Erasmus University Rotterdam

Erasmus School of Economics

Master Thesis [Economics and Business - Health Economics specialization]

Optimism, pessimism and cancer screenings: Understanding the role of health beliefs on preventive health behaviors

Student: Elissavet Mina

704986

Supervisor: Fanny Tallgren

Second assessor: Carlos Riumallo Herl

Rotterdam, 05/07/2024

Words:10,390

Abstract

This thesis investigates the impact of health beliefs on preventive health behaviors among Dutch women, focusing on breast and cervical cancer screening. This study explores how optimistic and pessimistic, as well as positive and negative, health beliefs can affect the individuals' decision to undergo cancer screening. Using data from the Longitudinal Internal studies for the Social Sciences (LISS), we examine two analyses: the first is a panel data analysis from 2008 to 2018 and the second one is a single wave study from 2010. The first analysis categorizes health beliefs as positive, negative or neutral based on how the individuals describe their health status. The second analysis categorizes health beliefs as optimistic, pessimistic or neutral based on the individuals' self-health assessment in comparison with others sharing similar socio-economic characteristics. Both analyses consider additional factors like employment status, age, educational attainment and the presence of kids as they may affect the association between health perceptions and screening behaviors. The findings of the first analysis suggest that positive outlooks are positively correlated with breast cancer screening uptake, while negative outlooks are negatively correlated. Regarding cervical cancer screening, there is a positive correlation between subjectively assessed health status, regardless if it is positive or negative, and screenings. The results of the second analysis showed that, for breast cancer, optimistic and pessimistic health beliefs are positively correlated with screening test uptake while for cervical screenings this correlation is negative.

Table of contents

1. Introduction

Breast and cervical cancers are a global health concern, constituting important causes of death among women each year. According to the WHO (2024), in 2022, 2.3 million women were diagnosed with breast cancer globally, resulting in 670,000 deaths, making it the most prevalent type of cancer among women. In addition to that, for the same year, there were 660,000 new cases of cervical cancer and around 350,000 deaths, which ranks it as the fourth most common cancer among females, (WHO, 2024). Despite the big burden of these two cancers, an early detection through regular cancer screenings is crucial. Early screening increases cancer's curability, providing the patient with a broader range of treatment options and therefore achieving higher survival rates. In this thesis, we will focus on cancer screenings such as mammograms, breast self-examinations and Pap smears. Data from Eurostat for 2021 for the Dutch females, state that the Netherlands belong to the top 10 European countries with the highest participation rates in preventive cancer screening programs for breast and cervical cancer, scoring 72.2% mammogram and 54.8% Pap smear test uptake. However, these percentages have faced a decrease of more than 10% since 2007, raising concerns among decision makers about the factors contributing to this decline.

The decision of undertaking screening depends on a multitude of factors such as education, age, socioeconomic status, health beliefs, time preferences, recommendations by doctors, risk aversion, health insurance and premiums. Particularly, an individual's health belief about their own health can be a significant psychological influence which can lead to avoidance of screenings, even when it is necessary (Arni et al. 2O21) or to an overwillingness to undergo them due to misperception of their health (Hoy et al. 2014) (Ghirardato et al. 2004). Nevertheless, it requires a wide spectrum of research to comprehend how health beliefs work. Here, we are aiming to shed more light on this topic by specifically investigating the association between the way that individuals self-assess their own health and cancer screening uptake. In addition, in this thesis we are looking at the relationship between health beliefs and cancer screening decisions of the Dutch female population. We will focus on optimistic and pessimistic beliefs which we assess in two ways. Firstly, based on how individuals perceive their own health, positively or negatively and secondly, compared to others, as they often remain overly optimistic believing that others might develop cancer but not themselves and vice versa, despite being aware of the risks (Arni et al. 2O21).

Women that misperceive their health status and are often characterized by an optimistic view, may avoid screening believing in their overall healthiness and ignoring any warning symptoms. Hence, they may overestimate the probability of not getting cancer in the next few years and choose to remain uninformed about their actual health status (Oster et al. 2013). This can be considered as an optimal decision, as individuals derive utility from having optimistic beliefs about their health status. Therefore, their anticipated consumption utility from avoiding cancer screening in the present, may be higher than the actual one of knowing their true health state in the future, leading to utility maximization (Oster et al. 2013). However, a pessimistic misperception of their health can have two potential outcomes. Either an excessive eagerness to undergo regular testing in order to be prepared for a potential cancer development, or an avoidance of cancer check-ups driven by the fear of expecting to get bad news due to the belief of being in poor health. From one perspective, individuals holding a pessimistic view about their health who are more prevention-driven, are more likely to engage in preventive health behaviors such as screening and vaccination programs as they intend to prevent negative outcomes fueled by their poor health status (Leder et al. 2014) (Hazlett et al. 2011). On the other hand, individuals may avoid collecting information about their health, driven by fear and anxiety, if they negatively evaluate their health status and expect to receive bad news by visiting a doctor or engaging in any preventive health behavior (Golman et al., 2017).

1.1 Aim and research question

This study aims to identify a non-causal effect between health beliefs and the decision to undergo testing by looking at how people perceive their own health. It assesses whether self-reported health status is correlated with screenings using a panel data analysis and whether optimistic and pessimistic self-perceptions of health are correlated with screenings via a single wave analysis. This subjectively self-health evaluation may affect their decision to get tested for breast and cervical cancer. This investigation focuses on women who are between 28 to 75 years old.

All these lead to the key question:

- *"How can subjectively assessed health status affect the uptake of cancer screening tests in the case of breast and cervical cancer for Dutch women aged 28-75 years?"*

Furthermore, this study will additionally investigate the following subquestion:

- *"How can optimistic and pessimistic health beliefs affect breast and cervical cancer screening test uptakes for Dutch women aged 28-75 years?"*

This thesis is structured into five main chapters, following up with the theoretical framework. Later on, the data are presented with the empirical methods and results and the last chapter is dedicated to the discussion part and conclusions.

2. Theoretical framework

This chapter will discuss the theoretical framework behind the research questions. Based on the literature, it is not clear whether optimism and pessimism, as well as positive and negative health beliefs, will increase or decrease screenings. The evidence behind each of them, as well the potential influence of the factors of age and education will be explained in the following sections.

2.1 Optimistic health beliefs

In the majority of the available literature, the term "optimism" is used to describe individuals with unrealistically optimistic or too positive expectations and beliefs. Schneider, S.L. (2001) proposes that there is another aspect of optimism which stems from having and pursuing a positive outlook while being aware of potential challenges and risks. In her research, she explains that individuals with just a positive point of view or a positive self-perception are not necessarily overconfident or overestimating themselves and their capabilities. Additionally, Levy and Myers (2004) in their research on how self-perception of aging can influence preventive health behaviors, demonstrate that individuals that just have a more positive self-perception of aging, are more likely to engage in preventive health behaviors, such as exercise, a balanced diet, regular doctor visits etc., over the next two decades. Hence, distinguishing between types of optimism can provide different results in cancer screening uptake. In our study, to clearly differentiate these two terms, we will refer to "optimistic health beliefs" in the context of unrealistic optimism where the individuals view their health too positive, when actually is not as good as they think it is. On the other hand, we will refer to "positive health beliefs" in the context where the individuals just have a positive outlook, which might be realistic. Although we are not able to objectively assess the realism of these beliefs, we can simply say that they have a positive point of view of their health and we aim to observe the potential differences in our results.

It is known that women are more likely and willing to engage in preventive health behaviors than men (Deeks et al. 2009) but among these women, the ones who hold an overly optimistic belief about their health, often seem to be reluctant to seek medical attention and undergo cancer screening testing. This stems from the perception of being in a good health status, leading to delayed doctor visits and underutilization of preventive care services (Spitzer and Shaikh, 2022). Spitzer and Shaikh (2022), in their study on the influence of health perception on doctor visits for the European population aged 50+, consider individuals with overly optimistic outlooks about their health to be overconfident. They assess this by comparing objective measures of performance with their self-reported equivalents, such as the ability to stand up from a chair. The findings of their research show that overconfident individuals are less likely to use preventive care and tend to visit the doctor 17% less frequently than those who accurately evaluate their health.

Research conducted by Oster et al. (2014) examines how an individual's optimistic perceptions about their risk for Huntington's disease can affect their decision to undergo genetic testing. They employ an “optimal expectations model” to explain behavior patterns and beliefs of untested individuals at risk for Huntington’s disease. It is divided into 3 periods. At t=0 the individuals decide if they want to know through genetic testing whether they carry the mutated gene or not and at t=1 they take a corresponding action. Following, it is the utility trade-off between their expected utility of their future consumption at this period (t=1) and the actual one received at t=2 when they find out which is their true health status. When the consumption utilities at t=1 and 2 align, they can maximize their utility. However, Oster et al. (2013) demonstrated that even if the untested individuals choose an overly optimistic belief, it can still be an optimal decision if a high expected consumption utility at t=1 offsets the low and actual one at t=2. If this requirement holds, then these individuals will prefer to remain optimistic and not undergo testing since they do not have to face the chance of finding out whether their actual health is better or worse than what they expected, meaning if they carry the gene or not.

Golman et al. (2017), building upon the foundations of Hazlett et al. (2011) on regulatory focus, demonstrated that promotion-driven individuals, in order to maintain optimistic expectations, are motivated to avoid information, which can be provided through screening. These individuals may develop strategies and take the risk of not undergoing testing in order to not receive a negative revision of their expectations. Therefore, they might avoid information about their health to maintain their optimistic point of view, since it can be shattered if they receive a positive cancer screening test. Hence, an optimistic outlook presents a regulatory alignment with promotion-focused individuals who may use various tools to ensure positive outcomes and avoid any additional information or factors that might alter their expectations.

Based on the information above, the first and second hypotheses concerning optimistic and positive health beliefs are formed as follows:

H1: Optimistic health beliefs are associated with lower cancer screening uptake

H2: Positive health beliefs are associated with higher cancer screening uptake

2.2 Pessimistic health beliefs

As in optimistic beliefs, pessimism can also have two aspects, the unrealistic or too negative expectations that can lead to underconfidence and underestimation of one's capabilities and another one that concerns of just having a negative outlook, which might be conformed to reality (Chipperfield et al., 2019). Most available studies focus on unrealistic pessimism and its effect on people's health behaviors, which can be either positive or negative. However, the aspect of simply having a negative point of view has not been sufficiently examined by researchers and therefore we have limited information about its impact on preventive health behaviors. We hypothesize though that, similar to positive health beliefs, individuals that just have a more negative self-perception about their health are more likely to engage in preventive care. In our research, to clearly differentiate these two terms, we will refer to "pessimistic health beliefs" in the context of unrealistic pessimism where the individuals view their health too negative, when actually is not as bad as they think it is. Alternatively, we will refer to "negative health beliefs" in the context where the individuals just have a negative outlook, which might be realistic. As previously mentioned, we are only able to say that they have a negative point of view of their health, without objectively assessing the realism in this belief and we aim to document the potential differences in our results.

