LUND UNIVERSITY
School of Economics and Management

THE EFFECT OF OVERWEIGHT AND OBESITY ON EARLY RETIREMENT IN EUROPE

Department of Economics
Author: Malin Bengtsson
Civilekonomprogrammet
Supervisor: Margareta Dackehag
August 2014

ABSTRACT

The aim of this paper is to examine the impact of overweight and obesity on early retirement. Data is gathered from the Survey of Health, Ageing and Retirement in Europe (SHARE) database and on this data a linear probability model is applied, using both a Pooled Ordinary Least Squares (POLS) and a Fixed Effect (FE) strategy. In addition, separate regressions are run for current and lagged Body Mass Index (BMI). The results indicate that being obese has a positive effect on the probability of early retirement for men, and lagged BMI has a greater effect than current. For females, no significant relationship between overweight or obesity on early retirement could be found. When it comes to health indicators connected to overweight/obesity, no significant relationship was found between these and early retirement among males, while the female sample shows ambiguous effect regarding self-rated health and early retirement. Including health indicators have mixed effects on the magnitude of the weight coefficients. Judging from our results, being obese per se affects early retirement among males, and this effect do not seem to run through different health indicators.

Key words: SHARE, early retirement, obesity, fixed effects, POLS

TABLE OF CONTENTS

1. INTRODUCTION

All across Europe, there is a growing concern regarding early exit from paid employment. Exiting employment before statutory retirement age is not only a risk on a personal financial and social level (Gallo, 2000), but it is also a societal concern. With increasing life expectancy and decreasing numbers of births, troubles with financing of the welfare state may arise if actions aimed to keep people at work are not completed (Ilmarinen, 2001). In order to make these actions effective, more studies need to be conducted on the subject of what make people leave paid employment in the first place, which is where this paper wishes to contribute.

Commonly, determinants of retirement are divided into four categories consisting of institutional factors, financial factors, personal employment experience and work attitudes and family characteristics (Hochman & Lewin-Epstein, 2013). Institutional factors cover mandatory or statutory retirement age, while financial factors are dependent on employment, as well as pension, payments. Personal employment experience and work attitudes cover job factors such as stress, self-fulfillment, job control, etc and family characteristics include marital status and personal health. Obesity and overweight could affect one or more of these factors, resulting in leaving paid employment earlier than expected. Previous research have time after time concluded that overweight and obesity is costly from a societal perspective. Müller-Riemenschneider et al (2008) estimated the relative economic burden of obesity to range from 0.09% to 0.61% of gross domestic product (GDP) in the 10 European countries they studied. To put this number in perspective, one can compare it to the estimated cost of cancer in the European Union which is 1.07% of GDP (Luengo-Fernandez et al, 2013). Obesity have been found to increase morbidity (Allison et al, 1999), increase absenteeism from work (Bungum, 2003) and predict unemployment (Garcia and Quintana-Domeque, 2006). Obesity does also significantly lower the probability of labor market participation amongst individuals (Lundborg et al, 2007). Both Renna & Thakur (2010) and Huston et al (2009) have previously found a significant relationship between obesity and early retirement. These significant results has however only been established using American data, while significance among European data is lacking.

On the basis of overweight affecting productivity, this paper aims to evaluate the impact of overweight and obesity on early retirement, mostly through their effect on health. The hypothesis is that being considered overweight or obese will increase the probability of early retirement. There have been some previous studies (Renna & Thakur, 2010; Robroek et al, 2013; Huston et al, 2009; Friis et al, 2007) evaluating this relationship, but they have not completely covered the area. First of all, no significant relationship has been found between early retirement and overweight/obesity when using European data. Furthermore, some of the methods used have been found to not control properly for unobserved heterogeneity (Renna & Thakur, 2010) and significant results are also lacking.

In order to evaluate how weight affects early retirement, data have been collected from The Survey of Health, Ageing and Retirement in Europe (SHARE) database. This dataset consists of four different waves containing over 60 000 individuals from several different European countries. In this paper, data from wave 2 and 4 (complemented with some variables from wave 1) will be used.

Following previous studies, three dummy variables based on Body Mass Index (BMI) were created, in order to reflect whether the individual is underweight, overweight or obese. Early retirement was measured as a binary variable, indicating early retirement or not. Using Pooled Ordinary Least Squares (POLS) and fixed effects (FE) strategies constructed with current and lagged BMI, this allows us to control for time-fixed unobserved heterogeneity, as well as reversed causality. Utilizing a linear probability model, our results shows that obesity affects the probability of early retirement positively for males, while overweight show few signs of having any significant effect. Lagged BMI has a larger effect on early retirement than current in the male sample. For females, no significant relationship between overweight or obesity on early retirement could be found. When it comes to health indicators connected to overweight/obesity, no significant relationship was found between these and early retirement in the male sample, while self-rated bad health seem to affect the probability of early retirement for women negatively in some cases. Including health variables in the estimations had ambiguous effects on the size of the weight coefficients.

The limitations of this study include many missing observation reducing the statistical power, potential missing covariates, self-reported values introducing potential bias and some of the variables' definition. Also, the methods used relies on fairly strong assumptions.

The paper is organized as follows: we start with introducing some previous research on the area, then we continue on to presenting our data and variables, followed by a section discussing our method of choice. After this our results are presented, followed by a discussion part and finally a conclusion.

2. THEORY AND PREVIOUS RESEARCH

In this section, a general literature review will be given. The first part will give some explanation as to how overweight and obesity potentially affects early retirement. The second part explains some findings in studies performed on weight and its effect on labor market productivity. The reason to why this is important in this study is because parts of the methods later discussed and used are taken from these studies. The third part will immerse further and review the research done on weight and early retirement.

2.1. Theoretical underpinnings

In previous literature, some theories regarding the connection between overweight and labor market participation have been presented. Baum and Ford (2004) divide these theories into internal and external factors. The internal factors cause the individual to be less productive because the overweight affects personal health, limiting possible work assignments. Overweight workers may also place higher value on current consumption (assuming they have myopic preferences), implying lower concern about future health and therefore higher risk of obesity. This may also suggest lower incentives to participate in work training, limiting work advancement. External factors include employers discriminating towards overweight individuals. This may be due to the employer's belief that overweight workers are less productive (Asgeirsdottir, 2011). According to Baum & Ford (2004), there might also be customer discrimination (depending on occupation) towards obese individuals, creating lower productivity because customers are unwilling to interact with overweight individuals.

In the case of early retirement, these theories suggest that overweight could affect one or more of the four factors of retirement (especially financial factors, personal employment experience and work attitudes and family characteristics). The financial factor is to be affected if overweight has a negative effect on wage, the personal employment experience and work attitudes factor would play a part if overweight, for example, causes dissatisfaction with current job and the family characteristics factor will be affected if overweight lowers individual health. Houston et al (2009) suggests that obese individuals may leave the labor market both due to lower achieved characteristics, as well as the negative health effects of obesity. Renna & Thakur (2010) also proposes that labor market discrimination towards obese individuals may cause them to leave the labor market early. In this paper, we focus mostly on how overweight affects early retirement through its effect on health, since the health data in SHARE is fairly complete compared to other factors of early retirement.

2.2. Weight, productivity and causality

During the course of the years, many studies have been performed on the subject of weight and different kinds of labor market outcomes. In this section we only wish to consider a few of these studies, namely those that use relevant methods for our upcoming estimations.

Several studies find a wage penalty for obese individuals (Baum and Frod, 2004; Register and Williams, 1990; Cawley et al., 2005). Often, the main issue is how to establish causality. Cawley (2004) examines this closer, looking at wage penalties for overweight individuals and utilizing different strategies. He points out that weight status is not exogenous, since there are both genetic and non-genetic (upbringing, culture, etc) unobserved factors that affect weight. Also, wages and personal characteristics will have an impact on weight, implying a reversed causality problem. To solve these potential issues he suggests three strategies. The first one is to replace weight with a lagged value and assume that the correlation between lagged weight and the wage residual is zero (see also Conley and Glauber, 2007). This would solve the problem of current wage status effect on current weight, but ignores the fact that lagged genetic and non-genetic factors can be correlated with current wages. The second strategy applies a fixed effect (FE) approach by using the difference of current and previous weight (see also Averett & Korenman, 1996). This will eliminate the fixed genetic and non-genetic unobserved factors, however reversed causality can still be an issue. In the last strategy, Cawley uses a sibling's BMI as an instrument (correlated with weight but uncorrelated with the other regressors) in order to remove endogeneity (see also d'Hombres & Brunello, 2005). Using these three different strategies, Cawley finds that weight lowers wage for white females. Instrumental variable regressions have also been performed by Lundborg et al (2007) when examining the relationship between obesity and the probability of being employed. They use SHARE data from wave 1 but since there is no sibling data in SHARE, Lundborg et al uses instruments that indicate whether or not the respondent was the oldest child, whether the respondent had only sisters and whether anyone else in the household is obese. In their estimations, they find that being obese is associated with a lower probability of being employed for both men and women (-0.05 and -0.10).

