Nijmegen School of Management
Department of Economics and Business Economics
Master's Thesis Economics (MAN-MTHEC)

Socioeconomic Determinants of Health in Europe

The Relation Between Health & Income, Education, Immigrant Status, and Unemployment in European Countries

By Carlijn Schumacher (s1065844)

Nijmegen, 21 July 2022

Program: Master's Program in Economics
Specialisation: Economic Behaviour & Policy EBP
Supervisor: Agapi Thaleia Fytraki

Radboud Universiteit

Preface

Before you lies the master's thesis "Socioeconomic Determinants of Health in Europe". This is my final project of the Radboud University master Economics, Behaviour and Policy. I was engaged in researching and writing this thesis from January to June 2022.

I came up with this specific research by combining my interests with the expertise of my supervisor. Health inequalities have interested me for a long time, and now, with the help of my supervisor, I was able to further research which socioeconomic determinants play a part in determining health. This was not without challenges, of course. The database chosen for this research proofed not to fit to the research question. Thus, it required some flexibility to change data sources and perform a different data analysis. Although it was a challenge, it also taught me a lot on how to be more thorough in exploring the available data, and how to not get stuck in the beaten track.

I would like to thank a number of people for their contributions to this thesis. First of all, my thesis advisor dr. Agapi Thaleia Fytraki of the Business Economics Department. Whenever there was an issue, I could ask for help and guidance. I would also like to thank the second reader of this thesis for their time and feedback.

Finally, I must express my gratitude to my family and friends for providing me with their unwavering support and continuous encouragement throughout the process of researching and writing this thesis.

This thesis would not have been possible without them all. Thank you.

Carlijn Schumacher Nijmegen, 2022

Table of Contents

Abstract

This master's thesis explores how subjective health is impacted by different socioeconomic variables: income, education, immigrant status, and unemployment. This relationship is analysed in 38 European countries, and health differences between those countries are also further explored. The main novelty is that the impact of these socioeconomic variables is now studied as a group instead of individually. Data from the European Social Survey from 2018 is analysed using a complex sample multinomial logit model. This reveals that there is a positive effect of income and education on subjective health, whereas immigrant status does not have a significant effect. Unemployment has a negative effect on health compared to those who follow education, perform housework, or perform paid work. Lastly, country of origin also significantly affects subjective health. Differences in this can also be identified between European regions. Future research is encouraged to dive deeper into the causes of these inequalities, in order to address them adequately in policymaking.

Chapter 1 Introduction

1.1 Introduction of the Topic

What once was a widely held consensus, namely that sickness hits the rich and poor alike, has now been refuted repeatedly. Health inequalities do exist, and they are deadly. They have existed for decades but became even more apparent during the COVID-19 pandemic (Bambra et al., 2020). The effects of this crisis were overwhelmingly felt by individuals from Black and minority ethnic groups, with lower incomes, or in urban and rurally deprived locations (Mishra et al., 2021). In the entirety of Europe, people with a lower level of education, occupational class, or income have more health problems and are more likely to die at a younger age (Mackenbach, 2006). This highlights the importance of socioeconomic characteristics in health. Health inequalities between people based on high and low education and income exist in all European countries (Mackenbach, 2006). These gaps have even worsened over the last decades (Bambra et al., 2020).

They also cost money. Each year, health inequalities cost an estimated €980 billion, which is 9.4% of European GDP (Forster et al., 2018). This is the result of lower productivity and higher healthcare and welfare costs. There are differences between countries, however. If we look at infant mortality for example, the differences between socioeconomic groups are much more pressing in Hungary and Croatia than in England, Wales, or Belgium (Mackenbach, 2006). Apart from costing the public money, these inequalities are unjust and avoidable. They affect the whole of society and are present in all life stages from infancy to old age (Acheson, 2001). Therefore, it is important to be aware of the relation between socioeconomic characteristics and health. This knowledge is necessary to construct a policy that addresses these issues. If the major determinants of health are social, so must be the remedies.

1.2 Research Problem & Motivation

The WHO has defined health inequalities as *"*avoidable inequalities in health between groups of people within countries and between countries*"* (World Health Organization, 2013*, p.2*). According to Mackenbach (2006), all European countries at the start of the 21st century are faced with substantial health inequalities. Inequalities in mortality have even increased in many European countries in the past decade. By taking action on social determinants, governments are not just improving health but also the broader circumstances in which people live and work (Marmot, 2005). Health inequalities were studied before, also in combination with socioeconomic indicators. The indicators of education, occupation and income are overwhelmingly dominant in the literature (*e.g.,* Dahl, 1994; Forster et al., 2018; Ross & Wu, 1995; Stronks et al., 1997). There is some critique on occupation as a determinant of health because the word in itself has many potential meanings (Illsley & Baker, 1991). The prestige, salary and working conditions of certain occupations differ between countries as well, making it difficult for cross country comparisons (Ravesteijn et al., 2013).

Other determinants that were discussed are land ownership and wealth, but this is a rare occurrence (Forster et al., 2018). Higher education can lead to better health because those people are exposed to fewer risk factors and have increased health literacy (Forster et al., 2018). A steady and satisfying job can support health by providing financial stability, social status, or even social networks to protect from psychosocial hazards (Forster et al., 2018). When a job is unsteady or unsatisfying, it has the opposite effect. Most research that looks at health as a whole, considers education and income as indicators (Jürges, 2010).

Another weakness in this field is the strong focus on Western European countries (Jutz, 2020). This is especially worrisome since life expectancy continues to be better in countries in Western than in Eastern Europe (Forster et al., 2018). The effect of health inequalities also differs depending on the country (Mackenbach, 2006). In Central and Eastern Europe, education based health inequality appear to be much higher than in Western Europe (Jutz, 2020). This difference is not visible for income based health inequalities. Even within these regions, there are big differences in size and determinants of health inequalities between countries (Jutz, 2020).

Health inequality is a big issue, both within and among countries. Among countries, there are differences of up to 48 years in life expectancy (Marmot, 2005). Within Denmark, men with a low education level experience an average of 0.7 years shorter lifetime due to cardiovascular disease alone, than men with a high education level (Brønnum-Hansen & Baadsgraad, 2007). In Spain, at age 65 the life expectancy difference between people with a high and low education was 2.1 years for men and 1.9 years for women (Majer et al., 2011).

It is well known that poverty in the form of material deprivation (dirty water, poor nutrition) can lead to bad health outcomes (Marmot, 2005). However, it is socially determined whoever gets access to those resources. Therefore, it is important to look at the cause of the causes: what causes poverty. Material deprivation or bad hygiene alone does not explain the size of the health inequalities (Marmot, 2005). That is where the socioeconomic determinants of health come in. The policies that are currently in place are not sufficient to eradicate this inequality. This is partly because the socioeconomic definition in academic research is limited because they often only include income and education. Thus, the research problem that is tackled in this master's thesis is to broaden the definition of socioeconomic determinants, and to identify their effect on health.

1.3 Research Objectives

This research aims to broaden the term of socioeconomic determinants as it is now often used in determining health. Only one or two indicators are often considered per research, mainly income and education. This research involves more determinants, in order to get a broader view on their effects. It also compares health between different European countries, and explores whether trends can be identified in European regions.

The focus is on European countries because a lot of previous research was done on this region, so the results can be compared. Other factors in this decision are the data availability, and the necessity to limit the scope of the research. A data analysis is performed to determine the effect of the individual socioeconomic determinants on health. The effect of all those indicators together can also be identified. In this case, subjective health is compared because it provides a more holistic view than objective health indicators. It is also more widely available for different European countries and was used in previous research which allows for comparisons. Subjective health entails self-perceived health and well-being, and functional and activity limitations (ESS, 2018).

That relationship between health and socioeconomic determinants can make policies more effective. By knowing what socioeconomic determinants are related to better or worse health, public health interventions can become more targeted (Forster, 2018). If it is known that certain groups are at risk, research can be done on how to gain access to that specific group, for example by public health campaigns. If income turns out to be a big determinant, universal health coverage or other social benefits may offer a solution. If education is more of an issue, health literacy and education may be given more attention. Although prevention has become a bigger part of health policy, there is too little attention from policy makers on the actual determinants of health risk (Adler & Newman, 2002). Those determinants are socioeconomic. The aim should not be to diminish socioeconomic disparities, but to diminish health inequality caused by socioeconomic factors. Not everyone should have to follow the same level of education, but everyone should have the same health opportunities regardless of education. Thus, knowing more about the relation between socioeconomic determinants and subjective health can improve informed policy making.

1.4 Research questions

This leads us to the following central research question that is answered in this thesis:

What is the Relation between Socioeconomic Determinants and Subjective Health In European Countries?

The following four sub questions are used to answer the central research question:

SQ1) What is the effect of **income** on subjective health?
SQ2) What is the effect of **education** on subjective health?
SQ3) What is the effect of **immigrant status** on subjective health?
SQ4) What is the effect of **unemployment** on subjective health?

1.5 Research Methodology

This research considers multiple socioeconomic determinants. With a literature review following from desk research, it is identified which determinants are relevant for this research. These are the explanatory variables of subjective health. Control variables and ways to measure health are also discussed in the literature review. The socioeconomic determinants, control variables and dependent variable all come from the European Social Survey (ESS). In this survey, participants are asked about their subjective health as well as about multiple socioeconomic determinants. The data is available for 38 European countries in 2018. So, the first part of the research is secondary: summarizing and synthesizing previous literature. The second part is primary. This concerns active participation and analyzing the data in a new way. The data analysis is performed in SPSS. To assess how and if the socioeconomic determinants significantly predict health, a complex sample multinomial logit model is used.

1.6 Thesis Outline

In the next chapter, a literature review is provided that substantiates the hypotheses for the previously mentioned sub and central research questions. This literature review discusses health, socioeconomic determinants of health, and health inequalities between European countries. Chapter 3 dives deeper into the methodology: explaining the research methodology, variable measurement, data collection and data analysis. Chapter 4 gives an overview of the results that follow from this data analysis. Lastly, chapter 5 closes with the conclusion, discussion, and suggestions for future research.

Chapter 2 Literature Review

The goal of this literature review is to form hypotheses based on previous research. Most of this information was selected from medical journals. Following the ranking of health economics journal by Haley (2016), The Lancet and Social Science and Medicine journal scored very well. Based on citations from articles in these journals, other useful and top-ranked journals were identified as well. These include, among others, the journal of Epidemiology and Community Health, and the American Journal of Epidemiology. Prof. Dr. Mackenbach is connected to the WHO and was another very helpful source, due to this research on health inequalities within Europe. Webster & Watson (2002) provided a useful way to organize this existing literature.

In the literature review, existing literature on health is discussed, as well as how health will be defined in this research. This is followed by health inequalities between countries. Then it is discussed what socioeconomic determinants impact health. The hypotheses on the sub questions are based upon this information. Lastly, a conceptual framework provides an overview of the relationships between the relevant variables.

2.1 Health

The World Health Organization WHO, defines health as a 'state of complete physical, mental and social well-being, not merely the absence of disease or infirmity' (WHO, 1948, p46). However, this definition has been heavily criticized. The number of people with chronic illnesses has been steadily increasing, so the vision of the WHO might be too utopian (Leonardi, 2018). A state of complete well-being means a state so extreme that it is impossible to achieve. Thus, there is a new idea to see health as a more dynamic process defined by someone's capabilities, and their ability to adapt to malaise and well-being conditions (Kingma, 2007). An option that aligns with this idea is subjective health.

Much of the research into health and socioeconomic determinants relies on subjective health (Hu et al., 2016; Jutz, 2020; Von Rueden, 2006). It can be very useful but has also been criticised, especially when subjective health surveys are used for cross-country comparisons (Hu et al., 2016; Jutz, 2020; Von Rueden, 2006). Subjective health is determined by the question: "How is your health in general" (Sarti & Zella, 2016, p121). Participants are asked to rank their answer on a Likert scale: very bad – bad – fair - good - very good. However, people in one culture might rank their health differently and with other standards than they do in another culture (Heine et al., 2002). Even people within the same country with different cultural backgrounds might exhibit large differences (Heine et al., 2002). This especially poses a problem for socioeconomic determinants of health when people with different cultural backgrounds also have different socioeconomic characteristics than natives to that country. On average, Italians assessed their health more positively than Finns, but more negatively than French people for example (Jylhä, 2009). The research does not mention whether this difference actually stems from their cultural

background, or whether French and Finnish people experience an objectively worse health. If the latter is the case, subjective health is still a suitable measurement tool.

Indeed, subjective health has been found to be a reliable health measure in multiple studies. Idler and Benyamini (1997) found that global self-rated health is an impressive predictor of mortality in nearly all of the 27 studies they examined. It correlates with objective measures, such as physician assessments and morbidity measurements (Pinquart, 2001; Sarti & Zella, 2016). Jylhä (2009) confirms this and claims that the reason why subjective health predicts mortality so well is because it is an inclusive measure of health. It is especially useful in recording subjective well-being, including life satisfaction, anxiety, and depression (Schneider et al., 2004). It combines the subjective experience of acute and chronic symptoms, and different feelings of well-being (Sarti & Zella, 2016). This could include feeling tired, having a backache and headaches. Thus, subjective health provides a more holistic view.

Overall, subjective health has its drawbacks but seems most suited for such a research. The effect of socioeconomic determinants on objective health has also been researched, but this was always on single categories of health. Their impact on life expectancy (Sede & Ohemeng, 2015), mortality (Mackenbach, 2017; Vleugelers et al., 2001) or depression (Margaretten et al., 2010) for example. When considering the general health of someone, there are simply too many objective parameters: how often are they sick, how sick do they get, are they disabled or chronically ill, does a CT scan show any abnormalities, what is their blood pressure, how are their blood values, and so on… (Wu et al., 2013). Even if all these factors could somehow be included in a research, it still does not take mental health into account, which is even more difficult to measure with objective parameters (Schedler et al., 1993). Thus, subjective health provides a holistic view on health and encompasses all parameters, which makes it suitable for this research.

2.2 Cross-country comparison of health inequalities

Now that health is discussed, it is time to address health inequalities. As mentioned before, there are health inequalities between countries, for example between Eastern and Western Europe. Stirbu et al. (2010) found the highest mortality in Central and Eastern European, and Baltic countries. This was followed by Northern and Western European countries, and smallest in the Southern European regions. Avoidable mortality contributed between 11 and 24% to the inequalities in life expectancy between higher and lower educated groups. This was confirmed by Di Girolamo et al. (2020), who considered changes in cardiovascular mortality between the 1990s and 2010s by gender, educational level, and occupational class. They found that overall mortality rapidly declined, but this decline occurred relatively faster for higher socioeconomic groups. Again, they found differences between countries. Mortality was lowest in Southern Europe, intermediate in Northern and Western Europe, and largest in Central Eastern Europe and Baltic countries.

