LUND UNIVERSITY

SCHOOL OF ECONOMICS AND MANAGEMENT

Health and Wealth: Understanding Regional Variations in Health, Income, Healthcare Costs, and Productivity

By

Tilda Mellberg
ti8515me-s@student.lu.se

Lund University School of Economics and Management

Bachelor Thesis in Economics (15 credits ECTS)

January 2025

Supervisor: Ulf Gerdtham

Abstract

While studies on the relationship between health, income, healthcare spending and productivity is generally consistent, little attention has been brought to the regional contexts. Meanwhile, a country like Sweden who is generally known for a well-established healthcare system and welfare state, experiences increasing regional disparities. This thesis investigates the relationship between health, income, healthcare spending and productivity in Swedish regions across 26 years. Utilizing panel data, the analysis employs several regression models and compares the results. The study finds notable differences between models who include and exclude control variables, where the results found rather counterintuitive relationships such as higher income is associated with higher mortality rates, and vice versa. One reason for this might be due to lifestyle factors in higher income regions, such as stress, which affect the mortality rates. The thesis concludes that it is fundamental to consider contextual factors when evaluating health, income, healthcare spending and productivity, in particular for policymakers and future research.

Keywords: health, income, healthcare spending, productivity, Sweden, regions

Table of Contents

List of Tables

1. Introduction

It may seem obvious to most people that improved health has positive economic and societal outcomes and that higher incomes are associated with better health and quality of life. In fact, research suggests that healthier populations tend to be more productive, with improved health contributing to lower healthcare spending, improved economic growth and better well-being (The Health Foundation, 2021).

However, the relationship between health, income, healthcare spending, and productivity is far more complex, particularly in high-income countries where country-specific factors create complicated relationships between these variables. Even less is understood when the regional context is considered, as limited attention has been given to studying these dynamics within regions of high-income countries. In Sweden, for instance, increasing regional disparities have been observed, in particular, driven by demographic changes where younger inhabitants move to larger cities (André et al., 2021). However, little is known about how these regional inequalities influence the dynamics between health, income, healthcare spending, and productivity, leaving a significant gap in our understanding of the regional context. This regional perspective is particularly relevant given that Swedish healthcare is governed at the regional level, with each region being responsible for funding and organizing healthcare.

As a result, studying these dynamics at the regional level allows for a more comprehensive understanding of how regional socioeconomic and demographic factors shape the dynamics between health, income, healthcare spending and regional productivity.

1.1. Aim & Research Question

As a result of the lack of understanding of the Swedish regional, there is a need for more research that studies this interplay. Thus, this study aims to investigate the relationships between health, income, healthcare spending, and regional productivity within the context of Swedish regions, particularly focusing on how these relationships are influenced by contextual factors such as regional differences and time-specific effects. The purpose of this thesis is to obtain a broader and more nuanced understanding of the intricate relationships to better support policymakers when implementing region-specific health policies. To help guide this research, the following research question is therefore proposed:

How do regional contexts and time-specific effects influence the relationships between health outcomes, income, healthcare spending, and regional productivity in Swedish regions?

Followed by the sub-question:

What are the implications of the regional contexts and time-specific effects for regional policy?

The research questions aim to examine how the relationship between income, healthcare spending, and regional productivity changes when confounding factors are considered and what it might mean for the broader setting and is based on well-established theories of health and economics, along with previous empirical research.

1.2. Scope

Analyzing the situation in Sweden regarding health outcomes, healthcare costs, income, and productivity makes for an insightful case study. Given that Sweden has a rather established welfare state and universal healthcare, one would expect that the access to healthcare and other services should be quite equal across the country. However, there are still significant regional disparities in health outcomes and economic performance, which means there is still a question of how these factors are connected at the regional level that needs to be answered.

This thesis considers contextual factors such as differences in healthcare spending, regional income levels, and demographic trends rather than individual or household level factors. This study tries to capture how structural and systemic factors affect the outcomes by focusing on these regional aspects.

The time period 1997–2023 is chosen both to enhance the sample size and to include all of the possible trends and changes in the data. These years cover more than two decades of change in Sweden's healthcare system and economic policies as well as its population. It is therefore possible to study how these factors have evolved and interacted over time.

This study attempts to identify the connections between health, income, healthcare costs, and productivity across Swedish regions and in doing so, contribute to the understanding of the causes of regional differences. The findings help to understand how these relationships can be useful in formulating policies that can lead to more unbiased health outcomes and the reduction of regional inequalities.

1.3. Structure

The thesis begins with the Background, introducing basic information and statistics about the current situation in Sweden regarding health, healthcare, income and productivity. The two subsequent sections—Theoretical Framework and Literature Review—present the theoretical model and previous empirical research utilized to support the thesis's findings. Thereafter, the Methodology is introduced, including the usage of variables and model specification. In the following section, called Result, the obtained findings are presented and thereafter interpreted in the Discussion section with theoretical and practical implications. Lastly, the thesis is tied together with some concluding remarks in the Conclusion section.

2. Background

The following section presents an overview of economics and health in Sweden. This includes healthcare utilization and the healthcare system, along with some statistics in the Swedish context. Moreover, it presents the economic standards for Swedish citizens and the current public health situation. The aim is to provide a general outlook of the present conditions in Sweden.

2.1. The Swedish Healthcare and Healthcare System

In Sweden, healthcare is accessible to all citizens and is essentially free of charge, except for a smaller fee paid per visit up until reaching a specific threshold per year (Socialstyrelsen, 2019). The healthcare system is regional, meaning that the 21 regional authorities are responsible for healthcare provision in each region (Nordic co-operation, 2025). About 80% of Swedish healthcare is provided by the regional authorities, and the remaining 20% is privately provided. However, 95% of the private healthcare providers have an agreement with the regional

healthcare providers, thus also essentially being free of charge for Swedish citizens (Statistiska Centralbyrån, 2022).

While Sweden is reported to be in the tenth place on the doctor-inhabitant ratio list and twelfth place on the nurse-inhabitant list amongst OECD countries, significant differences in access to and utilization of healthcare professionals have been reported across regions (Socialstyrelsen, 2024b). Region Kalmar, for instance, has twice as many nurses per thousand inhabitants as Region Stockholm (22.9 and 11, respectively). Similarly, Region Stockholm has about double the number of healthcare visits, 1900, compared to Region Västra Götaland, 900, per thousand inhabitants in 2021.

Moreover, a prevalent issue in Sweden is waiting times, which have been persistently long for several years (Socialstyrelsen, 2024c). Table 1 shows the reported mean days waited for the first context with healthcare in the Swedish regions during 2023. As seen, Region Jönköping has the lowest number of days, and Region Norrbotten has the highest number of days (highlighted in green and red, respectively). That is, in Region Norrbotten, patients had to, on average, wait 230 days between getting the decision that specialized care is needed and the first meeting with a specialized healthcare provider.

Table 1: Mean waiting times for first contact with healthcare in 2023

Region	Mean waiting time (days)	Region	Mean waiting time (days)
All regions	101	**Region Skåne**	101
Region Blekinge	87	**Region Sörmland**	57
Region Dalarna	93	**Region Stockholm**	98
Region Gävleborg	92	**Region Uppsala**	96
Region Gotland	84	**Region Värmland**	152
Region Halland	52	**Region Västerbotten**	86
Region Jämtland Härjedalen	137	**Region Västernorrland**	124

Region Jönköpings län	**49**	**Region Västmanland**	64
Region Kalmar län	71	**Region Örebro län**	105
Region Kronoberg	77	**Region Östergötland**	129
Region Norrbotten	**230**	**Västra Götalandsregionen**	114

Source: Sveriges Kommuner och Regioner, 2023b

2.2. Economics and Public Health of Sweden

Sweden generally has a good economic standard and public health, with a mean monthly salary of 39,900 SEK in 2023 (Statistiska Centralbyrån, 2024b). Many prerequisites for good health have improved but other prerequisites have stagnated or worsened and disparities between groups have grown (Folkhälsomyndigheten, 2023). For example, individuals born outside Europe and those with low socioeconomic status are more likely to experience lower economic standards, with long-term unemployment affecting these groups.

Due to variations in tax power across regions, that is, some regions are wealthier than others, Sweden has a tax equalization system to ensure that all residents in all regions and municipalities have equal economic standards (Sveriges Kommuner och Regioner, 2023a). This is done in order to guarantee the residents the same level of public services, such as healthcare and schools. Therefore, a part of wealthier regions' tax income is redistributed to less affluent regions and thus decreases inter-regional inequalities.

In addition, public health in Sweden has improved over time, with life expectancy increasing since 2006. The mean age in Sweden in 2021 was 83.0 years, however, significant disparities can be seen between groups, such as between men and women and between educational groups. Sweden spent about 4200EUR per capita on healthcare in 2021, which is slightly higher than the average for EU countries. Accessibility to healthcare is generally high, but differences in groups—especially between less and more remote regions—seem to prevail (OECD/European Observatory on Health Systems and Policies, 2023). Moreover, while mortality rates for common causes of death have decreased, significant inequalities persist. People with low socioeconomic status, for instance, have shorter life expectancies and higher mortality rates

before the age of 65. Data also reveals that public health trends show mixed progress. Positive developments include reduced smoking rates and lower alcohol consumption among certain groups. However, significant disparities persist (Folkhälsomyndigheten, 2023).

3. Theoretical Framework

The following section presents the theoretical framework utilized in the thesis. It introduces the Grossman Model and highlights the aspects relevant to the aim of the thesis. The aim is to get a theoretical understanding of health and economics, which later is used in the discussion.

3.1. The Grossman Model

Grossman's work on the human capital model of the demand for health was originally published in 1972 (Grossman, 1972) and has since then been revised in later works by Grossman (2000) and several scholars (see for example, Ehrlich & Chuma, 1990; Jacobson, 2000; Muurinen, 1982; Ried, 1998; Zweifel, 2012). This theoretical framework builds on the revised work from 2000 (Grossman, 2000) which is a comprehensive discussion of his original work, with some extensions. This version is chosen for this theoretical framework as it is the recent version and contains more comprehensible perspectives of his original model. Because this thesis focuses on health outcomes, income, healthcare spending, and productivity, these aspects of the Grossman model are focused on.

The Grossman model is derived from the human capital theory proposed by Becker (1964; 1967), Ben-Porath (1967), and Mincer (1974). Human capital theory suggests that increases in individuals' knowledge or human capital promote productivity in the market (where capital earnings are generated) as well as in the household sector (where goods are created and enter the utility function). Human capital theory focuses on the productivity a person generates through knowledge. Grossman emphasized instead that health establishes the aggregate time that can be spent on producing capital earnings and goods.

Grossman draws his work from the household function, where he includes the distinction between health as an output and healthcare as one of several inputs. The household production function is a model that depicts how households combine their time, resources, and goods to

produce things they value, such as health and knowledge. Of importance, he argues, is to distinguish between items of choice (which he calls commodities) and market goods and services, which all enter the utility function. Grossman notes that "Consumers produce commodities with inputs of market goods and services and their own time." (p.350). In other words, consumers use goods provided by the market (such as gym memberships) to produce commodities, such as enjoyment from exercising. Therefore, the demand for healthcare and other inputs to health is extracted from the demand for health.

In the model, health is both demanded and produced, and it is considered a choice variable because it provides satisfaction (demanded) and influences income or wealth (produced). Health is demanded because it is valuable for well-being, and it is produced because individuals take action to maintain or improve it. Specifically, people desire to improve their health because it reduces sick days, allowing them to avoid missing out on income due to illness. Moreover, Grossman argues that if the shadow price of health (the "true cost" of achieving better health) increases, the demand for health decreases. Meaning, that if attaining health becomes more difficult (e.g. through older age, more expensive healthcare) individuals are less willing to demand (i.e. achieve) health.

The full wealth model, including time and money constraints, is the following:

$$\sum_{t=0}^{n} \frac{P_t M_t + Q_t X_t + W_t(TL_t + TH_t + T_t)}{(1+r)^t} = \sum_{t=0}^{n} \frac{W_t \Omega}{(1+r)^t} + A_0 \qquad (1)$$

Where the left-hand side in equation (1) above represents total spending over a person's life, discounted to account for the value of money over time, and the right-hand side represents total lifetime wealth, discounted to present value. In its simplest form, the equation depicts that a person's total lifetime spending (on health, goods, and time costs) must equal their total lifetime resources (income from working all available time plus any savings). Adjusting the variables on either side, we can also detect what effects it may have on the other side. For example, if an individual spends less time being sick—meaning, is healthier—they have more time to work and increase their earnings. Vice versa, if an individual has higher total wealth, they can spend more time and resources on improving health.

Moreover, Grossman elaborates on how investment in health (in the form of healthcare) interacts in his model. Going back to Equation 1 above, a two-way relationship could be indicated. On the one hand, wealthier individuals may have more resources to spend on healthcare, such as cosmetic procedures or other costly preventative medical care, which in turn raises personal healthcare spending. On the other hand, if an individual has high total wealth, it could be argued that personal healthcare spending decreases, due to investments in other proactive factors that go into the health production function, such as improved diet, better living conditions, or similar. Perhaps most strongly emphasized in Grossman's work is the former argument. However, as is later discussed in the literature review (Section 4) the second argument may be more suitable in cases where healthcare is free - as in the case of Sweden, which is the focus of this study. Universal healthcare often focuses on prevention, providing free health screenings and vaccination programs to lower overall medical costs.

Grossman's model highlights that health is a dual commodity, meaning, both a consumption and investment good. The dual role of health means that individuals who invest in health (treating it as an investment good) experience higher productivity because they use resources (e.g., funding, labor) more efficiently. This can be interpreted as healthier individuals allocate resources more efficiently, leading to better productivity outcomes. In addition, the production function of healthy time shows that as health capital increases, healthy time also increases, albeit at a diminishing rate.

To summarize Grossman's model in alignment with the aim of this thesis, it is argued that individuals treat health as a form of capital that they can invest in to improve productivity and quality of life. Higher-income regions may have residents who can afford more investments in health (e.g., better nutrition, healthcare, education, or preventive care), leading to better health outcomes. Moreover, Grossman's model also suggests that wealthier individuals (who collectively contribute to the wealth of regions) may have better access to efficient health systems and preventive measures, reducing the need for more (and often more expensive) treatments. Lastly, it is shown that individuals (here, regions) with better health outcomes are likely to have higher stocks of health capital, resulting in more healthy time for productive activities.

4. Literature Review

This section presents previous empirical evidence on the relationships of interest. The relationship between health, income, healthcare spending, and productivity has been extensively studied and has generated a diverse body of literature. It is neither feasible nor practical to elaborate on every aspect of these relationships in detail. Instead, this section aims to provide an overview of the most widely recognized empirical studies, with a focus on the evidence most relevant to the research question.

4.1. Health, Income and Healthcare Spending

The relationship between health and income is broadly studied amongst scholars, where various perspectives have been undertaken. There is a general consensus amongst previous studies regarding the interplay between health and income in areas characterized by lower development and/or high-income inequality (Becker, 2007; Benzeval & Judge, 2001; Grossman, 2017; Institute For Research On Public Policy & Centre For The Study Of Living Standards, 2002; Rowntree & Orr, 1936). The overall agreement amongst researchers is that higher income results in better health outcomes. However, less agreement is found when the reverse relationship is considered, that is, if better health outcomes raise income.