2.3. Weight and early retirement

When looking at studies investigating the connection between weight and early retirement, it can be concluded that there has been less done in this area compared to other areas evaluating BMI and productivity.

Robroek et al (2013) use the first three waves of the same dataset that will be used in this study (the SHARE-dataset). Their aim is to evaluate how poor health, unhealthy behaviors, and unfavorable work characteristics affects exit from paid employment due to disability pension, unemployment, and early retirement among older workers. To model this, they use a Cox proportional hazard model. The basic concept of this model in current context is that employed respondents are followed from the first visit to the time when they retire early or have their last follow-up visit (whichever happens first). For all individuals that left for early retirement during the period, this event was compared with the people staying in the workforce until the end of the follow-up or until censored. Individuals that reach the end of the studied period (alternatively the statutory retirement age) are censored. The results are then presented by using a hazard ratio (HR). For example, if the HR of obesity is 2, this means that receiving the treatment (obesity) results in twice the risk of an event occurring (retiring early), proportionally to the control group. The findings indicate that individuals with poor health, lack of physical activity and obesity have a higher risk of leaving employment for disability pension, but there was no significant relation between any of these factors and early retirement.

Staying in Europe, Friis et al (2007) have a sample of nurses, collected from the Danish Integrated Database for Labor Market Research combined with an additional survey. They study these nurses during the period from 1993-2002 and wish to analyze the relationship between sociodemographic, work-related, health and lifestyle factors (including indicators for overweight and obesity) and the early retirement scheme Post Employment-Wage (PEW). To model this, they use discrete time survival analysis, complemented with a log-log link function to model the relationship between explanatory variables and joining the PEW. Friis et al (2007) included four variables measuring lifestyle (smoking, alcohol consumption, leisure-time physical activity and BMI), but found no association between these and early retirement status. However, they did find a connection between low self-rated health and joining the PEW (HR 1.28).

In the UK, Mein et al (2000) studies predictors of early retirement on British civil servants. They collected the data from the Whitehall II cohort of male and female civil servants, and performed a 7-year follow-up study between 1988 to 1995. They did not include a control for weight, however they did control for other health indicators. The results imply that individuals are more likely to retire early if they have bad or very bad perceived health (HR 1.63 for men and 1.20 for women) or if they have a long-term illness (HR 1.08 for men and 1.25 for women).

Moving on to the United States, Renna & Thakur (2010) study the impact of obesity on labor market outcomes of working age adults collected from US Health and Retirement Study. The labor market outcomes they choose to look at are "working", "not working due to a disability" and "not working due to early retirement" and they use two estimation strategies in order to conclude if obesity affects labor market outcome direct (via body impairments) or indirect (via medical conditions closely connected to obesity). The first strategy models employment status in 2002 on weight status in 1992. The second one studies the impact of obesity on body impairments and chronic diseases, followed by regressing the impairments and diseases on the different labor market outcomes. The results from the first strategy show significant impact of obesity on disability, but not on early retirement. However, once adding the controls for chronic diseases and body impairments, they get a significant effect of having suffered from a heart attack on early retirement. They interpret this as obese individuals being less prone to retire early, unless their weight considerably affects their health. In the second strategy, they get significant results in both regressions, which they state suggests that there is a causal relationship between obesity and labor market outcomes. In order to extract the magnitude of this relationship, they multiply the effect of obesity on the medical conditions by the effect of the medical conditions on labor market outcome. The results from these calculations are that obesity of class 2 and 3 (BMI>35) increases the probability of early retirement by 1.5 percentage points for men and by 2.5 percentage points for women. The probability of disability pension increases with 1.5 percentage points for men and 1.7 percentage points for women.

Another study performed on American data was completed by Houston et al (2009) and is presented using Cox proportional hazard model, in conformity with Robroek et al (2013). The sample (consisting of white and African-American individuals) where collected from the Atherosclerosis Risk in Community (ARIC) study. They examined associations between weight status at age 25, 45-55 and age at early retirement over a 9-year follow-up study. Being overweight or obese at 25 was significantly related to early retirement in all groups but white women (HR 1.62, 1.32 and 1.43 for African-American women, white men and African-American men). Obesity at 45-55 was only significantly associated with early retirement among white men (HR 1.23 for overweight and 1.32 for obese).

To conclude this section, it is reasonable to state that the main issue among previous studies has been to establish significant results. Another problem is the lack of controls for unobserved heterogeneity for several of the previous studies. If no attention is given to this problem, one cannot be certain that the results generated are actually causal (since there can be other unobserved factors driving the effect). Renna & Thakur (2010), however, are able to control for time-fixed unobserved heterogeneity by using a fixed effect estimator. We will, in this paper, contribute to the current research by presenting estimates focusing solely on early retirement. To our knowledge, no estimations in this field have been performed on SHARE data from wave 4.

Also, we will utilize the fixed effect approach in combination with lagged BMI (and compare with current BMI estimations), which has only been done on American data so far. This is done in order to control for unobserved fixed heterogeneity and reversed causality, which could otherwise bias our estimations.

3. DATA

The data used in this study is collected from The Survey of Health, Ageing and Retirement in Europe (SHARE) database. It is a cross-sectional, multidisciplinary database containing micro-data from more than 60.000 individuals aged 50 years or older and their partners in several European countries. The baseline study (wave 1) was conducted in 2004 and included 11 European countries (Denmark, Sweden, Austria, France, Germany, Switzerland, Belgium, Netherlands, Spain, Italy and Greece). After that, three different waves have been performed and more countries have been added. Wave 2 was collected in 2006-2007, wave 3 in 2008-2009 (and focused more on detailed retrospective life histories of the respondents, which means it will be ignored in this study) and wave 4 in 2010-2011. The SHARE database collects information about demographics, physical and mental health, employment and pensions, behavioral risks, cognitive function, children, social support, health care, financial transfers, housing, household income, consumption, assets, activities and expectations. The data used in this study originates from wave 2 and 4, with some variables collected from wave 1 (more information on this in the following sections). The countries that partook in all SHARE waves (and will therefore be included in this study) are all the ones mentioned above, except for Greece.

3.1. Sample

Since this study examines how weight affects early retirement, the sample needed to be trimmed in a step towards extracting causal results. The individual was asked to pick what best described their current employment situation out of the following: retired, employed or self-employed (including working for family business), unemployed, permanently sick or disabled, homemaker or other. For weight and early retirement, only individuals who stated that they were employed or self-employed in wave 1 were included in this study (9070 individuals), eliminating 21 746 individuals[1]. By doing this we eliminate individuals who were permanently sick, homemakers, etc., in the beginning of the study. We are aware of the fact that there are other pathways into

[1] We made sure to control our data and confirmed that the vast majority of the SHARE participants ending up in early retirement originated from the group that were employed in wave 1. Thereby we do not exclude an important group when looking at early retirement and the sample becomes more heterogeneous.

early retirement that we eliminate by doing this. An individual may for example go from employment (earlier than in wave 1), to unemployment and then end up as early retired. However, we think one of the most relevant societal aims should be to keep the employed in the labor force until reaching the statutory retirement age, and therefore this is the focus of our study. Then, we needed to make sure that all individuals included were participating in all three waves, leaving us with 3335 individuals. Two individuals were identified with the same id-number and some lacked health indicators and when we excluded them our final sample consisted of 1631 females and 1667 males. Since we want to control for income (where many values are missing) we loose individuals around 500 individuals for both sexes. Because lagged BMI is included as an explanatory variable in certain parts of this study, we could only run our regressions on observations from wave 2 and 4 (since lagged data on BMI is not available for wave 1). Finally, we are down to a sample of 2580 observations for males and 2520 observations for females, where all countries mentioned above are included.

3.2. Dependent variable

Since the main goal in this thesis is to find the impact of overweight and obesity on early retirement, the individual can either be early retired or not. To illustrate this, a binary outcome variable was constructed, taking on the value of one if the individual in question did retire early, and zero otherwise.

Following the same definition as Robroek et al (2013) and Renna & Thakur (2010), early retirement status was extracted in two steps. First, we took the retirement year stated in the SHARE questionnaire and from that the retirement age was calculated by subtracting the year of birth (if the individual is not retired, there is no retirement year). When doing this, no attention was given to which month the individual was born and/or retired, which might create some minor inconsistencies. Secondly, this age was compared to the statutory retirement age of the country where the individual is currently living (see Table 1 in the Appendix). If the retirement age was below the statutory retirement age, the individual was coded as early retired.

