

# **Education and Smoking Cessation: The Role of Skills Explained**

**A.M.T. Roest**

Erasmus School of Economics, Erasmus University Rotterdam

Master Thesis

September 2015

Supervisor: Hans van Kippersluis

Name: Alexander Roest

Student ID: 290406AR

Study: Economics & Business, specialization Health Economic

# Abstract

This thesis analyses the relationship between education and smoking cessation. It focuses on the soft skills as an interpretation of the relationship. This thesis tests the theory that education leads to better skills, and those skills are being used for smoking cessation, which has been considered as one of the plausible explanations. Longitudinal data, gathered from a survey among Dutch households, confirms a jointly significant influence of the various educational levels and the likelihood to quit smoking. The association of various soft skills, such as following schedules, is tested for their relationship with education using OLS regressions. To show the effect runs from education to skills, rather than the other way around, the interaction effect of age and being a student on smoking cessation is also tested. To complete the theory the relationship between skills and quitting smoking is tested using a fixed effects regression. The theory could be proven if any of the skills showed significant effects in all regressions, and the causality was significantly positive. This was not the case for any of the skills however. It was shown that skills were involved in the relationship between education and smoking, but this relationship isn't a causal one. Also shown was that income was both significantly positively associated with both education and quitting smoking.

# Contents

- Chapter 1: Introduction..... 3
  - Motivation ..... 3
  - Research question ..... 4
  - Outline ..... 5
- Chapter 2: Theoretical Framework ..... 6
  - Linking education and smoking..... 6
  - Explanations ..... 7
  - Criticism..... 8
  - Quitting smoking and education ..... 10
  - The role of skills..... 12
  - Smoking and Health ..... 15
  - Hypothesis development ..... 16
  - Hypotheses..... 17
- Chapter 3: Data & Methods ..... 19
  - Data ..... 19
  - Variables ..... 20
  - Methods ..... 21
- Chapter 4: Results ..... 25
  - Education and Quitting Smoking..... 25
  - Education and Quitting Smoking; the effect for men ..... 28
  - Education and Quitting Smoking; the effect for women ..... 29
  - Skills..... 30
  - Education and Income..... 31
  - Education and Skills..... 32
  - Age and Skills..... 34
  - Skills and Smoking Cessation..... 37
  - Education Skills and Smoking Cessation..... 39
- Chapter 5: Conclusions..... 42
- Chapter 6: Discussion ..... 44
  - Survey and variables..... 44
  - Methods and results..... 46
  - Looking ahead ..... 48
- Literature..... 49

# Chapter 1: Introduction

## Motivation

Smoking became more common through time. Its history started between 5000 and 3000 BC in South America (Tayade, 2014). Several civilizations and religions used it as part of rituals, but it was later used for pleasure (Tayade, 2014). International trade brought tobacco to Europe. In 1560 it was introduced to France by Jean Nicot, and its trade spread to England after that (Tayade, 2014). Its users were originally the elites (Vedøy, 2014). Cigarette use has not remained constant over time however. In the middle of the seventeenth century all major civilizations had access to tobacco, but many rulers tried to forbid the practice of smoking (Tayade, 2014). The use of cigarettes was extremely low by the end of the 1800s (Vedøy, 2014).

In 20th century more attention was given to the relationship between smoking and health. Strong associations between lung cancer and smoking were found (Doll & Hill, 1950). In 1954 its evidence was confirmed in the British Doctors Study, after which the government advised against the practice of smoking (Tayade, 2014). Cigarette smoking had its peak in the 1960s (Vedøy, 2014). In 1964, the United States Surgeon General Report published the dangers of smoking for its first time, and especially the highly educated people started smoking less (de Walque, 2004). The rates of smoking in the USA declined from 42% in 1965 to 20.8% in 2006 (Tayade, 2014).

Over time more has become known about the effects of smoking. Even now there are studies that provide more evidence that smoking is bad for health (Sasco, Secretan & Straif, 2004). Initially limited effects of smoking on diseases such as lung cancer were found. Recent evidence is provided to a wider spectrum of diseases (Cutler and Lleras-Muney, 2010). Smoking can also lead to strokes, blindness, heart disease, diabetes, chronic obstructive airways disease, tuberculosis, various sorts of cancer and many other diseases (Winkleby et al., 1992; Roche et al., 1996; USDHHS, 2014).