Higher healthcare spending has furthermore often been linked to improved health outcomes, particularly in lower-income settings (Newhouse, 1977; Singh, 2014; Hall & Jones, 2007). However, in more developed regions, the link between healthcare spending and outcomes becomes more complex, as diminishing returns on healthcare spending may limit its impact. This lack of consensus highlights a gap in understanding how investments in health (healthcare spending) influence income levels in high-income contexts, where health differences may remain despite extensive healthcare spending. Thus, the following subsection evaluates what previous empirical studies have found regarding income, health, and healthcare.

Two well-known studies focusing on income and health are provided by Ecob and Davey Smith (1999), along with Ettner (1996). Both studies draw data from the mid-80s and focus on morbidity, but the former investigates England, Wales, and Scotland and the latter focuses on America. With some slight differences, both studies conclude that higher incomes result in better health outcomes. What both articles highlight is the focus on morbidity (the disease rate)

instead of mortality, which added significantly to the body of literature at the time. Their differences lie in the way they look at income. Ettner (1996) shows a clear and significant relationship between income and health, with coefficients around 0.11 (dependent variable: income, independent variable: health rating), however with attention put on the fact that higher income increases alcohol consumption. Ecob and Davey Smith (1999) argue that the relationship is true, but with diminishing returns at higher levels of income, concluding that income inequality is associated with worse health.

Lacking in both studies, albeit touched upon but not thoroughly discussed, is the reverse relationship between income and health. This is another crucial aspect to consider, as examining the relationship both ways provide a more comprehensive understanding of the interplay between health and income. The earlier works have found that the relationship between health and income is positive (Barro, 2013; Bloom, Kuhn & Prettner, 2019; Bloom & Williamson, 1998). However, this view has more recently been challenged.

On the one hand, Lorentzen, McMillan, and Wacziarg (2008) demonstrate that health has a positive impact on income, by looking at how decreasing adult mortality has a positive impact on income and thus leads to increasing returns on human and physical capital. This view would be following Grossman's model of human capital (1972, 2000). On the other hand, Acemoglu and Johnson (2007) argue that the influence of improved health is negative on economic development. While Lorentzen, McMillan, and Wacziarg (2008) have a more micro-perspective, Acemoglu and Johnson (2007) focus on the macro view. The negative economic development, the authors argue, happens through rapid mortality decline without fertility decline which may lead to population growth that reduces capital per worker, limiting income per capita (a "Malthusian response").

However, both studies aim at lower-income areas, with a specific focus on African countries, and fail to address higher-income countries and their regions. This limitation is not exclusive to these studies but reflects a general tendency in the literature, where research focuses on low- and middle-income countries. Consequently, there is a lack of understanding into how the health-income relationship appears in higher-income contexts, where the dynamics may differ due to established healthcare systems, higher living standards, and lower mortality rates.

Shifting the focus to healthcare expenditures, studies nowadays are quite consistent. It is widely agreed that increased healthcare spending is associated with better health outcomes. For example, a study from Canada looking at data from 1978-1992 (Crémieux et al. 1999) finds through a regression analysis statistically significant results where life expectancy is expected to rise (coefficient of 0.003 and 0.001 for males and females, respectively) with higher healthcare spending, whereas infant mortality has decreased. Another study by Martin et al. (2008) performed in England, obtaining data from 2004-2005, has theorized two healthcare programs and their possible effect on health outcomes and finds that healthcare spending has a strong positive effect on the health outputs from the programs. Similarly, a more recent study utilizing 2015 data from OECD countries also finds that healthcare spending positively impacts health outcomes, where maternal and infant mortality along with life expectancy are measured as health outcomes (Karaman et al., 2020).

As seen, there is a consensus on the relationship between healthcare spending and health outcomes. However, an important limitation in these studies is the use of life expectancy and infant mortality as proxies for health outcomes. While these measures may be relevant in lower-income areas where these variables show significant deviations over time, they are of less relevance in the early 2020s, particularly in higher-income settings where life expectancy and infant mortality have remained stable.

In these contexts, the more central issue lies in understanding intra-country variations in health outcomes. Therefore, there is a gap in the existing literature which this thesis aims to address by focusing on deaths caused by illnesses as a more appropriate measure of health outcomes in higher-income settings. This approach offers a more nuanced perspective, reflecting current health challenges that persist even in economically developed areas.

Moreover, it is also of relevance to introduce literature looking at medical care spending and income. One who does this well is Newhouse (1977) who tries to estimate the relationship between medical expenditures and income.

Newhouse finds that there is a direct positive relationship between income and healthcare spending in developed countries. He shows that as GDP per capita increases, so does medical care spending. Excluding Greece from the regression, the results exhibit a coefficient of 0.0763. Newhouse suggests that as the price of healthcare to the consumer goes towards zero which,

according to the author, is true for most developed countries, income elasticity falls. Thus, at higher levels of income, medical care becomes a necessity. Another argument Newhouse makes is that this relationship is only true inter-country or intra-country over a period of time. In the context of this thesis, it would be between regions over several years.

4.2. Health and Productivity

The health-productivity dynamic can be similar to the health-income relationship as it can be argued that it is a two-way relationship. However, the difference between the two is that health-productivity regards how health is related to how much output is produced, while health-income considers how much income is generated. Although related, it is important to distinguish them.

Health and productivity are extensively considered in the business context nowadays, as healthier workers are assumed to be more productive, thus yielding better business performance (Wolf, 2010; Loeppke et al., 2009). However, in the context of this thesis, it is important to take the macro-perspective into account.

In general, it is found that improved health increases productivity. A large-scale study by Mitchell and Bates (2011), involving over one million participants and examining health conditions and lifestyle factors, quantifies the economic impact of poor health. The study finds that productivity losses among unhealthy workers ranged from $15 to $1,601 per year. Additionally, the substantial medical costs associated with these workers further burden society, enhancing the overall economic impact of poor health. This indicates that poor health not only affects individual productivity but also has broader implications for economic growth at the macro level. The implication is further illustrated by Arora (2001) who has performed a long-term study on European countries, including Sweden, investigating health, productivity, and economic growth. With data from the late 1900s to the mid-1990s, the analysis finds that improvements in health account for an economic growth rate of 30%-40%. The author argues that the growth is sourced from improvements in productivity since “Productivity and pace of growth then increases because a healthier labor force is also more able.” (Arora, 2001, p. 733).

The existing literature on health and productivity provides valuable insights into the relationship, particularly in the context of businesses and individual-level productivity. Studies show the economic impact of poor health, both in terms of lost productivity and increased social costs, as well as the broader macroeconomic benefits of improved health. However, an important gap in the literature lies in the lack of research focusing on regional-level disparities in health and productivity within high-income countries, particularly in more recent years. Most studies either use a global or national perspective or focus on historical data from the 20th century.

4.3. The Case of Sweden

Some studies have been conducted focusing solely on health, income, healthcare spending, and productivity in Sweden, although they are somewhat scarce. One study conducted by Fritzell et al. (2004) looks at the income-health relationship (that is, how income affects health) which, as previously discussed, lacks consensus amongst scholars. The authors include self-reported health data from 1996 and 1997 and statistics on income and illustrate that income strongly relates to health, including when controlling for structural factors. They thus conclude that higher income is associated with a lower frequency of poor health.

Furthermore, the reverse relationship has been investigated in Sweden. A comparative study of Sweden and England finds that, although England is "worse" than Sweden, income inequality still plays a role in health disparities, meaning that individuals with higher incomes tend to report better health, while those with lower incomes tend to report worse health and that income inequality is a factor in how people report their health (Yngwe, 2001). Moreover, a more recent study looking at healthcare utilization and mortality related to income disparities between the years 2004-2017 finds that lower incomes are associated with both higher healthcare utilization and higher mortality rates (Flodin et al., 2023). The authors conclude that despite the prevalence of universal healthcare, income-related health inequalities persist.

Moreover, various studies have been conducted focusing on productivity losses, healthcare expenditure, and health in Sweden (Bolin, 2018; Bolin et al., 2009; Bolin & Lindgren, 2007; Bolin et al., 2011; Neovius et al., 2012; OECD, 2013). Although focusing on different conditions and lifestyle habits, such as diabetes, epilepsy, smoking, obesity, physical inactivity,

and psychological health, all studies conclude that worse health outcomes lead to productivity losses and higher healthcare spending. Except the scholars arguing that improved health leads to the "Malthusian response", the result in Sweden is consistent with the majority of previous general studies and theories.

Although smaller opposing views, the overall conclusion for health, income, productivity, and healthcare spending in Sweden is consistent, both within Sweden and across more general studies. Common for several of the studies looking at the income-health relationship is using self-reports as the measure of health. This may yield biased results, because individuals may not state the true case, report what is true for the particular day the report is filled in or be influenced by the nature of the question (Rosenmann et al, 2011). In consequence, the validity of the results decreases. Therefore, it may be possible to detect a gap in the literature, which this thesis aims to fill. Instead, this study uses an alternative measure of health outcomes, deaths attributable to illnesses, which is further motivated in Section 5.

The existing literature on health, income, healthcare spending, and productivity is extensive and diverse, providing valuable insights into these relationships. However, much of the research focuses on low- and middle-income countries or examines national or global perspectives, portraying a gap in understanding how these variables interact at the regional level within high-income countries. In the Swedish context, while some studies address health-income relationships, healthcare utilization, and productivity losses, they seem to often rely on self-reported data or do not explore the complex dynamics between these factors in modern settings.

Moreover, traditional measures like life expectancy and infant mortality, commonly used in health-related studies, are less relevant in high-income regions with less fluctuating health outcomes. Thus, there is a need for research that encompasses health metrics which invole health outcomes related to illness and lifestyles choices, such as alcohol consumption and heart conditions. This thesis therefore aims to fill this gap by analyzing the relationship between health, income, healthcare spending and productivity across Swedish regions using metrics more specific to high-income contexts utilizing data over a longer period of time, including from the last couple of years. Thus, it provides more detailed insights into the role of regional health and income setting using more relevant variables.

Henceforth, given the various theories, empirical research, and results that previously have been conducted, the following hypotheses naturally arise for the regional Swedish context:

$H_{1.1}$: *Higher income is associated with lower deaths attributable to illnesses per capita within Swedish regions*.

$H_{1.2}$: *Higher deaths attributable to illnesses per capita are associated with lower income within Swedish regions*.

H_2: *Higher income is associated with lower healthcare spending per capita within Swedish regions*.

H_3: *Higher healthcare spending per capita is associated with lower deaths attributable to illness per capita*.

H_4: *Higher deaths attributable to illness per capita are associated with lower regional productivity*.

5. Methodology

The following section describes the methodology employed for the study. It starts by elaborating on the choice of research design and thereafter depicts the data collection procedure followed. Subsequently, the variables and data are explored. Lastly, it illustrates the data analysis process.

5.1. Research Design

The chosen methodology of this research is based on the aim of the thesis. To recapitulate, the aim is to investigate the relationships between health, income, healthcare spending, and regional productivity within the context of Swedish regions, particularly focusing on how these relationships are influenced by contextual factors such as demographic regional differences and time-specific effects. Because the study aims to investigate relationships between variables—testing theories—there is a need to collect data and statistically analyze them to

draw viable conclusions (Creswell, 2003). Therefore, a quantitative approach is most appropriate.

The quantitative approach is based on a set of predetermined theories and literature. This thesis aims to expand the preexisting knowledge of the dynamics of health outcomes, income, healthcare spending, and productivity and apply it to the Swedish regional setting. The main research question evaluated is instrument-based, meaning, can be measured numerically. It is based on secondary, observational data. The analytical approach in this thesis is statistical and its results are based on established statistical interpretations. The sub-research question is answered based on the discussion of the main question. By doing this, the study obtains solid analytical results which assist in understanding the proposed research question and fulfill the aim of the study.

In consequence, hypotheses are needed to evaluate the dynamics of health, income, healthcare spending, and productivity. These hypotheses are developed based on the existing literature explored in Sections 3-4. In other words, this study employs a deductive approach (Wilson, 2014). This approach is deemed to be the most suitable for the thesis since it conveys the reasoning of transforming the specific to the general. Since existing theory has established that the aforementioned variables are associated with each other, this relationship may be said for the regional Swedish perspective as well, which has not yet been studied.

5.2. Data Collection

The thesis relies on secondary aggregate national data from Kolada (Kolada, 2024.), Statistiska Centralbyrån (Statistiska Centralbyrån, 2024a), and Socialstyrelsen (Socialstyrelsen, 2023). Kolada is managed by the Swedish Council for the Promotion of Municipal Analysis (RKA), a partnership between the Swedish government and the Swedish Association of Local Authorities and Regions (SKR). Statistiska Centralbyrån (SCB) is Sweden's official statistics agency, while Socialstyrelsen, the National Board of Health and Welfare, focuses on social services, health, and medical data collection. These sources provide reliable and comprehensive data for this study.

The dataset includes 20 Swedish regions, excluding Gotland due to insufficient data entries. While regional heterogeneity exists, the regional-specific contexts of Gotland is not substantially different from the other regions, thereby the exclusion of it should not significantly bias the results. The data spans 1997–2023, but not all years are included in the econometric analyses due to non-consistent data entries for the different regions. After retrieval, the data is compiled in Microsoft Excel and imported into Stata for analysis. Stata, an integrated statistical software, is chosen due to its wide variety of statistical tools and the author's familiarity with it (STATA, 2023). The data is structured as long-format panel data with 540 entries, allowing the study to account for both cross-sectional and time-series variations across regions (Wooldridge, 2012).

The sample size for each regression includes at least 360 observations, which meets the reliability criteria (Sekaran & Bougie, 2016). This size captures variability across regions, such as differences in income, health outcomes, and healthcare spending, providing sufficient statistical power and precision for robust and reliable conclusions.

The collected variables initially had heterogeneous scaling, with, for example, healthcare spending measured per capita and deceased by illness for the aggregate region. Mixed scales can lead to explanatory issues, such as larger values dominating analyses or regression coefficients becoming difficult to compare. To address this, all variables are transformed into per capita measures for consistency (Wooldridge, 2012).

5.3. Variables

The key variables used in the econometric analysis are:

- Deaths attributable to illness per capita (mortality)
- Mean income per capita
- Healthcare spending per capita
- Gross Regional Product per capita

Deaths attributable to illness per capita is used to evaluate health outcomes, obtained from Socialstyrelsen's official statistics database (Socialstyrelsen, 2024a). This measure includes all

deaths caused by illnesses within Swedish regions. By looking at illness-related mortality across all age groups, this variable provides a comprehensive morbidity indicator of regional health, reflecting the quality of healthcare, preventive measures, and the overall health situation. Focusing solely on deaths under 65, for instance, would overlook chronic conditions and elderly health, providing an inadequate representation of population health. While neither being an absolute mortality or morbidity variable, for simplicity, this variable is referred to as mortality or mortality rates in this thesis. According to the Grossman model presented in the theoretical framework, this variable is used as both a health input and output in the regressions.