3.3. Independent variables

The main independent variables of interest in this thesis are three dummy variables, representing whether the individual is underweight, overweight or obese according to their BMI. The BMI is calculated according to normal standards, i.e. by dividing the person's weight (in kilograms) by the square of his/her height (in meters). If the individual had a BMI under 18,5, he/she was coded as underweight, if it was between 25-30, the code was overweight and if the BMI was ≥30 he/she was considered obese, all according to WHO's weight classifications. The reference category will thereby be individuals with a BMI between 18,5-25, which are considered "normal

weight". The BMI is calculated using self reported weight and height. Since height is assumed to stay fixed between years, no height is reported in wave 2 and 4. Therefore, the majority of the height measurements are taken from wave 1. The assumption was made that nobody of age 50 and above weighs less than 30 kilos and/or is shorter than 100 cm, and therefore these observations were considered to suffer from measurement errors and were consequently marked as missing. This was done in order not to produce unreasonable BMI values.
As stated before, parts of this study will exploit the effect of *lagged BMI* on early retirement. Therefore such variable was created using the weight dummies from previous wave (wave 2 for wave 4 and wave 1 for wave 2). This creates a lag of two to four years, depending on when the interviews were conducted.

When discussing how to measure weight, it should be noted that BMI limits for elderly might differ from the ones used for younger individuals. According to Douketis et al (2005), overweight individuals over 65 may not be at increased risk for health problems, while those that are underweight (BMI<18,5) are at a higher risk of health problems. This implies that elderly individuals may have a higher and wider range of the "normal" BMI limits, reaching from 22-29 (compared to 18,5-25 for younger adults). We will not perform estimations using these limits, but still we feel the need to inform the reader that such research exists.

A number of control variables are included in this study and they are to a large extent inspired by the controls included in Lundborg et al (2007) and Robroek et al (2013), since they use SHARE data and study roughly the same thing. *Age* (calculated as interview year minus year of birth) and *age squared* are controlled for, as well as education in the form of *years of schooling*.

To further control for other factors, several dummy variables were created. These control for *living with a partner* and lifestyle factors in the form of *smoking* and *drinking habits* (see description in Table 1). When health indicators are included in the regression, these consist of four different dummies. The first dummy indicates whether the individual is *depressed* or not. The second dummy portraits *self-reported health*, based on a raking scale of five steps reaching from "excellent" to "poor". The other two dummies represent presence of *mobility limitations* and *chronic diseases*. The mobility limitations include: walking 100 meters; sitting for about two hours; getting up from a chair after sitting for long periods; climbing seve
one flight of stairs without resting; stooping, kne
arms above shoulder level; pulling or pushing large objects like a living room chair; lifting or carrying weights over 10 pounds/5 kilos, like a heavy bag of groceries; picking up a small coin from a table. If the respondent experienced any of these limitations, the mobility dummy is one.

The chronic diseases included in the chronic dummy variable are all diseases closely related to overweight and obesity and include diabetes, hypertension, arthritis and heart failure (Renna & Thakur, 2010). This dummy is constructed on similar manners as the mobility dummy, i.e. it takes on the value of one if any of the chronic diseases are present.

The income measure in the SHARE dataset is rather problematic, considering the significant amount of missing data. Nevertheless, we chose to still control for income in the form of *annual income*. This measure includes earnings from employment, pensions and self-employment last year. To be able to include income from pensions, some additional calculations had to be performed. In SHARE, they ask for the typical payment of pension last year. We then multiplied this answer with the period it covers, for the sake of presenting the pension payment on a yearly basis. If the payment, for example, covered one month, this payment way multiplied by 12, etc. When income from employment, self-employment and pensions was added together, the logarithm of this amount was taken, which resulted in the variable annual income. Initially, the intention was to also include a variable controlling for having a physically demanding job. This, however, turned out to be impossible, since there were too many missing values. A summary and description of all variables included in this study is shown in Table 1. The development of our key variables (early retirement, obesity and overweight) over time is shown in Table 2 in the Appendix.

Table 1. Definition and descriptive statistics of all variables					
Variable	**Definition**	**Male**		**Female**	
		Mean	**Std. dev.**	**Mean**	**Std. dev.**
<u>Dependent variable</u>					
Early retirement	D.V. 1 for early retirement	.14	.35	.12	.33
<u>Independent variables</u>					
Underweight	D.V 1 for underweight (BMI<18.5) in current wave	.00	.06	.02	.13
Overweight	D.V. 1 for overweight (BMI 25-30)	.49	.50	.33	.47
Obese	D.V 1 for obese (BMI≥30) in wave 4	.17	.38	.15	.36
L.Underweight	D.V 1 for underweight (BMI<18.5) in previous wave	.00	.07	.02	.13
L.Overweight	D.V. 1 for overweight (BMI 25-30) in previous wave	.49	.50	.15	.36
L.Obese	D.V 1 for obese (BMI≥30) in previous wave	.16	.37	.15	.36
Income	Logarithmic value of annual income from pension and employment	10.04	1.16	9.65	1.19
Age	Age in years	59.06	5.25	57.65	5.68
Age squared	Age^2	3515.19	634.12	3355.34	658.33
Education	Total number of years of education	12.62	4.33	12.15	3.89
Partner	D.V. 1 for married individuals and individuals in registered partnership	.84	.37	.75	.43
Alcohol	D.V. 1 for consuming alcohol three or four days a week or more during the last 3 months	.49	.50	.30	.46
Smoker	D.V. 1 if answering "yes" to the question "Do you smoke at present time?"	.24	.43	.21	.40
Depression	D.V. 1 if answering "yes" to the question " In the last month, have you been sad or depressed?"	.24	.43	.41	.49
Health status	D.V. 1 for SRH lower than "good"	.17	.38	.17	.38
Mobility	D.V. 1 for individuals who have at least one mobility limitation	.24	.43	.35	.48
Chronic	D.V 1 for individuals who have at least one chronic disease	.40	.49	.40	.49

D.V.=Dummy variable

As mentioned before when describing how the BMI measure was calculated, the SHARE project assumes that some variables included in my regression do not change during the course of time. Therefore, these variables are only stated if the status from the previous wave has changed. Besides height, these variables include education and marital status. Consequently, many values from these variables are collected from wave 1.

4. METHOD

In this section, the method used will be further investigated. First, we will introduce the linear probability model used in this paper. Secondly, the two strategies of POLS and FE will be introduced and discussed. We finish this section off with a discussing regarding how to measure weight.

4.1. Linear probability model

To analyze the data, the same basic model will be used, but the strategy will differ. Therefore, we start by introducing the model of choice, which is the linear probability model. In the linear probability model, the outcome variable can only take on two values, 0 or 1. The general model will look like this:

$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + u$$

where y is the binary dependent variable and β_k is the intercept for the independent variable (x_k) in question. If we assume that the zero conditional mean assumption holds ($E(u|x_1,\ldots,x_k)=0$) then

$$E(y|X) = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k$$

where X is representing all explanatory variables. Since it is always true in a LPM that E(y|X)=P(y=1|X), we can interpret β_k as the probability of a certain state to occur (Wooldridge, 2009). For example, if x_1 would rise with one unit, the probability that y=1 would increase with β_1 (everything else held constant). Evidently, this makes the linear probability model easy to estimate and to interpret.

However, there are some downturns with using a linear probability model. First of all, the model can, for certain values of x, generate probabilities below zero or over one, which does not make sense. If one, for example, look at an individual that have predominantly higher x-values than the average individual, this will probably create a probability of y=1 that negative or exceeds one. In addition, the linear probability model assumes that the marginal effect is constant, irrespective of the value of x. This seems fairly unlikely, since, for example, the effect of going from zero to one child probably is different than the marginal effect of going from four to five children.
This being said, the model still works well, especially when looking at the marginal effect close to the sample averages of the regressors (Wooldridge, 2009; Verbeek, 2012). Angrist & Pischke (2009) states that using nonlinear models for limited dependent variables may fit the conditional expectation function (CEF) better, but when it comes to the marginal effect this matter little, since the CEF is linear in the middle. Therefore, the linear probability model will be the base of further estimations.