Tobacco smoking is one of the biggest preventable causes of death in many countries (Koning, Webbink & Martin, 2010). Smoking is therefore now seen as a problem that should be solved. As a result, one seeks solutions to counteract this problem. Examples include the labels on the packaging and bans on smoking in public.

## Research question

In this study I want to further interpret the relationship between education and smoking. There seems to be causal evidence, but the explanation is lacking (Cutler and Lleras-Muney, 2006). Without a solid explanation its causality will always be questioned. Literature seems to show a statistical relationship really well, but it is lacking an answer to the question “how can it be that education causes improved smoking behavior?”

Vida Maralani (2014) explains that there are 6 links between smoking and education. They are information, social networks, future expectations, analytical and non-cognitive skills, and a joint determination between education and smoking. She concluded that no analytical or cognitive skill in the dataset could predict smoking in adulthood. She did not measure non-cognitive skills.

Maralani (2014) is an important source for this research. In the discussion she mentions that the origin of the family and non-cognitive skills could be more important explanations than what has been investigated so far. Investigating this is among the most important added values of this research. Non- cognitive skills are being described by Maralani (2014) as soft skills, such as time management skills, self-efficacy, organization skills, and the ability of working together.

This research explains the background and thus indirectly improves the existing evidence of causality through interpretation. I argue that one way the relationship works is through its direct policy goal of increasing skills. Schooling teaches students better skills, think of thought processing skills, which is used to quit smoking. It could possibly explain why highly educated exhibit healthier smoking behavior than less educated. Furthermore, potential policy measures can be implemented in a better way.

There is an issue. We do not know whether, and how, investments in education lead to improvements in smoking behavior. Gaining health knowledge does not seem a logical explanation. An alternative theory is that education contributes to increased skills that reduce bad smoking behavior. I want to test this hypothesis.

There are several ways how skills, related to education, could lead to improved smoking behavior. Smart people could have a better education and use their IQ to handle smoking behavior better. In that case education doesn't cause better smoking behavior. In this case IQ leads to both education and better smoking behavior. Alternatively, education could improve problem solving skills, which in turn leads to less smoking. In that case improving their problem-solving capacity is an interpretation

of the relationship between education and smoking. This means an investment in education causes a health effect.

For reasons related to bias, which will be mentioned later, I would like to focus on the variable of quitting smoking, rather than starting smoking. On top of that, a lot of people start smoking at young ages, thus researching the effect of education on starting smoking might be more of a topic on its own (Maralani, 2013). The dataset I use only includes people of age 16 and older, meaning the initiation questions are more difficult to research.

The research question I want to answer is the following:

**To what extent does an increase in skills explain the relationship between education and quitting smoking?**

## Outline

The next chapter is the theoretical framework. Chapter 3 describes the data, variables and methods. Chapter 4 shows the results of the regressions. The conclusions are in chapter five. Chapter 6 is the discussion section. The next section is the list of used literature. This is followed by the appendix.

I try to limit the skills in this study to those that can interpret the relationship, if possible. I also limit it to quitting smoking. Another option is to research starting smoking, but this causes an issue because most people start smoking at a young age.

## Chapter 2: Theoretical Framework

### Linking education and smoking

Research has been done to reduce smoking behavior. Some of these highlighted the relationship between education and smoking. Following a longer or higher level education seems to lead to both a smaller likelihood to start smoking and a greater likelihood of quitting. Cutler and Lleras-Muney (2010) found that each extra year of education reduces the chance of smoking. Sander (1995) found that education had a positive effect on the chances of quitting smoking.

Many studies have found a link between smoking and education (Cutler and Lleras-Muney, 2006). The differences between education groups were small, but have become bigger (Pamuk et al, 1998). Over time the user group of cigarettes has changed. Before the 1950s, smoking was high among all educational groups (Maralani, 2013). At the same time smoking has been on the decline. The educational differences are for a large part related to birth cohort membership (Vedøy, 2014). Although the prevalence of the smoking status within the cohorts has been decreasing over time, it has been the rise of the amount of never smoking that has been explaining a large part of the decline (Maralani, 2013). The rise and fall of smoking seems to be consistent with the theory of diffusion of innovations (Rogers, 2010). This theory basically predicts that different groups pick up a product at different times. The innovators and early adopters were the highly educated, who also dropped the product earlier.