The measure for income is mean income. This variable is collected from SCB (Statistiska Centralbyrån, 2023a). This variable illustrates both general income levels and regional wealth distribution. It aims to reflect the economic conditions experienced by most residents. For simplicity, it is referred to as income in this thesis. According to the Grossman model presented in the theoretical framework, this variable is used as both a health input and output in the regressions.

Healthcare spending per capita is sourced from Kolada (Kolada, 2024). It measures how much, each resident "costs" in healthcare utilization. A high value would, for instance, indicate that residents on average more frequently utilizes (or uses more costly) healthcare services. According to the Grossman model presented in the theoretical framework, this variable is used as both a health input and output in the regressions.

Gross regional product per capita measures labor productivity, quantifying the value added from goods and services produced (Thomson et al., 2019). It is obtained from SCB (Statistiska Centralbyrån, 2023b). It is measured in current prices. GRP per capita is calculated by dividing the total GRP by the population. For simplicity, it is referred to as GRP in this analysis. According to the Grossman model presented in the theoretical framework, this variable is used as both a health input and output in the regressions.

Control variables include mean age, fraction of women, and fraction of higher educational attainment (university level or above). They are sourced from Kolada and SCB. These variables aim to represent differences in regional contextual factors. For example, regions with a higher fraction of women might have different health outcomes or healthcare spending due to women's longer life expectancy. In addition, time-fixed effects are added to control for factors

that vary across time but are constant across regions. Time-fixed effects are discussed in detail in section 5.5. These variables aim to see how confounding variables may shift the relationship and what implications the added variables may have on the overall relationship.

In the Grossman model, health outcomes (mortality), income, healthcare spending, and productivity can be seen as both inputs and outputs within the household production function. Health, for instance, acts as an input by impacting productivity and to produce income, while also functioning as an output through the accumulation of health capital. Likewise, income and productivity can influence investments in healthcare and lifestyle improvements, ultimately influencing health outcomes.

In summary, the selected variables align with the thesis's aims and hypotheses, providing a solid foundation for statistical analysis. All independent and dependent variables are log-transformed to address measurement non-normalities in the data (Feng et al., 2014). Table 2 displays a list of the variables included in the analysis.

Table 2: A detailed list of all variables employed in the analysis, including their type and what they measure

VARIABLE NAME	DENOTED	ABBREVIA-TION	TYPE	TRANSFORMED
Mean Income	Income	INC	Dependent & Independent	Log-transformed
Deaths Attributable to Illnesses Per Capita	Mortality	MORT	Dependent & Independent	Log-transformed
Healthcare Spending Per Capita	Healthcare spending	HCS	Dependent & Independent	Log-transformed
Gross Regional Product Per Capita	GRP per capita	GRP	Dependent	Log-transformed

Percentage of Higher Education	-	CV	Control	Standardized
Percentage of Females	-	CV	Control	Standardized
Mean Age	-	CV	Control	Standardized
Total Residents	-	-	Descriptive	Standardized
Year/Time Fixed Effects	-	δ / TE	Time Effects	-

5.4. Data

5.4.1. Data Analysis Procedure

The collected data is panel data. Panel data is a type of data form that includes several entities (here, regions) over a period of time (Wooldridge, 2012). The advantage of panel data is that it controls for unobserved heterogeneity—differences across regions that do not change over time. For example, differences in regional policies, cultural attitudes, or historical health investments may influence the outcomes but stay constant throughout the studied period. Panel data also increases the sample size and statistical power, making it easier to detect relationships between variables. This structure is useful in this study as it allows for examining time trends and region-specific variations in health and income outcomes, thus providing a broader understanding of their dynamics than a purely cross-sectional or time-series dataset could do. The general model of panel data regression looks as follows:

$$y_{it} = \beta_0 + \beta_1 X_{it1} + \beta_2 X_{it2} + \cdots \beta_k X_{itk} + \alpha_i + \epsilon_{it} \qquad (2)$$

Where:

y_{it} = Dependent variable

β_0 = Intercept

$\beta_1, \beta_2, \ldots, \beta_k$: Coefficients for the independent variables

$X_{it1}, X_{it2}, \ldots, X_{itk}$: Independent variables for entity i at time t

α_i : Unobserved time-invariant effect for entity i (fixed or random)

ϵ_{it}: Error term for entity i at time t.

5.5. Model Specification

Specifying a model when performing panel data regressions is a crucial step in generating reliable, valid, and generalizable results (Creswell, 2003). A simple linear regression, for example, is not a feasible choice when working with panel data, as it defies the assumption of a regular linear regression model, such as independence of observations and homoscedasticity (Wooldridge, 2012). To address these challenges, other models are needed, such as fixed-effects (FE) and random-effects (RE) models. The fixed-effects model controls for unobserved differences (heterogeneity) that do not change over time by allowing each entity (here, region) to have its own intercept while the random-effects model assumes that the unobserved variation is uncorrelated with the independent variables (Clark & Linzer, 2014). Because of the complexities associated with the data, it is essential to carefully choose which model to use. On the one hand, it is important to have a model which accurately captures the effects. On the other hand, it is important to not overcomplicate the model, as it may lose efficiency.

There are several factors to keep in mind when choosing the model specification. Hoechle (2007) has written a journal article that elaborates on the different models available in Stata for panel regressions and acts as the basis of choice of model for this thesis. The initial regression in this thesis used a simple panel data regression model (Equation 2) with robust standard errors. However, when running the diagnostics, it is found that autocorrelation is highly present. Thus, clusters are added which account for autocorrelation and heteroskedasticity (Hoechle, 2007). This, on the other hand, fails to account for cross-sectional dependence (data points from different regions are related to each other) which also is present in the regression models. Instead, Hoechle suggests using Driscoll and Kraay's standard errors (Driscoll and Kraay, 1998). It accounts for autocorrelation, heteroskedasticity, and cross-sectional dependence. Because all dependent and independent variables are log-transformed, this thesis employs a log-log model.

Because this study aims to investigate the relationships between health, income, healthcare spending, and regional productivity within the context of Swedish regions, focusing on how these relationships are influenced by contextual factors such as regional disparities and time-

specific effects, the regressions are first performed without any control variables. That is, the regression exclusively includes the dependent and independent variables. After this is performed, the regression models are re-performed two times: once where only the control variables are considered and once where control variables and year-fixed effects are added. Year-fixed effects (or time-fixed effects) control for factors that vary across time but are constant across regions within each year, such as macroeconomic trends, national policies, or technological developments (Baltagi, 2005, p. 6). This variable therefore captures and controls for the possible effects on the dependent variables from, for instance, structural shifts or national policies.

The first regression performed uses a simple panel regression equation, as seen in equation (2), with only one independent variable. It looks as follows:

$$\log(Dependent_{it}) = \beta_0 + \beta_1(Independent_{it}) + \alpha_i + \epsilon_{it} \qquad (3)$$

Where the variable on the left-hand side of the equal sign in equation (3) is the dependent variable in region *i*, year *t*, β_0 is the intercept, β_1 is the independent variable in region *i*, year *t*. All hypothesis suggests $\beta_1 < 0$. α_i is fixed or random effects and ϵ_{it} is the error term for region *i* at time *t*. The standard errors are robust, computed using Driscoll and Kraay's method, which accounts for potential issues with heteroscedasticity, autocorrelation, and cross-sectional dependence. Thereafter, control variables are added:

$$\log(Dependent_{it}) = \beta_0 + \beta_1(Independent_{it}) + \sum_{k=1}^{K}\beta_k(CV_{kit}) + \alpha_i + \epsilon_{it} \qquad (4)$$

Where The sum of $\beta_k(log(CV_{kit}))$ is the control variables (*k*) as described in section 5.3. in region *i*, year *t*. Lastly, the third model specification adds time-fixed effects:

$$\log(Dependent_{it}) = \beta_0 + \beta_1(Independent_{it}) + \sum_{k=1}^{K}\beta_k(CV_{kit}) + \alpha_i + \delta_i + \epsilon_{it} \qquad (5)$$

Where δ_i are the added time-fixed effects. All hypotheses are estimated using equations (3), (4) and (5), consecutively, to observe the effects of adding region-and time specific variables. $H_{1.1}$ and $H_{1.2}$ measure the years 1997-2022, H_2 measures 2001-2023, H_3 measures 2001-2023

and H_4 measures 1997-2022. The regression results are estimated using the "xtscc" command in Stata, implementing this robust error correction for panel data, which accounts for heteroscedasticity, autocorrelation, and cross-sectional dependence. The equations used for each of the hypotheses can be found in Appendix B.

The reason one model omits the year fixed effects, and one includes it is twofold. On the one hand, comparing the two models is an interesting aspect to add, since it is possible to detect whether time-specific factors affect the result or not. If the coefficients change substantially, this suggests that unobserved time-varying factors play a role in the studied relationships. On the other hand, as is discussed below, including year-fixed effects is accompanied by the risk of multicollinearity. Including both models acknowledges the tradeoffs contained in regression modeling. The model without year-fixed effects provides more precise estimates but risks omitting important time-specific factors. The model with year-fixed effects accounts for these factors but has potential multicollinearity. By presenting both, the analysis offers a balanced perspective, enhancing the reliability, validity, and credibility of the study.

However, before performing regression analysis to evaluate the hypotheses, an exploratory data analysis (EDA) is conducted. EDA is a crucial step in data analysis, as it assists in detecting possible patterns, and variances and provides an initial understanding of the relationship between variables. Thus, descriptive statistics are performed to get familiar with the data.

5.5.1. Exploratory Data Analysis and Descriptive Statistics

Table 3 summarizes the mean values of the original, non-transformed variables used in the analysis. The red highlights illustrate the lowest values, and the green highlights depict the highest values. As seen, all income-related variables show the highest values in Region Stockholm, perhaps unsurprisingly as it is the capital. Except for that, at first glance, there are no obvious outliers or patterns. Table 4 further illustrates basic descriptive statistics for the data for all regions across all years.

Table 3: Mean values for all regions included in the study for all utilized variables

REGION	MEAN INCOME	MORT-ALITY	GRP	HEALTH SPEND-ING	HIGHER EDU-CATION	% FEMALE	MEAN AGE	TOTAL RESID-ENTS
Region Blekinge	222 723	0,011	313 081	27 758	0,20	0,493	42,9	153 949
Region Dalarna	223 177	0,011	316 663	26 752	0,17	0,498	43,2	280 648
Region Gävleborg	223 146	0,012	302 835	26 951	**0,17**	0,499	43,1	280 603
Region Halland	239 527	0,009	308 196	**24 811**	0,21	0,501	41,5	302 917
Region Jämtland Härjedalen	**218 227**	0,012	315 365	26 937	0,20	0,497	42,9	**128 857**
Region Jönköping	227 827	0,010	339 923	25 377	0,18	0,498	41,2	342 342
Region Kalmar	218 569	0,012	306 265	26 008	0,18	0,499	**43,4**	238 085
Region Kronoberg	223 515	0,010	354 466	24 986	0,20	0,494	41,5	187 174
Region Norrbotten	232 819	0,011	390 611	28 473	0,20	**0,490**	42,9	251 787
Region Skåne	227 373	0,010	336 606	24 856	0,24	**0,504**	40,9	1 249 916
Region Stockholm	**279 508**	0,008	**531 263**	25 296	**0,30**	0,504	**39,1**	**2 082 215**
Region Sörmland	226 242	0,010	**289 051**	25 570	0,18	0,501	42,1	275 069
Region Uppsala	238 631	0,008	355 649	26 259	0,28	0,502	39,6	337 691
Region Värmland	220 765	0,011	306 724	26 524	0,19	0,500	43,1	276 760
Region Västerbotten	224 346	0,010	331 831	**29 005**	0,25	0,496	41,1	262 460
Region Västernorrland	230 169	0,012	346 630	27 651	0,19	0,498	43,3	244 818
Region Västmanland	232 000	0,010	317 036	26 125	0,20	0,499	41,9	262 326
Region Örebro län	224 615	**0,015**	328 357	26 677	0,20	0,502	41,6	285 634
Region Östergötland	225 785	**0,007**	331 739	25 439	0,22	0,497	41,1	435 161
Västra Götalands län	236 835	0,009	379 973	25 018	0,23	0,500	40,8	1 600 513
Total	229 790	0,010	340 113	26 324	0,210	0,499	41,858	473 946

Table 4: Observations, mean, standard deviation, minimum and maximum values for all variables in the study

VARIABLES	OBS	MEAN	STD. DEV.	MIN	MAX
Mean Income	520	229790	52110	140500	402700
Mortality	540	0,01	0,00	0,01	0,02
GRP	540	340113	100411	175115	795950
Health Spending	439	26324	6007	16295	43448
Higher Education	540	0,2	0,05	0,12	0,36
Percent Female	440	0,5	0,00	0,49	0,51
Mean Age	520	42	1	38	45
Total Residents	540	473946	517079	126201	2454821

The Exploratory Data Analysis found that several variables had skewed distributions. Skewed distributions can be problematic because they indicate that the data is not symmetrically distributed, which can affect statistical analyses and interpretations. Skewed data can lead to biased estimates and misleading conclusions, as well as hide the true relationship between variables, making it harder to draw reliable conclusions (Wilcox, 2003).

To address skewed distributions, the independent and dependent variables are log-transformed, and the control variables are standardized. Log-transformation reduces skewness by reducing larger values and expanding smaller ones, bringing the data closer to a normal distribution. Standardization further warrants comparability by changing variables to have a mean of zero and a standard deviation of one. The independent and dependent variables are not standardized as it requires to interpret the regression coefficients in terms of standard deviation, which is not

optimal for interpretation. These steps improve the reliability of statistical analyses and make it easier to interpret the relationships between variables (Feng et al., 2014). The variables before and after the log transformation can be found in Appendix A, Tables 7-10.

5.5.2. Diagnostics

Fixed vs Random Effects

Before performing the panel data regression, it is essential to determine whether the regression calls for fixed or random effects. Fixed effects control for unobserved, time-invariant characteristics of the regions that may influence the dependent variable, for instance, geography or cultural factors. A key assumption is that the unobserved characteristics are correlated with the independent variables. Random effects, on the other hand, assume unobserved characteristics are uncorrelated with the independent variables. That is, geography or cultural factors do not affect the independent variables.

Although there are tests to perform to aid in the decision, such as the Hausman test, it is important to consider the context of the analysis to be performed. In this thesis, it is reasonable to assume that unobserved effects are related to the independent variables. For example, income levels may be affected by the availability of natural resources or cultural differences. Similarly, healthcare spending can be influenced by healthcare infrastructure or regional policies. Therefore, this thesis employs fixed effects for all regressions.

Stationarity & Unit-Root

One of the main assumptions of regression analysis is stationarity. Stationarity refers to a property of a time series where its statistical properties, such as mean, variance, and autocorrelation, remain constant over time (Wooldridge, 2012). Non-stationary data can lead to biased regression results, such as untrue relationships, because the relationships between variables may only reflect general trends rather than actual associations. Thus, a Unit-Root test is performed.