Research on objective health inequalities, such as mentioned in mortality, is quite common. When it comes to cross-country subjective health inequalities, however, a lot less research has been done. According to Olsen & Dahl (2007), the research that has been done focusses almost exclusively on two or three countries. The studies that did include more countries focussed on the East-West division of Europe, and did not include other European regions. Olsen & Dahl (2007) did find that people in Nordic countries, but also countries such as Austria, and Switzerland show high levels of subjective health. Levels for Portugal and Greece were much lower. However, this research did not correct for the effect of socioeconomic determinants. Rathmann et al. (2015) looked at the effect of macro-level characteristics on subjective health, controlling for the effect of individual factors. They found that people in countries with higher income inequality and with liberal welfare tradition were associated with more subjective health complaints than those from countries with lower income inequality or the Social Democratic regime. This research was only focused on adolescents, however, so it is uncertain how this relationship holds for people of other ages.

Thus, there are differences in health inequalities across Europe. This makes it interesting to compare subjective health between countries, controlling for the effect of individual socioeconomic factors. The proposition that is further analysed in exploratory research is that there could be differences in subjective health between European countries, and between different European regions.

2.3 Socio-economic determinants of health

2.3.1 The Life-Course Approach

Following the explanation on health inequalities, the effect of socio-economic determinants on health inequalities can now be estimated. Populations that suffer health disparities are defined in the Health Care Fairness Act as those "With a significant disparity in the overall rate of disease incidence, morbidity, mortality, and survival rates in the population as compared to the health of the general population" (Shavers, 2007, p.1013). Or, according to the WHO, health inequalities are "avoidable inequalities in health between groups of people within countries and between countries" (World Health Organization, 2013, p3). Thus, it entails population groups that suffer a worse health than the general population, in spite of this being avoidable.

In the literature, there are two mechanisms that are used to explain how social determinants can cause such health inequalities: *materialistic mechanisms* and *psychosocial mechanisms*. The material mechanism works via material resources. The rationale is that income creates those resources, and those resources in turn provide possibilities for health advantages by investing income in (health promoting) goods and services (Jutz, 2020). This income is not only individual, but also includes access to public services such as education and transport. Absolute income or poverty loses its importance as a predictor of health when the most basic material conditions are met, such as hygiene, clean drinking water and

the prevention of hunger (Bambra, 2011). However, this is not the only explanation of health inequalities since these also exist in rich industrial nations where almost all material resources are met.

This brings us to the psychosocial mechanism. That entails that people compare themselves to others, to evaluate their position in society. When there is more social inequality, those comparisons are more distressing (Bambra, 2011). This could lead to chronic stress, which affects both your mental and physical health via the stress hormone cortisol (Jutz, 2020). Instead of revolving around absolute income and poverty, this mechanism works through status and relative income.

These two mechanisms combined form the life-course approach (Wilkinson, 1997). This approach considers the accumulation of health risks as the cause of health inequalities. Material, social and psychosocial (dis)advantages add up over a lifetime and manifest health inequalities. If the impact of education on health is considered for example, both the material and psychosocial mechanisms show very clearly. A higher education leads to a higher absolute income and better working conditions, both of which can lead to a healthier life (Yildirim, 2016). It can also increase health inequalities by signalling status. Following a high education reflects certain cognitive abilities and might be taken more seriously by health care employees. A higher education could also lead to more attention to health prevention and knowledge of a healthy lifestyle (Jutz, 2020).

Thus, the life-course approach provides mechanisms through which socioeconomic determinants might influence health. Following this, empirical research that has been done on this topic is discussed, to see if these mechanisms actually occur.

2.3.2 Empirical Research

When it comes to the socioeconomic determinants of health, quite some research has been done on this topic. Broadly, there have been two types of explanatory characteristics: contextual and individual. Much of the literatures focuses on the effects the individual characteristics, mainly education and income. In the research of Jutz (2020), 23 countries from Central and Eastern Europe and 20 countries in Western Europe were considered. In all those countries, income and education have a positive effect on health. This means that income and education are socioeconomic determinants of health in almost all European countries.

The reason why *income* is expected to increase health, is through bigger access to material resources. These resources can create access to sports, to fresh and healthy food products, to bigger social networks and most importantly, give access to good health care (Bambra, 2011). Gravelle & Sutton (2008) tested the relation between relative income and self-assessed health in Great Britain. They found a clear correlation of measures of individual income and health. This could be the case because of the lack of material resources that was discussed previously. The same positive relation between income and subjective health has been found in provincial China (Pei & Rodriguez, 2006), the US (Lillard et al., 2015), and Colombia (Hessel et al., 2018). Some moderating information did come from Ecob & Smith

(1999), who found that incidences of morbidity are linearly related to income, except for very high and low incomes. This could be because the research took place in England, where very low income households get government health support. People with very high incomes generally live unhealthier lifestyles (more calories, alcohol, smoking and stress) (Gage, 2006). These exceptions occurred in England and might be different in countries where low income households get lower health support. Still, it is important to consider that the relation between income and health is not necessarily linear. Income is an important indicator to include since its impact on health has a solid empirical foundation. It leads to the following hypothesis:

H1) Income has a positive effect on subjective health

When it comes to *education*, there is a widely studied positive effect of education on health, in every global region (Brunello et al., 2015). If we consider the impact of education on health via the life course approach, we see both the material and psychosocial mechanisms very clearly. A higher education leads to a higher absolute income and better working conditions, both of which can lead to a healthier life (Yildirim, 2016). It can also increase health inequalities by signalling status. Following a high education reflects certain cognitive abilities and might be taken more seriously by health care employees. A higher education could also lead to more attention to health prevention and knowledge of a healthy lifestyle (Jutz, 2020). Brunello et al. (2015) investigated the causal effect of education on health and found that health behaviours (smoking, drinking, exercising, BMI) account for a quarter of this effect. Thus, education might improve decision-making abilities as well. Empirical research also found such a positive relation. This effect has been found on both objective and subjective health parameters, all over the world (Arendt, 2005). A causal effect can still not be proven however. Hu et al. (2016) also considered 17 European countries from all over the continent and found that a less than good self-assessed health was more prevalent in lower education groups. This was particularly in Southern and Eastern Europe, and the Baltic states. Not only your own, but even your partner's education has an impact on your self-assessed health (Monden et al., 2003). This widely studied theoretical and empirical positive effect of education convinces me to include it as socioeconomic determinant. Although its effect is well-known, it is important to add education in order to know the full impact of socioeconomic determinants. This leads to the following hypothesis:

H2) Education has a positive effect on subjective health

Another one of the individual characteristics is *immigration*. Malmusi (2015) found that there are significant differences in self-reported health between immigrants and natives. Immigrants in Europe reported poorer health, even when adjusted by age, education, occupation, and socio-economic conditions. This may also be influenced by their environment. Lorant et al. (2008) found that compared to native-born Belgians, immigrant groups from Turkey and Morocco were more likely to have poorer self-rated health (Lorant et al., 2008). Disparities in mortality, morbidity and in health status have all

been found, for example in the US (Franks et al., 2003), in England and Wales (Wild & McKeigue, 1997) and in France (Guillot et al., 2019). Although migrants tend to have a lower socioeconomic status, early studies concluded that social class is not an important explanation of higher mortality (Wild & McKeigue, 1997).

Other research has shown contrary results, however, and finds a big influence of socioeconomic status on health disparities (Lindstrom et al., 2001). There are several possible explanations. Health disparities may be disease-specific, or the importance of the living environment might be overlooked. There might also be differences depending on which country or region someone migrated from. Discrimination has been found to have negative health effects among adult immigrant minorities. In Germany, the Netherlands, and Sweden, it was found that discrimination by the policy and security personnel was most common (Kauff et al., 2017). It also had the biggest negative effect on health outcomes. Thus, health may be dependent on immigrant status. Therefore, the following hypothesis is formed:

H3) Being an immigrant has a negative effect on subjective health

Another important factor is *unemployment*. Unemployment correlates negatively with health but could be both a cause and a consequence (Ross & Mirowsky, 1995). On the one hand, employment could improve the health of people via higher income and status. On the other hand, healthy people might get and keep jobs more than unhealthy people do. Ross & Mirowsky (1995) found, using longitudinal data, that full-time employment predicts slower declines in perceived health. This is the same for men as for women. A higher perceived health increased employment odds for women, but not for men. Nordtström et al. (2014) did a systematic review and found that most of the studies in the review showed a negative effect on health from unemployment. This was dependent on gender, age, geographic location, and education level. This relationship between employment and health is confirmed by Virtanen et al. (2002) in Finland. They concentrated on employment security. When this was perceived to be low, it was associated with lower self rated health, and more chronic disease and psychological distress. Van der Noordt et al. (2014) performed a systematic review on the health effects of employment. This confirms the psychological distress results, as shown by Virtanen et al. (2002). There was strong evidence for a protective effect of employment on depression and general mental health. They mention that there was insufficient evidence for general health due to a lack of studies, or inconsistent findings. However, it is known that subjective health is especially useful in recording subjective well-being, including life satisfaction, anxiety, and depression (Schneider et al., 2004). Thus, if employment has a protective effect on general mental health, it might also have such an effect on subjective health. Although there is a theoretical foundation, the empirical evidence is inconclusive. This makes it interesting to include it as a socioeconomic determinant of health. Most research indicates that when employment is lost, health decreases. Thus, the following hypotheses is formed:

H4) Unemployment has a negative effect on subjective health

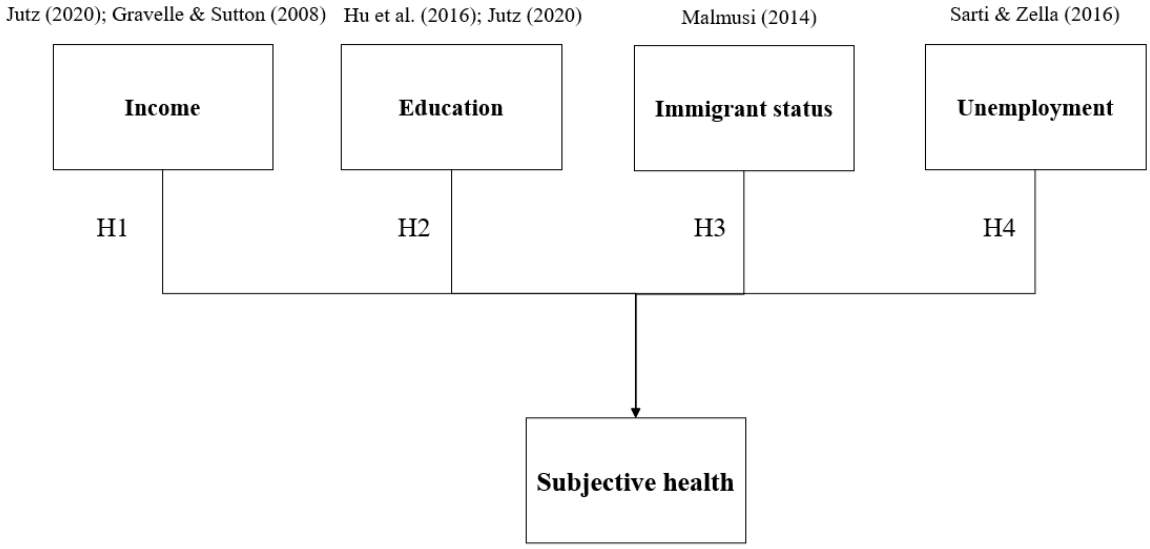
There has been empirical research on the effect of socioeconomic circumstances of children as well. Von Rueden et al. (2006) for example, considered parental educational status and family wealth as determinants of health for the children. They found that for children, a higher parental education status had a positive impact on health. For adolescents, family wealth was a greater indicator of health. This might have to do with status and access to social resources. Although this is an interesting approach, it often relies on recollection of adults about their childhood. These have proven to be unreliable (Chitkara et al., 2008; Offer et al., 2000). Furthermore, it lies outside of the scope of this research to consider childhood circumstances as well.

Apart from individual socioeconomic factors, there are also contextual factors. Those consider someone's social environment. Vleugelers et al. (2001) considered how the socioeconomic characteristics of someone's neighbourhood affect mortality in Canada. No significant association was found. However, when the effect of the social environment was considered on mortality in the U.S., there was a negative effect. The difference between the results might be due to the socio-cultural and political differences between the United States and Canada. Even when some individual characteristics such as income and education were controlled for, Yen & Kaplan (1999) found a significantly higher risk of death in neighbourhoods with a low social environment. However, this was based on a survey in the U.S. in 1983, which had a very different health care system than most European countries nowadays have. It also failed to consider individual socioeconomic factors such as employment or immigrant status. According to Stafford & McCarthy (2006), factors such as social capital and social cohesion in a neighbourhood might play a role in objective health, but both the definition and effects of them are vague. Because of this vagueness, the probable high correlation between individual factors and contextual factors, and to limit the scope of the research, the contextual factors are not further considered.

2.5 Conceptual framework

Figure 1

Conceptual Framework



In figure 1, the conceptual framework is shown. Income, education, immigrant status and unemployment are all socioeconomic determinants that impact self-perceived health. The relation between those determinants and subjective health is affected by the country of residence. The hypotheses and proposition following from the literature review have provided further insight in what such a relation might look like. In the next chapter, the methodology is discussed to either reject or support these hypotheses.

Chapter 3 Methodology

3.1 Research Methodology

To answer the central and sub research questions, desk research and data analyses are performed. The data that is used for this thesis comes from the European Social Survey, from now on abbreviated with ESS (ESS ERIC, 2021). This is an academically driven cross-national survey that includes 38 European countries (ESS, n.d.). Since ESS only provides European data, this research is confined to that continent. Since a lot of other research has also been done on European countries, it does allow for a good comparison with their results. An issue with this is that there could be missing data for some countries. In that case, this data or country might be left out. The countries can be grouped in Western, Central and Eastern, Northern and Southern Europe. Another issue is if the socioeconomic determinants are interrelated. Education, for example, also has a correlation with income. In general, the higher your education level, the higher your income (Manski, 1992). In that case, the predictive power of these variables is gone. Thus, it is important to run checks on the data that control for this correlation. The data analysis was performed in SPSS.

