Covariate Dummy	Coded 1 if the respondent hardly ever, or never does any activities that require a moderate level of energy, 0 otherwise.	(+)
Physical limit	Covariate Dummy	Coded 1 if the respondent has been limited in activities for the past six month before the interview, 0 otherwise.	(+)
Poor health	Covariate Dummy	Coded 1 if the respondent self-ranks their general health as poor, 0 otherwise.	(+)
Long-term illness	Covariate Dummy	Coded 1 if suffers from chronic or long-term health problem, 0 otherwise.	(+)

3.2 Methodology

As discussed in the previous sections, the main challenge of this study is reverse causality from mental health to the duration of work. Besides the FE regression whose purposes is to control for time-invariant individual characteristics, the study also attempts to answer the research question with two other identification strategies, which are an IV-FE approach and the method of regression kink design (RKD). Together with the data processing and summary statistics, the regressions are performed with the well-known STATA software.

3.2.1 Fixed effects regression

In order to control bias for time-invariant unobserved factors that affect both the duration of work and mental health conditions of individuals, the study first applies the FE as one the main research method. The regression models take the following form:

$$D_{it} = \beta_0 + \beta_1 WH_{it} + \beta_{\mathrm{x}} X_{it} + \varepsilon_i + u_{it} \qquad (1),$$

where D_{it} is the variable for mental health status of individual *i* at time *t*, WH_{it} is the variable that indicates working hours, X_{it} is a set of confounders, ε_i is the individual FE, and u_{it} is the time-variant error term. The group of confounders could include age, marital status, the working industry, risky behaviours as drinking, smoking, and inactiveness, the concern about household finance, the physical ability, and self-ranked general health as described in section 3.1. With D_{it} indicating the number of depression symptoms and WH_{it} being the absolute number of working hours, the estimated results could be interpreted as the change in depression symptom scales in older people as their working hours change. We also estimate two other model specifications in which the dummy WH_{it} represents the working hours exceeding 48 hours and 55 hours. To further confirming the odds of having intervention-required depression or depressive disorder, we later apply an FE logistic regression of the dummy variable indicating more than three symptoms of late-life depression on the number of working hours, with a similar set of control variables.

We check the robustness of our findings in several steps. Firstly, pool OLS regression is applied to test for the significance of expected confounders and the signs of those confounders. Secondly, even though the study aims to control for time-invariant unobserved characteristics of the observations, the appropriateness of FE regression is also compared with random effects (RE) regression with Hausman test. For pool OLS and RE estimations, we include time-invariant control variables of gender and educational attainment. Finally, we consider the possible selection bias in the data when individuals have severe depression or other mental problems, they stop working or even drop out of the next wave's surveys. This problem of data attrition is tackled by running the same regressions on a balanced sub-sample, with the observations that appear in all three waves.

3.2.2 IV-FE regression

In an effort to address the issue of reverse causality in the relationship between the duration of work and mental health, a combination of panel data techniques and IV approach is applied in the second part of the analysis. The IV approach is a common solution for casual inference when randomized controlled experiments are impossible and there is an apparent problem of endogeneity between the dependent and independent variables. The 2SLS model is estimated in two steps with the second stage takes a similar form as in the FE regressions:

$$D_{it} = \beta_0 + \beta_1 \widehat{WH}_{it} + \beta_3 X_{it} + \varepsilon_i + u_{it} \qquad (1)$$

Before entering (1), the statutory working hours are predicted in the first stage of the regression that has the following form:

$$\widehat{WH_i} = \alpha_0 + \alpha_z Z_i + \sigma_i \qquad (2),$$

where Z_i is the set of IVs and σ_i is the error term. We instrument the number of working hours using two variables of industry indicator and national average working hours as described earlier. Strong IVs are required to be relevant – that means to be correlated with the endogenous regressors, and exogenous to the dependent variable of the second stage, which is the EURO-D score in this regression. While the average hours employers work in the studied countries are more likely regulatory and cultural and should not be directly linked with the mental health status, the exogeneity of the industry variable and depression symptoms is more complicated to be confirmed. Cragg-Donald Wald F statistic (Cragg and Donald, 1993; Stock and Yogo, 2005), Sargan statistic (Arellano and Bond, 1991), and the Anderson LM test statistic (Anderson, 1951) are calculated for the first stage regressions using the command ***xtivreg2*** in STATA to test whether the IVs produce problems of underidentification, weak identification and overidentification. Robustness check is also performed with the same balanced subsample.

3.2.3 Regression kink design

The main concern of the IV approach is the validity and strength of the IVs while there is not official tests but mostly intuition in choosing the suitable ones. On another note, the summary statistics and the regressions of EURO-D scale and indicator of significant depression on the squared variable of age (detailed explanation is in Chapter 4) suggest that there might be an association between age and depression symptoms through the amount of work in the aged population. We suspect the main cause of such impact to be the official retirement where the working hours suddenly reduce to zero. Therefore, to further explore the relationship between work and mental health outcome and overcome the problems of unobserved time-variant variables, invalid or weak IVs, and reverse causality, we attempt to apply an RKD approach.

Pioneered by Nielsen, Sørensen and Taber (2010) and since then increasingly applied in evaluating social and economic policies, RKD has its conceptual basis and implementation evolved from the regression discontinuity design (RDD) – a quasi-experimental design developed by Thistlethwaite and Campbell (1960). While RDD seeks a discontinuity of the outcome variable at a cut-off point to determine the casual effect of a policy, RKD exploit the change of slope which leads to a discontinuity in the first derivative of the assignment variable. That change of slope makes the cut-off point a "kink" point where the treatment takes place or comes into effect (Nielsen, Sørensen and Taber, 2010; Card *et al.*, 2012, 2017; Cattaneo, Idrobo and Titiunik, 2019). The most crucial assumption of RKD is there should be no kinks for other covariates around the cut-off point; hence, the presence of the kink on the assignment available can be used to confirm the treatment effect (the policy). To implement RKD, the three key factors to decide are the kernel, the bandwidth, and the polynomial (Card *et al.*, 2017; Sohn and Lee, 2019).

We employ the age as the assignment variable to examine the effect of retirement – or more specifically the sudden reduction of working hours to zero – on the EURO-D score. We choose 65 years old as the first cut-off point as it is the most common regulatory retirement age in Europe and repeat the estimate with 67 years old as the cut-off point since we expect a delayed effect of retirement on the late-life depression. After observing the presence of a kink at the chose cut-off point, we also check for kinks on other covariates. The RKD applied in the study is not a sharp but fuzzy RKD design as the official retirement age is not universal among European nations, and early retirement is also an available option in many cases. The polynomial is set at 1 as a local linear regression, while the kernel and bandwidth are chosen automatically by the STATA software's command ***rdplot*** to produce a visualization of the estimated relationship between retirement (age) and depression symptoms.

CHAPTER 4 RESEARCH RESULTS & DISCUSSION

This chapter presents the descriptive statistics of variables used in the sample and the regression estimations. As the research question is whether or not long working hours impact the mental health outcome, the magnitude of the effects (if any) is not the focus of the below discussion.

4.1 Descriptive statistics

Table 4.1 shows the frequencies and percentages of 12 EURO-D depression symptoms reported in the sample. It is observable that feeling sad or depressed, having trouble with sleep, and feeling tired are the most three common symptoms. As mentioned in the variable description, significant depression with more than three symptoms is likely to be a depressive disorder that might need clinical diagnostic and intervention, which accounts for 29.42% of the sample (
Appendix 2).

Table 4.1
12 EURO-D scale symptoms distribution

Symptom	Frequency	Percentage
Sad/Depressed feelings	51,266	39.16
No hopes for future	19,880	15.19
Suicidal feelings	9,448	7.22
Guilt	29,033	22.39
Trouble sleeping	44,887	34.28
Loss of interest	9,933	7.59
Irritability	36,790	28.10
Diminution in appetite	10,817	8.26
Fatigue	45,152	34.48
Difficulty in concentrating on reading	16,884	12.90
No enjoyment	15,117	11.55
Tearfulness	30,827	23.55

Table 4.2 and Table 4.3 display the statistics of working hours in categories and the EURO-D scores according to those categories. 70% of the study sample is not employed and has the working hours recorded as zero. The second largest category is the group that works 40 to 47 hours weekly, which is higher than the average hours worked in most countries in the official reports of OECD statistics (OECD. Stat, 2020) (**Error! Reference source not found.**). While the proportion of observations with working hours higher than 55 hours is quite small at 2% of the sample, the

figure is even lower for people who only work less than 10 hours a week (Table 4.2). The summary statistics of EURO-D depression score do not show much discrepancy among the categories in all measures (Table 4.3). There are observations without any symptoms of depression and those with serious clinical depression (self-reported to have 10-12 symptoms) in every categories of working hours. Nonetheless, we notice that observations that work at least 40 hours weekly have the smaller mean values of the depression score, as well as smaller standard deviations. The group that has highest mean and highest standard deviation of the EURO-D score is actually the group that does not work at all.

