The effect of a pension reform on subjective well-being of older workers: A comparison of Denmark and Belgium

Jochem Pronk

436060

Supervisor: Jolanda Hessels

Second assessor: Niels Rietveld

Erasmus University Rotterdam

Erasmus School of Economics

MSc Strategy Economics

Master Thesis Economics

Rotterdam

22-4-2022

Table of Contents

Abstract

Many members of the European Union (EU) have implemented pension reforms due to the expected increase of the old age dependency ratio in the coming decades. A common reform was an increase in the state pension age, which led to longer working lives. The goal of this thesis is to analyze the often-overlooked effects of an increase in pension age on the subjective well-being (SWB) of older workers in Denmark.

Panel data gathered between 2004 and 2020 from the Survey of Health, Aging and Retirement in Europe (SHARE) is used. The sample consists of Belgian and Danish people born between 1950-1959. A difference-in-differences (DD) method embedded in ordered and multinomial logistic regression models is used to analyze the effect of an increase in pension age on subjective well-being. To relax the assumption of parallel trends in subjective well-being before the policy change, Belgium is added to the model and a linear triple differences (DDD) model is implemented.

Job satisfaction significantly decreases for Danish workers affected by the increase in pension age in the DD model, but not in the DDD model. For life satisfaction and depression, no significant deterioration was found.

The differences in outcome for job satisfaction might be caused by a multicollinearity issue in the DDD model, and therefore the DD model is deemed to provide the most trustworthy outcomes. Moreover, It is possible that life satisfaction and depression experience a short-term shock after the policy, but the effects are recovered within a few years.

Keywords: subjective well-being, pension reform, Denmark, difference-in-differences, ordered logit

1. Introduction

In the coming decades, the proportion of elderly people in most European countries will rise to unparalleled levels in the history of humankind (Börsch-Supan, 2022). As a consequence, most member states of the European Union (EU) have had to implement changes to their pension systems so they can be sustained in the future (Carone et al., 2016). A common reform measure was an increase in the state pension age. If people remain active in the labor force until a later age, they contribute to the pension system for a longer time. The government simultaneously has to start paying pension benefits at a later point in time. The state thereby enables itself to preserve a pension system that does not incur losses. For workers however, the consequence is a longer work life than previously expected. This consequence might negatively impact the well-being of older workers.

To quantify welfare, analyses in the twentieth century almost exclusively used economic factors such as gross domestic product (GDP) (Conceição & Bandura, 2008; Kahneman & Krueger, 2006). However, both researchers and politicians have observed that increases in economic performance have been matched with stagnant levels of well-being in the last decades. Therefore, the importance of subjective well-being (SWB) has been acknowledged and became more integrated in policy decisions. SWB consists of many aspects and has been divided in three different groups: hedonic, evaluative and eudaemonic SWB.

Nevertheless, pension reforms were motivated by changes in demographics and the desire to attain fiscal sustainability (Carone et al., 2016). Previous research has shown that millions of euros can be saved annually by changes in the pension system (Staubli & Zweimüller, 2013). The impact on the mental health of the working population, particularly on the older workers who are the first ones to receive pension benefits at a later age, is however often overlooked. In existing literature, it is argued that big life events can affect SWB of individuals (Pavot & Diener, 2009). A change in policy that results in extra working years could thus have a negative impact on SWB. Hence, the research question of this thesis is:

How is SWB of older workers affected by an increase in the pension age?

Though, data availability on SWB has historically been lower than data on economic measurements. Predicting SWB effects as a consequence of policy has therefore been challenging. However, aiming for higher levels of SWB is not merely an intangible ideal. Lower levels of well-being could namely be paired with higher costs in healthcare and insurances.

Previous research has found that lower levels of SWB could lead to physical health deterioration (Carrino et al., 2018). Moreover, workplace errors are more likely among employees with higher depression rates (Grip et al., 2012).

Empirical research about the effects of a pension reform on SWB of older people has been scarce. In the Netherlands, higher levels of depression and lower levels of job satisfaction were found on the among public sector employees who were affected by an increased pension age (Grip et al., 2012; Montizaan & Vendrik, 2014). Though, the findings were short-term results, and it is also uncertain if similar patterns can be found in other EU countries.

In this research, the effects of an increase in pension age of three proxies for SWB will be analyzed: job satisfaction, life satisfaction and depression. Panel data from the Survey of Health, Aging and Retirement in Europe (SHARE) will be used to analyze changes in SWB of older Danish workers after an increase in the state pension age. The database was created by the European Commission (EC), which recognized the importance of understanding the well-being of elderly because of the rising old-age dependency ratio. The data consists of 8 waves and was gathered between 2004 and 2020. For analysis on the data, linear regression methods are inefficient because of the ordered categorical nature of the dependent variables. Consequently, a difference-in-differences (DD) method will be embedded within ordered logit models. DD is a common method to evaluate policy effects in natural experiments (Dimick & Ryan, 2014). Two periods are divided: the first before the policy change, and the second afterwards. Both of the periods consist of a seven-year period. The average differences in probabilities of SWB outcomes of the first birth cohort of Danish workers to receive pension benefits at a later age between periods will be measured. Then those differences will be compared to the average differences in probabilities of SWB outcomes of the last birth cohort that experienced no increase in pension age. The comparison between the different birth cohorts is made to control for changes in SWB that would have happened anyway in absence of a change in policy. Since a systematic difference between birth cohorts might exist, Belgium will be added to the model afterwards. A linear triple difference (DDD) method will be implemented in this part of the study. Though less efficient than logistic models, a linear model will be used because of the high complexity of logistic DDD analyses and the lack of the usage of the model in existing literature. The DDD model relaxes the assumption that a parallel trend between the birth cohorts should exist for the SWB measures before the pension reform announcement. Control variables that may impact SWB, such as physical health and marital status are implemented in the models. For job satisfaction, the DD model shows for Danish cohort affected by the increased pension age a significantly higher probability to fall in the worse

categories of job satisfaction compared to the unaffected cohort. For the other dependent variables, no significant deterioration of SWB is found for Danish workers affected by a higher pension age.

This thesis will be structured as follows. The next section will provide the theoretical background to the study. It shows the relevance of the study and empirical findings on the topic. Afterwards, the data source and the sample selection will be discussed. It is followed by the methodology section, which will start off with a description of the variables. Then, the statistical methods used to test the hypotheses are presented. The result section follows, where assumptions of the models are tested, leading to the model choice and the selection of control variables. The results of the regressions of the preferred models will be interpreted subsequently. Finally, the discussion part will conclude this thesis. It consists of a summary of the results and compare them to other findings in existing literature. Next, the research question will be answered. Also, some limitations of the study will be discussed, and lastly some implications and recommendations for future research are provided.

2. Theoretical Framework

In this chapter, the theoretical background will be discussed. This part will give a better understanding on the relevance of the study, as well as an argumentative basis for the hypotheses that will be tested. First, different methods to analyze the effectiveness of public policy will be explained. A historical view will be provided, where a shift can be observed in the focus of policy outcomes. Politicians and scholars alike argue for a larger emphasis on policy effects on well-being, instead of policy effects on the economy. Then, an overview will be provided on the recent pension reform trends within the EU. Afterwards, a summary of the effects of the pension reforms in the EU will be shown on the basis of previous research outcomes. The effectiveness on pension reforms will be presented from an economic perspective, as well as a well-being perspective. Lastly, the Danish and Belgian pension reforms will be covered more in depth since these countries form the case of this study. In the final part of this chapter, the hypotheses will be formulated as well.

2.1. Public Policy Determinants

It is challenging to predict public policy outcomes, as it is hard to evaluate their effectiveness. For policy evaluation, it is important to have large sets of data on the country level as well as on the individual level. Moreover, one needs to control for changes that would have happened anyway without policy. The following part will give the historic course of policy evaluation methods, and how it has changed over time.

2.1.1 Economic Indicators

Economists have seen a crisis unfold where using economic indicators as leading determinants for public policy has led to unwanted outcomes (Forgeard et al., 2011). Simon Kuznets (1934), the economist who first started gathering economic data in the United States and promoting its use, then already argued that collecting economic information should be used to analyze how such indicators influence the well-being of nations, rather than be used as proxies for well-being. Indicators as gross domestic product (GDP) and income are easy to measure and have been readily available for a long time. Therefore, they have been predominant factors in policy decisions and as indicators of well-being (Conceição & Bandura, 2008; Kahneman & Krueger, 2006). Admittedly, in earlier stages of the rapid economic development over the last century, increasing economic output was important, because basic needs had to be met (Diener & Seligman, 2004). On top of that, the wealth of a nation is correlated with well-being.

However, in recent decades the critique towards the weight which is given to the economic indicators as a basis of public policy has drastically increased. Easterlin (1974) was the first economist who empirically showed that for more developed countries, a higher income per capita did not result in more happiness. A trend can be seen that output and income have boomed since the 1950s, while well-being has barely been affected (Diener & Seligman, 2004). Therefore, on the long run GDP appears to be a bad predictor of subjective well-being (SWB) (Bartolini & Sarracino, 2014). Even more concerning is the fact that mental disorders have boomed in the same time period. The highest average levels of well-being are not always found in the most economically prosperous countries. They can instead be found in countries with effective political institutions, social capital, high mutual trust, and low corruption (Helliwell, 2003). Political institutions encompass accountability, stability and lack of violence, government effectiveness, rule of law, control of corruption, and the regulatory framework. Direct democracy, where referenda are common, also leads to higher well-being (Frey & Stutzer, 2000). Higher social capital is found in countries where social activities and connections are high. Helliwell (2003) measures this by club membership. Additional criticism on GDP focus comes from Zencey (2009), who argues that the variable omits production which adds value where no transactions are made, such as volunteer work. Also, the value of natural capital cannot be included in GDP data, since setting a price on such amenities can lead to subjectivity. Moreover, spendings such as pollution abatement and health care spending have a positive impact on GDP, while it decreases welfare. These externalities of market activities cannot be monetized (Adler & Seligman, 2016). Thus, GDP measures activity rather than

benefit (e.g., for every tree that is cut down, GDP and thus welfare rises, which is counterintuitive because the intrinsic value of the tree gets ignored).

In the context of evaluation of public policy, economists have traditionally observed how policies affect behavior to measure their impact on social welfare. Then, through modeling, they link policies to changes in welfare. With these measures it is however impossible to capture changes in well-being (Di Tella & MacCulloch, 2006). Another issue is that changes in SWB are not always rational. Some macroeconomic changes, such as inflation, should have no consequence on the well-being of people (Di Tella & MacCulloch, 2006). In reality, however, it can be observed that inflation has a negative impact on happiness (Wolfers, 2003). Failing to adapt to irrational feelings can negatively impact welfare and reduce trust in governments and their economic statistics (Conceição & Bandura, 2008).

The previous findings have shown that measuring economic effects of policy change are inefficient, since better economic circumstances do not translate in higher levels of well-being. The next section will provide alternative methods of policy evaluation.

2.1.2. Subjective Indicators

The critique on economic measurements as policy determinants was accompanied with alternative indicators of welfare to focus on. Both academics and politicians have argued that subjective measures should be implemented in policy decisions (POST, 2012). Former President of France, Nicolas Sarkozy, wanted better measurements in well-being and use them for public policy (Arora, 2008). Moreover, in his time as prime minister of the United Kingdom (UK), David Cameron acknowledged that GDP should not be the leading driver of the country; general well-being should take its place (Stratton, 2010). Following these words, the UK Office for National Statistics started to collect data about well-being and debate began how to implement these measurements to policy (Seaford, 2011).

Multiple academics also recognize the importance of implementing SWB into public policy decisions. In a reaction to Sarkozy's action to improve measurements on well-being, Easterlin (2010) reported that the use of self-reported subjective feelings into politics is long overdue. Diener & Seligman (2004) assert that policymakers should make well-being a primary focus, because there are considerable differences between economic indicators and well-being. Similarly, Taylor (2011) argues that welfare economics has a focus on economic utility, whereas a focus on well-being could bring us to a 'fully rounded humanity'. For many of the key areas of public policy, direct analysis of SWB can give much better measurements for the benefits of a policy than objective economic variables (Layard, 2010). None of these academics

claim that traditional economic indicators have no use anymore. They argue for policies where both economic and subjective indicators are considered instead.

Three subcategories of SWB can be distinguished (Steptoe et al., 2015). The first category is evaluative well-being, commonly measured by life satisfaction. Life satisfaction consists of multiple domains of satisfaction. Some domains include satisfaction on job, relationships, health, finances, living area, time allocation, and children. Life satisfaction is an aggregate of the domain satisfactions, where each individual puts a different weight on each domain (Dolan & Metcalfe, 2011). Measurements do often not vary much over a brief period of time, which makes yearly questionnaires fairly trustworthy. Substantial changes in life satisfaction can be found after big life events (Pavot & Diener, 2009). The second SWB category is hedonic well-being, where feelings of positive (e.g., happiness and optimism) and negative emotions (e.g., stress and anger) are used to measure well-being. The negative emotions are subtracted from positive emotions and aggregated over time, which should represent a person's well-being (Seligman & Csikszentmihalyi, 2014). Hedonic well-being may vary between short periods of time, and data should be gathered frequently. Moreover, hedonic experiences only encompass small doses of well-being, which might not translate to satisfaction on the long term (Kahneman, 1999). The third SWB category is eudaemonic well-being, where the sense of purpose and meaning in life is used to measure well-being. Among older people, eudaemonic well-being is positively related to survival (Steptoe, et al., 2015). Of the SWB categories, eudaemonic well-being fluctuates the least over time.

Both scholars and politicians argue for SWB measures as a basis of policy making and policy evaluation. However, many policies are still formed through the expected effect on the economy, and many researchers still evaluate research through their economic outcome. The following section will shed light on pension reforms in the European Union, where economic considerations weighted heavily.