3.2 Measurements

3.2.1 Dependent variable

In this research, the effect of socioeconomic determinants on the dependent variable subjective health is considered. Subjective health is sensitive to cross-country comparisons, but delivers valuable information about someone's well-being, with regards to life satisfaction, anxiety, and depression (Schneider et al., 2004). The indicator used, and thus the operationalization of the dependent variable is "Subjective general health" in ESS. From now on, when there is spoken of health, this is what is referred to. Participants are asked the question "How is your health in general? Would you say it is…" (ESS, n.d.) There are multiple answers: very good – good – fair – bad – refusal - don't know - no answer. The last three are all registered as missing values. This is a standardized question recommended by the World Health Organization (WHO, n.d.). It refers to the participant's health in general, not to the present state of health, and it concerns physical, social, and emotional functions and biomedical signs and symptoms. Using this variable to measure the effect of socioeconomic determinants on health follows previous research (Hu et al., 2016; Jutz, 2020; Von Rueden et al., 2006).

3.2.2 Independent variables

The socioeconomic determinants of this research are the explanatory variables. These followed from the literature and include the following:

- Income
- Education
- Immigrant status

- Unemployment

This is in line with previous research, which also considered *income* (Jutz, 2020; Sede & Ohemeng, 2015), *education* (Hu et al., 2016; Jutz, 2020; Sede & Ohemeng, 2015; Von Rueden et al., 2006), *immigrant status* (Malmusi, 2015) and *unemployment* (Sarti & Zella, 2016).

Shavers (2007) identified methodological strengths and issues in measuring socioeconomic status in health inequality research. The first explanatory variable is income. Its strength is that it allows access to material goods and services that may influence health. However, this is age dependent and does not include all assets such as wealth or health insurance coverage. Therefore, age is a control variable. Respondents are asked for their year of birth, and based on this their age is calculated. Wealth is a difficult concept to measure, however, and very prone to reporting errors (Munda, 2015). Thus, income is used as explanatory variable. Participants are asked "Please tell me which letter describes your household's total income, after tax and compulsory deductions, from all sources? If you don't know the exact figure, please give an estimate. Use the part of the card that you know best: weekly, monthly, or annual income" (ESS, 2018, p58). Since 2008, a decile approach has been applied on income in the ESS. The categories are national, and based on deciles of the actual household income range in the given country (ESS, 2018). These deciles are derived from different sources, depending on the country. For the Netherlands, that source is the CBS. The median income is the reference point, and the ten deciles are calculated with the median as the top of the fifth decile. The incomes are displayed in the country's currency. Random letters were assigned to each of the deciles to make the question less painful. The possible answers were:

- J – 1st decile
- R – 2nd decile
- C – 3rd decile
- M – 4th decile
- F – 5th decile
- S – 6th decile
- K – 7th decile
- P – 8th decile
- D – 9th decile
- H – 10th decile

Other options were refusal, don't know and no answer which are all reported as missing values.

The second explanatory variable, education, has many advantages according to Shavers (2007). It is easy to measure, it captures lifestyle and behaviour aspects, and is predictive of your job, housing, neighbourhood, working conditions and income. A downside is that it has different social meanings in

different cultures, and that economic returns may differ significantly across gender groups (Shavers, 2007). To address the first issue, the International Standard Classification of Education is used. This divides education in multiple levels (UNESCO Institute for Statistics, 2012). Many other researchers on the effect of education on health use the same standard (Haas, 2008; Jutz, 2020; Stirbu et al., 2010; Von Rueden et al., 2006). This makes a comparison with their results easier and minimizes the cultural differences. Huisman et al. (2007) found that educational differences in self-assessed health cannot be expected to seriously overestimate educational differences in objective health. To address the second issue, differences in economic returns between gender groups, gender is added as control variable. To get the variable education, respondents are asked the question: 'What is the highest level of education you have successfully completed' (European Social Survey, 2016). The possible answers for highest level of education were as follows:

- ES-ISCED I – less than lower secondary
- ES-ISCED II – lower secondary
- ES-ISCED IIIa – lower tier upper secondary
- ES-ISCED IIIb – upper tier upper secondary
- ES-ISCED IV – advanced vocational, sub-degree
- ES-ISCED V1 – lower tertiary education, BA level
- ES-ISCED V2 – higher tertiary education over MA level

Vocational education includes educational programs that are designed to acquire the knowledge, skills, and competencies specific for a particular occupation or trade (ESS, 2016). Other options were other, refusal, don't know and no answer, which were reported as missing values.

Thirdly, immigrant status. As mentioned, there are often health differences between immigrants and natives (e.g., Malmusi, 2015). Immigrants in Europe reported poorer health, when adjusted by age, education, occupation, and socio-economic conditions. In ESS, the question is asked whether the participants were born in the country they are currently living in or not. The possible answers were either yes or no. Other options were refusal, don't know or no answer, which were all reported as missing values.

Fourthly, when it comes to unemployment, the following question is posed: "Using this card, which of these descriptions applies to what you have been doing for the last 7 days" (ESS, 2016, p34). There were several answering possibilities:

- In education (not paid for by employer) even if on vacation
- Paid work (Or away temporarily, either employee, self-employed or working for your family business)
- Unemployed

- Retired
- Doing housework, looking after children or other persons
- Doing community or military service
- Other

Due to the low respondence rate, the category of doing community or military service was included in the 'other' category. Respondents checked one of these boxes. When these boxes were not applicable, the 'other' box was checked.

Lastly, the country of residence. This is not a question on the survey, but something the ESS collects independently. The ESS itself keeps track of which surveys originate from which country, and attaches the connected label to it. There are 38 countries in total.

3.2.3 Control variables

To address the previously mentioned issues, the control variables age and gender are added. This focused on participants aged fifteen and up, following Stirbu et al. (2010) and Jutz (2020). Gender is a categorical variable and has male and female as options. The limitations of this binominal definition of gender are recognized, but the research is restricted to the data that is available. Both age and gender are obtained by asking participants in a survey about it. For age, respondents are asked for their year of birth and age is calculated based on this. Age has been found to correlation positively with both income and subjective well-being (Charles & Hurst, 2003; Shmotkin, 1990). The relationship between age and subjective health is more complicated, so it is crucial to add age as control variable. Gender can also influence the impact of education on health and is thus also added (Bertocchi & Bozzano, 2020). Mackenbach (2006) found mortality inequalities to be larger among men, with cardiovascular disease as the main culprit. Gender differences in subjective well-being in adults have been found to be small, although there were gender differences in health complaints (Shmotkin, 1990; Torsheim et al., 2006). Thus, both age and gender are included as control variables.

3.3 Data collection procedure

The way the data was procured was by downloading the entire dataset for the year 2018. From 2019 to 2021 there were no survey results due to COVID-19. For 2022, quite some data was missing since most of the data is collected at the end of the year. So 2018 was the most recent and complete year. The data comes from the European Social Survey ESS. This was primarily designed to monitor changing attitudes and values across Europe. Thus, it comprises a core module with questions that measure respondent's opinions on politics, social issues and other topics of interest to the social sciences. Opinions on climate change, democracy, immigration, gender equality and gay marriage are all included, for example. This core module also includes the most comprehensive set of socio-structural (background) variables of any cross-national survey. The respondents are all aged fifteen and over. In total, 32554 respondents answered the questions that were necessary for this research. ESS works together with research

organizations from within each country to perform these interviews. In the Netherlands, this is I&O Research for example. These organizations invite participants to an hour-long face-to-face interview that is computer assisted. This is a data collection method in which the interviewer reads questions to the respondents from the screen of a computer, laptop or a mobile device and enters the answers.

The aim of this face-to-face interview is to improve standards in the field of cross-national surveys. It is done by outlining the methods of ESS in a document, which all countries are required to adhere to. This includes information about the national questionnaire, preparing the data collection, conduction of the data collection and processing data (European Social Survey, n.d.). The survey involves strict random probability sampling, a minimum target response rate of 70%, and rigorous translation protocols.

Apart from ESS, another option to find health surveys was the World Health Survey Plus (WHS+) from the WHO (WHO, n.d.). Their aim is to address essential data gaps when it comes to health. They also monitor countries' progress towards health related SDGs. Although this is a database with a lot of indicators, most of them are extremely specific (e.g., whether alcohol content is displayed on containers) (WHO, n.d.). Additionally, all of them are objective parameters which makes it hard to get the general overview of mental and physical health. Furthermore, the values of these parameters are given per region but there is no additional information on the effect of socioeconomic determinants on this. Another option that was considered was Eurostat (Eurostat, n.d.). The issue with this database is the way in which the data is displayed. This makes it impossible to include multiple explanatory variables in the model. Therefore, although both databases provide invaluable information, they are not suited for this research.

3.4 Data analysis

3.4.1 Complex Sample

The data from ESS is drawn by complex sampling methods. The assumption of analytical procedures in SPSS is that the observations in a data file represent a simple random sample from the population of interest. Individual sampling units are then selected at random with equal probability and without replacement, directly from the entire population (IBM, 209). This is not the case for the surveys from the ESS, which makes it a complex sample. This means the analysis needs to be different, because the survey design needs to be accounted for to represent the population. If the complexity of the sample is not considered, it can lead to an under-estimate of the width of confidence intervals. This is an estimation bias that could also lead to wrong conclusions. The survey design for the ESS differs across countries, because they are designed to achieve a minimum effective sample size (Kaminska, 2020). This is done in the most cost-effective way for each country, considering each country context. Countries can differ in (1) magnitude of selection probabilities, (2) the variation among them, and (3) the clustering and stratification (Kaminska, 2020).

Especially the second difference is important since most countries use an address-based sample. This means that one person is selected at random at each address in a random sample of addressed. Therefore, people that live alone are twice as likely to be selected than people living in a household of two, and thrice as likely to be selected than people from a household of three, and so on. If this is not corrected for, the samples are heavily skewed towards people living alone. In Germany, a survey found that living alone was a good predictor of both physical and cognitive health (Beller & Wagner, 2017). Pasanen et al. (2021) found similar results from a survey in Finland, although they emphasize that people living along are a very heterogeneous group in terms of subjective health. Therefore, it is important to correct for the higher selection probability of people living alone.

Additionally, countries differ in nonresponse processes (Kaminska, 2020). This is reflected both in national response rates, but also in demographic differences among respondents. As we see in the dataset, three countries account for almost half of the respondents. Thus, post-stratification weights are introduced to reduce the impact of nonresponse error. They also correct for coverage and sampling errors with respect to the post-stratification variables. Those are based on gender, age, education and geographical region. A weighted analysis that uses post-stratification weights completely correct for errors in coverage, sampling and nonresponse.

The European social survey provides the analysis weight, which corrects for differential selection probabilities within each country as specified by sample design, nonresponse, noncoverage and sampling error, and it takes into account differences in population size across countries. This variable is provided in the integrated data file from ESS. The same goes for the variable 'psu' (the primary sampling units) which is filled in in clusters. For strata, the variable 'stratum' (the sampling stratum) is used. Lastly for sample weight, 'anweight' is used (the analysis weight).

In SPSS, multiple models can be constructed taking the complex sample into account. The models that are discussed below both have this option. For both models, alpha is set at 0.05 and the confidence intervals at 95%.

3.4.2 Complex sample ordinal logistic regression

An option in SPSS is the complex sample ordinal logistic regression. Many researchers that looked at individual objective parameters performed a logistic regression, among which Sede & Ohemeng (2015) and Vleugelers et al. (2001). They were able to do this since their dependent variables (life expectancy and mortality rates respectively) were ratio variables that were normally distributed. Sarti & Zella (2016) used a binominal regression model, which is done to predict the odds of seeing an event. In this case, those odds were that individuals changed their health status during their employment. However, since the dependent variable has 5 categories this is not relevant. Romero-Ortuno (2013) performed an ordinal regression to find the influence of education, income, smoking, alcohol intake and parental longevity on frailty. Frailty entails a state of vulnerability for elderly people. Although the dependent variable differs

between the research and this thesis, both are similar in their explanatory variables and the fact that the dependent variable is ordinal. Lelisho et al. (2022) also used ordinal regression, but this time to predict someone's status from socioeconomic determinants. Again, the dependent variable differs but the reasoning of the analysis applies. Lelisho et al. (2022) and Lall et al. (2002) conclude that ordinal regression is the best option to predict an ordinal form of health. However, there are certain assumptions that need to be met in order to do an ordinal regression:

1) The dependent variable is ordinal
2) One of more of the independent variables are either continuous, categorical, or ordinal
3) There is no multi-collinearity
4) There are proportional odds

The last assumption states that there are proportional odds. This means that between all health categories, there is a common slope for the effect of any of the explanatory variables. The effect of education on the likelihood of "very bad" health reports should be the same as the effect of education on the likelihood of "very good" health reports. This is quite a strong assumption. It is tested by the Wald test of parallel lines, which has the null hypothesis that regression parameters are equal for all categories (IBM, n.d.). This test compares the estimated model with one set of coefficients for all categories to a model with a separate set of coefficients for each category. If the test of parallel lines is significant, the general model with separate parameters for each category gives a significant improvement in the model fit (IBM, n.d.). In that case, the assumption of proportional odds is not met. This can be due to several things, including use of an incorrect link function or using the wrong model. If the ordinal regression model is not the right one, the multinomial logit model is an alternative According to Lall et al. (2002), this is the most appropriate method of analysis when one is presented with a grouped continuous response variable.

3.4.3 Complex sample multinomial logit model

This is a model that can be used if the assumption of proportional odds is not met. This model allows for different slopes, or different sets of odds ratios, for every category. The ordinal regression has a preference however, since this is more parsimonious. If the Wald test fails, however, a multinomial logit model is a good option. It models the response probabilities, that is the probability that a respondent is in a certain category. For example, how likely are unemployed people to rate their health as "very bad". There is always a reference category. For health, the reference category is "fair" since this is the neutral option in the middle. For income, the reference category is the lowest income decile. For education, it is the lowest level of education: ES-ISCED I , less than lower secondary. For immigrant status, the reference is being an immigrant and for unemployment the reference is being unemployed. Lastly, for country the reference category is the Netherlands since there is background knowledge on that country.

The model then estimates pairwise contrasts between the reference category and the other ones. So it runs all the different paired equations. This allows for a more precise estimation than would be possible in the ordinal regression model. A disadvantage is that it is harder to find a common effect across categories. The outcome of this model is the odds ratio. If this ratio is bigger than one, it means that those in the comparison category are more likely than those in the reference category to rate their health a certain way. If it is smaller than one, those in the comparison category are less likely to belong to that health category.