Table 4.2.
Weekly working hours in categories

Weekly working hours	Frequency	Percent
Not working	93029	70.31
Under 10	2458	1.86
10 to 19 hours	3023	2.28
20 to 29 hours	4354	3.29
30 to 34 hours	3180	2.40
35 to 39 hours	7290	5.51
40 to 47 hours	13291	10.04
48 to 54 hours	3031	2.29
55 hours and above	2665	2.01
Total	**132321**	**100.00**

Table 4.3
EURO-D score statistics by working hours

Working hours in categories	N	Mean	Sd	Min	Max
Not working	89872	2.636	2.328	0	12
Under 10	2430	2.084	1.899	0	10
10 to 19 hours	2996	2.45	2.125	0	12
20 to 29 hours	4304	2.327	2.047	0	12
30 to 34 hours	3144	2.184	1.996	0	11
35 to 39 hours	7210	2.092	2.006	0	11
40 to 47 hours	13090	1.788	1.832	0	11
48 to 54 hours	2981	1.843	1.854	0	12
55 hours and above	2612	1.969	1.956	0	11

The discrepancy of working hours and clinical depression in European regions are shown in Table 4.4 and Table 4.5. The Southern Europe with two countries – Spain and Italy – has the largest proportion of the population that is unemployed (79.2%), while this figure of the Nordic countries – Denmark and Sweden – is the smallest (59%). The percentage of working hours that area round the published average trend (35 to 39 hours) is also the largest in Nordic region at 9.9%. While older people in Continental countries – Austria, Belgium, Germany, France, and Switzerland – tend to be more familiar with extra hours than other areas with 2.5% working 48-54 hours and 2.9% working at least 55 hours, their average working hours are the lowest with 37.4 hours per week. Czech Republic and Slovenia – the two countries in the East of Europe – have the highest average weekly working hours at 42.1 hours, which is significant longer than the other three regions where the weekly averages are all under 40 hours. Looking at the distribution of observations with more than three symptoms in the EURO-D scale, we observe that the region with the lowest proportion of unemployed individuals – the Nordic countries – also has the smaller proportion of individuals suffering from clinical depression. Conversely, Continental Europe has the shortest weekly work duration at 37.4 hours but the highest proportion of people in significant depression. According to Artazcoz *et al.* (2016), the different patterns in the relationship between working hours and health status in those European regions are due to a combination of cultural dynamics as well as economic vulnerability and labour market deregulation.

Table 4.4
Categories of working hours in different parts of Europe

Weekly working hours	Country Typologies				
	Continental Europe	Eastern Europe	Southern Europe	Nordic countries	Total
Not working	68.18	76.06	79.16	59.01	70.31
Under 10	2.21	1.38	0.59	2.64	1.82
10 to 19 hours	3.13	1.15	0.87	2.32	2.22
20 to 29 hours	4.12	1.67	1.81	3.58	3.15
30 to 34 hours	2.66	0.73	1.33	4.49	2.33
35 to 39 hours	5.99	1.53	3.95	9.87	5.36
40 to 47 hours	8.32	13.08	8.62	13.02	9.99
48 to 54 hours	2.45	2.51	2.11	2.63	2.43
55 hours and above	2.94	1.89	1.55	2.44	2.41
Total	100.00	100.00	100.00	100.00	100.00

Table 4.5
Statistics of working hours in different parts of Europe in the currently-employed subsample

Country typologies	Summary statistic of working hours	Significant

	N	Mean	SD	depression
Continental Europe	17335	37.409	14.409	24.7%
Nordic countries	7049	38.712	10.978	16.9%
Eastern Europe	4664	42.086	11.127	18.9%
Southern Europe	4637	38.4	12.513	21.3%

Table 4.6 displays the mean and standard deviation of the variables used in the research regression models for the entire sample and a subsample of individuals that are still in the labour force. The main rationale of examining the subsample is to observe the average number of working hours, which is not a meaningful measurement for the whole sample as more than 70% does not in employment and has working hours as zero. The study also performs summary statistics separately for people with and without significant depression. For dummy variables, the mean value shows the percentage of observations with the assigned characteristics.

Table 4.6
Mean and standard deviation of the research variables

	Entire sample (N = 132,321)				**Currently-employed subsample (N = 33,685)**			
	No significant depression		**Significant depression**		**No significant depression**		**Significant depression**	
	Mean	**Sd**	**Mean**	**Sd**	**Mean**	**Sd**	**Mean**	**Sd**
Number of working hours	11.761	18.638	7.441	15.592	38.086	12.319	35.944	13.249
Long working hours	.055	.227	.033	.179	.185	.388	.167	.373
Extreme long working hours	.027	.162	.017	.131	.091	.287	.088	.283
Age	66.503	9.492	68.524	10.911	57.371	4.592	56.823	4.493
Male	.494	.5	.332	.471	.527	.499	.349	.477
Years of education	11.141	4.406	10.296	4.339	12.453	4.431	12.27	4.318
Private sector	.188	.391	.129	.335	.574	.495	.581	.493
Public sector	.082	.274	.058	.233	.258	.437	.277	.448
Self-employed	.063	.242	.036	.187	.168	.374	.141	.348
Only contributor	.811	.392	.813	.39	.77	.421	.767	.423
Make ends meet	.733	.442	.535	.499	.798	.401	.635	.481
Smoke at present	.24	.427	.251	.434	.319	.466	.36	.48
Heavy drinker	.073	.26	.071	.257	.077	.267	.07	.255
Drink weekly	.542	.498	.408	.492	.604	.489	.544	.498
No exercise	.382	.486	.587	.492	.23	.421	.297	.457
Inactive	.077	.267	.233	.423	.036	.185	.063	.243
Physical limit	.364	.481	.667	.471	.225	.417	.43	.495
Long-term illness	.422	.494	.669	.47	.309	.462	.486	.5

Poor health	.208	.406	.362	.481	.105	.307	.264	.441

We observe similar patterns in both sets of samples (Table 4.6). People who do not have more than three symptoms of depression tend to work longer hours than those with significant depression. For people who are still currently in the labour force, the mean difference is about 2 hours weekly (38.09 hours – no significant depression compared to 35.94 hours – with significant depression) while the difference is larger at 4.3 hours weekly for the entire sample (11.76 hours – no significant depression compared to 7.44 hours – with significant depression). Similarly, the percentages of people working at least 48 hours and 55 hours per week are higher in the subsets with no significant depression than in those with significant depression. Gender also appears to possess a discrepancy in the matter of mental health. The groups without significant depression have quite a balanced gender distribution, while in the groups that suffer from serious late-life depression, men only take up a third of the samples (33.2% and 34.9%). For the level of education, the mean and standard deviation values are slightly lower for the groups with significant depression, indicating that people with higher educational attainment are more likely to have better mental health. However, the differences are barely noticeable and unlikely to be significant. As mentioned above, the data of educational level is obtained from the question of the number of years staying in full-time education. The effects of education could be estimated more precisely with information about the highest degree achieved, which is ambiguous in the data release. The statistic measures of variable indicating sectors the individuals work in also have vague differences among the groups, but it seems that people who are self-employed is less likely to have serious depression compared to people who work in public and private sectors. For the currently-employed subsample, 58.1% of people report more than three symptoms of depression work in private sector, 27.7% work in public sector and only 14.1% of them are self-employed.

From the same table, we also observe that the majority of the respondents are the only income contributors in their households and while the intuition is that the situation might create stress and have negative effect on their mental health, their statistics do not present significant difference across the studied groups. About 81% of the entire sample and 77% of the currently-employed subsample state that they are there is no other income contributor in their family, regardless whether they report the signs of clinical depression. On the other hand, the other variable used to indicate the financial situation tells another story. For observations who do not show evidence of significant depression, more than 70% of them think their household can make ends meet easily or fairly easily (73.3% the entire sample and 79.8% in the sub-sample), while the figure is less than 70% for those who suffer from more symptoms of depression (53.5% in the entire sample and 63.5% in the sub-sample) (Table 4.4). Therefore, if long working hours could later be proved to adversely affect the mental health in the old population, a plausible mechanism might be through the income concern. As the individuals have the pressure to sustain their living, they are more prone to being depressed and willing to work longer hours.