2.2. Pension reforms in the European Union

European countries score high on the Global Pension Index (GPI), which rates the pension systems of countries in three separate domains (Mercer & CFA Institute, 2022). Adequacy is the domain that weighs most heavily, and it encompasses the system design, savings, and government support. It accounts for 40% of the overall GPI. Sustainability accounts for 35% of the GPI score and it consists of government debt, demography and economic growth. Integrity accounts for the last 25% and contains governance, protection and communication. However, Europe is facing a challenge in being the fastest aging continent worldwide, threatening the

durability of pension systems (United Nations, 2017). In 2007, 26 out of the 27 current members of the EU were part of the 40 most aged countries in the world, with an exception for Ireland (Beets, 2007). In all EU countries, under country-specific rules for eligibility, elderly people can receive state pension benefits. Rules for receiving a full state pension include working in the country for a minimum number of years and reaching a certain age (Carone et al., 2016). Often, people retire from work after they start receiving pension benefits. The aging trend has caused that the proportion of non-working people compared to working people has risen. Since pension benefits are paid by todays workers, every elderly person's pension benefit is paid by a lower number of people. Without intervention, pensions will become too costly in the future. Many European governments have reformed their pension systems, so that fiscal sustainability can be attained whilst the elderly may still receive sufficient pension benefits (Carone et al., 2016). Starting in the early 2000s, the intensity of reforms started to rise, with an acceleration at the global crisis of 2008-2009 (Van Vliet et al., 2012).

Within the EU, a number of common trends in pension reforms can be found. First, the pension age has been increased in all member states but Luxembourg. Additionally, most countries where women have a lower pension age than men plan to equalize the pension age for the genders. Second, incentives for early retirement are blocked. Multiple EU countries experienced high unemployment rates in the 1970s, whereafter early retirement became an easy option for the older unemployed through either unemployment or sickness insurances. Reforms have blocked these paths to retirement. Furthermore, some EU members increased the minimum years of contribution required to receive a full pension. Another policy has been penalizing people who retire early or provide bonuses to people who extend their work life. Also, some countries scrapped the formal retirement age, which more or less forced people to retire at a certain age (Carone et al., 2016). Third, systemic changes to the pension system have been implemented. In many countries, the public pension system has been the most important pillar of the pension system. Now, however, a switch towards private pension schemes as the most important pillar can be observed (Van Vliet et al., 2012). Another systemic reform is the inclusion of mechanisms that are aimed at automatically adjusting key pension parameters (Carone et al., 2016). Pension age, benefits and financing resources will be adjusted when life expectancy and old age dependency increase. These mechanisms are implemented to ensure sustainability of pension systems. Fourth, measures to increase revenues have been implemented. Some EU members increased the social contribution rate, while others increased taxes on pensions (Carone et al., 2016).

The financial crisis of 2008-2009 accelerated reforms. During the crisis, public support for pension systems fell, through temporary reduction of contribution, and reduction of tax-incentives and subsidies (Carone et al., 2016). Besides, the overall value of private pensions fell with 23% in 2008. Finally, older workers losing their jobs were more willing to retire early, increasing governments' expenditure on pensions. The solutions of EU member states fall in one of three categories. First, there were temporary measures in pension benefits. Some countries cut in pension payments, while others reduced indexation or increased contributions. Second, eligibility rules were changed in some countries. Some countries closed of early retirement access, while others rapidly increased pension age. Third, pension system architecture has been tackled. These measures were aimed at private pension schemes. The most common measure here was a reduction in tax incentives or subsidies to private pensions. Another measure was to reduce the contributions to mandatory private schemes.

The previous section provided an explanation why EU countries opted for pension reforms, and presented some common trends within reforms. The following section will provide the economic policy effects of the pension reforms within the EU.

2.3. Economic Evaluation of Pension Reforms

Several studies have investigated whether these policies have resulted in higher employment of older workers and postponement of retirement. Within Europe, the average retirement age has increased. In 2000, the average retirement age in Europe was 63.1, compared to 65.4 in 2018 (OECD, 2022). Eichhorst, et al., (2014) state that increasing the state pension age, as well as removing incentives to retire early are the main contributors to prolonging work life. These incentives include early retirement schemes through unemployment insurance. Failing to remove both might cause that the desired economic benefit may not be obtained. For instance, a study conducted in Austria showed that increasing the early retirement age resulted in significantly higher employment of older workers. However, an even larger spillover effect was found, where low-wage less healthy workers still tended to retire early through unemployment insurance (Staubli & Zweimüller, 2013). Nevertheless, the authors expect government net expenditures to be reduced by approximately 230 million euros per birth-year cohort yearly. Moreover, Hairlault et al., (2010) found evidence that social age is a better explainer of employment of an age cohort than biological age. The social age is determined by how many years one has left until they are old enough to receive pension benefits. They found that the interaction between the distance to pension and generosity of unemployment benefits significantly affect the employment rate. The last effective change in pension schemes to get

people working longer is to put a heavier weight on labor income earned as an older worker on pension assessment (Buyse et al., 2013). This implies that two people with similar lifetime earnings might get dissimilar pension benefits, because one of the two retired early. The one who worked until an older age will get a higher benefit in this case.

In the previous section, some consequences of pension reforms are discussed in terms of monetary benefits and extension of work life. However, many papers fail to research the mental effects of the people who have been affected by the reform. The next section will summarize some findings on effects of pension reforms on SWB of the older working population.

2.4. Subjective Evaluation of Pension Reforms

Even though economic effects of pension reforms appear to be positive, the effects on the impact of well-being of older workers may have been overlooked or underestimated. The goals of the pension reforms were to attain fiscal sustainability, implying that economic measurements were focused upon. Admittedly, many of the EU members have gradually increased pension age, indicating that there was some sort of awareness on mental effects of the working population. Some concerns on physical health and mental well-being of older workers after pension reforms have arisen, but a lack of empirical research is the reason that little is known about the topic (Pilipiec et al., 2021). One research found that UK pension reforms have had negative effect on the mental and physical healh of women in routine-manual occupations (Carrino et al., 2018). Moreover, two studies on the SWB of older workers after a pension reform in The Netherlands have brought some empirical results. One research was a study on Dutch employees in the public sector where the depression rate of people born in 1949 and 1950 was investigated. The 1950 cohort was the first age group that was affected by the new pension system, meaning that they could receive state pension benefits at a later age. The 1950 cohort had a significantly higher rate of depression than the 1949 cohort one year after the policy change (Grip et al., 2012). This is alarming for three reasons. First, depression is linked to other types of illnesses, such as diabetes, heart disease and a number of forms of cancer (Ranga, et al., 2002; Aromaa, et al., 1994). Second, workplace errors may occur more often and worker productivity may decline when a worker is depressed. Third, healthcare cost and disability insurance cost may rise (OECD, 2008).

Later, a followup study of the Dutch workers was conducted were job satisfaction was the new measurement of SJW. Another addition to this study was the implementation of a panel stuctured analysis. In this research, the effects were measured over a period of three years. For

the 1950 cohort, job satisfaction was significantly lower than for the 1949 cohort at every year after the pension reform (Montizaan & Vendrik, 2014). Grip et al. (2012) argued that one of the mechanisms that caused the deterioration in mental health was a sense of unfair treatment. People often compare themselves with people of similar demographics. When people of similar age have to work until a lower age than another, this can have negative effects on SWB. Moreover, Alesina et al. (2004) found that Europeans react more heavily on inequality than Americans.

However, the previous researches bear a few weaknesses. First, an external validity issue exists. Only the results for one European country are investigated, and results might differ among different countries. Second, there is no control country added to the models. A control group adds power to the model, since systematic different patterns in SWB may exist between age cohorts. By adding a control country, these heterogeneities are accounted for. Nevertheless, the authors claim that their model is strong, because no other policy changes between the age cohorts have been implemented. Additionaly, finding a control country is hard, because each country has a unique pension system. On top of that, a good control country should not have similar changes to the pension system as the country of interest. However, many countries have implemented pension reforms at similar points in time. To be able to compare the SWB of two different countries, it is important that the pension schemes are similar in the time period before the policy change, where only one country-age cohort experiences a policy change.

In this section, the limited empirical research on the effects of pension reforms on SWB of older working people is summarized. The evidence is limited in terms of articles as well as in terms of reliability. In the next section, the research case of this study will be introduced.

2.5. Research Case

This thesis will study the changes in SWB in Denmark, where Belgium will be implemented as a control country. In the following section, a brief summary will be given about the pension system and pension reforms in both countries. Furthermore, the hypotheses will be formulated.

2.5.1. Danish Pension System

The Danish pension system consists of three pillars: state pension; semi-mandatory labor pensions; and private pensions (Andersen J. G., 2016). The state pension consists of a flat standard amount, and some supplements exists for the poor and middle class. The labor market pensions have seen a shift from defined benefits towards defined contribution, which is a worldwide trend. The labor market pensions have grown in importance. Denmark experienced a deficit in state pensions, but since the 90s this has changed to a permanent surplus. The

replacement rate for lower income groups is close to 100%, which leads to low poverty among the elderly (Andersen T. M., 2015).

In December 2011, Denmark passed a law which will gradually increase the state pension age from 65 to 67 years between 2019-2022 for both men and women (ETUI, 2017). Each January 1st and each July 1st, the state pension age rises with six months. People born in the first half of 1954 are the first age cohort who are affected by the pension reform and will reach state pension age at 65.5 years old (Udbetaling Danmark, 2021). People born between 1 July 1955 and 31 December 1962 reach state pension age at 67. Before the reform, older workers could enter a voluntary early retirement scheme up to five years before the retirement age. Afterwards, this was reduced to three years, and it was made so unattractive that almost nobody uses it anymore (Andersen J. G., 2016).

The reforms in Denmark in terms of retirement age are radical compared to other countries (Andersen J. G., 2016). As a result, the old age dependency ratio within the country will remain similar over the coming decades. Denmark scores third best on the Global Pension Index (GPI), partly caused by a high sustainability score (Mercer & CFA Institute, 2022). Inequality and poverty are not expected to increase in the future with the current system (Andersen J. G., 2016).

This research aims to expand the knowledge on the effects of a pension reform on SWB of the working population. As previously mentioned, SWB consists of three different categories. These are hedonic, evaluative and eudaemonic. Within the lines of this study, the effect on evaluative SWB is expected to be influenced the most. Hedonic SWB can fluctuate a lot during short periods of time, and many measurements on regular basis are not available. Moreover, eudaemonic happiness does not fluctuate a lot after big life events, while evaluative happiness does (Pavot & Diener, 2009). One of the papers presented significant effects on job satisfaction of older workers after a pension reform (Montizaan & Vendrik, 2014). By measuring the same variable of interest used in other studies, knowledge on the topic is extended. Therefore, the effects job satisfaction will be further analyzed. Furthermore, since job satisfaction is a domain satisfaction of life satisfaction, this study will also investigate whether life satisfaction is significantly affected after a pension reform. Following the radical changes in retirement age in Denmark and the limited current findings on the effects of SWB, the next hypotheses on evaluative happiness are formulated:

Hypothesis 1a: Danish people born between 1954-1959 experience a larger fall in job satisfaction compared to Danish people born between 1950-1953.

Hypothesis 1b: Danish people born between 1954-1959 experience a larger fall in life satisfaction compared to Danish people born between 1950-1953.

Another article which analyzed the well-being effects of pension reforms in the EU has found significant effects on depression of older workers after a pension reform (Grip et al., 2012). By analyzing the effects of depression in this study as well, a better understanding on this side of SWB can be found as well. Moreover, another issue of SWB research is tackled by also analyzing changes in depression levels. Diener & Seligman (2004) namely stated that a main issue with SWB research is the unsystematic nature of current findings and measures. They claim that true understanding of SWB cannot be attained if life satisfaction and depression analysis is not provided within the same study. Following the findings in current literature, the next hypothesis on depression is formulated:

Hypothesis 2: Danish people born between 1954-1959 experience a larger rise in depression level satisfaction to Danish people born between 1950-1953.

2.5.2. Belgian Pension System

Identical to Denmark, Belgium uses the same three pillars of the pension system (Onelife, 2022). As per 2009, both men and women reach the statutory retirement age at 65 (OECD, 2009). Earlier, the retirement age for women was 62, but it has been gradually risen to reach gender equality. One of the main differences between the two countries is that Belgium has a much more flexible early retirement scheme compared to Denmark. With 44 years of working experience, a Belgian worker may retire at 60 years of age (Expatica, 2021). In 2008, only a third of people over 55 were still working, but this number has increased to 50% in 2022 (APG, 2022). This was due to government efforts to increase pension taxes on early retirement (Expatica, 2021).

Comparable to Denmark, Belgium is planning to increase their state pension age. However, in Belgium the state pension age will increase to 66 as late as January 1st, 2025 (Overheid Vlaanderen, 2022). This means that people born in 1960 are the first age cohort affected by the reform.

In the GPI, Belgium has similar scores in adequacy and integrity of their pension system as Denmark (Mercer & CFA Institute, 2022). On the sustainability part, Belgium scores lower than Denmark. This is a result of a slow increase in retirement age and low percentages of older people that are still working. Therefore, Belgium only holds 17th place on the GPI.

For the second set of hypotheses, the same types of SWB will be analyzed as in hypotheses 1 and 2. Belgian women experienced an increase in pension age during the period that will be analyzed in this study. This might infer with the outcomes, and therefore only the male respondents for both countries will be studied for this part of the research. Following the delayed increase of the retirement age in Belgium as compared to Denmark, these hypotheses are formulated:

Hypothesis 3a: Danish men born between 1954-1959 experience a larger fall in job satisfaction compared to Danish men born between 1950-1953 and the fall is larger than in Belgium, where no increase in pension age was implemented.