In SPSS, the model can be found under complex samples logistic regression. Of course, there are also assumptions that must be met for the multinomial logit model (Schreiber-Gregory et al., 2018):

1) Appropriate outcome structure – dependent variable is at the nominal or ordinal level
2) Observation independence – each independent variable has a single value for each case
3) Absence of multicollinearity
4) Large sample size

The first two assumptions are met, if the data structure of the ESS is considered. Assumption 3 states that explanatory variables, the socioeconomic determinants, do not demonstrate a linear relation between them. Regression coefficients are still consistent in such a case but are no longer reliable since the standard errors are inflated. The model's predictive power is not reduced, but the coefficients may not be statistically significant with a type II error. Thus, a check is needed for multi-collinearity. A correlation coefficient of 0.8 would indicate that there is a strong linear relationship between two variables (Alin, 2010). To see if that strong relation is multi-collinearity, a variance inflating factor VIF test is performed. This indicates the increase in variance of a regression coefficient as a result of collinearity. There is no formal cut-off value to use with VIF, but when it is higher than 4 it generally indicates that multicollinearity might exist, and that further investigation is required. When the value exceeds 10, there is significant multicollinearity that needs to be corrected (Craney & Surles, 2002; Senaviratna & Cooray, 2019). For the fourth assumption, a guideline is to multiply the number of explanatory variables by 10, and divide this by the expected probability of your least frequent outcome. If this is for example 500, then a minimum sample size of 500 is needed.

Thus, assumptions for the ordinal regression are tested first since this model is more parsimonious. If the assumptions are not met, a multinomial logit model needs to be used. Using these models, data from ESS is analysed. In the next chapter, the results of this are discussed.

Chapter 4 Results

4.1 Descriptives and assumptions

First, a description is given of the dependent variable Subjective general health.

Table 1

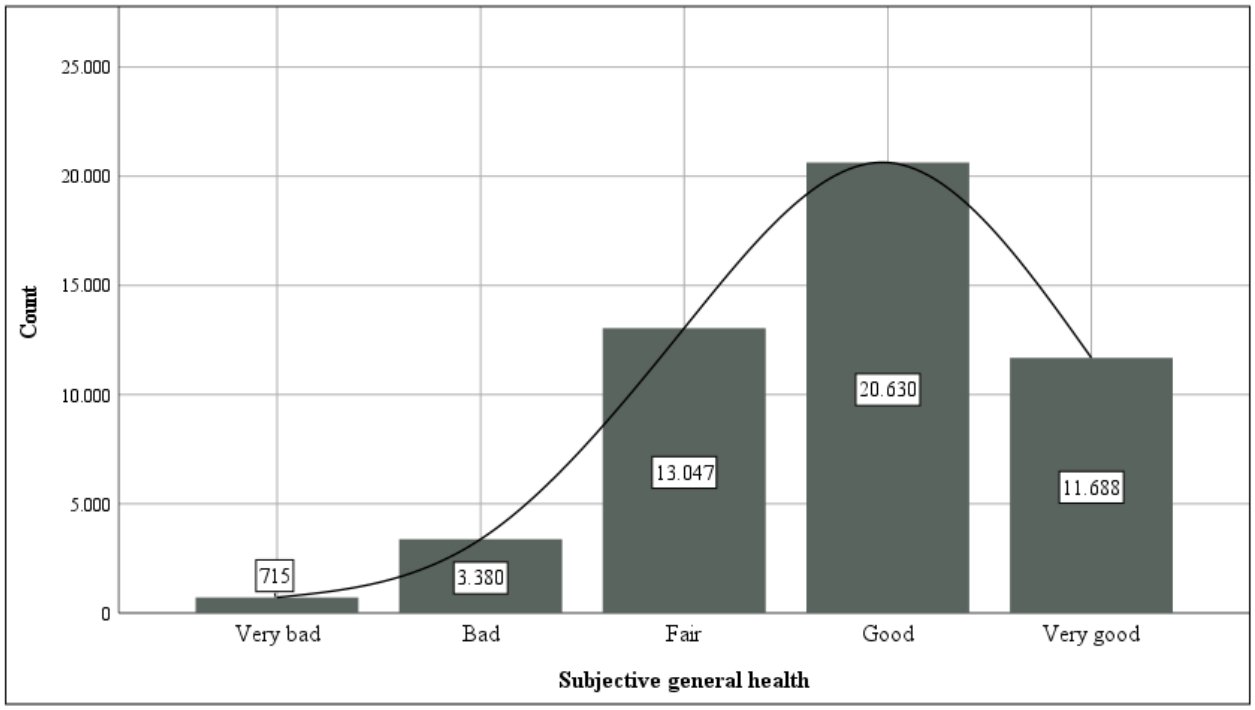
Descriptive Data of Subjective General Health

	N	Mean	Standard Deviation	Variance	Skewness statistic and s.e.	
Subjective general health	49460	2.21	.928	.861	.526	.011

As can be seen in table 1, 49460 respondents answered the question, of whom the average health was rate at 2.21 (s.d. = 0.93). For skewness, there is a value of 0.53 (s.e. = 0.01). This positive value for skewness indicates that the tail is on the right side of the distribution. This is confirmed in figure 2, which indicates the uneven distribution of health reports. Of all respondents, 65.3% claimed a “good” or “very good” health, whereas only 8.0% claimed a “bad” or “very bad” health.

Figure 2

Bar Count of Subjective General Health



Note. Numbers shown in the bars are the exact count.

After descriptive of the dependent variable, those of the explanatory variables are also given. First, it was checked whether any cells are empty or extremely small. This was not the case. Then, a missing value analysis was run. By far the most missing values are for income. 9654 values are missing, which is 19.5% of all respondents. This is quite high. Implications of this are discussed in chapter 5.

To continue, the frequencies of answer categories are found below in table 2. For some variables, points of emphasis are mentioned below the table.

Table 2

Variable Frequencies

		Weighted Count	Weighted Percent
Subjective general health	Very bad	462	1.4%
	Bad	2160	6.6%
	Fair	8654	26.6%
	Good	14134	43.4%
	Very good	7144	21.9%
Gender	Male	16055	49.3%
	Female	16499	50.7%
Household's total net income	1st decile	2828	8.7%
	2nd decile	3417	10.5%
	3rd decile	3254	10.0%
	4th decile	3492	10.7%
	5th decile	3354	10.3%
	6th decile	3375	10.4%
	7th decile	3550	10.9%
	8th decile	3358	10.3%
	9th decile	2821	8.7%
	10th decile	3105	9.5%
Highest level of education	ES-ISCED I	3358	10.3%
	ES-ISCED II	6109	18.8%
	ES-ISCED IIIa	6136	18.8%
	ES-ISCED III	5396	16.6%
	ES-ISCED IV	4001	12.3%
	ES-ISCED V1	3011	9.2%
	ES-ISCED V2	4418	13.6%
	Other	125	0.4%
Born in country	Yes	28647	88.0%
	No	3907	12.0%
Last7days	Education	2346	7.2%
	Housework	5258	16.2%
	Paid work	14779	45.4%
	Retired	7219	22.2%
	Unemployed	1958	6.0%

	Other	994	3.1%
Country	Austria	597	1.8%
	Belgium	871	2.7%
	Bulgaria	501	1.5%
	Switzerland	558	1.7%
	Cyprus	58	0.2%
	Czechia	578	1.8%
	Germany	6251	19.2%
	Denmark	406	1.2%
	Estonia	108	0.3%
	Spain	2866	8.8%
	Finland	426	1.3%
	France	4778	14.7%
	United Kingdom	4384	13.5%
	Croatia	257	0.8%
	Hungary	476	1.5%
	Ireland	251	0.8%
	Iceland	26	0.1%
	Italy	2896	8.9%
	Lithuania	201	0.6%
	Latvia	142	0.4%
	Montenegro	40	0.1%
	Netherlands	1180	3.6%
	Norway	383	1.2%
	Poland	1952	6.0%
	Portugal	682	2.1%
	Serbia	447	1.4%
	Sweden	759	2.3%
	Slovenia	153	0.5%
	Slovakia	326	1.0%

Note. N = 32554. Missing values were removed from the dataset.

As can be seen in table 2, most people rate their health as good (43.4%). This is followed by fair (26.6%), then very good (21.9%), then bad (6.6%) and lastly very bad (1.4%). The control variables gender and age are distributed quite even.

As for the explanatory variable, income answers are also quite evenly divided over the deciles. There is only a slightly lower response rate for the 1st (8.7%) and 9th decile (8.7%). This deviation is not large enough to cause concern. When it comes to the next explanatory variable, education, the distribution is slightly less even with a majority of people who finished ES-ISCED II or III as highest education. These are the lower and upper levels of secondary education.

For immigrant status, table 2 shows that the vast majority of the respondents (88.0%) was born in the country where the survey took place. As for last7days, almost half of the respondents performed paid work in the past week. Lastly, there seems to be a Western dominance in the dataset. The ones most present are Germany (19.2%), France (14.7%) and the United Kingdom (13.5%). Together, these three countries account for almost half of the entire dataset.

After this description of the data, the following assumptions are tested:

1) Normality
2) Multicollinearity
3) Proportional odds

First, to test for normality, a Kolmogorov-Smirnov test is run. This confirms the suspicions of abnormality that arose from figure 2 ($D(32554) = 0.24$, $p < 0.01$). Thus the distribution is positively skewed. However, since the sample size is large (N=32554), skewness is no real problem for statistical tests according to the central limit theorem. However, it is important to keep in mind that in the ordinal regression the complementary log-log function might fit better with the data than the logit function due to this positively skewed distribution.

After the normality tests, the ones for multicollinearity are run. If the explanatory variables are significantly correlated with each other, it affects coefficients and p-values. Since the aim of this thesis is to look at the effect of individual variables, this assumption is important. As table 3 shows, the VIF values are all below 5. This indicates that there is no reason to suspect multicollinearity between the variables.

Table 3

Collinearity Statistics

	Tolerance	VIF
Born in country	.99	1.01
Household's total net income. all sources	.89	1.13
Age of respondent. calculated	.93	1.08
Highest level of education. ES - ISCED	.95	1.05
Gender	.99	1.01

The last assumption to be tested is that of proportional odds. This is tested by Wald's test of parallel lines. This test is significant ($F(d1 = 159.0, d2 = 18642.0) = 4.44$, $p < 0.01$). The assumption has been tested on all available link functions. However, the test was significant with p values below 0.01 independent of which model was used. This leads to the conclusion that the assumption of parallel lines cannot be met for this data. This means that the slopes for the health categories significantly differ. If this is the case, the multinomial logit model should be used. As mentioned in chapter 4, this model accounts for different slopes depending on the health categories. It takes one health category as reference and estimates pairwise contrasts between the reference and the other categories. The results in the next sections follow from this multinomial logit model.

4.2 Results multinomial logit model

4.2.1 Complete model

After the complex sample multinomial logit model was run, the following classification table is given that compares the observed and predicted categories of Subjective general health.

Table 4

Comparison of the Observed and Predicted Categories of Subjective General Health

Observed	**Predicted**					
	Very bad	Bad	Fair	Good	Very good	**Percent Correct**
Very bad	146.02	1.57	218.96	87.25	8.42	**31.6%**
Bad	300.67	14.50	1033.29	742.70	68.98	**0.7%**
Fair	324.76	10.53	3168.90	4647.95	501.68	**36.6%**
Good	154.45	6.34	2122.25	10056.10	1794.95	**71.1%**
Very good	35.95	3.77	449.75	4688.57	1965.75	**27.5%**
Overall Percent	**3.0%**	**0.1%**	**21.5%**	**62.1%**	**13.3%**	**47.2%**

These results can be compared to those of a null model. This is without explanatory variables, and would classify all customers into the health category that occurs most often ("Good"). Given the observed frequency of that category, the null model would be correct 43.4% of the time. As can be seen in table 4, the multinomial logit model is correct 47.2% of the time. Thus, it correctly classifies 3.8% more of the respondents than the null model does. In particular, the model does considerably better at classifying those who have a "very bad", "fair", "good" or "very good" health. The model is, however, worse in predicting "bad" health. This is predicted much less often (0.1%) than it actually occurs (6.6%). The number of guesses for "good" health are slightly overstated (62.1% guessed, occurs 43.4%). After these general model effects, the effect of the independent variables is now discussed.

Table 5

Model Effects on Subjective General Health

Source	df1	df2	Wald F	Sig.
Gender	4.00	18797.00	1.42	.22
Immigrant status	4.00	18797.00	.41	.80
Education	28.00	18773.00	8.66	.00
Last 7 days	24.00	18777.00	9.24	.00
Income	36.00	18765.00	11.22	.00
Country	112.00	18689.00	17.82	.00
Age	4.00	18797.00	88.09	.00

Note. Dependent variable is Subjective general health (reference category is "fair").

Table 5 shows that the variables education ($F = 8.66$, $p < 0.01$), last 7 days ($F = 9.24$, $p < 0.01$), income ($F = 11.22$, $p < 0.01$), country ($F = 17.82$, $p < 0.01$), and age ($F = 88.09$, $p < 0.01$) have a significant effect on Subjective general health. Gender ($F = 1.42$, $p = 0.22$) and immigrant status ($F = 0.41$, $p = 0.80$) do not have a significant effect. The odds ratios of the control variables are discussed first. In all of the following results, "Fair" is the reference category for Subjective general health.

Table 6

Odds Ratios Gender

	Subjective general health	Odds Ratio	Confidence Interval	
			Lower	Upper
Male vs. Female	Very bad	.98	.71	1.35
	Bad	1.10	.96	1.27
	Good	1.11	1.01	1.21
	Very good	1.08	.97	1.19

As mentioned, gender is not a significant predictor of Subjective general health (F=1.42, p = 0.23). Table 6 displays the odds ratio of health for male and female. The only category where there is a significant difference between males and females is "good". Here, males are 1.11 times more likely to report "good health" (95% CI [1.01, 1.21]). However, considering the lower boundary of the confidence interval (1.01), this difference is barely significant. On the other hand, the other control variable age is significant (F = 88.10, p < 0.01).

Table 7

Odds Ratios Age

Subjective general health	Odds Ratio	Confidence Interval	
		Lower	Upper
Very bad	.10	.98	1.01
Bad	1.01	1.00	1.01
Good	.99	.98	.99
Very good	.96	.96	.97

Note. Odds ratios are displayed for one unit change in age, one year.

The effect for "very bad" health is not significantly different from one. For "bad" health, this is also the case. For "good" and "very good" health however, as age increases, respondents are significantly less likely to rate their health as "good" or as "very good". They are 0.99 times less likely to report "good" health (95% CI [0.98, 0.99]) and 0.96 times as likely to report "very good" health (95% CI [0.96, 0.97]).

Thus, there is a slight negative relationship between age and Subjective general health. Therefore, it is necessary to include this control variable in the model. The other control variable, gender, does not have a significant effect on the likelihood of a certain health category being chosen. The model was also run without this control variable, but this did not change the outcome. Following these control variables, the effect of each of the explanatory variables is now discussed.

