Erasmus
School of
Economics

ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Bachelor Thesis Strategy Economics

The effect of income on healthcare consumption in the Netherlands

Name: Bas Wortman
Student number: 505135

Supervisor: Bhaskarabhatla, A.S.
Second assessor: Ours, J.C. van

Date final version: 30-7-2021

Abstract

This research uses health and income data from 1800 Dutch individuals over the period 2019-2020. It attempts to prove a relation between income and healthcare consumption. This is done by performing multiple OLS regressions, negative binomial regression and tobit analysis on the dataset. It is concluded that there is no relation between income and healthcare consumption. This is partly as the relation between income and visits to a doctor ends up being rather insignificant. However, some other very significant relations are found regarding healthcare consumption.

Table of contents

Introduction

In this current world where diseases like COVID spread around, it is once again proven that a well-developed healthcare system is an important asset of a developed country. Where some developed countries such as the United States opt for private healthcare, other countries like the United Kingdom choose to provide 'free' healthcare for all citizens. The Netherlands is a country that has decided to implement universal healthcare as well. A system based on risk equalization was implemented by making insurance compulsory for all citizens above 18 years of age.

According to this system, healthcare should be equal for all individuals in the Netherlands. No matter their age, they should be able to receive the best care possible. An individual's income should not play a part in this. However, how equal is this system? To prevent abuse of the healthcare system, insurance companies charge a government mandated deductible that citizens will have to cover themselves before they are eligible for a reimbursement from their insurance. This raises the question if this could be a barrier for low-income citizens to see a doctor.

This raises the main research question of this thesis:
To what extent does net income per year have an effect on the number of times an individual visits a doctor?

To answer this question, there are multiple sub-questions that form the base of this research. Data will be used to analyze these sub questions, and they are as follows:

Does a individuals' income affect how much of an effect health has on their daily life?

Does a individuals' income affect how much they visit a general practitioner?

Does a individuals' income affect how much they visit a specialist doctor in a hospital?

Does a individuals' income affect how high their voluntary deductible is?

The remainder of this paper will consist of several parts. It will start with an overview from relevant and related literature. This will give a broader understanding as to why the system is designed in this way. After that, the focus will be on the data & methodology section where the data and how it will be analyzed will be discussed. Finally, the results will be analyzed and from that, the conclusions will be drawn. We will relate the answer of these sub questions to the main question and using that, a final conclusion to our main question will be drawn. As always, the paper will include a discussion where improvements for future research will be discussed.

Literature Review

The Dutch government, also known as Rijksoverheid, has mandated that everyone who lives or works in the Netherlands must take out basic health insurance. The national government determines what is included in the basic package. The basic insurance covers standard care provided by, for example, a general practitioner, hospital or pharmacy. According to the Dutch government: “The Dutch healthcare system guarantees good and accessible healthcare for everyone.” (Rijksoverheid, 2021). A logical question would be to ask if this is actually the case. Is healthcare really that equal?

Earlier research on this has been done by Davis, Doty and Ho in the year 2005. In their research, Davis et al. used a dataset from the 1970’s where individuals were randomly assigned a different deductible of either zero, 25 percent, 50 percent, or 95 percent coinsurance. This experiment is referred to as the RAND Health Insurance Experiment. This study found that individuals with a higher coinsurance had a lower utilization of physician services and a lower amount of hospital stays, and on top of that, they had an overall lower total health spending. This would prove that there is a negative effect between coinsurance and hospital visits.

According to this literature, the amount somebody pays out of pocket could have an influence on healthcare consumption. In the Netherlands, by law, each insured individual has a pre-defined deductible of €385, which can be increased up to €885 per year in return for a lower monthly premium. (Zorgwijzer, 2021)

The deductible exists for good reasons, it obliges individuals to deal more consciously with the amount of healthcare they receive. This is as the costs are not only for the insurance, but also your own. In this paper, there is a distinction between two types of doctors. These are general practitioners and medical specialists in a hospital. Each individual has a general practitioner completely covered by insurance, which means there are no extra costs for visiting one. Apart from that, there are also the medical specialists in a hospital, if one decides to visit a specialist doctor, there is a copayment in almost all cases.

In this paper, we will analyze how much effect an individual's income actually has on the amount of healthcare they consume. This paper will analyze what has been concluded in the RAND Health Insurance Experiment as written in the paper by and Davis, K., Doty, M., & Ho, A (2005) but in a different setting.

There are some significant differences between the situation in the paper and our dataset. An important difference is that this thesis will use data from Dutch individuals from recent years. This will be able to give more accurate and recent results than data used by Doty & Ho (2005) from the 1970's. It is also important to take into account that the RAND Health Insurance Experiment was performed in the United States. In the United States health insurance is voluntary and not government regulated, while it is in the Netherlands. Therefore, it could be expected that our analysis might reward different findings and results.

A later study (2006) by Van Doorslaer, Masseria, & Koolman concerning inequalities in access to medical care by income in developed countries reaffirmed income-related differences in healthcare consumption. This study was performed in countries that are members of the Organisation for Economic Co-operation and Development (OECD). It was found that there was a negative relation between income and doctors' visits. It suggested that individuals in the bottom quintile, on average, visited the doctor about 50% more times. This would suggest a relation between income and doctors' visits.