The initial test showed that the variable for log-transformed healthcare spending is non-stationary, regardless of using trends and lags. To address this issue, the variable is differentiated to remove the unit root. Differencing is a common method to handle non-

stationarity, as it removes trends by looking at changes between observations, year by year, rather than absolute levels (Hyndman & Athanasopoulos, 2018). By transforming the variable into a stationary series, it becomes possible to use it in regression analysis without going against the assumptions of the model. The consequence of this transformation is that the healthcare spending variable is interpreted in terms of growth rates.

Autocorrelation

Autocorrelation means the correlation of a variable with itself across different years, meaning that the value of a variable at one time may be influenced by its value at a previous time (Wooldridge, 2012). Autocorrelation can be problematic in regression analysis because it violates the assumption of independent errors, leading to biased standard errors and unreliable significance tests. However, since the regression uses Driscoll and Kraay standard errors, autocorrelation is already accounted for and needs therefore not to be taken into account.

Multicollinearity

Multicollinearity occurs when two or more independent variables in a regression model are highly correlated, making it difficult to evaluate the individual coefficient of each variable on the dependent variable (Murel & Kavlakoglu, 2023). High multicollinearity can increase standard errors and reduce the reliability of coefficient estimates.

When initially testing for multicollinearity, it is found that all regressions exhibit high levels of multicollinearity. To further investigate this, a correlation matrix with p-values is created, revealing that some control variables are highly correlated. One common solution is to omit one of the two highly correlated variables (Wooldridge, 2012). In this case, the fraction of the population aged 65+ is excluded from all regressions and instead only included mean age. After this adjustment, the test for multicollinearity is re-performed, and the results confirm that multicollinearity is no longer present. However, when including a time-fixed variable in the regression, all analyses are found to have multicollinearity. This, on the other hand, is not a major problem when using fixed effects. Multicollinearity caused by fixed effects is generally not a serious issue when the main focus is the key variables of interest. In this case, since the focus is solely on the dependent and independent variable, it does not bias the coefficient estimates but can make them less exact (Allison, 2012). It is also quite reasonable to see

multicollinearity in instances where variables have increased over time, such as income and healthcare spending.

Heteroscedasticity

Heteroscedasticity refers to a situation where the variance of the error terms in a regression model is not constant across observations, potentially leading to less efficient estimates and unreliable statistical outcomes (Wooldridge, 2012). Heteroscedasticity can be addressed by using Driscoll and Kraay standard errors.

The advantage of using these standard errors is that they adjust for heteroscedasticity, ensuring that the standard errors and significance tests stay valid even when the assumption of constant variance is violated. A potential disadvantage is that robust standard errors may lead to a loss of precision. This is, however, not an issue in this thesis.

6. Results

The following section presents the results obtained from the analysis. The section is divided by the hypotheses, where the results of the three models are compared. The aim is to portray what happens when gradually adding effects to the regressions, to later discuss its implications in the next section.

Table 5 illustrates the coefficient results of the three models used for each hypothesis. The p-values are reported in the adhering parenthesis. In this log-log model, the coefficients represent elasticities, indicating the percentage change in the dependent variable associated with a 1% change in the independent variable, holding other factors constant. For simplicity, the model without control variables is referred to as the exclusion model, whereas the two models including control variables—either with or without time-fixed effects—are referred to as inclusion models.

Table 5: Results of coefficients and p-values for all regressions performed

Model			No control variables (Eq. 3)	Control variables excl. Time (Eq. 4)	Control variables incl. Time (Eq. 5)
Hypothesis	**Dependent**	**Independent**	**Coefficient (p-value)**	**Coefficient (p-value)**	**Coefficient (p-value)**
H1.1	Mortality	Income	-0.194 (0.00)***	0.183 (0.05)**	0.312 (0.03)**
H1.2	Income	Mortality	-2.82 (0.00)***	0.096 (0.14)	0.020 (0.04)**
H2	HC Spending	Income	0.007 (0.528)	0.179 (0.00)***	-0.136 (0.13)
H3	Mortality	HC Spending	-0.230 (0.03)**	-0.051 (0.54)	-0.060 (0.19)
H4	Productivity	Mortality	-3.030 (0.00)***	0.157 (0.29)	0.150 (0.08)**

**Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level*

$H_{1.1}$: Higher income is associated with lower deaths attributable to illnesses per capita within Swedish regions

$H_{1.1}$ is first evaluated using log-transformed deaths attributable to illnesses (mortality) as dependent variable and log-transformed mean income as independent variable, without any control variables (exclusion model). The regression is significant at the 0.1 level and exhibits negative coefficient of -0.19. This means that for every 1% increase in income, mortality decreases by approximately 0.19%. In other words, higher income may lead to fewer illness-related deaths. Given the coefficient is less than one in absolute value, this represents an inelastic relationship which means that the change in mortality is not very large in relation to the change in income. Thus, when excluding control variables, we reject the null hypothesis and conclude that higher incomes are associated with fewer deaths.

However, when control variables are included, irrespective of incorporating time-fixed effects, the results instead portray statistically significant and positive coefficients of 0.18 and 0.31, without and with time effects respectively. This means that for every 1% increase in income,

mortality rates increase by approximately 0.18% (disregarding time effects) or 0.31% (regarding time effects). These relationships are also inelastic. These findings contradict the hypothesis that higher income is associated with fewer deaths and instead suggest that higher income is associated with more deaths attributable to illnesses.

$H_{1.2}$: Higher deaths attributable to illnesses per capita are associated with lower income within Swedish regions.

$H_{1.2}$ is examined using mean income as dependent variable and mortality as independent variable. Excluding the control variables, the regression model is statistically significant and depicts negative correlation between income and mortality, with a coefficient of -2.8. This implies that a 1% increase in mortality is associated with a decrease in income of 2.8%. Hence, it suggests that worse health outcomes result in lower income levels, as suggested by hypothesis 1.2. Because the absolute value of the coefficient is greater than one, this represents an elastic relationship. That is, changes in income levels are relatively responsive to changes in mortality. Therefore, in the exclusion model, we reject the null hypothesis.

On the other hand, while no significant results are depicted for the model omitting the time-variable, adjusting for time effects reach statistical significance. With a positive coefficient of 0.02, the regression illustrates that a 1% increase in mortality is associated with a 0.02% increase in income. This is an inelastic relationship, meaning that changes in income are not very responsive to the change in mortality. In other words, higher levels of deaths attributable to illnesses seem to increase income levels, although quite weakly. This finding contradicts the suggested hypothesis and instead proposes the reverse relationship.

H_2: Higher income is associated with lower healthcare spending per capita within Swedish regions.

H_2 is investigated where the dependent variable is the first difference of the log-transformed healthcare spending per capita, and income is the independent variable. Because healthcare spending is differentiated, the variable is interpreted as the change in growth rate. The results of the exclusion model do not show significant results ($p=0.538$) thus meaning that the regression fails to reject the null hypothesis. Therefore, it is not possible to conclude that higher income is associated with lower healthcare costs per capita within Swedish regions when excluding control variables.

However, the H_2 regression shows statistically significant results when omitting the time variable. A 1% increase in income (level) is associated with a 0.13% increase in the growth rate of healthcare spending. This suggest that higher income levels drive healthcare spending to grow slightly faster over time. This relationship is inelastic which means that the change in mortality is not very large in relation to the change in income. These findings are inconsistent with the hypothesis that higher income is associated with lower healthcare spending per capita. When the time variable is considered, statistically significance is not achieved.

H_3: Higher healthcare spending per capita is associated with lower deaths attributable to illness per capita

The third hypothesis, H_3, takes log-transformed mortality as the dependent variable and the first difference of the log-transformed healthcare spending as the independent variable. In the exclusion model, a significance level of 0.03 is obtained and yields a coefficient of -0.23, which means that a 1% increase in the growth rate of healthcare spending is associated with a 0.23% decrease in mortality rates. The negative sign indicates an inverse relationship: as healthcare spending grows faster, mortality rates decline. This relationship is inelastic, thus the change in mortality is not very large in relation to the change in healthcare spending.

When excluding control variables, the results thus support the hypothesis that higher healthcare spending per capita is associated with lower deaths attributable to illness per capita. However, neither regression in the inclusion models show significant results. Therefore, it is not possible to reject the null hypothesis and instead conclude that there is not enough evidence to state that higher healthcare spending per capita is associated with lower deaths attributable to illness per capita.

H_4: Higher deaths attributable to illness per capita are associated with lower regional productivity

Lastly, H_4 is evaluated by looking at log-transformed GRP per capita as the dependent variable and log-transformed mortality as the independent variable. In the exclusion model statistical significance is obtained ($p = 0.00$) and generates coefficient of -3.03, which means that a 1% increase in mortality is associated with a 3% decrease in GRP per capita. This represents an elastic relationship, thus indicating that the change in GRP per capita is relatively large in relation to the change in mortality. This finding is consistent with the hypothesis that higher

levels of death are associated with lower levels of productivity. This suggests that regions with worse health outcomes have lower productivity levels when excluding other external effects.

When further including control variables, omitting the time variable it is not possible to obtain statistical significance and thus fails to reject the null hypothesis. However, including the time variable, it significant at the 10% level. Here, the coefficient is positive at 0.15, which indicates that a 1% increase in mortality is associated with a 0.15% increase in GRP. This is an inelastic relationship. These findings contradict the hypothesis that more deaths are associated with lower productivity and instead show that higher levels of mortality are related to higher levels of productivity.

6.1. Sensitivity Analysis

To ensure the robustness of the results and assess the reliability of the findings, a sensitivity analysis is conducted. Given the complexities of regional differences in income, health, productivity, and healthcare costs, a sensitivity analysis investigates how different assumptions, and methodological alternatives affect the outcomes. It also identifies to which extent specific variables and model specifications influence the relationships. Employing this thereby ensures that the conclusions are not overly dependent on specific assumptions. The results of the sensitivity analysis can be found in Appendix B, Tables 11-12.

To address potential skewness and multicollinearity, the main analysis uses log-transformed variables. Sensitivity tests using raw and non-standardized data are conducted to examine the robustness of results to these transformations. The most interesting aspect to consider is whether the coefficients are positive or negative, compared to the primary model.

The first sensitivity analysis evaluated the same three models (no control variables, control variables excluding time effects and control variables including time effects) as in the main analysis but used non-transformed variables. That is, the raw data without log-transformations or standardization. Disregarding the values of the coefficients, the direction of the relationships are almost identical to the original model. The exceptions are not statistically significant in either the sensitivity analysis or the main model. This indicates that the applied transformations

(log-transformations and standardizations) do not drastically affect the relationships, meaning the model is robust to these adjustments.

Moreover, the sensitivity analysis performed is based on the key methodological assumptions. The used model uses Driscoll and Kraay's standard errors (assumes panel-specific autocorrelation and heteroskedasticity) and employs fixed effects. Sensitivity tests using an alternative model (xtreg in Stata), and random effects are conducted to evaluate the robustness of the results to these assumptions. Moreover, regressions are performed where only data measured after 2010 is tested, as well as one case where two extremes (highest and lowest income regions) are excluded. All regressions use the same transformed variables as in the original model.

The alternative model (xtreg) depicts the same relationship between the variables (in terms of the direction) as the primary model. These results further confirm the robustness of the results across methodological choices. This is also true for the regression which excludes the highest- and lowest income regions. Moreover, the model utilizing random effects shows almost identical results as the fixed-effects model, with the exception of H_2 which, on the other hand, is statistically insignificant in the sensitivity analysis. This means that the choice of fixed effects does not seem to significantly alter the conclusions. Lastly, the post-2010 model deviates from the primary model, where only H_3 and H_4 are consistent. This indicates that the relationships in the data may have changed over time, suggesting that results from the full dataset could be influenced by earlier years.

To conclude, this sensitivity analysis finds that the majority of the tests are consistent with the primary model which suggests that the main results are robust and reliable. However, the deviations in the post-2010 model may be due to shifts in economic circumstances, healthcare policies, or regional inequalities after 2010. These findings call for further research to investigate potential differences in the relationships between health, income, healthcare costs, and productivity. Future studies could benefit from focusing specifically on post-2010 data, exploring what factors might explain these deviations.

6.2. Limitations

To further investigate the results, it is necessary to introduce possible limitations and what has been done in the data analysis process to minimize them. By doing this, the thesis further ensures that the obtained results are valid, reliable and credible.

First, data-related limitations could appear. As mentioned in Section 5.2., not all variables have data for all years due to non-consistent data entries. This could potentially lead to gaps in the time trends and influence the robustness of the findings. The study further relies solely on secondary data sources, which may have limitations related to how accurate the data is. However, this limitation has been tried to be minimized by firstly ensuring a large sample size to account for non-consistent data entries and, secondly, only obtaining data from nationally recognized and official sources. While there is a possibility that the data might be biased regardless of using well-known sources, the measures taken minimize the risks.

Second, there could be limitations related to the methodological approach. While the methodology accounts for autocorrelation, heteroskedasticity, and cross-sectional dependence, The adjustments might not completely reflect all the complex, non-straightforward connections. However, these issues have been accounted for by employing the “xtscc” model in Stata, which provides robust standard errors that account for such complexities, thereby reducing the risk of misinterpreting the results. Moreover, there is a risk of the omitted variable bias. Omitted variable bias occurs when a regression model excludes one or more relevant variables that are correlated with both the dependent and independent variables (Hanck et al., 2024). This can lead to biased and inconsistent estimates of the coefficients, as the effects of the omitted variables are incorrectly associated with the included ones. This risk has however been minimized through the inclusion of fixed effects. The fixed effects model accounts for unobservable, time-invariant factors, thus controlling for potential causes of omitted variable bias.

Finally, there is the limitation of human error. There is always a risk that something has not been correctly performed, documented, or transferred into the software, which could affect the reliability of the results. However, this risk has been minimized through continuous checks and controls of the data to ensure the accuracy and rationality of the results throughout the process.

Moreover, the use of software familiar to the author, Stata, has further reduced the likelihood of errors arising from incorrect implementation or misinterpretation of results.

6.3. Summary of Results

To sum up, the exclusion model validates all hypotheses except for H_2, where there is not enough evidence to draw any conclusions. This suggests that higher levels of income and general wealth of regions are associated with better health outcomes, while regions that have higher (growth rate of) healthcare spending per capita also have healthier residents (lower mortality per capita). Further, it is found that regions with worse health outcomes (higher mortality) are less productive. Thus, it is possible to see that higher income may contribute to better health outcomes, which in turn increases productivity.