Table 4.6 also reflects observable heterogeneity in the summary statistics of health-related risky behaviours on mental health. While the percentage of respondents who still smoke at the survey periods is only 1% different in the full dataset between those have low and high risk of depression (24% and 25%), the difference is larger in the sub-sample of the currently-employed (32% and 36%). People who are still in employment are more likely to smoke and working people who have significant depression are more likely to smoke than who do not. For alcohol consumption, the picture is not as clear. People who suffer from serious depression seem to drink less frequently than mentally healthy people and people who are still at work seem to drink more often compared to the whole sample. However, the proportion of heavy drinkers who have six or more alcoholic beverages at once at least three or four days a week is almost identical among the groups, the figure is approximately 70% across the samples. The two variables indicating the inactiveness presents more synchronized results, as the percentage of people who rarely engage in any of sports or activity that require vigorous amount are significant higher in the groups with significant depression.

On the other hand, the set of physical ability and general health variables shows consistent results that the people who report to have conditions that limit their physicality, chronic or long-term illness or self-evaluate to have poor general health are much more likely to have at least three symptoms in the EURO-D scale. However, this percentage is lower for the group of people who are still in employment. For example, for observations that do not have enough symptoms to be diagnosed as being clinically depressive in the whole sample, 21% of them report to have poor general health while that percentage is 36% for people who have significant depression. In the currently-employed subsample, those figures stay lower at 11% and 26%, respectively (Table 4.4). In general, people who stay in employment seem to have better general health and less chronic illness than the whole population.

Among the control variables, age appears to have the most sophisticated relationship with depression symptoms. The age of observations ranges from 50 - the age eligible for the SHARE survey - to 106, the group of 75 years old or older makes up the largest part of the sample (24.14%) (Appendix 3). From Table 4.6, it is noticeable that while the mean age for people with significant depression is higher than that of those without in the entire sample (68.5 years old compared to 66.5 years old), the situation is the opposite in the sub-sample of employed people but the difference in this group is rather small and likely insignificant (57.4 years old compared to 56.8 years old). Table 4.7 shows the summary statistics of EURO-D score by age. While the scores all range from 0 to 12 points, and their mean and standard deviation values are not noticeably different across the age categories, both values seem to decrease from younger to older groups until 65 years old and increase along the remaining groups. This observation suggests a possible nonlinear association between age and mental health, leading to our decision but add the squared value of age into our regression models.

Table 4.7
EURO-D score summary statistics by age

Age Category	N	Mean	Sd	Min	Max
Under 55	13525	2.467	2.281	0	12
55 to 59	20682	2.374	2.202	0	12
60 to 64	23390	2.245	2.14	0	12
65 to 69	22686	2.205	2.12	0	12
70 to 74	18524	2.375	2.171	0	12
75 or older	30157	2.887	2.382	0	12

Table 4.8 illustrates a further look into the relationship between working hours, age and late-life depression. People who work extremely long hours are present in all age categories, there are individuals who are 75 years old or older and still work a maximum of 84 hours weekly. The mean values and standard deviation of working hours decreases as people move to older age groups, the largest jump is the reduction of 10 hours in the mean value and 8 hours in the standard deviation from the group of 60-64 years old to the group of 65-69 years old. The statistics of the variable indicating possible significant depression shows similar pattern to the variable of the total EURO-D score – decreasing along the younger groups and increasing after 65-69 years old.

Table 4.8
Working hours and significant depression by age

Age categories	Summary statistics of working hours					Significant depression
	N	Mean	Sd	Min	Max	Percent
Under 55	13673	29.134	19.004	0	160	29.5%
55 to 59	20930	25.752	20.317	0	152	27.7%
60 to 64	23712	13.763	19.387	0	110	25.3%
65 to 69	23103	3.739	11.272	0	112	24.6%
70 to 74	18958	1.355	6.67	0	105	27.6%
75 or older	31945	0.385	3.419	0	84	38.2%

4.2 Fixed effects regression results

4.2.1 Effects of working hours on mental health

Table 4.9 shows the regression results of the total EU-D depression score on the absolute number of weekly working hours across different model specifications. The first three columns present the results of pool OLS estimations for the whole sample, columns (4) and (5) are of RE, and the last two columns are of FE estimations. The coefficients of the working hours are all negative and significant across all the models. The negative sign indicates that time spent on work actually has positive effects on the mental health of people over 50 years old. Conditional on the covariates, results of the Pool OLS estimations in column (1) and (2) show that each extra working hour reduces the EURO-D score by 0.033 and 0.039 point, respectively. The same model specification of column (2) estimated with RE regression yields a similar result with the coefficient of working hours being -0.037 and standard error being 0.006 at 0.1% level of significance (column (5)). The magnitude of the effect is smaller at 0.018 and significant at 5% level with higher standard error in the FE estimations. The result of Hausman tests (Appendix 4) applied for estimations shown in column (5) and (6) proves that the FE estimation is more appropriate and there is presence of unobserved within-individual effects which the RE regression is not able to account for.

By including the variable indicating interaction of working hours and age, we can confirm that the association of working hours and mental health depends greatly on the age of individuals. While the magnitude of the interaction effect is extremely small, the coefficient is statistically significant with very low standard error in all regressions. Furthermore, the effect of working hours on the depression score decreases significantly to -0.006 in the estimations that exclude the interaction variable (Column (3) and (4)). In addition, the statistical significance of the squared variable of age confirms that the association between age and depression in the older population is nonlinear.

The majority of other covariates are highly significant at 0.1% level with expected signs in all model specifications. The Pool OLS and RE estimations show that men score significantly lower than women on the EURO-D scale and the longer people stay in full-time education, the lower they score on the EURO-D scale at the age of 50 or older. Being married or staying with a partner associates with healthier mental status while being widowed associates with higher depression scores. Individuals that have less concern about their household finance have lower depression scores than individuals that have difficulty in making ends meet. People who have physical problems that limit their activities severely tend to have higher scores than people who are not limited or limited but not severely. Chronic or long-term illness and self-evaluated poor general health also reversely affect the mental health as the presence of them significantly raises the score of depression. Among the health-related risky behaviours, while there is limited evidence in these estimations that drinking heavily or smoking can significantly influence the mental health, a lifestyle without exercise or sports or any kinds of vigorous physical activities consistently increases the EURO-D depression score in the old age.

However, inconsistency still presents across the models on the covariate that represents white-collar workers. While Pool OLS and RE results show that white-collar jobs increase the depression level with the coefficient being positive and statistically significant at 0.1% level (except for the RE estimation without the interaction variable (column (4)), FE estimations produce statistically insignificant and negative coefficient.

Table 4.9
Estimation Results of EU-D score on the number of working hours

	(1) Pool OLS	(2) Pool OLS	(3) Pool OSL	(4) RE	(5) RE	(6) FE	(7) FE
No. of working hours	-0.033***	-0.039***	-0.006***	-0.006***	-0.037***	-0.018*	-0.018*
	(0.006)	(0.007)	(0.001)	(0.001)	(0.006)	(0.009)	(0.009)
Working hours * Age	0.0004***	0.001***			0.001***	0.0003*	0.0003*
	(0.000)	(0.0001)			(0.0001)	(0.0001)	(0.0001)
Age	-0.204***	-0.215***	-0.199***	-0.194***	-0.261***	-0.322***	-0.349***
	(0.013)	(0.019)	(0.011)	(0.018)	(0.014)	(0.030)	(0.030)
Age^2	0.001***	0.001***	0.001***	0.001***	0.002***	0.003***	0.002***
	(0.000)	(0.0001)	(0.000)	(0.0001)	(0.0001)	(0.0002)	(0.0002)