Hypothesis 3b: Danish men born between 1954-1959 experience a larger fall in life satisfaction compared to Danish men born between 1950-1953 and the fall is larger than in Belgium, where no increase in pension age was implemented.

Hypothesis 4: Danish men born between 1954-1959 experience a larger rise in depression level compared to Danish men born between 1950-1953 and the rise is larger than in Belgium, where no increase in pension age was implemented.

3. Data

The following section will provide a description of the data source and the sample selection. The data used was retrieved from the SHARE database (Börsch-Supan, 2022). The EC has had strong interest in gathering scientific evidence on the elderly population for the last decades, since the old age dependency ratio is expected to reach the highest levels in the history of humankind in the coming decades. SHARE was therefore created by the EC in 2002 to compose a longitudinal database on the elderly European population. Since 2004, SHARE gathers survey data from European and Israelian civilians older than 50 and their spouses, regardless of their age. Providing micro-level data on economic, social and health factors that accompany and influence ageing processes at the individual and societal levels is the main goal (Börsch-Supan, et al., 2013). In total, 29 countries participate in the project and over 140,000 individuals are included in the data. Approximately once every two years, a new wave is added. In April 2022, 8 SHARE waves were in existence. Like most panel studies, SHARE suffers from attrition, which is why a refreshment sample was added in each wave, apart from wave 3. Wave 3 is different from other waves since this wave only includes retrospective interviews. Data on early life may provide better understanding on the current circumstances of older people. A large downside for longitudinal research is that no data was gathered on current life circumstances

within wave 3. Since data on SWB is not included in wave 3, no data from that wave will be used for this research. The gap between wave 2 and wave 4 is four to five years, depending on when the respondents were interviewed.

To analyze the effects of an increase in pension age on SWB of older workers first affected by the policy, a sample of Belgian and Danish people born between 1950 and 1959 will be used. People who permanently left the labor market before the announcement of the pension reform are excluded from the data, since this research focuses on the effects of SWB after increasing the pension age on the working population.[1] As a result, 13,404 observations from 4,092 unique individuals remain in the sample.

4. Methodology

The following part will cover the relevant variables gathered from the SHARE database, followed by the statistical analysis techniques that will be used to test the hypotheses. Then, the descriptive statistics are discussed.

4.1. Variables

4.1.1. Dependent Variables

The first dependent variable is job satisfaction, which is a domain satisfaction of life satisfaction. This means that job satisfaction partly determines life satisfaction, along with other domain satisfactions (e.g., satisfaction with social relationships). Both of these measurements are part of evaluative SWB. Within the SHARE surveys, job satisfaction is an ordered categorical variable with a score range from 1 to 4 (strongly agree; agree; disagree; strongly disagree).

The second dependent variable used in this study is life satisfaction. Within the SHARE surveys, respondents are asked how satisfied they are with life on a scale from 0 to 10. A score of 0 is given when the respondent is completely dissatisfied with life, and a score of 10 represents complete life satisfaction. The scores 0 until 3 have an accumulated frequency of around 1%. Even though a significant effect might exist between increasing the pension age and life satisfaction, it might not be found with logistic regression analysis if groups with very low frequencies exist. Therefore, the scores of 0 until 3 have been combined to a score of 1 for the analysis of the data. A score of 4 has been recoded to a score of 2 and so forth. Afterwards, the life satisfaction scale used in this research has a range from 1 to 8, ranked from least to

[1] People who were permanently out of the labor market before wave 5 were excluded. For an explanation, see 4.1.2. Independent Variables.

highest life satisfaction. Dolan & Metcalfe (2011) argue that life satisfaction is a good evaluation measurement of SWB for policy review. Using life satisfaction as a dependent variable leads to omission of wave 1, because SHARE started gathering data for that variable in wave 2.

The last dependent variable used in this study is the EURO-D depression scale. The scale is an EU generated initiative to compare depression symptoms of elderly within European countries (Prince et al., 1999). Within Europe, different types of scales are used to determine depression of elderly people. The EURO-D depression scale is created as a European universal internally valid scale of depression. The scale consists of twelve depressive symptoms, which provides an aggregated depression score, ranging from 0 to 12. The symptoms used are appetite; concentration; depression; enjoyment; fatigue; guilt; interest; irritability; pessimism; sleep; suicidality; and tearfulness (Maskileyson et al., 2021). If an individual has low appetite, bad sleep, a high level of tearfulness, and no other depressive symptoms, they score 3 on the Euro-D depression scale. A score of 0 represents "not depressed", and a score of 12 represents "very depressed". As can be observed, some symptoms are phrased positively (e.g., enjoyment) and others are phrased negatively (e.g., tearfulness). A score of 1 is given to a respondent with low levels of enjoyment, while a score of 1 is given to a person with high levels of tearfulness. The scores 8 until 12 have been combined to a score of 8, because of the low proportion of respondents falling in those categories (approximately 1% accumulated). Afterwards, the depression scale has a range from 0 to 8, ranked from least to most depressed. Earlier research found a higher depression level among the first age cohort affected by higher pension age in the Netherlands (Grip et al., 2012). Depression belongs to hedonic SWB. Diener & Seligman (2004) emphasize the importance of different kinds of measurements of SWB within a study, because different a policy change can result in different patterns across SWB measures (Lucas et al., 1996). Consequently, this study consists of three different dependent variables.

4.1.2. Independent Variables

The independent variables vary between hypotheses. The following set of hypotheses have the same independent variable:

Hypothesis 1a: Danish people born between 1954-1959 experience a larger fall in job satisfaction compared to Danish people born between 1950-1953.

Hypothesis 1b: Danish people born between 1954-1959 experience a larger fall in life satisfaction compared to Danish people born between 1950-1953.

Hypothesis 2: Danish people born between 1954-1959 experience a larger rise in depression level satisfaction to Danish people born between 1950-1953.

For these hypotheses, only the Danish respondents in the sample will be analyzed. The independent variable will be an interaction term between birth cohort and period. Birth cohort is a binary variable, which has a value of 0 for people born between 1950-1953, and a value of 1 for people born between 1954-1959. The cohorts are divided at this point because people born in 1954 are the first who have to work until a later age due to the Danish pension reform. Period is a dummy variable as well, with a value of 0 for waves 1-4, and a value of 1 for waves 5-8. In December 2011, the Danish government announced the reform in the pension system. In this study, it is assumed that rumors of an increase in pension age have had no effect on SWB. Rather, the effect only happened after the policy was confirmed. SHARE data from wave 4 in Denmark was gathered in 2011, and fieldwork was finished before December (Börsch-Supan, et al., 2013). Hence, the second period starts at wave 5. Wave 1 fieldwork was conducted throughout 2004 in Denmark, and throughout 2004 and 2005 in Belgium. The end of the first period is at wave 4, where fieldwork was done throughout 2011 for both countries. Period one thus comprises of six to seven years. The second period lasts from wave 5 until wave 8; or from 2013 until 2019/2020. Both periods consist of six to seven years.

Table 1 shows the percentual distribution of current job situation per wave. As can be observed, in wave 1-4 there are no observations for people in the categories retired; permanently sick; or homemaker. This study focuses on SWB effects of the Danish pension reform on the population that is active on the labor market. Therefore, individuals who were permanently out of the labor force before wave 5 were excluded from the data. If a respondent permanently left the labor force after wave 5, their observations were kept. The reasoning behind this, is that people who left the labor force after the policy change could still provide useful data. For instance, a Danish individual who was active in the labor force between wave 1 and 6 and stopped being active in the labor force afterwards, has been affected by the policy, and should therefore not be excluded from the sample. In total, there are 13,046 observations of current job situation.

The following set of hypotheses have a common independent variable as well, which differs from the previous one.

Table 1: Percentual distribution of job situation through different waves

	Wave								
Current Job Situation	1	2	4	5	6	7	8	Total	Observations
Retired				2.51	4.71	6.87	5.60	19.68	2,568
(Self-)employed	5.82	8.32	13.81	15.51	12.30	9.16	3.67	68.60	8,949
Unemployed	0.62	0.68	1.58	1.59	1.29	0.84	0.24	6.84	892
Permanently sick				1.05	1.27	0.92	0.34	3.58	467
Homemaker				0.37	0.45	0.37	0.11	1.30	170
Total	6.44	9.01	15.39	21.02	20.03	18.16	9.96	100.00	13,046

Hypothesis 3a: Danish men born between 1954-1959 experience a larger fall in job satisfaction compared to Danish people born between 1950-1953 and the fall is larger than in Belgium, where no increase in pension age was implemented.

Hypothesis 3b: Danish men born between 1954-1959 experience a larger fall in life satisfaction compared to Danish men born between 1950-1953 and the fall is larger than in Belgium, where no increase in pension age was implemented.

Hypothesis 4: Danish men born between 1954-1959 experience a larger rise in depression level compared to Danish men born between 1950-1953 and the rise is larger than in Belgium, where no increase in pension age was implemented.

For these hypotheses, the independent variable is an interaction term between birth cohort, period, and country. Compared to the previous set of hypotheses, country is added to the model. Denmark is the treatment country, because it is the country where an increase in pension age was implemented. Belgium is used as a control, because no similar policy exists for the same age cohort for Belgian men. This part of the study only investigates differences in the male population, because Belgian women were affected by an increase in pension age during the early stages of the research period. Belgium is a fitting control since Denmark and Belgium had similar pension systems before the Danish reform. Moreover, the male population in Belgium did not experience a similar change in pension age.

4.1.3. Control Variables

Several control variables are added to the model. These controls all have empirical backup to support that they significantly impact SWB. The first covariate is the categorical variable self-rated health (excellent; very good; good; fair; and poor). In previous literature, it is found that poor physical health has an advert relationship with mental health (Easterlin, 2004). Moreover, deterioration of health appears to have a permanent negative effect on SWB (Easterlin, 2003).

Another confounding factor is marital status, a binary variable with value 0 if the respondent is 'never married,' 'divorced,' or 'widowed'. A value of 1 is given to respondents who are either 'married, living with spouse,' 'married, not living with spouse,' or 'registered partnership'. Previous studies have shown that married people have higher levels of SWB (Helliwell, 2003). Compared to being widowed or separated, Blanchflower and Oswald (2000) found that a lasting marriage is worth $100,000 per year. Moreover, big life events, such as divorce or widowhood have a significant effect on happiness (Pavot & Diener, 2009).

An additional variable that influences SWB is social capital (Bartolini & Sarracino, 2014; Diener & Seligman, 2004). Social capital consists of multiple aspects, such as trust in other people; community activity (voluntary work or club membership); government trust. The first variable to control for social capital is trust, a categorical variable with scores ranging from 0 to 10. A score of 0 represents a minimum level of trust, and 10 represents the maximum. Helliwell (2003) found that a higher level of trust could lead to higher levels of SWB. The second social capital covariate is the binary variable club membership, where value 0 represents "not a member of a club" and value 1 represents "member of a club". Communities with higher levels of voluntary work and club membership experience higher levels of SWB (Putnam, 2000). The last part of social capital is government trust. In this study, the perceived chance by the respondent whether the government will reduce the respondent's pension will be used as a proxy for government trust. It is a discrete variable, ranging from 0 to 100 per cent chance. Countries with better political institutions, report higher levels of SWB (Frey & Stutzer, 2000).

A further factor which influences SWB is religion. Overall, religious people tend to be happier than non-religious people (Diener et al.; 1999 Ferriss, 2002). As a proxy of religion, the categorical variable "frequency of praying" will be used (more than once a day; daily; multiple times a week; once a week; never).

Also, on the individual level, household income has a positive relationship with SWB (Easterlin, 1974). Though, the benefits in SWB are diminishing with higher income. After a certain treshold, virtually no increases in SWB can be found with a higher level of income (Helliwell, 2003). Income is a variable that commonly suffers from item non-response within survey data (Moore, et al., 2000). SHARE experiences the same issue, however they use the fully conditional specification (FCS) method of van Buuren et al. (1999) to impute missing observations. To account for non-normality within this variable, a logarithmic transformation will be implemented.

Finally, it is expected that current job situation is a confounding factor within this research. The different categories of current job situation are presented in Table 1. An increase

in pension age is expected to have negative SWB effect on the people who are in the labor force. Someone could for instance retire at an earlier age than the state pension age. Then the effect of increasing the state pension age should have little to no effect on that individual. For the hypotheses where the dependent variable is job satisfaction this control is omitted since all respondents in those tests fall in the category (self-)employed.

4.2. Descriptive statistics

A summary of the dependent variables is presented in Table 2. Job satisfaction has a mean score of 1.538, where 1 is the best score and 4 is the worst. The mean score of life satisfaction is 6.138, where the best score is 8, while the combined worst scores are combined to value 1. The mean of Euro-D depression score is 2.058 where 0 is the best score and the worst scores are combined to value 8. In terms of SWB, the respondents on average fall in the better categories.

The descriptive statistics of the control variables are shown in Appendix A. Religion is measured in frequency of praying, and 22.39% of the people that responded prays at least once a week. A further 19.11% of respondents pray less than once weekly, and 58.50% never prays. Furthermore, of the observations, 46.2% is Danish and 54.8% is Belgian. Moreover, 47.7% of observations is male, and 53.3% is female. SHARE provides no explanation for the higher response rate for females compared to males. However, previous research has found that women are more likely than men to participate in surveys (Curtin et al., 2000). The mean birthyear is 1955, indicating no severe skew towards one end of the sample. Of the observations, 69.2% of interviews were conducted in the second period. The skew in observations can partly be explained by the omission of the third wave.[2] The mean of self-rated health is 2.55, where 1 is the best score and 5 is the worst. Therefore, most people fall in the better categories of health. Around 75% of respondents is married, and approximately 45% of respondents is a member of a club. Mean household income is 56,688 euros, with a minimum of 8,400 and a maximum of 772,221. Through the high standard deviation of 64,677 and to diminish the power of outliers, a logarithmic transformation will be used to analyze the data. Government trust experiences a relatively high rate of item non-response, with 5,159 observations. The mean level of trust is 6.492 on a scale from 0 to 10. The mean level of government trust is 51.2%.