Several studies suggested that there is a causal link in the relationship between education and smoking. Both de Walque (2007) as Grimard & Parent (2007) use used the Vietnam War draft avoidance as a method to determine causality. Attending college allowed people to not go to war. Those who had followed education for this reason less likely to smoke. They also were more likely to quit smoking.

The central question we must ask is why, and if, education leads to better smoking behavior. Until we know how it's difficult to determine which policy measures can be effective (Cutler and Lleras-Muney, 2006).

Vida Maralani (2014) has set up her own framework mentioning 6 sources of possible links between education and smoking. Two of them are information as described earlier and social networks, which

were described earlier. She found that having a parent that smokes or a friend that smokes significantly increases the likelihood of smoking.

Another option she mentions is future expectations (Maralani, 2014). Thinking forward and following an education are related. Not only that, but education may also teach one to think forward. It is likely that they cause each other. The results do not support this claim however.

She also mentions skills as a possible link (Maralani, 2014). She makes a difference between analytical skills, like numeracy, and non-cognitive skills, like time management skills. She found that effects of courses algebra and earth science could not explain the relationship. She concluded that no analytical or cognitive skill in the dataset is a significant predictor of smoking in adulthood when making the comparison among similarly educated adults. Non-cognitive skills weren't measured however.

Yet another explanation could be that there is a joint determination of education and smoking (Maralani, 2014). Both factors may be caused by the same factor. This could be early experiences related to school or personality. She finds that the likelihood of initiation of regular smoking is largest between ages 11 and 19. The differences between the educational groups in smoking initiation probabilities are also largest in this age group. Also, the variables smoking status of age 16 and having a parent that did not complete high school are both predictors of both education and smoking in adulthood. The most important link between education and smoking seems to be a cause in the early life, before the extra years of education have started. Most other research incorrectly treat the smoking decision as if it was a decision that is made in adulthood, and end up potentially drawing the wrong conclusions. This puts the suggested causal relationship between education and starting smoking into question.

## Explanations

Social contacts were considered as an explanation. Not only does education relate to your own smoking behavior, but also to the smoking behavior of others. The parents of children who attended more years of education are more likely to quit smoking (Field, 2005). Social contacts who smoke appear to cause others to smoke more, but the effect is only partly causal (Sacerdote, 2000). Also having a highly educated spouse seems to be related to the reduction of smoking behavior, but this effect could be caused by finding the right spouse (Monden, 2003).



Income was also considered as an explanation. Having a better education is related to a higher salary. Although this could explain other health effects, this does not explain why low educated people smoke because smoking is expensive (Cutler and Lleras-Muney, 2006). It is possible that highly educated people are more sensitive to the price of cigarettes, but this is unlikely (Cutler and Lleras-Muney, 2010). Other studies (Lahelma et al, 2004; Cutler and Lleras-Muney, 2006) corrected for differences in jobs, but this doesn't make the effect of education on smoking insignificant.

Also, information and cognitive skills have been considered. According to Grossman (1972) education allows people to produce more health. Educated people would better understand the relationship between health behavior and health (Sander, 1995). The informational link can be related to school activities. Schools can provide classes with content related to health (Maralani, 2014). Also school policies can teach children what is bad for them by punishing bad smoking behavior (Maralani, 2014).

Several researchers suggest that the differences in smoking behavior are caused by better information and knowledge about smoking (Kenkel, 1991, Taubman & Rosen, 1980). Some knowledge is only learned at school (Cutler and Lleras-Muney, 2010). Education gives you access to more information and it improves critical information processing (Cutler and Lleras-Muney, 2006). This may also explain why the highly educated were the first group after the release of the 1964 Surgeon General's Report to stop smoking (de Walque, 2004). Differences in information however, explain little (Meara, 2001). Many of the effects of smoking are well known, even among low-skilled. People who know more about the effects of smoking are less likely to smoke (Cutler and Lleras-Muney, 2010). This effect is limited however (Cutler and Lleras-Muney, 2010), and possibly endogenous (Kenkel, 1991).

Cutler and Lleras-Muney (2010) investigated the effects of personal characteristics as discounting, value of the future, and risk aversion in the relationship between education on health behaviors, but found only small effects.