Male	-0.595***	-0.550***	-0.594***	-0.614***	-0.622***		
	(0.018)	(0.025)	(0.018)	(0.028)	(0.021)		
Schooling years	-0.013***	-0.002	-0.017***	-0.010***	-0.022***		
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)		
Married		-0.231***	-0.235***	-0.235***	-0.290***	-0.251***	
		(0.025)	(0.018)	(0.029)	(0.021)	(0.064)	
Widowed	0.250***						0.517***
	(0.025)						(0.070)
White collar	0.048***	0.084***	0.132***	0.050	0.110***	-0.114	-0.120
	(0.025)	(0.034)	(0.029)	(0.040)	(0.029)	(0.069)	(0.069)
Heavy drinker		0.080					
		(0.046)					
Smoke at present				0.068*			
				(0.030)			
No exercise	0.296***	0.355***	0.415***	0.294***	0.339***	0.102***	0.102***
	(0.019)	(0.026)	(0.018)	(0.028)	(0.018)	(0.020)	(0.019)
Inactive	0.744***			0.648***			
	(0.030)			(0.046)			
Physical limit	0.755***	0.660***	0.800***	0.683***	0.660***	0.305***	0.306***
	(0.022)	(0.030)	(0.022)	(0.033)	(0.021)	(0.022)	(0.021)
Poor health	0.302***	0.453***	0.289***	0.302***	0.262***	0.141***	0.142***
	(0.021)	(0.031)	(0.021)	(0.032)	(0.020)	(0.021)	(0.021)
Long-term illness	0.514***	0.451***	0.525***	0.506***	0.489***	0.253***	0.254***
	(0.021)	(0.029)	(0.021)	(0.031)	(0.020)	(0.021)	(0.021)
Able to make ends meet	-0.691***	-0.618***	-0.710***	-0.656***	-0.578***	-0.167***	-0.167***
	(0.019)	(0.027)	(0.019)	(0.028)	(0.020)	(0.023)	(0.023)
Constant	9.854***	10.210***	9.695***	9.578***	12.007***	14.777***	14.328***
	(0.451)	(0.657)	(0.389)	(0.621)	(0.500)	(1.053)	(1.050)
N	57106	26722	57107	24962	57107	89205	89205
R^2	0.205	0.167	0.197	0.203	0.196	0.122	0.122

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

To endow additional credibility to our findings, we also apply a logistic regression with the independent dummy variable indicating whether the individuals are suffering from over three depression symptoms. The logistic results with Pool OLS, RE, and FE estimations are presented in Table 4.10. The number of observations is significantly smaller for the FE estimation as 41,201 groups were dropped out because those individuals either stayed mentally healthy or clinically depressed in the whole research period. While the explanatory variable of working hours has negative signs in all regressions, it is highly significant at 0.1% level in the Pool OLS and RE estimations but only significant at 10% level in the FE estimations. The FE regression also produce coefficients with remarkably higher standard errors. However, the negative signs of their coefficients confirm the findings in the previous model specifications that as the number of working hours increases, the observations in our sample tend to have lower risk of clinical depression or depressive disorder.

Table 4.10
Logit regression results of significant depression on the number of working hours

	(1) Pool OLS	(2) RE	(3) FE
No. of working hours	-0.020***	-0.037***	-0.038⁺
	(0.006)	(0.009)	(0.020)
Working hours*Age	0.0002**	0.0004**	0.0006*
	(0.000)	(0.0001)	(0.0003)
Country typologies	-0.035***	-0.052***	
	(0.007)	(0.012)	
Age	-0.202***	-0.332***	-0.541***
	(0.011)	(0.018)	(0.057)
Age^2	0.001***	0.002***	0.004***
	(0.000)	(0.0001)	(0.0004)
Male	-0.618***	-0.910***	
	(0.017)	(0.028)	
Married	-0.150***	-0.234***	-0.297*
	(0.017)	(0.027)	(0.117)
No exercise	0.292***	0.387***	0.117**
	(0.018)	(0.025)	(0.038)
Inactive	0.581***	0.774***	0.276***
	(0.024)	(0.036)	(0.053)
Physical limit	0.661***	0.865***	0.368***
	(0.020)	(0.029)	(0.041)
Poor health	0.306***	0.408***	0.155***
	(0.018)	(0.026)	(0.037)
Long-term illness	0.424***	0.592***	0.334***
	(0.019)	(0.028)	(0.040)
Able to make ends meet	-0.610***	-0.780***	-0.170***
	(0.017)	(0.025)	(0.040)
Constant	6.439***	10.226***	
	(0.396)	(0.637)	
lnsig2u			
_cons		1.091***	
		(0.030)	
N	90919	90919	21323

Standard errors in parentheses

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

4.2.2 Effects of long working hours and extremely long working hours

In order to confirm the effects of long working hours on the mental health of the older population and specify the duration of long working hours and extremely long working hours, the study replaces the numerical variable of working hours by two dummy variables representing people who work at least 48 hours and those who work at least 55 hours (Table 4.11). The coefficients of the main explanatory variables have positive signs in all regression models, indicating that working longer than the regulatory limit of 48 hours associates with depressive problems in the elderly. The FE estimation results show that working 48 hours a week or longer raises the depression score by 0.17 compared to people who work shorter hours, holding other covariates constant. The coefficient is statistically significant at 1% level. For other covariates, the signs of their coefficients stay consistent with previous estimations. The Hausman test still confirms that FE models are more appropriate, once again suggesting that individual-specific unobservables may bias simple associations and RE results. However, the significantly small value of R-squared suggests that there is a consistent problem of omitted variables across all the models, including the FE regressions.

Table 4.11
Regression results of the EURO-D score on long working hours and extremely long working hours

	Working at least 48 hours			Working at least 55 hours		
	(1) Pool OLS	(2) RE	(3) FE	(4) Pool OLS	(5) RE	(6) FE
Working hours dummy	0.173***	0.177***	0.170**	0.244***	0.232***	0.163*
	(0.038)	(0.038)	(0.057)	(0.050)	(0.049)	(0.072)
Working hours*Age	-0.00008***	-0.00009***	-0.0001	-0.00008***	-0.00009***	-4.36e-06
	(9.63e-06)	(0.00001)	(0.0001)	(9.12e-06)	(9.59e-06)	(0.00001)
Country typologies	-0.035***	-0.035***		-0.035***	-0.035***	
	(0.006)	(0.007)		(0.006)	(0.007)	
Age	-0.191***	-0.210***	-0.336***	-0.191***	-0.211***	-0.338***
	(0.009)	(0.010)	(0.028)	(0.009)	(0.010)	(0.028)
Age2	0.001***	0.002***	0.002***	0.001***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male	-0.655***	-0.687***		-0.654***	-0.686***	
	(0.014)	(0.017)		(0.014)	(0.017)	
Married	-0.156***	-0.194***	-0.253***	-0.156***	-0.194***	-0.253***
	(0.014)	(0.017)	(0.064)	(0.014)	(0.017)	(0.064)
No exercise	0.427***	0.347***	0.102***	0.427***	0.347***	0.102***

	(0.015)	(0.014)	(0.020)	(0.015)	(0.014)	(0.020)
Physical limit	0.807***	0.663***	0.306***	0.807***	0.663***	0.306***
	(0.017)	(0.016)	(0.022)	(0.017)	(0.016)	(0.022)
Poor health	0.319***	0.279***	0.141***	0.318***	0.279***	0.141***
	(0.017)	(0.016)	(0.021)	(0.017)	(0.016)	(0.021)
Long-term illness	0.476***	0.446***	0.253***	0.476***	0.445***	0.253***
	(0.017)	(0.016)	(0.021)	(0.017)	(0.016)	(0.021)
Able to make ends meet	-0.756***	-0.609***	-0.167***	-0.756***	-0.610***	-0.167***
	(0.015)	(0.015)	(0.023)	(0.015)	(0.015)	(0.023)
Constant	9.239***	9.879***	13.831***	9.263***	9.903***	13.896***
	(0.308)	(0.346)	(0.953)	(0.308)	(0.346)	(0.953)
N	89205	89205	89205	89205	89205	89205
R^2	0.192	0.190	0.022	0.192	0.190	0.022

Standard errors in parentheses

* p < 0.05, ** p < 0.01, *** p < 0.001

Table 4.12 shows the results of the main FE regressions but on the balanced subsample of individuals who appear in all three waves of the panel as a test for model robustness. From left to right, the columns display the estimates of the EURO-D score on the number of working hours, the logistic estimates of the binary variable of being clinically depressed, the estimates of the depression level on the dummy variable indicating weekly working duration of at least 48 hours and of at least 55 hours. The coefficients are statistically significant at 5% and 10% respectively and have the same signs of those in estimations on the full sample. As shown in column (1) and (2), the effect magnitudes are even larger with lower standard errors compared to the same model specifications applied on the full panel (Column (7) of Table 4.9 and column (3) of Table 4.10). Column (3) shows that working at least 48 hours a week increases the depression score by 0.14 point and working extremely long above 55 hours increases the score by 0.16 point comparing to people who follow the regulatory hour limit. In general, the results of estimations on the balanced subsample confirms our above findings that there is a positive link between work and the mental health of people aged 50 and older, the longer the individuals work, the less likely they suffer from serious depression that require clinical intervention. Working too long, however, could adversely affect the mental health, as working at least 48 hours tends to increase the EURO-D depression score, and working at least 55 hours increases the score in larger magnitude.