For government trust, the proxy used was the perceived chance that the government will reduce the pensions. The answer to this question can vary among respondents, regardless of their overall trust in the government. The proxy for religion was frequency of praying, but a person who never prays is not non-religious by definition.

[2] See 3. Data for an explanation

Table 2: Descriptive statistics of dependent variables

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
Job satisfaction	6,195	1.538	0.648	1	4
Life satisfaction	12,548	6.138	1.351	1	8
Depression level	11,640	2.058	2.005	0	8

4.3. Model Specification

A common method to measure policy effect is DD analysis (Dimick & Ryan, 2014). For a DD estimation, an interaction term should be implemented which divides the control group for the treatment group and the first period for the second period. The interaction term for hypotheses 1 and 2 is between cohort and period. The standard DD model, while letting vector *X* represent a constant term including the control variables, appears as follows:

$$y = X\beta + \beta_1 Period + \beta_2 Cohort + \beta_{12}(Period \times Cohort) + u \quad (1)$$

Where u denotes an error term. Furthermore, expected outcomes appear as follows:

$$E(y|X, Cohort = 1, Period = 1) = X\beta + \beta_1 + \beta_2 + \beta_{12}$$
$$E(y|X, Cohort = 1, Period = 0) = X\beta + \beta_2$$
$$E(y|X, Cohort = 0, Period = 1) = X\beta + \beta_1$$
$$E(y|X, Cohort = 0, Period = 0) = X\beta$$

Letting *x* denote the vector of covariates, the difference in *E(y|x)* from the pretreatment *(Period=0)* to the posttreatment *(Period=1)* for the treated group *(Born between 1955 and 1959)* is $\beta_1 + \beta_{12}$. For the control group *(Born between 1950 and 1954)*, the difference in *E(y|x)* between the pretreatment to the posttreatment period is β_1. What follows is that the differences in *E(y|x)* between the treatment and control groups from the pretreatment to the posttreatment period is β_{12}. Then, β_{12} is an estimate of the policy effect on the treated.

However, the standard DD model is a linear regression method. The outcome variables in this study are categorical, and therefore the propensity scores for each have to be calculated. Using standard linear regression methods would allow the outcome variable to exceed their numerical limit of 0-1. Hence, a nonlinear model should be implemented. Given the ordered nature of the dependent variables in all hypotheses, an ordered logit model following a DD structure as proposed by Karaca-Mandic et al. (2012) will be used in this study.

In this model, let the conditional probability that $y = 1$ be expressed as a function as the same linear index shown in equation (1). Then, the nonlinear DD model can be written as follows:

$$P(y = 1|x) = F(X\beta + \beta_1 Period + \beta_2 Cohort + \beta_{12}(Period \times Cohort)) \qquad (2)$$

And the same logic for DD outcomes can be applied:

$$P(y = 1|X, Cohort = 1, Period = 1) = F(X\beta + \beta_1 + \beta_2 + \beta_{12})$$
$$P(y = 1|X, Cohort = 1, Period = 0) = F(X\beta + \beta_2)$$
$$P(y = 1|X, Cohort = 0, Period = 1) = F(X\beta + \beta_1)$$
$$P(y = 1|X, Cohort = 0, Period = 0) = F(X\beta)$$

In these equations, the parameter β_1 allows the linear index to be different for all subjects in the period after treatment compared to the pretreatment period. Furthermore, β_2 allows that the linear index can be different for treatment subjects compared to the control group. Finally, β_{12} allows the linear index to vary in the posttreatment period, and thus the conditional probability that $P(y = 1|x)$ to vary over and above the difference attributable to the nonlinearity of the model for respondents in the treatment group versus the control group. This additional difference in the differences will provide a measure of the policy effect on the treated.

To test hypotheses 3 and 4, a linear DDD method will be implemented. DDD is an extension of the DD method and tackles one of its issues. One of the major limitations of DD is the fact that changes in SWB between the treatment group and the control group between periods could be systematically different through other causes than increasing pension age (Angrist & Pischke, 2008). To control for potential systematically different trajectories, Belgium is added as a control country. Belgium is a fitting control, because for the male population it had a similar pension system as Denmark before the Danish pension reform. Furthermore, the Belgian age cohort 1955-1959 experienced no change in pension age during both periods. In this model, only the male respondents will be used.[3] A linear DDD is chosen rather than a logistic implementation, due to the lack of empirical implementation of the method. The DDD model relaxes the assumption that a parallel trend between the birth cohorts should exist for the SWB measures before the pension reform announcement. Moreover, multiple previous research on SWB have used linear regression methods to analyze effects on

[3] See 2.5.2. Belgian Pension System

outcomes such as life satisfaction and depression (Della Giusta et al., 2011; Diener & Diener, 2009; Oishi et al., 2009).

The mode uses similar logic as equation (1), with an extra interaction *Country* (D) added. D is chosen as abbreviation because the variable *country* has a value of 1 if the country is Denmark. Then, the nonlinear DDD model can be written as follows:

$$y = X\beta + \beta_1 Period + \beta_2 Cohort + \beta_3 D + \beta_{12} Period \times Cohort + \beta_{13} Period \times Cohort + \beta_{23} Period \times D + \beta_{123} Period \times Cohort \times D + v \quad (3)$$

Following this formula, the subsequent expected outcomes appear:

$$E(y|X, Cohort = 1, Period = 1, D = 1) = X\beta + \beta_1 + \beta_2 + \beta_3 + \beta_{12} + \beta_{13} + \beta_{23} + \beta_{123}$$
$$E(y|X, Cohort = 1, Period = 1, D = 0) = X\beta + \beta_1 + \beta_2 + \beta_{12}$$
$$E(y|X, Cohort = 1, Period = 0, D = 1) = X\beta + \beta_1 + \beta_3 + \beta_{13}$$
$$E(y|X, Cohort = 0, Period = 1, D = 1) = X\beta + \beta_2 + \beta_3 + \beta_{23}$$
$$E(y|X, Cohort = 1, Period = 0, D = 0) = X\beta + \beta_1$$
$$E(y|X, Cohort = 0, Period = 1, D = 0) = X\beta + \beta_2$$
$$E(y|X, Cohort = 0, Period = 0, D = 1) = X\beta + \beta_3$$
$$E(y|X, Cohort = 0, Period = 0, D = 0) = X\beta$$

The outcome of the model is equivalent to the difference between the two DD. The first DD is between Period and Cohort, as in equation (1). The second DD is that between Period and Denmark.

5. Results

In this section, the results of the statistical analyses will be presented. Before tests are conducted, it is important to select the best fitting model. Therefore, this chapter starts with the testing of model assumptions. Afterwards, for the different logistic models, likelihood ratio (LR) tests are performed to test whether control variables add strength to the model. If a control variable has no significant impact, or omits a large part of the sample, they will be excluded from the final model. Then the results are presented, and the hypotheses will be answered. Finally, to assess whether the results are robust, some minor changes in models are executed to test if outcomes differ between them.

5.1. Assumption and control variable testing

Each statistical method has assumptions that need to hold to ensure that the model choice is valid. The first assumption of ordered logistic regression is that the outcome variables are ordered by nature (e.g., from worst to best). All three dependent outcomes within this study have an ordered nature. The second assumption is that one or more of the independent variables are either continuous, categorical or ordinal, which is met as well.

Another assumption states that no multicollinearity should exist (Senaviratna & Cooray, 2019). Before testing for multicollinearity, the set of variables within the models needs to be determined first. To determine which control variables significantly impact the model, LR tests for all the different controls are conducted. The results of the LR tests are represented in Table 3. For the model with Job satisfaction as outcome, marital status has no significant impact. Therefore, the control is excluded from the model. For the other outcome variables, marital status is significant and is left in the models. Furthermore, some controls that do significantly impact the model are still omitted from the models since they heavily reduce the number of observations in the sample. The variables religion, trust, and government trust all dropped more than half of the total observations for all dependent variables. For the remaining variables, multicollinearity tests are performed. The most common threshold for multicollinearity is $|r|<0.7$ (Dormann, et al., 2013). Though, more relaxed thresholds of 0.85 have been used in the past as well (Elith, et al., 2006) The correlations with job satisfaction as dependent variable are presented in Appendix B, Table B1. The highest correlation within the table is 0.815, which normally is alarmingly high. However, this is the correlation between a double interaction and a triple interaction (period x country with period x cohort x country). The triple interaction is a function of the double interaction variable, and therefore a high correlation between such variables is common. This does thus not necessarily imply that a correlation problem exists. For all variables that are not a function of each other, correlations remain well below 0.5. The highest correlation for those variables is 0.310 (country with household income). The dependent variable job satisfaction has the highest correlation with self-rated health, with a value of 0.179. For the other two dependent variables, a similar trend can be observed. For life satisfaction and depression, the highest correlations are also with self-rated health (-0.359 and 0.366 respectively). Moreover, correlations higher than 0.7 are only found between variables that are functions of each other, and always remain below 0.8.

Table 3: Likelihood ratio test for control variables

Variables	Job Satisfaction 3,165 obs.	Life Satisfaction 5,860 obs.	Depression Level 5,501 obs.
Self-Rated Health	0.000***	0.000***	0.000***
	[3,165]	[5,860]	[5,501]
Married	0.640	0.000***	0.000***
	[3,165]	[5,860]	[5,501]
Trust	0.000***	0.000***	0.000***
	[1,037]	[2,340]	[2,340]
Club Membership	0.001***	0.000***	0.000***
	[3,161]	[5,824]	[5,464]
Government Trust	0.000***	0.000***	0.000***
	[1,322]	[2,003]	[2,329]
Religion	0.000***	0.000***	0.000***
	[1,036]	[2,340]	[2,340]
Household income (log)	0.000***	0.000***	0.000***
	[2,850]	[5,850]	[5,156]
Gender	0.026**	0.000***	0.000***
	[3,165]	[5,850]	[5,501]
Job Status		0.000***	0.000***
		[5,696]	[5,351]

Note. obs. stands for the number of observations in the base model. The coefficients represent P-values. The codes for significance are as follows: *** indicates $p < 0.01$, ** indicates $p < 0.05$, * indicates $p < 0.1$. Number of observations after adding the control are reported below the likelihood ratio results within brackets.

For further analysis on multicollinearity, a variance inflation factor (VIF) test has been conducted for all hypotheses (see Appendix C). A frequently used threshold to divide small from large linear dependency of variables is a VIF value of 10 (Alin, 2010). In the VIF Table where job satisfaction is used as the dependent variable, and the DDD model is tested, two VIF values are above 10. This is worrying, and bias in the results could exist. The highest VIF value within the other tables is 7.16, which gives no indication of multicollinearity issues. From the results of the correlation matrices and the VIF Tables, it is concluded that the only model where a collinearity issue could exist, is the triple interaction model with job satisfaction as the dependent variable.

The proportional odds assumption is another assumption for the ordered logit model. It states that the relationship between outcome groups has to be the same. This implies that a parallel line should exist between outcomes. With the Wolfe Gould test for proportional lines, the null hypothesis which states that the slopes do not differ is tested. The alternative hypothesis is that the slopes are different. The null-hypothesis is rejected if $p > 0.050$, and then it is concluded that the proportional odds assumption is reasonable, meaning that the ordinal logistic regression can be used. However, when the null hypothesis cannot be rejected ($p < 0.050$), the

proportional odds assumption is violated, and a multinomial logistic regression is preferred. The outcomes of Wolfe Gould tests are presented in Table 4. For hypothesis 1a, the p-value is 0.192 > 0.050, meaning that the null-hypothesis is rejected. Thus, the proportional odds assumption is reasonable, and the ordered logit model is preferred. For hypotheses 1b and 2, the p-value is 0.000 < 0.050, meaning that the null-hypothesis cannot be rejected. Thus, the proportional odds assumption is violated, and the multinomial logit model is preferred instead.

Table 4: Wolfe Gould test of proportional lines

Hypothesis	Dependent Variable	P-value	Assumption Violated	Preferred Model
1a	Job Satisfaction	0.192	No	Ordered Logit
1b	Life Satisfaction	0.000	Yes	Multinomial Logit
2	Depression Level	0.000	Yes	Multinomial Logit

However, the multinomial logit model has a strict assumption which should be met. This is the assumption of independence of irrelevant alternatives (IIA). It requires that the inclusion or exclusion of categories does not affect the relative risks associated with the regressors in the remaining categories (Hausman & McFadden, 1984). A Hausman test is performed to determine whether this assumption holds and is presented Table 5. For both hypotheses, the assumption is violated. Since the structure of the outcome variables has a better fit with an ordered logit model, this model will be used for all three outcomes. A bias in the results for life satisfaction or depression as outcome is expected.

Table 5: Hausman test of independence of irrelevant alternatives

Hypothesis	Dependent Variable	P-value	Assumption Violated	Preferred Model
1b	Life Satisfaction	0.000	Yes	Ordered Logit
2	Depression Level	0.000	Yes	Ordered Logit

Additionally, for the DD method to give unbiased results, the assumption of parallel trends should be met. It requires that the difference between the ‘treatment’ and ‘control’ group is constant over time, in absence of treatment (Abadie, 2005). Thus, in the case of this study, the trend of the three outcome variables should be parallel before the Danish government implemented the change in pension age. The results of this assumption are graphically presented in Appendix D. For job satisfaction, though the lines are not exactly parallel before the break point at wave four, the direction and slopes are similar. First, an increase of the average score of job satisfaction is observed between 1 and 2. Between wave 2 and for, a downward trend is observed. For life satisfaction, the measurements start at wave 2 since wave 1 did not include

data on life satisfaction. The trends are both downward with a similar slope. With depression, for the birth cohort 1950-1953 an upward trend with a stable slope can be observed between periods 1 and 4. For the age cohort 1954-1959, the slope for depression becomes steeper after period 2. In all three figures, the slopes of the lines are in the same direction, and the differences in steepness are small. Therefore, no bias in the results is expected for any of the models. Moreover, in the DDD models, the assumption of parallel lines is further relaxed.