This paper further distinguishes between 2 types of healthcare. These are primary care (GP) and secondary care (medical specialists). It is concluded that visits to a general practitioners are generally unrelated to income, as they seem to be equal between all income groups. However, a significant pro-rich inequality is found in the likelihood of contacting a specialist. It is concluded that higher-income people are more likely to seek specialist care than their lower-income counterparts.

An interesting find is that the degree of pro-rich inequity in healthcare consumption is found to be highest in the United States and Mexico, which are the only 2 countries without universal healthcare coverage of their population in this study.

In the paper “Socioeconomic inequalities in waiting times for primary care across ten OECD countries” by Martin, Siciliani and Smith (2020), another possible explanation for differences in healthcare consumption is uncovered. In the paper a possible relation between income and waiting times is researched. This research was done based on data from 10 OECD countries (Australia, Canada, France, Germany, Netherlands, New Zealand, Norway, Sweden, Switzerland, and the United Kingdom). Their results suggest a socio-economic inequality in waiting times for primary care in several countries. It is suggested that this inequality is a result of better-off individuals being able to afford to pay fees charged by private providers or the fact that they are able to travel to providers with shorter waiting times. However, something to note is that the Netherlands was one of the countries where no significant difference was found.

When going to more recent times, there are also some interesting studies that relate to COVID. Research by Oronce, Scannel, Kawachi & Tsugawa (2020) in the United States has shown that there is also a relation between income inequality and COVID infections / mortality. In their research they made use of a state-level Gini index, which is an index for income inequality and compared that to the amount of COVID infections and deaths.

It was observed that states with a larger income inequality, had significantly more infections and deaths than states with a lower income inequality. The researchers concluded that this was most likely a result of economic segregation, decreased social mobility, and lower access to medical care in the larger lower-class in states with high income inequality.

As can be seen in articles above, income is related to healthcare in multiple different ways. This can be copayments, healthcare distribution, waiting times and much more. In this paper, we will continue this research by analyzing the relation between income and healthcare consumption in the Netherlands in 2019. This is an interesting point of research as all the other data-driven studies use outdated data that is not from the Netherlands.

Data

To answer our research questions. We will make use of Dutch empirical data over 2019, collected by the LISS panel over the period of November and December 2020. This dataset contains large amount of health, income and education related data from over 5000 individuals in the Netherlands. These variables are as follows:

- ID (The unique identifier to track 1 individual trough the multiple datasets)
- Family Physician (The amount of times an individual has visited a family physician in the past 12 months)
- Medical Specialist (The amount of times an individual has visited a medical specialist at a hospital in the past 12 months)
- Handicap (Dummy variable: If the individual suffers from any kind of long-standing disease, affliction or handicap, or suffers from the consequences of an accident. The answer was either yes (1) or no (0))
- Physical Health (Dummy variable: To which extent physical health or emotional problems hinder an individual's daily activities over the past month, for instance in going for a walk, walking up stairs, dressing itself, washing itself, visiting the toilet. The answers were: 1. Not at all 2. Hardly 3. A bit 4. Quite a lot 5. Very much)
- Taking Medicines (Dummy variable: If the individual has any problems with taking medicines. The answers were: 1. Without any trouble 2 . With some trouble 3. With a lot of trouble 4. Only with an aid or the help of others 5. Not able to)
- Health Self (Dummy variable: How the induvial would rate their own health on a 1-5 scale. The answers were. 1. Poor 2. Moderate 3. Good 4. Very good 5. Excellent)
- Income (Individuals total net income of the household over the period from 1 January 2019 to 31 December 2019)
- Income SQ (Square term of income)
- Deductible (Individuals voluntary deductible in 2020 in euro's)
- Education (Rated on a 1-27 scale where 1 is the lowest, and 27 is the highest form of education)

After all the data was carefully selected there were a total of 1800 feasible observations that could be used for empirical analysis. A large amount of observations was removed as the individuals reported a negative income or did not report their income. A small amount of observations was also removed as respondents had non-existing answer options (e.g. response was 0 when the dummy variable was on a 1-5 scale).

Descriptive Statistics

Table 1: Descriptive statistics from health and income data from 1800 individuals in 2019

Variable	***Obs***	***Mean***	***Std. Dev***	***Min***	***Max***
Family Physician	1800	1,742	2,617	0	48
Medical Specialist	1800	1,167	5,258	0	200
Handicap	1800	0,374	0,480	0	1
Physical Health	1800	1,712	1,025	1	5
Taking Medicines	1800	1,13	0,632	1	5
Health Self	1800	3,114	0,759	1	5
Income	1800	44805,11	198706,6	0	8000000
Deductible	1800	73,111	166,98	0	500
Education	1800	16,148	6,579	1	26

Notes: Table 1 shows the average statistics from all variables that are included in the dataset. It shows the amount of observations, the mean, the standard deviation and the minimum- and maximum of the selected variables.