## Criticism

There is more criticism. As previously mentioned, the effects of income, social contacts and information are limited. They can't explain the relationship between education and smoking in its entirety.

Another problem is that education could be correlated with non-observable variables that are related to health (Sander, 1995), and in this case to smoking behavior. Chaloupka (1990) argued that schooling is related to time preferences. People with high time preferences invest less in both education and in health. This means that there isn't necessarily a causal relationship between health and education. The relationship may be the result of self-selection (Sander, 1995). The effect of time preferences later proved to be small however (Cutler and Lleras-Muney, 2010).

The effect of health knowledge is being doubt. It might be endogenous (Kenkel, 1991). Viscusi (1992) states, that both smokers and non-smokers overestimate the effects of smoking.

In addition, it is possible that education is a result of health (Grossman, 1976). If smoking is correlated with education it's still not sure whether education has an effect on smoking, or smoking has an effect on education (Sander, 1995). This is an endogeneity issue. Instrumental variables are sometimes used to solve this. In the article of Sander (1995) the education of the parents was used as an IV, replacing education. Another issue of IVs is that smoking decisions have a longitudinal character, and most IVs used do not take this into account (Koning, Webbink & Martin, 2010). Also, IVs are always based on belief that they are valid. Using an instrumental requires its assumptions to be true. I'll give an example. An IV, parental education, is assumed to affect smoking only through education. If it affects smoking directly it's incorrect to use this IV, because you cannot prove that your own education has an effect on smoking.

Another problem is the fact that many people begin smoking at an early age (Koning, Webbink & Martin, 2010). If people have started smoking before their additional years of study, the students will have the same amount of education, thus this cannot explain decision to start smoking at young ages. In addition, Koning, Webbink & Martin (2010) argue the variable "years of education" is biased. Young people may start smoking due to group behavior; this effect is likely strong at young ages when the amount of years of education completed is low. This could lead to overestimations of the effect of education on smoking hazards.

The relationship between education and smoking is a complex problem. There are many correlations. The relationship is hard to explain. However, it is undisputed that there is a correlation between having a high level of education and smoking behavior.

## Quitting smoking and education

The correlation between education and smoking has been researched in various ways. Most common seems research on the smoking hazard, like the work of de Walque (2007). Other papers included information on starting (Maralani, 2014) or quitting smoking (Sander, 1995).

William Sander (1995) found that schooling has a significant positive effect on quitting smoking. To get around endogeneity issues he used parental smoking as an instrumental variable. Koning, Webbink & Martin (2010) used a Mixed Proportional Hazard specification to get around endogeneity issues. They find that an additional year of education leads to a reduction of the duration of smoking by 9 months.

A positive association between education and quitting hasn't always been found, despite the amount of research done. For example, Eisinger (1966) found no significant association between education and changing smoking behavior. In 1982 no association was found between education and smoking cessation among men in intervention programs (Ockene et al., 1982). Most modern research does find this association however (Koning, Webbink & Martin, 2010).

Quitting smoking can be related to the assistance a smoker can get. Getting high amounts of support from a partner is associated with both short term abstinence and long term quitting, as is the perceived availability of general support (Mermelstein et al., 1986). On the other hand, having smokers in the social circle has a negative effect on abstinence from smoking. Their presence also significantly differs between the relapsing group and the long term abstainers (Mermelstein et al., 1986).

Not all socioeconomic groups have an equal chance to quit smoking (Wetter et al., 2004). Gender used to be the leading sociodemographic factor associated with smoking. Later on this became education. It is now the biggest factor in both the likelihood to be smoking and the likelihood to quit smoking. An American longitudinal study showed that people with less than a High School education had 0.13 times the odds of quitting smoking compared to those with a college degree (Wetter et al., 2004). The writers consider the most important predictors of quitting smoking to be characteristics related to work, tobacco related variables, demographics and household variables, which they correct for in a multivariate analysis. Only 2 of these were significantly associated with both education and smoking; cigarettes per day, and "pros of smoking", derived from a Decisional Balance form (Wetter et al., 2004).