The main argument that researchers have found about why people holding overly pessimistic beliefs for their health may avoid screening testing stems from fear and anxiety for receiving bad news from the tests concerning their health status. Karlsson et al. (2009) incorporated the concept of selective attention to information into the "ostrich effect" term by Galai and Sade (2003), by introducing a decision-making model to forecast selective attention behaviors where individuals derive utility from information and can partially control when to acquire it. The model presents that people may selectively avoid collecting information if they expect to receive bad news and therefore aim to shield themselves by pretending that this information does not exist. Examining this effect in the setting of testing and doctor's visits, Golman et al. (2017) provide two illustrative examples. One applies to individuals at risk of AIDS who may avoid medical testing due to anxiety and fear of receiving a positive diagnosis and thus overlook potential symptoms. Another example corresponds to people over fifty years old who subjectively perceive themselves to be at high risk for developing cancer in the future, reflecting a pessimistic point of view for their health. Despite this perception, they are more likely to avoid doctor's visits in order to abstain from any additional information that may confirm it.

Another evidence supporting this argument shows that individuals often tend to be reluctant to seek information about their health or visit a doctor, even if they have alarming symptoms and it is cost-free. This can be due to fear of receiving bad news driven by their personal health beliefs about being in a poor health in the next period. This perspective has been analyzed as well by Koszegi, (2003) by using a patient's utility function in two periods. At first, the patient makes a decision related to his/her health, such as whether to go to the doctor or not. Depending on this decision, may either receive treatment from the doctor or follow a health-related lifestyle based on the patient's choice. In that case, if the decision-maker believes that he/she will receive bad news, might avoid the doctor's visit as his/her future well-being is going to be lower if she will get diagnosed with cancer.

On the contrary, few researchers have demonstrated that individuals who have a pessimistic point of view might be more prone to engage in preventive health behaviors such as undergoing cancer screening tests. Spitzer's and Shaikh's (2022) research on health perception and doctor's visits, as previously mentioned, also indicates that people who have a more negative perception about their health status tend to have 21.4% more doctor's visits. Their findings show that these individuals are more likely to engage in earlier screenings and subsequent diagnosis of potential diseases, stemming from their increased frequency of doctor visits. In addition, they concluded that women are more inclined to hold a pessimistic belief about their health than men, which can lead to higher cancer screening uptake.

Leder et al. (2014), following the steps of Hazlett et al. (2011), demonstrate that prevention orientation is associated with people holding a pessimistic point of view. Specifically, in their research they conducted four main studies using different measures of regulatory focus to assess health behaviors in vaccination uptake. They propose that individuals with a pessimistic mindset that they expect to face losses, such as health deterioration due to a disease, may be more motivated to pursue strategies to guard against these losses. These strategies can include participating in immunization programs to decrease the possibility of getting a particular disease or the severity of the symptoms. Hence, a pessimistic outlook presents a regulatory alignment with prevention-focused individuals who may use available tools to prevent anticipated negative outcomes and protect their health and well-being.

Overall, according to the majority of empirical evidence, the third hypothesis concerning pessimistic health beliefs is developed as follows:

H3: Pessimistic health beliefs are associated with lower cancer screening uptake

Lastly, supporting our previously mentioned assumption, the fourth hypothesis concerning negative health beliefs is formed as follows:

H4: Negative health beliefs are associated with higher cancer screening uptake

2.3 Influence of education and age

Evidence has shown that the factors of education and age can influence both the decision to undergo cancer screening and people's health beliefs. In this section, we provide the rationale behind this influence and in the following chapters we control these confounders and perform sub-group analysis on them.

The educational attainment of individuals is an important factor in their engagement in preventive health behaviors, such as participation in cancer screening programs. Picone et al. (2004) suggests that least educated individuals are less likely to engage in preventive health behaviors whereas more educated ones are more inclined to undergo mammograms and Pap smear tests. The rationale behind the influence of education stems from the fact that more educated people are better able to comprehend health information and consequently, they might be more inclined to participate in preventive health activities (Eibich and Goldzahl, 2020). Moreover, Dickie and Gerking (1997) conducted a study on genetic risk factors and beliefs and offsetting actions in the context of skin cancer, including education as an explanatory variable in their model. In particular, it is about a utility maximization model, which examines how individuals can maximize their utility through precautionary behaviors that can reduce the risk of developing skin cancer. Their results suggest that those with higher levels of formal education are more inclined to participate in preventive health behaviors, such as using sun protection and spending less leisure time outdoors during summer.

On the other side, Bago d'Uva, et al. (2020) demonstrated that the least educated individuals are more likely to incorrectly evaluate their decisions that have impacts on their longevity. In their study, they compared the subjective probability of the participants to live till the age of 75 with the actual survival rates in order to evaluate how accurate their expectations were. The results, in relation to the individuals' education levels, suggest that people with the least education and cognitive ability tend to report the least accurate survival probabilities and are more prone to exhibit both optimistic and pessimistic behaviors.

The age of individuals is an important factor influencing their decision to engage in preventive health behaviors. Numerous studies have developed the hypothesis that there is a positive association between age and screening behaviors, indicating that older individuals are more likely to undergo testing (Leder et al. 2014). Nevertheless, Deeks et al. (2009) in their study on "The effects of gender and age on health-related behaviors" found that this pattern applies to the majority of cancer screening tests but might vary with respect to cervical and breast cancer across different ages. They performed a "chi-squared" analysis on data that was collected through a random probability sampling method using self-completion surveys from 1456 Australian residents. Their findings show that survey participants over 51 years old were more likely to undergo screening tests such as mammograms compared to younger individuals. However, in the case of pap smear screening for cervical cancer, their results indicate that younger women, aged 31 to 40 years, were more inclined to undergo testing than older women over 61 years old. Overall, their research observed that younger participants are generally less likely to engage in preventive health behaviors and perform annual health checkups compared to older individuals.

The other perspective suggests that the individuals' age can affect their health beliefs and as they might be more or less likely to have an overly optimistic or pessimistic outlook. Bago d'Uva, et al. (2020) also mention that older individuals are less able to correctly assess their longevity. Specifically, those aged 54-65 are more likely to incorrectly evaluate their survival probabilities of living to 75 by adopting a pessimistic approach.

2.4 Contribution to the literature

The contribution of this thesis to the literature on the role of health beliefs on preventive health behaviors is significant. According to Spitzer's and Shaikh's (2022), the impact of these biases on healthcare utilization has been underexplored by researchers. Studies by Leder et al. (2014), Koszegi (2003), Golman et al. (2017), Karlsson et al. (2009) and Hazlett et al. (2011) include the influence of health beliefs as supplementary analysis and/or mainly provide a rationale for them by following a psychological perspective. In addition, there is no application of these biases to specific health behaviors beyond doctor's visits, which applies also in the study of Spitzer's and Shaikh's (2022). In our opinion, this approach might not be very representative since the decision to visit the doctor for a flu is very different from seeking information about cancer risk. Hence, this study aims to fill this gap by providing important insights in the application to breast and cervical cancer while exploring this influence in practice, building upon the theory that was established from previous studies.

Furthermore, Spitzer's and Shaikh's (2022), explore the influence of health misperceptions in health care utilization, focusing only on older individuals over 50 years old while keeping a broad scope to the European population. In contrast, our study includes younger age groups as well and focuses on cancer screenings in the Netherlands. Given the national invitations offered to women being in the age eligibility for screenings, it is important to observe individuals younger than 50 years old, who have completed their education and analyze their behaviors in the context of a specific and life-threatening aspect of healthcare utilization.

Oster's et al. (2014) research on genetic testing for Huntington's disease followed a similar approach to ours in some aspects. Even though their insights were valuable to our research, there are some crucial differences and limitations which we would like to address with our contribution. Firstly, they cover only the optimistic outlook, arguing how it could still be an optimal decision, without addressing the pessimistic perspective. While this may be due to the nature of their topic, it is crucial to have a holistic view on both sides of misperceptions. This thesis aims to accomplish that by examining both optimistic and pessimistic perceptions in the context of cancer screening tests. Moreover, it is essential to mention that the nature of the genetic testing for Huntington disease is very different from cancer screenings and in general tests for more common ailments. The rationale behind this is that it is not a curable disease and therefore undergoing a genetic test can only detect if the individuals will develop the disease with no available treatment thereafter. On top of that, the age factor has an inverse influence, meaning that the probability of carrying the disease decreases as people age without symptoms. These two variations may be more applicable to rare diseases with low prevalence. However, breast and cervical cancer are very common types of cancer with a higher prevalence and available treatments, provided that they are detected in early stages, while the risk of developing cancer increases with age. Hence, there is a need for a study to bridge the gap in understanding how health beliefs can influence the decision to undergo screening for breast and cervical cancer.

Last but not least, given the limited evidence on the positive and negative health beliefs, in this study, we are aiming to shed light on their potential impacts on participation in cancer screening programs. We intend to investigate if these impacts have different influences in the decision to undergo breast and cervical cancer screening, in comparison to overly optimistic and pessimistic health beliefs.

3. Data and empirical strategy

3.1 Dataset description

Two different datasets are utilized for two analyses from the Netherland's Longitudinal Internet studies for the Social Sciences (LISS). The first analysis addresses the main research question and the second analysis addresses the subquestion. The first dataset is a panel dataset from the Core LISS study, which ran from 2007 to 2022, with the exemption of the year 2014, from which we use 22,783 observations. This dataset’s questionnaire focuses on health, health perception and health related job situations. The main focus of this analysis is in the waves 2-11, covering the years 2008-2018, in order to examine people’s beliefs and screening behaviors over this time period. We excluded the years 2007 and 2019-2022 as they do not include our dependent variables, which are presented later on. From this dataset, we extracted dependent and independent variables for the first analysis, as well as the covariates, which are used for both analyses, as they correspond to the same individuals. Lastly, the second dataset, from a 2010 single wave study, includes 1,160 observations and focuses on health prevention from which we use two variables in order to construct our main independent variable for the second analysis.