4.2.2 Income

In this section, the first explanatory variable income is analysed. This is done to either reject or accept the first hypothesis:

Hypothesis 1) Income has a positive effect on subjective health

As mentioned before, income has a significant effect on Subjective general health (F = 11.22, p <0.01). To further specify this effect per income decile, the odds ratios are shown below in table 8.

Table 8

Odds Ratios Income

	Subjective general health	Odds Ratio	Confidence Interval Lower	Upper
2nd decile	Very bad	.19	.12	.32
	Bad	.88	.68	1.14
	Good	1.09	.87	1.36
	Very good	1.07	.83	1.40
3rd decile	Very bad	.17	.10	.28
	Bad	.64	.48	.86
	Good	1.12	.89	1.39
	Very good	1.25	.97	1.62
4th decile	Very bad	.11	.05	.26
	Bad	.45	.34	.60
	Good	1.38	1.09	1.73
	Very good	1.30	1.00	1.69
5th decile	Very bad	.23	.14	.38
	Bad	.42	.30	.58
	Good	1.36	1.08	1.70
	Very good	1.42	1.09	1.86
6th decile	Very bad	.17	.09	.32
	Bad	.48	.34	.68
	Good	1.47	1.17	1.85
	Very good	1.57	1.21	2.05
7th decile	Very bad	.13	.06	.28
	Bad	.47	.33	.65
	Good	1.64	1.30	2.06
	Very good	1.66	1.25	2.19
8th decile	Very bad	.13	.05	.33
	Bad	.56	.39	.81
	Good	1.61	1.26	2.06
	Very good	1.81	1.38	2.38
9th decile	Very bad	.15	.06	.37
	Bad	.47	.33	.68
	Good	1.55	1.21	1.98
	Very good	1.77	1.33	2.37
10th decile	Very bad	.15	.07	.31
	Bad	.52	.35	.75
	Good	1.84	1.43	2.38
	Very good	2.38	1.77	3.21

Note. The reference category is the first income decile.

As can be seen in table 8, compared to those in the first decile, respondents in the second are 0.20 times as likely to rate their health as “very bad” (95% CI [0.12, 0.32]). Moving to the third decile, both the categories “very bad” and “bad” health become less likely than for those in the first decile. (OR = 0.17, 95% CI [0.10, 0.28] and OR = 0.64, 95% CI [0.48, 0.86] respectively). For the fourth income decile, respondents are again less likely to report a ”very bad” or “bad” health (OR = 0.11, 95% CI [0.05, 0.26] and OR = 0.45, 95% CI [0.34, 0.60] respectively). They also become 1.38 times more likely to report a “good” health (95% CI [1.09, 1.73]).

From the fifth decile onwards, all categories show significant differences in how likely they are for respondents from the first versus the other deciles. Those in the fifth income decile are 0.23 times as likely to report a “very bad” health (95% CI [0.14, 0.38]) and 0.42 times as likely to report a “bad” health (95% CI [0.30, 0.58]). They are 1.36 times more likely to report “good” health (95% CI [1.08, 1.70]) and 1.42 times more likely to report “very good” health (95% CI [1.09, 1.86]). Respondents in the sixth income decile are 0.17 times as likely to report “very bad” health (95% CI [0.10, 0.32]) and 0.48 times as likely to report “bad” health (95% CI [0.35, 0.68]). They are 1.47 times more likely to report a “good” health and (95% CI [1.17, 1.85]) and 1.57 times more likely to report a “very good” health (95% CI [1.21, 2.05]).

People in the seventh income decile are 0.13 times as likely to report a “very bad” health (95% CI [0.06, 0.28]) and 0.47 times more likely to report “bad” health (95% CI [0.33, 0.65]). They are 1.64 times more likely to report “good” health (95% CI [1.30, 2.06]) and 1.66 times more likely to report a “very good” health (95% CI [1.25, 2.19]). Respondents in the eighth decile are 0.13 times as likely to report a “very bad” health (95% CI [0.05, 0.33]) and are 0.56 times as likely to report a “bad” health (95% CI [0.39, 0.81]). They are 1.55 times more likely to report “good health” (95% CI [1.21, 1.99]) and 1.77 times more likely to report “very good” health (95% CI [1.33, 2.37]).

Respondents from the ninth income decile are 0.15 times as likely to report “very bad” health (95% CI [0.07, 0.31]) and 0.47 times as likely to report “bad” health (95% CI [0.33, 0.68]). They are 1.55 times more likely to report “good” health (95% CI [1.21, 1.99]) and 1.774 times more likely to report “very good” health (95% CI [1.33, 2.37]). Lastly, the tenth decile. The ones in this income decile are 0.15 times as likely to report “very bad” health (95% CI [0.07, 0.31]) and 0.52 times as likely to report “bad” health (95% CI [0.35, 0.75]). They are 1.84 times as likely to report “good” health (95% CI [1.43, 2.38]) and 2.38 times as likely to report “very good” health (95% CI [1.77, 3.21]).

To conclude, as the income gap increases, so does the health gap. People from the first income decile are significantly more likely to report “very bad” and “bad” health than any other income decile. When

it comes to being more likely to report “good” and “very good” health, no significant difference is found for the second and third income decile. For the fourth, there is only a significant difference for the category “good”. From the fifth decile onwards, however, this difference is significant and increasing almost everywhere. Only the odds ratio for “good” from the seventh to the eight income decile is decreasing. People from the fifth income decile are 1.42 times as likely to report “very good” health than those from the first income decile. This difference has risen to 1.81 times more likely for the eight decile and even to 2.38 for the tenth. Thus, the hypothesis of a positive effect of income on subjective health is accepted.

4.2.3 Education

The second explanatory variable, education, is analyzed in this section. This is done to either reject or accept the second hypothesis:

Hypothesis 2) Education has a positive effect on subjective health

As mentioned before, education has a significant effect on Subjective general health ($F=8.656$, $p<0.01$). To further specify this effect per education level, the odds ratios are shown below in table 9.

Table 9

Odds Ratios Education

	Subjective general health	Odds Ratio	Confidence Interval	
			Lower	Upper
ES-ISCED II	Very bad	.19	.11	.33
	Bad	.59	.45	.78
	Good	1.13	.91	1.40
	Very good	.93	.71	1.21
ES-ISCED IIIa	Very bad	.21	.11	.41
	Bad	.70	.54	.91
	Good	1.33	1.08	1.63
	Very good	1.12	.88	1.44
ES-ISCED IIIb	Very bad	.23	.12	.43
	Bad	.58	.43	.77
	Good	1.41	1.14	1.73
	Very good	1.31	1.02	1.68
ES-ISCED IV	Very bad	.25	.13	.48
	Bad	.59	.44	.79
	Good	1.47	1.18	1.83
	Very good	1.31	1.01	1.69
ES-ISCED V1	Very bad	.26	.13	.49
	Bad	.69	.50	.96
	Good	1.66	1.33	2.08
	Very good	1.65	1.26	2.15
ES-ISCED V2	Very bad	.32	.18	.56
	Bad	.60	.44	.81
	Good	1.71	1.37	2.14
	Very good	1.99	1.53	2.58

Note. Reference category is ES-ISCED I.

The lowest level of education on the ES-ISCED scale is taken as reference category. This is ES-ISCED I and includes people who reported less than lower secondary as their highest level of education. The first comparison is between the education levels I and II. People from the second education level are 0.19 times as likely to report "very bad" health (95% CI [0.11, 0.33]) and 0.59 times as likely to report "bad" health (95% CI [0.45, 0.78]) compared to those with education level I. There is no significant difference for the "good" and "very good" categories. Respondents with education level IIIa are 0.21 times as likely to report "very bad" health (95% CI [0.11, 0.41]) and 0.7 times as likely to report "bad" health (95% CI [0.54, 0.91]). There are again no significant differences for the other two categories.

From education level IIIb onwards, there are indeed significant differences for every category. People that followed IIIb education are 0.23 times as likely to report "very bad" health (95% CI [0.12, 0.43]) and 0.58 times as likely to report "bad" health (95% CI [0.43, 0.77]). In this case, however, there is a significant difference for the other two categories. These people are 1.41 times more likely to report "good" health (95% CI [1.14, 1.73]) and 1.31 times more likely to report "very good" health (95% CI [1.02, 1.68]). Respondents with education level IV are 0.25 times as likely to report "very bad" health (95% CI [0.13, 0.48]) and 0.59 times as likely to report "bad" health (95% CI [0.44, 0.79]). They are 1.47 times more likely to report "good" health (95% CI [1.18, 1.83]) and 1.31 times more likely to report "very good" health (95% CI [1.01, 1.69]).

As for the tertiary education, this is divided up into lower and higher tertiary education. People in the lower tertiary education (level V1) are 0.26 times as likely to report "very bad" health (95% CI [0.13, 0.48]) and 0.69 times as likely to report "bad" health (95% CI [0.50, 0.96]). They are 1.66 times more likely to report "good" health (95% CI [1.33, 2.08]) and 1.65 times more likely to report "very good" health (95% CI [1.26, 2.15]). Lastly, the people that followed higher tertiary education (level V2). Compared to people from education level I, they are 0.32 times as likely to report "very bad" health (95% CI [0.18, 0.56]) and 0.60 times as likely to report "bad health" (95% CI [0.44, 0.81]). They are 1.71 times more likely to report "good" health (95% CI [1.37, 2.14]) and 1.99 times more likely to report "very good" health (95% CI [1.53, 2.58]).

To conclude, there is a similar effect of education to that of income. Just as in income, there is a difference in how likely someone is to rate their health as "bad" or "very bad" in all education levels compared to the first. The odds ratio for "bad" health seems quite stable among the categories. Similar to income, once the gap between the first and the other education levels is big enough, there is a consistent difference in how likely someone is to rate their health "good" or "very good". This gap starts from education level IIIb onwards, the upper tier of upper secondary education. Contrary to the likelihood of "bad" and "very bad", there is a clearer trend here. The higher the education level of

comparison, the higher the odds ratio. Thus, the hypothesis of a positive effect of education level on subjective health is accepted.

4.2.4 Immigrant status

Following, the explanatory variable immigrant status is analyzed. This is done to either reject or accept the third hypothesis:

Hypothesis 3) Being an immigrant has a negative effect on subjective health

As mentioned before, the explanatory variable immigrant status does not significantly predict health ($F=0.41$, $p = 0.80$). To further consider this effect, the odds ratios for immigrant status are given below in table 10.

Table 10

Odds Ratios for Immigrant Status

	Subjective general health	**Odds Ratio**	**Confidence Interval**	
			Lower	**Upper**
Born in country	Very bad	.88	.55	1.42
	Bad	1.01	.77	1.32
	Good	.94	.80	1.10
	Very good	.92	.77	1.09

Note. Reference category is not being born in the country of respondence.

The reference category is being an immigrant. As can be seen in table 10, none of the health categories are significantly more or less likely to be reported by immigrants than by non-immigrants. It can also be seen in the wide dispersion of the confidence intervals that the estimates of the odds ratio are not very precise. Thus, the hypothesis on the negative effect of being an immigrant on subjective health is rejected.

4.2.5 Unemployment

The next explanatory variable to be analysed is unemployment. This is done to either reject or accept the fourth hypothesis:

Hypothesis 4) Unemployment has a negative effect on subjective health

As mentioned before, the main occupation in the last 7 days has a significant effect on Subjective general health ($F = 9.24$, $p<0.01$). To further specify this effect per occupation in the last week, the odds ratios are shown below in table 11.

Table 11

Odds Ratios of Last 7 Days

	Subjective general health	**Odds Ratio**	**Confidence Interval**	
			Lower	**Upper**
Education	Very bad	.41	.05	3.08
	Bad	.80	.42	1.52
	Good	1.75	1.29	2.36
	Very good	2.71	1.96	3.73
Housework	Very bad	1.57	.58	4.26
	Bad	.85	.58	1.26
	Good	1.49	1.20	1.84
	Very good	1.54	1.19	1.99
Paid work	Very bad	.58	.21	1.61
	Bad	.59	.42	.83
	Good	1.50	1.23	1.83
	Very good	1.61	1.27	2.05
Retired	Very bad	2.05	.89	4.76
	Bad	.85	.58	1.24
	Good	1.05	.84	1.31
	Very good	1.24	.92	1.67
Other	Very bad	14.90	5.63	39.43
	Bad	3.63	2.41	5.47
	Good	.67	.48	.94
	Very good	.80	.53	1.22

Note. Reference category is unemployed people

As table 11 indicates, compared to unemployed people, people who followed education for the last seven days did not significantly differ in their likelihood to either report “very bad” or “bad” health. They were, however, 1.75 times more likely to report “good” health (95% CI [1.29, 2.36]) and 2.71 times more likely to report “very good” health (95% CI [1.96, 3.73]). The same thing occurs when comparing unemployed people to those who primarily did housework in the last 7 days. There was no significant difference in the likelihood to report “very bad” or “bad” health. People who did housework were, however, 1.49 times more likely to report “good” health (95% CI [1.20, 1.84]) and 1.54 times more likely to report “very good” health (95% CI [1.19, 1.99]).

Compared to unemployed people, those performing paid work in the last seven days did also not significantly differ in the likelihood to report a “very bad” health (95% CI [0.21, 1.61]). They were, however, 0.59 times as likely to report a “bad health” (95% CI [0.42, 0.83]). They were also 1.50 times more likely to report “good” health (95% CI [1.23, 1.83]) and 1.61 times more likely to report “very good” health (95% CI [1.27, 2.05]). Lastly, retired people do not significantly differ from unemployed people in how they rate their health in any category. Their likelihood of rating their health as “very bad” (95% CI [0.89, 4.76]), as “bad” (95% CI [0.58, 1.24]), as good (95% CI [0.84, 1.31]) and as “very good” (95% CI [0.92, 1.67]) is all insignificant. Lastly, those in the ‘other’ category are compared to unemployed people. They are 14.90 times more likely to report “very bad” health (95% CI [5.63, 39.43]). and 3.63 times more likely to report “bad” health (95% CI [2.41, 5.47]). They are 0.67 times as likely to report “good” health (95% CI [0.48, 0.94]), but there is no significant difference in likelihood of reporting “very good” health (95% CI [0.53, 1.22]).

To conclude, there are mixed results when it comes to the effect of past week unemployment and health. Compared to unemployed people, people who followed education were more likely to report good and very good health but did not differ in how likely they were to report very bad or bad health. This is also true for people who did housework in the last week. People that performed paid work did not differ from unemployed people in their rating of very bad health, but they were less likely to rate their health as bad and more likely to rate it as good or very good. There were no significant differences between retired and unemployed people. People from ‘other’ occupations, however, were more likely to report very bad and bad health, and less likely to report good and very good health than unemployed people. So people who followed education, did housework, or performed paid work seem to have a better health than unemployed people. People in the ‘other’ category perform worse, however, and those in retirement do not differ. Therefore, the hypothesis of a negative effect of unemployment on subjective health cannot be either fully rejected or accepted.

