university of groningen

Universitat Pompeu Fabra Barcelona

The Great Recession in Spain and the short- and longer-term health effects of this crisis for people aged over 50.

Using SF-12 physical and mental health scores in SHARE

Master Thesis

Final version – 28 December 2021

Author:	**Thom Hidde Klein**
Program:	**DDM Social Demography**
Supervisors:	**Dr. Tobias Vogt (UG) & Dr. Joan Benach de Rovira (UPF)**
Word count:	**17,507 (excluding acknowledgments, references and appendices)**

Table of contents

List of tables and figures

Acknowledgements

This thesis marks the end of my time of being a student. Six years and a three months in Groningen, Berlin and Barcelona.

During the bachelor Human Geography & Urban and Regional Planning and the masters Sociology & Demography and Population Studies, I have learned a lot about a variety of subjects. I learned what I like to do and where I'm good at, as well as what I don't like. This thesis combines some of my favorite research subjects: health, economics and social inequalities.

I would like to thank my first supervisor Tobias Vogt of the University of Groningen for the good comments he gave me and the fruitful meetings we had throughout the entire process of writing this thesis. I would also like to thank my second supervisor Joan Benach of the Universitat Pompeu Fabra for the useful comments, especially on my proposal and during the first steps of writing the thesis. I also want to thank Lara Bister, who helped me create the SF-12 variables that I have used as main dependent variables in this research project.

Obviously, I want to thank friends and family that motived me to finish the thesis but who also distracted me to do enough fun things when I needed this.

I hope reading this thesis will be interesting for you!

Abstract

Prior to the Great Recession that started in 2008, Spain had a growing economy and a relatively low unemployment rate. Spain however has been hit very hard by the economic crisis and recovery took long. Spain has an ageing population and older people are often more vulnerable in times of crisis. It is unclear how the health of people aged above 50 has been affected by the economic crisis on the short- and longer-term. Linear regression analyses have been performed using data from different waves of the SHARE dataset. Health is measured by calculating SF-12 scores for all individuals. On the short-term, job loss and concerns about the economic situation of children are not related to worse health outcomes. Regarding income loss, it depends on the severity of the loss and on the income prior to the crisis. Individuals with a high income prior to the crisis don't have worse health outcomes on the short-term when they experience income loss, whereas this is the opposite for individuals with a low income before the crisis. Available data is insufficient for a solid longer-term analysis, therefore the longer-term effects of the economic crisis on the health of older Spanish individuals remain unclear.

Keywords: **Spain; Elderly; Health; Economic crisis; SHARE dataset; SF-12 health scores.**

1. Introduction

1.1. Economic crisis and health in Spain

Spain had a growing economy prior to the crisis, with a booming construction industry and housing prices rapidly increasing (Busch et al., 2013). The Spanish economy was heavily reliant on the construction industry and a real-estate bubble arose. The situation abruptly changed in 2008, when housing prices went down. In the following years, the government gross debt increased (Beker, 2014; Panico & Purificato, 2013; Busch et al., 2013; Roca et al., 2013).

Just as in other European countries, the economy of Spain fell into a large economic crisis in 2008. However, the crisis in Spain has been particularly severe, and the country was hit harder than other countries. Unemployment has for instance grown from about 8% prior to the crisis to 27.16% five years later, in 2013 (Bosch et al., 2014) and has remained relatively high even in 2017, with an unemployment rate of about 17% (INE, 2021a). Public debt has also grown substantially in Spain (Bosch et al., 2014). After the crisis, the economies of European countries recovered again, yet this happened in different speeds. Northern and Eastern European countries saw a much more rapid recovery than Southern European countries such as Spain, Italy, Greece and Portugal (Eurostat, 2019). The relatively slow recovery from the large economic crisis in Spain shows again that Spain was very hard hit by the crisis and that the crisis lasted especially long in the country (Carballo-Cruz, 2011). The recovery in Spain has been especially slow due to the high public debt, high unemployment rate and adverse effects of austerity policies (Carballo-Cruz, 2011).

Spain had, relatively to other European countries, a low government gross debt in 2007, before the start of the European Debt Crisis (Beker, 2014; Panico & Purificato, 2013). The Spanish public debt was about 36% of its GDP in 2007, whereas the debt in the rest of the Euro area was around 66% (Gordo et al., 2013). However, the large and severe economic crisis caused the Spanish public debt to increase. In 2012, in the middle of the crisis, the government deficit reached 10.6% of GDP in one year only. This was the worst figure in Europe (Bosch et al., 2014). The Spanish debt climbed to over 88% of GDP in the first quarter of 2013 (Gordo et al., 2013). This shows the severity of the government deficit in Spain, which also forms an explanation for the austerity measures that have been taken.

In periods of economic decline, such as during the European Debt Crisis, there are two broad options that can be implemented. Governments could choose to invest, boost the economy and trigger economic growth, or they could choose to reduce government spending and try to lower the debt. The latter are austerity measures (Stuckler et al., 2017). In Europe, many countries implemented austerity measures. Spain was also one of the countries that implemented austerity measures (Dellepiane & Hardiman, 2012;

Busch et al., 2013; Karanikolos et al., 2013; Bosch et al., 2014). The austerity measures that followed the economic recession have, mainly in Southern European countries, primarily affected wages, pensions and social and health services (Busch et al., 2013; Frangakis, 2015). Because of the negative impact on for instance social and health services, but also on the slow recovery speed of an economy after much austerity, many scholars are critical on the implementation of austerity measures (Ayuso-Mateos et al., 2013; Kinsella, 2012; Karanikolos et al., 2013, McKee et al., 2012). Not much is known about the impact of austerity on health and well-being, although it seems that austerity measures have had an impact on health and healthcare systems (Stuckler et al., 2017; McKee et al., 2012) According to Stuckler et al. (2017), mortality among elderly people in several European countries such as the UK and Italy increased. Much information and data on health of elderly people during economic crisis is still lacking, but the most deprived and vulnerable groups in society often suffer the most of economic recession and subsequent austerity measures (Stuckler et al., 2017). This makes it plausible that elderly people are one of the more vulnerable groups and their health could potentially be affected strongly in times of economic decline. Especially because elderly rely more on the health care system and receive pensions. The Spanish healthcare system performs well, though public spending is high (OECD, 2019). The Spanish healthcare system can be included into the Beveridge model healthcare system group. (Gaeta et al., 2017). This type of healthcare system is characterized by three main defining features: universal single-payer insurance; public healthcare provision; and free care (Bhattacharya et al., 2014; Lameire et al., 1999). During the economic crisis, healthcare reforms have been implemented in Spain, partly in the form of austerity. Large cuts in public healthcare spending have been made during the crisis: 13.7% in 2012 and 16.2% in 2013, with some regions going even further (Legido-Quigley et al., 2013). This has had negative effects for especially elderly that need more care and medication. Pensioners now have to pay a larger part of the costs of their medication (Legido-Quigley et al., 2013). The measures that have been taken during the economic crisis are seen as a threat to the universal provision of healthcare, moreover the UN discourages such measures.

It will be hard to test in this research how austerity and health budget cuts have affected health of individuals, but these cuts and other insecurities during crises can cause people to be stressed, which is not good for their health. Especially if these individuals are already in more vulnerable groups such as the elderly. (Whitehead & Bergeman, 2017; Karanikolos et al., 2016). In Spain for instance the number of mental health problems have increased during the economic crisis (Roca et al., 2013). Moreover, austerity has not been helpful in providing a quick recovery of the Spanish economy, but did the opposite and had as an effect that the Spanish economy has been in the crisis even longer (Carballo-Cruz, 2011), with all its consequences.

1.2. Relevance and motivation

The main motivation to start this research is that Spain has been hit exceptionally hard by the economic crisis and recovery took long. Moreover Spain has an ageing population and elderly are often seen as a vulnerable group. Research is lacking especially on the longer-term health effects of the crisis and this thesis would like to contribute to the knowledge about this.

In and after the economic crisis, the healthcare system in Spain has suffered (Bosch et al., 2014). Large cuts in funding have been implemented, not only on health but also on pensions and other forms of social security (Legido-Quigley et al., 2013; Dellepiane & Hardiman, 2012; Busch et al., 2013; Karanikolos et al., 2013; Bosch et al., 2014). Elderly (65+) are a vulnerable group (Stuckler et al., 2017; Levy & Sidel, 2009; Zavras et al., 2012), of which in Spain currently one in eight lives at-risk of poverty (Eurostat, 2021). Economic insecurity and financial insecurity can lead to stress and unhealthy behavior in some cases, especially for the more vulnerable groups (Whitehead & Bergeman, 2017; Karanikolos et al., 2016). Numerous studies with regard to e.g. individual employment status and health have been carried out. Studies of the relationship between more macro-level conditions, such as economic crises, and health are much more limited (Janlert, 2009), see also the literature review in the following chapter. Moreover, the economic crisis starting in 2008 is a relatively recent event, the health effects on the longer run are still unclear and understudied, whereas short-term health effects are already more established (Bosch et al., 2014). Now, a few years after the crisis with a new crisis (COVID-19 and the subsequent economic slow-down) currently happening in Spain, it makes it relevant to find out if elderly are suffering from longer-term health effects of the last economic crisis. The research objective is therefore to determine whether older people in Spain suffer from health effects that can be related to the economic crisis, on both the short-term as the longer-term. Is there a relationship between how individuals have been affected during the economic crisis and their current health outcomes? The research will look at different economic indicators that are related to economic crises and will seek for an innovative way to measure health.

It is academically relevant to research this because not much research has been done on the case of elderly specifically, especially not on the longer-term. This age group is often overlooked in research and the focus is more on people in working ages. The societal relevance is to identity when people, in this case elderly, are at risk of facing health consequences in the future. By defining these situations more clearly, programs or implementation strategies could be designed and policy advice can be made.

2. Theoretical framework

2.1. Literature review

There have been many studies that try to find a relationship between health and economy or economic crises. There have been studies in Spain, but also in other countries all over the world. This chapter differentiates between studies and findings on the macro-level (societies) and studies and findings on the micro-level (individuals or households). In this research, the measure for health is the SF-12 variable. This is a calculated health score for each individual in the dataset that consists of both a physical health score and a mental health score. Later, in the research design chapter, SF-12 will be explained in more depth. However it is important to know that this research thus looks at the broad concept of health, and includes both physical and mental health. Therefore studies on both types of health (mental and physical) and their relation to economic crises are discussed in this literature review.

2.1.1. Macro level studies

Economic crises have a large impact on societies and individuals. On a societal level, an economic crisis negatively impacts the development of wages, it increases unemployment rates and there is a decline of economic output (Carballo-Cruz, 2011; Frangakis, 2015; Arechavala et al., 2015). Moreover, an economic crisis can have a negative effect on public spending of the government on pensions and social and health services due to for instance decreasing tax revenues (Busch et al., 2013; Frangakis, 2015; Arechavala et al., 2015). These spending cuts can subsequently influence the health of a population (Stuckler et al., 2017; McKee et al., 2012) and can have a negative impact on the quality of life of individuals (Arechavala et al., 2015).

There have been multiple studies that try to find a relationship between health and economy. Mental health is one of the two main forms of health that are researched. A possible measure for mental health is suicide rates, although this does not tell the whole story. In the East Asian economic crisis for instance, unemployment rapidly grew, leading to an increase in suicide mortality on a societal level. Also, in the European Union and the United States unemployment has previously led to an increase of suicides (Uutela, 2010). Gudmundsdottir (2013) states that unemployed individuals are often unhappier than employed individuals, she is drawing this conclusion from a wide range of academic sources. However, she states that the relationship between happiness and economic crisis is unclear. Persons that are not directly affected by a crisis, and do not fall into unemployment, will therefore not be unhappier during times of economic crisis. Fountoulakis et al. (2012) goes even further and concludes that there is no causal link between economic crisis and suicide, even though they also acknowledge the negative effects

of poverty and economic uncertainty on mental health. Roca et al. (2013) confirm this and mention that poorer mental health does not have to lead to more suicides.

Another example of a study that does not find a negative relationship between health and economic crisis is the research by Granados & Rodriguez (2016). The authors found improved health in Finland, Iceland and Greece during the economic crisis, even though Greece was hit much harder by the crisis, and therefore no signs of something they call a 'health tragedy' were found. In his well-known and classic paper, Preston (1975) describes the relation between mortality and the level of economic development. He finds that at a certain level of economic development, displayed in income per capita, life expectancy of a country does not substantially rise anymore: there is no linear relationship, but a so-called Preston curve (Dalgaard & Strulik, 2012; Bloom & Canning, 2007). This could mean that in countries with a relatively high income per capita (e.g. also Spain) income loss would not make much of a difference for health, displayed in life expectancy. This thus suggest that for richer countries, there is not so much a relationship between health and economy (income levels) anymore on a societal level. The Preston curve can, when zoomed in on richer countries at the end of the curve, thus be seen as an example that economic downfalls do not necessarily have to lead to worsening health outcomes in a society.

This paragraph shows that scholars report conflicting on the societal (macro) level. Some researchers state that health on a societal level worsens during crises, while others suggest the opposite or do not see any effects. The interest of this thesis is however on the impact of a macro-level event (economic crisis) on the micro-level (health of individuals). Therefore not only macro-level studies are consulted, but also studies that focus more on individuals.

2.1.2. Micro level studies

On an individual or household level, individuals will make different expenditure choices when they are hit by an economic crisis. They do this because of e.g. loss of income or unemployment. Changing expenditure patterns can negatively impact education, health and nutrition. This can have longer-term negative effects, especially for vulnerable groups such as children (UN DESA, n.d.). Moreover, economic hardship can lead to stress (Whitehead & Bergeman, 2017) and unhealthy behavior among groups that are most vulnerable (Karanikolos et al., 2016). The elderly population is often also described as such a vulnerable group (Stuckler et al., 2017; Levy & Sidel, 2009; Zavras et al., 2012).

Whitehead & Bergeman (2017) show that stress due to problems with finances or employment can lead to poorer physical health. This shows that mental and physical health are related. The authors however mention, just as Fountoulakis et al. (2012), that perception matters: if one does not experience job loss

or income insecurity themselves, they are less likely to report poor physical health than individuals that did experience job or income loss. Although there are also scholars that mention that concerns about a worse economic outlook or concern about family members could still affect stress levels and therefore health (e.g. Maitoza, 2019). However, as has become clear, in times of economic crises employment can be less secure, which impacts many persons. Individuals that experience unstable employment or unemployment are more likely to report poorer self-rated health, and moreover will be more likely to be diagnosed with chronic diseases (Virtanen et al., 2002; Davalos & French, 2011). Examples of these chronic diseases are asthma, rheumatoid arthritis, diabetes and cardiovascular disease (Virtanen et al., 2002).

There are also various studies that mention opposite findings: studies that find a positive effect of economic crises on health. Or at least do not find negative effects. An example is Fountoulakis et al., (2012) that found no signs of worse health. In the case of Iceland for instance, the economic crisis has led to healthier behavior among its citizens (Ásgeirsdóttir et al., 2014).

Even though scholars do not agree on all details, broadly speaking, it is clear that economic crises can have a large impact on the lives of individuals. People's health can be impacted negatively by economic crises, both mental and physical health. Negative effects are clearer on the individual level than on a societal level as described in the former paragraph. Since vulnerable groups are often harder hit, and elderly are seen as a vulnerable group, it is interesting to research how this group's health is affected in times of economic crises, and how their health is affected in the longer run. In this study, the case of Spain will be researched. Spain was hit very hard by the economic crisis that started in 2008 and has a large elderly population.

2.2. Theories

2.2.1. Coleman's bathtub

In social research, Coleman's (1990) bathtub model has become an important diagram to represent macro-micro-micro-macro links (Raub & Voss, 2017). The diagram shows the relationship of macro-conditions, to micro-conditions, then of these micro-conditions to micro-outcomes, and subsequently the relationship of the micro-outcomes to macro-outcomes. In figure 1, Coleman's bathtub is pictured. In the case of this research, the focus will be on process 1 and 2 in the figure. The effects of a macro-condition (economic crisis) on micro-outcomes (health outcomes of individuals).

This model by Coleman will form the basis for the theoretical framework that will be used in this thesis. The major processes, possible causes and outcomes can be displayed in the model. The overarching

macro condition that is the basis of this research is the Great Recession in Spain. This is point A in the model. Micro conditions (point B in the model) are the direct consequences of this economic crisis. These can be things such as the previously mentioned income loss, financial uncertainty and increased unemployment rates. These are all clear effects of economic crises that are well-established in existing literature. Point C is the part that is of main interest in this thesis: the outcome on the individual level (micro). This is the health of an older individual. That is where the use of Coleman's bathtub model stops in this thesis. The line between point C and point D will not be included in this research, because the interest is on the effects of the macro condition on the individual level, and not on a societal level.