However, the results shift in the inclusion models. When omitting the time variable, enough evidence (statistical significance) is found for $H_{1.1}$ and H_2 and both have positive correlations. That is, regions with higher incomes are associated with worse health outcomes (higher mortality) and have higher healthcare spending per capita. When the time variable later is introduced, the relationship between the variables shifts somewhat further. First, $H_{1.2}$ and H_4 also become statistically significant. $H_{1.2}$. finds that contrary to the inclusion model, regions with worse health outcomes portray higher income levels. Similarly, H_4 finds that regions with worse health outcomes seem to have higher productivity. Second, while the observed relationship for $H_{1.1}$ does not change, H_2 instead illustrates that when time-fixed effects are included, wealthier regions seem to have lower healthcare expenditures per capita. A summary of the observed relationships can be found in Table 6 below.

The sensitivity analysis further reveals that the observed results in the primary model seem to be robust and reliable for most tests. However, the post 2010-data warrants for future studies where more independent variables can be included to see what might yield the adverse effects compared to the primary model.

Table 6: Simplified summary of results with direction of relationship for all regression models

HYPOTHESIS	EXCLUSION MODEL (EQ. 3)	INCLUSION MODEL EXCL. TIME (EQ. 4)	INCLUSION MODEL INCL. TIME (EQ. 5)
H1.1	↑ Income → ↓ Mortality	↑ Income → ↑ Mortality	↑ Income → ↑ Mortality
H1.2	↑ Mortality → ↓ Income	-	↑ Mortality → ↑ Income
H2	-	↑ Income → ↑ HCS	↑ Income → ↓ HCS
H3	↑ HCS → ↓ Mortality	-	-
H4	↑ Mortality → ↓ GRP	-	↑ Mortality → ↑ GRP

7. Discussion

The following section discusses and interprets the results from the previous section by considering the theoretical framework and previous research. It starts by reviewing the differences in findings and thereafter presents possible theoretical and practical implications.

In this section, for simplicity, $H_{1.1}$ is named income-health relationship, $H_{1.2}$ is denoted as the health-income relationship, H_2 is called income-healthcare spending, H_3 is called healthcare spending-health and H_4 is referred to as health-productivity.

7.1. The dynamics of Health, Income, Healthcare Spending and Productivity

While the regression results show weak (inelastic) relationships between health, income, healthcare spending and productivity, this can be attributed to several factors, including the complexities of these variables and the challenges of including all relevant regional differences.

This therefore warrants the question: Why do we see these differences, and what do they mean?

At first glance, when solely looking at the relationship between health, income, healthcare spending, and productivity without external influences (the exclusion model), it seems as if both positive growth rate of healthcare spending per capita and income have a positive influence on mortality. That is, these results suggest that improvements in income and faster (in terms of growth rate) healthcare spending could positively influence the health outcomes of the population or, conversely, lower the mortality rates. The results further show that higher levels of mortality have negative effects on income levels, meaning, regions with worse health outcomes are less wealthy. Similarly, higher mortality rates seem to harm GRP per capita which implies that worse health outcomes are negative for regional productivity.

These findings are consistent with both the Grossman model (1972, 2000) and previous empirical studies (see for example Becker, 2007; Benzeval & Judge, 2001; Mitchell & Bates, 2011). Grossman's wealth model (Equation 1) showed, for example, that higher total wealth results in more time to spend on improving health. Here, we see that that is true. Regions with higher incomes are seen to have lower mortality thus aligning with Grossman's model. Similarly, Ettner's study (1996) in England finds a comparable result, highlighting that increased income is associated with improved health.

The observed reverse relationship—health-income—is somewhat consistent with the earlier studies, such as Barro (2013) Bloom, Kuhn, and Prettner (2019), and Bloom and Williamson (1998), who argue that better health outcomes (here, lower mortality) are associated with higher incomes. Similarly, it is consistent with the theoretical framework, which suggests that individuals desire to be healthier to have a greater opportunity to generate income. That is, better health is associated with higher income. That would thus mean that healthier regions experience higher wages. Due to the choice of variables, the result of this thesis indicates that poorer health is related to decreased income. Although the relationship is not identical, the interpretations and implications align, as both suggest a link between health outcomes and income levels. However, the result of this thesis from the exclusion model finds that this relationship also holds on the regional level in higher-income countries. This finding adds significance to the current body of research, as it highlights that the "pure" relationship between income and health is true in the regional context as well.

The exclusion model further suggests consistencies with the healthcare spending-health and health-productivity relationships and previous studies. First, similar to the studies by Crémieux et al (1999) and Martin et al. (2008), where their empirical findings show that increased healthcare inputs have positive impacts on health outcomes, the exclusion model suggests that increases in healthcare spending per capita are associated with better health outcomes. Second, the analysis finds results consistent with Mitchell and Bates (2011) and Arora (2001), which also find evidence that better health outcomes are associated with better productivity in high-income countries. This thesis further contributes to the body of literature by confirming the pure (i.e., non-control variable) relationship on the Swedish regional level.

However, the result of this thesis finds that when including the control variables (the inclusion model), these relationships are no longer true. While the relationships seem weak, it is not much different from previous studies. Ettner (1996), for example, depicts coefficients of 0.11 and Ecob and Davey Smith (1999) show coefficients around the same values. Thus, while the relationships are weak, they are within the standards of previous studies.

First, it is found that higher income is associated with higher mortality (income-health relationship). In other words, wealthier regions seem to experience worse health outcomes. This finding is interesting because it to some extent contradicts Grossman's model, along with the classical literature on the health-income relationship. Although the Grossman model is mostly considered on the micro level, the result of this thesis can argue that it does not hold for the macro perspective. Instead, the findings of the income-health relationship in this thesis may be more aligned with the research suggested by Ecob and Davey Smith (1999), which argue that the income-health relationship has diminishing returns. For a high-income country like Sweden, this becomes evident when comparing the results from the inclusion and exclusion model. When time effects are excluded, the proposed hypothesis suggesting higher income leads to lower mortality rates seems to be true. This is further emphasized when evaluating the sensitivity analysis conducted where only post-2010 data is included as it, similar to the exclusion model, found a negative correlation (see Appendix B, Tables 11-12). This highlights that over a shorter period of time, the conventional income-health relationship is true. However, over time, the return on health from income is diminishing.

On the one hand, it may be argued that there is a paradox in the income-health relationship. Beyond a certain income threshold, income increases may lead to diminishing returns in terms of health improvements, as suggested by Ecob and Davey Smith (1999). For instance, in wealthy regions like Stockholm, where healthcare access is good and income is high, the benefits of wealth in improving health outcomes are over time diminished. Increases in income within regions might initially improve health outcomes, as suggested by $H_{1.1}$ and previous research (Lorentzen, McMillan, & Wacziarg, 2008). As the population grows, the pressure on resources increases. In turn, the system may not be able to support the greater demand for healthcare, and thus illness-related mortality rates may increase. While the findings of this thesis contradict the Grossman model, the theoretical framework also highlights that if healthcare is more difficult to obtain, demand for healthcare decreases. In turn, more illness may prevail and thus, despite higher incomes, health outcomes may worsen.

Moreover, as mentioned in the Background (Section 2), Sweden's tax equalization system redistributes tax income to reduce regional inequalities. While this system aims to ensure equal access to public services, these findings suggest that it may not effectively address health inequalities. If the system fully equalized health outcomes, the observed relationship would generate a coefficient of -1 (unitary elasticity). That is, if income increases with 1%, mortality decreases with 1%. While the relationship is seen to be inelastic, the statistically significant coefficient indicates that wealth does not necessarily translate into better health. This highlights a potential fault in the equalization system. While it ensures equal access to public services, it may not account for other factors driving unequal health outcomes such as stressful lifestyle habits.

Thus, on the other hand, urbanization and stress may also play an important role in this dynamic. For example, higher-income jobs are often associated with a more sedentary behavior and higher levels of stress which consequently can give rise to worse health outcomes in the long run. As suggested by the Grossman model, health is produced through market goods (e.g. healthcare) and time spent on improving health (e.g. resting). If individuals spend more time working (generating income) the Grossman model would suggest that they have less time to use market goods and time on improving health, which may therefore decrease health outcomes. Therefore, regions with higher incomes may see higher mortality. This issue is not captured or solved by the tax equalization system since it primarily focuses on redistributing financial resources between regions rather than addressing behavioral factors. Thus, in larger

cities such as Stockholm, stress-related and poor lifestyle choices may contribute to higher illness-related mortality despite high incomes. While the results to some extent contradict previous studies, they help fill the gap in understanding income and health by showing that, while the relationships seem true at first glance (inclusion model), including time effects and contextual factors is crucial for understanding the relationship.

Second, the thesis finds that regions with higher levels of mortality, that is, regions with worse health outcomes, seem to have higher income levels when considering the unobserved time effects such as inflation rates and technological advances. Similarly, the regression implies a similar dynamic for the health-productivity relationship where regions with higher mortality seem to be associated with better productivity.

This observation could be reflected in the "Malthusian Response" Acemoglu and Johnson (2007) argue for, however, somewhat reversed. While Acemoglu and Johnson suggest that improved health would lead to worse economic outcomes through mortality decline but no fertility decline (that is, population growth) which then "diluted" income per capita, the case for Swedish regions could be that regions with higher levels of deaths attributable to illnesses may rather "concentrate" income per capita. In other words, higher illness-related mortality rates may lead to a reallocation of wealth or resources which can increase income for the remaining population. In return, increased income may assist households to invest in productivity-enhancing features such as education and technology which consequently could improve productivity.

Moreover, one explanation for the negative health-productivity relationship may be demographic. Since the data reflects mortality for all ages, the relationship might suggest that a region's workforce remains highly productive despite these deaths, as the economy is pushed by a younger, healthier labor force. For example, in larger cities, high illness-related mortality due to pollution or other stress may exist with high productivity because urban areas often draw highly skilled, economically active workers and industries.

Another reason for this observed relationship in the inclusion model (with the time variable) might also be technological advances such as AI, captured by the time variable, may allow production to be more efficient without the need for the human workforce. In that way, productivity could increase regardless of the health of individuals and rather be more dependent

on technology. These findings diverge from much of the previous literature, which generally emphasizes that improved health is positively associated with improved productivity. The divergence in results could be explained by the focus of previous studies on national-level or low-income populations, where the health-income and health-productivity relationships are more prominent due to less differences in socioeconomic and technological factors. On the contrary, this thesis examines these dynamics in a regional context within a high-income country, where disparities across regions and other confounding factors play a more significant role.

Indeed, the exclusion model in this study depicts results more in line with previous research, as the relationships are less influenced by confounding variables. However, when control variables and time-specific effects are accounted for, the results reveal more complex dynamics, highlighting how regional factors and technological progress may form these relationships differently. This thesis thus contributes to the current body of knowledge by providing a regional and contextual perspective on the health-productivity relationship, particularly within higher-income settings, where traditional observations may not fully apply.

The inclusion model further finds ambiguous results for H_2. When not controlling for the time variable, the observed coefficient does not support the hypothesis that higher income is associated with lower healthcare spending per capita. Instead, it is suggested that higher income accelerates the growth of healthcare spending. This finding is, on the one hand, consistent with Newhouse's (1977) study, which suggests that higher income levels are associated with higher medical spending. Although the current result reflects an effect on the growth rate of healthcare spending rather than its level, it aligns with the literature showing that higher income levels are linked to increased healthcare spending. When not considering the time variable, the reason higher income leads to higher healthcare expenditure may be similar to why higher income seems to be associated with higher mortality rates, that is, regions with higher-income individuals may take on lifestyles that could lead to health issues like stress and chronic conditions (such as diabetes, cardiovascular diseases). These conditions require long-term healthcare, contributing to higher costs.

Although this is true when excluding the time variable, the results shift when time is included in the model, and the reverse relationship is highlighted. While the regression showed statistically insignificant results, the regression was just slightly above the 10% threshold.

Therefore, it is still of relevance to discuss why the relationship shift. That is, when one also considers year-specific effects that might influence income levels across all regions, such as inflation or national events, the result suggests that healthcare spending per capita decreased. As regions' incomes increase, healthcare spending per capita decreases.

This finding is inconsistent with Newhouse. However, it is consistent with the Grossman model (1999, 1972), which suggests that wealthier individuals put investments into other proactive factors into the production function, such as improved diet or living conditions. Because Sweden has free universal healthcare, it is reasonable to assume that some proportion of individuals' income could be put on preventive measures instead of utilizing healthcare. These investments could lead to a less need for public healthcare services, thereby lowering per capita healthcare spending at the regional level. Thus, higher-income regions may have more resources to put on preventive health measures, such as better school lunches and outdoor gym facilities, which could reduce the need for consuming healthcare.

In addition, when time effects are included for and the relationship instead highlights that higher income lowers healthcare spending per capita, the time variable possibly shows the effects of technological improvements. This could decrease healthcare spending over time and wealthier regions may adopt these technologies faster, making healthcare distribution more efficient and less expensive. Wealthier regions may also experience diminishing returns to healthcare utilization. For example, wealthier regions may initially experience high healthcare spending because of high healthcare utilization. Over time, as populations become healthier, the demand for expensive healthcare treatments falls.

While the results to some extent contradict current studies, this thesis contributes to the existing literature by emphasizing the complexity and context-specific characteristics of the relationship between income and healthcare spending. While much of the previous research has focused on national or cross-country analyses, this study provides a regional perspective within a high-income context, highlighting how socioeconomic and time-fixed effects influence healthcare spending. By accounting for time-specific effects and considering the analysis within Sweden's healthcare system, these findings offer a more detailed understanding of how income and healthcare spending interact over time, highlighting the importance of regional and time variations in health and economic studies.

To sum up, the results are quite obscure and could be argued to suggest varying perspectives. While the results of the thesis suggest counterintuitive relationships between all variables, the findings of this thesis highlight several important factors that need to be considered. First, the weak relationship observed may be explained by unaccounted-for external factors, such as socioeconomic variables or environmental influences that were not included in the model. However, the weak coefficients also align with previous studies. Furthermore, the methodological approach, including the exclusion of time effects or other region-specific factors, may have contributed to the ambiguity in the results.

Then, the research question of this thesis resurfaces: How do regional contexts and time-specific effects influence the relationships between health outcomes, income, healthcare spending, and regional productivity in Swedish regions?

First, regions with higher incomes may experience higher levels of mortality, and regions that experience increased levels of income may see worse health outcomes. Second, regions with higher levels of mortality may experience higher income levels. Third, regions that also see higher incomes may find that the growth rate of healthcare spending per capita decreases. This means that improved income could lead to lower growth rate of healthcare spending, giving room to reallocate spending to other public resources—such as education and infrastructure. Lastly, regions that experience higher levels of deaths attributable to illnesses may see higher levels of productivity.

Second, the importance of accounting for confounding variables is seen when comparing the inclusion and exclusion models. As seen, the relationships shift when control variables are included, suggesting that the relationship between health, income, healthcare spending, and productivity is not as evident as initially thought. Rather, it implies a far more complex depiction, highlighting the second point: context is of utter importance when estimating these relationships. While it seems that higher income is associated with higher mortality rates, or that higher mortality is associated with higher levels of productivity, the discussion above highlights that other contextual factors, such as demographic structures and technological advances, should be considered when establishing conclusions for the relationships. As seen, the obtained results are likely not casual but should rather be interpreted as correlational findings. Therefore, decision-makers must account for these points when formulating policies.