Table 4.12
FE estimation results on the balanced subsample

	(1) Number of working	(2) FE-logit	(3) Long working hours	(4) Extremely long

	hours			working hours
Explanatory variable of working hours	-0.022*	-0.055*	0.139*	0.157+
	(0.011)	(0.024)	(0.069)	(0.088)
Working hours*Age	0.0004*	0.001*	-1.59e-06	3.67e-06
	(0.0002)	(0.0004)	(0.00001)	(0.00001)
Age	-0.330***	-0.488***	-0.301***	-0.302***
	(0.035)	(0.066)	(0.032)	(0.032)
Age^2	0.002***	0.003***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Married	-0.203**	-0.284*	-0.205**	-0.205**
	(0.074)	(0.139)	(0.074)	(0.074)
No exercise	0.091***	0.100*	0.090***	0.090***
	(0.023)	(0.044)	(0.023)	(0.023)
Physical limit	0.302***	0.373***	0.304***	0.304***
	(0.025)	(0.048)	(0.025)	(0.025)
Poor health	0.139***	0.154***	0.139***	0.139***
	(0.025)	(0.043)	(0.025)	(0.025)
Long-term illness	0.229***	0.286***	0.229***	0.229***
	(0.025)	(0.047)	(0.025)	(0.025)
Able to make ends meet	-0.146***	-0.139**	-0.146***	-0.147***
	(0.027)	(0.049)	(0.027)	(0.027)
_cons	13.788***		12.687***	12.741***
	(1.220)		(1.098)	(1.098)
N	49761	15681	49761	49761
R^2	0.020		0.020	0.020

Standard errors in parentheses

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.3 IV-FE regression results

The instrumental variable estimations are performed under the assumption that the working industry only affects the mental health via its influence on the number of working hours. In other words, the IVs employed need to be independent of other unobservable characteristics. This is a very strong assumption as the industry a person works in can affect their health in other channels, such as the working environment, the working hazards that are specific to such industry. The results of Cragg-Donald Wald F statistic (Cragg and Donald, 1993; Stock and Yogo, 2005), Sargan statistic (Arellano and Bond, 1991) and the Anderson LM test statistic (Anderson, 1951) computed in the first stage regression show that the two IVs are suitable to use in our models (Appendix 5).

Table 4.13 displays the results of IV-FE regressions employing either the country-specific average working hours or the industry or both variables as IVs. The first three columns are results of estimations with the absolute number of working hours, columns (4) and (5) are results of regressions on the dummy indicating long working hours from 48 hours, and the last two columns are on the extremely long hours from 55 hours. In general, the estimates from the IV models are consistent in sign with those of the previous models, indicating an overall positive effect of work on the mental health of the older population in Europe, and a negative effect of long and extremely long working hours on the likelihood of significant depression. Compared to the normal FE models with similar specifications, all IV-FE estimations produce coefficients of the interested explanatory variables with higher magnitudes. Each extra hours of work decrease the EURO-D score by 0.2 point. For people who work 48 hours or longer, their depression scores tend to increase by 4.5, indicating serious depression that may require clinical intervention. For people who work at least 55 hours, their depression scores are likely to be greater than 5.

Table 4.13
IV-FE regression results

	Number of working hours			**Long working hours**		**Extreme long hours**	
IV(s)	**(1) Country average hours**	**(2) Industry**	**(3) 2 IVs**	**(4) Industry**	**(5) 2 IVs**	**(6) Industry**	**(7) 2 IVs**
Effect of working hours on health	0.304	-0.226***	-0.218***	4.469***	4.487***	5.767***	5.789***
	(0.463)	(0.051)	(0.051)	(1.073)	(1.071)	(1.391)	(1.387)
Working hours * Age	-0.005	0.004***	0.004***	-0.001***	-0.001***	-0.0004***	-0.0004***
	(0.008)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.0001)
Age	0.067	-0.638***	-0.628***	-0.299***	-0.299***	-0.349***	-0.349***
	(0.616)	(0.073)	(0.073)	(0.031)	(0.031)	(0.030)	(0.030)
Age^2	-0.000	0.004***	0.004***	0.002***	0.002***	0.002***	0.002***
	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Married	-0.246***	-0.259***	-0.259***	-0.235***	-0.235***	-0.238***	-0.237***
	(0.066)	(0.065)	(0.065)	(0.069)	(0.069)	(0.069)	(0.069)
No exercise	0.108***	0.098***	0.098***	0.097***	0.096***	0.096***	0.096***
	(0.022)	(0.020)	(0.020)	(0.021)	(0.021)	(0.021)	(0.021)
Physical limit	0.315***	0.299***	0.300***	0.307***	0.307***	0.315***	0.315***
	(0.026)	(0.022)	(0.022)	(0.023)	(0.023)	(0.023)	(0.023)
Poor health	0.140***	0.142***	0.142***	0.136***	0.135***	0.143***	0.143***
	(0.021)	(0.021)	(0.021)	(0.023)	(0.023)	(0.023)	(0.023)
Long-term illness	0.260***	0.248***	0.248***	0.244***	0.244***	0.243***	0.243***
	(0.024)	(0.021)	(0.021)	(0.023)	(0.023)	(0.023)	(0.023)

Able to make ends meet	-0.174***	-0.163***	-0.163***	-0.149***	-0.149***	-0.153***	-0.153***
	(0.025)	(0.023)	(0.023)	(0.025)	(0.025)	(0.025)	(0.025)
Constant	-1.216	25.071***	24.694***	12.635***	12.630***	14.499***	14.501***
	(22.978)	(2.707)	(2.686)	(1.061)	(1.061)	(1.034)	(1.034)
N	89205	89205	89205	89205	89205	89205	89205
R^2	0.048	0.090	0.092	0.034	0.033	0.036	0.035

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.14 displays the estimates from the same model specifications using both IVs on the subsample of observations that participate in all three survey waves as a test for attrition bias. The sizes of the sample are reduced by half, but the signs of all coefficients remain consistent, including those of the covariates. The effects of working hours on the depression level have even higher magnitudes in all estimations compared to those on the entire panel.

Table 4.14
IV-FE regression results on the balanced subsample

	Number of working hours	Long working hours	Extreme long working hours
Effect of working hours	-0.2649***	4.9674***	6.4949***
	(0.071)	(1.359)	(1.784)
Working hours * Age	0.0043***	-0.0005***	-0.0004***
	(0.001)	(0.000)	(0.000)
Age	-0.6511***	-0.2653***	-0.3191***
	(0.099)	(0.036)	(0.035)
Age^2	0.0044***	0.0019***	0.0022***
	(0.001)	(0.000)	(0.000)
Married	-0.2077**	-0.1611*	-0.1435
	(0.074)	(0.081)	(0.082)
No exercise	0.0895***	0.0813***	0.0841***
	(0.023)	(0.025)	(0.025)
Physical limit	0.2948***	0.3094***	0.3116***
	(0.025)	(0.027)	(0.027)
Poor health	0.1398***	0.1323***	0.1336***
	(0.025)	(0.027)	(0.027)
Long-term illness	0.2233***	0.2168***	0.2208***
	(0.025)	(0.027)	(0.027)
Make ends meet	-0.1411***	-0.1256***	-0.1351***
	(0.027)	(0.030)	(0.029)
_cons	25.7318***	11.5089***	13.5317***
	(3.645)	(1.231)	(1.211)
N	49761	49761	49761

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.4 RKD results

Figure 4.1 and 4.2 are the graphic presentations of the RKD estimates. Even though 65 years old appears to be where the retirement takes place, the change of slope or the kink appears visually at 67 years old. The number of depression symptoms individuals suffer from and age has a negative association until they reach 67 years old and a positive relationship with greater magnitude after that. It is consistent with the mentioned descriptive statistics that the EURO-D score tends to decrease along the younger age groups, from 50-54 years old to 65-69 years old and increase greatly among the older ones (Table 4.8). The method is proved appropriate as there are no similar kinks found at the cut-off point of 67 years old on other covariates. The RKD plots for the relationship between retirement age and health-risk behaviours, self-reported general health, and financial concern can be found in Appendix 6.