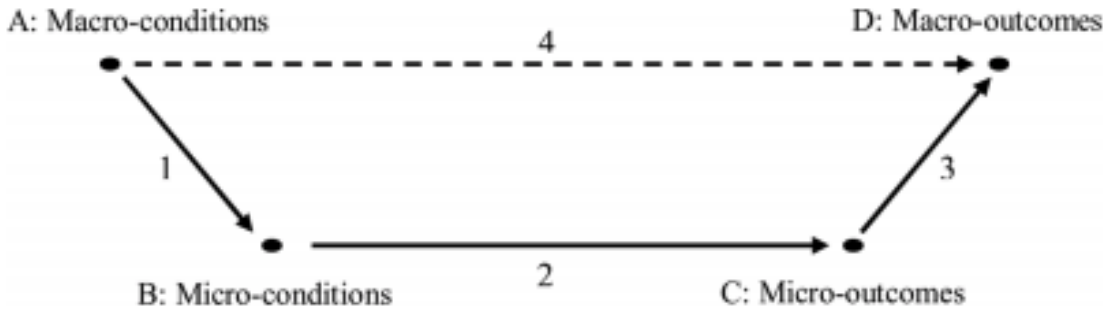


Figure 1 Coleman's bathtub model (Coleman, 1990; Raub et al., 2011).

The model is relevant to include in this research because it clearly represents that main links of interest (1 and 2), existing theory (1) and the research gap (2) can easily be included into the model. Moreover, the model is simple and easy to understand. Therefore, this model serves as the basis for the conceptual model that will be introduced later in this chapter.

2.2.2. Life course health development framework

Following the life course health development framework, health is a consequence of multiple determinants and contexts that change and impact health during a lifetime as someone develops. These determinants can be biological, genetic, social, behavioral and moreover also economic (Halfon & Hochstein, 2003). Especially in utero conditions are important determinants for later life health outcomes (Arcaya et al., 2015), also for the elderly this is the case (Gluckman et al., 2005). Good economic circumstances during the year of birth have a positive effect on later-life health outcomes, such as cognitive functioning, and vice versa (Doblhammer et al., 2013). It is however also studied that not only the economic circumstances at birth are important for health at later life. Also, economic circumstances at other points in life can influence health outcomes. For instance, a bad income during adulthood or a low retirement income for elderly (Arcaya et al., 2015; Adler et al., 2007). This follows the life course health development framework, which assumes that the cumulative impact of conditions

throughout the lifetime determine someone's health (Hazen & Anthamatten, 2020; Halfon & Hochstein, 2003). The life course health development framework offers a conceptual model to understand development of health outcomes. It is an interdisciplinary framework that can guide research on health, ageing and development and the framework has gained popularity in research (Kuh et al., 2003). The model is widely applicable, both for chronic disease epidemiology as for research on infectious diseases, moreover the framework can be used in research about broader notions of health and wellbeing (Ben-Schlomo & Kuh, 2002), such as in this thesis. The concept of time is very important in the life course health development framework (Kuh et al., 2003), as health develops continuously over time in life (Geller et al., 2018). The component time of the life course health development framework should therefore be included in the conceptual model of this thesis. Different versions of the framework by Kuh et al. (2003) can be found in appendix 1.

The life course health development framework could be seen as unimportant for this research, because not the cumulative effect of different circumstances during the lifetime is researched, but merely the effect of just one event (economic crisis). However, I would argue that the life course health development framework is important. In the first place to make clear that, besides economic crises, there are many other important factors that could influence health, that are not included in this research. It is important to acknowledge this and elucidate that not only the economic crisis determines health outcomes. Moreover, the life course health development framework includes the important concept of time, and thus provides a way to introduce the concept of time and its importance into the conceptual model. This is especially important because events (such as the economic crisis) can have different effects over time, on the short- and longer-term.

2.2.3. Family stress model of economic hardship

This thesis aims at finding out whether there is a relationship between the Great Recession in Spain and health of people aged over 50. In order to measure this, different concepts need to be determined that can be used as a measure for economic crises. From existing literature and previous studies on the subject it has become clear that income loss and unemployment are omnipresent during crises (Frangakis, 2015; Arechavala et al., 2015). These should therefore be included into a theoretical framework for this study. However, it are not only these quite straightforward and clear variables that could influence health. Moreover, concerns on the economic status of loved ones has been mentioned in existing literature as possible factor that contributes to health outcomes (e.g. Mucci et al., 2016; Chalari, 2014).

This can be related to the family stress model. Originally, the family stress model of economic hardship mainly focusses on adolescents that grow up in families that are going through economically stressful periods (Conger et al., 2000). The stress that comes with economic hardship can have a negative impact

on relationships within a family and increase family members' vulnerability. This can cause depressions, anxiety and adversely influence psychological wellbeing (Conger et al., 2000; Layte & McCrory, 2018, Whitehead & Bergeman, 2017). The family stress model however assumes that stressful situations occur in families due to income loss, unemployment or other forms of economic insecurity of parents, which can subsequently influence children's wellbeing and health. However, this research seeks for the opposite: older individuals that are concerned about the economic situation of their children.

Therefore, a "reverse" family stress model of economic hardship is proposed as a theory that applies to this research. Meaning that stress is not only induced on children by economic hardship of parents, but that this could also occur vice-versa: economic hardship of children could cause stress for parents. Maitoza (2019) for instance states that unemployment not only has an effect on the unemployed themselves, but moreover an effect could be felt by people close to them as well. This could thus also be parents and therefore creating a "reverse" family stress model. Chalari (2014) shows that in Greece many older parents are worried and concerned about their children's economic situation and wellbeing. Consequences of having unemployed children could be for instance an increased dependence on parents, making both the parents and the child more financially vulnerable (Brydsten et al., 2016). Another consequence could be increased alienation and anti-social behavior (Hammarström, 1994). This, and economic uncertainty for the future (of their children) in general, can thus cause parents to be concerned and worried (Maitoza, 2019; Chalari, 2014). The stress that older individuals can experience in situations such as the aforementioned could be bad not only for mental but also for physical health, as Evers et al. (2010) describe that stress 'can get under the skin'. With this phrase, these researches mean that increased levels of stress may on the long term contribute to worse health outcomes such as an increased severity of chronic diseases (Evers et al., 2010; Mucci et al., 2016).

2.3. Conceptual model

Based on all the concepts and theories that have been discussed in this chapter, a conceptual model is created. The main concepts that are of importance for this research and that therefore are included into this model are Coleman's bathtub model, the life course health development framework, the (reversed) family stress model, unemployment and income change. In figure 2, the conceptual model that is discussed in this paragraph can be found.

As is visible in the figure, Coleman's bathtub model forms the basis of the conceptual model that is created for this research. Only link 1 and 2 and points A, B and C are included. Meaning that link 3 (and 4) and point D from the original model by Coleman are omitted. The macro-condition that is at the start of the model is the economic crisis. This macro-condition leads to three main micro-condition that are experienced on the individual level and have been derived from theory. These include income change,

employment status change and family stress/concerns. This link (link 1) will not be researched in this thesis and is well-established in existing research. These three micro-conditions subsequently can lead to micro-outcomes on the short- and longer-term regarding the health of the individual. This link (link 2) is the main research focus of this thesis.

On top of the conceptual model, a large arrow that represents life course health development is depicted to include the concept of time into the model. Time is important in the model, among others because of the division between short- and longer-term micro-outcomes, but moreover because of the assumption that many conditions can have an effect on health over an individual's lifetime. Economic crisis is only one of these conditions within this lifetime. Therefore the arrow of time does not begin and stop in the model, but starts before and continues after the conceptual model.

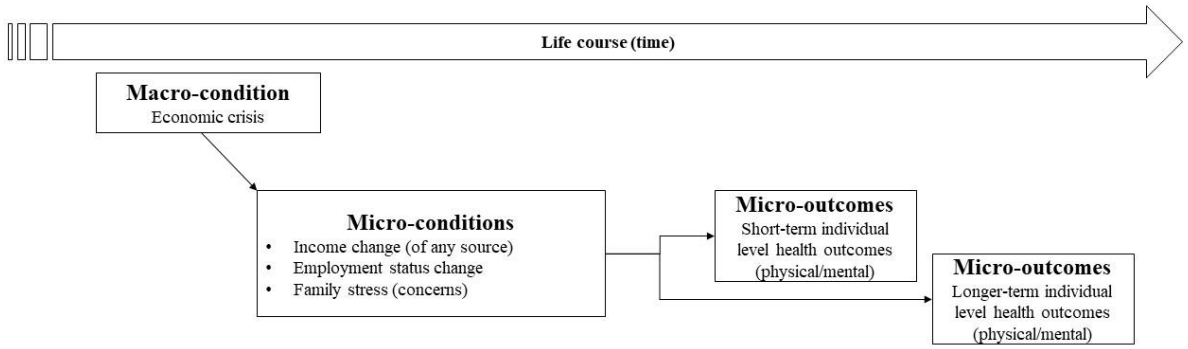


Figure 2 Conceptual model.

2.4. Research objective and research questions

The objective of this thesis is to find out whether the health of Spanish people that were 50 years or older at the start of the Great Recession has been influenced by this crisis, on both the short- and longer-term, where existing research is still lacking.

The main research question of this thesis, related to the research objective that has been defined, is:

What are the short- and longer-term effects of the economic crisis in Spain on the health of people aged above 50?

To answer this relatively broad question, several sub-questions have been formulated. These sub question help guide the research and are moreover based on the main concepts and focus points that are included in the conceptual model as introduced in the previous paragraph. The sub-questions that have been defined are the following:

1. Is there a relationship between income change and health outcomes of Spanish individuals above 50 on the short-term?
2. Is there a relationship between income change and health outcomes of Spanish individuals above 50 on the longer-term?
3. Is there a relationship between employment status change and health outcomes of Spanish individuals on the short-term?
4. Is there a relationship between employment status change and health outcomes of Spanish individuals on the longer-term?
5. Is there a relationship between having concerns about the economic situation of their children and health outcomes of Spanish individuals above 50 on the short-term?
6. Is there a relationship between having concerns about the economic situation of their children and health outcomes of Spanish individuals above 50 on the longer-term?

3. Research design

This chapter is divided into two main paragraphs. The first gives a detailed description of the data, the data collection and the data quality. The second paragraph is on the methodology of the research and how the research questions are operationalized using the data.

3.1. Data

3.1.1. Data collection

For this thesis, the SHARE dataset will be used. SHARE is the abbreviation of the Survey of Health, Ageing and Retirement in Europe. The survey is designed to gather a wide range of information to study the effects of health, social, economic and environmental policies and changes over the life-course of European citizens, more specifically older citizens because this group will grow rapidly in size in the coming decades (Börsch-Supan et al., 2013). The target population of the survey are therefore individuals aged above 50 from 28 European countries and Israel. Because the survey consists of multiple waves, starting in 2004, the survey provides researchers and policymakers longitudinal data. SHARE is the largest pan-European social science study that provides internationally comparable interdisciplinary longitudinal micro data (SHARE, 2021a).

The main funding for the SHARE comes from the European Commission. Data collection for the SHARE is coordinated by the Max Planck Institute for Social Law and Social Policy and is done through interviews (SHARE, 2021a). In the different countries that are included in the survey, affiliated academic institutions are responsible for the data collection and management. The Spanish partner organization that is responsible for collecting and managing the Spanish data is the CEMFI (Center for Monetary and Financial Studies) (SHARE-CEMFI, n.d.).

3.1.2. Data quality, representativeness and concerns

As already mentioned, the SHARE dataset consists of multiple waves, starting in 2004 (Wave 1). In total, seven waves have been completed and made available to download. The data collection for the most recent available wave (Wave 7) started in 2017 and was finished in 2018 (SHARE, 2021b). Spain was among the first countries to participate in the SHARE survey, already in the first wave, and has participated in all waves that have been held consequently. This means that there are waves available with data from before, during and after the large economic crisis in Spain. The most recent wave (Wave 8) with data from 2020 is not yet available (SHARE, 2021b).

For the first wave, 2324 interviews have been completed in Spain with eligible individuals and, if applicable, their partners. The individuals that took part in these interviews have been sampled through a two-stage random stratified selection from the census in all Spanish municipalities (SHARE-CEMFI, n.d.).

For the second wave, not all respondents returned in the sample. There are various reasons for this, some died, some refused to participate. Therefore, a refreshment sample was included, with new individuals to compensate for the loss of previous participants. In this refreshment sample the younger ages were oversampled to make sure that the representativeness of people aged above 50 would be maintained. In the third wave no new refreshment sample was added, hence a smaller samples size in this wave due to loss of previous participants. In the fourth wave a refreshment sample was included again, increasing the number of interviews to 3872 and maintaining the surveys representativeness for people aged over 50 (SHARE-CEMFI, n.d.).

In the fifth, sixth and seventh wave of the SHARE survey, a refreshment sample has been added once again. However, these samples are not representative for Spain as a whole. The interviews from these refreshment samples have only been held with individuals from the northeastern province Girona. The cases from Girona have been given a unique ID in the dataset to be able to identify and possibly separate these (SHARE-CEMFI, n.d.). The data from Girona has been collected by a separate team of researchers from Catalonia, instead of by the larger Spanish team (SHARE, 2021c). Because there are no new representative refreshments for the entire Spanish population in the more recent waves, the dataset is not representative anymore for individuals above 50 in wave 5, but for people above 52. In wave 6 and 7 the dataset is only representative for individuals aged over 54 and 56 representatively (SHARE-CEMFI, n.d.).

For this research, the aforementioned issue with representativeness is not a concern. Only people above 50 at the beginning of the crisis need to be included, because the target population of this research is Spanish individuals that were already 50 years or older during the crisis, and therefore were 50 years or older in Wave 2, for which the data was collected mostly in 2007 (SHARE, 2020), just before the start of the economic crisis in 2008 (Regidor et al., 2014; Éltető, 2011). This means that this research will use the sample of Wave 2 as baseline sample and only works with individuals that participated in the research the entire time between Wave 2 and the other wave of interest, in this research either Wave 4 or Wave 7.

In figure 3, the sample size of all waves (*Ola 1-7*) is visualized. Also, the number of individuals that left the study (in dark orange, *Salida*) after each wave is shown. In green, the size of the refreshment samples

is shown (*Refresco*). From Wave 5 onwards only refreshment samples from Girona are available (in blue tones).

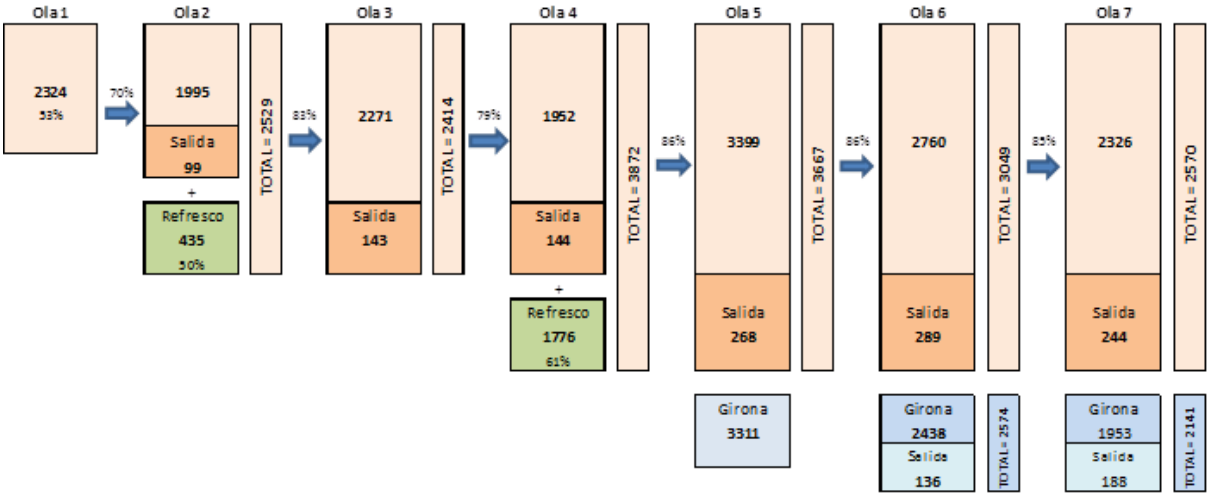


Figure 3 Number of interviews in each wave of the SHARE survey in Spain, including refreshment samples and the number individuals that left the study (SHARE-CEMFI, n.d.).

A concern with the data that does remain is the sample size. The sample is representative for the population and in each wave the sample size is quite large. The issue here is that people leave the survey sometimes, for various reasons. This has as a consequence that the sample size of the longitudinal dataset decreases, even if the size of the cross-sectional data increases. There has been quite some loss of respondents over time, which can be a problem for the sample size of this research, especially because this research is not only looking for short-term but also for longer-term effects of the economic crisis on health of Spanish individuals now, using economic indicators from early waves and looking at health indicators in the most recent wave (Wave 7). There are thus many years in-between the first and last interview and therefore also many years that a respondent could have left the interview, making a longitudinal analysis harder.

3.1.3. Description of the dataset

The dataset is not only divided into separate waves, but also in many different modules. This makes the data files smaller, because unnecessary variables can be excluded easily from the analysis and therefore it makes working with the dataset easier. Depending on the wave, there is a different number of modules available. In the later waves, more modules have been included, because also more different variables are included here. The most important modules for this research are *Mental Health*, *Physical Health*, *Activities*, *Social Network*, *Employment and Pensions* and *Demographics and Networks*, but also the

more general *GV_imputations* module (Generated Variables). Online, SHARE provides an easy tool to search through all waves and all different modules: the Data and Documentation Tool.

3.2. Operationalization

The objective of this research is to investigate whether there is a relationship between the large economic crisis in Spain that started in 2008, and health of people aged above 50. This research differentiates between health on the short-term and health on the longer-term.