7.2. Theoretical and Practical Implications

Because of the ambiguous results obtained, highlighting the theoretical and practical implications is beneficial for future research and implementation of policy, given the following sub-research question: What are the implications of the regional contexts and time-specific effects for regional policy?

First, one important theoretical implication is that the results show the complexity and ambiguity of the income-health-productivity relationship, highlighting that relationships observed at the macro level may differ from those at the micro level (as observed from Grossman's model). This suggests the need for further research to develop theoretical models that look at this complexity in high-income contexts.

Second, the comparison between the inclusion and exclusion model shows to the importance of integrating external variables in empirical research. Future studies should thus prioritize data collection to control for demographic, technological, and systemic factors. Moreover, future research could focus on the inclusion of several independent variables to examine how other variables may influence the relationships. Similarly, the findings suggest that previous studies may not fully explain the dynamics in high-income countries which emphasizes the importance of considering contextual factors when implementing ideas from previous studies. For example, the diminishing returns of income on health outcomes, as highlighted by Ecob and Davey Smith, may provide a more suitable theoretical perspective for understanding these relationships in similar contexts. The weak relationship observed in this analysis suggests a need for future research to better account for regional and time-specific factors that could provide a better understanding of the health-income-productivity relationship. Further consideration into the impact of confounding variables, such as social determinants of health, would be beneficial. Future research should therefore account for the contextual and country-specific factors when performing analyses.

Third, the observed relationships challenge future researchers to distinguish between correlation and causation. This further supports the need for robust and suitable econometric techniques with careful consideration for model specificity and diagnostic precision. For example, this study does not have enough evidence to conclude anything about healthcare spending and deaths attributable to illnesses, which thus could be better modeled and studied

in future studies. Lastly, the findings of this thesis suggest an increasing importance of considering technological advancements and structural differences when establishing the relationship between income, health, healthcare spending, and productivity. Future studies should therefore integrate these factors to better reflect the current reality. Moreover, an interesting aspect for future research could be to study the influence of the tax equalization system more closely by, for instance, seeing which effects its removal in the regression analysis would have on the overall outcome.

For policy makers, the results suggest that preventive measures may ease potential negative effects of income-related health outcomes, which highlights that policymakers should have a proactive approach to improving health when implementing policies and strategies. Moreover, the findings in this thesis imply that regions with higher incomes, but worse health outcomes, may need targeted health interventions such as stress management programs and public resources promoting active lifestyles. Thus, policymakers should aim to investigate closer into the income-health relationship in various regions to improve the quality of life for the residents.

Decision makers should further also examine whether high-income regions experience overstrained healthcare systems, as possibly found in this thesis, and actively work to invest in measures that improve the size and efficiency of the healthcare system in that specific region. In addition, while the tax equalization system balances the financial service power of the regions, this thesis found that it may not equalize the health inequalities. Therefore, policy makers could consider implementing strategies which, alongside with tax equalization, create more equitable health outcomes across regions. This could, for instance, be to impose stress-management programs or lifestyle-enhancing policies in regions who experience higher stress levels.

This thesis further reveals the importance of accounting for context-specific factors when implementing regional policies. Policies should aim to address underlying issues, such as resource distribution inefficiencies or unequal access to preventive measures, rather than drawing conclusions from correlational findings, as well as consider region-specific characteristics when implementing policies. Ultimately, while this thesis provides valuable insights into regional variations in income, health outcomes, healthcare spending, and productivity, it emphasizes that correlations should not be mistaken for causation. This

highlights the importance of policymakers carefully considering the findings with the context of other evidence to create strategies that promote equal economic growth, improved health outcomes, and sustainable productivity across all regions.

To conclude, these implications highlight that while the findings may seem counterintuitive, they suggest valuable insights for policymakers, researchers, and theorists. By emphasizing context, complexity, and the need for detailed modeling, this thesis contributes to a more nuanced understanding of the interplay between health, income, healthcare spending, and productivity on the regional level in high-income countries like Sweden.

8. Conclusion

This thesis aimed to investigate the relationships between health, income, healthcare spending, and regional productivity within the context of Swedish regions, focusing on how these relationships are influenced by contextual factors such as regional disparities and time-specific effects. This field of study was important because, although Sweden is generally a high-income country with low mortality rates, increasing inequalities between regions have become more pronounced. Simultaneously, few studies have been dedicated to investigating the relationship between income, health, healthcare spending and productivity in the regional context. Therefore, this thesis tried to fill this gap by examining the variables on the Swedish regional level, thus trying to obtain more clarity in how income, health, healthcare spending and productivity interact in the regional setting. By doing so, this thesis contributes to the literature by providing new empirical insights into these dynamics, specifically focusing on the less examined context of Swedish regions. This was directed by answering the following research question:

How do regional contexts and time-specific effects influence the relationships between health outcomes, income, healthcare spending, and regional productivity in Swedish regions?

Followed by the sub-question:

What are the implications of the regional contexts and time-specific effects for regional policy?

The study undertook a quantitative methodology with secondary panel data to answer the research question. The data consisted of inputs related to income, health, healthcare spending and productivity as the key variables, along with control variables for 20 Swedish regions between 1997-2023. Based on the Grossman model and previous research, five hypotheses were formulated. They were as follows:

$H_{1.1}$: *Higher income is associated with lower deaths attributable to illnesses per capita within Swedish regions.*

$H_{1.2}$: *Higher deaths attributable to illnesses per capita are associated with lower income within Swedish regions*.

H_2: *Higher income is associated with lower healthcare spending per capita within Swedish regions*.

H_3: *Higher healthcare spending per capita is associated with lower deaths attributable to illness per capita*.

H_4: *Higher deaths attributable to illness per capita are associated with lower regional productivity*.

The hypotheses were studied using panel regressions. First, the relationships were analyzed directly, without any confounding variables. Thereafter, control variables were included in order to see what effects external factors had on the relationships. Lastly, time-fixed effects were added to assess which impact time trend might have had on the dynamics. In addition, a sensitivity analysis was conducted to evaluate the robustness of the results.

This study found counterintuitive results. When excluding the confounding variables, all regressions except for H_2, that lacked significance, obtained results aligned with the proposed hypothesis and the majority of previous empirical studies. That is, higher income regions seemed to experience lower mortality rates, while regions with higher mortality portrayed lower levels of health and productivity. Similarly, regions with higher growth rate of healthcare spending per capita had lower mortality rates.

However, when the control variables were included the direction of the results shifted considerably. While only $H_{1.1}$ and H_2 were statistically significant when excluding the time variable, the model including the time variable found contrary results. On the one hand, H_2 was now significant and in accordance with the suggested hypothesis and previous research. That is, higher income regions seemed to have lower healthcare expenditure (in terms of growth rate) per capita. On the other hand, regions with higher incomes also now showed to have higher levels of mortality. Similarly, regions with higher mortality levels seemed to have higher income and productivity. The sensitivity analysis thereafter conducted confirmed the robustness of the results.

The thesis proposed several reasons why this relationship may have been observed. One reason, also highlighted in previous studies, is the prevalence of diminishing returns. That is, as income rises, the return on health decreases. Another reason attributed to the higher income—higher mortality relationship may have been that healthcare systems are overstrained. Moreover, one explanation for the high mortality-high income dynamic was that the income is "diluted" over a smaller population—also known as a "Malthusian Response".

While the results may seem weak or illogical, they reflect the complexity of income, health, and productivity, which are shaped by various contextual, regional, and time-specific factors which complicates the ability to establish straightforward relationships. Therefore, this thesis contributes methodologically by emphasizing the value of incorporating detailed models and control variables when analyzing these relationships. This thesis highlights the importance of comprehensive models in capturing these dynamics. Future research should refine theoretical models to better account for this complexity.

This thesis further highlights the need for policymakers to critically assess results and account for regional contexts when designing health policies. Strategies should be tailored to address underlying issues like preventive care and resource distribution to ensure equitable healthcare outcomes across Swedish regions. Ultimately, achieving balanced health, income, productivity, and healthcare spending requires a multifaceted approach that integrates empirical evidence with regional-specific considerations.

To sum up, while many studies have found clear positive relationships between income, health, and productivity at the national level, this thesis contributes to the literature by highlighting the

complexity of these relationships at the regional level within a high-income context. The results suggest that, rather than a direct, linear relationship, a more complex dynamic exists, influenced by regional characteristics and time-specific effects. This challenges the relevance of existing models, such as Grossman's, to regional data in high-income countries and creates new opportunities for future research.

Bibliography

Acemoglu, D. & Johnson, S. (2007). Disease and Development: The Effect of Life Expectancy on Economic Growth, *Journal of Political Economy*, [e-journal] vol. 115, no. 6, pp.925–985, Available Online: https://economics.mit.edu/files/4478 [Accessed 17 December 2024]

Allison, P. (2012). When Can You Safely Ignore Multicollinearity?, *Statistical Horizons*, Available Online: https://statisticalhorizons.com/multicollinearity/ [Accessed 17 January 2025]

André, C., Beom, J., Pak, M. & Purwin, A. (2021). Keeping Regional Inequality in Check in Sweden, *OECD Economics Department Working Papers*, [e-journal] no. 1689, Available Online: https://www.oecd-ilibrary.org/docserver/e4bec28f-en.pdf?expires=1707875623&id=id&accname=guest&checksum=8298E7089D40E8C89F1BC2DE29023CD7 [Accessed 13 January 2025]

Arora, S. (2001). Health, Human Productivity, and Long-Term Economic Growth, *The Journal of Economic History*, [e-journal] vol. 61, no. 3, pp.699–749, Available Online: https://www.jstor.org/stable/2698133 [Accessed 18 December 2024]

Baltagi, B. H. (2005). Econometric Analysis of Panel Data, Chichester ; Hoboken, Nj: J. Wiley & Sons

Barro, R. J. (2013). Health and Economic Growth, *Annals of Economics and Finance*, [e-journal] vol. 14, no. 2, pp.329–366, Available Online: https://ftp.aefweb.net/WorkingPapers/w572.pdf [Accessed 17 December 2024]

Becker, G. S. (1964). Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, First Edition, *Www.nber.org*, 1st edn, [e-book] NBER, Available Online: https://www.nber.org/books-and-chapters/human-capital-theoretical-and-empirical-analysis-special-reference-education-first-edition [Accessed 10 December 2024]

Becker, G. S. (1967). Human Capital and the Personal Distribution of Income, Institute of Public Administration

Becker, G. S. (2007). Health as Human Capital: Synthesis and Extensions, *Oxford Economic Papers*, vol. 59, no. 3, pp.379–410

Ben-Porath, Y. (1967). The Production of Human Capital and the Life Cycle of Earnings, *Journal of Political Economy*, [e-journal] vol. 75, no. 4, pp.352–365, Available Online: https://www.jstor.org/stable/1828596 [Accessed 12 December 2024]

Benzeval, M. & Judge, K. (2001). Income and Health: The Time Dimension, *Social Science & Medicine*, [e-journal] vol. 52, no. 9, pp.1371–1390, Available Online: https://www.sciencedirect.com/science/article/pii/S0277953600002446?casa_token=B6JqOgV71UMAAAAA:f5huBd5JV6jcSMAAEB3Y3TNH2pLzeA5fQm8qb0FHtrNqW2MDJFy87q8--jmCoJ9soEP9iEw_UA [Accessed 21 October 2024]

Bloom, D. E., Kuhn, M. & Prettner, K. (2019). Health and Economic Growth, *Oxford Research Encyclopedia of Economics and Finance*

Bloom, D. E. & Williamson, J. G. (1998). Demographic Transitions and Economic Miracles in Emerging Asia, *The World Bank Economic Review*, vol. 12, no. 3, pp.419–455

Bolin, K. (2018). Physical Inactivity: Productivity Losses and Healthcare Costs 2002 and 2016 in Sweden, *BMJ Open Sport & Exercise Medicine*, [e-journal] vol. 4, no. 1, p.e000451, Available Online: https://bmjopensem.bmj.com/content/4/1/e000451 [Accessed 26 December 2024]

Bolin, K., Gip, C., Mörk, A.-C. . & Lindgren, B. (2009). Diabetes, Healthcare Cost and Loss of Productivity in Sweden 1987 and 2005-a Register-Based Approach, *Diabetic Medicine*, vol. 26, no. 9, pp.928–934

Bolin, K. & Lindgren, B. (2007). Smoking, Healthcare Cost, and Loss of Productivity in Sweden 2001, *Scandinavian Journal of Public Health*, vol. 35, no. 2, pp.187–196

Bolin, K., Lundgren, A., Berggren, F. & Källén, K. (2011). Epilepsy in Sweden: Health Care Costs and Loss of Productivity—a Register-Based Approach, *The European Journal of Health Economics*, [e-journal] vol. 13, no. 6, pp.819–826, Available Online: https://link.springer.com/article/10.1007/s10198-011-0361-8 [Accessed 26 December 2024]

Clark, T. S. & Linzer, D. A. (2014). Should I Use Fixed or Random Effects?, *Political Science Research and Methods*, vol. 3, no. 2, pp.399–408

Crémieux, P.-Y., Ouellette, P. & Pilon, C. (1999). Health Care Spending as Determinants of Health Outcomes, *Health Economics*, [e-journal] vol. 8, no. 7, pp.627–639, Available Online: https://onlinelibrary.wiley.com/doi/abs/10.1002/(SICI)1099-1050(199911)8:7%3C627::AID-HEC474%3E3.0.CO;2-8?casa_token=Mf49a4WeQbEAAAAA%3Af1Bwvo9By64lussYLrcKc2aq33_ecSZaU147dloTdLSvhWKsBCUAXxHgd9yfStkYJvkfYd6deREUag [Accessed 17 December 2024]

Creswell, J. W. (2003). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches, 2nd edn, London: Sage Publications Ltd

Driscoll, J. C. & Kraay, A. C. (1998). Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data, *The Review of Economics and Statistics*, [e-journal] vol. 80, no. 4, pp.549–560, Available Online: https://www.jstor.org/stable/2646837 [Accessed 10 December 2024]

Ecob, R. & Davey Smith, G. (1999). Income and Health: What Is the Nature of the Relationship?, *Social Science & Medicine*, [e-journal] vol. 48, no. 5, pp.693–705, Available Online: https://www.sciencedirect.com/science/article/pii/S0277953698003852?casa_token=rn9zaxGQthoAAAAA:UP68QQXKWkkaGfLTLfZajK9-XQFfcaoeXCmXPUmWR7NjDMmd9NipUvIGhFuaSytqarOA8jXRhw [Accessed 21 October 2024]

Ehrlich, I. & Chuma, H. (1990). A Model of the Demand for Longevity and the Value of Life Extension, *Journal of Political Economy*, [e-journal] vol. 98, no. 4, pp.761–782, Available Online: https://www.journals.uchicago.edu/doi/10.1086/261705 [Accessed 11 December 2024]

Ettner, S. L. (1996). New Evidence on the Relationship between Income and Health, *Journal of Health Economics*, [e-journal] vol. 15, no. 1, pp.67–85, Available Online: https://www.sciencedirect.com/science/article/pii/0167629695000321?via%3Dihub [Accessed 21 October 2024]