4.2. Pooled Ordinary Least Squares (POLS)

The first strategy applied in this paper is to ignore the fact that we have a panel dataset and present a linear probability in the form of a pooled ordinary least squares (POLS) regression. Here, the binary variable *early_ret*$_{it}$ is randomly drawn from the population at time *t* and is determined by weight (W_{it}) and a vector of control variables (C_{it}).

$$early\_ret_{it} = \beta_1 W_{it} + \beta_2 C_{it} + \varepsilon_{it}$$

For observation (i) at time (t), the error term is assumed to have a mean of zero ($E(\varepsilon_{it})=0$). If we draw a random individual from the population, this implies that ε_{it} is independent and identically distributed (IID), so at any t for $i \neq j$, $E(\varepsilon_{it}, \varepsilon_{jt})=0$ (Wooldridge, 2002). This independence is, however, usually not true for the same individual across time in our panel data set. It is, for example, likely that an individuals weight is affected by unobserved factors in ε_{it} that varies little over time (such as genetics). If the variance between different time periods would differ, this would not cause problems regarding the consistency of our estimates, however the heteroskedasticity and serial correlation would imply that the estimated standard errors would be incorrect (Verbeek, 2012). Therefore robust standard errors will be used in all the upcoming estimations.

Even though robust standard errors are used, there are other reasons to why the POLS results may be biased. Cawley (2004) argues that there are genetic and nongenetic factors (such as individual choices and environment) that potentially affect both the dependent variable and the independent variable of interest (in this case early retirement and weight). He goes on by stating that there might be reverse causality, i.e. the effect between the dependent and independent variable works in both directions. Even though this problem is partly eliminated by only including working individuals at the baseline of this study, there might still be an effect of early retirement on weight. If this is the case, the results extracted show a correlation, rather than a casual effect. Making the fairly strong assumption that the residual is uncorrelated with lagged weight, we run our POLS with lagged values of weight from the previous wave, in order to reduce reverse causality.

4.3. Fixed Effects Model

In order to deal with the potential unobserved heterogeneity mentioned above, the second strategy used in this study will take advantage of the fact that the data contains several observations across time for the same individual. We specify a model as follows:

$$early\_ret_{it} = \beta_1 W_{it} + \beta_2 C_{it} + \alpha_i + \varepsilon_{it} \tag{1}$$

where αi is the unobserved individual fixed effects, such as genetics or ability. Then we average this equation over time for each individual, and subtract this from (1). Because we assume that α_i is fixed over time, subtracting the average of this variable will eliminate it and leave us with the following specification:

$$early\_ret_{it} - \overline{early\_ret_i} = \beta_1(W_{it} - \overline{W_i}) + \beta_2(C_{it} - \overline{C_i}) + \varepsilon_{it} - \overline{\varepsilon_i} \qquad (2)$$

where all variables now are time-demeaned version of the variables in (1). This is called within-transformation and since the unobserved α_i now is eliminated, we can run an OLS estimation and get the fixed effects (FE) estimator (Wooldridge, 2009).

In order for the FE model to provide consistent estimates, the regressors need to be strictly exogenous, meaning that ε_{it} needs to be unrelated to the independent variables (X_{it}) at all times ($E(\varepsilon_{it}|X_{i1},\ldots,X_{it}, \alpha_i)=0$) (Angrist & Pischke, 2009). Compared to the assumption made in POLS, this assumption is less strict and thereby more reasonable. Still, considering weight, this assumption may be unreasonable since weight probably is affected by some unobserved characteristics such as genes, etc.. However, the other assumption of unobserved characteristics being unchanged over time seems more plausible.

Using a FE approach eliminates the impact of unobserved individual fixed effects. However, similar to the POLS regression, it does not rectify endogeneity problems. These problems can be present due to omitted variables, measurement errors in our variables or reversed causality (Wooldridge, 2008). The optimal way of solving the endogeneity problem would be to find an instrument that is correlated with our weight variables and uncorrelated with any other determinants of early retirement (as done by Lundborg et al (2007) and Cawley (2000)). This, however, could not be done in our study, since no significant first stage relationship was found for potential instruments we have data for (spouse's BMI, oldest child, only sisters). Nevertheless, in an attempt to reduce potential endogeneity, we will complement our FE model with one containing lagged weight values, just as in the POLS regression.

When looking at the results from our study, one should be aware that there are a few downsides with using the FE estimator. One issue is that FE models tend to increase the impact of measurement errors, since it only measures the variation within individuals (Angrist & Pischke, 2009). Even though unreasonable observations were removed during the process of trimming the sample in this study (see Data), it would be rather over-optimistic to claim that all measurement errors are eliminated. Another possible downside with FE is that it can, by construction, only estimate the effect of weight on early retirement for those who actually change weight status.

This, in turn, may cause a selection bias, since individuals who change weight status may differ in characteristics compared to the ones who do not. One should have this in mind when comparing the results from the POLS (which include the effect of all individuals) with the FE results.

4.4. Measuring weight

For this paper, we have decided to measure weight according to the BMI standard. At the outset of this, we then continue to divide the individuals into three weight categories based on their BMI (underweight, overweight and obese), with "normal" weight as the reference category. However, when considering BMI as a measurement of weight, one should be aware of its limitations. Gosse (2014) points out that self reported weight is often under-reported, especially when studying obese individuals. Also, alternative measures of BMI that focus more on the actual body fat has been suggested by Cawley & Burkhauser (2008), since BMI risks to categorize individuals with high muscle mass as overweight. They suggest measures of percent body fat, total body fat, and fat-free mass for the sake of generating more robust results. However, the SHARE dataset unfortunately offers no such measurements. Also, several studies (Lakdawalla & Philipson, 2002; Zagorsky & Smith, 2009) have concluded that the difference between self-reported height and weight and actual height and weight are not significant. Therefore, BMI based on self reported weight and height will be used in this study.

When using weight values based on lagged BMI, lags are collected from previous wave, leaving us with a time span of two to four years. Ideally, we would like to perform our regressions including lagged weight variables collected further back in time. As we want to eliminate endogeneity by including lags, longer time periods would execute this more efficiently. Cawley (2004), for example, uses a time lag of seven years, and Conley and Glauber, (2007) expands this to 13-15 years. Increasing the time span between lags like this make it more plausible to argue that the error term do not capture omitted variables related to both present and lagged BMI. However, since the SHARE project has only been conducted four times so far (in which one time include different questions), including a longer time lag would only be possible for wave 4 (using data from wave 1). Nevertheless, this would make it impossible to perform a fixed effect study, which we do in this study. Consequently we have to settle for a shorter time lag.

5. RESULTS

In this section, estimation results will be reported. The results will be presented in four different tables; two using the POLS strategy (Table 2 and 3) and two using the FE strategy (Table 4 and 5). One table in each strategy contains estimates performed with current BMI (Table 2 and 4) and the other one is performed using lagged BMI (Table 3 and 5). Each table is divided into estimations for men and women separately.

5.1. POLS

Starting with the POLS estimation using current BMI (Table 2), column 1 shows that the male coefficient for obesity is 0.075, meaning that being obese increases the probability of early retirement with 7,5 percentage points[2]. Age and age squared are both significant and implies that getting a year older increases the probability of early retirement with 31,1 percentage points, but age squared shows that the effect is decreasing with age, which is expected. Having a partner and consuming alcohol both have positive effect on early retirement for males. When adding health variables in column 2, some of the effect of obesity disappears for males but the same covariates stay significant with approximately the same effect. This could be because some of the effect of obesity on early retirement goes through the different health variables, even though they are insignificant.

Moving on to females in column 3, none of the weight variables are significant. In contrast to males, income is significantly positively related to early retirement with a coefficient of 0.019. Having a partner and being female has about the same effect as in the male sample (4.8 percentage points compared to 5.6). When adding health variables in column 4, the effects do not change perceptibly. The added health variables are not significant among females or males.

[2] It should be brought to the readers' attention that all these effects presented are at the sample averages of the regressors (Wooldridge, 2008). One cannot generalize this marginal effect to values at the beginning or end of the sample distribution.

Table 2. Effect of current weight on early retirement. POLS regression with robust standard errors				
Variable	**Male**		**Female**	
	(1) Health variables excluded	(2) Health variables included	(3) Health variables excluded	(4) Health variables included
Underweight	-.027	-.018	-.043	-.046
	(.104)	(.105)	(.053)	(.053)
Overweight	.029	.026	.003	.000
	(.017)	(.017)	(.017)	(.017)
Obese	**.075****	**.065****	.033	.026
	(.024)	**(.025)**	(.021)	(.022)
Age	**.311*****	**.307*****	.039	.039
	(.027)	**(.027)**	(.022)	(.022)
Age squared	**-.002*****	**-.002*****	.000	.000
	(.000)	**(.000)**	(.000)	(.000)
Income	.006	.006	**.019****	**.019****
	(.007)	(.007)	**(.006)**	**(.006)**
Education	-.003	-.002	-.003	-.003
	(.002)	(.001)	(.002)	(.002)
Partner	**.056****	**.056****	**.048****	**.049****
	(.020)	**(.020)**	**(.018)**	**(.018)**
Alcohol	**.052****	**.050****	.018	.018
	(.016)	**.015**	(.017)	(.017)
Smoker	-.003	-.002	-.030	-.030
	(.019)	(.019)	(.018)	(.018)
Depression	-	-.024	-	.003
	-	(.018)	-	(.016)
Health status	-	-.005	-	.001
	-	(.023)	-	(.022)
Mobility	-	.033	-	.022
	-	(.021)	-	(.017)
Chronic	-	.016	-	.005
	-	(.018)	-	(.016)
R2	0.11	0.11	0.14	0.14
Observations	2570	2570	2502	2502

Notes: Robust standard errors are given in parentheses.
*, ** and *** represent $p<0.05$, $p<0.01$ and $p<0.001$ respectively.