3.2 Data transformation and variables

For the first analysis, several data transformations were performed. These transformations include converting the dataset from a wide to a long format, merging categories and renaming them in our covariates, as well as transforming the variable of age from continuous to categorical. Table 1 presents the variables of the final dataset for the first analysis.

Table 1: Variables’ description-1st analysis

Variable	Type	Categories	Description
breast_screening	Binary,Dependent	1=YES, 0=NO	Indicates whether the respondent had an X-ray taken of one or both breasts over the past two years.
cervical_screening	Binary,Dependent	1=YES, 0=NO	Indicates whether the respondent had a smear test taken over the past five years for cervical cancer.
des_H	Categorical	1=Neutral 2=Negative 3=Positive	Describes the way the respondent perceives her own health.

education	Categorical	1=Low 2=Medium 3=High	Represents the respondent's level of education based on the Dutch educational system. Respondents with unclear educational information were excluded.
kids	Binary	1=YES, 0=NO	Indicates whether the respondent has kids or not.
job	Binary	1=Paid job 0=No paid job	Indicates whether the respondent has a paid job or not.
age	Categorical	1=[28-39] 2=[40-51] 3=[52-63] 4=[64-75]	Represents the age intervals of the respondents.
survey_years	Categorical	2008, 2009, 2010, 2011, 2012, 2013, 2015, 2016, 2017, 2018	Represents the years of the survey, excluding 2014, constructed through wide to long reshaping.

The variable "des_H" corresponds to the question "How would you describe your health, generally speaking?". It takes value 1 for Neutral, if the respondents described their health as "Good", value 2 for Negative if they described it as "Poor" or "Moderate" and value 3 for Positive if they described it as "Very good" or Excellent. The variable "kids" was included since giving birth can affect the risk of developing breast and cervical cancer (Breast Cancer Organization, 2024) (American Cancer Society, 2020). Moreover, we included the variable "job" for employment status, as having a job or not might affect the screening test uptake due to their affordability. In addition, the variable for gender was restricted to females, as these types of cancer screening tests are specific to women. Lastly, the age interval was constructed in a way that it corresponds to women who have completed their educational attainment and are within the age range suitable and capable enough to participate in this study

For the second analysis, we use the same covariates from the core study, besides "survey_years", focusing only on the year 2010. Regarding the dependent variables, we created two binary ones, "breast_screening_dummy" and "cervical_screening_dummy", that take value 1 if the sum of the times that the individual has answered "YES" to the core study variables "breast_screening" and "cervical_screening" is equal or bigger than 1, meaning if the woman went for screening at least once during the time period 2010 to 2018. Moreover, they take the value 0 for "NO", if the woman did not go for screening from 2010 to 2018 and therefore, the sum equals 0. In order to achieve this, we used survey data from 2012 to 2018 for "breast screening" and from 2015 to 2018 for "cervical screening". This approach is based on the definition of those variables, as breast screening corresponds to whether the respondent had an X-ray taken of one or both breasts over the past two years and cervical screening corresponds to whether the respondent had a smear test taken over the past five years. Hence, this adjustment ensures that observations of the new variables “breast_screening_dummy” and “cervical_screening_dummy” begin precisely from 2010. Additionally, we excluded from the analysis the first age category, as this single wave dataset corresponds to individuals who are older than 41 years old. For the main independent variable, we use two different variables, “c_10y” and “c_10y_others”, in order to construct it. These variables correspond to the following questions: “What do you think is the percentage chance that you will develop cancer in the next 10 years?” and “If you consider 100 people that are identical to you in terms of socio-economic characteristics (such as age, education, income, etc.), how many of them do you think will develop cancer in the next 10 years?” respectively. Therefore, the main independent variable “H_beliefs” is calculated as the difference between the “c_10y” and the c_10y_others” divided by 100. The method is inspired from the study of Arni et al. (2021), where they measured the subjective health ranking of the individual in the population health distribution through the difference between 100 randomly selected residents in the same age of the respondents and the answer that they provide in the question” How many of those 100 people would be in better health than you?”. Hence, following a similar approach, we generate our independent variable and categorize it such that, after dividing by 100, it takes value 1 for neutral if difference=0, 2 for pessimism if 0<difference≤1 and 3 for optimism if -1≤difference<0. The benefit of this analysis is that we are able to identify if people have self-perceptions that are overly optimistic or pessimistic.

Overall, it would be ideal to include several additional covariates in our analysis in order to have a more holistic view of our research questions and avoid any potential omitted variable bias in the second analysis of 2010. Even though we address omitted variable bias from time invariant factors in the panel data analysis by using a fixed effects model, for which a detailed reference of the empirical strategy is presented below, there might still be omitted time variant factors. In particular, it would be useful to have information of the individual's family history regarding breast and cervical cancer, as it could also explain a part of their decision to undergo screenings or not. Furthermore, a variable that objectively measures the true risk of these individuals to develop cancer could have been valuable to accurately define them as optimists or pessimists, but still this information was not provided. Moreover, we were unable to construct the variable of “race” due to the questionnaire's limitations as it provides only options for Dutch, Western or non- Western background, which we believe is not very representative. Additionally, obtaining information about the area where the respondents live would be useful for our analysis. However, the only available options range from extremely urban to not urban, making it difficult to assess which category corresponds to rural residence. This uncertainty also creates concerns of the respondents’ capability to correctly respond to this question. Lastly, we aimed to control the numeracy skills of the respondents to better determine the accuracy of their answers in questions involving probabilities. Unfortunately, this information was not available in these questionnaires. Nevertheless, the datasets we use are sufficient enough to address our research questions, as we still control for various variables and can perform two different analyses to explore optimistic and pessimistic as well as positive and negative health beliefs.

3.3 Descriptive statistics

Several descriptive statistics are presented in this section in order to examine the first trends of our variables and gain a general understanding of our sample. It is split into two different groups that each one corresponds to each analysis.

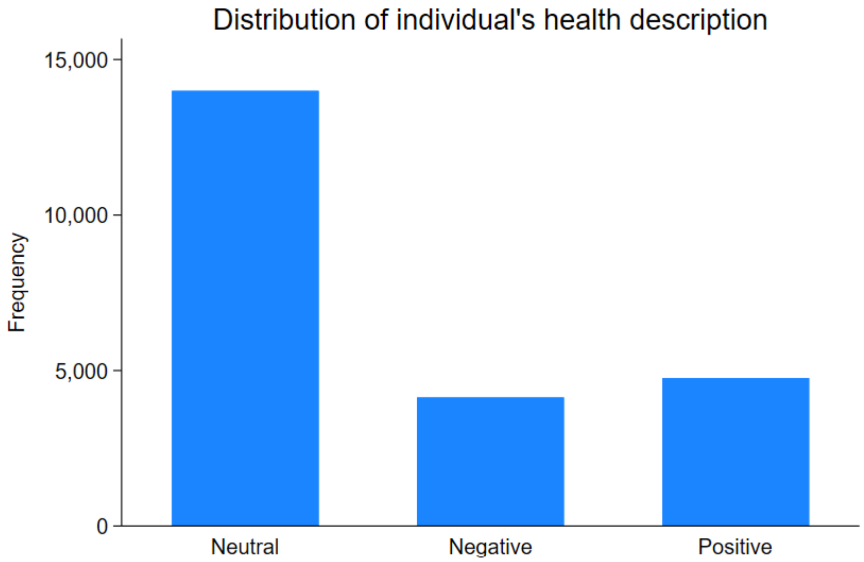
Table 2: Descriptive statistics of the variables from the 1st analysis-panel data analysis. The frequency and percent are presented.

		Freq.	**Percent**
cervical_screening	yes	12,021	52.76
	no	10,762	47.24
breast_screening	yes	13,576	59.59
	no	9,207	40.41
des_H	Neutral	13,994	61.13
	Negative	4,138	18.08
	Positive	4,761	20.80
agecat	1=[28-39]	4,928	21.53
	2=[40-51]	6,161	26.91
	3=[52-63]	6,799	29.70
	4=[64-75]	5,005	21.86
education	Low	893	3.90
	Middle	14,495	63.32
	High	7,505	32.78
job	no	7,593	33.17
	yes	15,300	66.83
survey_year	8	2,516	10.99

9	2,524	11.03
10	2,339	10.22
11	2,088	9.12
12	2,389	10.44
13	2,226	9.72
15	1,882	8.22
16	2,198	9.6
17	2,462	10.75
18	2,269	9.91

The total number of observations in the first analysis is 22,783. A comparison between the two dependent variables, shows that women engage more in screening behaviors for breast cancer, 59.59 % of the sample, than cervical cancer, 52.76% of the sample. The age categories appear relatively equally distributed, though the first category, including the youngest age group, has the fewest observations than the others. As expected, the majority of Dutch women have a middle to high education level, with only 3.9% of our sample having a low level of education. It is observed that the majority of the participants have a paid job and do not have children. Lastly, on average there are around 2.300 observations in each survey year. However, the survey year 2015 has the lowest number of observations with 1,882, while the year 2009 has the highest number with 2,524 observations.

Graph 1: Distribution of individual's health description- 1st analysis-panel data analysis



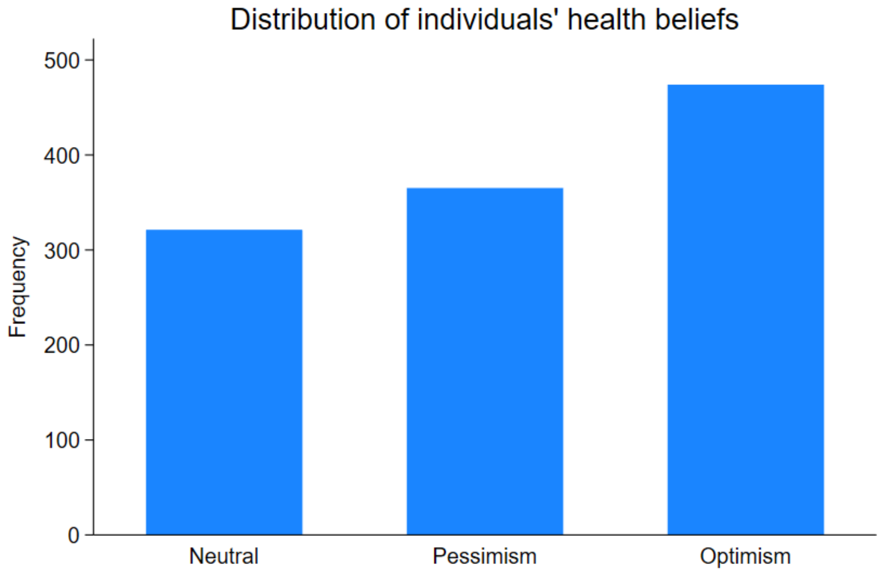
Graph 1 illustrates the categories of our main independent variable in the 1st analysis and it combines all the self-reported health descriptions in the period 2008 to 2018. The majority of the observations, 61.13%, are within the "Neutral" category in which the respondents described their health as "Good". However, the frequency of the "Negative" and "Positive" categories is almost equally distributed, 18.08% and 20.80% respectively, corresponding to people that described their health as "Poor" or "Moderate" and "Very good" or "Excellent" respectively.