In the previous section, the assumptions for the different models were tested. Marriage was dropped as a control in the model with job satisfaction as dependent variable, because the variable did not significantly impact the model. Furthermore, trust, government trust and religion were dropped from all models because of the high item non-response rates for these variables. Additionally, for the triple interaction model with job satisfaction as dependent variable, the results may be biased due to a multicollinearity issue. Moreover, a bias in the results for life satisfaction and depression as outcome variables is expected since the assumption of proportional lines was violated. The issue could not be resolved by using a multinomial logit model due to violation of the IIA assumption. Finally, the assumption of parallel lines, which belongs to the DD method, was not violated in any of the models.

5.2. Logistic DD results

The results for the ordered logistic regression with job satisfaction as the dependent variable are presented in Appendix E. The coefficients are presented in odds ratios. The first column presents the base model, while the second column represents the full model. In the full model, worse levels of self-rated health leads to higher odds of falling in the higher categories of job satisfaction. This implies that, when controlled for club membership, household income and gender, people with worse health on average score are less satisfied with their job, significant at the 1% level. The interpretation might seem counterintuitive, but higher levels of job satisfaction represent lower levels of job satisfaction. Moreover, males have lower odds of falling in the higher categories of job satisfaction. This means that men are on average more satisfied with their jobs compared to women, significant at the 1% level. In both models, the interaction term between the second period and the cohort affected by an increase in pension age is above 1 and significant at the 5% level. This implies a significant DD effect towards the worse categories of job satisfaction, as expected.

To get a better understanding of the DD effects, the probability margins are presented in Table 6. For each of the outcomes, the predicted probabilities are calculated. To obtain the average DD effect, the difference within the treatment group (young cohort) between the

Table 6: Probability Margins - Job satisfaction

	Outcome			
Period x Cohort	Strongly Agree	Agree	Disagree	Strongly Disagree
First period x Old cohort	0.534***	0.401***	0.054***	0.011***
First period x Young cohort	0.526***	0.408***	0.056***	0.011***
Second period x Old cohort	0.624***	0.331***	0.038***	0.007***
Second period x Young cohort	0.529***	0.405***	0.055***	0.011***

Note. The codes for significance are as follows: *** indicates p < 0.01, ** indicates p < 0.05, * indicates p < 0.1.

periods has to be calculated first. For the outcome 'strongly agree,' this would mean 52.9 – 52.6 = an increase of 0.3 percentage points. Then, the difference within the control group between the time periods is calculated. In this case, we have 62.4 – 53.4 = an increase of 9.0 percentage points. Finally, the difference of the control group is subtracted from the treatment group. In this case, we get 0.3 – 9.0 = - 8.7. Then, the average DD effect of outcome 'strongly agree' is 8.7 percentage points. This is the DD estimate of the effect of the treatment. Table 7 shows the average DD in probability. As can be observed, the DD effect for 'strongly agree' is the same as the manually calculated outcome. The DD effect for 'strongly agree' is negative and is positive for the worse categories of job satisfaction. This is the effect direction that was hypothesized, and the results are significant at the 1% confidence level. The trend is a fall in probability for 'strongly agree,' and a rise in probability for the lower categories of job satisfaction for Danish workers born between 1954-1959. Therefore, hypothesis 1a, which stated: *Danish people born between 1954-1959 experience a larger fall in job satisfaction compared to Danish people born between 1950-1953.*, cannot be rejected.

Table 7: average DD in probability - Job Satisfaction

Outcome	DD	P-value
Strongly Agree	-0.087***	0.003
Agree	0.068***	0.003
Disagree	0.015***	0.004
Strongly Disagree	0.003***	0.006

Note. The codes for significance are as follows: *** indicates p < 0.01, ** indicates p < 0.05, * indicates p < 0.1.

The ordered logistic regression output for life satisfaction is presented in Appendix E and are presented in odds ratios. In the base model, the coefficient for the interaction term is 0.761, and significant at the 5% level. It implies a significant DD effect towards the worse categories of life satisfaction, as expected. However, no statistically significant result for the interaction can be found in the full model. Furthermore, the coefficient for married is 1.655, and significant at the 1% level. The coefficient implies that married people have 65.5% higher odds of being in the higher rather than lower categories of life satisfaction compared to non-

married people. Moreover, the coefficients for the categories of self-rated health are below one and are decreasing in magnitude for worse categories of health. This implies a trend that people with poorer health score lower on average on life satisfaction Moreover, men have lower odds of falling in the higher categories of life satisfaction compared to women, significant at the 1% level. Also, unemployed and permanently sick people have a higher risk of falling in the lower categories of life satisfaction compared to retired people, both significant at the 1% level.

For a better understanding of the DD effects for life satisfaction as outcome variable, the average DD in probability is presented in Table 8. Just as with the ordered logit model, the differences within the control group between periods are subtracted from the treatment group differences between periods. Table 1 in Appendix F shows the probability margins of life satisfaction. For a score of 8, the average DD effect on probability is -2.1 percentage points. The negative sign indicates a smaller rise in the highest level of life satisfaction for Danish people born between 1955 and 1959 compared to Danish people born between 1950 and 1954. This is the expected sign of the effect, but the effect does not statistically differ from zero. For scores 1-6, the DD outcomes are positive, implying a DD effect towards the worse categories of life satisfaction for people affected by a higher pension age. A trend can thus be seen, but this trend does not statistically differ from 0 Therefore, hypothesis 1b, which stated: *Danish people born between 1954-1959 experience a larger fall in life satisfaction compared to Danish people born between 1950-1953.*, is rejected.

Table 8: Average DD in probability - Life Satisfaction

Outcome	DD	P-value
1 (minimum life satisfaction)	0.000	0.567
2	0.000	0.558
3	0.002	0.536
4	0.001	0.513
5	0.005	0.472
6	0.013	0.330
7	-0.000	0.967
8 (maximum life satisfaction)	-0.021	0.323

Note. The codes for significance are as follows: *** indicates $p < 0.01$, ** indicates $p < 0.05$, * indicates $p < 0.1$.

Appendix E also shows the ordered logit regression results with depression as outcome. In neither model a significant result for the interaction term can be found, implying that depression score is not significantly affected by an increase in pension age. A trend can be observed that worse levels of physical health are associated with higher levels of depression. Moreover, being married results in 23% lower odds to fall in the worse categories of depression

compared to non-married people. Being a member of a club also significantly lowers the odds of falling in the higher categories of depression at the 1% level. An interesting result is that males on average have lower odds of falling in worse categories of depression compared to women. This result is remarkable, since men also have lower odds of falling in the better categories of life satisfaction compared to women, which contradicts each other. Additionally, higher income significantly lowers the odds of falling in the worse categories of depression at the 1% level. Similarly, people who are permanently sick are at a higher risk of falling in the higher levels of depression compared to people who are retired.

The average DD effects in probability are represented in Table 9. None of the DD effects are statistically significant. For a depression score of 0, the average DD effect is -0.4 percentage points. This indicates that Danish people born between 1954 and 1959 experience a smaller rise in a depression score of one compared to Danish people born between 1950 and 1953. This is the expected sign of the effect. For a score of 8, however, the sign is also negative. This effect is opposite from expected. In-between the scores of 0 and 8, some DD outcomes are negative, while others are positive, and most are close to 0. A DD pattern from higher to lower levels of depression cannot be found. Furthermore, none of the effects statistically differ from zero. Therefore, hypothesis 2, which stated: *Danish people born between 1954-1959 experience a larger rise in depression level satisfaction to Danish people born between 1950-1953.*, is rejected.

Table 9: average DD in probability - Depression

Outcome	DD	P-value
0 (not depressed)	-0.004	0.875
1	0.003	0.283
2	0.002	0.738
3	0.000	0.953
4	-0.000	0.965
5	-0.000	0.913
6	-0.000	0.874
7	-0.000	0.844
8 (highest levels of depression)	-0.000	0.819

Note. The codes for significance are as follows: *** indicates $p < 0.01$, ** indicates $p < 0.05$, * indicates $p < 0.1$.

5.3. Linear triple difference (DDD) results

Table 10 shows the regression results of the triple difference estimations. The first two columns show the results with job satisfaction as the dependent variable. Column 1 represents the base model without controls. In the base model, the triple interaction has a negative sign, but the result does not statistically differ from zero. The coefficient for Denmark is -0.170, indicating

that Danish male respondents score lower categories on job satisfaction compared to Belgian male respondents. This thus implies that Danish men born between 1950 and 1959 have higher levels of job satisfaction compared to Belgian men of the same birth cohort, statistically significant at the 1% level. The result might seem counterintuitive, but a score of 1 on job satisfaction represents 'strongly agree', while a score of 4 represents 'strongly disagree'. A negative sign thus implies a higher level of job satisfaction. The control variables are added in column 2. The coefficients of self-rated health are all positive, and the magnitudes for the lower categories are higher compared to the better categories of health. This implies that poorer levels of health are associated with the worse categories of job satisfaction for Belgian and Danish men born between 1950 and 1959. Moreover, the coefficient of household income is negative and significant at the 5% level, implying that on average, men with higher income are more satisfied with their job. The interaction young cohort x Denmark does not statistically vary from 0 in either model. This result contradicts the outcome of hypothesis 1a, but it could be caused by a multicollinearity issue within this model. The coefficient in the triple interaction is negative, implying that the fall in job satisfaction for the birth cohort 1950-1953 is lower in Denmark compared to Belgium. This result is opposite from expected, but it does not statistically differ from zero. Hypothesis 3a, which stated: *Danish men born between 1954-1959 experience a larger fall in job satisfaction compared to Danish men born between 1950-1953 and the fall is larger than in Belgium, where no increase in pension age was implemented.* is therefore rejected.

Columns 3 and 4 of Table 10 present the triple difference results with life satisfaction as the dependent variable. In both models, the coefficient for Denmark is positive, indicating that men in Denmark born between 1950 and 1959 score higher on life satisfaction than Belgian men from the same age. The result is significant at the 1% confidence level in both models. added to the model. With self-rated health the reference category is 'Excellent,' and all coefficients are negative and statistically significant the 1% confidence level. The magnitude of the coefficients increases towards worse levels of physical health, which implies that on average, Danish and Belgian men born between 1950 and 1959 with worse physical health score lower on life satisfaction. Moreover, married men within the sample score higher on life satisfaction on average compared to men who are not. Furthermore, for club membership a positive relationship is found with life satisfaction, significant at the 10% level. Similarly, household income has a positive effect on life satisfaction within the sample, which is significant at the 10% level. Compared to being retired, men who are either (self-)employed; permanently sick; or unemployed all score lower on life satisfaction on average.

Table 10: Triple difference estimator

	Job Satisfaction		Life Satisfaction		Depression Level	
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Triple difference						
Period 2	-0.132*	-0.057	0.051	0.046	0.056	-0.124
	[0.077]	[0.080]	[0.085]	[0.084]	[0.107]	[0.122]
Young cohort	-0.048	0.013	-0.133	-0.160*	0.125	0.102
	[0.041]	[0.048]	[0.089]	[0.083]	[0.106]	[0.116]
Denmark	-0.170***	-0.037	0.802***	0.622***	-0.267**	-0.114
	[0.045]	[0.055]	[0.102]	[0.095]	[0.118]	[0.134]
Period 2 x Young Cohort	0.191**	0.111	0.108	0.170*	0.073	0.133
	[0.087]	[0.090]	[0.105]	[0.099]	[0.137]	[0.144]
Period 2 x Denmark	0.045	-0.028	0.011	0.012	-0.166	-0.042
	[0.094]	[0.098]	[0.121]	[0.112]	[0.153]	[0.161]
Young cohort x Denmark	0.049	-0.016	0.086	0.057	-0.074	0.019
	[0.062]	[0.070]	[0.133]	[0.122]	[0.162]	[0.171]
Period 2 x Young Cohort x Denmark	-0.087	-0.023	-0.248	-0.195	-0.004	-0.168
	[0.111]	[0.115]	[0.155]	[0.143]	[0.203]	[0.206]
Self-rated health (Excellent ref.)						
Very good		0.107***		-0.278***		0.267***
		[0.035]		[0.047]		[0.069]
Good		0.220***		-0.533***		0.642***
		[0.037]		[0.048]		[0.071]
Fair		0.334***		-0.955***		1.628***
		[0.048]		[0.057]		[0.088]
Poor		0.326***		-1.733***		2.853***
		[0.097]		[0.097]		[0.156]
Married						
Yes				0.462***		-0.303***
				[0.037]		[0.057]
Club Member						
Yes		0.018		0.059*		-0.049
		[0.025]		[0.031]		[0.047]
Household income						
Household income (log)		-0.046**		0.053*		-0.045
		[0.021]		[0.028]		[0.040]
Job Situation (Retired ref.)						
(Self-)Employed				-0.091**		-0.027
				[0.045]		[0.070]
Unemployed				-0.603***		0.440***
				[0.074]		[0.114]
Permanently sick				-0.523***		0.731***
				[0.102]		[0.162]
Homemaker				-0.092		0.767
				[0.348]		[0.513]
Constant	1.626***	1.871***	5.783***	5.499***	1.740***	1.860***
	[0.031]	[0.234]	[0.072]	[0.296]	[0.080]	[0.431]
Observations	3,031	2,653	5,991	5,811	5,531	4,980
R-squared	0.017	0.042	0.084	0.231	0.019	0.178

Note. Ref. stands for reference category. The codes for significance are as follows: *** indicates $p < 0.01$, ** indicates $p < 0.05$, * indicates $p < 0.1$. Standard errors are reported beneath the regression estimates between brackets.