When current BMI is replaced by BMI from previous wave, the results change slightly. As presented in Table 3, we can see that lagged obesity seem to have an even greater effect on the probability of early retirement. This could be because we, by using weight measures from two to four years back, managed to exclude at least some of the possible reversed causality between early retirement and weight. Another possible explanation is that the effect of becoming obese is delayed and by that obesity displays greater effect as time passes. In column 1, one can see that the effect of becoming obese increases the probability of early retirement with 10.4 percentage points for males. It thereby seems like being obese in the previous wave has an additional 2.9 percentage point effect on the probability of early retirement. The effect of age and age squared are still significant, as well as having a partner and consuming alcohol frequently. When adding health variables in column 2, same thing happens as with current BMI, i.e. the effect of obesity decreases slightly. Still, none of the health variables are significant. In the female sample, starting with column 3, income and partner are still significant and the effect is similar as when using current BMI. New is that our education variable now is significant, implying that an additional year of schooling decreases the probability of early retirement for females by 0.4 percentage points. When adding health variables, nothing happens to our estimations.

Looking at the tables utilizing the POLS strategy, problems of biased coefficients can arise, since POLS does not control for unobserved heterogeneity. In an attempt to solve this problem we move on to our next strategy, namely the FE approach. Here we cannot control for constant variables (since the FE model by construction is depending on the variation within the variables) and therefore education is left out.

Table 3. Effect of lagged weight on early retirement. POLS regression with robust standard errors				
Variable	**Male**		**Female**	
	(1) Health variables excluded	(2) Health variables included	(3) Health variables excluded	(4) Health variables included
L.Underweight	-.012	-.022	.042	.040
	(.120)	(.119)	(.058)	(.058)
L.Overweight	.022	.019	-.013	-.017
	(.017)	(.017)	(.017)	(.017)
L.Obese	**.104*** **	**.095*** **	.017	.026
	(.025)	**(.026)**	(.023)	(.022)
Age	**.309*** **	**.306*** **	.039	.039
	(.027)	**(.027)**	(.022)	(.022)
Age squared	**-.002*** **	**-.002*** **	.000	.000
	(.000)	**(.000)**	(.000)	(.000)
Income	.006	.006	**.018** **	**.018** **
	(.007)	(.007)	**(.006)**	**(.006)**
Education	-.002	-.002	**-.004* **	**-.004* **
	(.002)	(.002)	**(.002)**	**(.002)**
Partner	**.060** **	**.060** **	**.048** **	**.048** **
	(.019)	**(.019)**	**(.018)**	**(.018)**
Alcohol	**.055*** **	**.054** **	.016	.016
	(.016)	**(.016)**	(.017)	(.017)
Smoker	-.003	-.002	-.031	-.030
	(.019)	(.019)	(.018)	(.018)
Depression	-	-.025	-	.003
	-	(.018)	-	(.016)
Health status	-	-.011	-	.000
	-	(.018)	-	(.022)
Mobility	-	.034	-	.021
	-	(.021)	-	(.017)
Chronic	-	.011	-	.012
	-	(.018)	-	(.016)
R2	0.12	0.12	0.13	0.14
Observations	2568	2568	2512	2512

Notes: Robust standard errors are given in parentheses.

*, ** and *** represent p<0.05, p<0.01 and p<0.001 respectively.

5.2. Fixed Effects

Turning to Table 4, we can immediately see that most of the previously significant variables now have turned insignificant. For males, in column 1 and 2, none of the weight variables are significant and the significant age coefficient is reduced compared to previous estimations. For females, we get insignificant results for the weight variables in this estimation (however, note that the coefficients are smaller than in previous estimations). The age coefficient was never significant in previous estimations, but using the FE strategy we get significant results. However, looking at column 3 and 4, the effect of age seems to work in the opposite direction compared to the male sample. One reason for this contradictory result could be that as females are getting closer to the country's statutory retirement age, staying at work a couple extra years might not matter. One should also remember that early retirement is only one of several possible ways to exit employment early. Even though the probability of early retirement might decrease with age, the probability of unemployment or sick pension may increase. This also goes for column 4 that now shows that having bad self-rated health decreases a woman's probability of early retirement by 10 percentage points. With the exception for female self-rated health, not other health variables are significant and including them in the regression now have little impact on the weight effect (compared to the decreasing effect seen in the POLS).

Table 4. Effect of current weight on early retirement. Fixed effects regression with robust standard errors				
Variable	**Male**		**Female**	
	(1) Health variables excluded	(2) Health variables included	(3) Health variables excluded	(4) Health variables included
Underweight	.158	.149	-.119	-.111
	(.121)	(.117)	(.104)	(.105)
Overweight	-.014	-.011	.052	.059
	(.037)	(.037)	(.041)	(.040)
Obese	-.018	-.016	.123	.132
	(.063)	(.063)	(.070)	(.071)
Age	**.114****	**.109***	**-.174*****	**-.170*****
	(.035)	**(.035)**	**(.036)**	**(.036)**
Age squared	**-.001***	-.002	**.002*****	**.002*****
	(.000)	(.000)	**(.000)**	**(.000)**
Income	.006	.006	-.013	-.013
	(.009)	(.009)	(.010)	(.010)
Partner	.054	.053	-.017	-.027
	(.061)	(.062)	(.069)	(.068)
Alcohol	.040	.040	.003	.001
	(.033)	(.033)	(.036)	(.036)
Smoker	-.020	-.016	-.081	-.086
	(.041)	(.041)	(.047)	(.046)
Depression	-	.013	-	.001
	-	(.026)	-	(.024)
Health status	-	.006	-	**-.100****
	-	(.035)	-	**(.038)**
Mobility	-	.026	-	.014
	-	(.030)	-	(.029)
Chronic	-	.031	-	-.052
	-	(.033)	-	(.030)
R2	0.16	0.16	0.22	0.23
Individuals	1546	1546	1497	1497
Observations	2577	2577	2509	2509

Notes: Robust standard errors are given in parentheses.

*, ** and *** represent p<0.05, p<0.01 and p<0.001 respectively.

Replacing current BMI with lagged BMI, we get somewhat better results in Table 5. For males, all weight variables are now significant. Being male and underweight leads to a 26.7 percentage points decrease in the probability of retiring early. If he, on the other hand, is overweight or obese, this increases the probability of early retirement with 8.8 or 17.8 percentage points. In other words, our FE estimations (in conformity with our POLS estimations) confirm that lagged BMI has a greater effect on early retirement, compared to current BMI. This may imply that we eliminated some of the reversed causality. The age coefficient is lower than with current BMI (this is true for both men and women) and the age square variable is no longer significant for males. For women, the only significant weight coefficient is for women who became underweight and thereby increase their probability of early retirement with 14,8 percentage points. As when using current BMI, being female with bad health still lowers the probability of early retirement, but the rest of the health variables stays insignificant. This could be interpreted as obesity "by itself" affects the probability of early retirement (obese individuals experiencing discrimination, etc.), and the effect do not run through bad health or medical conditions that the weight might bring.

Table 5. Effect of lagged weight on early retirement. Fixed effects regression with robust standard errors

Variable	Male		Female	
	(1) Health variables excluded	(2) Health variables included	(3) Health variables excluded	(4) Health variables included
L.Underweight	**-.279***	**-.267***	**.163***	**.148***
	(.133)	**(.130)**	**(.073)**	**(.075)**
L.Overweight	**.088***	**.088***	- .083	-.082
	(.041)	**(.041)**	(.044)	(.043)
L.Obese	**.177****	**.178****	-.156	-.155
	(.066)	**(.065)**	(.080)	(.080)
Age	**.106****	**.101****	**-.167*****	**-.168*****
	(.034)	**(.034)**	**(.036)**	**(.036)**
Age squared	-.001	.000	**.002*****	**.002*****
	(.000)	(.000)	**(.000)**	**(.000)**
Income	.007	.007	-.013	-.013
	(.009)	(.009)	(.010)	(.010)
Partner	.056	.054	-.005	-.014
	(.060)	(.060)	(.068)	(.067)
Alcohol	.037	.038	-.001	-.002
	(.033)	(.033)	(.035)	(.035)
Smoker	-.021	-.017	-.098	-.104
	(.040)	(.040)	(.047)	(.047)
Depression	-	.010	-	.004
	-	(.026)	-	(.024)
Health status	-	.000	-	**-.097****
	-	(.035)	-	**(.037)**
Mobility	-	.030	-	.018
	-	(.030)	-	(.028)
Chronic	-	.029	-	-.057
	-	(.033)	-	(.030)
R2	0.17	0.17	0.22	0.23
Individuals	1543	1543	1497	1497
Observations	2576	2576	2520	2520

Notes: Robust standard errors are given in parentheses.