In order to achieve the goal of this research and perform a statistical analysis to investigate the relationship between economic crises and health, different variables are needed. In the first place, a dependent variable that can measure health is essential. In the following paragraph (3.2.1.) this variable will be introduced and explained. Other important variables in order to perform the analyses are economic variables that capture the economic crisis. One could think of income loss and unemployment. These variables are the explanatory variables in the analysis. The explanatory variables that are used in this research are introduced and explained in paragraph 3.2.3.

3.2.1. Measuring health using SF-12

For this research, health is a key concept. Health is the concept that needs to be explained, therefore health is the dependent variable in the regression analyses that will be performed. According to the definition of the WHO, health is *'a state of complete physical, mental and social wellbeing and not merely the absence of disease or infirmity'* (WHO, 2021). This is a broad definition and shows that there are multiple aspects that are important for an individual's health. Because of this, I would argue that a broad measure of health is also important in this research. The SHARE dataset contains many health variables. One of these is the variable *Self-perceived health (sphus)*. Respondents of the SHARE survey have been asked to rate their health on a scale from one to five, with one being *Excellent* and five being *Poor*. One could argue that this is a good measure for health to use in the analysis of this research. Self-perceived or self-rated health is a widely used, simple, clear and easy to understand measure (Bombak, 2013; Berger et al., 2015; Lundberg & Manderbacka, 1996; Idler & Benyamini, 1997). However, the variable is rather limited, meaning that it is not a broad measure of health. Moreover, the variable is somewhat abstract. It can be unclear what exactly is measured with this variable (Jylhä, 2009). The measure is subjective and a respondent themselves can make the decision on what their answers is based. One respondent may for instance only base their answer on their mental health status, whereas another only does this on the absence of disease. This means that it thus can be very different than the more objective health-rating from a doctor or physician (Jylhä, 2009). Using a measure such as self-rated or self-perceived health therefore requires substantial interpretative work from respondents (Knäuper &

Turner, 2003). Another problem can be that the scale from excellent to poor is not the same in every respondent's mind (Jylhä, 2009). Considering the aforementioned issues with the limitations of the variable *Self-perceived health*, another measure for health is used in this research: the 12-item short form health survey score (SF-12 score).

Originally, the SF-12 score was developed as the 36-item short form (SF-36) health survey. The SF-36 is a short health survey with only 36 questions, first made available in a standard form in 1990 (Ware, 2000). These questions together result in eight different item scores (Physical Functioning; General Health; Bodily Pain; Role Physical; Mental Health; Role Emotional; Social Functioning; Vitality) and two summary scores (Physical Health and Mental Health; abbreviated to PCS and MCS) (Krahn et al., 2009; Ware, 2000). This original SF-36 score has led to the development of an even shorter version: the SF-12 score, which is used in this research. The SF-12 provides the same eight item scores and the same two summary scores, but used only 12 questions instead of 36 to calculate these scores. (Ware et al., 1995). In figure 4, the original 36 questions, eight items and two summary measures are shown. In the rectangles in the left column, the 12 questions from the SF-36 that are used to create the SF-12 are highlighted.

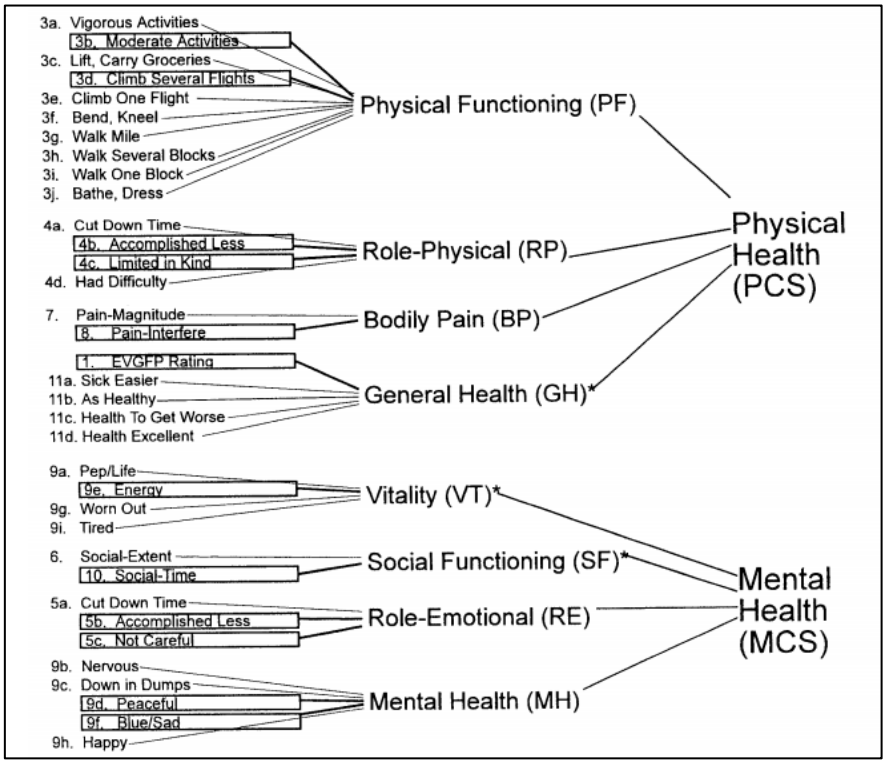


Figure 4 Original SF-12 Measurement Model: construction of the shorter SF-12 using SF-36 questions (Ware et al., 1995).

To calculate the SF-36 or SF-12 scores, one would normally let respondents fill out the short-form health survey. However, in some cases, when working with an existing secondary dataset, this is not possible.

In these situations, variables have to be found or created that are the same, similar or comparable to the original 36 or 12 questions of the short-form health survey. In this thesis, the already existing SHARE dataset was used and therefore existing variables had to be used to calculate the SF-12 scores.

The algorithm to calculate SF-12 scores in the German SOEP dataset by Nübling et al. (2006) and Andersen et al. (2007) has been used as an example for the calculation of the SF-12 scores with the SHARE dataset in this thesis. The variables from the SOEP dataset that are used to create the SF-12 scores can be found in appendix 2. These variables are in large part the same as the necessary variables to create the SF-12 scores as defined by Ware et al. (1995). There are only some minor differences (Nübling et al., 2006). In table 1, the final variables from the SHARE dataset (Wave 4) that were chosen to calculate the SF-12 scores can be seen. If these SHARE variables are compared to the variables that are used in the example from SOEP (appendix 2), one can see that there are many differences. Variable 1b, whether someone's health has an effect on their ability to perform tiring task does not exist in the SHARE data. To proxy this variable, the variable *ph049d12* was chosen. The variable is not the same, but measures something relatively similar. The same situation occurs for variables 4a, 5b, 6a, 6b, and 7. All these variables are different, but do measure something similar. The only variable that did not exist in the SHARE dataset and could also not be replaced by another existing variable from the dataset is variable 3: whether the respondent has felt physical pain in the last 4 weeks. In order to proxy this measure, a new variable had to be created. Variables ph011d7 and ph011d8, on whether the respondent used drugs against pain in the body or in muscles/joints, were aggregated and a new variable was created on whether the respondent uses drugs against physical pain. This obviously leaves out people that experience physical pain, but do not (yet) use drugs, and therefore may not catch all respondents with bodily pain. However, since there is no other measure available in the SHARE dataset, and this variable only accounts for 1/8 of the SF-12 calculation, it was still included.

Table 1 Variables from the SHARE dataset (Wave 4) used to create SF-12 scores.

SF-12 item and variable number	*Name in SHARE*	*Description in SHARE*
Item 1: PF (Physical Functioning)		
Variable 1a	ph048d4	Difficulties: climbing several flights of stairs
Variable 1b	ph049d12	Difficulties: doing work around the house or garden
Item 2: GH (General Health)		
Variable 2	ph003_	Health in general question 2
Item 3: BP (Bodily Pain)		
Variable 3	ph011d7 + ph011d8	Drugs for: joint pain + Drugs for: other pain
Item 4: RP (Role Physical)		
Variable 4a	ph061_	Health problem that limits paid work
Variable 4b	ph005_	Limited in activities because of health
Item 5: MH (Mental Health)		
Variable 5a	mh002_	Sad or depressed last month
Variable 5b	mh014_	Concentration on entertainment
Item 6: RE (Role Emotional)		
Variable 6a	mh022_	Ever told affective or emotional disorders
Variable 6b	ac017_	Do the things you want to
Item 7: SF (Social Functioning)		
Variable 7	sn012_	Network satisfaction

Item 8: VT (Vitality)		
Variable 8	ac023_	Feel full of energy

Over time, the questions that are included in the SHARE questionnaire change, and therefore not all variables are available in all waves. For this reason, some different variables had to be used to create SF-12 scores in Wave 7 as compared to Wave 4. In table 2, the necessary variables from the SHARE dataset in Wave 7 are shown. It can be seen that variables 4a, 5a, 5b, 6a and 7 are different from the variables in Wave 4. For variable 5a, this is due to a very limited number of respondents that had observations for the original variable. Since only respondents that have valid answers to all 12 necessary questions will be included, this would lead to a loss of too many observations. For the other variables, new variables had to be sought because of the complete absence of the original variables in this wave. Variable 7 was newly generated using two existing variables, just as has happened for variable 2.

Table 2 Variables from the SHARE dataset (WAVE 7) used to create SF-12 scores.

SF-12 item and variable number	***Name in SHARE***	***Description in SHARE***
Item 1: PF (Physical Functioning)		
Variable 1a	ph048d4	Difficulties: climbing several flights of stairs
Variable 1b	ph049d12	Difficulties: doing work around the house or garden
Item 2: GH (General Health)		
Variable 2	ph003_	Health in general question 2
Item 3: BP (Bodily Pain)		
Variable 3	ph011d7 + ph011d8	Drugs for: joint pain + Drugs for: other pain
Item 4: RP (Role Physical)		
Variable 4a	hs052_	Ever had physical injury to disability
Variable 4b	ph005_	Limited in activities because of health
Item 5: MH (Mental Health)		
Variable 5a	ph011d10	Drugs for: anxiety or depression
Variable 5b	ac704_	I see myself as someone who is relaxed, handles stress well
Item 6: RE (Role Emotional)		
Variable 6a	ph006d18	Other affective/emotional disorders: ever diagnosed
Variable 6b	ac017_	Do the things you want to
Item 7: SF (Social Functioning)		
Variable 7	hs054_ + hs063d4	Number of periods of ill health + Consequences of illness period: made social life more difficult
Item 8: VT (Vitality)		
Variable 8	ac023_	Feel full of energy

After the 12 necessary variables have been selected from the dataset, multiple steps need to be taken in order to create the two summary measures for mental and physical health. First, for some variables the scale needs to be inverted. This has to be done in order to have variables where a higher score always means a better health, and vice versa. The second step is to create the items. For items 2, 3, 7 and 8, this is straightforward: these items only consist of one variable. If an item consists of two variables, the mean of both variables will be taken to build the item. In some cases, the scales of two variables within one item are different and therefore have to be aligned before the mean can be taken. After the aforementioned steps have been finished, eight new variables are created.

After this, four further steps need to be taken. All the Stata codes for these steps can be found in appendix 3. This syntax is based on the syntax by Nübling et al. (2006). Step one is to make sure that all eight variables are in the same scale, with a minimum of 0 and a maximum of 100. Step two is to standardize the scores using z-score standardization: this step will create variables with a mean of 0 and a standard deviation of 1. Step three is to transform the scales using norm-based scoring. This will result in variables that all have a mean of 50 and a standard deviation of 10. Interpreting the results (the health scores) is much easier when mean scores and standard deviations are standardized. Norm-based scoring has been proven to be the useful to help interpreting the results (Ware, 2000). With this transformation, it is clear that when a scale score is below 50, health status is below average, and each point less or more is 1/10 of the standard deviation. Step four is to split the eight variables into two groups of four variables that will be used to generate the separate physical and mental component scores by performing a factor analysis. The factors that roll out of this analysis will eventually be used to calculate the PCS and MCS health scores. The obtained factor coefficients that are used for the calculation of PCS and MCS in Wave 4 and Wave 7 can be found in the tables of appendix 7.

It is important to note that the technique is new and used for the first time, to my knowledge, in the SHARE dataset. This makes the technique to calculate these health scores new and unverified. One could argue that this is a weakness of this research, however, one could also argue that this is a strength, because it is innovative. In my opinion, it makes the research unique and interesting, and creates a broad multidimensional health score that has never been created before with SHARE data. The scores are reliable and trustworthy because the technique has been inspired by peer-reviewed documents by Nübling et al. (2006) and Andersen et al., (2007), and is thus not entirely made-up and new. This thesis could therefore possibly even be useful and serve as an inspiration to other researchers that would like to work with SF-12 in SHARE. This will be further discussed in the discussion chapter.

3.2.2. Main explanatory variables

From theory, three main economic mechanisms were defined: unemployment, income loss and concerns about children (related to the family stress model). These can also be found in the conceptual model (see figure 2). Therefore, three explanatory variables, measuring these main mechanisms, should be found in the SHARE dataset, or should be created using the data from the dataset. The effect of these three mechanisms that will be researched in this thesis is a lagged effect, using a dependent variable from a later moment than these three explanatory variables. This means that the predictors (explanatory or independent variables) have not been measured at the same time as the dependent variable. In the case of this research this means that the economic variables, such as information on income (loss) or employment have been measured during the crisis, whereas health variables after the crisis. Because of this, variables from different waves are needed. This will become clear in this paragraph.

Income loss: The first mechanism that could influence an individual's health is income change (e.g. UN DESA, n.d.; Stuckler et al., 2017; McKee et al., 2012; Carballo-Cruz, 2011). In an economic crisis, loss of income can occur. In the SHARE dataset, many different separate variables are available to measure income. There are variables that measure income from employment, income from pensions, income from rent and more possible sources of income. In the generated variables dataset, the an income variable is included. This variable measures the total household income, from any source. Using this variable from both Wave 2 as from Wave 4, makes it possible to calculate the change of income during the economic crisis, between 2007 and 2011. Multiple variables measuring income change are generated using this information: 1) Absolute income change between 2007 and 2011, a ratio variable; 2) Income loss between 2007 and 2011: yes/no, a binary variable; 3) Income change categories, an ordinal variable.

Employment change: The second mechanism that could influence an individual's health is change in employment (e.g. Uutela, 2010; Whitehead & Bergeman, 2017). During economic crises, people can lose their job. In Spain, this was a particularly large problem (INE, 2021a; Bosch et al., 2014), during and after the crisis, the country had one of the highest unemployment rates of the European Union. Employment status can affect income, but moreover can cause stress and uncertainty, which could possibly lead to poorer health (Virtanen et al., 2002; Davalos & French, 2011). To test whether changes in the employment status of older individuals in Spain has had an effect on their short-term and longer-term health outcomes, a variable on change in employment status should be included into the statistical model. The necessary variable is made using the variable on the respondent their current job situation. This variable is taken for both Wave 2 as for Wave 4 and several dummy variables are generated.

Concerns about children: A third mechanism that could influence older individual's health is concerns about loved ones. According to the family stress model of economic hardship (Conger et al., 2000), this mainly goes in the direction from parents to children, but in this research I would like to test the opposite direction: concerns of parents about the economic situation of their children. This measure is especially relevant for the situation in Spain, where youth unemployment rates have reached exceptionally high levels (OECD, 2021), possibly causing economic hardship, stress and uncertainty for both children as their parents (Chalari, 2014; Mucci et al., 2016). The SHARE dataset does not contain a variable that measures concerns for the economic situation of children, therefore a proxy measure is needed. The occupational status of children is taken as a proxy. This variable includes information on the occupational status of children of the respondent. The original dataset contains 20 separate variables with information about the occupational status of child number 1 up to number 20. Only 14 of these variables are eventually used to create the new variable, because no participant in the dataset indicates to have more than 14 children. These 20 original variables contain categorical information about the occupational status of children, divided into 10 categories, such as unemployed, in education, part-time employed, sick or disabled. Moreover there are three categories that do not provide any good data to

work with: ‘don’t know’, ‘other’ and ‘refusal’. Besides these ‘empty’ categories, there is missing data for a large number of respondents. There could be multiple reasons for this. The most obvious reason for missing data is that the respondent does not have children, because then the question about occupational status of children won’t be asked. However there could also be other reasons why data is missing, such as a refusal to participate or nonresponse to this specific question. The variable that has been created is a dummy variable that shows whether a child in unemployed or not. The variable has two options: the participant has an unemployed child (=1) or not (=0). In the category of not having an unemployed child, all other original categories are included, and this thus ranges from employed children, to children in education, but also to not having children at all. The created binary variable has advantages and disadvantages. An advantage is that it is clear and focusses on the most important information that is needed: does the respondent have an unemployed child. Disadvantages are mainly the wide range of different categories that are included in the only other category. An issue is for instance that not all individuals in the dataset are parents. This is solved by marking those individuals as not having an unemployed child, and therefore including them into the same group as a parent with employed children or children that are in e.g. education. This keeps the variable binary. However, one could say that by doing this important information is left out. It could thus be argued that it would be better to create a categorical variable with three categories (unemployed child; no child; all other). This is a still a small number of categories and therefore the variable is easy to interpret, just as the dummy variable and moreover this variables contains a bit more information than its binary counterpart. However, it is hard to create this variable without mistakes. The ‘no child’ category should be created with the missing cases, it is however unsure whether these cases are missing because the respondent has no child, or for other reasons such as nonresponse. Moreover, a tabulation shows that many respondents without children still answered the question on occupational status of children but answered ‘don’t know’ or ‘other’. Therefore it is better for the trustworthiness of the newly created variable to only focus on respondents that actively stated to have unemployed children. Besides, the main goal of the variable is to check whether parents of unemployed children have different health outcomes than all other respondents. Therefore other categories such as the category of not being a parent or having employed children are not as important (less relevant for the research question) and don’t necessarily need their own category in the newly created variable. Therefore, the created dummy variable is the most accurate and preferred variable to use. The effect of having an unemployed child is tested by including this dummy variable into the statistical model.