Table 3: Descriptive statistics from the year 2010 of the variables from the 2nd analysis- 2010 single wave analysis. The frequency and percent are presented.

		Freq.	**Percent**
cervical_screening_dummy	no	197	16.98
	yes	963	83.02
breast_screening_dummy	no	390	33.62
	yes	770	66.38
H_beliefs	Neutral	321	27.67
	Pessimism	365	31.47
	Optimism	474	40.86
agecat	2=[40-51]	387	33.36
	3=[52-63]	484	41.72
	4=[64-75]	289	24.91
education	Low	61	5.26
	Middle	793	68.36
	High	306	26.38
job	no	452	38.97
	yes	708	61.03
kids	no	768	66.21
	yes	392	33.79

The total number of observations in the second analysis is 1,160. In both types of cancer screenings, we can see that the majority of people underwent screening from 2010 to 2018. The highest number of observations is among the third age category, with 484 respondents, and in the middle level of education, with 793 respondents. Lastly, as before, most of the participants have a paid job and do not have children.

Graph 2: Distribution of individual's health beliefs- 2nd analysis -2010 single wave



Graph 2 illustrates the three categories of our main independent variable in the second analysis. The majority of the observations are within the "Optimism" category, 40.86%, where the respondents described their chance of developing cancer in the next 10 years as lower than that of other individuals with similar socio-economic characteristics. Whereas the "Neutral", where the respondents described their chance of developing cancer as the same as of other individuals, and "Pessimism" categories are almost equally distributed, 27.67% and 31.47% respectively.

3.4 Empirical strategy

The study employs a Fixed Effects regression model for the first analysis, as we are dealing with panel data with multiple observations for each individual over time, and it helps us to utilize the structure of our dataset in an effective way. In a Fixed effects model, the following assumptions are crucial in order to ensure the validity of the results. Strict exogeneity is essential, meaning that the idiosyncratic shock is uncorrelated with the independent variable des_H at any time point. It is not necessary to know if the individual heterogeneity is correlated with des_H, as we account for that correlation. This model addresses all observed and unobserved time-invariant sources of bias and estimates the effect of the control variables changing over time on going to screenings.

For the second analysis, we employed a logistic regression model, as it is the most appropriate to use due to the binary nature of the outcome variables concerning breast and cervical cancer screenings. This model is able to capture non-linearities and predict the probability of occurring events. By employing a logistic regression, it can capture the size of the association that health beliefs have on the probability of undergoing screening, allowing for a better understanding of the relationship.

Regarding the assumptions of a logistic regression model, the following ones should be satisfied. First, the assumption of appropriate outcome structure states that the dependent variables are appropriately used for the model. Second, the assumption of linearity of independent and log odds requires that there is a linear relationship between the independent variables and the log odds of the outcome. Third, the assumption of independence assumes that each observation should be independent of each other. Fourth, the assumption of absence of multicollinearity requires that the independent variables of the model should not be highly correlated with each other. The last assumption requires that there should be a large sample size.

Furthermore, we are clustering at an individual level. In that way, we account for correlated data within the same individuals and non-independence, which can increase the accuracy of our estimates and efficiency of the regression outcomes.

The following equations will be used to estimate the non-causal effect of health behaviors and health beliefs across the two analyses. Equations (1) and (2) correspond to the first panel data analysis, which examines the association between breast and cervical cancer screening and how people describe their health. Equations (3) and (4) correspond to the second analysis for the 2010 single wave study that analyzes the association between breast and cervical cancer screening and the individuals' health beliefs based on how they perceive themselves and others that share the same socio-economic characteristics.

(1) $\text{breast\_screening}_{it}=\alpha_i+\varphi_t+\sum_{c=3}^{c}\beta_c D_{cit}+\sum_{c=2}^{c}\beta_c J_{cit}+\sum_{c=2}^{c}\beta_c K_{cit}+\sum_{c=10}^{c}\beta_c S_{cit}+\sum_{c=4}^{c}\beta_c A_{cit}+\sum_{c=12}^{c}\beta_c D_{cit}A_{cit}+\sum_{c=3}^{c}\beta_c E_{ci}+\sum_{c=9}^{c}\beta_c D_{cit}E_{ci}+\varepsilon_{it}$

(2) $\text{cervical\_screening}_{it}=\alpha_i+\varphi_t+\sum_{c=3}^{c}\beta_c D_{cit}+\sum_{c=2}^{c}\beta_c J_{cit}+\sum_{c=2}^{c}\beta_c K_{cit}+\sum_{c=10}^{c}\beta_c S_{cit}+\sum_{c=4}^{c}\beta_c A_{cit}+\sum_{c=12}^{c}\beta_c D_{cit}A_{cit}+\sum_{c=3}^{c}\beta_c E_{ci}+\sum_{c=9}^{c}\beta_c D_{cit}E_{ci}+\varepsilon_{it}$

(3) $\text{breast\_screening\_dummy}_i=\beta_0+\sum_{c=3}^{c}\beta_c H_{ci}+\sum_{c=2}^{c}\beta_c J_{ci}+\sum_{c=2}^{c}\beta_c K_{cit}+\sum_{c=3}^{c}\beta_c A_{ci}+\sum_{c=9}^{c}\beta_c H_{ci}A_{ci}+\sum_{c=3}^{c}\beta_c E_{ci}+\sum_{c=9}^{c}\beta_c H_{ci}E_{ci}+\varepsilon_i$

(4) $\text{cervical\_screening\_dummy}_i=\beta_0+\sum_{c=3}^{c}\beta_c H_{ci}+\sum_{c=2}^{c}\beta_c J_{ci}+\sum_{c=2}^{c}\beta_c K_{cit}+\sum_{c=3}^{c}\beta_c A_{ci}+\sum_{c=9}^{c}\beta_c H_{ci}A_{ci}+\sum_{c=3}^{c}\beta_c E_{ci}+\sum_{c=9}^{c}\beta_c H_{ci}E_{ci}+\varepsilon_i$

In these regressions, variable D stands for how people describe their health, A represents the age categories, E the education categories, J indicates if the individual has a paid job or not and K if she has kids, S corresponds to the survey years and lastly H corresponds to the individuals' health beliefs according to their self-perception and perceptions of others with socio-economic characteristics. In addition, φ_t represents the year fixed effects, α_i is the individual fixed effects, εi is the error term and $\beta 0$ is the intercept term.

As previously mentioned, this study follows a non-causal approach, since establishing causation is challenging. In the model, there may still be variables, both observed and unobserved, as well as age and education, that could influence both the decision to undergo cancer screening and health beliefs, potentially introducing biases in the results. Despite these challenges, it is still interesting to see if those who view themselves as less or more likely to be healthy are more or less likely to go to screenings and employing a non-causal approach can still provide valuable insights on this association.

4. Results

4.1 Interpretation of the findings

4.1.1 First analysis- panel data analysis, 2008-2018

The results of the first and second regression from the first analysis are presented in Table 5. It includes the findings of the Fixed Effects model for both breast and cervical cancer screening.

Table 5: Results from the Fixed Effects model for breast and cervical cancer screening- 1st analysis

FIXED EFFECTS

	breast_screening (1)	**cervical_screening (2)**
des_H		
Negative	-0.0012 (0.0487)	0.0344 (0.0405)
Positive	0.098* (0.0559)	0.048 (0.0465)
Job	0.0217* (0.0123)	0.0484*** (0.0103)
Kids	0.0573*** (0.0152)	-0.0052 (0.0126)
des_H#education		
Neutral#Middle	-0.0263 (0.0764)	0.0564 (0.0636)
Neutral#High	0.0094 (0.0867)	0.0175 (0.0721)
Negative#Middle	-0.0101 (0.0801)	0.0633 (0.0667)
Negative#High	0.0748 (0.0912)	0.0356 (0.0759)
Positive#Middle	-0.147 (0.0916)	0.0040 (0.0763)
Positive#High	-0.0899 (0.1005)	-0.0356 (0.0836)
Agecat		
2=[40-51]	0.0864*** (0.0159)	-0.0344*** (0.0132)
3=[52-63]	0.0995*** (0.022)	0.333*** (0.0183)

4=[64-75]	-0.1664*** (0.0278)	0.2748*** (0.0231)
des_H#agecat		
Negative#2	-0.053** (0.0256)	-0.0245 (0.0213)
Negative#3	-0.0377 (0.0258)	-0.0411* (0.0214)
Negative#4	0.0051 (0.0275)	-0.0698*** (0.0229)
Positive#2	0.0346* (0.0209)	-0.0089 (0.0174)
Positive#3	0.0563** (0.022)	0.0088 (0.0184)
Positive#4	0.01685 (0.0251)	-0.0093 (0.0209)
survey_year		
9	-0.0052 (0.0099)	0.0167** (0.0083)
10	-0.0329*** (0.0103)	0.0252*** (0.0086)
11	-0.0286*** (0.0107)	0.0338*** (0.0089)
12	-0.032*** (0.0108)	0.0502*** (0.009)
13	-0.0514*** (0.0111)	0.0638*** (0.0093)
15	-0.0909*** (0.012)	0.0775*** (0.01)
16	-0.0838*** (0.0123)	0.0742*** (0.0103)
17	-0.1063*** (0.0127)	0.0961*** (0.0106)
18	-0.1116*** (0.0132)	0.1045*** (0.011)
_cons	0.6044*** (0.0766)	0.2529*** (0.0638)
sigma_u	0.3855	0.3611
sigma_e	0.3257	0.2712
rho	0.5834	0.6394

*SE in parentheses, ***p<0.01 **p<0.05 *p<0.1*

Association of positive health beliefs and cancer screening uptake

The models include 22,783 observations and the within R-squared is 0.0581 for breast screening and 0.1054 for cervical screening. The results reveal that on average, people who change from neutral or negative to positive health beliefs are 9.8 percentage points more likely to undergo a breast cancer screening, ceteris paribus. The effect is statistically significant for breast screening at 10% significance level. For cervical screening, changing from neutral or negative to positive health beliefs is associated with 4.8 percentage points increase in cancer screening uptake, ceteris paribus. However, the effect is not statistically significant at 10% significance level. Based on our findings, we fail to reject the 2nd hypothesis stating that positive health beliefs are associated with higher cancer screening uptake. This outcome aligns with the evidence from the available literature suggesting a positive relationship between positive health beliefs and preventive health behaviors.