Existing literature on retirement and mental health draws different conclusions on this relationship. Many studies do not achieve inclusive findings about a signifcant correlation between retirement and mental health (Ross and Drentea, 1998; Drentea, 2002). Using data from the first wave of SHARE, Coe and Zamarro (2011) find evidence of a positive association between retirement and depression but cannot confirm any casual mechanism. The study of Vo and Tran (2019) reports the most similar findings to our study with the same panel of SHARE. They also find a continuous reduction of EURO-D score of newly retirees in the first two years and an increase back to the same level before retirement after that two-year period. They refer such change as an Ashenfelter's dip and argue that it exists due to the complete adaptation of individuals to their new non-working life which diminishes the positive effects of retirement on the mental health. Conversely, the study of Kim and Feldman (2000) and Dingemans and Henkens (2019) support the the ideas that retirement leads to a decline in mental health and life satisfaction at the old age. According to Atchley's continuity theory of aging (Atchley, 1989), older individuals psychologically desire to maintain their daily routines, life structure, social contact. For those who identify themselves through their career achievements, continuing work at the old age is even more crucial for life satisfaction. The observable gradual decrease of the depression score until the age of 67 of our RKD estimations is in line with those studies that confirm the adverse effects of retirement on mental health. The retirees struggle to adjust their usually routines and life structure of daily work to a new life of zero working hours, such struggle does not have immediate impact but gradually create distress and other symptoms of depression which might need therapeutic and clinical intervention.

Figure 4.1
RKD estimate of EURO-D sore on age – Cut-off point is 65 years old

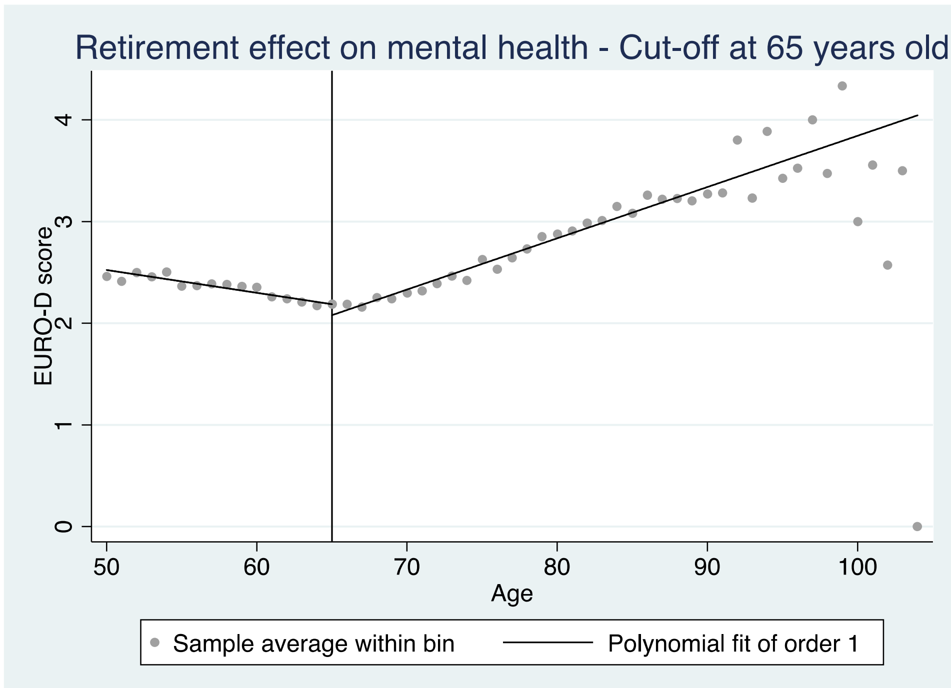
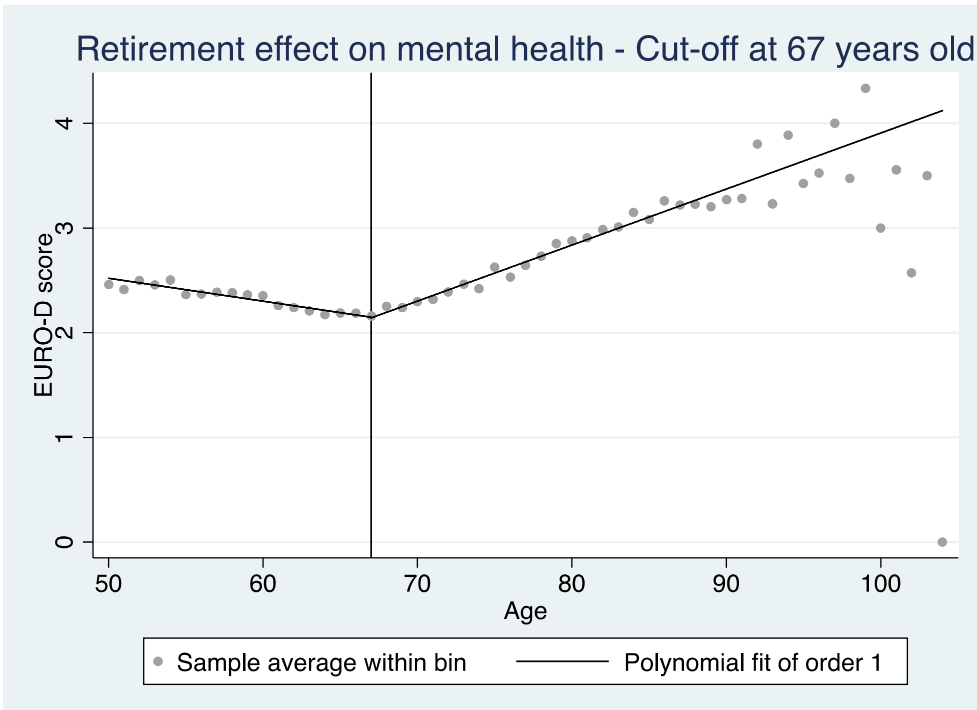


Figure 4.2
RKD estimate of EURO-D sore on age – Cut-off point is 67 years old



CHAPTER 5 CONCLUSION

"There is no health without mental health."(European Commision, 2005, p. 4). Without a baseline of mental wellness, human beings are incapable in pursuing endeavours that lead to the well-being in other spectrums. Using data from SHARE – the Survey of Health, Ageing and Retirement in Europe, the research applies three different approaches to investigate the association between long working hours and mental health conditions of Europeans who are 50 years or older, based on the 12 items of the EURO-D depression scale. The study answers the call of Caruso *et al.* (2006) for study of the relationship between working hours and health outcomes in a specifically high-risk groups – the older generation.

While not all model specifications produce significant results and there is heterogeneity in the effect magnitudes between estimations, the findings show evidence that it is more beneficial for the older population to work moderately than too little work or not work at all. The longer hours people work, the lower their mental health scores on the EURO-D scale. However, working longer than the weekly regulatory limit which is 48 hours could adversely affect the mental health, increasing risk of clinical depression. Those who work extremely long hours, over 55 hours per week, face even greater risk.

Another important finding is that the impact of working hours on the mental health status of older people depends significantly on their age. The likelihood of suffering from depression symptoms declines over time until the age of 67 and significantly increases after that, which also gives hints of a negative effect of retirement on the mental health of the elderly.

Despite its existing limitations of omitted variables, the strength of the IVs and a more proper implementation of the RKD approach, the study still sheds some light on labour policy and legislation. The EU has been done impressive work on the enforcement of working time limits, while still need to achieve other essential objectives such as increasing employment rate, sustaining productivity, or providing equal employment opportunities. With the understanding that the older population is fully capable of contributing, labour laws and policy might be better catered for this group, making working at old age more desirable and productive.

References

Afonso, P., Fonseca, M. and Pires, J. F. (2017) 'Impact of working hours on sleep and mental health', *Occupational Medicine*, 67(5), pp. 377–382. doi: 10.1093/occmed/kqx054.

Anderson, T. W. (1951) 'Estimating Linear Restrictions on Regression Coefficients for Multivariate Normal Distributions', *The Annals of Mathematical Statistics*, 22(3), pp. 327–351.