Hypotheses

Using the literature discussed in this paper, it is possible to make some careful assumptions when it comes to the main- and sub questions. The expected results are in the form of hypotheses:

1. *Does an individuals' income affect how much of an effect health has on their daily life?*

   According to the RAND study mentioned before in the literature section, while income inequality leads to different consumptions in healthcare, the mortality rates remained the same across all groups. There is a possibility for socio-economic inequality in waiting times, which result in different effects on impact of health, but as no significant effects were found in the study by Martin, Siciliani and Smith (2020), a large effect is not expected. Therefore I would expect that an individual's income does not influence how much impact health has on their daily life.

   H1: The net income of an individual has no effect on how much health impacts their daily life

2. *Does an individuals' income affect how much they visit a general practitioner?*

   As discussed before, going to a general practitioner is almost completely covered by your health insurance in the Netherlands. Rich or poor, the costs are the same, they are non-existent. This was further reaffirmed in the study by Doorslaer, Masseria, & Koolman (2006). It was concluded that visits to a general practitioner were generally unrelated to income. Therefore, no significant difference is expected.

   H2: The amount of income has no effect on the amount of visits to a general practitioner

3. *Does a individuals' income affect how much they visit a specialist doctor in a hospital?*

As discussed previously, when visiting most specialist doctors (in a hospital) in the Netherlands, there is a part of the costs that you have to pay yourself, this has to do with your yearly insurance deductible. Because of this, it could be argued that there is a relation between the amount of income a individuals has, and how fast they visit a specialist doctor. This is also reaffirmed by the study of Doorslaer, Masseria, & Koolman (2006) which found a pro-rich inequality in medical specialist visits. Therefore it could be expected that here is a relation between income and specialist doctor visits.

H3: There is a positive relation between income and the amount of specialist doctor visits

4. *Does a individuals' income affect how high their voluntary deductible is?*

For this last sub-question, it could be argued that it is logical that low-income individuals prefer high deductibles in exchange for lower monthly premium, as a way to lower their monthly costs. I would argue that the small reduction in premium is not interesting or relevant for high-income individuals, as health income expenses are relatively much smaller to them.

H4: There is a negative relation between income and a individuals deductible.

Methodology

Is this part of the paper we will go in-depth on how we will analyze our data to get to the answer of our research questions. As each sub-question will have a different method of researching, we will go in-depth on each hypotheses. Finally, the chosen control variables will be discussed.

1. *Does a individuals' income affect how much of an effect health has on their daily life?*

The effect of an individual's income on how much of an effect health has on their daily life will be measured using an ordinary least squares (OLS) regressions. The dependent variable used is 'PhysicalHealth', which is defined as 'the extent to which physical health or emotional problems hinder an individual's daily activities', this variable is a self-assessment of how much their health impacts their daily life. The independent variables will be income, education, general practitioner visits, medical specialist visits, handicap and problems taking medicines. The regression is as follows:

- $PhysicalHealth = \alpha + \beta 1 * Income + \beta 2 * Education + \beta 3 * FamilyPhysician + \beta 4 * MedicalSpecialist + \beta 5 * Handicap + \beta 6 * TakingMedicines + \varepsilon$

Apart from this, we are also testing for any nonlinear effects. This is done by including the square term of income in a separate regression. This regression is as follows:

- $PhysicalHealth = \alpha + \beta 1 * Income + \beta 2 * IncomeSQ + \beta 3 * Education + \beta 4 * FamilyPhysician + \beta 5 * MedicalSpecialist + \beta 6 * Handicap + \beta 7 * TakingMedicines + \varepsilon$

Control variables

The regression will make clear if there is an effect or significant relationship between income and an individuals' Physical Health. In an attempt not to overstate the effect of income on an individual's health rating, control variables will be added. These are, the level of education,

the amount of general practitioner visits 'FamilyPhysician', the amount of visits to a medical specialist in a hospital 'MedicalSpecialist', if the individual suffers from any kind of long-standing disease, affliction or handicap, or suffers from the consequences of an accident 'Handicap', and the degree to which the individual has any problems taking medicines 'TakingMedicines'. The last control variable used is education.

The second and third sub-question are closely related, therefore, they will be discussed in the same paragraph. The sub-questions are as follows:

2. *Does a individuals' income affect how much they visit a general practitioner?*
3. *Does a individuals' income affect how much they visit a specialist doctor in a hospital?*

To analyze them, we will use the following OLS regressions:

2. $FamilyPhysician = \alpha + \beta1 * Income + \beta2 * Education + \beta3 * Handicap + \beta4 * PhysicalHealth + \beta5 * TakingMedicines + \varepsilon$
3. $MedicalSpecialist = \alpha + \beta1 * Income + \beta2 * Education + \beta3 * Handicap + \beta4 * PhysicalHealth + \beta5 * TakingMedicines + \varepsilon$

As before, a second regression will be ran to test for any nonlinear effects. This will include the square term of income:

2. $FamilyPhysician = \alpha + \beta1 * Income + \beta2 * IncomeSQ + \beta3 * Education + \beta4 * Handicap + \beta5 * PhysicalHealth + \beta6 * TakingMedicines + \varepsilon$
3. $MedicalSpecialist = \alpha + \beta1 * Income + \beta2 * IncomeSQ + \beta3 * Education + \beta4 * Handicap + \beta5 * PhysicalHealth + \beta6 * TakingMedicines + \varepsilon$

The regressions will make clear if there is an effect or significant relationship between income and healthcare consumption. As there is separate data for visits to the general practitioner and medical specialist visits, it is possible to run individual regressions and this allows us to compare the results. We will once again test for any non-linear effects. Therefore, 2 separate OLS regressions are done for both types of doctor's visits. For both sub questions, the second OLS regression contains the square term of income.