Finally, The coefficient of the triple interaction is negative, indicating that men born between 1954 and 1959 who live in Denmark show a stronger fall in life satisfaction compared to men from Belgium of the same age group. The effect does not statistically differ from zero,

however. Hypothesis 3b, which stated: *Danish men born between 1954-1959 experience a larger fall in life satisfaction compared to Danish men born between 1950-1953 and the fall is larger than in Belgium, where no increase in pension age was implemented.*, is therefore rejected.

Columns 5 and 6 of Table 10 represent the triple interaction results with depression level as the dependent variable. In the base model, men from Denmark and Belgium born between 1954 and 1959 are on average more depressed than men from the same countries born between 1950 and 1953. These results are statistically significant at the 5% level. In the full model however, the relationship does not statistically differ from zero. Column 6 shows that in the sample, men with worse levels of self-rated health score higher on the depression scale on average. Moreover, men who are married on average score lower on depression than men who are not married, at the 1% significance level. Household income has no significant effect on depression level. Unemployed or permanently sick score higher on average on the depression scale compared to retired men. The triple interaction has a negative sign, which implies a smaller increase in depression level for Danish men born between 1954 and 1959 compared to Belgian men from the same cohort. This effect is opposite from what was expected, but the results do not statistically differ from zero. Hypothesis 4, which stated: *Compared to Belgian men, Danish men in the age cohort 1955-1959 experience a higher increase in depression rates.*, is therefore rejected.

5.4. Robustness checks

For the DDD outcomes in the previous section, linear models were used. But since the outcome variables are categorical, linear regressions are not the most efficient models to analyze the effects of a policy. Linear models are chosen for the triple interaction because of the complexity of interpretation and the lack of empirical papers with a triple interaction in a logistic model. It is however still interesting to investigate whether the output of the linear regressions is similar to the outcomes of the logistic models. Therefore, ordered logistic regression has been performed for the triple interaction models and its output is presented in appendix G. As in the linear model, worse physical health has a positive relationship with job satisfaction at the 1% confidence level. This is concluded because the odds ratios are increasing with worse levels of physical health. It implies that worse health is linked with less job satisfaction. Moreover, household income has a negative relationship with job satisfaction, just as in the linear model. This can be concluded, because the odds ratio is 0.880, which is lower than 1. The relationship of the triple interaction with job satisfaction is below 1, which indicates a negative relationship,

though not statistically significant. All these results are similar to the linear model. Therefore, no reason is found to assume that the linear model gave results that differ from the logistic models. The coefficients from the linear model should however not be interpreted.

For life satisfaction, no significant effect can be found for the triple difference estimator in either model. Moreover, the trend that worse levels of physical health increase the odds of falling in the lower categories of life satisfaction is similar with the outcome of the linear model. An interesting difference in the models is that club membership appears to have no significant relationship with life satisfaction in the ordered logit model. In the linear model, the effect was positive and statistically significant at the 10% level. Overall, the results for both models are similar, and therefore it is concluded that the answer on the hypothesis is trustworthy.

Finally, the linear DDD model with depression level as outcome will be compared with the ordered logit DDD model. Both models show a negative, yet statistically insignificant DDD effect for depression. More similarities can be found with physical health, where worse levels of health leads to higher odds of falling in the worse categories of depression. Moreover, being married reduces the odds of falling in the worse categories of depression significantly in either model. The only difference in outcome can be found in household income, where a higher income significantly reduces the odds of falling in the worse categories of depression at the 10% level in the ordered logit model. In the linear model, on the other hand, no statistically significant relationship was found. Overall, the results of the models were very similar. It is therefore concluded that the linear models gave trustworthy results, and the hypothesis should still be rejected.

6. Discussion

The final part of this study will provide an overview on the research question, methodology and the results. Afterwards, the research question is answered. The results are then compared to findings in existing literature. Next, some limitations of the study are discussed. Finally, suggestions for further research and the SHARE database are presented.

This research analyzed the effects of an increase of the state pension age on SWB on older workers from Denmark. Different types of SWB have been defined within SWB literature, consisting of evaluative, hedonic and eudaemonic SWB. The effects on job satisfaction, life satisfaction and EURO-D depression scale were analyzed, where job satisfaction and life satisfaction fall into evaluative SWB. Previous literature argued that evaluative happiness fluctuates little over the course of a year and is therefore a suiting measurement for a study where respondents were interviewed approximately once every 2 years. Moreover, previous

literature has argued that different domains of SWB can react differently to big life events and policies and should therefore be studied within the same research. Hence, depression, which falls under hedonic SWB, was added as a variable of interest. Though, hedonic SWB fluctuates more over time and is thus less fitting in a longitudinal study with new observations added every few years. A sample of Belgian and Danish respondents who were born between 1950 and 1959 was used from the SHARE database. Danish people from the birth cohort 1954-1959 were confronted with a gradual increase of the state pension age from 65 towards 67 years old, while the pension age for the birth cohort 1950-1953 remained at 65 years old.

The research question of the study was:

How is SWB of older workers affected by an increase in the pension age?

Before answering the question, a summary of the results is provided and will be compared to previous findings. For job satisfaction, the probability to fall in the worse categories of job satisfaction was significantly larger for the birth cohort that was affected with the higher pension age at the 5% level. This finding is in line with previous research in The Netherlands. It extends the existing knowledge on the topic in twofold. First, the previous results appear to be externally valid, implying that the results are similar for populations in different countries. Second, the results appear to be steady over a longer period of time. The research in The Netherlands focused on very short-term effects, while this study found that DD effects on job satisfaction are significant over a period of more than five years. Furthermore, a DD trend was found towards the lower levels of life satisfaction. Nevertheless, the results were not statistically significant. This result implies that the negative DD effects on job satisfaction are not large enough to significantly deteriorate overall life satisfaction. For depression, the DD effects did not statistically differ from zero and no trend towards better or worse levels of depression could be found. This result is not in line with previous research in The Netherlands, where higher levels of depression were found among public sector workers who were affected by an increase in pension age, compared to unaffected people. It could be possible that an increase of depressive symptoms is a short-term effect of an increase in the pension age that recovers within a few years.

To relax the parallel line assumption, Belgium was later added to the model and a triple difference method was used. For job satisfaction, a negative coefficient was found for the triple interaction. This effect was opposed from the hypothesis, but the outcome did not statistically differ from zero. The effect also differed from the DD model outcome. A possible reason for

this could be a multicollinearity issue in the triple interaction model. For life satisfaction, the sign of the triple interaction was negative as hypothesized, but the effect did not statistically differ from zero. For depression, the sign of depression was negative and opposite from what was hypothesized in the base model as well as the full model. The triple difference effect did not statistically differ from 0 for depression either.

It is concluded that older Danish workers experienced a negative SWB effects following an increase of the state pension age in terms of job satisfaction. Towards other domains of SWB, no significant effect could be found. Most control variables showed similar results with previous research findings. For instance, people with better self-rated health, scored significantly better in all researched domains of SWB. Moreover, married people showed higher levels of life satisfaction and lower levels of depression. The answer to the research question is that job satisfaction has significantly deteriorated for older works in Denmark who are affected by an increase in the pension age. Life satisfaction and depression are not significantly affected.

The models used in this study bear some limitations. For instance, the study encompassed a fixed period of time. No distinctions were made between short- or long-term effects. Also, no distinctions were made between people who were self-employed, a public employee or a private firm employee. A further limitation of this study is the omission of wave 3 within the data. The gap between wave 2 and wave 4 consists of four to five years, leading to a large gap in data close to the policy change of the Danish government. It is therefore less certain that assumptions of parallel trends pre-treatment held for the different SWB variables. The triple interaction relaxed the assumption, however. Another limitation of this study is that some control variables were omitted because of high item non-response rates. Trust, government trust, and religion all have empirical support for being confounding factors of SWB, but individually all dropped over half of the observations. Lastly, the assumption of parallel lines was violated for the models where life satisfaction or depression were used as dependent variable, leading to possible bias in the results.

To conclude this thesis, some implications and recommendations will be provided. The findings of this study suggest that policy change can significantly deteriorate domains of SWB that are closely linked to the policy, and those effects may be felt for years. Therefore, it is important for lawmakers to consider the mental effects of a policy, or compensate the people affected that could diminish those effects. However, the existing literature on SWB effects of an increase in pension age is scarce. More research should be conducted, within Denmark as well as for other countries. For research within Denmark, it is recommended that short-term SWB effects are studied. By studying policy effects on SWB on the short-term, a better

understanding could be found in short-term shocks of SWB parameters after policy. Moreover, a study where birth cohorts consist of a fewer number of years might give different results. A sense of unfairness can be felt at a higher level for the very youngest people affected by a policy. For external validity purposes, it is important that similar studies are conducted for other countries. However, finding a suiting country that could be added as a control for a triple difference analysis might not always be possible. Furthermore, for SHARE it is recommended that new questions that could measure "government trust" and "religion" are formulated. Moreover, a more frequent gathering of data would improve the power of studies that use it.

REFERENCES

Abadie, A. (2005). Semiparametric Difference-in-Difference Estimators. *The review of economic studies*, 1-19.

Adler, A., & Seligman, M. E. (2016). Using wellbeing for public policy: Theory, measurement, and recommendations. *International journal of wellbeing*, 1-35.

Alesina, A., Di Tella, R., & MacCulloch, R. (2004). Inequality and happiness: are Europeans and Americans different? *Journal of public economics*, 2009-2042.

Alin, A. (2010). Multicollinearity. *Wiley Interdisciplinary reviews: Computational statistics*, 370-374.

Andersen, J. G. (2016). The Danish pension system. *Policy Network.*

Andersen, T. M. (2015). Robustness of the Danish pension system. *CEFifo DICE Report*, 25-30.

Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist's companion.* Princeton university press.

APG. (2022, February 11). *Belgian pension system is on a collision course.* Retrieved from APG: https://apg.nl/en/publication/belgian-pension-system-is-on-a-collision-course/

Arora, R. (2008, January 17). A Well-Being Report Card for President Sarkozy. Washington D.C. Retrieved February 25, 2022, from https://news.gallup.com/poll/103795/wellbeing-report-card-president-sarkozy.aspx

Bartolini, S., & Sarracino, F. (2014). Happy for how long? How social capital and economic growth relate to happiness over time. *Ecological economics*, 242-256.

Beets, G. (2007). *Wereld tot 2050 steeds grijzer*. Demos.

Blanchflower, D. G., & Oswald, A. J. (2004). Well-Being Over Time in Britain and the USA. *Journal of public economics*, 1359-1386.

Börsch-Supan, A. (2022). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 8. Release version: 8.0.0. Data set.

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., . . . Zuber, S. (2013). Data resource profile: the Survey of Health, Ageing and Retirement in Europe (SHARE). *International journal of epidemiology*, 992-1001.

Carone, G., Eckefeldt, P., Giamboni, L., Laine, V., & Pamies, S. (2016). Pension reforms in the EU since the early 2000's: Achievements and challenges ahead. *European economy discussion paper*.

Carrino, L., Glaser, K., & Avendano, M. (2018). Later pension, poorer health? Evidence from the new State Pension age in the UK. *Harvard Center for Population and Development Studies Working Paper Series*.

Conceição, P., & Bandura, R. (2008). Measuring subjective wellbeing: A summary review of the literature. *United nations development programme (UNDP) development studies, working paper*.

Curtin, R., Presser, S., & Singer, E. (2000). The effects of response rate changes on the index of consumer sentiment. *Public opinion quarterly*, 413-428.

Della Giusta, M., Jewell, S. L., & Kambhampati, U. S. (2011). Gender and Life Satisfaction in the UK. *Feminist economics*, 1-34.

Di Tella, R., & MacCulloch, R. (2006). Some uses of happiness data in economics. *Journal of economic perspectives*, 25-46.

Diener, E., & Diener, M. (2009). Cross-Cultural Correlates of Life Satisfaction and Self-Esteem. In E. Diener, *Culture and well-being: The collected works of Ed Diener* (pp. 71-91). Dordrecht: Springer.

Diener, E., & Seligman, M. E. (2004). Beyond money: Toward an economy of well-being. *Psychological science in the public interest*, 1-31.

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. E. (1999). Subjective wellbeing: Three decades of progress. *Psychological Bulletin*, 276–302.

Dimick, J. B., & Ryan, A. M. (2014). Methods for evaluating changes in health care policy: the difference-in-differences approach. *Jama*, 2401-2402.

Dolan, P., & Metcalfe, R. (2011). *Measuring subjective wellbeing for public policy: Recommendations on measures.* The Office for National Statistics .

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., . . . Lautenbach, S. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their peformance. *Eography*, 27-46.

Earterlin, R. A. (2010). Well-being, front and center: A note on the Sarkozy report. *Population and Development Review*, 119-124.

Easterlin, R. A. (1974). Does economic growth improve the human lot? Some empirical evidense. *Nations and households in economic growth*, 89-125.

Easterlin, R. A. (2003). Explaining Happiness. *Proceedings of the National Acacamy of Sciences*, 11176-11183.

Easterlin, R. A. (2004). The economics of happiness. *Daedalus*, 26-33.

Elith, J., Graham, C. H., Anderson, R. P., Dudík, M., Ferrier, S., Guisan, A., . . . Zimmerman, N. E. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 129-151.

ETUI. (2017, April 11). *Pension reforms in Denmark: background summary*. Retrieved from European Trade Union Institute: https://www.etui.org/covid-social-impact/denmark/pension-reforms-in-denmark-background-summary

Expatica. (2021, June 1). *A guide to the Belgian pension system.* Retrieved from Expatica: https://www.expatica.com/be/finance/retirement/pensions-in-belgium-762019/

Ferriss, A. L. (2002). Religion and the quality of life. *Journal of Happiness Studies*, 199-215.