3.2.3. Statistical analysis

To answer the research question, two multiple linear regression analyses will be executed. This is a linear regression analysis that includes one continuous dependent variable and multiple continuous or categorical independent variables (Mehmetoglu & Jakobsen, 2017). The dependent variable will be the

health score, either PCS or MCS. The independent variables will be the ones mentioned in the previous chapter. Additionally, other commonly used independent variables such as age, gender and years of education are included into the analysis.

Sometimes the effect of a predictor in a regression model is not immediate (Hyndman & Anthanasopulos, 2018). In the case of this research, where the focus is on both short-term and longer-term effects of economic crisis on health of older individuals, this is also the case. Therefore, lagged predictors will be introduced into the statistical mode, as has been explained in the former paragraph. The health score is measured later than the explanatory variables, thus the explanatory variables are the predictors that are tested to have a lagged effect on the later health core.

The software package that will be used to perform the analysis is Stata (17.0 SE-Standard Edition).

4. Results

The results chapter of this thesis is divided into different paragraphs. First, some descriptive statistics and information about key variables and the dataset in general will be given. Hereafter, the results from the analysis of the short-term health effects of the economic crisis for Spanish individuals aged over 50 will be discussed. Both physical health and mental health will be covered. The last paragraph will dive into the longer-term health effects of the economic crisis.

4.1. Descriptive results

4.1.1. Income and health over time

In this paragraph, the relationship between income and health over time among the respondents of the survey in Spain is explored. This has been done on a societal level (full dataset) instead of at an individual level to get a broad picture of a possible trend in the relationship between income and health. From the SHARE data Wave 2, 4, 5, 6 and 7 (2007, 2011, 2013, 2015 and 2017), the generated household income variable is used. Only individuals aged above 50 in the full dataset are included. Of the income variable, the average is taken. The average yearly household income of Spanish individuals aged above 50 that are included in the SHARE dataset can be found in figure 5, displayed on the left vertical axis. In this same graph, on the right vertical axis, the average self-perceived health rating of all respondents of 50 years and older is displayed. This is the self-perceived health of respondents on a scale from one to five, with one being *Excellent* and five being *Poor*.

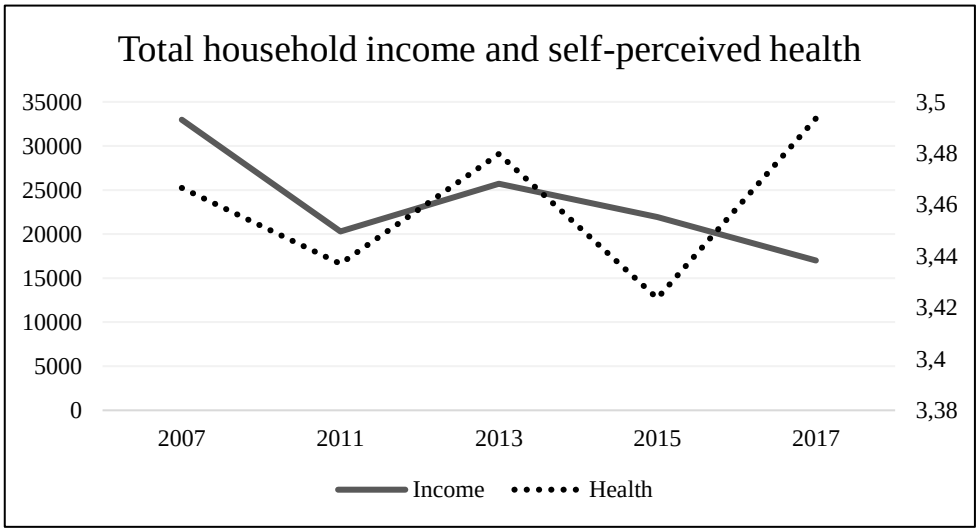


Figure 5 Average annual total household income in euros and self-perceived health among individuals aged above 50 years in Spain over time, data from the SHARE dataset.

As mentioned in the previous chapter, the preferred measure for health is the newly generated SF-12 variable, divided into PCS and MCS. However, this variable has only been generated for Wave 4 and Wave 7, and not for the other years, because not for all years the same data is available. Moreover, the SF-12 variable cannot be compared over time between different waves, because the score is relative to other respondents within the same dataset (so within the same wave) and also differs from wave to wave when it comes to the variables that have been used to create the score. Therefore, in this descriptive graph of health compared to income change over time *'self-perceived health'* will be used instead. Besides this graph, this measure won't be used further in this thesis. In the statistical regression analysis, the SF-12 variables are the preferred health measures, as has been explained in the previous chapter.

As can be seen in the graph, there is no clear relationship between self-perceived health and average income in the dataset over the entire timeframe. Possibly, a pattern between income and health can be noticed between 2007 and 2015, but after 2015 this pattern is absent. Nevertheless, there are some conclusions that can be drawn from the data that is visualized in the graph. E.g. the trend of income is downward. Especially between 2007 and 2011, the first years of the economic crisis, average household income decreased sharply, from €32975 annually in 2007 to €20309 annually in 2011. Another factor that could have played a role in this decrease of average income is the average age of the respondents in the dataset. In 2007 this was 67.47 years, whereas in 2011 the average age was 68.35 years. This means that more people have probably retired, possibly affecting incomes. In 2013, the average income increased slightly again. After this year the average income slowly decreased each new wave. Once again, age can play a role here: see table 3.

Table 3 Average age in the Spanish SHARE dataset (Waves 2-7, without Wave 3).

	Mean	**Std. Dev.**	**Low**	**High**
2007 (Wave 2)	67.47	10.71	50	104
2011 (Wave 4)	68.35	10.87	50	102
2013 (Wave 5)	68.67	10.91	50	104
2015 (Wave 6)	70.30	10.56	50	106
2017 (Wave 7)	71.57	10.29	50	102

Looking at figure 5 again, no real trend can be seen in self-perceived health. The scores are very volatile, changing each year, although yearly differences are small. Lower scores are observed in 2011 and 2015. Results are unclear without executing a proper statistical analysis. Therefore, a regression analysis has been used in this thesis to research a possible link between health and different economic variables.

4.1.2. Physical and Mental Component Scores

As explained in the previous chapter of this thesis, an SF-12 score, divided into a Physical Component Score (PCS) and a Mental Component Score (MCS), has been calculated for all individuals in the dataset. This has been done for both Wave 4 as for Wave 7. Because not all individuals in the dataset had observations for all necessary items that are used to build the SF-12 scores, the number of observations in the dataset has declined after the creation of these scores. In table 4, descriptive summary statistics can be seen for the generated PCS and MCS in Wave 4 and Wave 7 of the SHARE dataset.

Table 4 Summary statistics for the calculated SF-12 scores.

	Wave 4 N = 3872		**Wave 7** N = 4711	
	PCS	**MCS**	**PCS**	**MCS**
N	2535	2535	3092	3092
Mean	49.98	49.98	50	50
Std. dev.	8.37	6.17	8.33	6.86
Low	24.92	28.67	19.02	18.79
High	63.22	62.45	63.29	61.09

For histograms of the PCS and MCS in Wave 4 and Wave 7, see appendix 4.

A list of all important variables and their summary statistics can be found in appendix 8.

4.2. Short-term health effects of the economic crisis

To study the short term health effects of the economic crisis on people aged above 50 in Spain, two linear regression analyses have been performed. In the first regression, the SF-12 Physical Component Score (PCS) has been put into the model as dependent variable. In the second regression, the SF-12 Mental Component Score (PCS) is the dependent variable. All other explanatory and control variables in both regression models are the same. This way, the effects of the same (economic) variables are tested for both physical and mental health separately.

4.2.1. Physical health (PCS)

See appendix 5 for the regression analysis table that is discussed in this paragraph. Five different models have been created. Model 1 (r2=0.142) only includes control variables and variables on demographics such as age and gender, whereas the subsequent models also include the independent variables that are of special interest for this research, such as income loss and changes in employment status. In model 1, the number of observations is 2535. In models 2, 3, 4 and 5, the number of observations has lowered to 1004. This can be explained by the fact that not all individuals in the dataset have observations for the

more specific (economic) variables, but all individuals in the dataset do have observations for the more basic variables that contain demographic information.

In model 2, besides the variables that were already included in model 1, the first economic variables have been introduced. Two simple variables on income change were included, both without statistically significant effects on the physical health score and only resulting in a small increase of the proportion of explained variance (r2=0.146, an increase of 0.004). In model 3, three new variables were included, next to the previously included variables from models 1 and 2. The first variable contains information about in which income quantile an individual was placed prior to the crisis, in 2007. The other is on how the individual's income has changed (seven categories, ranging from severe income loss to very strong income growth). The third variable that was added in this model is not a normal variable, but an interaction between the two previously mentioned variables. Some of the categories from the newly added variables or interactions do show statistically significant effects. Model 3 results in a relatively larger increase of the explained variance, with an R-squared of 0.184. In model 4, a few new variables have been added to the regression equation once again, resulting in an increase of the R-squared to 0.202. The variables that have been added in model 4 contain information on employment status of the respondents prior to the crisis in 2007, and on the short-term during the crisis in 2011. The variables are binary variables. Interactions have also been included. These interactions contain information on whether the individual has maintained their job or lost their job. Retirement is not included.

Looking at the R-squared, model 5 explains the largest proportion of the variance in the dependent variable, namely 20.3% (r2=0.203). This is approximately the same value as model 4 (r2=0.202), so the newly added variable does not make a large difference. Model 5 also includes the most independent variables and interaction into the regression of all five models. Thus, in this paragraph the results that we can derive from model 5 are mentioned and discussed, because this model provided the highest R-squared value and is the most complete model.

The following equation is the linear regression equation of model 5:

$$\begin{aligned} PCS_{it} = {} & \beta_0 + \beta_1 age_{it} + \beta_2 marital\ status_{it} + \beta_2 number\ of\ children_{it} + \beta_3 gender_{it} \\ & + \beta_4 years\ of\ education_{it} + \beta_5 total\ household\ income_{it} + \beta_6 income\ change_{it} \\ & + \beta_7 income\ loss_{it} + \beta_8 income\ quantile_{it-1} + \beta_9 income\ change\ categories_{it} \\ & + \beta_{10} income\ quantile_{it-1} income\ change_{it} + \beta_{11} unemployed_{it-1} \\ & + \beta_{12} unemployed_{it} + \beta_{13} employed_{it-1} + \beta_{14} employed_{it-1} unemployed_{it} \\ & + \beta_{15} unemployed_{it-1} unemployed_{it} + \beta_{15} occupational\ status\ child_{it} + \varepsilon_{it} \end{aligned} \quad (5)$$

In this equation, β_0 represents the constant. The other betas are the coefficients of the corresponding independent variables or interactions. PCS is the dependent variable (the physical health score). The i is the individual and t is time. Time is normally the current year (2011, Wave 4), but when '-1' is added, it is a variable from 2007 (Wave 2). Lastly, ε represents the error term.

One of the basic assumptions for a linear regression is that there is no correlation between the independent variables in the model. To check this, a regression check has been performed in Stata, which computes, among others, Variance Inflation Factors (VIF) to check for multicollinearity. No multicollinearity problems were found. Hence, the use of model 5 can be continued. This check has been particularly important for this specific model because multiple variables that, in theory, are somewhat related to each other are included into the model, such as the different variables for income change.

It can be seen that age, gender and years of education all have a significant effect on someone's PCS. Model 5 shows that an individual's PCS decreases significantly ($p<0.05$) with age: a decrease of 0.103 points per year. Model 5 also shows that gender plays a role in someone's PCS. Women's PCSs are significantly lower ($p<0.001$) than those of men: 4.083 points lower. Individuals that have enjoyed more years of education have a significantly higher ($p<0.001$) PCS, and therefore a significantly better health. For each additional year that someone has been in education, the PCS increases with 0.327 points. Of the control variables, marital or relationship status (binary: being single or not) and the respondent's number of children do not have a significant effect on the physical health score.

After looking further into the regression table, it becomes clear that not many variables or interactions have a statistically significant effect on someone's PCS. However, there are still some results from model 5 that can be discussed.

The binary variable *'Income loss between 2007 and 2011'* shows whether an individual has experienced income loss (of any kind) between the years 2007 to 2011 (during the economic crisis) or not. The effect of this variable is negative, meaning that individuals that experienced income loss in this time period have a lower health score on the short-term. However, this effect is not statistically significant. When looking at income in 2007, prior to the crisis, two interesting results should be mentioned. As compared to the middle income group (the third quintile, which is around the median income in the dataset), individuals with the lowest and individuals with the highest income have a significantly lower PCSs ($p<0.05$). Looking at the hypotheses that income and health are related, this is a surprising result: the result for the lowest incomes was expected, whereas the result for the highest income group was not. The next variable in the regression is *'Income change between 2007 and 2011'*. This variable is a categorical variable with seven categories. All categories (except for one) of income change show a negative effect of the income change on the PCS when compared to the group that had the strongest

income growth in the period 2007-2011, but not all effects are statistically significant. The strongest effect of income change is observed among individuals that have experience some income change in the period 2007-2011. Individuals that experienced some income loss in this time period are expected to have a PCS that is 8.277 points lower than those of individuals that experienced a very strong income growth over the same time period, on a statistical significance level of 1% ($p<0.01$).

Model 5 also includes an interaction between *'Income quantile in 2007'* and *'Income change between 2007 and 2011'*. This interaction between the two categorical variables leads to 20 interaction variables. Out of these 20 interactions, only tree interactions have a p-value of lower than 0.05, of which just one interaction has a p-value lower than 0.01. Additionally, three interactions give a test result with a p-value between 0.1 and 0.05, which is generally not considered to be statistically significant and therefore also remains undiscussed in this thesis. The largest (positive) effect, that is moreover the most statistically significant ($p<0.01$), is seen in the interaction between the fifth quintile and some income loss. Individuals from the highest income group in 2007 that experienced some income loss during the economic crisis are expected to have a 10.583 points higher PCS on a 1% significance level. The most important result from this interaction is however that there is no clear pattern visible in the regression results.

The last set of variables in the regression model relate to employment status, as already briefly mentioned previously in this chapter. It can be seen that both employed individuals in 2007 as unemployed individuals in 2007 have a significantly ($p<0.001$ and $p<0.05$) higher PCS in 2011 than other groups. Other groups are in this case for instance people that are retired, inactive or permanently sick. An interaction has also been included to check whether people that lost their job during the crisis and people that remained unemployed during the crisis have different health outcomes than other groups. Both interactions show relatively large negative effect, however no statistically significant effect was found. Only one new variable has been added in model 5, as compared to the previous four models. This variable also related to employment status, however not of the respondent themselves but of their children: *'Occupational status of the child: unemployed'*. The variable is a binary variable that shows whether the respondent has one or more unemployed children. Having an unemployed child has a small negative effect on someone's physical health, however also this effect is not statistically significant.

4.2.2. Mental health (MCS)

See appendix 6 for the regression table that corresponds to the results that will be discussed in this paragraph. As can be seen in the regression table, the exact same variables and interactions have been included into this regression analysis on mental health on the short-term as in the previous analysis on physical health on the short-term. Therefore, this paragraph will be shorter and more concise: not all

variables will be discussed, only relevant and statistically significant effects will be mentioned. Just as in the regression with PCS, the last model that includes most variables (model 10) has the highest R-squared (r2=0.123). This is following the expectation that larger models tend to explain a larger proportion of the variance in the dependent variable. Model 10 will be used in this paragraph.

Below, the equation of the linear regression of model 10 can be found. This is the same as model 5, except for the dependent variable:

$$\begin{aligned} MCS_{it} = \beta_0 + \beta_1 age_{it} + \beta_2 marital\ status_{it} + \beta_2 number\ of\ children_{it} + \beta_3 gender_{it} \\ + \beta_4 years\ of\ education_{it} + \beta_5 total\ household\ income_{it} + \beta_6 income\ change_{it} \\ + \beta_7 income\ loss_{it} + \beta_8 income\ quantile_{it-1} + \beta_9 income\ change\ categories_{it} \\ + \beta_{10} income\ quantile_{it-1} income\ change_{it} + \beta_{11} unemployed_{it-1} \\ + \beta_{12} unemployed_{it} + \beta_{13} employed_{it-1} + \beta_{14} employed_{it-1} unemployed_{it} \\ + \beta_{15} unemployed_{it-1} unemployed_{it} + \beta_{15} occupational\ status\ child_{it} + \varepsilon_{it} \end{aligned} \quad (10)$$

For model 10, a check for multicollinearity has been performed in the same way as has been done for model 5. Variance Inflation Factors (VIF) were computed using the regcheck command in Stata. No multicollinearity problems were found.