Association of negative health beliefs and cancer screening uptake

For negative health beliefs, the sign of the results does not align in both types of cancer. It is presented that on average, people who change from neutral or positive to negative health beliefs are 0.12 percentage points less likely to undergo a breast cancer screening, ceteris paribus. For cervical screening, changing from neutral or positive to negative health beliefs is associated with 3.4 percentage points increase in cancer screening uptake, ceteris paribus. However, both effects are not statistically significant at 10% significance level. It is observed that the magnitude for breast screening is very small and with a negative sign while for cervical screening the magnitude is bigger with a positive sign. Therefore, for breast screening, the 4th hypothesis, stating that negative health beliefs are associated with higher cancer screening uptake, is rejected, while for cervical screening holds. This discrepancy in the results could be due to asymmetric and unclear information or restricted promotion about the benefits of breast screening by medical staff or national campaigns which may prevent the stimuli-responsive impact of negative beliefs from taking action.

Covariates

The most striking findings in our covariates are observed in the variable “kids” and in the two interaction terms of “agecat” and “education” with “des_H”. Regarding the variable “kids” there are big differences between the two types of cancer screening. For breast screening, the change from not having kids to having kids over the time period has a positive effect that is statistically significant at 1% while for cervical screening the effect is almost 10 times smaller, negative and not statistically significant at 10% significance level. In addition, there is a big difference in the magnitude and sign of the effect that individuals with a middle level of education, holding positive health beliefs, have in breast versus cervical cancer screenings. We observe that in breast screening the effect is negative and big with 14.7 percentage points while on cervical screening it is positive and small with 0.4 percentage points. Lastly, it is observed that for cervical cancer, the effect of individuals who hold positive health beliefs and belong in the second, third and fourth age category, corresponding to the ages from 40 till 75 years old in total, is very small, less than 1 percentage point, in comparison with the majority of the rest of the results from this interaction term.

4.1.2 Second Analysis- single wave analysis, 2010

The results of the third and fourth regressions from the second analysis are presented in Table 6. Table 6 presents the results of the marginal effects of the logistic regression model. We provide the marginal effects table in order to get measurable coefficients, since by just providing the findings of the logit model we can only talk about the sign and not the magnitude or significance. The results of the logistic regression model for breast and cervical cancer are presented in the first appendix.

Table 6: Results from margins of the Logit model for breast and cervical cancer screening- 2nd analysis

	MARGINS	
	breast_screening_dummy (3)	**cervical_screening_dummy (4)**
H_beliefs		
Pessimism	0.0176 (0.031)	-0.0686** (0.0271)
Optimism	0.0425 (0.0295)	-0.0776*** (0.0253)
Job	0.0522* (0.0281)	0.0233 (0.0274)
Kids	0.1078*** (0.0344)	0.0568* (0.0304)
Education		
Middle	-0.0726 (0.0571)	0.0589 (0.0552)
High	-0.0674 (0.0608)	0.0503 (0.0584)
Agecat		
3=[52-63]	-0.0767** (0.0335)	0.1681*** (0.0341)

4=[64-75]	-0.5235*** (0.0541)	0.0455 (0.0497)

*SE in parentheses, ***p<0.01 **p<0.05 *p<0.1*

Association of optimistic health beliefs and cancer screening uptake

The models include 1,160 observations and the Pseudo R2 is 0.2375 for breast screening and 0.0635 for cervical screening. From Table 6 we can conclude that on average, people having optimistic beliefs about their health are 4.2 percentage points more likely to undergo a breast cancer screening compared to those who hold a neutral health belief, ceteris paribus. The effect is not statistically significant at 10% significance level. For cervical screening, optimistic health beliefs are associated with 7.7 percentage points decrease in cancer screening uptake compared to neutral beliefs, ceteris paribus. The effect is statistically significant at 1% significance level. According to the results, we fail to reject the 1st hypothesis stating that optimistic health beliefs are associated with lower cancer screening uptake for cervical cancer but we reject it for breast cancer.

Association of pessimistic health beliefs and cancer screening uptake

For pessimistic health beliefs, the results show that on average people with pessimistic beliefs about their health are 1.7 percentage points more likely to undergo a breast cancer screening compared to those who hold a neutral health belief, ceteris paribus. The effect is not statistically significant at 10% significance level. For cervical screening, having a pessimistic outlook is associated with a 6.8 percentage points decrease in cancer screening uptake compared to neutral health beliefs, ceteris paribus. The effect is statistically significant at 5% significance level. Based on these outcomes, we fail to reject the 3rd hypothesis stating that pessimistic health beliefs are associated with lower cancer screening uptake for cervical cancer but we reject it for breast cancer.

The reason why for cervical cancer in both optimistic and pessimistic health beliefs we get statistically significant results might be due to the high number of respondents who engaged in cervical cancer screening, 963 observations. Additionally, we attribute the discrepancies in the different signs of the effect in breast and cervical cancer to potential campaigns and excessive promotion of breast cancer screening in 2010. These actions might have increased awareness and health consciousness and overshadowed the effects of optimistic and pessimistic health beliefs while influencing the individuals to get tested.

Covariates

It is worth mentioning the significant differences that are observed in the age and education categories. Firstly, the direction of the effect in the third and fourth age categories is opposite for breast and cervical cancer screening. In addition, the magnitude of the effect in the fourth category of breast screening is very large, 52.3 percentage points, which is almost ten times bigger than in cervical cancer, 4.5 percentage points. Hence, we can say that the influence of the 64-75 age category on breast cancer screening is substantial. Lastly, even though the magnitude of the education categories is similar in both types of cancer screening, the effect for breast screening has a negative sign, while for cervical screening it is positive.

5. Discussion and limitations

This section will summarize the results presented in the previous chapter in relation to our research questions. Further information on the credibility and limitations of the analyses will be provided and lastly suggestions for future research will be described.

From the first analysis we can assess the first and main research question: *"How can subjectively assessed health status affect the uptake of cancer screening tests in the case of breast and cervical cancer for Dutch women aged 28-75 years?"*. The answer for breast cancer is not entirely clear and it varies whether the individuals assess their health status in a positive or negative way. Hence, positive outlooks are positively correlated and negative outlooks are negatively correlated with breast cancer screening uptake. For cervical cancer, the answer is less complicated, as subjectively assessed health status, regardless if it is positive or negative, is positively correlated with cervical cancer screening uptake.

From the second analysis we can address the second research question: *"How can optimistic and pessimistic health beliefs affect breast and cervical cancer screening test uptakes for Dutch women aged 28-75 years?"*. The results of this analysis showed that, in cervical cancer, both optimistic and pessimistic health beliefs are negatively correlated with screening and their effects are statistically significant at 1% and 5% significance level, respectively. However, in breast cancer optimistic and pessimistic health beliefs are positively correlated with screening and their effects are not statistically significant at 10% significance level.

5.1 Credibility and reliability of the results

To ensure the reliability of our results, we checked the accuracy of our model of choice. For the first analysis, we conducted a Hausman test in order to determine whether a random effects or a fixed effects model is more appropriate. By running the Hausman test for both the regressions of breast and cervical screening, we get a p-value equal to 0. This allows us to reject the null hypothesis, implying that the difference in the coefficients between the Fixed Effects and Random Effects models is not systematic, at the 1% significance level, ceteris paribus. Hence, the coefficients are systematically different and the unobserved heterogeneity matters and influences the independent variable des_H. Therefore, the Fixed Effects model is the best estimator, as it can account for the correlation between this unobserved heterogeneity and des_H. The results of the Random Effects are presented in the second appendix. In the Random Effects results, there is variability in the magnitude of the coefficients, as well as some differences in the sign and statistical significance. Moreover, the within R2 for breast screening is 0.0564 and for cervical screening is 0.0943 which are lower than in the Fixed Effects model. This additional information indicates that the Fixed Effects model is a better fit for our data.

Since the unobserved heterogeneity matters and it is correlated with des_H, we cannot assume that there is exogeneity. However, we can assume that there is strict exogeneity, corresponding to the previously mentioned assumption of the Fixed Effects model. Hence, for the Fixed Effects to be unbiased, the idiosyncratic shock should not be correlated with the parameter of our interest, des_H. This could include a sudden health event for the individual, a change in the guidelines of cancer screening or of health insurance coverage, a new advanced method or technology and many more. These scenarios are very likely to happen, especially personal health and life events. For instance, the information obtained by RIVM (2023) indicates that, in 2010, the HPV vaccination program was implemented, which protects against cervical cancer. Therefore, it is unlikely that the strict exogeneity assumption holds, as these factors could possibly be correlated with the independent variable but cannot be measured. Nevertheless, it is still beneficial to use Fixed Effects as it can control for time-invariant characteristics such as genetic predisposition which can affect both health behaviors and the decision to undergo cancer screening and it can mitigate the impact of omitted variable bias in our study.

For the second analysis, to determine whether a Logit or an Ordinary Least Squared model is more appropriate, we ran the regressions using an OLS model and then we examined the predicted values of our dependent variables from this model. Based on the results, for breast screening we obtained values bigger than 1 and for cervical screening the minimum value was very high. Hence, a Logit model is more appropriate as it can predict values between 0.01 and 0.9, making the range of predictions closer to the actual values. The results of the OLS model are presented in the third appendix.

Regarding the validity of the assumptions of the logistic regression model, the following ones are satisfied. Firstly, the model has 1,160 observations, which can be considered as a sufficiently large sample compared to the small size of the population of the Netherlands. Secondly, the structure of the outcome is appropriate since our dependent variables are binary and take values 0 and 1. Thirdly, in order to address any potential violation of independence of the observation's assumption, we allowed for clustering at an individual level in our regressions. For the absence of multicollinearity assumption, we performed a Pairwise correlation test and the results showed that the correlation coefficients are smaller than 0.8. Therefore, the assumption holds. Lastly, the assumption of linearity of the independent variables and log odds is most likely to hold since we included interaction terms between des_H and age and education categories which can account for any potential nonlinear interactions. Additionally, the binary nature of the "job" and "kids" variables makes them suitable with the logit model as they are coded as 0 and 1.