Angrave, D., Charlwood, A. and Wooden, M. (2014) 'Working time and cigarette smoking: Evidence from Australia and the United Kingdom', *Social Science & Medicine*, 112, pp. 72–79. doi: 10.1016/j.socscimed.2014.04.031.

Arellano, M. and Bond, S. (1991) 'Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations', *The Review of Economic Studies*, 58(2), pp. 277–297. doi: 10.2307/2297968.

Artazcoz, L. *et al.* (2016) 'Long working hours and health in Europe: Gender and welfare state differences in a context of economic crisis', *Health & Place*, 40, pp. 161–168. doi: 10.1016/j.healthplace.2016.06.004.

Atchley, R. C. (1989) 'A Continuity Theory of Normal Aging', *The Gerontologist*, 29(2), pp. 183–190. doi: 10.1093/geront/29.2.183.

Bannai, A. and Tamakoshi, A. (2014) 'The association between long working hours and health: A systematic review of epidemiological evidence', *Scandinavian Journal of Work, Environment & Health*, 40(1), pp. 5–18. doi: 10.5271/sjweh.3388.

Becker, G. S. (1965) 'A Theory of the Allocation of Time', *The Economic Journal*, 75(299), pp. 493–517. doi: 10.2307/2228949.

Börsch-Supan, A. (2020a) 'Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 4'. SHARE–ERIC. doi: 10.6103/SHARE.W4.710.

Börsch-Supan, A. (2020b) 'Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 5'. SHARE–ERIC. doi: 10.6103/SHARE.W5.710.

Börsch-Supan, A. (2020c) 'Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 6'. SHARE–ERIC. doi: 10.6103/SHARE.W6.710.

Burger, A. S. (2015) 'Extreme Working Hours in Western Europe and North America: A New Aspect of Polarization', *SSRN Electronic Journal*. doi: 10.2139/ssrn.2609012.

Card, D. *et al.* (2012) *Nonlinear Policy Rules and the Identification and Estimation of Causal Effects in a Generalized Regression Kink Design*. w18564. National Bureau of Economic Research. doi: 10.3386/w18564.

Card, D. *et al.* (2017) *Regression Kink Design: Theory and Practice*, *Regression Discontinuity Designs*. Emerald Publishing Limited. doi: 10.1108/S0731-905320170000038016.

Caruso, C. C. *et al.* (2006) 'Long working hours, safety, and health: Toward a national research agenda', *American Journal of Industrial Medicine*, 49(11), pp. 930–942. doi: 10.1002/ajim.20373.

Cattaneo, M. D., Idrobo, N. and Titiunik, R. (2019) *A Practical Introduction to Regression Discontinuity Designs: Foundations*. 1st edn. Cambridge University Press. doi: 10.1017/9781108684606.

Coe, N. B. and Zamarro, G. (2011) 'Retirement effects on health in Europe', *Journal of Health Economics*, 30(1), pp. 77–86. doi: 10.1016/j.jhealeco.2010.11.002.

Collewet, M. and Sauermann, J. (2017) 'Working hours and productivity', *Labour Economics*, 47, pp. 96–106. doi: 10.1016/j.labeco.2017.03.006.

Cragg, J. G. and Donald, S. G. (1993) 'Testing Identifiability and Specification in Instrumental Variable Models', *Econometric Theory*, 9(2), pp. 222–240.

Cronin, C., Forsstrom, M. and Papageorge, N. (2017) *Mental Health, Human Capital and Labor Market Outcomes*, *CINCH Working Paper Series*. 1704. Universitaet Duisburg-Essen, Competent in Competition and Health. Available at: http://ideas.repec.org/p/duh/wpaper/1704.html (Accessed: 19 July 2020).

Dembe, A. E. (2005) 'The impact of overtime and long work hours on occupational injuries and illnesses: new evidence from the United States', *Occupational and Environmental Medicine*, 62(9), pp. 588–597. doi: 10.1136/oem.2004.016667.

Dingemans, E. and Henkens, K. (2019) 'Working After Retirement and Life Satisfaction: Cross-National Comparative Research in Europe', *Research on Aging*, 41(7), pp. 648–669. doi: 10.1177/0164027519830610.

Drentea, P. (2002) 'Retirement and Mental Health', *Journal of Aging and Health*, 14(2), pp. 167–194. doi: 10.1177/089826430201400201.

European Commision (2005) *Green Paper - Improving the mental health of the population - Towards a strategy on mental health for the European Union*. Brussels: Publications Office of the EU. Available at: https://op.europa.eu/en/publication-detail/-/publication/5420b88c-690b-47df-8c89-c0d7112134c3/language-en (Accessed: 21 November 2020).

European Commision (2012) *White Paper - An Agenda for Adequate, Safe and Sustainable Pensions*. COM(2012) 55 final. Brussels.

European Parliament (2003) 'Directive 2003/88/EC of the European Parliament and of the Council of 4 November 2003 concerning certain aspects of the organisation of working time'.

Eurostat (2020) *Labour market and Labour force survey (LFS) statistics*, *Eurostat*. Available at: https://ec.europa.eu/eurostat/statistics-explained/index.php/Labour_market_and_Labour_force_survey_(LFS)_statistics.

Finnish Centre for Pensions (2020) *Retirement Ages*, *Finnish Centre for Pensions*. Available at: https://www.etk.fi/en/work-and-pensions-abroad/international-comparisons/retirement-ages/#:~:text=In%20the%20EU%20Member%20States,being%20linked%20to%20life%20expectancy. (Accessed: 18 July 2020).

Grossman, M. (1972) 'On the Concept of Health Capital and the Demand for Health', *Journal of Political Economy*, 80(2), pp. 223–255.

Grossman, M. (2000) 'Chapter 7 The human capital model', in *Handbook of Health Economics*. Elsevier, pp. 347–408. doi: 10.1016/S1574-0064(00)80166-3.

Guerra, M. *et al.* (2015) 'Psychometric properties of EURO-D, a geriatric depression scale: a cross-cultural validation study', *BMC Psychiatry*, 15. doi: 10.1186/s12888-015-0390-4.

Hanna, A. S., Taylor, C. S. and Sullivan, K. T. (2005) 'Impact of Extended Overtime on Construction Labor Productivity', *Journal of Construction Engineering and Management*, 131(6), pp. 734–739. doi: 10.1061/(ASCE)0733-9364(2005)131:6(734).

Hanna Awad S., Taylor Craig S., and Sullivan Kenneth T. (2005) 'Impact of Extended Overtime on Construction Labor Productivity', *Journal of Construction Engineering and Management*, 131(6), pp. 734–739. doi: 10.1061/(ASCE)0733-9364(2005)131:6(734).

Holden, L. *et al.* (2010) 'Work Performance Decrements Are Associated With Australian Working Conditions, Particularly the Demand to Work Longer Hours':, *Journal of*

Occupational and Environmental Medicine, 52(3), pp. 281–290. doi: 10.1097/JOM.0b013e3181d1cdbb.

Keller, S. M., Berryman, P. and Lukes, E. (2009) 'Effects of Extended Work Shifts and Shift Work on Patient Safety, Productivity, and Employee Health', *AAOHN Journal*, 57(12), pp. 497–504. doi: 10.1177/216507990905701204.

Khwaja, A. (2010) 'Estimating willingness to pay for medicare using a dynamic life-cycle model of demand for health insurance', *Journal of Econometrics*, 156(1), pp. 130–147. doi: 10.1016/j.jeconom.2009.09.011.

Kim, S. and Feldman, D. C. (2000) 'Working in Retirement: The Antecedents of Bridge Employment and Its Consequences for Quality of Life in Retirement', *The Academy of Management Journal*, 43(6), pp. 1195–1210. doi: 10.2307/1556345.

Kim, W. *et al.* (2016) 'Effect of working hours and precarious employment on depressive symptoms in South Korean employees: A longitudinal study', *Occupational and Environmental Medicine*, 73, p. oemed-2016. doi: 10.1136/oemed-2016-103553.

Loerbroks, A., Karrasch, S. and Lunau, T. (2017) 'The longitudinal relationship of work stress with peak expiratory flow: a cohort study', *International Archives of Occupational and Environmental Health*, 90(7), pp. 695–701. doi: 10.1007/s00420-017-1232-0.