Forgeard, M. J., Jayawickreme, E., Kern, M. L., & Seligman, M. E. (2011). Doing the right thing: Measuring wellbeing for public policy. *International journal of wellbeing*, 1-35.

Frey, B. S., & Stutzer, A. (2000). Happiness prospers in democracy. *Journal of Happiness Studies*, 79-102.

Grip, A., Lindeboom, M., & Montizaan, R. (2012). Shattered dreams: the effects of changing the pension system late in the game. *The Economic Journal*, 1-25.

Hausman, J. A., & McFadden, D. L. (1984). Specification tests for the multinomial logit model. *Econometrica*, 1219-1240.

Helliwell, J. (2003). How's life? Combining individual and national variables to explain subjective well-being. *Economic Modelling*, 331-360.

Kahneman, D. (1999). Objective Happiness. *Well-being: The foundations of hedonic psychology*, 1-23.

Kahneman, D., & Krueger, A. B. (2006). Developments in the measurement of subjective well-being. *Journal of Economic perspectives*, 3-24.

Kuznets, S. (1934). National Income, 1929-1932. In *National Income, 1929-1932* (pp. 1-12). NBER.

Layard, R. (2010). Measuring subjective well-being. *Science*, 534-535.

Lucas, R. E., Diener, E., & Suh, E. (1996). Discriminant validity of well-being measures. *Journal of personality and social psychology*, 616-628.

Maskileyson, D., Seddig, D., & Davidov, E. (2021). The EURO-D Measure of Depressive Symptoms in the Aging Population: Comparability Across European Countries and Israel. *Frontiers in Political Science*, 90.

Mercer, & Institute, C. (n.d.). *Global Pension Index 2021.* Retrieved February 28, 2022

Montizaan, R., & Vendrik, M. (2014). Misery loves company: Exogenous shocks in retirement expectations and social comparison effects on subjective well-being. *Journal of Economic Behavior & Organization*, 1-26.

Moore, J. C., Stinson, L. L., & Welniak, E. J. (2000). Income measurement error in surveys: A review. *Journal of Official Statistics-Stockholm*, 331-362.

OECD. (2008). *OECD Employment Outlook 2008.* Organisation for Economic Co-operation and Development.

OECD. (2009). Pension Country Profile: Belgium. In *OECD Private Pensions Outlook 2008* (pp. 156-172). OECD.

Oishi, S., Diener, E., Lucas, R. E., & Suh, E. M. (2009). Cross-Cultural Variations in Predictors of Life satisfaction: Perspectives from needs and values. In E. Diener, *Culture and well-being: The collected works of Ed Diener* (pp. 109-127). Dordrecht: Springer.

Onelife. (n.d.). *Het Belgische Pensioensysteem.* Retrieved February 28, 2022, from Onelife: https://www.onelife.com/our-solutions/european-presence/belgium/?lang=nl#:~:text=In%20Belgi%C3%AB%20wordt%20het%20wettelijk,minstens%2042%20jaar%20gewerkt%20hebben

Pavot, W., & Diener, E. (2009). Review of the satisfaction with life scale. In E. Diener, *Assessing Well-Being* (pp. 101-117). Dordrecht: Springer.

POST. (2012). *Measuring National Wellbeing.* London: POST.

Prince, M. J., Reischies, F., Beekman, A. T., Fuhrer, R., Jonker, C., Kivela, S. L., . . . Copeland, J. R. (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*, 330-338.

Putnam, R. (2000). *Bowling alone: The collapse and revival of American community.* New york: Simon ad Schuster.

Seaford, C. (2011). Time to legislate for the good life. *Nature*, 532-533.

Seligman, M. E., & Csikszentmihalyi, M. (2014). Positive psychology: An introduction. In M. Csikszentmihalyi, *Flow and the Foundations of Positive Psychology* (pp. 279-298). Dordrecht: Springer.

Senaviratna, N. A., & Cooray, T. M. (2019). Diagnosing multicollinearity of logistic regression model. *Asian Journal of Probability and Statistics,* 1-9.

Staubli, S., & Zweimüller, J. (2013). Does raising the early retirement age increase employment of older workers? *Journal of public economics*, 17-32.

Steptoe, A., Deaton, A., & Stone, A. A. (2015). Subjective wellbeing, health, and aging. *The Lancet*, 640-648.

Stratton, A. (2010, November 14). David Cameron aims to make happiness the new GDP. *The Guardian*.

United Nations, D. o. (2017). *World Population Ageing 2017 - Highlights.* New York.

Van Vliet, O., Been, J., Caminada, K., & Goudswaard, K. (2012). Pension reform and income inequality among older people in 15 European countries. *International Journal of Social Welfare*, 8-29.

Wolfers, J. (2003). Is business cycle volatility Costly? Evidence from surveys of subjective well-being. *International Finance*, 1-31.

Zencey, E. (2009, August 9). G.D.P. R.I.P. *New York Times*.

APPENDIX

Appendix A: Descriptive Statistics

Table A1: Descriptive statistics of Religion

Variable	Definition	Frequency	Percent
Religion	Praying: More than once a day	208	3.75
	Praying: Once daily	451	8.13
	Praying: A couple of times a week	304	5.48
	Praying: Once a week	279	5.03
	Praying: Less than once a week	1,064	19.11
	Praying: Never	3,245	58.50
Total		5,547	100

Table A2: Descriptive statistics of continuous, discrete, and binary variables

Variables	Number of Observations	Mean	Standard Deviation	min	max
Country	13,404	0.462	0.499	0	1
Gender	13,404	0.477	0.499	0	1
Birthyear	13,404	1,955	2.757	1950	1959
Period	13,404	0.692	0.462	0	1
Self-Rated Health	13,404	2.556	1.028	1	5
Married	13,404	0.746	0.435	0	1
Club Membership	13,263	0.449	0.497	0	1
Household Income	12,509	56,688	64,677	8,400	772,221
Trust	5,550	6.492	2.439	0	10
Government Trust	5,159	51.20	32.70	0	100

Appendix B: Correlation Tables

Table B1: Correlation Matrix - Job Satisfaction

Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
(1) Job satisfaction	1.000						
(2) Period	-0.008	1.000					
(3) Country	-0.106***	0.164***	1.000				
(4) Cohort	0.042*	0.164***	-0.090***	1.000			
(5) Period x Cohort	0.031	0.794***	0.080***	0.511***	1.000		
(6) Period x Country	-0.071***	0.676***	0.604***	0.054**	0.481***	1.000	
(7) Period x Cohort x Country	-0.034	0.544***	0.486***	0.351***	0.686***	0.805***	1.000
(8) Self-rated health	0.179***	0.010	-0.223***	0.037	0.045*	-0.103***	-0.057**
(9) Club member	-0.021	0.048*	0.142***	0.036	0.042*	0.131***	0.113***
(10) Household income (log)	-0.087***	0.115***	0.310***	-0.012	0.063**	0.240***	0.184***
Variables	(8)	(9)	(10)				
(8) Self-rated health	1.000						
(9) Club member	-0.139***	1.000					
(10) Household income (log)	-0.158***	0.109***	1.000				

Note. . The codes for significance are as follows: *** indicates p < 0.01, ** indicates p < 0.05, * indicates p < 0.1.

Table B2: Correlation Matrix - Life Satisfaction

Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
(1) Life Satisfaction	1.000						
(2) Period	0.028*	1.000					
(3) Country	0.295***	0.034**	1.000				
(4) Cohort	-0.056***	0.039**	-0.063***	1.000			
(5) Period x Cohort	-0.020	0.572***	-0.013	0.723***	1.000		
(6) Period x Country	0.231***	0.435***	0.790***	-0.019	0.214***	1.000	
(7) Period x Cohort x Country	0.140***	0.314***	0.571***	0.397***	0.550***	0.723***	1.000
(8) Self-rated health	-0.359***	0.075***	-0.224***	0.008	0.049***	-0.142***	-0.093***
(9) Married	0.199***	-0.023	0.019	-0.004	-0.012	0.007	0.017
(10 Club member	0.126***	0.032*	0.175***	0.022	0.029*	0.169***	0.139***
(11) Household income (log)	0.202***	0.003	0.329***	0.069***	0.061***	0.265***	0.232***
(12) Job status	-0.144***	-0.164***	0.020	0.236***	0.142***	-0.032*	0.101***
Variables	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**		
(8) Self-rated health	1.000						
(9) Married	-0.104***	1.000					
(10 Club member	-0.169***	0.045***	1.000				
(11) Household income (log)	-0.201***	0.245***	0.147***	1.000			
(12) Job status	0.089***	-0.060***	-0.064***	0.015	1.000		

Note. . The codes for significance are as follows: *** indicates p < 0.01, ** indicates p < 0.05, * indicates p < 0.1.

Table B3: Correlation Matrix – Depression level

Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
(1) Depression level	1.000						
(2) Period	-0.005	1.000					
(3) Country	-0.130***	0.054***	1.000				
(4) Cohort	0.068***	0.045**	-0.061***	1.000			
(5) Period x Cohort	0.045**	0.500***	-0.009	0.744***	1.000		
(6) Period x Country	-0.109***	0.477***	0.760***	-0.015	0.201***	1.000	
(7) Period x Cohort x Country	-0.040**	0.299***	0.477***	0.446***	0.599***	0.628***	1.000
(8) Self-rated health	0.366***	0.053***	-0.226***	-0.009	0.021	-0.142***	-0.091***
(9) Married	-0.121***	-0.019	0.011	-0.025	-0.022	-0.000	0.003
(10 Club member	-0.094***	0.029*	0.172***	0.024	0.026	0.162***	0.127***
(11) Household income (log)	-0.120***	0.020	0.308***	0.063***	0.064***	0.241***	0.183***
(12) Job status	0.133***	-0.168***	0.010	0.200***	0.138***	-0.044**	0.100***

Variables	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
(8) Self-rated health	1.000				
(9) Married	-0.104***	1.000			
(10 Club member	-0.173***	0.041**	1.000		
(11) Household income (log)	-0.186***	0.231***	0.137***	1.000	
(12) Job status	0.095***	-0.074***	-0.070***	0.001	1.000

Note. . The codes for significance are as follows: *** indicates $p < 0.01$, ** indicates $p < 0.05$, * indicates $p < 0.1$.

Appendix C: Variance Inflation Factor Tables

Table C1: Variance Inflation Factor - Job satisfaction DD

Variable	VIF	1/VIF
Period x Cohort	4.72	0.212
Period	3.36	0.298
Cohort	1.91	0.522
Self-rated health	1.05	0.954
Club member	1.04	0.957
Household income (log)	1.02	0.981
Mean VIF	2.18	

Table C2: Variance Inflation Factor - Life satisfaction DD

Variable	VIF	1/VIF
Period x Cohort	5.88	0.170
Cohort	4.00	0.250
Period	2.68	0.373
Job status	1.17	0.853
Household Income (log)	1.12	0.890
Married	1.09	0.917
Self-rated health	1.08	0.929
Club member	1.05	0.955
Mean VIF	2.26	

Table C3: Variance Inflation Factor - Depression level DD

Variable	VIF	1/VIF
Period x Cohort	5.36	0.187
Cohort	3.56	0.281
Period	2.66	0.376
Job status	1.17	0.857
Household Income (log)	1.11	0.902
Married	1.08	0.925
Self-rated health	1.08	0.929
Club member	1.05	0.956
Mean VIF	2.13	

Table C4: Variance Inflation Factor – Job satisfaction DDD

Variable	VIF	1/VIF
Period x Cohort	10.85	0.092
Period x Country	10.83	0.092
Period	9.82	0.102
Period x Country x Cohort	9.29	0.108
Country	2.07	0.483
Cohort	1.75	0.571
Household Income (log)	1.13	0.887
Self-rated Health	1.08	0.926
Club member	1.04	0.958
Mean VIF	5.32	

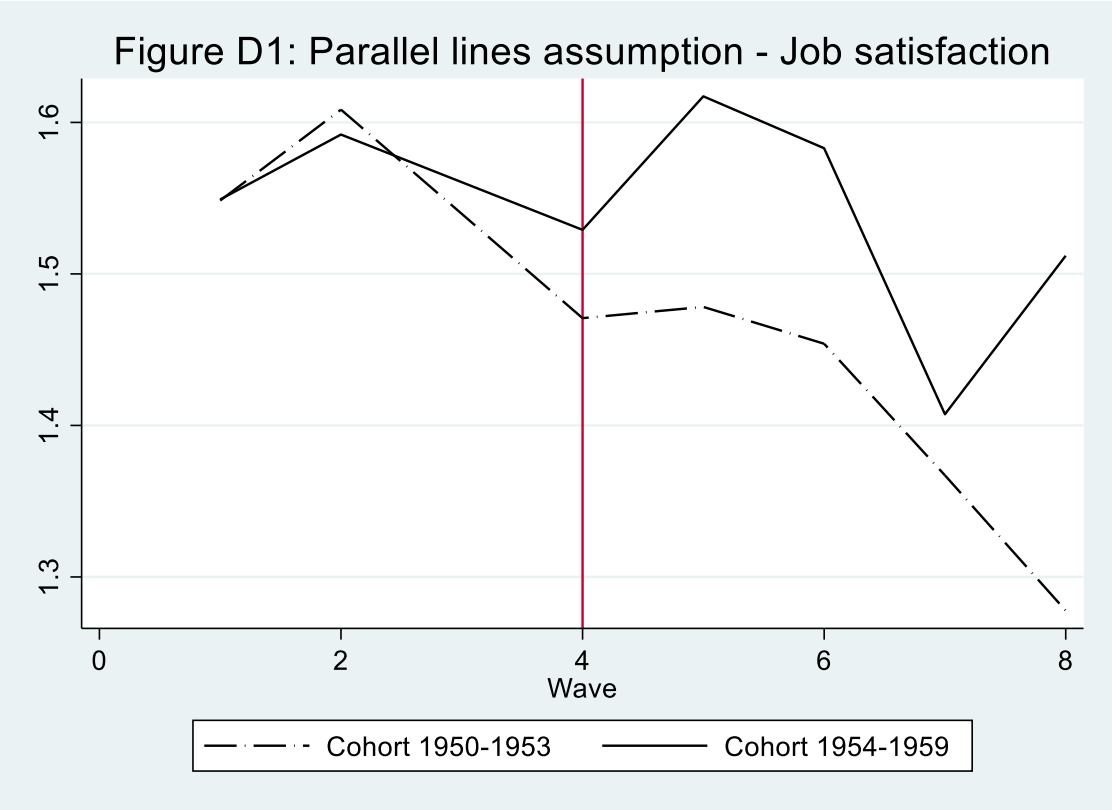
Table C5: Variance Inflation Factor – Life satisfaction DDD