To start, age is not significantly related with mental health in this sample. Single individuals have a statistically significant ($p<0.01$) lower mental health score (MCS). Their MCS is expected to be 1.576 points lower than those that are not single. People with more years of education tend to have a significantly ($p<0.01$) better MCS, of 0.139 points per year of education. Women have a significantly ($p<0.001$) lower MCS as compared to men, with 3.203 points less.

Individuals that were in the two lowest income quintiles in 2007, have a lower MCS score in 2011, as compared to individuals in the median income group: the third quintile. This is statistically significant on the 10% significance level only. The effect is negative and relatively large. Perhaps surprisingly, individuals that experienced a very strong income loss during the economic crisis, between 2007 and 2011, have a significantly better ($p<0.05$) mental health than people with very strong income growth in the same time period. When looking at the interaction between income quintiles and income change, it becomes clears that this effect is mainly true for individuals in the highest income groups in 2007. Hardly any significant interactions were found, besides these interactions between individuals in the fifth income quintile that experienced strong income loss and those that experienced some income loss. Different explanations can be possible, and these will be discussed in the discussion chapter of this thesis. It could for instance be that those individuals have already accumulated a relatively large wealth throughout their lifetime and now retired. They therefore lost income, but do not suffer from work-

related stress anymore and due to their own wealth do not have monetary issues. This can have a large positive effect on their mental health and thus could explain the observed large positive effect of the interaction between the fifth income quintile and income loss on individual's MCSs.

Respondents from the sample that were employed in 2007 have a significantly ($p<0.01$) higher MCS than people that were not employed. For unemployed individuals in 2007 the effect on MCS is negative, however this effect is not statistically significant. The same goes for people that lost their employment during the economic crisis and did not find a new job yet in 2011. A negative effect on MCS was observed here, but the effect was once again not statistically significant. Moreover, the occupational status of children also does not seem to have a significant effect of mental health scores of parents.

4.2.3. Justification of chosen models and the exploration of alternative models

In addition to the two main models that have been discussed so far in this chapter (model 5 and 10), various other models and variants of models have been tested as well. Among these various models there are models that include different variables for the income quantiles and the income change categories. In the original and used models these categorical variables are divided into five and seven categories. This has as a result that the interaction between these two variables leads to a large amount of interaction variables in the final model. Consequences are that this can lead to confusing and unclear regression tables, but moreover that the number of cases in each category or interaction are small and thus will not lead to significant results. Looking at the regression tables, this is true, and not many significant results were found (although some were found, as mentioned in the previous paragraphs). One could therefore suggest to create similar variables with less categories, for instance three categories for the income group variable and five for the income change category variable, instead of the previous five and seven categories. This has therefore been done, however these models did not lead to improved results. Still the number of significant results and interactions was low among these variables. Moreover, the R-squared declined by a few percentage points. Hence, the changes did not lead to improvements in the fit of the model and also did not increase the number of significant results. The new models however did lead to a smaller and slightly clearer regression table, though still many separate categories and interactions were included.

4.3. Longer-term health effects of the economic crisis

To research the longer-term health effects of the economic crisis on the elderly population in Spain, a similar research design as used for the short-term is needed. This means that two health variables are necessary: one for physical health and one for mental health. As previously explained in the research

design chapter, the SF-12 variables (PCS and MCS) have therefore not only been made for Wave 4 of the SHARE dataset (2011), but also for the most recent wave: Wave 7 (2017).

Moreover, the same explanatory variables as in the short-term analysis are preferably needed. Therefore the various variables on loss of income between 2007 and 2011 and employment change in this same period have been used again in a new regression analysis. This has however caused problems. The sample size ranges, depending on the used model, between 13 and 97 observations. This is too small to draw conclusions that apply to the Spanish elderly population. In 2017, the population aged 50 and over in Spain was namely over 18 million people (INE, 2021b). Mehmetoglu and Jakobsen (2017) and Krejcie & Morgan (1970) argue that for a population size over 100.000, the minimum sample size should be 385. Other sources are even stricter and require larger sample sizes (e.g. Israel, 1992). Therefore, a regression analysis is not a suitable method to research the longer-term health effects of the economic crisis. All things considered, this makes it thus impossible to carry out a solid regression to research longer-term effects of the economic crisis on health of elderly in Spain, by assessing all economic variables that are of interest to this research. A larger sample size and more values for the necessary variables would have been needed for this. On a European scale for instance this research would plausibly be possible using the SHARE dataset, but on this smaller country-level scale it is not due to lack of data. Therefore other research techniques have been explored, among which descriptive statistics. These include simple graphs and tables and can only be used to describe the data and carefully possible relationship, but no strong statements can be made.

4.3.1. Physical health (PCS)

The scatterplots and tables that are discussed in this chapter are not included in the text, nor in the appendices, because they are not of great importance for understanding and clarification. They can be obtained by contacting the author of this thesis, if this is wished for.

As explained in the former paragraph, the number of cases for some of the key variables that are needed for the analysis on the longer-term is too low to perform a regression analysis. Therefore different ways to analyze the available data need to be explored. For the first main explanatory variable, the binary income loss variable, plots and tabulation have been made using Stata. Only 13 observations are available, which makes it hard – if not impossible – to make statements about the influence of income loss on longer-term health outcomes. Analyzing the graphs and tables, no clear difference or trend can be observed between both groups (income loss between 2007 and 2011: yes/no) in terms of their PCS in 2017. Additionally, a t-test has been performed to check whether there is a statistical difference in the mean of the PCS of both groups (income loss: yes/no). There was no difference according to the t-test, but even more importantly so, the conditions to perform this t-test had not been met.

A scatterplot has also been made to explore a relationship between income change in euro's between 2007 and 2011 and PCS in 2017. The income change variable also lacks a sufficient number of cases (n=13) and no clear relationship between income change and the health score can be observed in the plot.

Nothing meaningful can be concluded regarding the longer-term effects of job loss in the economic crisis on someone's PCS. Again, the number of cases is very low, especially for the category of interest: the individuals in the dataset in 2017 that lost their job between 2007 and 2011 (n=3). No good longer-term analysis can thus be performed and the effects of job loss in the economic crisis on physical health on the longer-term.

Table 5 The relationship between health in 2017 and the Great Recession in Spain of 2008.

PCS and MCS linear regression models (longer-term)				
	PCS 2017		*MCS 2017*	
	MODEL 11a	**MODEL 12a**	**MODEL 11b**	**MODEL 12b**
Single				
Yes	-1.574***	-0.613	-0.299	-0.001
	(-4.52)	(-0.96)	(-0.99)	(0.00)
Total household income in 2017	0.000	0.000	-0.000	-0.000
	(0.35)	(0.26)	(-0.00)	(-0.02)
Gender				
Female	-1.193***	-1.550**	-2.142***	-2.953***
	(-4.18)	(-2.91)	(-8.63)	(-6.36)
Age	-0.223***	-0.248***	-0.052***	-0.050+
	(-15.06)	(-8.49)	(-4.07)	(-1.96)
Years of education	0.199***	0.178***	0.095***	0.125**
	(6.95)	(3.33)	(3.83)	(2.68)
Number of children	0.003	0.020	0.000	-0.002
	(0.18)	(0.83)	(0.01)	(-0.07)
Occupational status of child				
Unemployed		-1.324+		-0.751
		(-1.77)		(-1.15)
Constant	66.100***	70.521***	56.119***	57.498***
	(31.54)	(20.38)	(30.79)	(19.08)
N	3092	1066	3092	1066
r2	0.1416	0.1229	0.0420	0.0574

T statistics in parentheses
*+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001*

Having dealt with the longer-term effects of income loss and employment status change on physical health, only one of the three major variables of interest to this thesis remains unmentioned: concerns about children and their employment status. In table 5, it can be seen that a small regression analysis has been performed. The economic variables that are included into the regression models for the short-term have been left out of this regression due to the low number of cases, however the variable on employment

status of children does contain more observations and could therefore be tested in model 12a. Having unemployed children in 2011, during the economic crisis, has a statistically significant (90% confidence level; p>0.1) negative effect on PCS in 2017. This variable functions as a proxy for concerns about children, and therefore it can carefully be argued that parents that had concerns about their children in 2011 had a significantly lower physical health score in 2017. The addition of this variable in model 12a however does not improve proportion of explained variance and the goodness-of-fit of the model, as compared to model 11a. The R-squared even declines.

4.3.2. Mental health (MCS)

In this paragraph, the exploration of a possible relationship between mental health and the economic crisis on the longer-term is discussed. The same analysis has been done for MCS as for PCS in the former paragraph, therefore not much additional explanation is needed and thus this paragraph is short and concise. Just as for PCS, no relationships between MCS and income loss (binary variable), income change in euros (continuous variable) and job loss (binary variable) can be observed in the graphs and tables, mainly also because of the small n. Moreover, in the regression analysis of MCS and concerns about children (employment status of children), as can be seen in table 5, no statistically significant relationship is visible. It can thus be said that this research cannot conclude – nor deny – whether the economic crisis in Spain has had an effect on mental health of Spanish elderly in the longer-run.

5. Discussion

In this chapter, results of this thesis are discussed and explained. Moreover, a critical assessment of the used data and methods is made. Limitations of this research and recommendations for further research are also included.

5.1. Key findings, their interpretation and their implications

5.1.1. Income

Individuals in the lowest and the highest income groups in 2011 both have a significantly lower PCS in 2011. For the lowest income groups this is in line with previous research: as seen in the introduction and in existing literature, many scholars have found relationships between income and physical health. (e.g. Stuckler et al., 2017; McKee et al., 2012; Carballo-Cruz, 2011). This does however not tell us anything about the impact of the crisis, but solely shows a correlation between current income and current PCS (in 2011). For this, income loss during the crisis has to be looked at. This research shows that on the short-term income loss itself is not a predictor for a lower health status on the short term (PCS and MCS). This changes a bit when the severity of this income loss is taken into account. It was found that individuals with some income loss have a much lower PCS than individuals that saw a strong income growth. This effect was not found for MCS. This can be seen as surprising, because mental and physical health are often related, especially when the effect on physical health is so strong as in this specific situation.

Another interesting finding is that very strong income loss has a significantly positive effect on mental health. This is remarkable. When noticing the interaction however, this results becomes a bit more understandable, because the effect is especially positive and significant for individuals in the fifth income quintile, meaning the highest income group. This could possibly be because of retirement of people in the highest income group, meaning that they will lose a large amount of their income, but still not be in an economically stressful situation because they were able to save up during their lifetime due to their high income. Moreover, this group will be more relaxed and less stressed after retirement because the pressure of work is gone and they can enjoy their free time.

5.1.2. Employment status

Job loss is one of the main predictors for health status in the conceptual framework of this thesis. The effects of job loss in the economic crisis (between 2007 and 2011) has been researched, but no significant effects on health status were found. Not on PCS and not on MCS. The effects of job loss on PCS are

relatively large and negative, but as said, not significant. This thesis can thus not conclude whether job loss has an effect on short-term health outcomes of Spanish people aged over 50. This is in line with findings by Lorenzo-Carascosa (2018) and Blázquez-Fernández et al. (2021), but not with most existing literature, that suggests that job loss causes stress and is bad for someone's health (Layte & McCrory, 2018; Whitehead & Bergeman, 2017), both physical, but especially mental health due to e.g. the feeling of being left-out. Why could this be? One reason could be because the target group of this thesis is older individuals. These might act and think differently than the general population, with people of all ages, who are generally the research target of existing research. Older people may be less stressed about losing their job because they will retire relatively soon anyway. Another reason could be that they may have been able to save up more during their life already, which could mean that job loss does not necessarily lead to stress about paying groceries, rent, mortgage or other bills. Often, older people also do not have to provide for their children anymore, because they have jobs for themselves already and are less likely to still live at home. Another reason for the divergent results could be the context of Spain, which can be different than that of other countries that were researched in existing literature. In Spain and other Mediterranean countries, family is for instance a very important provider of support and personal and social wellbeing, more so than in Northern European countries (Del Valle et al., 2013). One of the main reasons that no effects of job loss was found on the short-term however, is probably because the number of people that indicated to have lost their job in the dataset is simply very low.

An important issue with the job loss variable that was used in this research, is the way it was created. People that were marked as having lost their job, are individuals that had a job in 2007 at the moment of the interview and did not have a job in 2011 at the moment of the interview (and were looking for a job at that moment, i.e. not being in education or retirement). This means that the variable does not catch all people that for instance lost their job in this period and have been jobless for a while, but found a new job again at the moment of the interview in 2011. Contrarily, individuals that lost their job in 2011, just before the interview, were marked as having lost their job, even if they may have found a job again right after the interview. The issue with the variable for job loss is thus that the duration of the jobless period is not taken into account. There is no available data for this in the SHARE dataset. This could have affected results.

Another interesting finding concerning employment status is that, on the short-term, both employed and unemployed individuals tend to have a better PCS than other groups. One could say this is surprising, but it does make sense. The other groups namely also include sick or disabled people and retired people. These in theory indeed often have worse health than younger and more fit people that are still in the workforce, regardless of their employment status.

5.1.3. Concerns about children

On the short-term, no statistically significant relationship was found between health (both mental and physical) of parents and employment status of their children. This opposes some of the qualitative results found by Chalari (2014), who found that parents had concerns about their children, which had implications for their health, mainly mental health. However, besides this Greek study, not many similar results have been found in existing academic literature. This means that the 'reversed' family stress model of economic hardship that was proposed by me in the theory chapter of this thesis cannot be substantiated using the results from this thesis. There is no proof of a possible 'reversed' version of the model, however, I would still not rule out that in some cases concerns on children's economic wellbeing may have implications for health of parents. Reasons why no relationship could be found on the short-term could be that the number of unemployed children was low, even though this in unlikely due to the high youth unemployment rate in Spain (OECD, 2021). This high unemployment rate for youth could also be a reason why no relationship had been found: simply because so many young people were unemployed, a parent can be less stressed about it because the situation of their child does not negatively differ strongly from those of other children. Another reason why no significant relationship could have been found specifically on the short-term could be because it is researched on the short-term. An unemployed child in 2011 can lead to stress in 2011, but this stress may not yet have influenced health yet. This could be a delayed or lagged effect. Therefore the longer-term analysis is even more interesting. The question is however how long the time difference between the triggering situation and the effect of the subsequent stress on health should be. Research has suggested that effects are not immediate, but Ibrahim et al. (2009) found that also on the longer-term the effect diminishes. The effect was in their research namely larger after two, than after six years.

In this research, the longer-term effect is research after six years (2011-2017). In the longer-term regression model an effect was found of unemployment of the child in 2011 and physical health in 2017, however the effect was only significant on a 10% significance level. Following the research results of Ibrahim et al. (2009) this effect could have thus been larger if the longer-term would have been measured a bit closer to 2011, but this has not been tested in the research. No effect was found on the MCS of parents on the longer-term.

5.2. Limitations

In this paragraph, some of the limitations of this research and the used data are discussed.

The SHARE data provides lots of opportunities for research, as the dataset contains many variables on various topics. In the Spanish sub dataset however, there were many missing values on some variables.

Theses missing values are in large part explained by the nonresponse on certain questions. Bergmann et al. (2019) show in their paper on the data quality of the SHARE data (Wave 7) that in Spain nonresponse is exceptionally high on monetary and economic questions. On questions about the value of their house, more than half (54%) of all Spanish respondents did not give an answer. Similarly, a bit less than 40% of Spanish respondents did not answer the question about annual household income and a comparable percentage of nonresponse was observed on a question related to their capital. These numbers are much higher than of any other country that was included in the dataset. Southern European countries showed higher nonresponse in general on economic and monetary questions than Northern and Eastern European countries (Bergmann et al., 2019). These monetary and economic variables are of special interest for this research, and include some of the important independent variables that were included in the regression analyses, therefore it can cause problems for the number of cases. This was indeed the case in the data of Wave 7, where there were too few cases to execute a regression analysis. Another problem with the high nonresponse level in these exact type of variables (economic and monetary) is that this could signal that people find questions related to money or wealth awkward to answer. Important data could therefore be lacking. The high nonresponse on these questions could therefore also possibly lead to biases and subsequently affect results. Especially if one group is over- or underrepresented. E.g. if poorer people that were hit harder by the economic crisis tend to be less likely to answer these questions, this could lead to biases. But also the contrary could happen, when people that are in a financially comfortable situation tend to not answer these questions it could lead to biases in the other direction.

The low number of cases for the analysis on the longer-term cannot only be explained by high levels of nonresponse. Issues with longitudinal data and people leaving the survey play an even larger role. For the short-term this is not a particularly large issue. Between 2007 and 2011 there is only a four-year gap. Respondents can definitely leave the survey in this time period for various reasons, but in only four year the outflow of respondents can remain modest. Therefore the analysis on the short-term had still been performed without any large issues. On the longer-term however, this is different. To research longer-term effects of the crisis, people that had been in the survey in 2007, still had to be participating in 2017. This is ten years, which is a lot of time for people to leave. Especially older individuals are more likely to leave the survey, because many of them could die in this time period.