5.2 Limitations

Our study has several limitations that should be identified and discussed. Firstly, we do not have information about what causes the individuals to change their positive or negative health beliefs over time. This means that there are probably some factors correlated with both the independent variable des_H and the error term, which can lead to biased estimates. We try to control for this by using the fixed effects model which takes into account time invariant factors. Yet, there may be time-variant factors that lead to a change in the self-assessed health status and cancer screenings behaviors, such as a health shock. This issue could be addressed by using an Instrumental Variable model; however, we were not able to find an appropriate instrument for this study.

Moreover, as already mentioned in chapter 2, evidence from the literature has shown that education and age can influence both the decision to undergo cancer screening and individuals' health beliefs. Despite the fact that we control for education in the fixed effects model, as we assume that it remains constant, we performed multiple subgroup analyses, in both panel data analysis and the 2010 single wave analysis, for each age and education category and for both types of cancer screening. The results presented in the appendices 4-8 are completely or partially different from our main analyses, both the first and the second types of analysis, except for those concerning the second education category, corresponding to middle level of education, for cervical cancer screening. We were aware of the endogeneity issues caused by the variables of age and education and we attempted to address them by including interaction terms. However, the subgroup analyses showed that age and education can modify the effect of our independent variables, des_H for the first analysis and H_beliefs for the second analysis, on breast and cervical screening. Even though this creates bias to our estimates, it also explains that this effect varies across the different levels of age and education. Regarding the similarity in the results with the second education category on cervical screening, this might be attributed to the high number of observations, since 63% and 68% of the observations in the first and second analyses, respectively are accumulated in this category. Nevertheless, it is unclear why we do not observe the same similarity in breast cancer screening. Lastly, despite this limitation, the scope of this study is to look at what happens to cancer screening when self-perceived health status changes rather than identifying the reasons behind these changes.

Furthermore, attrition is a crucial concern when dealing with panel data, as participants may drop out over time from the survey. This can cause bias in our results as the remaining sample might not be representative for the Dutch population. As expected, there was attrition in our sample since for both breast and cervical cancer screening, the median length of observation is 3 waves, which is considered relatively high attrition. In order to check whether attrition is correlated with our dependent variables we performed the next wave test. The results of the next wave coefficients have positive signs, meaning that participants who go to breast and cervical cancer screening are correlated with being in a higher wave. Hence, those who do not go to screenings are more likely to drop out over time from the survey. The coefficient corresponding to breast screening analysis is not statistically significant at 10%. This indicates that the attrition is at random, suggesting a weak relationship between attrition and undergoing breast cancer screening. However, the coefficient for the cervical screening analysis is statistically significant at 5%, indicating that attrition is not at random, which suggests a strong relationship between attrition and undergoing cervical cancer screening. This can introduce bias in our results, as attrition can also be correlated with other components linked to health beliefs. For instance, individuals with positive or negative health beliefs might trust less their physicians or the healthcare systems and may be more likely to drop out from the survey and not undergo screening.

Lastly, the issue of potential reverse causality may arise as well, as evidence from Skinner's et al. (1998) study suggest that women who had at least one mammogram in the past, were more likely to be optimists and less likely to be pessimists in comparison to women that had never had a mammogram. However, we do not have any information whether the individuals have undergone screening outside of our researcher's time interval. Additionally, Skinner's et al. (1998) do not specifically address whether this applies in the case of just positive and negative self-perceptions. Yet, even if reverse causality is a factor and respondents have gone screening in the past and received negative results, they should technically still go for the next screening, as there is a chance that they may have developed cancer in between.

5.3 Future research

This study's findings and limitations can inspire several interesting paths and ideas for future research. The following suggestions can lead to targeted policies that address potential obstacles and increase participation in cancer screening tests.

Particularly, incentives for conducting surveys that account for the objective measures of the true health risk that individuals could develop cancer are very valuable in this field. This approach allows researchers to give more accurate definitions of overly optimistic and pessimistic as well as just positive and negative health beliefs, which might be realistic, by comparing the individuals' self-assessed health status to their actual health status. Hence, the policymakers might identify a specific category of health beliefs that needs extra attention and, in general, decide to focus their efforts on areas that can achieve a higher increase in cancer screenings.

Additionally, it could be useful if the survey was not conducted only online but also through personal interviews. As a consequence, older individuals who are not familiar with technology will be able to participate. Given that they have a higher probability of developing cancer, it is essential that they regularly undergo screenings, (WHO, 2022).

Last but not least, this thesis provides evidence for people with low, middle and high education levels. However, in the Netherlands, the rate of secondary education enrollment of the Dutch population is very high, ranking it in the 5th position worldwide (The world economic forum 2018), so the results may be representative for countries with similar educational attainment like Norway and the UK (World population review, 2024). Therefore, this study could serve as an example for conducting similar research in other countries that might have broader external validity than the Netherlands. For instance, conducting this study in a country with a low to medium level of education might achieve results that are more representative for the majority of countries in Central and Western Europe.

6. References

American Cancer Society (3rd January 2020). Risk factors for cervical cancer. Retrieved from: https://www.cancer.org/cancer/types/cervical-cancer/causes-risks-prevention/risk-factors.html

Arni P., Dragone D., Goette L., & Ziebarth N. R.(2021) Biased health perceptions and risky health behaviors—Theory and evidence

Bago d'Uva T., O'Donnell O. & van Doorslaer E. (2020) Who can predict their own demise? Heterogeneity in the accuracy and value of longevity expectations. The Journal of the Economics of Ageing.

Breast Cancer Organization (18th April 2024). Breast cancer risk factors. Retrieved from: https://www.breastcancer.org/risk/risk-factors

Chipperfield, J.G. et al. (2019) 'A healthy dose of realism: The role of optimistic and pessimistic expectations when facing a downward spiral in health,' Social Science & Medicine, 232, pp. 444–452

Deeks, A. et al. (2009) 'The effects of gender and age on health-related behaviors,' BMC Public Health, 9(1).

Dickie, M., & Gerking, S. (1997). Genetic Risk Factors and Offsetting Behavior: The Case of Skin Cancer. Journal of Risk and Uncertainty, 15(1), 81–97.

Eibich, P., & Goldzahl, L. (2020). Health information provision, health knowledge and health behaviors: Evidence from breast cancer screening. Social Science & Medicine, 265

Eurostat (12th July 2023). Preventive cancer screenings - program data 2021. Retrieved from: https ://ec.europa.eu/eurostat/databrowser/view/HLTH_PS_PREV__custom_9971053/default/table?lang=en

Galai, D., & Sade, O. (2003). The 'ostrich effect' and the relationship between the liquidity and the yields of financial assets, SSRN working paper.

Ghirardato, P., Maccheroni, F., Marinacci, M. (2004). Differentiating ambiguity and ambiguity attitude. Journal of Economic Theory, 118(2), 133–173

Golman, R., Hagmann, D., & Loewenstein, G. (2017). Information Avoidance. Journal of Economic Literature, 55(1), 96–135.

Hazlett, A., Molden, D.C. and Sackett, A.M. (2011) 'Hoping for the best or preparing for the worst? Regulatory focus and preferences for optimism and pessimism in predicting personal outcomes,' Social Cognition, 29(1), pp. 74–96.

Hoy, M., Peter, R., & Richter, A. (2014). Take-up for genetic tests and ambiguity. Journal of Risk and Uncertainty, 48(2), 111–133

Karlsson, N., Loewenstein, G., & Seppi, D. (2009). The ostrich effect: Selective attention to information. Journal of Risk and Uncertainty, 38(2), 95–115.

Kőszegi, B. (2003). Health anxiety and patient behavior. Journal of Health Economics, 22(6), 1073–1084.

Leder, S., Florack, A. and Keller, J. (2014) 'Self-regulation and protective health behavior: How regulatory focus and anticipated regret are related to vaccination decisions,' Psychology & Health, 30(2), pp. 165–188.

Levy, B.R. and Myers, L.M. (2004) 'Preventive health behaviors influenced by self-perceptions of aging,' Preventive Medicine, 39(3), pp. 625–629.

Oster, E., Shoulson, I., & Dorsey, E. R. (2013). Optimal Expectations and Limited Medical Testing: Evidence from Huntington Disease. American Economic Review, 103(2), 804–830.

Picone, G., Sloan, F., & Taylor, D. (2004). Effects of Risk and Time Preference and Expected Longevity on Demand for Medical Tests. Journal of Risk and Uncertainty, 28(1), 39–53.

RIVM (2023). HPV vaccination. Retrieved from: https://www.rivm.nl/en/hpv-vaccination

Schneider, S.L. (2001b) 'In search of realistic optimism: Meaning, knowledge, and warm fuzziness.,' American Psychologist/the American Psychologist, 56(3), pp. 250–263.

Skinner, C.S., Kreuter, M.W., Kobrin, S., & Strecher, V.J. (1998) 'Perceived and actual breast cancer risk,' Journal of Health Psychology, 3(2), pp. 181–193.

Spitzer, S., & Shaikh, M. (2022). Health misperception and healthcare utilization among older Europeans. The Journal of the Economics of Ageing, 22

The world economic forum (2018). Country Profiles-Netherlands 2017-2018. Retrieved from: https://www3.weforum.org/docs/GCR2017-2018/03CountryProfiles/Standalone2-pagerprofiles/WEF_GCI_2017_2018_Profile_Netherlands.pdf

WHO (3rd February 2022). Cancer-Facts 2022. Retrieved from: https://www.who.int/news-room/fact-sheets/detail/cancer

WHO (5th March 2024). Cervical cancer 2022. Retrieved from: https://www.who.int/news-room/fact-sheets/detail/cervical-cancer

WHO (13th March 2024). Breast cancer 2022. Retrieved from: https://www.who.int/news-room/fact-sheets/detail/breast-cancer

World Population Review (2024). Most Educated Countries 2024. Retrieved from: https://worldpopulationreview.com/country-rankings/most-educated-countries

7. Appendices

Appendix 1: Results from the Logit model for breast and cervical cancer screening- 2nd analysis