*, ** and *** represent p<0.05, p<0.01 and p<0.001 respectively.

As can be seen in all tables above, the effect of age on early retirement seem to be unreasonably high in the POLS regressions (in the FE regressions, the age coefficients are in line with the ones extracted in Lundborg et al (2007)). One possible explanation to this may be that there is a yearly trend affecting early retirement and that the coefficient of age catches this effect. However, when adding a control for being in wave 4, the affect of age only decreases by a few percentage points, leaving the effect at a high level. Another possible explanation is that our model specification (i.e. the linear probability model) is incorrect and is therefore causing our age variable to take on excessive values. The same thing goes for the male coefficient for alcohol in the POLS. The dummy takes on one if the individual consumes alcohol more than three or four days a week[3] and according to our estimations, this consumption would increase the probability of early retirement by around 5 percentage points, which can be discussed.

We can, based on our estimations, conclude that lagged BMI plays a bigger role in affecting early retirement than current. In the POLS regression, the difference between using current and lagged BMI is around 3 percentage points. When estimating a FE model with current BMI, unobserved heterogeneity seem to be the underlying reason behind the results in the POLS. Conversely, the coefficients and significance of the lagged weight variables for males increases once FE is used (compared to POLS). In other words, with both strategies, using lagged BMI increases the effect of overweight/obesity on early retirement, compared to using current. Generally regarding the fixed effect model, it can be concluded that unobserved heterogeneity seem to play a role in this regression. All covariates (except for the ones concerning age) turn insignificant once unobserved fixed effects are eliminated.

When it comes to the different samples, the male sample exhibits more robust results than the female sample. All the male obesity estimates (except for the FE with current BMI) implies that there will be a positive impact on the probability of early retirement if the individual is obese. Also, all the covariates have signs according to what is usually expected. The women sample, on the contrary, is more ambiguous. Few coefficients are significant, and the once that are show different signs using POLS and FE.

5.3. Robustness

To expand our knowledge of to what extent weight affects early retirement, various robustness checks will now be performed. Robustness checks are always interesting to perform in order to

[3] Lundborg et al (2007) use the same definition, except their limit for frequent alcohol consumption is lower. They consider drinking once a week or more as frequent alcohol consumption. We, however, feel like that limit includes too many people with normal drinking habits and that is why we increase it to only include individuals who are drinking three to four days a week or more.

study how general our results are and if there are any misspecifications keeping us from extracting causal estimations. Robustness will be checked by firstly eliminating the lifestyle variables smoke and alcohol, secondly eliminating the income measure and lastly by creating new ways of measuring mobility and chronic issues. All results from the robustness checks are presented in Table 6.

For the first robustness check, we exclude our lifestyle variables alcohol and smoke. Even though alcohol abuse and tobacco smoking probably do have an effect on early retirement, the construction of this kind of variables can be difficult. For example, our smoke variable catches if the individual smokes today, but ignores the duration or the quantity. Also, our alcohol variable may include people that do not have a drinking problem (since it measures frequency, not amount of drinks). To see if this bias our estimations, we try running the regressions without the lifestyle variables. For the male sample, not much changes when dropping these variables, but for the female sample, we now get some significant results on our weight variables. Therefore, it could be that some of the effect of alcohol and smoking goes through obesity. It should be noted, however, that the effects of obesity moves in different directions in the FE model depending on if current or lagged BMI is used. Therefore, the one should interpret this with caution.

In our second robustness check, we exclude the income measure. As can be read in the "Data" section, our income measure is difficult since it includes many missing values. Also, since we are performing a fixed effect model, we need to include income both before and after retirement, which might bias our estimations. However, as can be seen in Table 6, eliminating the income variable does not change our results drastically. In fact, the only deviation from our original results is found when looking at the FE strategy using lagged weight for females. When eliminating the income measure, we get significant results that both overweight and obesity should negatively affect the probability of early retirement (-9.8 and -16.5 percentage points). It is hard to say if this is the actual effect or if the weight now takes on some of the income effect. Considering no change in other estimations, one can suspect that the income measure do not play a significant role in the regressions performed in this paper.

Table 6. Robustness check				
	POLS		FE	
	(1) Current BMI	(2) Lagged BMI	(3) Current BMI	(4) Lagged BMI
Exclude lifestyle variables (alcohol and smoke)				
Males				
Underweight	-.026 (.105)	-.026 (.117)	.156 (.118)	**-.276* (.133)**
Overweight	.028 (.017)	.020 (.017)	-.017 (.037)	**.085* (.040)**
Obese	**.064* (.025)**	**.090** (.026)**	-.033 (.062)	**.186** (.066)**
Depression	-.026 (.018)	-.027 (.018)	.005 (.026)	.003 (.025)
Health status	-.012 (.023)	-.017 (.023)	.002 (.035)	-.005 (.034)
Mobility	.033 (.021)	.034 (.021)	.032 (.030)	.035 (.029)
Chronic	.018 (.017)	.014 (.018)	.029 (.033)	.027 (.032)
Females				
Underweight	-.050 (.054)	.026 (.057)	-.127 (.106)	**.167* (.080)**
Overweight	-.001 (.017)	-.018 (.017)	.068 (.039)	**-.094* (.043)**
Obese	.022 (.022)	.003 (.023)	**.142* (.066)**	**-.163* (.079)**
Depression	.005 (.015)	.004 (.015)	-.003 (.023)	-.005 (.024)
Health status	.000 (.022)	-.002 (.022)	**-.097* (.037)**	**-.093* (.037)**
Mobility	.025 (.017)	.023 (.017)	.017 (.028)	.021 (.028)
Chronic	.006 (.016)	.013 (.016)	-.046 (.030)	-.049 (.029)
Exclude income measure				
Males				
Underweight	-.025 (.104)	-.022 (.119)	.146 (.114)	**-.258* (.130)**
Overweight	.025 (.017)	.018 (.017)	-.020 (.037)	**.087* (.041)**
Obese	**.065* (.025)**	**.094*** (.026)**	-.034 (.063)	**.181** (.066)**
Depression	-.024 (.018)	-.024 (.018)	.012 (.026)	.010 (.026)
Health status	-.007 (.023)	-.012 (.023)	.005 (.035)	-.002 (.035)
Mobility	.031 (.021)	.031 (.021)	.024 (.030)	.027 (.030)
Chronic	.017 (.018)	.012 (.018)	.024 (.033)	.022 (.033)
Females				
Underweight	-.040 (.053)	.046 (.057)	-.110 (.106)	.142 (.074)
Overweight	.000 (.017)	-.018 (.017)	.068 (.040)	**-.098* (.043)**
Obese	.024 (.022)	.008 (.023)	.138 (.070)	**-.165* (.081)**
Depression	.004 (.016)	.003 (.016)	-.004 (.024)	-.007 (.024)
Health status	-.001 (.022)	-.003 (.022)	**-.100 (.037)****	**-.097* (.037)**
Mobility	.024 (.017)	.023 (.017)	.018 (.029)	.022 (.028)
Chronic	.006 (.016)	.012 (.016)	-.049 (.030)	-.054 (.030)
New measurement: chronic & mobility				
Males				
Underweight	.002 (.111)	-.017 (.120)	.147 (.115)	**-.267* (.126)**
Overweight	.026 (.017)	.018 (.017)	-.016 (.037)	**.086* (.041)**
Obese	**.066** (.025)**	**.096*** (.026)**	-.031 (.063)	**.177** (.066)**
Depression	-.022 (.018)	-.022 (.018)	.012 (.026)	.010 (.026)
Health status	-.001 (.024)	-.005 (.024)	.006 (.036)	-.001 (.036)
Mobility	.002 (.008)	.001 (.007)	-.005 (.012)	-.003 (.012)
Chronic	.015 (.012)	.012 (.012)	.027 (.023)	.027 (.023)
Females				
Underweight	-.047 (.053)	.038 (.058)	-.103 (.103)	**.142* (.072)**
Overweight	.000 (.017)	-.017 (.017)	.062 (.039)	**-.089* (.043)**
Obese	.023 (.022)	.008 (.023)	**.138* (.070)**	-.158 (.081)
Depression	.004 (.015)	.003 (.016)	-.005 (.024)	-.007 (.024)
Health status	-.001 (.024)	-.001 (.024)	**-.091* (.038)**	**-.086* (.038)**
Mobility	.006 (.006)	.005 (.006)	-.012 (.009)	-.010 (.009)
Chronic	.008 (.012)	.011 (.012)	.004 (.023)	.000 (.023)

Notes: Robust standard errors are given in parentheses.
All regressions are also controlled for covariates included in previous estimations.
*, ** and *** represent $p<0.05$, $p<0.01$ and $p<0.001$ respectively.