4.3 Exploratory research on country of residence

Apart from the effect of the previously mentioned explanatory variables, the country of residence might also play a role in health inequalities. As mentioned before, it did have a significant impact on subjective health ($F=17.82$, $p<0.01$). To see the differences between countries, the odds ratios are calculated. The Netherlands is chosen as reference category, due to personal knowledge of the country. In appendix B, one can find the odds ratios per health category for every country. Since a lot of countries are included in the analysis, only the significant effects are discussed here. The countries are divided up into the European regions to explore if there is commonality within these regions. Appendix A shows an overview of which countries belong to which region.

Baltic Countries

In Estonia, they are 0.48 times as likely to rate their health as "good" (95% CI [0.40, 0.59]) and 0.43 times as likely to rate it as "very good" (95% CI [0.34, 0.54). In Lithuania they are 0.60 times as likely to rate their health as "good" (95% CI [0.48, 0.76]) and 0.47 times as likely to rate it as "very good" (95% CI [0.35, 0.63]). In Latvia, they are 0.37 times as likely to rate their health as "good" (95% CI [0.28, 0.48]) and 0.29 times as likely to rate it "very good" (95% CI [0.21, 0.41]). Thus, compared to the Netherlands, people in the Baltic countries are less likely to rate their health as "good" and less likely to rate it as "very good". So a positive health is less likely.

Central & Eastern Europe

In Czechia, people are 1.47 times more likely to rate their health as "bad" (95% CI [1.01, 2.13]) and 0.74 times as likely to rate it as "good" (95% CI [0.60, 0.91]). In Hungary they are 0.76 times as likely to rate their health as "good" (95% CI [0.59, 0.97]). In Poland, they are 1.48 times more likely to rate their health as "bad" (95% CI [1.01, 2.16]) and 0.69 times as likely to rate it as "good" (95% CI [0.55, 0.88]). In Serbia, they are 4.42 times more likely to rate their health as "very bad" (95% CI [1.24, 15.80]), 1.86 times more likely to rate it as "bad" (95% CI [1.31, 2.65]), 0.5 times as likely to rate it as "good" (95% CI [0.40, 0.63]). In Slovenia, they are 0.73 times as likely to rate their health as "good" (95% CI [0.59, 0.91]). In Slovakia, they are 2.15 times as likely to rate their health as "bad" (95% CI [1.47, 3.14]). There are no significant differences between and Austria and Bulgaria, and the Netherlands. For the countries that do show a significant effect, a negative health is thus more likely and a positive health less likely.

Northern Europe

In Denmark, they are 2.71 times more likely to rate their health as "very good" (95% CI [2.06, 3.56]). In Finland, they are 0.79 times as likely to rate their health as "good" (95% CI [0.65, 0.96]). In Norway, they are 2.05 times more likely to rate their health as "very good" (95% CI [1.58, 2.65]). In Sweden,

they are 2.39 times more likely to rate their health as “very good” (95% CI [1.85, 3.10]). In Iceland, they are 2.34 times more likely to rate their health as “very good” (95% CI [1.77, 3.08]). Thus, in most countries people are over two times as likely to rate their health as “very good”, except for Finland.

Southern Europe

On Cyprus, people are 4.20 times more likely to rate their health as “very good” (95% CI [3.05, 5.77]). In Spain, they are 0.76 times as likely to rate their health as “good” (95% CI [0.61, 0.93]). In Croatia they are 2.20 times more likely to rate their health as “bad” (95% CI [1.51, 3.22]), 0.63 times as likely to rate their health as “good” (95% CI [0.48, 0.83]) and 2.54 times as likely to rate their health as “very good” (95% CI [1.93, 3.35]). In Italy they are 1.38 times more likely to rate their health as “very good” (95% CI [1.06, 1.80]). In Montenegro, they are 1.88 times more likely to rate their health as “bad” (95% CI [1.20, 2.95]) 0.65 times as likely to rate their health as “good” (95% CI [0.49, 0.86]), and 2.54 times more likely to rate it as “very good” (95% CI [1.91, 3.38]). In Portugal, they are 0.48 times as likely to rate their health as “good” (95% CI [0.37, 0.60]) and 0.65 times as likely to rate it as “very good” (95% CI [0.49, 0.87]). Thus, there are mixed results.

Western Europe

In Switzerland, people are 1.29 times more likely to rate their health as “good” (95% CI [1.03, 1.63]) and 3.59 times more likely to rate it as “very good” (95% CI [2.80, 4.62]). In Germany, they are 0.63 times as likely to rate their health as “good” (95% CI [0.52, 0.78]) and 0.75 times as likely to rate it as “very good” (95% CI [0.60, 0.94]). In France, they are 0.68 times as likely to rate their health as “good” (95% CI [0.55, 0.85]). In the UK they are 1.99 times more likely to rate their health as “very good” (95% CI [1.56, 2.53]). In Ireland they are 3.70 times more likely to rate their health as “very good” (95% CI [2.84, 4.83]). They are no significant differences between the Netherlands and Belgium. Thus, there are again mixed results. Health in Switzerland is rated more positively, whereas in Germany it is rated more negatively.

To conclude, in the Baltic countries and Central and Eastern Europe, people are less likely to rate their health positively. In Central and Eastern Europe, a negative health is more likely. In Northern Europe, except for Finland, people are more likely to rate their health positively. Southern Europe shows mixed results with countries like Cyprus, Croatia, Italy and Montenegro where people are more likely to rate their health as “very good”. In two of those countries however, Croatia and Montenegro, they were also more likely to rate their health as “bad”. Lastly in Western Europe, there are again mixed results. Overall, it does seem to make sense to group the countries into regions when comparing their subjective health to that in the Netherlands. Although there are mixed results, especially for Southern and Western Europe, regional trends can be clearly identified for the other regions. It should, however, always be kept in mind that there are intraregional differences as well.

Chapter 5 Conclusion & Discussion

In this final chapter, the conclusion and discussion of the results are presented. The answers to the research questions are given in the conclusion, and the previous research, limitations, and future research possibilities are given in the discussion.

5.1 Conclusion

In this conclusion, the main research question will be answered:

What is the Relation between Socioeconomic Determinants and Subjective Health In European Countries?

To answer it, the sub research questions must first be answered. Those were as follows:

SQ1) What is the effect of **income** on subjective health?

There is a positive effect of income on subjective health. As the income gap increases, so does the gap in health. People from the fifth income decile are for example 1.42 times as likely to report "very good" health than those from the first income decile. This difference has risen to 1.81 times more likely for the eight decile and even to 2.38 for the tenth. Thus, this effect seems to be linear.

SQ2) What is the effect of **education** on subjective health?

There is also a positive effect of education on subjective health. Compared to all other education levels, people from education level I are more likely to rate their health as "bad" or "very bad". The likelihood of someone rating their health as "bad" stays relatively stable, whereas the likelihood of someone rating it as "very bad" goes down fast once education level goes up. Once the gap between education levels is big enough, there is also a difference in the likelihood of rating their health as "good" or "very good". There is a clearer trend here. The higher the education level of comparison, the higher the odds ratio. Thus, there is a positive effect of education, but this does not show as clear a linear trend.

SQ3) What is the effect of **immigrant status** on subjective health?

Based on the results, none of the health categories are significantly more or less likely to be reported by immigrants than by non-immigrants. Thus, there is no significant effect of immigrant status on subjective health.

SQ4) What is the effect of **unemployment** on subjective health?

People who followed education, did housework, or performed paid work seem to have a better health than unemployed people. People in the 'other' category perform worse, however, and those in retirement do not differ. Therefore, the effect of unemployment on subjective health depends on which category it is being compared to.

As for country of origin, the exploratory research did indicate some trends. In the Baltic countries and Central and Eastern Europe, people are less likely to rate their health positively. In most of Northern Europe people are more likely to rate their health positively. Both in Southern and Western Europe, results are more mixed.

To answer the central research question, income and education have a positive relation with subjective health, immigrant status has no significant effect, and unemployment has a negative relation when compared to education, housework, and paid work, but no significant relation compared to retirement.

5.2 Discussion

In the discussion, the results of this research are compared to previous research. Limitations and suggestions for future research are also mentioned.

5.2.1 Previous Research

When looking at income, this outcome is in line with previous research. There is a clear increase in the likelihood of reporting positive health outcomes as the income deciles go up. This positive relation between income and subjective health is found in China (Pei & Rodriguez, 2006), the US (Lillard et al., 2015) and Colombia (Hessel et al., 2018). The moderating effects found by Ecob & Smith (1999) do not match with this outcome. They found incidences of morbidity to be linearly related to income, except for very high and very low incomes. This could be because low-income households get government support, and because people with high incomes generally live unhealthier lives (Gage, 2006). In this research, however, the negative ratings of health in the 10^{th} decile do not significantly differ from those in the 7^{th}, 8^{th} or 9^{th}. The positive ratings of health are significantly higher. The difference could be in the comparison between morbidity and subjective health. Although those two do show a strong relation, subjective health is much more comprehensive. It for example also includes mental health. Funk et al. (2012) found a strong link between poverty and mental disorders. This could be a reason people of the first income decile rate their health so low. Subjective health also includes smaller ailments such as a bad back. People with a high income are better able to treat those ailments, or to invest in preventive care. Thus, when it comes to subjective general health, it does seem likely that there is a linear positive relation between income and health.

For education, a similar positive relation is found. This is in line with, among others, Brunello et al. (2015) and Jutz (2020). Hu et al. (2016) found that in 17 European countries a less than good self-assessed health was more prevalent in lower education groups. This is in line with the outcome of this

research. When comparing the odds ratios for income and education, there seems to be a similar gap between the lowest level and the other levels. For example, a person following higher tertiary education over master level is 1.99 times more likely to rate their health as very good. A person in the 10th income decile instead of the first, is 2.38 times more likely to rate it that way. This effect on health could be because a higher education leads to better working conditions. Those with a lower education are more likely to have physically demanding jobs, and to get less mental stimulation (Yildirim, 2016). A higher education could also lead to more attention to health prevention and knowledge of a healthy lifestyle (Jutz, 2020). Dickson-Spillmann & Siegrist (2010) found that knowledge of how to consume a healthy diet plays a very important role in dietary behaviour. Especially lower education sub-groups in the population hold misconceptions about how healthy, or unhealthy, certain foods are. Thus, it seems indeed likely that the relation on education found in this research holds true.

When it comes to immigrant status, the results are more surprising. Malmusi (2015) found that there are significant differences in subjective health between immigrants and natives. They adjust for age, gender, education, occupation, and socio-economic conditions. Within these conditions they include household income, material deprivation, ability to make ends meet, and living in an overcrowded household. The difference could come from the fact that Malmusi (2015) makes a distinction between three different types of countries and their views on immigration. They also only included immigrants from outside of the European Union that have lived in the new country of residence for less than ten years. This could explain the difference between the results. In the European Social Survey, the question "Are you born in this country" is asked. This includes immigrants from within the EU, and also those who have lived there for longer than ten years. Their health might be closer to that of natives in a country. Domnich et al. (2013) compared subjective health within Italy. They found the immigrant population to be strongly heterogeneous and found significant differences depending on the human development index of the country where immigrants came from. The reason for immigration also played a significant role. Immigration for work was associated with lower scores of physical health, whereas immigration for religious and family reasons displayed a lower probability of lower scores of mental health. The way the data is offered in ESS does not accurately reflect this heterogeneity, which could explain why the variable was not significant in this research.

As for unemployment, previous research found a clear negative relationship with health, which is at least partly in line with the results of this research (e.g., Nordtström et al , 2014; Ross & Mirowsky, 1995; Virtanen et al., 2002). In previous research, unemployment has often been compared only to full-time employment (Nordtström et al, 2014; Ross & Mirowsky, 1995) or the focus was on employment security (Virtanen et al., 2002). Thus, the two options were either having paid work, or being unemployed. By using the variable that considers multiple categories of what you did most last week, more comparisons can be made. This way, it can be seen that people performing paid work do indeed report a better health, but that it also the case for those performing housework and those following

education. When unemployment is compared to retired people there are no significant differences. However, when compared to the 'other' category, unemployed people report a significant better health. This is difficult because it is unclear what respondents view as 'other'. However, an explanation could be that people who are disabled or chronically ill report in this category. The other options are not applicable, since unemployed would imply you are looking for a job. Those people are very likely to report a bad health, which is why it would be significantly lower than unemployed people. There are also very wide confidence intervals for the 'other' category, which would imply that the estimate is not very precise. This could be the case because so many different types of people fall under this category. Disabled people, but also people who do community service. It might be interesting to further specify what would be included in the 'other' category.

Lastly, the country of residence. Stirbu et al. (2010) and Di Girolamo et al. (2020) considered mortality, and both found that it was highest in Central and Eastern Europe, and Baltic countries. This was followed by Northern and Western regions, and smallest in Southern European regions. This is at least partly in line with the results from this research, although results were mixed in Western and Southern European regions. The difference could originate from the difference between mortality and subjective health. If previous research on subjective health is considered, as done by Olsen & Dahl (2007), Nordic countries score high. This is in line with the findings of this research. They also found that Austria and Switzerland displayed high levels of subjective health, and Portugal lower. Although this research found no significant difference between the Netherlands and Austria, people from Switzerland were 1.29 times more likely to rate their health as good and even 3.59 times more likely to rate it as very good. In Portugal, people are indeed 0.48 times as likely to rate their health as good and 0.65 times as likely to rate it as very good. Thus, although Olsen & Dahl (2007) only focused on adolescents, their findings seem to be in line with those of this research.

5.2.2 Limitations

As for the limitations of this research, most of them are based on quality of the available data. First, there were quite a lot of missing values for income. 9654 values were missing, which was 19.5% of all answers. The second highest variable with missing values was age, for which only 0.4% was missing. This shows the sensitivity of the subject. People could be unsure of their household income, or they could be ashamed. Tourangeau and Yan (2007) found that intentional non-response is most evident when the data is personally sensitive in nature. This way, a bias might be present. If a very low income is systematically not reported, for example, the analysis misses that data and is skewed in favor of higher incomes. For example, people report income in the first income decile slightly less than in the other income deciles (8.7%). This is not a drastic difference with other income deciles however, so it is not expected to cause a big bias, but it should be mentioned.