Another, more theoretical, limitation of this thesis is the limited role that the life course health development framework has played in the operationalization of this thesis. The life course health development framework is important in research on health and how different determinants influence health over the life course. Therefore this framework has been added to the conceptual model that is used in this thesis. However, only one of the many determinants that could be part of this framework is tested in this research: the Great Spanish Recession of 2008. Other important determinants that could have been of importance have been left out, such as the Spanish Civil war or previous or later crises

(e.g. economic, political, nutritional). Moreover individual level determinants are lacking as control variables in the regression models. Only a few have been added, such as age and education. Important in utero or childhood conditions could not be added due to lack of data. Adding all important life course determinants of health would have been too ambitious or even impossible. It is however good to mention that the models in this research could never include the entire framework, and only look at the added effect on PCS and MCS of the economic crisis and its corresponding variables.

5.3. Recommendations

Recommendations for further research are briefly mentioned in this paragraph. A first important recommendation is to research the effect of economic crises on health of elderly on a different scale, because of the data issues on this country-level research with the SHARE dataset. If a European-scale was chosen, this could have possibly been avoided. Another recommendation is to involve elderly more into academic research problems. Most existing literature is on younger individuals or individuals in the working ages. With ageing populations in Spain, and the rest of Europe, it is important to not neglect the growing group of older individuals in our societies. Another recommendation for future research is to find a new broad measure of 'being affected by economic crisis'. Maybe inspired by the SF-12 score, except in this case it is not a health score but a score that can show how strongly an individual had been affected by an economic crisis. This could make research on the effects of economic crises on health, or other subjects, easier in future research. A last recommendation is to research the effects of concerns about children on the health of older individuals a bit better. Youth unemployment is extremely high in Spain, possibly leaving many parents concerned. What does this do to their health? And what can we do about it? This thesis suggests that on the short term, health is not really affected, whereas on the longer term a small effect is seen on physical health. This could be a topic that deserves further investigation.

6. Conclusions

The main research question of this thesis, related to the research objective that has been defined in the introduction, is: *What are the short- and longer-term effects of the economic crisis in Spain on the health of people aged above 50?* To answer this relatively broad question, several sub-questions have been formulated as well, which can be found in chapter 2. These questions have helped to guide the operationalization of the research.

On the short-term no clear conclusion can be drawn. There are definitely some effects of economic indicators on health of older Spanish individuals, but only in some situations and for some groups. This thesis can however not say that this is all due to the economic crisis as a broad concept, or merely due to these separate economic indicators such as income loss for a certain income group. Concerns about the employment status of children seems to have no immediate effect on physical and mental health of older Spanish individuals (50+) on the short-term. On the longer-term this thesis found even less results. This is mainly due to the lack of sufficient (longitudinal) data, because too few individuals stayed in the SHARE survey between 2007 and 2017. Moreover many important economic variables contain high percentages of missing values. It does seem however that having an unemployed child during the economic crisis can have small negative physical health effects on the longer-term. Because of high youth unemployment rates and a large elderly (and moreover ageing) population in Spain, it could be interesting to research this relationship further.

References

Andersen, H.H., Mühlbacher, A., Nübling, M., Schupp, J., & Wagner, G.G. (2007). Computation of standard values for physical and mental health scale scores using the SOEP version of SF-12v2. *Schmollers Jahrbuch, 127*(1), 171-182.

Arcaya, M.C., Arcaya, A.L., & Subramanian, S.V. (2015). Inequalities in health: definitions, concepts, and theories. *Global health action*, 8(1), 27106.

Arechavala, N.S., Espina, P.Z., & Trapero, B.P. (2015). The economic crisis and its effects on the quality of life in the European Union. *Social Indicators Research*, 120(2), 323-343.

Ásgeirsdóttir, T. L., Corman, H., Noonan, K., Ólafsdóttir, Þ., & Reichman, N. E. (2014). Was the economic crisis of 2008 good for Icelanders? Impact on health behaviors. *Economics & Human Biology*, 13, 1-19.

Ayuso-Mateos, J.L., Barros, P.P., & Gusmão, R. (2013). Financial crisis, austerity, and health in Europe. *The Lancet*, 382(9890), 391-392.

Beker, V.A. (2014). The European debt crisis: Causes and consequences. *J Stock Forex Trad*, 3(2), P1-9.

Ben-Schlomo, Y. & Kuh, D. (2002). A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives. *International Journal of Epidemiology*, 31, 285-293.

Berger, N., Van der Heyden, J., & Van Oyen, H. (2015). The global activity limitation indicator and self-rated health: two complementary predictors of mortality. *Archives of Public Health, 73*(1), 1-7.

Bergmann, M., Scherpenzeel, A., & Börsch-Supan, A. (2019). *SHARE Wave 7 Methodology: Panel innovations and life histories.* Munich: MEA, Max Planck Institute for Social Law and Social Policy.

Bhattacharya, J., Hyde, T. & Tu, P. (2014). *Health Economics*. New York: Palgrave Macmillan.

Blázquez-Fernández, C., Cantarero-Prieto, D. & Pascual-Sáez, M. (2021). Quality of Life, Health and the Great Recession in Spain: Why Older People Matter? *International Journal of Environmental Research and Public Health*, 18(4), 1-8.

Bloom, D.E. & Canning, D. (2007) Commentary: The Preston Curve 30 years on: still sparking fires. *International Journal of Epidemiology*, 36(3), 498-499.

Bombak, A.E. (2013) Self-Rated Health and Public Health: A Critical Perspective. *Frontiers in Public Health*, 1, 15.

Bosch, X., Moreno, P., & López-Soto, A. (2014). The painful effects of the financial crisis on Spanish health care. *International Journal of Health Services*, 44(1), 25-51.

Brydsten, A., Hammarström, A., & San Sebastian, M. (2016). The impact of economic recession on the association between youth unemployment and functional somatic symptoms in adulthood: a difference-in-difference analysis from Sweden. *BMC Public Health*, 16(1), 1-8.

Busch, K., Hermann, C., Hinrichs, K., & Schulten, T. (2013). Euro crisis, austerity policy and the European Social Model. *International Policy Analysis, Friedrich Ebert Foundation, Berlin.*

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., Schaan, B., Stuck, S. & Zuber, S. (2013). Data resource profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). *International journal of epidemiology*, 42(4), 992-1001.

Carballo-Cruz, F. (2011). Causes and consequences of the Spanish economic crisis: Why the recovery is taken so long? *Panoeconomicus*, 58(3), 309-328.

Chalari, A. (2014). The Subjective Experience of Three Generations During the Greek Economic Crisis. *World Journal of Social Science Research*, 1(1), 89-109.

Coleman, J.S. (1990). *Foundations of Social Theory*, Cambridge: Harvard University Press.

Conger, K.J., Rueter, M.A., & Conger, R.D. (2000). The Family Stress Model of Economic Hardship. In Crockett, L.J. & Silbereisen, R.K. (Ed.) *Negotiating Adolescence in Times of Social Change* (pp. 204-206). New York: Cambridge University Press.

Dalgaard, C.J., & Strulik, H. (2014). Optimal aging and death: understanding the Preston curve. *Journal of the European Economic Association*, 12(3), 672-701.

Davalos, M.E. & French, M.T. (2011) This recession is wearing me out! Health-related quality of life and economic downturns. *The journal of Mental Health Policy and Economics*, 14, 61-72.

Doblhammer, G., Van Den Berg, G.J. & Fritze, T. (2013). Economic Conditions at the Time of Birth and Cognitive Abilities Late in Life: Evidence from Ten European Countries. *PLOS* one, 8(9).

Dellepiane, S., & Hardiman, N. (2012). The new politics of austerity: Fiscal responses to the economic crisis in Ireland and Spain. *Geary Institute Discussion Paper Series.*

Del Valle, J.F., Cinali, C., Bravo, A. & Vecchiato, T. (2013). Child protection in Italy and Spain: Influence of the family supported society. *Psychosocial Intervention*, 22, 227-237.

Éltető, A. (2011). The economic crisis and its management in Spain. *Eastern Journal of European Studies*, 2(1), 41-55.

Evers, A.W.M., Verhoeven, E.W.M., Kraaimaat, F.W., De Jong, E.M.G.J., De Brouwer, S.J.M., Schalkwijk, J., Sweep, F.C.G.J. & Van De Kerkhof, P.C.M. (2010). How stress gets under the skin: cortisol and stress reactivity in psoriasis. *British Journal of Dermatology*, 163(5), 986-991.

Eurostat (2019). *Economic recovery of EU regions after 2008*. Retrieved online on 24 February 2021, from https://ec.europa.eu/eurostat/web/products-eurostat-news/-/DDN-20191023-1. Luxembourg: Eurostat.

Eurostat (2021). *At-risk-of-poverty rate for pensioners - EU-SILC and ECHP surveys.* Retrieved online on 23 February 2021, from https://ec.europa.eu/eurostat/databrowser/view/ilc_pns6/default/table?lang=en. Luxembourg: Eurostat.

Fountoulakis, K.N., Grammatikopoulos, I.A., Koupidis, S.A., Siamouli, M., & Theodorakis, P.N. (2012). Health and the financial crisis in Greece. *The Lancet*, 379(9820), 1001-1002.

Frangakis, M. (2015). Public debt crisis, austerity and deflation: the case of Greece. *Review of Keynesian Economics*, 3(3), 295-313.

Gaeta, M., Campanella, F., Capasso, L., Schifino, G.M., Gentile, L., Banfi, G., Pelissero, G. & Ricci, C. (2017). An overview of different health indicators used in the European Health Systems. *Journal of Preventive Medicine and Hygiene*, 58(2), 114-120.

Geller, A., Jaeger, K., & Pace, G. (2018). Using the Fragile Families and Child Wellbeing Study (FFCWS) in life course health development research. *Handbook of life course health development*, 601-620.

Gluckman, P. D., Hanson, M. A., & Pinal, C. (2005). The developmental origins of adult disease. *Maternal & child nutrition*, 1(3), 130-141.

Gordo, L., Hernández de Cos, P., & Pérez, J. (2013). Developments in Spanish public debt since the start of the crisis. *Economic Bulletin*, (JUL).

Granados, J.A.T., & Rodriguez, J.M. (2015). Health, economic crisis, and austerity: A comparison of Greece, Finland and Iceland. *Health Policy*, 119(7), 941-953.

Gudmundsdottir, D.G. (2013). The impact of economic crisis on happiness. *Social indicators research*, 110(3), 1083-1101.

Halfon, N. & Hochstein, M. (2003). Life Course Health Development: An Integrated Framework for Developing Health, Policy, and Research. *The Milbank Quarterly*, 80(3), 433-479.

Hammarström, A. (1994). Health consequences of youth unemployment—review from a gender perspective. *Social science & medicine*, *38*(5), 699-709.

Hazen, H., & Anthamatten, P. (2020). *An introduction to the geography of health*. Routledge.

Hyndman, R.J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.

Ibrahim, S., Smith, P., & Muntaner, C. (2009). A multi-group cross-lagged analyses of work stressors and health using Canadian National sample. *Social science & medicine*, *68*(1), 49-59.

Idler, E.L. & Benyamini, Y. (1997). Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies. *Journal of Health and Social Behavior*, 38(1), 21-37.

INE (2021a). *Encuesta de población activa. Últimos datos.* Retrieved online on 16 September 2021 from https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176918&menu=ultiDatos&idp=1254735976595. Madrid: Instututo Nacional de Estadística (INE).

INE (2021b). *Population figures. Latest data*. Retrieved online on 7 September 2021, from https://www.ine.es/dyngs/INEbase/en/operacion.htm?c=Estadistica_C&cid=1254736176951&menu=ultiDatos&idp=1254735572981. Madrid: Instituto Nacional de Estadística (INE).

Israel, G.D. (1992). Determining sample size. *Fact Sheet PEOD-6*, 1-5.

Janlert, U. (2009) Editorial: Economic crisis, unemployment and public health. *Scandinavian Journal of Public Health*, 37, 783-784.

Jylhä, M. (2009). What is self-rated health and why does it predict mortality? Towards a unified conceptual model. *Social science & medicine*, *69*(3), 307-316.

Karanikolos, M., Heino, P., McKee, M., Stuckler, D., & Legido-Quigley, H. (2016). Effects of the global financial crisis on health in high-income OECD countries: a narrative review. *International journal of health services*, 46(2), 208-240.

Karanikolos, M., Mladovsky, P., Cylus, J., Thomson, S., Basu, S., Stuckler, D. & McKee, M. (2013). Financial crisis, austerity, and health in Europe. *The Lancet*, 381(9874), 1323-1331.

Kinsella, S. (2012). Is Ireland really the role model for austerity? *Cambridge Journal of Economics*, 36(1), 223-235.

Knäuper, B., & Turner, P.A. (2003). Measuring health: improving the validity of health assessments. *Quality of Life Research, 12*(1), 81-89.

Krahn, G.L., Fujiura, G., Drum, C.E., Cardinal, B.J., Nosek, M.A., & RRTC Expert Panel on Health Measurement. (2009). The dilemma of measuring perceived health status in the context of disability. *Disability and Health Journal*, *2*(2), 49-56.

Krejcie, R.V., & Morgan, D.W. (1970). Determining sample size for research activities. *Educational and psychological measurement*, 30(3), 607-610.

Kuh, D., Ben-Schlomo, Y., Lynch, J., Hallqvist, J. & Power, C. (2003). Glossary: Life Course Epidemiology. *Journal of Epidemiological Community Health*, 57, 778-783.

Lameire, N., Joffe, P., & Wiedemann, M. (1999). Healthcare systems - an international review: an overview. *Nephrology Dialysis Transplantation*, 14(6), 3-9.

Layte, R., & McCrory, C. (2018). Fiscal crises and personal troubles: the great recession in Ireland and family processes. *Social psychiatry and psychiatric epidemiology*, *53*(9), 987-1001.

Legido-Quigley, H., Urdaneta, E., Gonzalez, A., La Parra, D., Muntaner, C., Alvarez-Dardet, C., Martin-Moreno, J.M. & McKee, M. (2013). Erosion of universal health coverage in Spain. *The Lancet*, 382(9909), 1977.

Levy, B.S., & Sidel, V. (2009). The economic crisis and public health. *Social medicine*, 4(2), 82-87.

Lorenzo-Carascosa, L. (2018). La salud de las personas mayores en España durante la crisis económica [The health of the elderly in Spain during the economic crisis]. *Rev Esp Geriatr Gerontol*, 53(1), 12-14.

Lundberg, O., & Manderbacka, K. (1996). Assessing reliability of a measure of self-rated health. *Scandinavian journal of social medicine*, *24*(3), 218-224.

Maitoza, R. (2019). Family challenges created by unemployment. *Journal of Family Social Work*, *22*(2), 187-205.

McKee, M., Karanikolos, M., Belcher, P., & Stuckler, D. (2012). Austerity: a failed experiment on the people of Europe. *Clinical medicine*, 12(4), 346-350.

Mehmetoglu, M. & Jakobsen, T.G. (2017). *Applied Statistics Using Stata. A Guide for the Social Sciences.* London*:* Sage.

Mucci, N., Giorgi, G., Roncaioli, M., Perez, J. F., & Arcangeli, G. (2016). The correlation between stress and economic crisis: a systematic review. *Neuropsychiatric disease and treatment*, 12, 983-993.

Nübling, M., Andersen, H.H. & Mühlbacher, A. (2006). *Entwicklung eines Verfahrens zur Berechnung der körperlichen und psychischen Summenskalen auf Basis der SOEP-Version des SF 12 (Algorithmus)* (No. 16). DIW Data Documentation.

OECD (2019). *Spain: Country Health Profile 2019, State of Health in the EU.* Paris: OECD Publishing.

OECD (2021). *Youth unemployment rate*. Retrieved online on 27 September 2021, from https://data.oecd.org/unemp/youth-unemployment-rate.htm. Paris: Organization for Economic Cooperation and Development.

Panico, C., & Purificato, F. (2013). Policy coordination, conflicting national interests and the European debt crisis. *Cambridge Journal of Economics*, 37(3), 585-608.

Preston, S. H. (1975). The changing relation between mortality and level of economic development. *Population studies*, 29(2), 231-248.

Raub, W., Buskens, V. & Van Assen, M.A.L.M. (2011). Micro-Macro Links and Microfoundations in Sociology. *The Journal of Mathematical Sociology*, 35(1-3), 1-25.

Raub, W. & Voss, T. (2017). Micro-macro models in sociology: antecedents of Coleman's diagram. In Jann, B. & Prsepiorka, W. (Ed.) *Social Dilemmas, Institutions, and the Evolution of cooperation* (pp. 11-36). Berlin: De Gruyter.

Regidor, E., Barrio, G., Bravo, M.J., & de la Fuente, L. (2014). Has health in Spain been declining since the economic crisis? *Journal of Epidemiological Community Health*, 68(3), 280-282.

Roca, M., Gili, M., Garcia-Campayo, J., & García-Toro, M. (2013). Economic crisis and mental health in Spain. *The Lancet*, 382(9909), 1977-1978.

SHARE (2020). *Data Documentation: Waves Overview: Wave 2.* Retrieved online on 26 May 2021 from http://www.share-project.org/data-documentation/waves-overview/wave-2.html.

SHARE (2021a). *Organization: About SHARE.* Retrieved online on 26 May 2021 from http://www.share-project.org/organisation.html.

SHARE (2021b). *Dates and Facts about SHARE.* Retrieved online on 26 May 2021 from http://www.share-project.org/organisation/dates-facts.html?L=0.

SHARE (2021c). *Organization: SHARE Country Teams: Spain*. Retrieved online on May 26 from http://www.share-project.org/organisation/share-country-teams/spain.html.