Feng, C., Wang, H., Lu, N., Chen, T., He, H., Lu, Y. & Tu, X. (2014). Log-Transformation and Its Implications for Data Analysis., *Shanghai archives of psychiatry*, [e-journal] vol. 26, no. 2, pp.105–109, Available Online: http://europepmc.org/articles/PMC4120293 [Accessed 6 January 2025]

Flodin, P., Allebeck, P., Gubi, E., Burström, B. & Agardh, E. (2023). Income-Based Differences in Healthcare Utilization in Relation to Mortality in the Swedish Population between 2004–2017: A Nationwide Register Study, *PLOS Medicine*, [e-journal] vol. 20, no. 11, pp.e1004230–e1004230, Available Online: https://pubmed.ncbi.nlm.nih.gov/37971955/ [Accessed 15 December 2024]

Folkhälsomyndigheten. (2023). Folkhälsan I Sverige, Available Online: https://www.folkhalsomyndigheten.se/contentassets/a448b27d603c44f590fc1aff741b0d5d/folkhalsan-sverige-arsrapport-2023.pdf [Accessed 13 January 2025]

Fritzell, J., Nermo, M. & Lundberg, O. (2004). The Impact of Income: Assessing the Relationship between Income and Health in Sweden, *Scandinavian Journal of Public Health*, vol. 32, no. 1, pp.6–16

Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health, *Journal of Political Economy*, [e-journal] vol. 80, no. 2, pp.223–255, Available Online: https://www.jstor.org/stable/1830580 [Accessed 10 December 2024]

Grossman, M. (2000). Handbook of Health Economics, in A. J. Culyer & J. P. Newhouse (eds), *Handbook of Health Economics*, [e-book] Elsevier, pp.347–408, Available Online: https://cphs.huph.edu.vn/uploads/anpham/Handbook_of_Health_Economics__Volume_1A.pdf [Accessed 11 December 2024]

Grossman, M. (2017). The Demand for Health, Columbia University Press

Hall, R. E. & Jones, C. I. (2007). The Value of Life and the Rise in Health Spending, *The Quarterly Journal of Economics*, [e-journal] vol. 122, no. 1, pp.39–72, Available Online: https://academic.oup.com/qje/article/122/1/39/1924761?login=true [Accessed 26 December 2024]

Hanck, C., Arnold, M., Gerber, A. & Schmelzer, M. (2024). 6.1 Omitted Variable Bias | Introduction to Econometrics with R, *Www.econometrics-With-R.org*, Available

Online: https://www.econometrics-with-r.org/6.1-omitted-variable-bias.html [Accessed 8 January 2025]

Hoechle, D. (2007). Robust Standard Errors for Panel Regressions with Cross-Sectional Dependence, *The Stata Journal: Promoting communications on statistics and Stata*, [e-journal] vol. 7, no. 3, pp.281–312, Available Online: https://journals.sagepub.com/doi/10.1177/1536867X0700700301 [Accessed 10 December 2024]

Hyndman, R. J. & Athanasopoulos, G. (2018). Forecasting : Principles and Practice, 2nd edn, Heathmont, Vic.: Otexts

Institute For Research On Public Policy & Centre For The Study Of Living Standards. (2002). The Review of Economic Performance and Social Progress, 2002 : Towards a Social Understanding of Productivity, edited by A. Sharpe, F. St-Hilaire, & K. G. Banting, Montréal, Quebec: Institute For Research On Public Policy

Jacobson, L. (2000). The Family as Producer of Health — an Extended Grossman Model, *Journal of Health Economics*, [e-journal] vol. 19, no. 5, pp.611–637, Available Online: https://doi.org/10.1016/S0167-6296(99)00041-7 [Accessed 13 December 2024]

Karaman, S., Ürek, D., Demir Bilgin, İ., Uğurluoğlu, Ö. & Işık, O. (2020). The Impact of Healthcare Spending on Health Outcomes: New Evidence from OECD Countries, *Erciyes Medical Journal*, [e-journal] vol. 42, no. 2, pp.218–222, Available Online: https://jcpres.com/storage/upload/pdfs/EMJ_42_2_218_222.pdf [Accessed 17 December 2024]

Kolada. (2024). Kolada, *Www.kolada.se*, Available Online: https://www.kolada.se/verktyg/fri-sokning/ [Accessed 10 November 2024]

Loeppke, R., Taitel, M., Haufle, V., Parry, T., Kessler, R. C. & Jinnett, K. (2009). Health and Productivity as a Business Strategy: A Multiemployer Study, *Journal of Occupational and Environmental Medicine*, vol. 51, no. 4, pp.411–428

Lorentzen, P., McMillan, J. & Wacziarg, R. (2008). Death and Development, *Journal of Economic Growth*, [e-journal] vol. 13, no. 2, pp.81–124, Available Online: https://doi.org/10.1007/s10887-008-9029-3 [Accessed 17 December 2024]

Martin, S., Rice, N. & Smith, P. C. (2008). Does Health Care Spending Improve Health Outcomes? Evidence from English Programme Budgeting Data, *Journal of Health Economics*, vol. 27, no. 4, pp.826–842

Mincer, J. (1974). Schooling, Experience, and Earnings, [e-book] New York: National Bureau Of Economic Research, Available Online: https://www.nber.org/books-and-chapters/schooling-experience-and-earnings [Accessed 13 December 2024]

Mitchell, R. J. & Bates, P. (2011). Measuring Health-Related Productivity Loss, *Population Health Management*, [e-journal] vol. 14, no. 2, pp.93–98, Available Online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3128441/ [Accessed 22 October 2024]

Murel, J. & Kavlakoglu, E. (2023). What Is Multicollinearity? | IBM, *Www.ibm.com*, Available Online: https://www.ibm.com/topics/multicollinearity [Accessed 6 December 2024]

Muurinen, J.-M. (1982). Demand for Health, *Journal of Health Economics*, [e-journal] vol. 1, no. 1, pp.5–28, Available Online: https://doi.org/10.1016/0167-6296(82)90019-4 [Accessed 11 December 2024]

Neovius, K., Rehnberg, C., Rasmussen, F. & Neovius, M. (2012). Lifetime Productivity Losses Associated with Obesity Status in Early Adulthood, *Applied Health Economics and Health Policy*, [e-journal] vol. 10, no. 5, pp.309–317, Available Online: https://link.springer.com/article/10.1007/BF03261865 [Accessed 26 December 2024]

Newhouse, J. P. (1977). Medical-Care Expenditure: A Cross-National Survey, *The Journal of Human Resources*, [e-journal] vol. 12, no. 1, p.115, Available Online: https://www.jstor.org/stable/145602 [Accessed 10 November 2024]

Nordic co-operation. (2025). Healthcare System in Sweden | Nordic Cooperation, *Www.norden.org*, Available Online: https://www.norden.org/en/info-norden/healthcare-system-sweden [Accessed 13 January 2025]

OECD. (2013). Mental Health and Work: Sweden, OECD Publishing, Available Online: http://dx.doi.org/10.1787/9789264188730-en [Accessed 26 December 2024]

OECD/European Observatory on Health Systems and Policies. (2023). Sweden: Country Health Profile 2023, State of Health in the EU, Paris/European Observatory on Health Systems and Policies, Brussels: OECD Publishing, Available Online: https://health.ec.europa.eu/document/download/38953059-f96f-48d4-8998-9e3e5db306ba_en?filename=2023_chp_se_english.pdf [Accessed 14 January 2025]

Ried, W. (1998). Comparative Dynamic Analysis of the Full Grossman Model, *Journal of Health Economics*, [e-journal] vol. 17, no. 4, pp.383–425, Available Online: https://doi.org/10.1016/S0167-6296(97)00038-6 [Accessed 11 December 2024]

Rosenman, R., Tennekoon, V. & Hill, L. G. (2011). Measuring Bias in Self-Reported Data, *International Journal of Behavioural and Healthcare Research*, [e-journal] vol. 2, no. 4, pp.320–332, Available Online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4224297/ [Accessed 26 December 2024]

Rowntree, B. S. & Orr, J. (1936). Food, Health and Income. Report on a Survey of Adequacy of Diet in Relation to Income., *The Economic Journal*, [e-journal] vol. 46, no. 184, p.749, Available Online: https://www.jstor.org/stable/2224711 [Accessed 26 December 2024]

Sekaran, U. & Bougie, R. (2016). Research Methods for Business: A Skill-Building Approach, 7th edn, Chichester, West Sussex, United Kingdom: John Wiley & Sons

Singh, S. R. (2014). Public Health Spending and Population Health, *American Journal of Preventive Medicine*, vol. 47, no. 5, pp.634–640

Socialstyrelsen. (2019). About the Swedish Healthcare System, *Socialstyrelsen*, Available Online: https://www.socialstyrelsen.se/en/about-us/healthcare-for-visitors-to-sweden/about-the-swedish-healthcare-system/ [Accessed 13 January 2025]

Socialstyrelsen. (2023). Statistik, *Socialstyrelsen*, Available Online: https://www.socialstyrelsen.se/statistik-och-data/statistik/ [Accessed 20 November 2024]

Socialstyrelsen. (2024a). Statistikdatabaser - Dödsorsaksstatistik - Val, *Sdb.socialstyrelsen.se*, Available Online: https://sdb.socialstyrelsen.se/if_dor/val.aspx [Accessed 15 November 2024]

Socialstyrelsen. (2024b). Stora Skillnader I Tillgång till Vårdpersonal I Landet – Nya Arbetssätt Kan Behövas I Framtiden, *Socialstyrelsen*, Available Online: https://www.socialstyrelsen.se/om-socialstyrelsen/pressrum/press/stora-skillnader-i-tillgang-till-vardpersonal-i-landet--nya-arbetssatt-kan-behovas-i-framtiden/ [Accessed 13 January 2025]

Socialstyrelsen. (2024c). Statistik Om Tillgänglighet I Hälso- Och Sjukvård, *Socialstyrelsen*, Available Online: https://www.socialstyrelsen.se/statistik-och-data/statistik/alla-statistikamnen/tillganglighet-i-halso-och-sjukvard/ [Accessed 13 January 2025]

STATA. (2023). Stata | Why Stata?, *Stata.com*, Available Online: https://www.stata.com/why-use-stata/

Statistiska Centralbyrån. (2022). Privata Företag Utförde 18 Procent Av Verksamheten Inom Vård, Skola Och Omsorg 2022, *Statistikmyndigheten SCB*, Available Online: https://www.scb.se/hitta-statistik/statistik-efter-amne/offentlig-ekonomi/finanser-for-den-kommunala-sektorn/finansiarer-och-utforare-inom-varden-skolan-och-omsorgen/pong/statistiknyhet/finansiarer-och-utforare-inom-vard-skola-och-omsorg-2022/ [Accessed 13 January 2025]

Statistiska Centralbyrån. (2023a). Genomsnittlig Grund- Och Månadslön Samt Kvinnors Lön I Procent Av Mäns Lön Efter Region, Sektor, Yrke (SSYK 2012) Och Kön. År 2023-Statistikdatabasen, *Statistikdatabasen*, Available Online: https://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START__AM__AM0110__AM0110A/LonYrkeRegion4AN/?loadedQueryId=151293&timeType=top&timeValue=1 [Accessed 15 November 2024]

Statistiska Centralbyrån. (2023b). Bruttoregionprodukt (BRP), Sysselsatta Och Löner (ENS2010) Efter Region (Län, Riksområde). År 2000 - 2023-Statistikdatabasen, *Statistikdatabasen*, Available Online: https://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START__NR__NR0105__NR0105A/NR0105ENS2010T01A/ [Accessed 13 November 2024]

Statistiska Centralbyrån. (2024a). PxWeb - Välj Tabell, *Www.statistikdatabasen.scb.se*, Available Online: https://www.statistikdatabasen.scb.se/pxweb/sv/ssd/ [Accessed 13 November 2024]

Statistiska Centralbyrån. (2024b). Medellöner I Sverige, *Statistiska Centralbyrån*, Available Online: https://www.scb.se/hitta-statistik/sverige-i-siffror/utbildning-jobb-och-pengar/medelloner-i-sverige/ [Accessed 13 January 2025]

Sveriges Kommuner och Regioner. (2023a). Vanliga Frågor Om Utjämningssystemet, *Skr.se*, Available Online: https://skr.se/skr/ekonomijuridik/ekonomi/utjamningssystem/vanligafragoromutjamningssystemet.11892.html [Accessed 23 January 2025]

Sveriges Kommuner och Regioner. (2023b). Utökad Uppföljning Av Specialiserad Vård, *Skr.se*, Available Online: https://skr.se/vantetiderivarden/vantetidsstatistik/utokaduppfoljningavspecialiseradvard.43536.html [Accessed 7 January 2025]

The Health Foundation. (2021). Relationship between Income and Health, *The Health Foundation*, Available Online: https://www.health.org.uk/evidence-hub/money-and-resources/income/relationship-between-income-and-health [Accessed 17 January 2025]

Thomson, K., Williams, L., Malam, K. & West, E. (2019). Measuring Gross Regional Product, Available Online: https://www.bitre.gov.au/sites/default/files/Measuring_Gross_Regional_Product-Information_Sheet_100.pdf [Accessed 6 December 2024]

Wilcox, R. R. (2003). Probability and Related Concepts, in *Applying Contemporary Statistical Techniques*, [e-book] Elsevier BV, pp.15–53, Available Online: https://www.sciencedirect.com/book/9780127515410/applying-contemporary-statistical-techniques [Accessed 6 December 2024]

Wilson, J. (2010). Essentials of Business Research a Guide to Doing Your Research Project., London: Sage Publications

Wolf, K. (2010). Making the Link between Health and Productivity at the Workplace ―a Global Perspective, *Industrial Health*, [e-journal] vol. 48, no. 3, pp.251–255, Available Online: https://pubmed.ncbi.nlm.nih.gov/20562499/ [Accessed 26 December 2024]

Wooldridge, J. M. (2012). Introductory Econometrics : A Modern Approach., 5th edn, [e-book] Mason Cengage Learning, Available Online: https://www.unife.it/economia/lm.economics/lectures/applied-econometrics/materiale-didattico-2018-2019/Jeffrey_M._Wooldridge_Introductory_Econometrics_A_Modern_Approach__2012.pdf [Accessed 6 December 2024]

Yngwe, M. A. (2001). The Role of Income Differences in Explaining Social Inequalities in Self Rated Health in Sweden and Britain, *Journal of Epidemiology & Community Health*, vol. 55, no. 8, pp.556–561

Zweifel, P. (2012). The Grossman Model after 40 Years, *The European Journal of Health Economics*, [e-journal] vol. 13, no. 6, pp.677–682, Available Online: https://doi.org/10.1007/s10198-012-0420-9 [Accessed 12 December 2024]

Appendix A

Table 7: Histogram Comparison Before and After Log-Transformation, Mean Income Per Capita