This shame might be enforced by the survey set up. These were face-to-face interviews. Although this does ensure that respondents do not miss any questions and that the survey is filled in the right way, it also brings along issues. According to Booth-Kewley et al. (2007) and Knapp and Kirk (2003) electronic survey formats bring along more truthful self-reports and lower non-responses to items. By being interviewed in person, they might feel the need to give socially desirable answers (Booth-Kewley et al., 2007). This can also be influenced by cultural expectations, and may thus differ between countries (Kays et al., 2012). Subconsciously, participants might want to make a good impression on the interviewer and are thus less likely to claim a lower income.

Apart from income, age or unemployment may also be such a sensitive issue. Age, however, is calculated based on the year of birth. That may be less sensitive information to provide than age. As for unemployment, table 3 shows that 6.0% of the respondents were unemployed for the last week. According to Eurostat (2018), unemployment in October 2018 was 6.7% in the EU 28. Since the difference between the response rate and the actual unemployment rate is not that high, it can be assumed that the effect of the sensitivity of the issue is not that big. Indeed, Kays et al. (2012) found that when it comes to sensitive topics, survey format only has a moderate effect. This could be because people do not only want to impress the interviewer, but also themselves. This sensitive information is even hard to disclose to oneself. It could also be the case that respondents simply do not know the answer. If they are not the primary breadwinner or not occupied with the financials of the household, they might not know how high the income is. Since the main goal of the ESS is to standardize surveys across countries, it makes sense to have an interviewer in the room to oversee the process. Although the issue of socially desirable answers arises, the pros of the face-to-face interviews seem to outweigh the cons.

Secondly, a Western dominance can be found in the answers. People from Germany, France and the UK make up almost half of all respondents. Using the weighting for ESS accounts for differences between countries in survey design, differential selection probabilities, nonresponses, and population sizes. Thus, comparisons between countries can still be made. However, since the multinomial logit model was run on individuals that mainly came from Western Europe, the relationship between health and the socioeconomic variables that was described might mostly be in Western Europe. This is an issue, since there is already a strong focus on this region in research (Jutz, 2020). This is especially worrisome since life expectancy and mortality continue to be better in Western European countries than in Eastern European countries (Forster et al., 2018). The effect of health inequalities also differs depending on the country (Mackenbach, 2006). In Central and Eastern Europe, education based health inequality appears to be much higher than in Western Europe (Jutz, 2020). This difference is not there for income based health inequalities. Even within these regions, there are big differences in size and determinants of health inequalities between countries (Jutz, 2020). Unfortunately, this research was limited due to the data that is available.

5.2.3 Future Research

This research attempted to create a more cohesive view of socioeconomic variables and their impact on subjective health. In this section, future research is proposed to continue this line of research. First, it would be interesting to study the effect of COVID-19 on this relationship between health and socioeconomics. The data comes from 2018, since ESS did not publish any data from 2019 to 2021. Not only did COVID-19 worsen some socioeconomic inequalities, but it also hit different socioeconomic groups in diverse ways. Wachtler et al. (2020) performed a literature review on this topic and found that for both risk of infection and for severity of the disease, the socioeconomically less privileged populations were hit harder. It would be interesting to use the same model and see if the relation between certain socioeconomic variables and subjective health has changed.

Secondly, a bigger focus on other European regions apart from Western Europe is recommended. This exploratory started examining subjective health between countries. A good addition to this would be to further specify what socioeconomic variables are the main cause of health inequalities in different regions, or even in specific countries. If that is done, policy can be made much more specific to address these issues. To do that, the underlying societal/political/economic/historical causes of health inequalities need to be researched as well. If that is done, they can be addressed in policymaking. By knowing what socioeconomic determinants are related to a bigger likelihood of illness, it might help for more targeted public health interventions as well (Forster, 2018). If income is, for example, a big cause of health inequalities, a subsidized public health care system might offer a solution. If all socioeconomic variables play a part, it could also be interesting to further dissect subjective health. Right now, participants are asked about their general health. This provides a comprehensive picture of health, but it might also provide interesting insights to further specify health into the categories mental and physical health, for example.

Health inequalities are avoidable. The end goal is not to undo socioeconomic disparities, since that is impossible. The aim is to find out how and why those disparities lead to health inequalities. People should not need an enormous income or a master's level in education to feel as healthy as all others. Gathering more knowledge on this topic is crucial to make policies against health inequalities more effect. Who knows, maybe someday we can bring back the saying that sickness hits the rich and poor alike

Bibliography

Acheson, D. (2001). Independent inquiry into inequalities in health report.

Adler, N. E., & Newman, K. (2002). Socioeconomic disparities in health: pathways and policies. Health affairs, 21(2), 60-76.

Alin, A. (2010). Multicollinearity. Wiley Interdisciplinary Reviews: Computational Statistics, 2(3), 370-374.

Arendt, J. N. (2005). Does education cause better health? A panel data analysis using school reforms for identification. Economics of Education review, 24(2), 149-160.

Bambra, C., Riordan, R., Ford, J., & Matthews, F. (2020). The COVID-19 pandemic and health inequalities. J Epidemiol Community Health, 74(11), 964-968.

Bambra, C. (2011). Health inequalities and welfare state regimes: theoretical insights on a public health 'puzzle'. Journal of Epidemiology & Community Health, 65(9), 740-745.

Bertocchi, G., & Bozzano, M. (2020). Gender gaps in education (pp. 1-31). Springer International Publishing.

Booth-Kewley, S., Larson, G., Miyoshi, D., . Social desirability effects on computerized and paper-and-pencil questionnaires. Computers in Human Behavior. 23, 463477. doi: 10.1016j.chb.2004.10.020 2005.

Britton, M., Fox, A., Goldblatt, P., Jones, D. and Rosato, M. (1990) The influences of socio-economic and environmental factors on geographic variation in mortality. In Britton, M. (ed) Mortality and Geography: a Review in the mid-1980s England and Wales. London: HMSO, 57–78

Brønnum-Hansen, H., & Baadsgaard, M. (2007). Increasing social inequality in life expectancy in Denmark. European Journal of Public Health, 17(6), 585-586.

Brunello, G., Fort, M., Schneeweis, N., & Winter-Ebmer, R. (2016). The causal effect of education on health: What is the role of health behaviors?. Health economics, 25(3), 314-336.

Centraal Bureau voor de Statistiek (n.d.). Gezondheidsenquete vanaf 2014. Retrieved from https://www.cbs.nl/nl-nl/onze-diensten/methoden/onderzoeksomschrijvingen/korte-onderzoeksbeschrijvingen/gezondheidsenquete-vanaf-2014

Charles, K. K., & Hurst, E. (2003). The correlation of wealth across generations. Journal of political Economy, 111(6), 1155-1182.

Chitkara, D. K., Talley, N. J., Schleck, C., Zinsmeister, A. R., Shah, N. D., & Locke III, G. R. (2009). Recollection of childhood abdominal pain in adults with functional gastrointestinal disorders. Scandinavian journal of gastroenterology, 44(3), 301-307.

Cox, D. T., Shanahan, D. F., Hudson, H. L., Fuller, R. A., & Gaston, K. J. (2018). The impact of urbanisation on nature dose and the implications for human health. Landscape and Urban Planning, 179, 72-80.

Craney, T. A., & Surles, J. G. (2002). Model-dependent variance inflation factor cutoff values. Quality engineering, 14(3), 391-403.

Dahl, E. (1994). Social inequalities in ill-health: the significance of occupational status, education and income-results from a Norwegian survey. *Sociology of Health & Illness*, *16*(5), 644-667.

Di Girolamo, C., Nusselder, W. J., Bopp, M., Brønnum-Hansen, H., Costa, G., Kovács, K., ... & Mackenbach, J. P. (2020). Progress in reducing inequalities in cardiovascular disease mortality in Europe. Heart, 106(1), 40-49.

Dickson-Spillmann, M., & Siegrist, M. (2011). Consumers' knowledge of healthy diets and its correlation with dietary behaviour. Journal of Human Nutrition and Dietetics, 24(1), 54-60.

Ecob, R., & Smith, G. D. (1999). Income and health: what is the nature of the relationship?. Social science & medicine, 48(5), 693-705.

Eder, A., Volk, H., & Haller, M. (2017). Patterns of Social Ties, Trust, and Participation after the Fall of the Iron Curtain: New Findings from Central and Southeast European Countries. International Journal of Sociology, 47(3), 147-161.

Ekgolm, I. (1971). Higher Education in the Soviet Union. Higher Education in a Changing World: The World Year Book of Education, 72, 294.

European Social Survey European Research Infrastructure (ESS ERIC). (2021). ESS9 Data Documentation. Sikt - Norwegian Agency for Shared Services in Education and Research. https://doi.org/10.21338/NSD-ESS9-2018

Eurostat (2018). Newsrelease Euroindicators. *Eurostat*. Retrieved from https://ec.europa.eu/eurostat/documents/2995521/9422707/3-30112018-BP-EN.pdf/fb7929e7-5e59-4a28-ba16-46a0beb188dc

EU SILC (2021). Health variables in SILC – methodology. Retrieved from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Health_variables_in_SILC_-_methodology

European Social Survey (n.d.). Data portal summary. Retrieved from https://ess-search.nsd.no/en/study/bdc7c350-1029-4cb3-9d5e-53f668b8fa74

European Social Survey (2016). Appendix A1 Education. *European Social Survey.* Retrieved from https://www.europeansocialsurvey.org/docs/round8/survey/ESS8_appendix_a1_e02_2.pdf

European Social Survey (2018). Appendix A2 Income. *European Social Survey*. Retrieved from https://www.europeansocialsurvey.org/docs/round9/survey/ESS9_appendix_a2_e03_0.pdf

Fang, C., & Wang, J. (2013). A theoretical analysis of interactive coercing effects between urbanization and eco-environment. Chinese Geographical Science, 23(2), 147-162.

Funk, M., Drew, N., & Knapp, M. (2012). Mental health, poverty and development. Journal of public mental health.

Forster, T., Kentikelenis, A., & Bambra, C. (2018). Health inequalities in Europe: setting the stage for progressive policy action.

Franks, P., Gold, M. R., & Fiscella, K. (2003). Sociodemographics, self-rated health, and mortality in the US. Social science & medicine, 56(12), 2505-2514.

Gage, R. (2006). Why You're Dumb, Sick and Broke... And How to Get Smart, Healthy and Rich!. Internet Profit Kit.

Guillot, M., Khlat, M., & Wallace, M. (2019). Adult mortality among second-generation immigrants in France: Results from a nationally representative record linkage study. Demographic research, 40, 1603.

Haas, S. (2008). Trajectories of functional health: The 'long arm'of childhood health and socioeconomic factors. Social science & medicine, 66(4), 849-861.

Haley, M. R. (2016). A ranking of journals for the aspiring health economist. Applied Economics, 48(18), 1710-1718.

Heine, S. J., Lehman, D. R., Peng, K., & Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert scales?: The reference-group effect. Journal of Personality and Social Psychology, 82(6), 903–918. https://doi.org/10.1037/0022-3514.82.6.903

Hessel, P., Avendano, M., Rodríguez-Castelán, C., & Pfutze, T. (2018). Social pension income associated with small improvements in subjective health of poor older men in Colombia. Health Affairs, 37(3), 456-463.

HSE (2021). Health and safety at work: Summary statistics for Great Britain 2021. Retrieved from https://www.hse.gov.uk/statistics/overall/hssh2021.pdf

Huang, K. S., & Gale, F. (2009). Food demand in China: income, quality, and nutrient effects. China Agricultural Economic Review.

Hu, Y., van Lenthe, F. J., Borsboom, G. J., Looman, C. W., Bopp, M., Burström, B., ... & Mackenbach, J. P. (2016). Trends in socioeconomic inequalities in self-assessed health in 17 European countries between 1990 and 2010. J Epidemiol Community Health, 70(7), 644-652.

Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: a review of twenty-seven community studies. Journal of health and social behavior, 21-37.

Illsley, R., & Baker, D. (1991). Contextual variations in the meaning of health inequality. Social science & medicine, 32(4), 359-365.

Jagger, C., Gillies, C., Moscone, F., Cambois, E., Van Oyen, H., Nusselder, W., & Robine, J. M. (2008). Inequalities in healthy life years in the 25 countries of the European Union in 2005: a cross-national meta-regression analysis. The Lancet, 372(9656), 2124-2131.

Jürges, H. (2010). Health inequalities by education, income and wealth: a comparison of 11 European countries and the US. Applied Economics Letters, 17(1), 87-91.

Jutz, R. (2020). Health inequalities in Eastern Europe. Does the role of the welfare regime differ from Western Europe?. *Social Science & Medicine*, *267*, 113357.

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.

Knapp, H., & Kirk, S. A. (2003). Using pencil and paper, Internet and touch-tone phones for self-administered surveys: does methodology matter?. Computers in Human Behavior, 19(1), 117-134.

Kauff, M., Wölfer, R., & Hewstone, M. R. C. (2017). Impact of discrimination on health among adolescent immigrant minorities in Europe: The role of perceived discrimination by police and security personnel. Journal of Social Issues, 73(4).

Kays, K., Gathercoal, K., & Buhrow, W. (2012). Does survey format influence self-disclosure on sensitive question items?. Computers in Human Behavior, 28(1), 251-256.

Kingma, E. (2007). What is it to be healthy?. Analysis, 67(2), 128-133.

Lall, R., Campbell, M. J., Walters, S. J., Morgan, K., & MRC CFAS Co-operative. (2002). A review of ordinal regression models applied on health-related quality of life assessments. Statistical methods in medical research, 11(1), 49-67.

Leonardi, F. (2018). The definition of health: towards new perspectives. International Journal of Health Services, 48(4), 735-748.

Lillard, D. R., Burkhauser, R. V., Hahn, M. H., & Wilkins, R. (2015). Does early-life income inequality predict subjective health in later life? Evidence from the United States. Social science & medicine, 128, 347-355.

Lindström, M., Sundquist, J., & Östergren, P. O. (2001). Ethnic differences in self reported health in Malmö in southern Sweden. Journal of Epidemiology & Community Health, 55(2), 97-103.

Loopstra, R., Reeves, A., & Tarasuk, V. (2019). The rise of hunger among low-income households: an analysis of the risks of food insecurity between 2004 and 2016 in a population-based study of UK adults. J Epidemiol Community Health, 73(7), 668-673.

Lorant, V., Van Oyen, H., & Thomas, I. (2008). Contextual factors and immigrants' health status: double jeopardy. Health & place, 14(4), 678-692.

Mackenbach, J. P. (2006). Health inequalities: Europe in profile. Produced by COI for the Department of Health.