When looking at previous studies performed (Friis et al, 2007; Renna & Thakur, 2010; Mein et al, 2000), they find a significant relationship between some of the health indicators connected to overweight/obesity and early retirement. For us however, such relationship is lacking in all cases but self-rated health for women in the FE model. With an aim to extract significant results from our estimation, we change the construction on our chronic and mobility variables. Instead of the variables being dummies, we turn them into count measures to portrait how many mobility issues or chronic conditions the individual suffer from. For mobility issues, the maximum amount of conditions for an individual to experience is then ten and for the chronic conditions the maximum is six[4]. After running the regressions with the new variables, we can conclude that this do not change the statistical power of our health variables, compared to previous estimations. For women in the FE model, the health status is still negatively related to the probability of early retirement (although the coefficients are slightly smaller), and no significant relationship is found in the male sample. Also, positive relationship between obesity and the probability of early retirement is now found once again in the female sample.

6. DISCUSSION

In this paper, we have studied the effect of overweight/obesity on early retirement. We used a linear probability model with two different strategies (POLS and FE). When combining these two approaches, we are able to compare two estimations strategies and obtain results without interference from unobserved fixed effects. In order to reduce reversed causality, we performed our estimations with both current and lagged BMI. We also wanted to evaluate the possible effect of health indicators connected to overweight/obesity. For the male sample, our results are fairly unanimous that obesity has a positive effect on the probability of early retirement, with an effect reaching from 6.5 to 17.8 percentage points. This is higher than the 1.5 percentage Renna & Thakur (2010) estimated, but the signs are still the same. Our results conform even better to Houston et al (2009)'s results where white males experience a HR of 1,32 of early retirement if they are obese. These two studies are both performed on American data, and none of the previous studies focusing on only obesity and early retirement have found a significant relationship between overweight/obesity and early retirement. Lundborg et al (2007), however, did estimate that obesity affects the probability of being employed by -0.10 for men, which is close to our estimates. These results should of course not be directly compared to our results, since we measure slightly different things, but we still feel the importance in pointing out the similarities. When it comes to overweight, significant relationships were only found in the lagged BMI FE model. Considering Douketis et al (2005)'s observations that overweight may not be as harmful for individuals over 65, this is not surprising. No relationship was found between health indicators connected to overweight/obesity and early retirement.

[4] See Data section for more information regarding which conditions that are included.

For the female sample, the results are more ambiguous. No significant relationship was found between overweight/obesity and early retirement, and the insignificant coefficients lack precision and vary extensively. This is consensual with both Robroek (2013) and Fris et al (2007), who both never managed to find a significant relationship between any of the sexes' overweight/obesity and early retirement. On the contrary, Renna & Thakur (2010) find an even greater effect between obesity and early retirement (2,5 percentage points) among females, compared to when they look at the male sample. Using the FE model, we found that bad self-rated health decreases the probability of early retirement with around 10 percentage points. Including health indicators had mixed impact on the weight effect depending on strategy applied. Using POLS lowered the weight effect for both sexes when including health variables, while the FE strategy shows similar weight effects with and without health indicators. Therefore, we do not see any clear evidence that the effect of obesity on early retirement runs through the health of an individual. True for all estimation is, however, that the health indicators remained insignificant (with the exception of self-rated health for women in the FE model).

Furthermore, our results show that lagged BMI has a greater effect on the probability of early retirement than current BMI in the male sample. In the POLS model the difference is around 3 percentage points and in the FE model, the current BMI estimations are not significant, while the lagged estimation indicates an effect of 17.8 percentage points. The greater coefficient achieved when using lagged BMI could be because we managed to get rid of some of the reversed causality, i.e. the effect running from early retirement to weight. Our results here are, however, hard to compare to previous research in this field, since former studies have been performed using *only* lagged BMI or *only* current BMI. For the female sample, no comparison between current and lagged BMI could be made since no weight variables were significant.

In contrast to other studies (Friis et al, 2007; Renna & Thakur, 2010; Mein et al, 2000), we find no significant connection between our included health indicators and early retirement (with the exception for women and self-rated health). Not even when trying different measures for mobility limitations and chronic conditions did we get a significant result. This can be interpreted as obesity per se affects early retirement, and not through different channels such as bad health, mobility issues, etc.. Also, this could mean that other factors that we did not control for in our estimation (such as discrimination, unsatisfying job, productivity differences, etc.) have greater impact than health state on an obese individual's decision on leaving the labor market early. Another possible reason for our insignificant health variables is that our health variables might be too broad and include people who are not suffering from their conditions. For example, our variable for depression represents if the person has been sad or depressed during the last month. This is a fairly arbitrary question that could be interpreted very different among different people, and possibly that results in covering too many people without actual depression problems. It should also be mentioned that our results regarding health indicators are in line with the findings of Robroek et al (2013), who uses data from the same source as the one utilized in this paper.

Bad self-rated health reduces the probability of early retirement among women, according to our FE results. This is on the contrary to what Friis et al (2007) conclude. However, one should keep in mind that early retirement is only one of several exists from paid employment. Just because bad self-rated health lowers the probability of early retirement does not mean it lowers the probability of exiting the labor market. In fact, Lundborg et al (2007) concludes that this is not the case, and state that women with a bad self-rated health have a lower probability of being employed. This may, of course, also be the case for our other health indicators, i.e. people suffering from any health problems may not find an increased risk of early retirement, but that does not mean that the probability of leaving paid employment is unaffected.

Generally, the results for women in our study were more imprecise than the ones collected from the male sample. Either overweight and obesity have no effect on women's early retirement or there are misspecifications in our model. One possible explanation is that there are covariates missing that could help explain early retirement. Although we have included control variables that are broadly in line with Lundborg et al (2007) and Robroek et al (2013), there is always a risk that there are additional covariates that need to be controlled for. For example, Hochman & Lewin-Epstein (2013) finds that having grandchildren affects elderly's willingness to retire positively. Also, due to the fact that some variables contain too many missing values (for example physically demanding job) or are simply not included in the SHARE questionnaire (for example sibling/parent's BMI, culture, motivation or ability), these have to be excluded from our regression. If these unobserved regressors have a significant effect on early retirement, this may lead to bias in our estimations.

When interpreting the effects of our independent variables, one should be aware of the sample selection. This paper only looks at individuals who are employed in wave 1 in order to make sure they all individuals once were part of the labor market. This, however, means that our coefficient results cannot be interpreted as the general effect of obesity on early retirement, but the effect of obesity on early retirement *for working individuals*. It cannot be generalized to other groups. In other words, homemakers, unemployed, permanently disabled, etc., can also enter into early retirement, but then the effects of overweight/obesity might be different.

Our estimations in this paper are not without limitations. As mentioned above, there might be missing covariates, resulting in incorrect estimations. Also, we are aware of that there are many missing values in some of our included variables. If the missing values are randomly distributed among individuals, this will not be an issue. However, if this is not the case and a certain "type" of individuals leave out their answers, this will lead to internal validity problems (Angrist & Pischke, 2009). Additionally, assumptions made when constructing our models are fairly strong. For example, the linear probability model assumes that the error term in any given period is uncorrelated with the explanatory variables in all time periods. Considering we are looking at BMI, this is rather unreasonable since unobserved genes, upbringing, etc., are likely to affect weight. Also, the fixed effect model only controls for fixed unobserved heterogeneity, meaning that time-variant unobserved variables could still bias our estimations. In this regard, using an instrumental approach might have been better, however, no suitable instruments were found in our sample.

Another issue could be the definition of early retirement. We have gathered all countries' statutory retirement age to the best of our ability, but there could of course still be different policies in different countries making it more profitable to retire before statutory retirement age. These country specific policies have not been accounted for in this study and may skew our results. However, as the aim of this study is to evaluate if overweight/obesity causes individuals to retire before statutory retirement age (compared to their normal weighted counterparts), we do not consider this an immediate problem. Lastly, we would ideally wish to have a longer and bigger panel dataset, i.e. including more than two years, in order to get make sure unobserved fixed heterogeneity is properly excluded.