As our dependent variable is a count variable, it is possible to run a negative binomial regression. This is done in an attempt to get to get more accurate and statistically significant results. The negative binomial regression will follow the regression without the square term of income.

Control variables

In the regression the following control variables were chosen:

- *Handicap*
- *Physical Health*
- *Taking Medicines*
- *Education*

These variables are chosen as they are expected to correlate with the amount of doctors' visits. By including these control variables, we attempt to make the regression as accurate as possible.

The last research question is one concerning deductibles. It is as follows:

4. *Does an individuals' income affect how high their voluntary deductible is?*

As deductibles have been discussed numerous times before, it would make sense to perform a regression on them given the data is available. The purpose of this regression is to attempt to identify a relation between income , visits to (any) doctor and a voluntary deductible

4. $Deductible = \alpha + \beta 1 * Income + \beta 2 * Education + \beta 3 * HealthSelf + \beta 4 * FamilyPhysician + \beta 5 * MedicalSpecialist + \varepsilon$

Apart from an OLS regression, it is also possible to run an analysis using tobit, as the voluntary deductible is never negative and can be at most €500,- per individual. The tobit regression will follow the same equation as above, with a lower limit of 0 and an upper limit of €500,-

Control variables

As a higher voluntary deductible is a choice a person makes based on his or her expectations of healthcare costs for that year, the self-reported value of how people rate their health on a 1-5 scale was included as control variable. Income, and visits to any doctor are added as well.

Results

In this part of the paper, we will discuss the results of the regressions as described in the Methodology. These results will be used to answer our research questions.

1. *Does a individuals' income affect how much of an effect health has on their daily life?*

Table 2: The result of a regression of income (x1000), general practitioner visits, medical specialist visits, an individual being handicapped and effort required taking medicines on an individual's physical health.

PhysicalHealth	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Incomex1000	.0002149	.0001057	2.03	0.042	7.60e-06	.0004222
Education	-.0215252	.0032006	-6.73	0.000	-.0278025	-.0152478
FamilyPhysician	.0658014	.0084063	7.83	0.000	.0493142	.0822886
MedicalSpecialist	.0173089	.0041032	4.22	0.000	.0092614	.0253565
Handicap	.7637914	.0451327	16.92	0.000	.6752732	.8523096
TakingMedicines	.1359606	.0331865	4.10	0.000	.0708723	.2010489
_cons	1.482978	.071662	20.69	0.000	1.342428	1.623528

Notes: Table 2 gives the results of the regression as described above. The robust standard errors are next to the coefficients. The total amount of observations (n) is 1800. The stars behind the coefficients refer to the significance level (*p<0.10 , **p<0.05, ***p<0.01)

Table 3: The result of a regression of income (x1000), squared income, general practitioner visits, medical specialist visits, an individual being handicapped and effort required taking medicines on an individual's physical health.

PhysicalHealth	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Incomex1000	-.0003735	.0004142	-0.90	0.367	-.0011858	.0004389
IncomeSQ	7.97e-08	5.43e-08	1.47	0.142	-2.67e-08	1.86e-07
Education	-.0206992	.0032486	-6.37	0.000	-.0270706	-.0143277
FamilyPhysician	.0655077	.008406	7.79	0.000	.0490211	.0819942
MedicalSpecialist	.0172559	.004102	4.21	0.000	.0092107	.0253012
Handicap	.7621684	.0451316	16.89	0.000	.6736522	.8506846
TakingMedicines	.1348138	.033185	4.06	0.000	.0697285	.1998992
_cons	1.495173	.0721181	20.73	0.000	1.353729	1.636617

Notes: Table 3 gives the results of the regression as described above. The robust standard errors are next to the coefficients. The total amount of observations (n) is 1800. The stars behind the coefficients refer to the significance level (*p<0.10 , **p<0.05, ***p<0.01)

To answer the first hypotheses, two regressions are done with 'Physical Health' as dependent variable. As a reminder, this was a dummy variable of: "To which extent physical health or emotional problems hinder an individual's daily activities over the past month, for instance in going for a walk, walking upstairs, dressing itself, washing itself, or visiting the toilet.". Individuals had to respond on a 1-5 scale where 1 was no issues, and 5 was a lot of issues.

The regression without the square term of income (table 2) is 99.99%+ significant for all control variables, and 95%+ significant for our dependent variable income. The second regression which includes the square term of income (table 3) is as significant for the control variables, but much less significant for the dependent variable income. It is therefore possible to conclude that there is no nonlinear effect when it comes to income in our data. As the regression without the square term of income is the most accurate, this will be used to analyze the hypotheses.

According to the model, each €1000 euro difference in income, is responsible for a small amount of 0.0002149 extra issues with Physical Health on a 1-5 scale. As the average income in our dataset is €45000,- euro, this would be an 0,0096705 increase with physical health

problems on this 1-5 scale. This is a very low amount and the income-effect can therefore be seen as very small.