Mackenbach, J. P., Kunst, A. E., Cavelaars, A. E., Groenhof, F., Geurts, J. J., & EU Working Group on Socioeconomic Inequalities in Health. (1997). Socioeconomic inequalities in morbidity and mortality in western Europe. The lancet, 349(9066), 1655-1659.

Majer, I. M., Nusselder, W. J., Mackenbach, J. P., & Kunst, A. E. (2011). Socioeconomic inequalities in life and health expectancies around official retirement age in 10 Western-European countries. J Epidemiol Community Health, 65(11), 972-979.

Malmusi, D. (2015). Immigrants' health and health inequality by type of integration policies in European countries. The European Journal of Public Health, 25(2), 293-299.

Manski, C. F. (1992). Income and higher education. Focus, 14(3), 14-19.

Mansyur, C. L., Amick III, B. C., Franzini, L., & Roberts, R. E. (2009). Culture and the social context of health inequalities. International Journal of Health Services, 39(1), 85-106.

Margaretten, M., Barton, J., Julian, L., Katz, P., Trupin, L., Tonner, C., ... & Yelin, E. (2011). Socioeconomic determinants of disability and depression in patients with rheumatoid arthritis. Arthritis care & research, 63(2), 240-246.

Marmot, M. (2005). Social determinants of health inequalities. *The lancet*, *365*(9464), 1099-1104.

Mishra, V., Seyedzenouzi, G., Almohtadi, A., Chowdhury, T., Khashkhusha, A., Axiaq, A., ... & Harky, A. (2021). Health inequalities during COVID-19 and their effects on morbidity and mortality. Journal of healthcare leadership, 13, 19.

Monden, C. W., Van Lenthe, F., De Graaf, N. D., & Kraaykamp, G. (2003). Partner's and own education: does who you live with matter for self-assessed health, smoking and excessive alcohol consumption?. Social science & medicine, 57(10), 1901-1912.

Munda, G. (2015). Beyond GDP: An overview of measurement issues in redefining 'wealth'. Journal of Economic Surveys, 29(3), 403-422.

Norström, F., Virtanen, P., Hammarström, A., Gustafsson, P. E., & Janlert, U. (2014). How does unemployment affect self-assessed health? A systematic review focusing on subgroup effects. BMC public health, 14(1), 1-13.

Offer, D., Kaiz, M., Howard, K. I., & Bennett, E. S. (2000). The altering of reported experiences. Journal of the American Academy of Child & Adolescent Psychiatry, 39(6), 735-742.

Olsen, K. M., & Dahl, S. Å. (2007). Health differences between European countries. Social science & medicine, 64(8), 1665-1678.

Pei, X., & Rodriguez, E. (2006). Provincial income inequality and subjective health status in China during 1991–7. Journal of Epidemiology & Community Health, 60(12), 1065-1069.

Pinquart, M. (2001). Correlates of subjective health in older adults: A meta-analysis. Psychology and Aging, 16(3), 414–426. https://doi.org/10.1037/0882-7974.16.3.414

Ravesteijn, B., van Kippersluis, H., & van Doorslaer, E. (2013). The contribution of occupation to health inequality. In Health and inequality. Emerald Group Publishing Limited.

Ross, C. E., & Mirowsky, J. (1995). Does employment affect health?. Journal of Health and social Behavior, 230-243.

Ross, C. E., & Wu, C. L. (1995). The links between education and health. American sociological review, 719-745.

Sarti, S., & Zella, S. (2016). Changes in the labour market and health inequalities during the years of the recent economic downturn in Italy. Social science research, 57, 116-132.

Shedler, J., Mayman, M., & Manis, M. (1993). The illusion of mental health. American psychologist, 48(11), 1117.

Schneider, G., Driesch, G., Kruse, A., Wachter, M., Nehen, H. G., & Heuft, G. (2004). What influences self-perception of health in the elderly? The role of objective health condition, subjective well-being and sense of coherence. Archives of gerontology and geriatrics, 39(3), 227-237.

Schreiber-Gregory, D., Jackson, H. M., & Bader, K. (2018). Logistic and linear regression assumptions: Violation recognition and control. Henry M Jackson Foundation.

Sede, P. I., & Ohemeng, W. (2015). Socio-economic determinants of life expectancy in Nigeria (1980–2011). Health economics review, 5(1), 1-11.

Senaviratna, N. A. M. R., & Cooray, T. M. J. A. (2019). Diagnosing multicollinearity of logistic regression model. Asian Journal of Probability and Statistics, 5(2), 1-9.

Shavers, V. L. (2007). Measurement of socioeconomic status in health disparities research. Journal of the national medical association, 99(9), 1013.

Shmotkin, D. (1990). Subjective well-being as a function of age and gender: A multivariate look for differentiated trends. Social indicators research, 23(3), 201-230.

Stafford, M., & McCarthy, M. (2006). Neighbourhoods, housing, and health. Social determinants of health, 2, 297-317.

Stirbu, I., Kunst, A. E., Bopp, M., Leinsalu, M., Regidor, E., Esnaola, S., ... & Mackenbach, J. P. (2010). Educational inequalities in avoidable mortality in Europe. Journal of Epidemiology & Community Health, 64(10), 913-920.

Stronks, K., Van De Mheen, H., Van den Bos, J., & Mackenbach, J. P. (1997). The interrelationship between income, health and employment status. *International journal of epidemiology*, *26*(3), 592-600.

Sun, S., Chen, J., Johannesson, M., Kind, P., & Burström, K. (2016). Subjective well-being and its association with subjective health status, age, sex, region, and socio-economic characteristics in a Chinese population study. Journal of Happiness Studies, 17(2), 833-873.

Titma, M., Tuma, N. B., & Roosma, K. (2003). Education as a factor in intergenerational mobility in Soviet society. European Sociological Review, 19(3), 281-297.

Tobiasz-Adamczyk, B., & Zawisza, K. (2017). Urban-rural differences in social capital in relation to self-rated health and subjective wellbeing in older residents of six regions in Poland. Annals of Agricultural and Environmental Medicine, 24(2).

Torsheim, T., Ravens-Sieberer, U., Hetland, J., Välimaa, R., Danielson, M., & Overpeck, M. (2006). Cross-national variation of gender differences in adolescent subjective health in Europe and North America. Social science & medicine, 62(4), 815-827.

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. Psychological bulletin, 133(5), 859.

UNESCO Institute for Statistics. (2012). International standard classification of education: ISCED 2011. Int. Stand. Classif. Educ.(ISCED) 2011.

Van der Noordt, M., IJzelenberg, H., Droomers, M., & Proper, K. I. (2014). Health effects of employment: a systematic review of prospective studies. Occupational and environmental medicine, 71(10), 730-736.

Veugelers, P. J., Yip, A. M., & Kephart, G. (2001). Proximate and contextual socioeconomic determinants of mortality: multilevel approaches in a setting with universal health care coverage. American Journal of Epidemiology, 154(8), 725-732.

Virtanen, P., Vahtera, J., Kivimäki, M., Pentti, J., & Ferrie, J. (2002). Employment security and health. Journal of Epidemiology & Community Health, 56(8), 569-574.

Wachtler, B., Michalski, N., Nowossadeck, E., Diercke, M., Wahrendorf, M., Santos-Hövener, C., ... & Hoebel, J. (2020). Socioeconomic inequalities and COVID-19–A review of the current international literature. Journal of Health Monitoring, 5(Suppl 7), 3.

Wamani, H., Tylleskär, T., Åstrøm, A. N., Tumwine, J. K., & Peterson, S. (2004). Mothers' education but not fathers' education, household assets or land ownership is the best predictor of child health inequalities in rural Uganda. International journal for equity in health, 3(1), 1-8.

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. MIS quarterly, xiii-xxiii.

Wild, S., & McKeigue, P. (1997). Cross sectional analysis of mortality by country of birth in England and Wales, 1970-92. Bmj, 314(7082), 705.

Wilkinson, R. G. (1997). Socioeconomic determinants of health: Health inequalities: relative or absolute material standards?. Bmj, 314(7080), 591.

World Health Organization (1946). Preamble to the Constitution of World Health Organization as adopted by the International Health Conference. New York, NY: World Health Organization.

World Health Organization (2013). Social determinants of health: Key concepts. Retrieved from https://www.who.int/news-room/questions-and-answers/item/social-determinants-of-health-key-concepts

World Health Organization (n.d.). World Health Survey Plus. Retrieved from https://www.who.int/data/data-collection-tools/world-health-survey-plus

Wu, S., Wang, R., Zhao, Y., Ma, X., Wu, M., Yan, X., & He, J. (2013). The relationship between self-rated health and objective health status: a population-based study. BMC public health, 13(1), 1-9.

Yen, I. H., & Kaplan, G. A. (1999). Neighborhood social environment and risk of death: multilevel evidence from the Alameda County Study. American journal of epidemiology, 149(10), 898-907.

Yildirim, Y. (2016). Healthier and more educated society improves multifactor productivity: time varying relationships.

Appendix A

European regions and their countries

Baltic Countries	Central & Eastern Europe	Northern Europe	Southern Europe	Western Europe
Estonia	Bulgaria	Denmark	Greece	Belgium
Latvia	Czechia	Finland	Spain	Germany
Lithuania	Hungary	Sweden	France	Ireland
	Austria	Iceland	Croatia	Luxembourg
	Poland	Norway	Italy	Netherlands
			Cyprus	Switzerland
	Slovakia		Portugal	United Kingdom
	Serbia		Slovenia	France
	Slovenia		Albania	
			Turkey	
			Kosovo	
			Malta	
			Montenegro	

Appendix B

Odds ratios Countries

	Subjective general health	Odds Ratio	Confidence Interval	
Austria vs. Netherlands	Very bad	1.45	.35	6.06
	Bad	1.24	.84	1.84
	Good	.88	.70	1.09
	Very good	2.52	1.97	3.22
Belgium vs. Netherlands	Very bad	1.04	.19	5.65
	Bad	1.07	.72	1.58
	Good	1.23	.99	1.52
	Very good	1.35	1.05	1.72
Bulgaria vs. Netherlands	Very bad	2.87	.78	10.51
	Bad	1.49	1.04	2.13
	Good	.84	.68	1.04
	Very good	1.10	.85	1.43
Switzerland vs. Netherlands	Very bad	1.37	.33	5.64
	Bad	1.21	.81	1.80
	Good	1.29	1.03	1.63
	Very good	3.59	2.80	4.62
Cyprus vs. Netherlands	Very bad	1.29	.28	5.88
	Bad	1.02	.66	1.58
	Good	.78	.56	1.09
	Very good	4.20	3.05	5.77
Czechia vs. Netherlands	Very bad	1.18	.29	4.87
	Bad	1.47	1.01	2.13
	Good	.74	.60	.91
	Very good	1.17	.92	1.49
Germany vs. Netherlands	Very bad	1.60	.41	6.22
	Bad	1.39	.96	2.03

	Good	.63	.52	.78
	Very good	.75	.60	.94
Denmark vs. Netherlands	Very bad	2.18	.55	8.72
	Bad	1.25	.81	1.92
	Good	.97	.75	1.26
	Very good	2.71	2.06	3.56
Estonia vs. Netherlands	Very bad	1.22	.31	4.72
	Bad	1.32	.93	1.87
	Good	.48	.40	.59
	Very good	.43	.34	.54
Spain vs. Netherlands	Very bad	1.51	.40	5.67
	Bad	1.44	.99	2.08
	Good	.76	.61	.93
	Very good	1.00	.78	1.29
Finland vs. Netherlands	Very bad	.67	.17	2.56
	Bad	.86	.59	1.25
	Good	.79	.65	.96
	Very good	1.01	.80	1.28
France vs. Netherlands	Very bad	1.03	.24	4.49
	Bad	.94	.65	1.36
	Good	.68	.55	.85
	Very good	.83	.64	1.06
United Kingdom vs. Netherlands	Very bad	1.58	.41	6.08
	Bad	1.45	.98	2.15
	Good	.93	.75	1.15
	Very good	1.99	1.56	2.53
Croatia vs. Netherlands	Very bad	1.84	.49	6.85
	Bad	2.20	1.51	3.22
	Good	.63	.48	.83
	Very good	2.54	1.93	3.35

Hungary vs. Netherlands	Very bad	2.04	.51	8.16
	Bad	1.15	.77	1.72
	Good	.76	.59	.97
	Very good	.93	.69	1.25
Ireland vs. Netherlands	Very bad	.91	.18	4.49
	Bad	1.04	.70	1.55
	Good	1.02	.80	1.30
	Very good	3.70	2.84	4.83
Iceland vs. Netherlands	Very bad	1.79	.29	10.90
	Bad	1.28	.78	2.09
	Good	.77	.59	1.00
	Very good	2.34	1.77	3.08
Italy vs. Netherlands	Very bad	1.92	.53	6.98
	Bad	1.00	.70	1.42
	Good	.87	.70	1.08
	Very good	1.38	1.06	1.80
Lithuania vs. Netherlands	Very bad	.47	.05	4.70
	Bad	.76	.49	1.17
	Good	.60	.48	.76
	Very good	.47	.35	.63
Latvia vs. Netherlands	Very bad	.42	.03	6.61
	Bad	.90	.59	1.38
	Good	.37	.28	.48
	Very good	.29	.21	.41
Montenegro vs. Netherlands	Very bad	1.48	.36	6.03
	Bad	1.88	1.20	2.95
	Good	.65	.49	.86
	Very good	2.54	1.91	3.38
Norway vs. Netherlands	Very bad	1.86	.35	9.90
	Bad	1.26	.82	1.91

	Good	.97	.77	1.23
	Very good	2.05	1.58	2.65
Poland vs. Netherlands	Very bad	1.78	.46	6.84
	Bad	1.48	1.01	2.16
	Good	.69	.55	.88
	Very good	1.15	.87	1.51
Portugal vs. Netherlands	Very bad	.57	.14	2.27
	Bad	1.05	.68	1.61
	Good	.48	.37	.60
	Very good	.65	.49	.87
Serbia vs. Netherlands	Very bad	4.42	1.24	15.80
	Bad	1.86	1.31	2.65
	Good	.50	.40	.63
	Very good	1.07	.82	1.38
Sweden vs. Netherlands	Very bad	1.19	.21	6.73
	Bad	1.31	.87	1.98
	Good	1.04	.82	1.31
	Very good	2.39	1.85	3.10
Slovenia vs. Netherlands	Very bad	1.52	.39	5.93
	Bad	1.38	.95	2.01
	Good	.73	.59	.91
	Very good	1.14	.88	1.46
Slovakia vs. Netherlands	Very bad	1.99	.43	9.23
	Bad	2.15	1.47	3.14
	Good	.79	.61	1.03
	Very good	1.22	.82	1.82