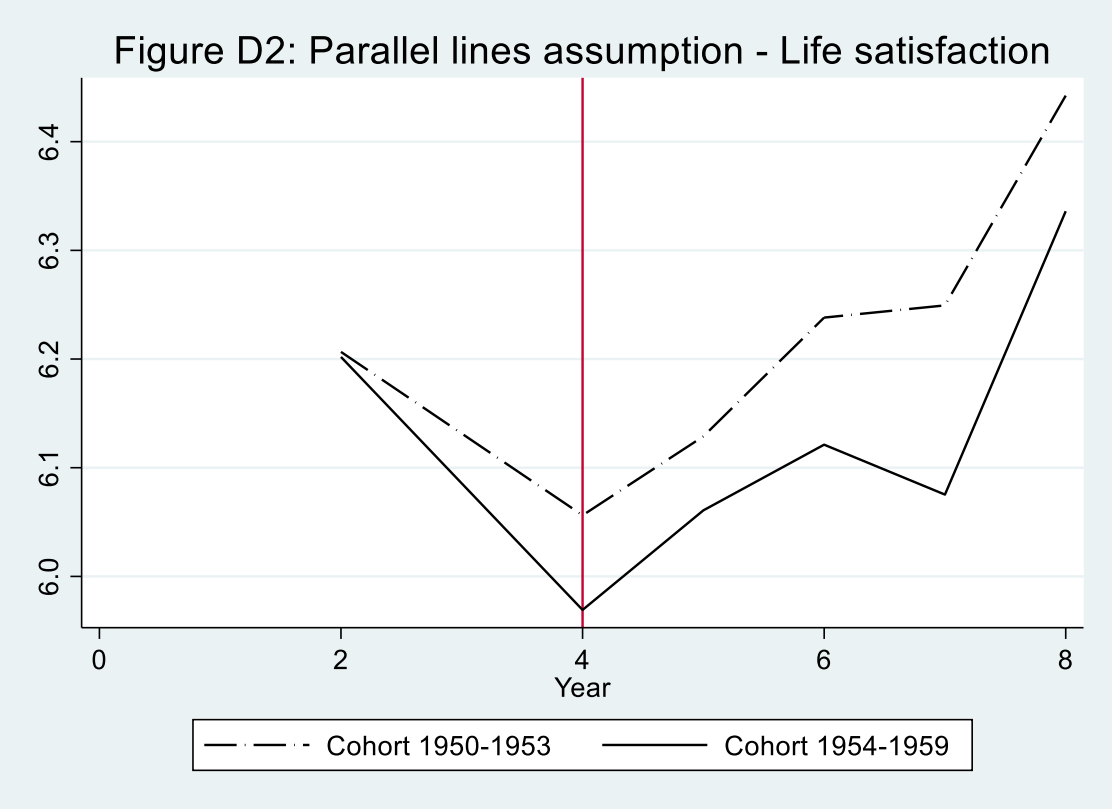
Variable	VIF	1/VIF
Period x Cohort	7.16	0.140
Period x Country	7.15	0.140
Period	4.41	0.227
Period x Country x Cohort	4.09	0.244
Country	4.09	0.245
Cohort	3.76	0.267
Household Income (log)	1.23	0.811
Self-rated health	1.15	0.867
Married	1.12	0.896
Club member	1.08	0.927
Job status	1.07	0.938
Mean VIF	3.30	

Table C6: Variance Inflation Factor – Depression level DDD

Variable	VIF	1/VIF
Period x Cohort	6.71	0.149
Period x Country	6.64	0.151
Country	4.49	0.223
Period	4.12	0.243
Cohort	3.53	0.283
Period x Country x Cohort	3.27	0.306
Household Income (log)	1.20	0.832
Job status	1.14	0.875
Self-rated health	1.11	0.898
Married	1.07	0.931
Club member	1.07	0.938
Mean VIF	3.12	

Appendix D: Parallel lines assumptions





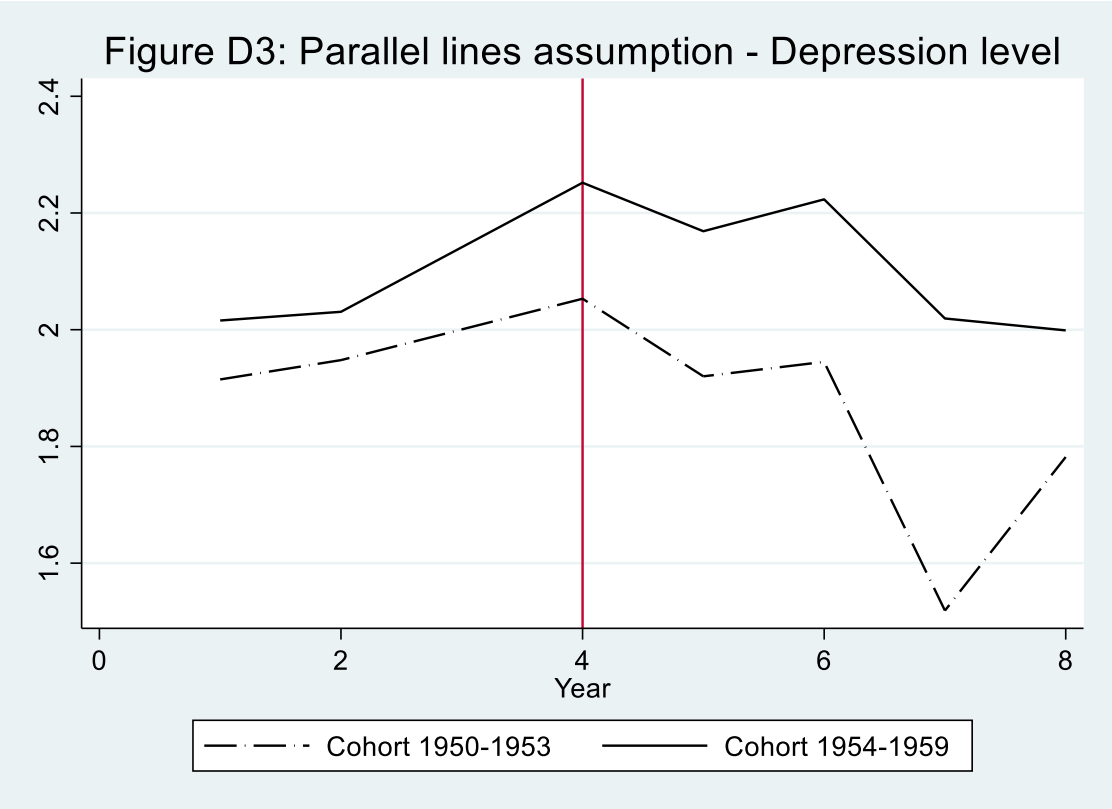


Figure D3: Parallel lines assumption - Depression level
2.4
2.2
2
1.8
1.6
0
2
4
6
8
Year
Cohort 1950-1953
Cohort 1954-1959

Appendix E: Ordered logit – Hypotheses 1a, 1b and 2

Table E1: Ordered logistic regression output – odds ratios

	Job Satisfaction		Life Satisfaction		Depression Level	
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Interaction Terms						
Period 2	0.713**	0.706**	1.288***	1.282**	0.859*	0.828*
	[0.094]	[0.100]	[0.111]	[0.125]	[0.068]	[0.082]
Young cohort	0.952	0.890	1.011	0.881	1.076	1.223**
	[0.093]	[0.100]	[0.097]	[0.087]	[0.091]	[0.120]
Period 2 x Young Cohort	1.421**	1.398**	0.761**	0.902	1.155	1.008
	[0.226]	[0.236]	[0.085]	[0.107]	[0.120]	[0.122]
Self-rated health (Excellent ref.)						
Very good		1.764***		0.523***		1.548***
		[0.171]		[0.033]		[0.102]
Good		2.642***		0.339***		2.467***
		[0.299]		[0.025]		[0.189]
Fair		2.694***		0.181***		6.745***
		[0.396]		[0.016]		[0.653]
Poor		3.166***		0.106***		11.822***
		[0.992]		[0.019]		[2.085]
Married						
Yes				1.655***		0.870**
				[0.101]		[0.055]
Club membership						
Yes		0.995		0.970		0.868***
		[0.077]		[0.049]		[0.046]
Household income						
Household income (log)		0.977		1.106**		0.869***
		[0.067]		[0.053]		[0.042]
Gender						
Male		0.779***		0.777***		0.604***
		[0.060]		[0.038]		[0.031]
Job Situation (Retired ref.)						
(Self-)Employed				0.905		1.044
				[0.071]		[0.088]
Unemployed				0.345***		2.209***
				[0.049]		[0.319]
Permanently sick				0.446***		1.703***
				[0.078]		[0.298]
Homemaker				0.884		0.826
				[0.347]		[0.370]
Observations	3,031	2,653	5,991	5,811	5,531	4,980
R-squared	0.017	0.042	0.083	0.230	0.016	0.176

Note. Ref. stands for reference category. The codes for significance are as follows: *** indicates $p < 0.01$, ** indicates $p < 0.05$, * indicates $p < 0.1$. Standard errors are reported beneath the regression estimates between brackets.

Appendix F: Probability margins – Hypotheses 1b and 2

Table F1: Probability Margins - Life Satisfaction

	A higher number indicates more life satisfaction							
Period x Cohort	1	2	3	4	5	6	7	8
First period x Old cohort	0.006***	0.006***	0.024***	0.019***	0.078***	0.308***	0.292***	0.267***
First period x Young cohort	0.007***	0.007***	0.027***	0.021***	0.085***	0.323***	0.286***	0.244***
Second period x Old cohort	0.005***	0.005***	0.019***	0.015***	0.065***	0.279***	0.299***	0.314***
Second period x Young cohort	0.006***	0.006***	0.024***	0.019***	0.077***	0.306***	0.293***	0.270***

Note. The codes for significance are as follows: *** indicates $p < 0.01$, ** indicates $p < 0.05$, * indicates $p < 0.1$.

Table F2: Probability Margins – Depression level

	A higher number indicates more depressed								
Period x Cohort	0	1	2	3	4	5	6	7	8
First period x Old cohort	0.313***	0.244***	0.180***	0.101***	0.072***	0.039***	0.027***	0.012***	0.012***
First period x Young cohort	0.274***	0.239***	0.189***	0.110***	0.082***	0.045***	0.032***	0.015***	0.015***
Second period x Old cohort	0.351***	0.247***	0.170***	0.091***	0.064***	0.034***	0.023***	0.010***	0.010***
Second period x Young cohort	0.309***	0.244***	0.181***	0.102***	0.073***	0.039***	0.027***	0.013***	0.012***

Note. The codes for significance are as follows: *** indicates $p < 0.01$, ** indicates $p < 0.05$, * indicates $p < 0.1$.

Appendix G: Robustness Checks

Table G1: Ordered logistic regression results - Triple interaction – Odds ratios

Variables	Job Satisfaction	Life Satisfaction	Depression Level
Triple difference			
Period 2	0.864	1.014	0.949
	[0.218]	[0.131]	[0.124]
Young cohort	1.122	0.760**	1.118
	[0.169]	[0.100]	[0.145]
Denmark	0.915	3.125***	0.927
	[0.160]	[0.476]	[0.138]
Period 2 x Young Cohort	1.321	1.374**	1.111
	[0.375]	[0.214]	[0.177]
Period 2 x Denmark	0.883	1.063	0.868
	[0.278]	[0.189]	[0.156]
Young cohort x Denmark	0.940	1.066	0.960
	[0.208]	[0.207]	[0.184]
Period 2 x Young Cohort x Denmark	0.923	0.781	0.892
	[0.340]	[0.177]	[0.205]
Self-rated health (Excellent Ref.)			
Very good	1.692***	0.577***	1.384***
	[0.202]	[0.043]	[0.108]
Good	2.402***	0.364***	2.157***
	[0.297]	[0.028]	[0.173]
Fair	3.072***	0.188***	5.675***
	[0.478]	[0.018]	[0.566]
Poor	3.035***	0.069***	18.177***
	[0.940]	[0.011]	[3.258]
Married			
Yes		2.043***	0.750***
		[0.124]	[0.048]
Club Member			
Yes	1.040	1.052	0.948
	[0.083]	[0.052]	[0.050]
Household Income			
Household Income (log)	0.880*	1.096**	0.921*
	[0.061]	[0.049]	[0.041]
Job Situation (Retired ref.)			
(Self-)Employed		0.823***	0.986
		[0.059]	[0.078]
Unemployed		0.428***	1.422***
		[0.053]	[0.185]
Permanently sick		0.431***	2.056***
		[0.072]	[0.369]
Homemaker		1.032	2.370
		[0.580]	[1.281]
Observations	2,653	5,811	4,980

Note. Ref. stands for reference category. The codes for significance are as follows: *** indicates $p < 0.01$, ** indicates $p < 0.05$, * indicates $p < 0.1$. Standard errors are reported beneath the regression estimates between brackets.