LOGIT MODEL

	breast_screening_dummy	**cervical_screening_dummy**
H_beliefs		
Pessimism	-1.5843 (1.3021)	-2.0459 (1.265)
Optimism	-1.2466 (0.9926)	-1.2235 (1.1475)
Job	0.3339* (0.1805)	0.1755 (0.207)
Kids	0.689*** (0.2219)	0.428* (0.2292)
H_beliefs#education		
Neutral#Middle	-1.4898* (0.7715)	-0.6333 (1.0145)
Neutral#High	-1.0628 (0.8048)	-0.5751 (1.052)
Pessimism#Middle	0.0325 (0.955)	1.0585 (0.6677)
Pessimism#High	-0.1304 (0.9834)	1.1567 (0.7197)
Optimism#Middle	-0.2274 (0.434)	0.2414 (0.4348)
Optimism#High	-0.3274 (0.4805)	0.0122 (0.4727)
Agecat		
3=[52-63]	-1.0235*** (0.382)	1.1005*** (0.4453)
4=[64-75]	-2.6163*** (0.4473)	0.7897* (0.4671)
H_beliefs#agecat		
Pessimism#3	0.8252* (0.4809)	0.4015 (0.5614)
Pessimism#4	-0.0635 (0.5833)	-0.6684 (0.5544)
Optimism#3	0.71 (0.4807)	0.1796 (0.5277)
Optimism#4	0.3236	-0.6414

	(0.5227)	(0.5074)
_cons	2.6903*** (0.8718)	1.7587 (1.0868)

*SE in parentheses, ***p<0.01 **p<0.05 *p<0.1*

Appendix 2: Results from the Random Effects model of the Hausman test for breast and cervical cancer screening-1st analysis

RANDOM EFFECTS

	breast_screening	**cervical_screening**
des_H		
Pessimistic	-0.0018 (0.041)	0.0806** (0.0337)
Optimistic	0.0758 (0.0499)	0.009 (0.0413)
job	0.0407*** (0.0094)	0.0088 (0.0077)
kids	0.0977*** (0.0095)	-0.0245*** (0.0077)
des_H#education		
Neutral#Middle	0.0269 (0.0295)	0.0545** (0.0234)
Neutral#High	0.0374 (0.0306)	0.0175 (0.0243)
Negative#Middle	0.0294 (0.0355)	0.0266 (0.0285)
Negative#High	0.0804** (0.0381)	-0.0061 (0.0305)
Positive#Middle	-0.068 (0.0507)	0.0394 (0.0413)
Positive#High	-0.0366 (0.0514)	-0.0038 (0.0418)
agecat		
2=[40-51]	0.071*** (0.0116)	0.1341*** (0.0095)
3=[52-63]	0.0468*** (0.0132)	0.6866*** (0.0107)

4=[64-75]	-0.2998*** (0.0163)	0.6936*** (0.0131)
des_H#agecat		
Negative#2	-0.0452* (0.0232)	-0.0248 (0.0192)
Negative#3	-0.0305 (0.0229)	-0.0615*** (0.019)
Negative#4	0.0164 (0.0245)	-0.0886*** (0.0203)
Positive#2	0.0296 (0.0188)	-0.0136 (0.0156)
Positive#3	0.0447** (0.0197)	0.0048 (0.0163)
Positive#4	0.0164 (0.0223)	-0.0169 (0.0184)
survey_year		
9	-0.006 (0.0096)	0.0056 (0.008)
10	-0.0288*** (0.0099)	0.0029 (0.0083)
11	-0.0214 ** (0.0102)	-0.0019 (0.0085)
12	-0.0198** (0.01)	0.0062 (0.0084)
13	-0.035*** (0.0102)	0.007 (0.0086)
15	-0.064*** (0.0108)	0.0006 (0.009)
16	-0.0492*** (0.0106)	-0.0202** (0.0088)
17	-0.0724*** (0.0105)	-0.0076 (0.0087)
18	-0.0723*** (0.0108)	-0.0112 (0.009)
_cons	0.5543*** (0.0318)	0.096*** (0.0254)
sigma_u	0.3097	0.2311
sigma_e	0.3257	0.2712
rho	0.4748	0.4206
Observations	22,783	22,783

*SE in parentheses, ***p<0.01 **p<0.05 *p<0.1*

Appendix 3: Results from the OLS model for breast and cervical cancer screening-2nd analysis

	OLS	
	breast_screening_dummy	**cervical_screening_dummy**
H_beliefs		
Pessimism	-0.2434 (0.1927)	-0.3455* (0.1808)
Optimism	-0.1913 (0.142)	-0.1516 (0.1089)
job	0.0558 (0.0311)	0.02 (0.0255)
kids	0.1022*** (0.0317)	0.0508* (0.028)
H_beliefs#education		
Neutral#Middle	-0.2457** (0.1133)	-0.0525 (0.0647)
Neutral#High	-0.1751 (0.1186)	-0.0471 (0.0699)
Pessimism#Middle	0.0049 (0.1507)	0.2365 (0.1616)
Pessimism#High	-0.0189 (0.1546)	0.2466 (0.1646)
Optimism#Middle	-0.0406 (0.0773)	0.0407 (0.0757)
Optimism#High	-0.0565 (0.0837)	0.0062 (0.0821)
agecat		
3=[52-63]	-0.1628*** (0.0544)	0.1162** (0.0464)
4=[64-75]	-0.54*** (0.0648)	0.0955* (0.0569)
H_beliefs#agecat		
Pessimism#3	0.1369** (0.0683)	0.0612 (0.0602)
Pessimism#4	-0.0327 (0.0803)	-0.1039 (0.0858)
Optimism#3	0.1244* (0.066)	0.0508 (0.061)
Optimism#4	0.0409 (0.0731)	-0.099 (0.0739)
_cons	0.9874***	0.8285***

	(0.12)	(0.0799)
Observations	1,160	1,160

*SE in parentheses, ***p<0.01 **p<0.05 *p<0.1*

Appendix 4: Results from the sub-group analysis of the education categories - 1st analysis

FIXED EFFECTS

	Education Categories					
	(1)		**(2)**		**(3)**	
	breast screening	cervical screening	breast screening	cervical screening	breast screening	cervical screening
des_H						
Negative	-0.0767 (0.2059)	-0.0143 (0.195)	0.0023 (0.0261)	0.0418* (0.0223)	0.0724* (0.0368)	0.0562** (0.0287)
Positive	0.1686 (0.1317)	-0.0087 (0.1247)	-0.0513** (0.0228)	-0.0059 (0.0195)	0.029 (0.0244)	-0.0036 (0.019)
job	-0.0634 (0.0511)	0.0376 (0.0484)	0.0181 (0.0152)	0.0405*** (0.013)	0.0213 (0.0239)	0.0547*** (0.0186)
kids	-0.2379* (0.1338)	-0.067 (0.1267)	0.0095 (0.0197)	-0.0031 (0.0168)	0.1053*** (0.0247)	-0.017 (0.0192)
agecat						
2=[40-51]	0.2728** (0.1154)	0.0419 (0.1093)	0.0986*** (0.0195)	-0.0332** (0.0167)	0.0547* (0.0282)	-0.0433** (0.0219)
3=[52-63]	0.3096** (0.1442)	0.2871** (0.1365)	0.1332*** (0.0268)	0.3593*** (0.0229)	0.0256 (0.0402)	0.2784*** (0.0313)
4=[64-75]	-0.0087 (0.1649)	0.4159*** (0.1561)	-0.1008*** (0.0337)	0.2981*** (0.0289)	-0.2868*** (0.0523)	0.197*** (0.0407)
des_H#agecat						
Negative#2	0.0515 (0.2342)	-0.0277 (0.2217)	-0.0592* (0.0309)	-0.0225 (0.0264)	-0.0224 (0.0469)	-0.0235 (0.0366)
Negative#3	0.0519 (0.2166)	0.1651 (0.2051)	-0.0161 (0.0313)	-0.0576** (0.0267)	-0.074 (0.0473)	-0.0344 (0.0368)
Negative#4	0.0973 (0.2159)	-0.061 (0.2044)	0.0267 (0.033)	-0.0593** (0.0283)	-0.0235 (0.0531)	-0.0885** (0.0414)
Positive#2	-0.1647 (0.2124)	0.104 (0.2011)	0.0606** (0.0281)	0.0016 (0.024)	0.0098 (0.0325)	-0.0272 (0.0253)
Positive#3	-0.0845	0.1733	0.0971***	-0.0098	0.0118	0.0308

	(0.1544)	(0.1462)	(0.0293)	(0.025)	(0.0353)	(0.0275)
Positive#4	0.0337 (0.1541)	-0.0196 (0.1459)	0.0694** (0.0316)	0.0064 (0.0270	-0.1073** (0.0457)	-0.0384 (0.0356)
survey_year						
9	0.0189 (0.0479)	-0.0186 (0.0454)	-0.0184 (0.0121)	0.0186* (0.0103)	0.0177 (0.019)	0.0173 (0.0148)
10	-0.095* (0.0496)	-0.042 (0.047)	-0.0466*** (0.0125)	0.0227** (0.0107)	0.0038 (0.0196)	0.0381** (0.0153)
11	-0.0878* (0.0529)	-0.0432 (0.0501)	-0.0497*** (0.013)	0.0297*** (0.0111)	0.0177 (0.0204)	0.0502*** (0.0159)
12	-0.0563 (0.0554)	-0.0542 (0.0525)	-0.0726*** (0.0131)	0.0414*** (0.0112)	0.0468** (0.0204)	0.0821*** (0.0159)
13	-0.0954 (0.0591)	-0.0604 (0.0559)	-0.0757*** (0.0135)	0.0618*** (0.0116)	-0.0025 (0.0209)	0.0807*** (0.0163)
15	-0.0976 (0.0637)	-0.0463 (0.0603)	-0.1237*** (0.0146)	0.0721*** (0.0125)	-0.0306 (0.0225)	0.102*** (0.0175)
16	-0.1597** (0.0668)	-0.0855 (0.0632)	-0.1305*** (0.0151)	0.0657*** (0.0129)	0.0056 (0.0228)	0.1059*** (0.0178)
17	-0.1717** (0.0693)	-0.0825 (0.0656)	-0.1528*** (0.0156)	0.0922*** (0.0134)	-0.0171 (0.0233)	0.124*** (0.0182)
18	-0.2366*** (0.0719)	-0.0509 (0.068)	-0.1783*** (0.0164)	0.0904*** (0.014)	0.0105 (0.024)	0.1471*** (0.0187)
_cons	0.4536*** (0.1319)	0.4098*** (0.1249)	0.6057*** (0.0234)	0.3204*** (0.02)	0.5779*** (0.0327)	0.2328*** (0.0255)
sigma_u	0.4284	0.3597	0.3996	0.3539	0.381	0.3752
sigma_e	0.3262	0.3089	0.3184	0.2724	0.3368	0.2625
rho	0.6329	0.5756	0.6117	0.6279	0.5613	0.6715
Observations	889	889	14,438	14,438	7,456	7,456