Both the visits to a medical specialist, and a general practitioner have a positive relation on the amount of physical health problems. This is expected as individuals that go to a doctor, usually go for a reason. The same can be said for the dummy for handicapped, it is to be expected that individuals with long-standing diseases, afflictions or handicaps, have more problems with their physical health.

The last significant positive variable is the degree to which an individual has problems taking medicines by themselves. This outcome seems simple to explain as individuals that have problems taking medicines by themselves, are usually more in need of care, and therefore it seems to be expected that they also have more problems with their physical health.

An interesting find is that education seems to be negatively correlated with physical health. This leads us to the conclusion that according to this dataset, a higher educational degree seems to be related to a better physical health.

To conclude our answer to question 1: While there is a 5% significant relation between income and physical health, the effect is so small that it is almost negligible. We can conclude that according to this data, income has almost no effect on an individual's physical health.

2. *Does a individuals' income affect how much they visit a general practitioner?*

Table 4: The result of a regression with income (x1000), education, an individual being handicapped, their physical health, and effort required taking medicines on the amount of general practitioner visits.

FamilyPhysician	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Incomex1000	-.0004994	.0002955	-1.69	0.091	-.001079	.0000802
Education	.0018046	.0090534	0.20	0.842	-.0159517	.0195609
Handicap	.8675802	.133933	6.48	0.000	.6048992	1.130261
PhysicalHealth	.5572744	.0643468	8.66	0.000	.431072	.6834769
TakingMedicines	-.0584399	.0931406	-0.63	0.530	-.2411155	.1242356
_cons	.5186017	.2220966	2.34	0.020	.0830065	.954197

Notes: Table 4 gives the results of the regression as described above. The robust standard errors are next to the coefficients. The total amount of observations (n) is 1800. The stars behind the coefficients refer to the significance level (*p<0.10 , **p<0.05, ***p<0.01).

Table 5: The result of a regression with income (x1000), squared income, education, an individual being handicapped, their physical health, and effort required taking medicines on the amount of general practitioner visits.

FamilyPhysician	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
Incomex1000	-.0013587	.0011579	-1.17	0.241	-.0036296	.0009122
IncomeSQ	1.16e-07	1.52e-07	0.77	0.443	-1.81e-07	4.14e-07
Education	.0029721	.0091813	0.32	0.746	-.015035	.0209793
Handicap	.8662267	.1339599	6.47	0.000	.6034927	1.128961
PhysicalHealth	.5553065	.0644052	8.62	0.000	.4289894	.6816236
TakingMedicines	-.0598676	.0931699	-0.64	0.521	-.2426005	.1228654
_cons	.5389201	.2236938	2.41	0.016	.1001921	.977648

Notes: Table 5 gives the results of the regression as described above. The robust standard errors are next to the coefficients. The total amount of observations (n) is 1800. The stars behind the coefficients refer to the significance level (*p<0.10 , **p<0.05, ***p<0.01).

In this regression, Family Physician is the amount of times an individual visits a general practitioner. As with the previous hypotheses, two separate regressions are done. The regression which does not include a square term of income is much more significant, and therefore we are able to conclude that in this situation, there is, again, no nonlinear effect of income. Therefore, the first regression (table 4) will be used to analyze this hypothesis.

Our main variable of interest, income, is almost 10% significant, but the same situation rises as with the previous regression. While the effect is negative, which means that a higher income results in less visits to a general practitioner, the coefficient for income is very small, to a point where it almost has no effect. Using the average income of €45000,- from our dataset, income is on average only responsible for only 0.022473 less visits to a general practitioner.

Our data shows that an individual's education level, is largely unrelated with the amount of visits to a general practitioner. While the effect is very minor, it is also largely insignificant. However, an individual being handicapped seems to have a big positive impact on the amount of visits to a general practitioner, together with state of the physical health of an individual. A somewhat surprising result is the fact that more problems with medicine intake is negatively

related with going to a doctor. It could be assumed that this surprising relationship is not entirely correct as the significance level is very low.

Table 6: The result of a negative binomial regression with income (x1000), education, an individual being handicapped, their physical health, and effort required taking medicines on the amount of general practitioner visits.

FamilyPhysician	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
Incomex1000	-.0005148	.0003095	-1.66	0.096	-.0011214	.0000919
Education	-.0022772	.0042726	-0.53	0.594	-.0106514	.006097
Handicap	.5052664	.0613515	8.24	0.000	.3850197	.625513
PhysicalHealth	.2496555	.0281477	8.87	0.000	.1944869	.304824
TakingMedicines	-.0574975	.0467891	-1.23	0.219	-.1492025	.0342074
_cons	-.0381415	.1046305	-0.36	0.715	-.2432135	.1669304
/lnalpha	-.2929147	.0670956			-.4244196	-.1614097
alpha	.7460858	.0500591			.6541493	.8509434

Notes: Table 6 gives the results of the regression as described above. The robust standard errors are next to the coefficients. The total amount of observations (n) is 1800. The stars behind the coefficients refer to the significance level (*p<0.10 , **p<0.05, ***p<0.01).