7. CONCLUSION

This paper aimed to estimate the effect of overweight/obesity on early retirement. This was done by using a linear probability model and two different strategies; POLS and FE. Data was gathered from the SHARE survey and the sample was divided into a male and female group. Regressions were run using both current and lagged BMI. Our results establish that obesity affects the probability of early retirement positively for males, while overweight show few signs of having any significant effect. Lagged BMI had a larger effect of early retirement than current in the male sample. For females, no significant relationship between overweight or obesity on early retirement could be found. When it comes to health indicators connected to overweight/obesity, no significant relationship was found between these and early retirement in the male sample, while self-rated health seem to affect the probability of early retirement in some cases among women. When it comes to the health variables impact on the weight effect, we get mixed results depending on which strategy we used. Considering this, it seems like obesity "alone" affects early retirement, and the effect do not run through the health variables we included.

For future research, it would be interesting to emerge deeper into through which channels weight affects early retirement. Focusing more on lagged weight variables may increase the theoretical robustness and, judging from our results, also the effect. Also, performing estimations similar to the ones in this study, but using a different method could generate deeper understanding. Another thought is to use other measures of weight, such as percentage of body fat, instead of BMI and see how the results changes. It would also be interesting to see the results if Douketis et al (2005)'s BMI limits were used instead of conventional BMI limits.

When looking at our results, it is important to keep in mind that obesity is only one out of several channels into early retirement. Furthermore, early retirement is only one out of multiple ways to leave the labor market prematurely. Still, we know that obesity is harmful for many reasons and if it, on top of all other complications, leads to early retirement, reduction of weight will only bring positive results. Since current public debate often discusses how to keep our labor force at work longer, this paper brings further light on the question. If actions towards reducing the occurrence of obesity, primarily among males, would be taken, our results suggest that the rate of early retirement would decrease. But of course, one can always ask if obesity really is the problem, or if it is the underlying reasons for obesity that, in turn, affects early retirement.

REFERENCES

Allison D.B., Fontaine K.R., Manson J.E., Stevens J., VanItallie T.B. 1999. Annual deaths attributable to obesity in the United States. *JAMA*. 282: 1530–38.

Angrist J.D & Pischke J.S. 2009. *Mostly Harmless Econometrics, An Empiricist's Companion.* Oxfordshire: Princeton University Press

Averett, S., & Korenman, S. 1996. The economic reality of the beauty myth. *Journal of Human Resources*, 31: 304–330.

Asgeirsdottir, T. L. 2011. Do body weight and gender shape the work force? The case of Iceland. *Economics and Human Biology*, Vol. 9, 148 – 56.

Baum C.L, & Ford, W. 2004. The wage effects of obesity: a longitudinal study. *Health Economics*, 13: 885-99.

Bungum, T., Satterwhite, M., Jackson, A.W., Morrow, J.R.J. 2003. The relationship of body mass index, medical costs, and job absenteeism. *Am J Health Behav.* 27:456 – 62.

Cai, L., Kalb, G. 2006. Health status and labour force participation: evidence from Australia. *Health Econ.* 15(3):241– 61.

Cawley, J. 2004. *Impact of obesity on wages.* The Journal of Human Resources, Vol. 42,51 – 74.

Cawley, J., & Burkhauser, R. V. 2008. Beyond BMI: the value of more accurate measures of fatness and obesity in social science research. *Journal of Health Economics*, 27: 519-29.

Cawley, J. Grabka, M. & Lillard, D. 2005. A comparison of the relationship between obesity and earnings in the U.S. and Germany. *Journal of Applied Social Science Studies* (*Schmollers Jahrbuch*), 125: 119-129.

Conley, D., & Glauber, R. 2007. Gender, body mass, and socioeconomic status. In: K. Bolin, & J. Cawley (Eds), Advances in health economics and health services research (Vol. 17, 255–78), *The Economics of Obesity.* Amsterdam: Elsevier.

d'Hombres, B., & Brunello, G. 2005. Does obesity hurt your wages more in Dublin than in Madrid? Evidence from the ECHP. IZA, *Discussion Paper No. 1704*.

Douketis, J., D., Paradis, G., Keller, h., and Martineau, C. 2005. Canadian guidelines for body weight classification in adults: application in clinical practice to screen for overweight and obese and to assess disease risk. *American Journal of Clinical Nutrition*, Vol. 172, 8: 995 – 98.

Friis, K., Ekholm, O., Hundrup, Y.A., Obel, E.B., Gronbaek, M. 2007. Influence of health, lifestyle, working conditions, and sociodemography on early retirement among nurses: the Danish Nurse Cohort Study. *Scand J Public Health*. 35(1):23–30.

Gallo, W.T., Bradley, E.H., Siegel, M, Kasl, S.V. 2000. Health effects of involuntary job loss among older workers: finding from the health and retirement survey. *J Gerontol B Psychol Sci Soc Sci.* 55(3): 314–40.

Garcia, J. & Quintana-Domeque, C. 2007. Obesity, employment and wages in Europe. *Advances* *in Health Economics and Health Services Research*, Vol.17. The Economics of Obesity, K. Bolin and J. Cawley, (eds.). Amsterdam: Elsevier.

Gosse, M.A. 2004. How accurate is self-reported BMI? *Nutrition Bulletin*. Vol. 39 Issue 1, 105-14.

Hochman, O., & Lewin-Epstein, N. 2013. Determinants of early retirement preferences in Europe: The role of grandparenthood. *International Journal of Comparative Sociology* 54(1): 29-47.

Houston, D.K., Cai, J., Stevens, J. 2009. Overweight and obesity in young and middle age and early retirement: the ARIC study. *Obesity (Silver Spring)*. 17(1):143–49.

Ilmarinen, J.E. 2001. Aging workers. *Occup Environ Med*. 58(8):546–52.

Jusot, F., Khlat, M., Rochereau, T., Serme, C. 2008. Job loss from poor health, smoking and obesity: a national prospective survey in France. *J Epidemiol Community Health*. 62(4):332–37.

Lakdawalla, D.N., Philipson, T. 2002. The Growth of obesity and Technological change: A theoretical and empirical examination, *NBER working paper No. 8946*

Luengo-Fernandez, R., Leal, J., Gray, A., Sullivan, R. 2013. Economic burden of cancer across the European Union: a population-based cost analysis. *Lancet Oncology*. 14(12): 1165-74

Lundborg, P., Bolin, K., Hojgard, S., and Lindgren, B. 2007. Obesity and occupational attainment among the 50+ of Europe. *The Economics of Obesity. Advances in Health Economics and Health Services Research*. Vol. 17, 219 – 51.

Mein, G., Martikainen, P., Stansfeld, S.A., Brunner, E.J., Fuhrer, R. & Marmot, M. 2000. Predictors of early retirement in British civil Servants. *Age and Ageing*. 29: 529-36.

Müller-Riemenschneider, F., Reinhold, T., Berghofer, A., & Willich, S. N. 2008. Health-economic burden of obesity in Europe. *European Journal of Epidemiology*. 23(8): 499-509.

Palta, M., Prineas, R. J., Berman, R., & Hannan, P. 1982. Comparison of self-reported and measured height and weight. *American Journal of Epidemiology*. 115(2), 223-30.

Register C.A. and Williams D. R. 1990. Wage effects of obesity among young workers. *Social Science Quarterly*, 71: 130-41.

Robroek, S.J., Schuring, M., Croezen, S., Stattin, M., Burdorf, A. 2013. Poor health, unhealthy behaviors, and unfavorable work characteristics influence pathways of exit from paid employment among older workers in Europe: a four year follow-up study. *Scand J Work Environ Health*. 39(3):125–33.

Verbeek, M. 2012. *A Guide to Modern Econometrics*, 4th edition, John Wiley & Sons

Wooldridge, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge, Massachusetts.

Wooldridge, J. M. 2009. *Introductory Econometrics: A Modern Approach*. 4th edition, South Western, Mason, Ohio.

Zagorsky, J. L., & Smith, P. K. 2009. Does the U.S. Food Stamp program contribute to adult weight gain? *Economics & Human Biology*. 7(2): 246-58.

Electronic sources:

SHARE database

APPENDIX

Table 1. European Statutory Retirement Ages		
	Male	**Female**
Austria	65	60
Belgium	65	65
Denmark	65	65
France	60	60
Germany	65	65
Italy	65	60
Netherlands	65	65
Spain	65	65
Sweden	65	65
Switzerland	65	64

Notes: Statutory retirement ages were gathered from the European Communities (2009) and Robroek (2013).

If the values conflict, the higher value was used.

Table 2. Development of key variables over time		
Percentage of individuals coded as:	Wave 2	Wave 4
Early retired	9.6%	30.1%
Obese	16.7%	17.1%
Overweight	39.8%	42.3%