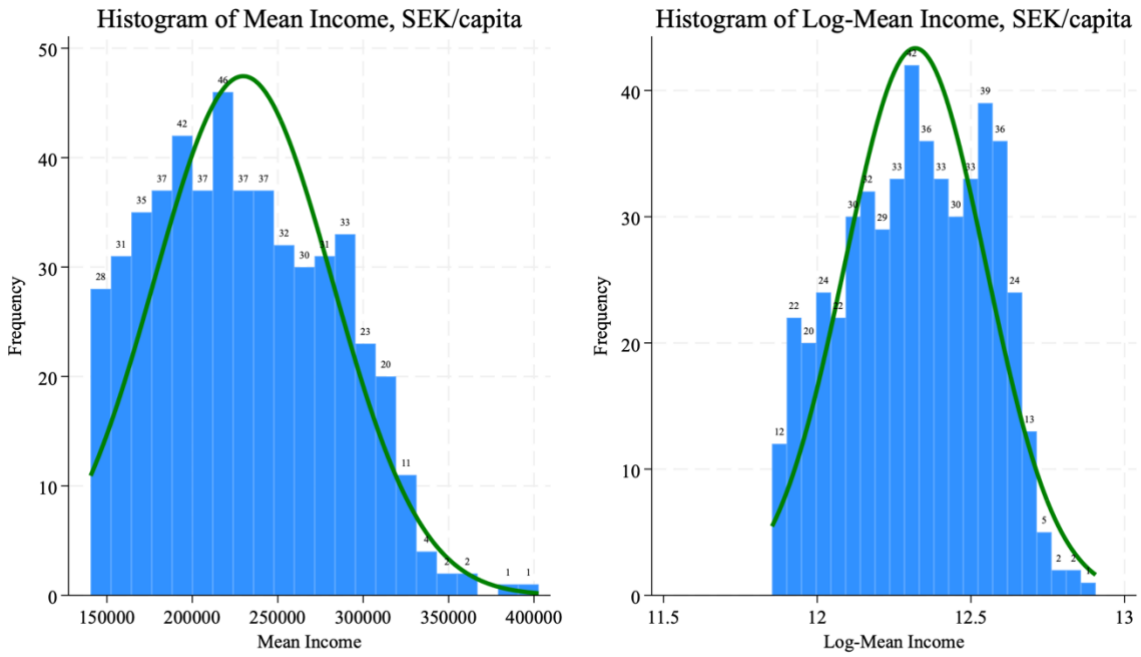


Table 8: Histogram Comparison Before and After Log-Transformation, Deaths Attributable to Illness Per Capita

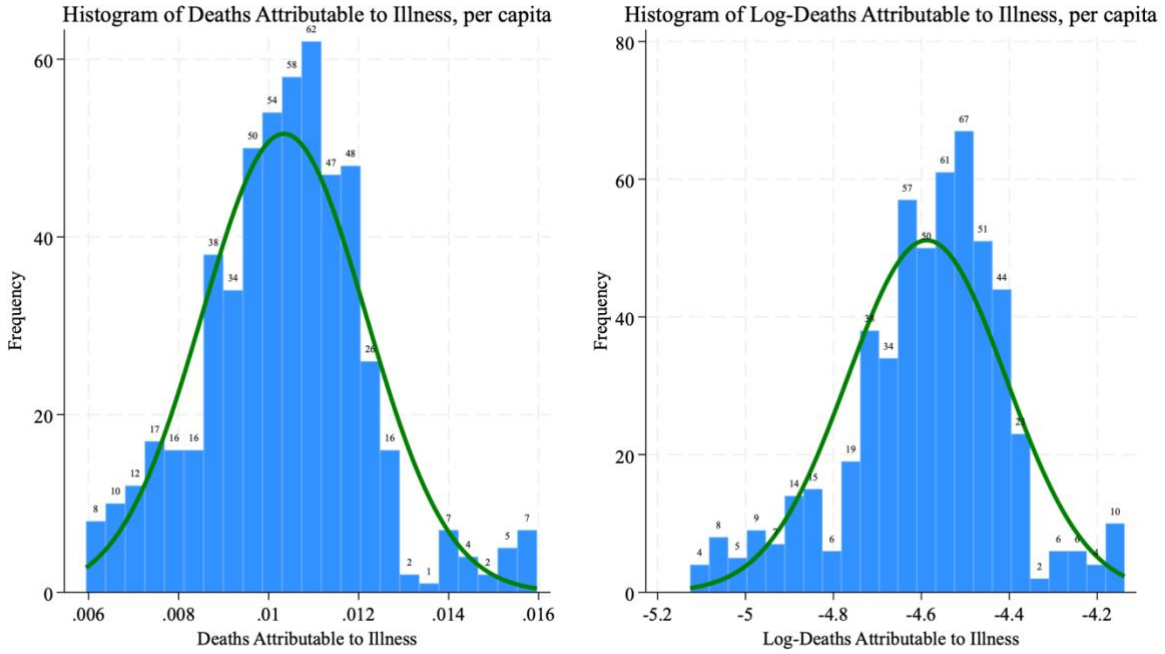


Table 9: Histogram Comparison Before and After Log-Transformation, Healthcare Spending Per Capita

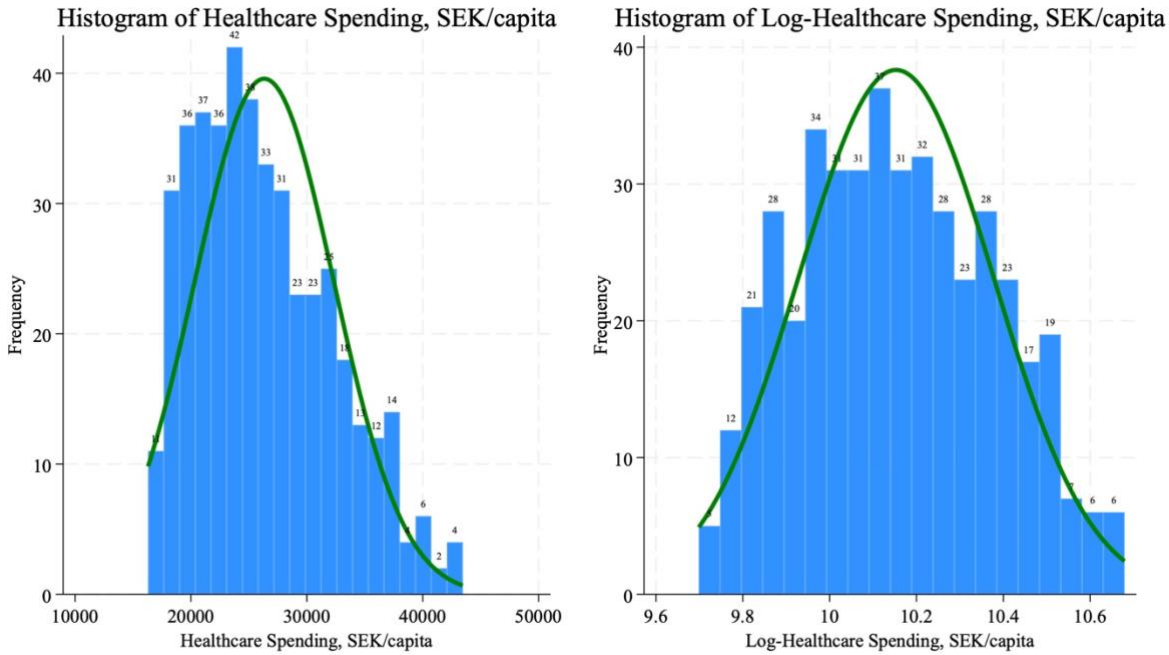
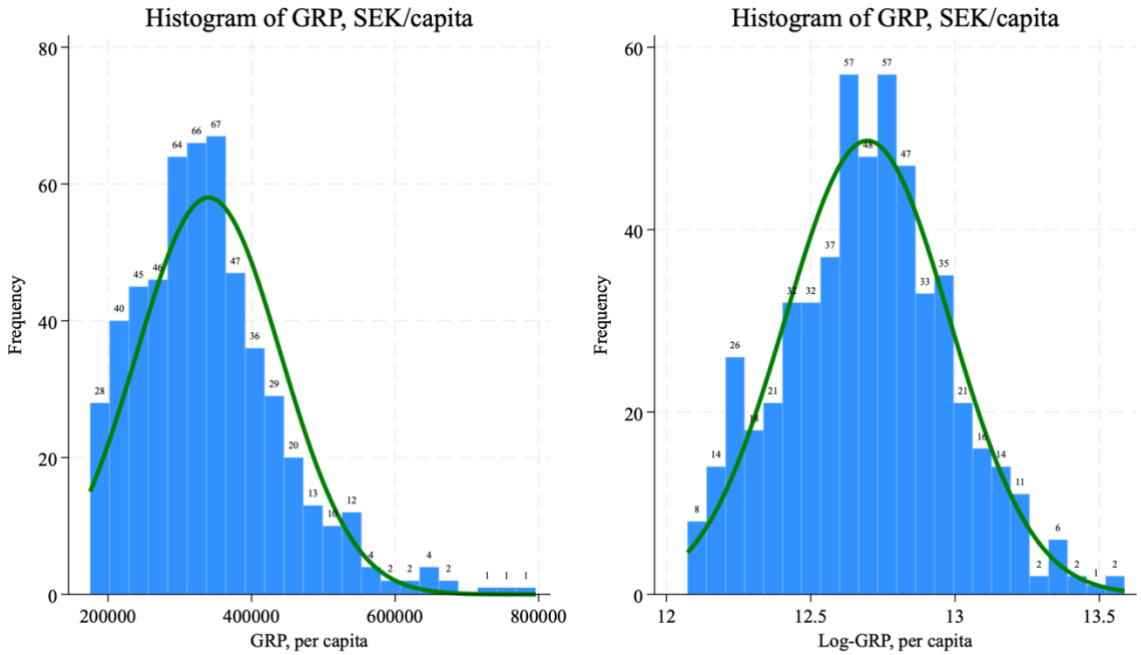


Table 10: Histogram Comparison Before and After Log-Transformation, GRP Per Capita



Appendix B

Equation Specification for Hypotheses

Equations used for each of the hypotheses. The controll variables are the same for each hypothesis and are mean age, percentage of women and percentage of higher education in the regional population.

$H_{1.1}$: Higher income is associated with lower deaths attributable to illnesses per capita within Swedish regions.

$$log(MORT_{it}) = \beta_0 + \beta_1\big(log(INC_{it})\big) + \sum_{k=1}^{K}\beta_k(CV_{kit}) + \alpha_i + \delta_t + \epsilon_{it} \qquad (6)$$

$H_{1.2}$: Higher deaths attributable to illnesses per capita are associated with lower income within Swedish regions.

$$log(INC_{it}) = \beta_0 + \beta_1\big(log(MORT_{it})\big) + \sum_{k=1}^{K}\beta_k(CV_{kit}) + \alpha_i + \delta_t + \epsilon_{it} \qquad (7)$$

H_2: Higher income is associated with lower healthcare spending per capita within Swedish regions.

$$log(HCS_{it}) = \beta_0 + \beta_1\big(log(INC_{it})\big) + \sum_{k=1}^{K}\beta_k(CV_{kit}) + \alpha_i + \delta_t + \epsilon_{it} \qquad (8)$$

H_3: Higher healthcare spending per capita is associated with lower deaths attributable to illness per capita

$$log(MORT_{it}) = \beta_0 + \beta_1\big(log(HCS_{it})\big) + \sum_{k=1}^{K}\beta_k(CV_{kit}) + \alpha_i + \delta_t + \epsilon_{it} \qquad (9)$$

H_4: Higher deaths attributable to illness per capita are associated with lower regional productivity.

$$log(GRP_{it}) = \beta_0 + \beta_1\big(log(MORT_{it})\big) + \sum_{k=1}^{K}\beta_k(CV_{kit}) + \alpha_i + \delta_t + \epsilon_{it} \qquad (10)$$

Appendix C

Sensitivity Analysis

Table 11: Sensitivity Analysis Comparing Three Models with Unadjusted Variables

	Model		No Control Variables (Eq. 3)	Control Variables excl. Time (Eq. 4)	Control Variables incl. Time (Eq. 5)
Hypothesis	**Dependent**	**Independent**	**Coefficient (p-value)**	**Coefficient (p-value)**	**Coefficient (p-value)**
H1.1	Mortality	Income	-8.49e-09 (0.00)***	6.56e-09 (0.09)*	4.56e-09 (0.194)
H1.2	Income	Mortality	-6.56e+07 (0.00)***	3819805 (0.1)*	423136 (0.13)
H2	HC Spending	Income	0.122 (0.00)***	0.125 (0.00)***	-0.087 (0.00)***
H3	Mortality	HC Spending	-5.38e-08 (0.00)***	2.16e-08 (0.00)***	3.96e-08 (0.04)**
H4	Productivity	Mortality	-1.03e+08 (0.00)***	1.15e+07 (0.09)*	7836609 (0.05)**

**Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level*

Table 12: Sensitivity Analysis Comparing Different Methods with Adjusted Variables

	Model		Xtreg	Random Effects	Post 2010-data	Excluding highest & lowest income region
Hypo-theses	**Depen-dent**	**Indepen-dent**	**Coefficient (p-value)**	**Coefficient (p-value)**	**Coefficient (p-value)**	**Coefficient (p-value)**
H1.1	Morta-lity	Income	.312 (0.12)	0.175 (0.05)**	-.046 (0.88)	.521 (0.00)***

H1.2	Income	Mortality	.020 (0.12)	0.080 (0.12)	-.001 (0.81)	.0037 (0.00)***
H2	HC Spen-ding	Income	-.136 (0.41)	0.022 (0.15)	.028 (0.86)	-.155 (0.10)*
H3	Morta-lity	HC Spending	-.060 (0.36)	-0.051 (0.58)	-.042 (0.19)	-.040 (0.46)
H4	Produc-tivity	Mortality	.150 (0.02)**	0.147 (0.21)	.176 (0.01)	.142 (0.07)*

**Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level*

Appendix D

The Usage of Generative AI

Tools:

ChatGPT

Usage:

Generative AI (ChatGPT) has been utilized in this thesis to the greatest extent for language purposes. It has been used for synonyms throughout the entire text (with the exception of the first page, titles and subtitles) where formal synonyms and other ways of describing text were requested. For instance, in the result section, re-formulating of sentences such as “the results find” or “is statistically significant” has been supported by generative AI. It has also been used to find grammatical errors.

In addition, generative AI has been used to assist the author in using the statistical software (Stata). This was done by receiving the correct code needed for, for instance, performing unit-root tests and to create two histograms in one graph (Appendix A). However, the idea and performance of the analysis and diagnostics is not AI generated.

Lastly, while all analyses, diagnostics and methodical approaches have been conducted using the author’s knowledge and published material, some explanations found in books and articles have been clarified using generative AI by saying, “Explain xxxx simply such that someone who has no knowledge about it could understand”. This has been done for the Methodology, in particular for the diagnostics (section 5 and 5.5.2.). However, AI has not performed the tests nor told how to conduct them. It has solely assisted in clarifying difficult concepts.