As our dependent variable is a count variable, a negative binomial regression is also performed, to analyze is there are any major differences. All of the significance levels are the same, apart from the constant, which is now highly insignificant. An individual being handicapped has gotten a slightly smaller coefficient together with the state of an individuals physical health, but generally speaking, there are no major differences.

To conclude: Based on both the OLS regression and the negative binominal regression, there seems only a very minor relation between an individual's income and visits to a general practitioner. No nonlinear effects have been found.

3. Does a individuals' income affect how much they visit a specialist doctor in a hospital?

Table 7: The result of a regression with income (x1000), education, an individual being handicapped, their physical health, and effort required taking medicines on the amount of medical specialist in a hospital visits.

MedicalSpecialist	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Incomex1000	-.0003989	.0006126	-0.65	0.515	-.0016005	.0008026
Education	.0234403	.0187692	1.25	0.212	-.0133715	.0602521
Handicap	1.120098	.2776656	4.03	0.000	.5755163	1.66468
PhysicalHealth	.7449235	.1334016	5.58	0.000	.4832846	1.006562
TakingMedicines	.1961169	.1930962	1.02	0.310	-.1826002	.574834
_cons	-1.114039	.4604436	-2.42	0.016	-2.017101	-.2109771

Notes: Table 7 gives the results of the regression as described above. The robust standard errors are next to the coefficients. The total amount of observations (n) is 1800. The stars behind the coefficients refer to the significance level (*p<0.10 , **p<0.05, ***p<0.01)

Table 8: The result of a regression with income (x1000), squared income, education, an individual being handicapped, their physical health, and effort required taking medicines on the amount of medical specialist in a hospital visits.

MedicalSpecialist	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Incomex1000	-.0011883	.0024008	-0.49	0.621	-.0058969	.0035203
IncomeSQ	1.07e-07	3.15e-07	0.34	0.734	-5.10e-07	7.24e-07
Education	.0245128	.0190369	1.29	0.198	-.012824	.0618497
Handicap	1.118855	.2777581	4.03	0.000	.5740912	1.663618
PhysicalHealth	.7431157	.1335404	5.56	0.000	.4812045	1.005027
TakingMedicines	.1948054	.1931823	1.01	0.313	-.1840807	.5736915
_cons	-1.095374	.463816	-2.36	0.018	-2.005051	-.1856973

Notes: Table 8 gives the results of the regression as described above. The robust standard errors are next to the coefficients. The total amount of observations (n) is 1800. The stars behind the coefficients refer to the significance level (*p<0.10 , **p<0.05, ***p<0.01)

As with the regression for visits to a general practitioner, two separate regressions have been done. They are equal, except the second regression includes the square term of income. The results prove that there are no nonlinear effects when it comes to income. This is because the significance levels for income are lower in the regression that includes the square term of income.

Interpreting the results of the OLS regression shows us that the income effect is small and very unsignificant. The same can be said for education, which is also nowhere near significant. Both an individual being handicapped and a worse physical health have a very significant effect on the amount of visits to a medical specialist according to the data. The issues an individual has with taking medicines are insignificant, so therefore no conclusion can be drawn from that

Table 8: The result of a negative binomial regression with income (x1000), education, an individual being handicapped, their physical health, and effort required taking medicines on the amount of medical specialist in a hospital visits.

MedicalSpecialist	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
Incomex1000	-.0002127	.000202	-1.05	0.292	-.0006085	.0001832
Education	.0081004	.0067272	1.20	0.229	-.0050848	.0212855
Handicap	1.104723	.0937378	11.79	0.000	.9209999	1.288445
PhysicalHealth	.3686002	.0414225	8.90	0.000	.2874136	.4497867
TakingMedicines	.011036	.0714233	0.15	0.877	-.128951	.151023
_cons	-1.371623	.1611879	-8.51	0.000	-1.687545	-1.0557
/lnalpha	.7102572	.0677702			.5774301	.8430842
alpha	2.034514	.1378794			1.781454	2.323522

Notes: Table 9 gives the results of the regression as described above. The robust standard errors are next to the coefficients. The total amount of observations (n) is 1800. The stars behind the coefficients refer to the significance level (*p<0.10 , **p<0.05, ***p<0.01

As with the previous sub question, a negative binomial regression is performed as well. Almost all the significance levels are exactly the same, apart from the constant, which went from a 99%, to a 95% significance level. The results can be interpreted in the same manner as the OLS regression. There is almost no difference when it comes to the effect of income on medical specialist visits.

To conclude: There seems to be no relation between an individual's income and the amount of visits to a specialist doctor in a hospital. This has been shown in a both an OLS and a negative binomial regression. No nonlinear effects have been observed either.

4. *Does a individuals' income affect how high their voluntary deductible is?*

For the last question, a further look will be taken on the effect of income on the voluntary deductible. Education, a self-perceived rating of health and the visits to a doctor or medical specialist in a hospital are used as control variables.

Table 10: The result of a regression with income (x1000), education, an individual rating of their own health, the amount of visits to either a general practitioner or medical specialist in a hospital on the voluntary deductible.