SHARE (2021d). *Data Documentation: Waves Overview: Wave 3 (SHARELIFE): SHARELIFE – The Retrospective Survey (Wave 3).* Retrieved online on 31 May 2021 from http://www.share-project.org/data-documentation/waves-overview/wave-3-sharelife.html.

SHARE (2021e). *Data Documentation: Waves Overview: Wave 7*. Retrieved online on 31 May 2021 from http://www.share-project.org/data-documentation/waves-overview/wave-7.html.

SHARE-CEMFI (n.d.). *Inicio: ¿Qué es SHARE?* Retrieved online on 21 May 2021 from http://www.share.cemfi.es/.

Stuckler, D., Reeves, A., Loopstra, R., Karanikolos, M., & McKee, M. (2017). Austerity and health: the impact in the UK and Europe. *European journal of public health*, 27(suppl4), 18-21.

UN DESA (n.d.). *The Social Impact of the Economic Crisis*. Retrieved online on 21 April 2021, from https://www.un.org/development/desa/dspd/world-social-report/2011-2/the-social-impact-of-the-economic-crisis.html#:~:text=During%20times%20of%20financial%20and,children%2C%20and%20thus%20perpetuate%20the. New York: UN DESA.

Uutela, A. (2010). Economic crisis and mental health. *Current opinion in psychiatry*, 23(2), 127-130.

Virtanen, P., Vahtera, J, Kivimäki, M., Pentti, J. & Ferrie, J. (2002). Employment security and health. *Journal of Epidemiological Community Health*, 56, 569-574.

Ware, J.E. (2000). SF-36 health survey update. *Spine*, *25*(24), 3130-3139.

Ware, J.E., Kosinski, M. & Keller, S.D. (1995). SF-12: *How to score the SF-12 Physical and Mental Health Summary Scales.* Second edition. Boston, MA: The Health Institute, New England Medical Center.

Whitehead, B.R. & Bergeman, C.S. (2017). The effect of the financial crisis on physical health: Perceived impact matters. *Journal of Health Psychology*, 22(7), 864-873.

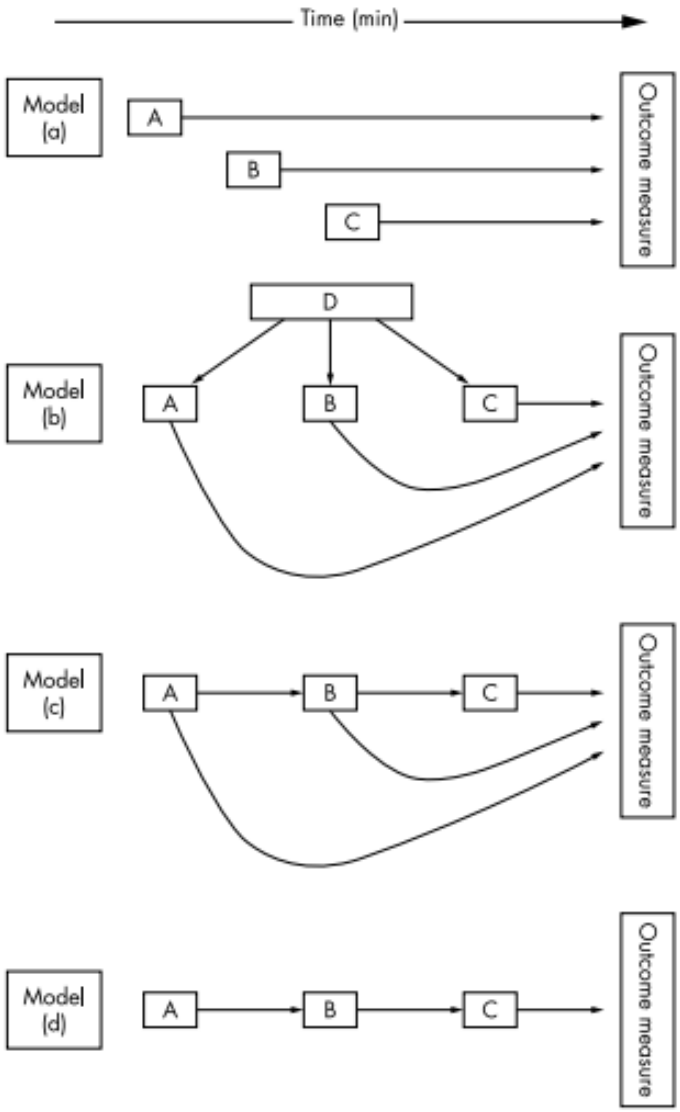
WHO (2021) *Constitution*. Retrieved online on 14 March 2021, from: https://www.who.int/about/who-we-are/constitution. Geneva: World Health Organization.

Zavras, D., Tsiantou, V., Pavi, E., Mylona, K., & Kyriopoulos, J. (2013). Impact of economic crisis and other demographic and socio-economic factors on self-rated health in Greece. *The European Journal of Public Health*, 23(2), 206-210.

Appendices

Appendix 1

Different models of the Life Course Health Development Framework by Kuh et al. (2003).



Appendix 2

SOEP 2004 variables to create SF-12 scores.

Source: https://paneldata.org/soep-core/data/up and Nübling et al. (2007)

SF-12 item and variable number	*Name in SOEP 2004*	*Description in SOEP*
Item 1: PF (Physical Functioning)		
Variable 1a	up84	State of health affects ascending stairs
Variable 1b	up85	State of health affects tiring tasks
Item 2: GH (General Health)		
Variable 2	up83	Current general health
Item 3: BP (Bodily Pain)		
Variable 3	up8605	Strong physical pain last 4 weeks
Item 4: RP (Role Physical)		
Variable 4a	up8606	Accomplished less due to physical problems
Variable 4b	up8607	Limitations due to physical problems
Item 5: MH (Mental Health)		
Variable 5a	up8602	Run-down, melancholy last 4 weeks
Variable 5b	up8603	Well-balanced last 4 weeks
Item 6: RE (Role Emotional)		
Variable 6a	up8608	Accomplished less due to emotional problems
Variable 6b	up8609	Less careful due to emotional problems
Item 7: SF (Social Functioning)		
Variable 7	up8610	Limited socially due to health last 4 weeks
Item 8: VT (Vitality)		
Variable 8	up8604	Used energy last 4 weeks

Appendix 3

The full Stata codes to create SF-12 scores in the SHARE dataset (example here is Wave 7, Wave 4 is same method, different numbers). This syntax is based on the syntax by Nübling et al. (2006).

```
*** dofile to create sf12 vars ***

use "X:\MASTER_THESIS\CreatingSF12\PH_es7.dta"
cd "X:\MASTER_THESIS\CreatingSF12\"

merge 1:1 mergeid using "AC_es7"
rename _merge merge1
merge 1:1 mergeid using "HS_es7"
rename _merge merge2
merge 1:1 mergeid using "MH_es7"
rename _merge merge3

drop if ph048d4 ==-2
drop if ph048d4 ==-1
rename ph048d4 var1a

drop if ph049d12 ==-1
rename ph049d12 var1b

drop if ph003_ ==-1
rename ph003_ var2

rename ph084_ var3

rename hs052_ var4a

drop if var4a ==-1
rename ph005_ var4b

drop if mh002_ ==-2
drop if mh002_ ==-1
rename mh002_ var5a

drop if ac704_ ==-1
drop if ac704_ ==-2
rename ac704_ var5b

drop if ac037_ ==-1
rename ac037_ var6a

drop if ac017_ ==-1
drop if ac017_ ==-2
rename ac017_ var6b

gen soclif=.
replace soclif = 1 if hs063d4==1
replace soclif = 0 if hs063d4==0
replace soclif = 0 if  hs054_==0
label define sociallife 0 "not selected" 1 "selected"
label values soclif sociallife
rename soclif var7
label variable var7 "Consequences of illness period: made social life more
difficult"

drop if ac023_ ==-1
rename ac023_ var8
```

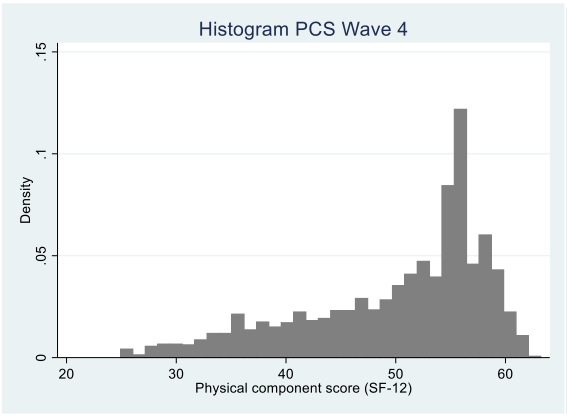
```
sum var1a var1b var2 var3 var4a var4b var5a var5b var6a var6b var7 var8

drop if ph006d18 ==-1
rename var6a ac037_
rename ph006d18 var6a

sum var1a var1b var2 var3 var4a var4b var5a var5b var6a var6b var7 var8

* issues with obersations for var3, var5a and also some issues with var4a and var7

drop if var1a ==.
drop if var1b ==.
drop if var2 ==.
drop if var4a ==.
drop if var4a ==.
drop if var5b ==.
drop if var6a ==.
drop if var6b ==.
drop if var7 ==.
drop if var8 ==.

* I did not delete missing obervations for var3 and var5a, for the rest I did. This
leaves us with 3099 individuals that have no missing values for 10 of the 12
variables we need. All individuals have missing values for variables 3 and 5a,
therefore we need to find replacing variables for these two.

rename var3  ph084_
rename var5a mh002_

gen pain=.
replace pain = 2 if ph011d7==1
replace pain = 2 if ph011d8==1
drop if ph011d8==-1
drop if ph011d7==-1
replace pain = 1 if ph011d7==0
replace pain = 1 if ph011d8==0
replace pain = 2 if ph011d7==1
replace pain = 2 if ph011d8==1
label define pain 1 "no" 2 "yes"
label values pain pain
label variable pain "Person uses drugs against pain in body"
rename pain var3
drop if var3==.

rename ph011d10 var5a

sum var1a var1b var2 var3 var4a var4b var5a var5b var6a var6b var7 var8

* now we have variables for all items! in total we have 3092 observations

* let's invert scales if necessary:

recode var1a (0=2)
recode var1b (0=2)
recode var2 (1=5) (2=4) (3=3) (4=2) (5=1)
label define cats 1"poor" 5"excellent"
label values var2 cats
recode var3 (1=2) (2=1)
label define druguse 1"selected" 2"not selected"
label values var3 druguse
recode var4a (1=1) (5=2)
label values var4a druguse
```

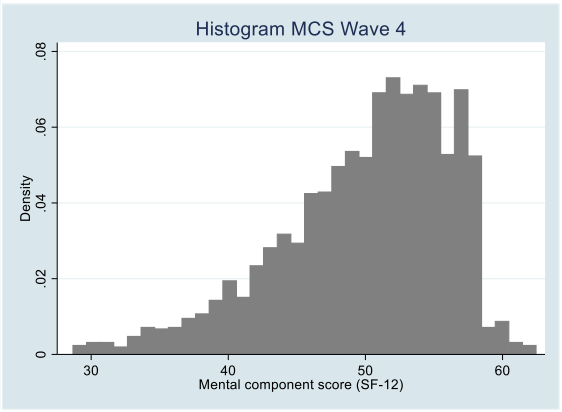
```
recode var5a (0=2) (1=1)
label values var5a druguse
recode var6a (0=2)
label values var6a druguse
recode var6b (1=4) (2=3) (3=2) (4=1)
label define oftennever 1"never" 2"rarely" 3"sometimes" 4"often"
label values var6b oftennever
recode var7 (0=2)
label values var7 druguse
recode var8 (1=4) (2=3) (3=2) (4=1)
label values var8 oftennever
label define poorhealth 1"poor" 2"fair" 3"good" 4"very good" 5"excellent"
label values var2 poorhealth

sum var1a var1b var2 var3 var4a var4b var5a var5b var6a var6b var7 var8

label values var1a druguse
label values var1b druguse

*now that all 12 variables are created and also have similar scales (meaning a low
score is bad and a high score is good) we can create the 8 items:

gen item1 = (var1a + var1b)/2
gen item2 = var2
gen item3 = var3
recode var4a (2=3)
gen item4 = (var4a +var4b)/2
recode var5a (2=5)
gen item5 = (var5a + var5b)/2
recode var6a (2=4)
gen item6 = (var6a + var6b)/2
gen item7 = var7
gen item8 = var8

sum item1 item2 item3 item4 item5 item6 item7 item8

*all the items are ready, they have different scales, all start at 1, some go up to
2 and are thus binary (item 1, 3 and 7), others go up to either 3 (item 4), 4 (item
6 and 8)  or 5 (item 2 and 5)

rename item1 PF
rename item2 GH
rename item3 BP
rename item4 RP
rename item5 MH
rename item6 RE
rename item7 SF
rename item8 VT

label variable PF "Physical Functioning"
label variable GH "General Health"
label variable BP "Bodily Pain"
label variable RP "Role Physical"
label variable MH "Mental Health"
label variable RE "Role Emotional"
label variable SF "Social Functioning"
label variable VT "Vitality"

sum PF GH BP RP MH RE SF VT

*all 8 items have been given the right names and descriptions

*now start with the rest of the calculations
```

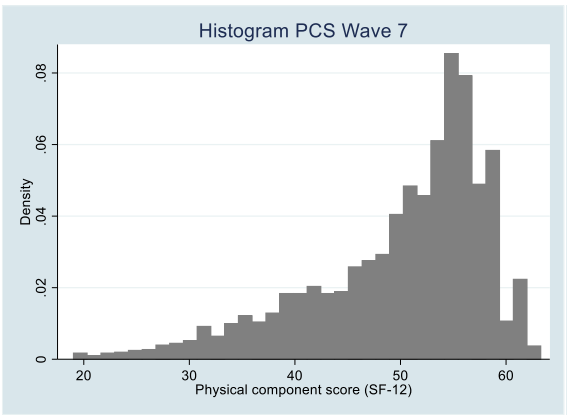
```
gen SF_PF = ((PF-1)/1)*100
egen SF_z_PF = std(SF_PF)
gen SF_nbs_PF = (SF_z_PF*10)+50

gen SF_GH = ((GH-1)/4)*100
egen SF_z_GH = std(SF_GH)
gen SF_nbs_GH = (SF_z_GH*10)+50

gen SF_BP = ((BP-1)/1)*100
egen SF_z_BP = std(SF_BP)
gen SF_nbs_BP = (SF_z_BP*10)+50

gen SF_RP = ((RP-1)/2)*100
egen SF_z_RP = std(SF_RP)
gen SF_nbs_RP = (SF_z_RP*10)+50

gen SF_MH = ((MH-1)/4)*100
egen SF_z_MH = std(SF_MH)
gen SF_nbs_MH = (SF_z_MH*10)+50

gen SF_RE = ((RE-1)/3)*100
egen SF_z_RE = std(SF_RE)
gen SF_nbs_RE = (SF_z_RE*10)+50

gen SF_SF = ((SF-1)/1)*100
egen SF_z_SF = std(SF_SF)
gen SF_nbs_SF = (SF_z_SF*10)+50

gen SF_VT = ((VT-1)/3)*100
egen SF_z_VT = std(SF_VT)
gen SF_nbs_VT = (SF_z_VT*10)+50

factor SF_z_PF SF_z_GH SF_z_BP SF_z_RP SF_z_MH SF_z_RE SF_z_SF SF_z_VT, factor(2)

rotate, varimax
estimates store factor12
loadingplot
predict factor12

gen PCS_agg = (SF_z_GH*0.27755) + (SF_z_PF*0.26254) + (SF_z_BP*0.15271) +
(SF_z_RP*0.28601) + (SF_z_MH*-0.06808) + (SF_z_RE*-0.03659) + (SF_z_SF*0.06421) +
(SF_z_VT*0.14489)

gen MCS_agg = (SF_z_GH*-0.00036) + (SF_z_PF*-0.04872) + (SF_z_BP*-0.04147) +
(SF_z_RP*-0.04180) + (SF_z_MH*0.33462) + (SF_z_RE*0.40358) + (SF_z_SF*0.05051) +
(SF_z_VT*0.14982)

gen PCS = (PCS_agg*10)+50
label variable PCS "Physical component score (SF-12)"

gen MCS = (MCS_agg*10)+50
label variable MCS "Mental component score (SF-12)"
```

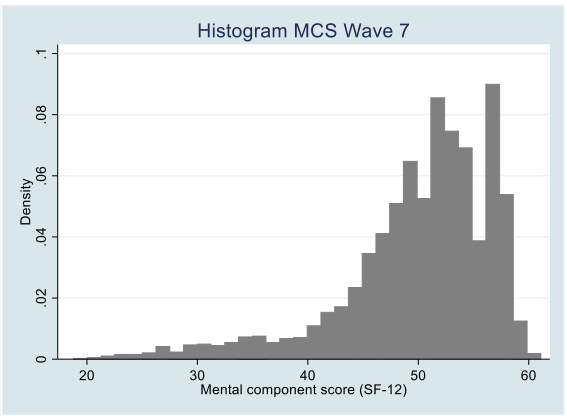
Appendix 4

Histograms for PCS and MCS in Wave 4 and Wave 7.









Appendix 5

Short-term: PCS linear regression models.