Milner, A., Smith, P. and LaMontagne, A. (2015) 'Working hours and mental health in Australia: Evidence from an Australian population-based cohort, 2001-2012', *Occupational and environmental medicine*, 72. doi: 10.1136/oemed-2014-102791.

Mullan, J., Vargas Llave, O. and Wilkens, M. (2017) *Working conditions of workers in different ages: European Working Conditions Survey 2015*. Luxembourg: Publications Office of the European Union (EF, 17/47).

Müller, G., Tisch, A. and Wöhrmann, A. M. (2018) 'The impact of long working hours on the health of German employees', *German Journal of Human Resource Management: Zeitschrift für Personalforschung*, 32(3–4), pp. 217–235. doi: 10.1177/2397002218786020.

Nielsen, H. S., Sørensen, T. and Taber, C. (2010) 'Estimating the Effect of Student Aid on College Enrollment: Evidence from a Government Grant Policy Reform', *American Economic Journal: Economic Policy*, 2(2), pp. 185–215. doi: 10.1257/pol.2.2.185.

OECD. Stat (2020) *Level of GDP per capita and productivity : Average hours worked per person employed*, *OECD. Stat*. Available at: https://stats.oecd.org/Index.aspx?QueryId=95895# (Accessed: 19 November 2020).

OECD.Stat (2020) *Average usual weekly hours worked on the main job*, *OECD.Stat*. Available at: https://stats.oecd.org/ (Accessed: 1 July 2020).

Papageorge, N. W. (2016) 'Why medical innovation is valuable: Health, human capital, and the labor market: Why medical innovation is valuable', *Quantitative Economics*, 7(3), pp. 671–725. doi: 10.3982/QE459.

Prince, M. *et al.* (1999) 'Development of the EURO-D Scale–A European, Union Initiative to Compare Symptoms of Depression in 14 European Centres', *The British journal of psychiatry : the journal of mental science*, 174, pp. 330–8. doi: 10.1192/bjp.174.4.330.

Prince, M. J. *et al.* (1999) 'Depression symptoms in late life assessed using the EURO–D scale: Effect of age, gender and marital status in 14 European centres', *The British Journal of Psychiatry; London*, 174(4), pp. 339–345. doi: http://dx.doi.org.eur.idm.oclc.org/10.1192/bjp.174.4.339.

Radon, K., Goldberg, M. and Becklake, M. (2002) 'Healthy worker effect in cohort studies on chronic bronchitis', *Scandinavian Journal of Work, Environment & Health*, 28(5), pp. 328–332. doi: 10.5271/sjweh.682.

Ross, C. E. and Drentea, P. (1998) 'Consequences of Retirement Activities for Distress and the Sense of Personal Control', *Journal of Health and Social Behavior*, 39(4), pp. 317–334. doi: 10.2307/2676341.

Shepard, E. and Clifton, T. (2000) 'Are longer hours reducing productivity in manufacturing?', *International Journal of Manpower*, 21(7), pp. 540–553. doi: 10.1108/01437720010378999.

Shimizu, T. *et al.* (2004) 'Relationship between self-reported low productivity and overtime working', *Occupational Medicine*, 54(1), pp. 52–54. doi: 10.1093/occmed/kqg099.

Sohn, H. and Lee, S.-W. (2019) 'Causal Impact of Having a College Degree on Women's Fertility: Evidence From Regression Kink Designs', *Demography*, 56(3), pp. 969–990. doi: 10.1007/s13524-019-00771-9.

Spierenburg, E. A. J. *et al.* (2015) 'Healthy worker survivor analysis in an occupational cohort study of Dutch agricultural workers', *International Archives of Occupational and Environmental Health*, 88(8), pp. 1165–1173. doi: 10.1007/s00420-015-1047-9.

Stock, J. H. and Yogo, M. (2005) *Testing for Weak Instruments in Linear IV Regression*. t0284. National Bureau of Economic Research. doi: 10.3386/t0284.

Thistlethwaite, D. L. and Campbell, D. T. (1960) 'Regression-discontinuity analysis: An alternative to the ex post facto experiment', *Journal of Educational Psychology*, 51(6), pp. 309–317. doi: 10.1037/h0044319.

Virtanen, M. *et al.* (2018) 'Long working hours and depressive symptoms: systematic review and meta-analysis of published studies and unpublished individual participant data', *Scandinavian Journal of Work, Environment & Health*, 44(3), pp. 239–250. doi: 10.5271/sjweh.3712.

Vo, T. T. and Tran, D. T. P. (2019) *Mental Health Around Retirement: Evidence of Ashenfelter's Dip Based on SHARE Panel Data, 2004−2017*. SSRN Scholarly Paper ID 3357265. Rochester, NY: Social Science Research Network. doi: 10.2139/ssrn.3357265.

Wong, K., Chan, A. H. S. and Ngan, S. C. (2019) 'The Effect of Long Working Hours and Overtime on Occupational Health: A Meta-Analysis of Evidence from 1998 to 2018', *International Journal of Environmental Research and Public Health*, 16(12), p. 2102. doi: 10.3390/ijerph16122102.

World Bank (2019) *Population ages 65 and above (% of total population)*, *World Bank*. Available at: https://data.worldbank.org/.

Appendices

Appendix 1
Average working hours in 11 studied countries (OECD.Stat, 2020)

Country typology	Country	Wave	Average hours worked weekly (hours)
Continental	Austria	4	36.5
		5	36.1
		6	35.7
	Germany	4	34.6
		5	34.4
		6	34.4
	Belgium	4	35.1
		5	35.2
		6	35.1
	France	4	36.6
		5	36.2
		6	36.0
	Switzerland	4	34.7
		5	34.5
		6	34.5
Eastern	Czech Republic	4	39.9
		5	39.7
		6	39.5
	Slovenia	4	39.2
		5	39.4
		6	39.2
Southern	Italy	4	36.0
		5	35.5
		6	35.5
	Spain	4	37.2
		5	36.5
		6	36.6
Nordic	Sweden	4	35.8
		5	35.8
		6	35.8
	Denmark	4	32.6
		5	32.6
		6	32.5

Appendix 2
Frequency and percentage of observations with significant depression

	Freq.	Percent	Cum.
No significant depression	93386	70.58	70.58
Significant depression	38935	29.42	100.00
Total	132321	100.00	

Appendix 3
Age in categories

Age category	Freq.	Percent	Cum.
Under 55	13673	10.33	10.33
55 to 59	20930	15.82	26.15
60 to 64	23712	17.92	44.07
65 to 69	23103	17.46	61.53
70 to 74	18958	14.33	75.86
75 or older	31945	24.14	100.00
Total	132321	100.00	

Appendix 4
Hausman test results of FE (Column 6) and RE (Column 5) estimations shown on Table 4.9

```
                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     est17        est18        Difference          S.E.
-------------+----------------------------------------------------------------
no_working~s |   -.0177821    -.0374517        .0196696        .0067353
noworkingh~e |    .0003006     .0005442       -.0002436        .0001066
         age |   -.3617798    -.2614431       -.1003366        .0266696
        age2 |    .0025865     .0018344        .0007521        .0001922
     married |     -.25069    -.2902769        .0395869        .0604994
 whitecollar |   -.1135304    -.0101328       -.1033976        .0629065
  noexercise |    .1016937     .3388926       -.2371989        .0075595
physicalli~t |     .305288     .6590801       -.3537922        .0063436
  poorhealth |    .1414551     .2618683       -.1204132        .0061389
longtermil~s |    .2525144     .4889031       -.2363888        .0071101
makeendsmeet |   -.1670269    -.5781139         .411087        .0128849
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
            B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(11) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                           =     4604.49
                 Prob>chi2 =      0.0000
```

Appendix 5
IVs test results (Column (3), Table 4.13)

```
Underidentification test (Anderson canon. corr. LM statistic):         1291.665
                                                 Chi-sq(2) P-val =       0.0000
-------------------------------------------------------------------------------
Weak identification test (Cragg-Donald Wald F statistic):               667.057
Stock-Yogo weak ID test critical values: 10% maximal IV size              19.93
                                         15% maximal IV size              11.59
                                         20% maximal IV size               8.75
                                         25% maximal IV size               7.25
Source: Stock-Yogo (2005).  Reproduced by permission.
-------------------------------------------------------------------------------
Sargan statistic (overidentification test of all instruments):            1.158
                                                 Chi-sq(1) P-val =       0.2818
```

Appendix 6
RKD plots to check for kinks on covariates around the cut-off point