Deductible	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Incomex1000	-.0064971	.0193128	-0.34	0.737	-.0443751	.0313809
Education	3.553043	.5955099	5.97	0.000	2.385077	4.721009
HealthSelf	31.26396	5.485699	5.70	0.000	20.50492	42.02299
FamilyPhysician	-4.366604	1.563361	-2.79	0.005	-7.432804	-1.300404
MedicalSpecialist	1.102847	.7518155	1.47	0.143	-.3716794	2.577373
_cons	-75.00612	19.26144	-3.89	0.000	-112.7833	-37.22891

Notes: Table 10 gives the results of the regression as described above. The robust standard errors are next to the coefficients. The total amount of observations (n) is 1800. The stars behind the coefficients refer to the significance level (*p<0.10 , **p<0.05, ***p<0.01)

Table 11: The result of a regression with income (x1000), squared income, education, an individual rating of their own health, the amount of visits to either a general practitioner or medical specialist in a hospital on the voluntary deductible.

Deductible	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Incomex1000	.0170185	.0758429	0.22	0.822	-.1317313	.1657682
IncomeSQ	-3.19e-06	9.94e-06	-0.32	0.749	-.0000227	.0000163
Education	3.522278	.6033378	5.84	0.000	2.338958	4.705597
HealthSelf	31.14028	5.500613	5.66	0.000	20.35199	41.92856
FamilyPhysician	-4.362334	1.563809	-2.79	0.005	-7.429413	-1.295255
MedicalSpecialist	1.103456	.752006	1.47	0.142	-.3714447	2.578356
_cons	-75.05381	19.26683	-3.90	0.000	-112.8416	-37.266

Notes: Table 11 gives the results of the regression as described above. The robust standard errors are next to the coefficients. The total amount of observations (n) is 1800. The stars behind the coefficients refer to the significance level (*p<0.10 , **p<0.05, ***p<0.01)

Initially, 2 OLS regressions are performed where the first does not include the square term of income, and the second has a variable of the square term of income. As observed, the significance levels for income are higher in the model without the square term of income. It is therefore possible to conclude that there is no nonlinear effect when it comes to income, and that the first model (table 10) will be used for further analysis.

It should be immediately obvious that while the coefficient for income is very small, it is also largely insignificant. Therefore, an income effect cannot be concluded from this model. The coefficient for education is positive, which leads to the conclusion that a higher education is related to a higher deductible. An individual's rating of their own health seems to have the largest impact on an individual's voluntary deductible. A higher rating of an individual's own health leads to a much higher voluntary deductible.

There are some small effects for visiting a general practitioner and medical specialist, but these coefficients are rather minor. The amount of visits to a family physician is negatively correlated with a voluntary deductible, which is expected. However, the amount of visits to a medical specialist is positively correlated with a voluntary deductible. This would lead to the conclusion that more visits to a doctor is related with a higher deductible, which would not seem logical and as this variable is not significant, it can be suspected it is not entirely correct.

As explained in the literature review, the maximum voluntary deductible is €500,-. This means an individual is able to have a voluntary deductible anywhere from 0 to €500. We can use this data to analyze the same data using a tobit model. The boundaries of this model are a lower limit of 0, and an upper limit of 500. The model is described on the following page.

Table 12: The result of a tobit model with income (x1000), education, an individual rating of their own health, the amount of visits to either a general practitioner or medical specialist in a hospital on the voluntary deductible.

Deductible	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Incomex1000	-.2027986	.5169433	-0.39	0.695	-1.216672	.8110753
Education	56.15448	10.55622	5.32	0.000	35.45071	76.85825
HealthSelf	428.5833	89.94571	4.76	0.000	252.174	604.9926
FamilyPhysician	-138.4255	38.16968	-3.63	0.000	-213.2872	-63.56382
MedicalSpecialist	14.1061	8.382488	1.68	0.093	-2.33436	30.54656
_cons	-3662.327	448.6081	-8.16	0.000	-4542.176	-2782.478
var(e.Deductible)	2718365	479933.2			1922759	3843179

Notes: Table 12 gives the results of the mode as described above. The robust standard errors are next to the coefficients. The total amount of observations (n) is 1800. The stars behind the coefficients refer to the significance level (*p<0.10 , **p<0.05, ***p<0.01). The model had a lower limit of 0, and an upper limit of 500.

Analysis using a tobit model yields very different results than the standard OLS regression. It is noteworthy that while most significance levels remain the same, as only medical specialist became 10% significant, the coefficients had some major changes.

The main variable of interest: income, is still very insignificant and the effect of income is still very minor. However, the effect of education has become much larger. As the average level of the dummy variable for education was 16, a coefficient of 56 results in a very large impact on the voluntary deductible. The same can be said for the scale individuals rate their own health on. With a coefficient of 428 for a 1-5 scale, the impact on the deductible is also large. It is also noteworthy that both of these variables are 99.99%+ significant.

The amount of visits to a general practitioner has a negative effect on an individual's voluntary deductible. This is expected as somebody with a lot of medical issues, and therefore high costs, has no benefit in a high voluntary deductible. The last coefficient is visits to a medical specialist, but the effect of this is very minor, and the significance is also low. We can therefore declare this is not very relevant.