*SE in parentheses, ***p<0.01 **p<0.05 *p<0.1*

Appendix 5: Results from the sub-group analysis of the age categories for breast screening- 1st analysis

FIXED EFFECTS

	Age Categories			
	(1)	**(2)**	**(3)**	**(4)**
des_H				
Negative	-0.0652	-0.0527	-0.0523	-0.0814

	(0.2013)	(0.1212)	(0.0795)	(0.0645)
Positive	0.1085 (0.1496)	-0.0187 (0.1882)	0.0523 (0.0967)	0.1838** (0.0774)
job	0.0033 (0.0413)	-0.0522 (0.0339)	-0.034 (0.0244)	0.0718*** (0.0219)
kids	0.0687** (0.0331)	0.0051 (0.0427)	-0.0146 (0.0245)	0.1389** (0.065
des_H#education				
Neutral#Middle	-0.0713 (0.2629)	0.338 (0.2628)	-0.0474 (0.1622)	-0.0524 (0.1264)
Neutral#High	0.0663 (0.2727)	0.1387 (0.2829)	-0.117 (0.1809)	-0.3796 (0.235)
Negative#Middle	-0.0209 (0.2765)	0.3642 (0.263)	-0.0045 (0.1666)	0.0482 (0.1264)
Negative#High	0.1655 (0.2875)	0.247 (0.2848)	-0.0356 (0.1873)	-0.2853 (0.2368)
Positive#Middle	-0.204 (0.2973)	0.3227 (0.3234)	-0.0504 (0.1897)	-0.2455* (0.1402)
Positive#High	-0.004 (0.3051)	0.1789 (0.3407)	-0.1093 (0.2059)	-0.5761** (0.245)
survey_year				
9	0.017 (0.0221)	0.0046 (0.018)	-0.0089 (0.0178)	-0.0579** (0.0238)
10	0.0256 (0.0236)	-0.0088 (0.0189)	-0.0679*** (0.0186)	-0.0907*** (0.0241)
11	0.0852*** (0.0252)	-0.0228 (0.0199)	-0.0842*** (0.0193)	-0.0989*** (0.0246)
12	0.0912*** (0.0262)	-0.0167 (0.0201)	-0.0817*** (0.0195)	-0.1196*** (0.0245)
13	0.118*** (0.0273)	-0.0378* (0.0208)	-0.1212*** (0.0201)	-0.1493*** (0.0246)
15	0.0964*** (0.0308)	-0.0771*** (0.0227)	-0.162*** (0.0216)	-0.191*** (0.0252)
16	0.195*** (0.0315)	-0.0797*** (0.0233)	-0.1605*** (0.0224)	-0.2251*** (0.0256)
17	0.2312*** (0.0328)	-0.0832*** (0.0242)	-0.1865*** (0.0229)	-0.2834*** (0.0263)
18	0.2578*** (0.0338)	-0.1075*** (0.0252)	-0.1893*** (0.0242)	-0.3126*** (0.027)
_cons	0.5148** (0.2627)	0.5647** (0.2636)	0.8843*** (0.1586)	0.4892*** (0.1266)
sigma_u	0.4496	0.3858	0.4164	0.4095

sigma_e	0.3487	0.3126	0.3095	0.2893
rho	0.6244	0.6037	0.6441	0.6672
Observations	4,887	6,120	6,780	4,996

*SE in parentheses, ***p<0.01 **p<0.05 *p<0.1*

Appendix 6: Results from the sub-group analysis of the age categories for cervical screening- 1st analysis

		FIXED EFFECTS		
		Age Categories		
	(1)	**(2)**	**(3)**	**(4)**
des_H				
Negative	-0.0907 (0.1154)	-0.0297 (0.1232)	0.1686*** (0.057)	-0.1126* (0.06)
Positive	-0.0048 (0.0858)	0.0418 (0.1913)	0.0473 (0.0693)	0.0448 (0.072)
job	0.0829*** (0.0237)	0.0112 (0.0345)	0.0039 (0.0175)	-0.0017 (0.0204)
kids	0.0094 (0.019)	-0.0267 (0.0434)	-0.0097 (0.0175)	0.0054 (0.0604)
des_H#education				
Neutral#Middle	-0.3429** (0.1507)	0.4293 (0.2671)	0.0935 (0.1163)	0.1445 (0.1175)
Neutral#High	-0.3908** (0.1563)	0.4118 (0.2876)	0.0948 (0.1297)	0.1041 (0.2185)
Negative#Middle	-0.2145 (0.1585)	0.4968* (0.2674)	-0.104 (0.1195)	0.2672** (0.1176)
Negative#High	-0.2401 (0.1648)	0.5062* (0.2895)	-0.0605 (0.1343)	0.1718 (0.2202)
Positive#Middle	-0.3513** (0.1704)	0.3831 (0.3287)	0.0387 (0.136)	0.0864 (0.1304)
Positive#High	-0.3916** (0.1749)	0.347 (0.3463)	0.045 (0.1477)	0.0148 (0.2278)
survey_year				
9	0.0004 (0.0126)	0.0562*** (0.0183)	0.0084 (0.0128)	0.0068 (0.0221)
10	0.0009 (0.0135)	0.1393*** (0.0192)	0.0019 (0.0133)	-0.0236 (0.0224)

11	0.0126 (0.0144)	0.1864*** (0.0202)	-0.0026 (0.0138)	-0.0371 (0.0229)
12	0.0365** (0.015)	0.2369*** (0.0205)	-0.0157 (0.014)	-0.0355 (0.0228)
13	0.038** (0.0157)	0.2801*** (0.0212)	-0.0208 (0.0144)	-0.0255 (0.0229)
15	0.0596*** (0.0177)	0.3548*** (0.0231)	-0.0317** (0.0155)	-0.0651*** (0.0234)
16	0.0516*** (0.018)	0.3484*** (0.0237)	-0.0532*** (0.0161)	-0.0916*** (0.0238)
17	0.0669*** (0.0188)	0.3856*** (0.0246)	-0.0327** (0.0164)	-0.0778*** (0.0245)
18	0.0776*** (0.0194)	0.4244*** (0.0256)	-0.0463*** (0.0174)	-0.086*** (0.0251)
_cons	0.3309** (0.1506)	-0.3935 (0.2679)	0.8079*** (0.1137)	0.7766*** (0.1177)
sigma_u	0.2539	0.4374	0.2871	0.3141
sigma_e	0.1999	0.3178	0.2219	0.2689
rho	0.6174	0.6546	0.626	0.577
Observations	4,887	6,120	6,780	4,996

*SE in parentheses, ***p<0.01 **p<0.05 *p<0.1*

Appendix 7: Results from the sub-group analysis of the education categories, margins- 2nd analysis

LOGIT

	Education Categories					
	(1)		**(2)**		**(3)**	
	breast screening	**cervical screening**	**breast screening**	**cervical screening**	**breast screening**	**cervical screening**
H_beliefs						
Pessimism	-0.165 (0.1611)	-0.2972** (0.1494)	0.0561 (0.037)	-0.0569* (0.0322)	-0.0384 (0.059)	-0.0497 (0.0524)
Optimism	-0.106 (0.1039)	-0.152* (0.0924)	0.0763** (0.0353)	-0.0621** (0.0304)	-0.0098 (0.0596)	-0.1032** (0.0517)
job	0.0464* (0.0256)	0.0264 (0.0308)	0.0524* (0.0281)	0.023 (0.0271)	0.0526* (0.0286)	0.0237 (0.0282)
kids	0.0958*** (0.0326)	0.0643* (0.0352)	0.1081*** (0.0345)	0.0562* (0.0301)	0.1085*** (0.035)	0.0579* (0.0314)

education						
Middle	-0.0726 (0.0571)	0.0589 (0.0552)	-0.0726 (0.0571)	0.0589 (0.0552)	-0.0726 (0.0571)	0.0589 (0.0552)
High	-0.0674 (0.0608)	0.0503 (0.0584)	-0.0674 (0.0608)	0.0503 (0.0584)	-0.0674 (0.0608)	0.0503 (0.0584)
agecat						
3=[52-63]	-0.0458 (0.0299)	0.2037*** (0.0491)	-0.0805** (0.0343)	0.1657*** (0.0343)	-0.0722** (0.0342)	0.171*** (0.038)
4=[64-75]	-0.4594*** (0.0678)	0.039 (0.057)	-0.5277*** (0.0548)	0.0459 (0.0492)	-0.5212*** (0.0545)	0.0457 (0.0513)
Observations	1,160	1,160	1,160	1,160	1,160	1,160

*SE in parentheses, ***p<0.01 **p<0.05 *p<0.1*

Appendix 8: Results from the sub-group analysis of the age categories, margins- 2nd analysis

LOGIT

	Age categories					
	(2)		**(3)**		**(4)**	
	breast screening	**cervical screening**	**breast screening**	**cervical screening**	**breast screening**	**cervical screening**
H_beliefs						
Pessimism	-0.0357 (0.0532)	-0.0785 (0.058)	0.1036** (0.052)	-0.0048 (0.0322)	-0.074 (0.0795)	-0.1519** (0.06)
Optimism	-0.0201 (0.0523)	-0.0838 (0.0581)	0.1026** (0.0495)	-0.0229 (0.0312)	0.0252 (0.0689)	-0.1514*** (0.048)
job	0.0459* (0.0265)	0.0325 (0.0392)	0.0593* (0.0323)	0.0137 (0.0164)	0.0678* (0.0386)	0.0278 (0.0313)
kids	0.0947*** (0.0368)	0.0793* (0.0458)	0.1224*** (0.0373)	0.0334* (0.0178)	0.14*** (0.0479)	0.0678** (0.0339)
education						
Middle	-0.0449 (0.0502)	0.073 (0.0705)	-0.0796 (0.0604)	0.035 (0.0389)	-0.1132 (0.084)	0.0798 (0.064)
High	-0.0443 (0.0535)	0.0622 (0.0756)	-0.0708 (0.0646)	0.0295 (0.0405)	-0.1062 (0.0882)	0.0688 (0.0678)
agecat						
3=[52-63]	-0.0767** (0.0335)	0.1681*** (0.0341)	-0.0767** (0.0335)	0.1681*** (0.0341)	-0.0767** (0.0335)	0.1681*** (0.0341)

4=[64-75]	-0.5235*** (0.0541)	0.0455 (0.0497)	-0.5235*** (0.0541)	0.0455 (0.0497)	-0.5235*** (0.0541)	0.0455 (0.0497)
Observations	1,160	1,160	1,160	1,160	1,160	1,160

*SE in parentheses, ***p<0.01 **p<0.05 *p<0.1*