Physical Component Score (PCS) linear regression models (short-term)

	MODEL 1	MODEL 2	MODEL 3	MODEL 4	MODEL 5
Age	-0.224***	-0.219***	-0.205***	-0.101*	-0.103*
	(-10.29)	(-5.50)	(-4.84)	(-2.11)	(-2.15)
Single					
Yes	-0.194	0.573	0.669	0.931	0.995
	(-0.43)	(0.82)	(0.95)	(1.32)	(1.40)
Number of children	-0.208+	-0.335+	-0.349*	-0.340+	-0.296
	(-1.85)	(-1.96)	(-2.00)	(-1.95)	(-1.64)
Gender					
Female	-2.603***	-3.352***	-3.296***	-4.123***	-4.083***
	(-8.29)	(-6.59)	(-6.47)	(-7.43)	(-7.34)
Years of education	0.316***	0.331***	0.325***	0.328***	0.327***
	(9.27)	(5.80)	(5.55)	(5.65)	(5.62)
Total household income in 2011	0.000*	0.000	0.000	0.000	0.000
	(2.46)	(1.47)	(0.74)	(0.81)	(0.79)
Income change between 2007 and 2011		-0.000	-0.000	-0.000	-0.000
		(-1.11)	(-0.37)	(-0.56)	(-0.55)
Income loss between 2007 and 2011					
Yes		-0.570	-0.598	-0.723	-0.745
		(-0.97)	(-0.35)	(-0.42)	(-0.43)
Income quantile in 2007 (reference: third quintile)					
Lowest incomes (first quintile)			-4.067+	-4.394*	-4.368*
			(-1.88)	(-2.04)	(-2.03)
Lower incomes (second quintile)			-0.184	-0.228	-0.129
			(-0.06)	(-0.07)	(-0.04)
Higher incomes (fourth quintile)			-4.601+	-4.318+	-4.267+
			(-1.79)	(-1.70)	(-1.68)
Highest incomes (fifth quintile)			-7.023*	-6.948*	-6.793*
			(-2.41)	(-2.40)	(-2.35)
Income change between 2007 and 2011, in categories (reference: very strong income growth)					
Very strong income loss			3.671	3.382	3.297
			(1.12)	(1.04)	(1.01)
Strong income loss			-1.549	-1.829	-1.821
			(-0.48)	(-0.57)	(-0.57)
Some income loss			-8.246**	-8.398**	-8.277**
			(-2.84)	(-2.92)	(-2.87)
Stable income			-5.720*	-5.734*	-5.740*
			(-2.09)	(-2.11)	(-2.11)
Some income growth			-4.875*	-4.721*	-4.690*
			(-2.32)	(-2.26)	(-2.25)

Strong income growth	-2.444	-2.483	-2.482
	(-1.08)	(-1.11)	(-1.11)
Interaction between income quantile in 2007 and Income change between 2007 and 2011			
First quantile # Some income loss	6.615+	6.683+	6.674+
	(1.82)	(1.85)	(1.85)
First quantile # Stable income	5.255	5.551+	5.590+
	(1.57)	(1.67)	(1.68)
First quantile # Some income growth	4.168	4.004	4.069
	(1.52)	(1.47)	(1.50)
First quantile # Strong income growth	0.633	0.672	0.699
	(0.24)	(0.26)	(0.27)
Second quintile # Strong income loss	-4.937	-4.224	-4.538
	(-0.92)	(-0.79)	(-0.85)
Second quintile # Some income loss	6.556+	6.201	6.006
	(1.72)	(1.64)	(1.59)
Second quintile # Stable income	1.400	1.100	0.907
	(0.35)	(0.28)	(0.23)
Second quintile # Some income growth	-0.163	-0.524	-0.674
	(-0.05)	(-0.15)	(-0.20)
Second quintile # Strong income growth	-2.674	-2.789	-2.853
	(-0.74)	(-0.78)	(-0.80)
Fourth quintile # Severe income loss	2.088	3.556	3.600
	(0.34)	(0.58)	(0.58)
Fourth quintile # Strong income loss	1.899	1.962	2.019
	(0.57)	(0.59)	(0.61)
Fourth quintile # Some income loss	5.835+	5.810+	5.688+
	(1.89)	(1.90)	(1.86)
Fourth quintile # Stable income	8.607*	8.139*	8.085*
	(2.55)	(2.43)	(2.41)
Fourth quintile # Some income growth	3.579	3.301	3.268
	(1.21)	(1.12)	(1.11)
Fourth quintile # Strong income growth	0.821	0.899	0.817
	(0.25)	(0.28)	(0.25)
Fifth quintile # Strong income loss	4.356	4.605	4.477
	(1.24)	(1.32)	(1.28)
Fifth quintile # Some income loss	10.846**	10.810**	10.583**
	(3.09)	(3.10)	(3.03)
Fifth quintile # Stable income	7.137	7.019	6.830
	(1.64)	(1.63)	(1.58)
Fifth quintile # Some income growth	8.197*	7.960*	7.811*
	(2.16)	(2.11)	(2.06)
Fifth quintile # Strong income growth	3.759	3.917	3.770
	(1.08)	(1.13)	(1.09)
Unemployed in 2007			
Yes		3.269*	3.224*
		(2.13)	(2.10)

Unemployed in 2011					
Yes				5.291	5.903
				(0.93)	(1.03)
Employed in 2007					
Yes				2.807***	2.767***
				(4.16)	(4.10)
Interaction between employed in 2007 and unemployed in 2011					
Yes # Yes				-3.596	-4.147
				(-0.62)	(-0.71)
Interaction between unemployed in 2007 and unemployed in 2011					
Yes # Yes				-4.599	-5.259
				(-0.76)	(-0.86)
Occupational status of child					
Unemployed					-0.769
					(-1.00)
Constant	62.988***	63.037***	66.978***	58.665***	58.761***
	(41.89)	(22.51)	(19.08)	(14.97)	(14.99)
N	2535	1004	1004	1004	1004
r2	0.142	0.146	0.184	0.202	0.203

T statistics in parentheses

*+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001*

Appendix 6

Short-term: MCS linear regression models.

Mental Component Score (MCS) linear regression models (short-term)

	MODEL 6	MODEL 7	MODEL 8	MODEL 9	MODEL 10
Age	-0.32+	-0.080**	-0.086**	-0.050	-0.051
	(-1.92)	(-2.61)	(-2.64)	(-1.35)	(-1.36)
Single					
Yes	-1.875***	-1.930***	-1.703**	-1.593**	-1.576**
	(-5.40)	(-3.62)	(-3.12)	(-2.90)	(-2.86)
Number of children	-0.023	0.055	0.096	0.125	0.136
	(-0.27)	(0.42)	(0.71)	(0.92)	(0.97)
Gender					
Female	-1.886***	-2.675***	-2.688***	-3.213***	-3.203***
	(-7.81)	(-6.86)	(-6.83)	(-7.46)	(-7.41)
Years of education	0.171***	0.128**	0.137**	0.140**	0.139**
	(6.53)	(2.92)	(3.03)	(3.10)	(3.09)
Total household income in 2011	0.000	0.000	0.000	0.000	0.000
	(1.08)	(0.27)	(0.31)	(0.24)	(0.24)
Income change between 2007 and 2011		-0.000	-0.000	-0.000	-0.000
		(-0.99)	(0.49)	(0.45)	(0.45)
Income loss between 2007 and 2011					
Yes		-0.502	-1.099	-0.947	-0.953
		(-1.12)	(-0.83)	(-0.71)	(-0.71)
Income quantile in 2007 (reference: third quintile)					
Lowest incomes (first quintile)			-2.926+	-3.045+	-3.039+
			(-1.75)	(-1.82)	(-1.82)
Lower incomes (second quintile)			4.089+	-4.125+	-4.100+
			(-1.69)	(-1.71)	(-1.70)
Higher incomes (fourth quintile)			-1.941	-1.798	-1.785
			(-0.98)	(-0.91)	(-0.90)
Highest incomes (fifth quintile)			5.509*	-5.395*	-5.356
			(-2.45)	(-2.40)	(-2.38)
Income change between 2007 and 2011, in categories (reference: very strong income growth)					
Very strong income loss			5.960*	5.727*	5.706*
			(2.36)	(2.27)	(2.25)
Strong income loss			-0.557	-0.796	-0.794
			(-0.22)	(-0.32)	(-0.32)
Some income loss			-1.160	-1.135	-1.104
			(-0.52)	(-0.51)	(-0.49)
Stable income			-1.694	-1.679	-1.680
			(-0.80)	(-0.79)	(-0.79)
Some income growth			-0.164	0.097	0.105

	(-0.10)	(0.06)	(0.06)
Strong income growth	-2.773	-2.613	-2.613
	(-1.59)	(-1.50)	(-1.50)
Interaction between income quantile in 2007 and Income change between 2007 and 2011			
First quantile # Some income loss	2.443	2.356	2.354
	(0.87)	(0.84)	(0.84)
First quantile # Stable income	1.608	1.596	1.606
	(0.62)	(0.62)	(0.62)
First quantile # Some income growth	-0.119	-0.114	-0.098
	(-0.06)	(-0.05)	(-0.05)
First quantile # Strong income growth	3.807+	3.653+	3.660+
	(1.88)	(1.81)	(1.81)
Second quintile # Strong income loss	-2.627	-2.400	-2.480
	(-0.63)	(-0.58)	(-0.60)
Second quintile # Some income loss	4.148	3.908	3.859
	(1.41)	(1.33)	(1.31)
Second quintile # Stable income	4.115	4.092	4.043
	(1.35)	(1.34)	(1.32)
Second quintile # Some income growth	2.083	1.816	1.778
	(0.79)	(0.69)	(0.67)
Second quintile # Strong income growth	4.721+	4.726+	4.710+
	(1.69)	(1.70)	(1.69)
Fourth quintile # Severe income loss	-7.000	-6.277	-6.266
	(-1.46)	(-1.31)	(-1.30)
Fourth quintile # Strong income loss	0.474	0.495	0.509
	(0.18)	(0.19)	(0.20)
Fourth quintile # Some income loss	1.428	1.298	1.267
	(0.60)	(0.55)	(0.53)
Fourth quintile # Stable income	2.883	2.547	2.533
	(1.10)	(0.98)	(0.97)
Fourth quintile # Some income growth	0.427	-0.602	-0.610
	(-0.19)	(-0.26)	(-0.27)
Fourth quintile # Strong income growth	3.356	3.153	3.132
	(1.35)	(1.27)	(1.26)
Fifth quintile # Strong income loss	5.952*	5.998*	5.965*
	(2.19)	(2.21)	(2.20)
Fifth quintile # Some income loss	6.243*	6.086*	6.028*
	(2.31)	(2.25)	(2.22)
Fifth quintile # Stable income	4.346	4.036	3.988
	(1.29)	(1.20)	(1.19)
Fifth quintile # Some income growth	4.506	4.038	4.000
	(1.53)	(1.38)	(1.36)
Fifth quintile # Strong income growth	4.921+	4.745+	4.707+
	(1.83)	(1.77)	(1.75)
Unemployed in 2007			
Yes		-0.713	-0.724

				(-0.60)	(-0.61)
Unemployed in 2011					
Yes				0.149	0.304
				(0.03)	(0.07)
Employed in 2007					
Yes				1.462**	1.452**
				(2.79)	(2.77)
Interaction between employed in 2007 and unemployed in 2011					
Yes # Yes				-0.566	-0.705
				(-0.13)	(-0.16)
Interaction between unemployed in 2007 and unemployed in 2011					
Yes # Yes				1.884	1.717
				(0.40)	(0.36)
Occupational status of child					
Unemployed					-0.195
					(-0.33)
Constant	51.824***	55.798***	57.799***	54.682***	54.707***
	(44.81)	(25.99)	(21.32)	(17.96)	(17.95)
N	2535	1004	1004	1004	1004
r2	0.066	0.085	0.113	0.123	0.123

T statistics in parentheses

*+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001*

Appendix 7

Factor analyses obtained coefficients used for the calculation of MCS and PCS. Factor 2 is used to calculate the MCS score and factor 1 for the PCS score.

Factor analysis scoring coefficients (method = regression; based on varimax rotated factors) Wave 4:

Variable	Factor 1	Factor 2
SF_z_PF	**0.25554**	-0.08407
SF_z_GH	**0.25039**	0.06593
SF_z_BP	**0.18707**	-0.12796
SF_z_RP	**0.30332**	-0.05507
SF_z_MH	0.01161	**0.25374**
SF_z_RE	0.00027	**0.29471**
SF_z_SF	-0.04041	**0.10455**
SF_z_VT	0.09672	**0.26953**

Factor analysis scoring coefficients (method = regression; based on varimax rotated factors) Wave 7:

Variable	Factor 1	Factor 2
SF_z_PF	**0.26254**	-0.04872
SF_z_GH	**0.27755**	-0.00036
SF_z_BP	**0.15271**	-0.04147
SF_z_RP	**0.28601**	-0.04180
SF_z_MH	-0.06808	**0.33462**
SF_z_RE	-0.03659	**0.40358**
SF_z_SF	0.06421	**0.05051**
SF_z_VT	0.14489	**0.14982**

Formulas to calculate PCS and MCS:

$$PCS_{agg} = (SF\ z\ GH\ * factor1\ GH\ ) + (SF\ z\ PF * factor1\ PF) + (SF\ z\ BP * facor1\ BP) + (SF\ z\ RP * factor1\ RP) + (SF\ z\ MH * factor1\ MH) + (SF\ z\ RE * factor1\ RE) + (SF\ z\ SF * factor1\ SF) + (SF\ z\ VT * factor1\ VT)$$

$$PCS = (PCS_{agg} * 10) + 50$$

$$MCS_{agg} = (SF\ z\ GH\ * factor2\ GH\ ) + (SF\ z\ PF * factor2\ PF) + (SF\ z\ BP * facor2\ BP) + (SF\ z\ RP * factor2\ RP) + (SF\ z\ MH * factor2\ MH) + (SF\ z\ RE * factor2\ RE) + (SF\ z\ SF * factor2\ SF) + (SF\ z\ VT * factor2\ VT)$$

$$MCS = (MCS_{agg} * 10) + 50$$

Appendix 8

Summary statistics of the most important variables in Wave 4 and Wave 7 and a description of the variable.

Variable	Wave 4 (2011)			Wave 7 (2017)			Existing or newly created variable	Information about the variable
	N	Mean	Std. dev.	N	Mean	Std. dev.		
PCS	2,535	49.9932	8.370839	3,092	50	8.331639	Created	*SF-12 Physical health Component Score.*
MCS	2,535	49.99123	6.169537	3,092	50	6.859366	Created	*SF-12 Mental health Component Score.*
age	3,624	67.72682	11.05996	4,711	71.3719	10.50712	Existing	*Age of respondent in years.*
single	3,624	0.2135762	0.4098875	4,711	0.2742518	0.4461838	Existing	*Binary variable: Is the respondent single? Yes = 1, No = 0.*
nchild	3,624	2.455298	1.539323	4,711	-70.80492*	45.4905	Existing	*Number of children of respondent.*
gender	3,624	1.543046	0.4982123	4,711	1.559541	0.4964948	Existing	*Binary variable: Male = 1, Female = 2.*
yedu	3,624	7.591611	5.016825	4,711	8.861176	5.190658	Existing	*Total number of years of education*
thinc	3,624	20626.86	25067.59	4,711	4900.087	14512.13	Existing	*Total yearly household income in euros of the respondent in 2011 (Wave 4) or 2017 (Wave 7).*
incomechange	1,556	2568.447	27595.02	421	-238062.8	4873641	Created	*Continuous variables of which the value represents the change of the respondent's yearly income between 2007 and 2011, can be negative and positive.*
incomeloss	1,556	0.3766067	.4846907	421	0.3444181	.4757435	Created	*Binary variable: has the respondent experienced income loss between 2007 and 2011? Yes = 1, No = 0.*

incomecatw2	2,338	2.995295	1.417077	2,338	2.995295	1.417077	Created	*Categorical variable that divides income in 2007 in 5 categories (Lowest, Lower, Median. Higher, Highest).*
incomechangecat	1,556	4.384319	1.709098	1,556	4.384319	1.709098	Created	*Categorical variable: How did the income of the respondent change between 2007 and 2011? 7 categories from "severe income loss" (=1) to "strong income growth" (=7).*
unemployedw2	2,338	0.0299401	.1704586	2,338	0.0299401	.1704586	Created	*Binary variable: was the respondent unemployed (still in the workforce but not employed) in 2007? Yes = 1, No = 0.*
unemployedw4	3,624	0.0587748	.2352352	3,624	0.0587748	.2352352	Created	*Binary variable: was the respondent unemployed (still in the workforce but not employed) in 2011? Yes = 1, No = 0.*
employedw2	2,338	0.5201027	.4997026	2,338	0.5201027	.4997026	Created	*Binary variable: was the respondent employed in 2007? Yes = 1, No = 0.*
occchild	4,406	1.10985	.3127384	4,406	1.10985	.3127384	Created	*Binary variable: Is at least one child of the respondent unemployed? Yes = 1, No, unknown, missing, in education or employed = 0.*

**: in Wave 7, the mean is negative, which obviously can't be true, This can be explained by the fact that "not applicable" has been labeled as -99 In the dataset. If this category would not be available, almost one in four observations were deleted. This issue does not create an issue for the research because this variable from Wave 7 is not used in any regression model or elsewhere in the research.*