To conclude: The amount of income an individual has, has a very minor effect on an individual's own deductible. However, it seems to be much more related to education levels and a perceived rating of their own health. There are no nonlinear effects for income.

Discussion

As with any research, there is always a possibility for improvements. In this paragraph limitations for our research, and suggestion and improvements for future research will be discussed.

One of the biggest limitations of this study is that most income-related variables were not significant. As this is the case, it is hard to draw strong conclusions from the data. One of the suggestions here is to find a bigger dataset, as this would most likely improve significance drastically. Another way to make the research more reliable, would be to do a difference-in-difference comparison between groups. As discussed in the theoretical part of this thesis, the RAND experiment where groups of people had different deductibles, was a very efficient way to analyze this subject. Unfortunately, as this is only a bachelor thesis, nor the funds, nor the time is available to do an experiment of this scale.

When analyzing the sub-question regarding deductibles, it became obvious that the income variable was not significant at all. A suggestion here would be to look at the total amount of income spent on healthcare for an individual, as it could be assumed that an individual would base a possible voluntary deductible based on that. Unfortunately, this kind of data was not available, but it would be interesting to include this if any further research is ever done.

Something that should be taking into account is the global covid pandemic. This dataset was gathered during the period of 2019-2020. The global pandemic could have had an effect on our dataset in multiple logical ways. As hospitals were filled to the peak during periods where covid reigned the land, it could have been possible that operations were cancelled, or individuals were not able to make appointments at the hospital as less urgent care was scaled down. On the other hand, there is also the chance that due to isolations and quarantine, people went less to the hospital than they usually would.

Another possible improvement for future research is to use wealth instead of income for certain parts of the population. A small amount of our dataset contained elderly people that were already retired. This gives them a lower income than they had when they would have worked. It could be argued that total wealth would be a better way to analyze their data than income as wealth would give a better representation of their consumption power.

Conclusion

In this paper, multiple regressions regarding income and healthcare consumption have been done. As a result of that, there is now a broader understanding of the effect from income on healthcare consumption. The main question of this paper was as follows:

To what extent does net income per year have an effect on the number of times an individual visits a doctor?

The first regression proved that while a numerous amount of factors had a significant amount of impact on an individual's physical health, income was not one of them. This can lead to the conclusion that somebody's physical health is not related to their income.

The same conclusion can be drawn with the one for visits to either a medical specialist at a hospital, or a general practitioner. In both these regression it became obvious that income played a very minor part on how many times an individual visits a doctor and was therefore not significant.

At the beginning of this paper, an assumption was made that voluntary deductibles were based on an individual's income. By analyzing the voluntary own deductible, it was shown that they were largely unrelated with income, and that there was therefore no direct effect.

To draw a final conclusion to this paper based on our main question, income does not seem to relate to the amount of times an individual visits any doctor, let it be a general practitioner or a medical specialist. Income does also not seem to play a significant factor in an individual's physical health. While income might have an indirect effect on health trough for example education, no direct effect of income has been found. Therefore, the final conclusion is that when looking at income, the Dutch healthcare system turns out to be a fair and equal system.

Literature

Davis, K., Doty, M., & Ho, A. (2005). How high is too high? Implications of high-deductible health plans. *The Commonwealth Fund*. Published. https://www.commonwealthfund.org/sites/default/files/documents/___media_files_publications_fund_report_2005_apr_how_high_is_too_high__implications_of_high_deductible_health_plans_816_davis_how_high_is_too_high_impl_hdhps_pdf.pdf

Martin, S., Siciliani, L., & Smith, P. (2020). Socioeconomic inequalities in waiting times for primary care across ten OECD countries. *Social Science & Medicine*, *263*, 113230. https://doi.org/10.1016/j.socscimed.2020.113230

Ministerie van Algemene Zaken. (2019, 2 oktober). *Hoe is de zorgverzekering in Nederland geregeld?* Zorgverzekering | Rijksoverheid.nl. https://www.rijksoverheid.nl/onderwerpen/zorgverzekering/zorgverzekeringsstelsel-in-nederland

Ministerie van Algemene Zaken. (2021, 15 januari). *Wanneer betaal ik een eigen risico voor mijn zorg?* Rijksoverheid.nl. https://www.rijksoverheid.nl/onderwerpen/zorgverzekering/vraag-en-antwoord/eigen-risico-zorgverzekering

Oronce, C. I. A., Scannell, C. A., Kawachi, I., & Tsugawa, Y. (2020). Association Between State-Level Income Inequality and COVID-19 Cases and Mortality in the USA. *Journal of General Internal Medicine*, *35*(9), 2791–2793. https://doi.org/10.1007/s11606-020-05971-3

Van Doorslaer, E. (2006). Inequalities in access to medical care by income in developed countries. *Canadian Medical Association Journal*, *174*(2), 177–183. https://doi.org/10.1503/cmaj.050584

Zorgwijzer. (2021, 29 april). *Eigen risico zorgverzekering (2021): wat is het?* https://www.zorgwijzer.nl/faq/eigen-risico

Acknowledgements as mandatory by using the LISS dataset

In this paper we make use of data of the LISS (Longitudinal Internet Studies for the Social sciences) panel administered by CentERdata (Tilburg University, The Netherlands)