Kabat & Wynder (1987) attempted to find the determinants of smoking cessation in order to predict what type of person would be the most likely to break their smoking habits. They criticized earlier research for mostly focusing on men, lacking a degree in variation in their sample group with regards to race, socioeconomic status and race, and their populations having a high risk of heart disease. Due to the timing issues, these studies often cannot provide information on long-term smoking cessation. Kabat and Wynder (1987) found that the lifetime quit rate was higher for men than for women and increased with age. Being white and being a Jew had had a significant positive effect on the quit rates. Quit rates also increase with level of education, ranking from Grammar School to Graduate School. Taking the first cigarette of the day long after waking up had a significant association with the quit rate. Arguably the most questionable variable was “occupational level”, ranking from Professional to Housewife, which had a significant positive effect (Kabat & Wynder, 1987).

Women have more difficulty quitting smoking than men do (US Public Health Service, 1980; Blake et al., 1989; Bjornson et al., 1995). Between 1965 and 1980 the amount of men who smoke dropped from 52.1% to 37.9%, while the proportion of women smoking dropped from 34.2 to 29.8% in that same period (Blake et al., 1989). The higher quit rates of men have been reported from the 50s onward (Haensel et al., 1956). The rates of quitting is similar however when the duration of smoking is controlled for (Blake et al., 1989). Each age cohort seems to show different rates of smoking hazard (Maralani, 2013) and quit ratio (Harris, 1983), but controlling for the duration of smoking and the amount of cigarettes smoked is often lacking in these types of research (Blake et al., 1989). Blake and his colleagues (1989) speculate that this is caused by the self-image associated with smoking is more important for woman than for men. They state that for the gender differences in quitting may be caused by women not being persistent, having different motivations to quit, and social influences between genders being different. Furthermore they suggest that “woman may be less confident about their ability to quit smoking than men” (Blake et al., 1989, p 410). They find that only 9% of women attempt to quit at least 3 times for at least a week, compared to 11% of men. A possible explanation could be that woman that attempt to quit think this is caused by the ability to quit rather than external qualities such as trying hard enough (Blake et al., 1989).

The Lung Health Study Research Group (1995) also found gender differences in the quit rates. Males seem to quit at a significantly higher rate. There were no significant differences in 12 month quit rates, but there were significant differences in 36 month quit rates, *ceteris paribus* (odds ratio of 1.24). Among significant factors were education, age and being married, which were all positively associated with quitting smoking. The researchers conclude that smoking cessation programs need to deal the coping skills of women and the unmarried and that skills training is required for gender specific cessation programs (Bjornson et al., 1995).

Smoking decisions have a longitudinal character (Koning, Webbink & Martin). This also applies to quitting smoking (Kabat & Wynder, 1987). In an American survey in 1986, 54% of smokers who had quit for a day the year before the survey had relapsed by the time of the survey (Hatziandreu et al., 1990). Among those who quit for at least 7 days this was 28.3%. The research concluded that being younger and having a higher education were the biggest predictors of attempting quitting for at least 7 days. Older age and being white were predictors of successfully abstaining for at least three months. A possible explanation for this could be the need to quit as a result of health problems. Correcting for other factors, the researchers found that although black and white had equal likelihoods of quitting temporarily, but were less likely to remain abstinent (Hatziandreu et al., 1990). Other researchers also found racial differences in quitting behavior (Novotny et al., 1988). Hatziandreu and his colleagues (1990) discussed that many people want to quit, but either lack motivation or the required skills to maintain a behavioral change. The research also reported that “years of education” is associated with the likelihood to attempt quitting. There was however no difference in the likelihood of quitting between genders. The differences in smoking prevalence between genders are mostly the result of different uptake rates (Hatziandreu et al., 1990).

Fiore and his colleagues (1990) used the same 1986 survey to analyze differences in quitting methods. Around 90 percent of people who successfully quit and 80 percent of unsuccessful quitters used individual methods to stop smoking, most commonly the "cold turkey" approach (quitting in one go). Groups who are more likely to use a cessation program are the more educated, woman, middle-aged people, people who attempted to quit multiple times before, and heavy smokers. Those with at least 13 years of education had 1.21 times the odds of using assisted methods of cessation, compared to those with up to 12 years of education (Fiore et al., 1990).

## The role of skills

It can be difficult to quit smoking as smoking is addicting. This is caused by a drug, namely nicotine (USDHHS, 2014). Certain skills are needed to overcome this addiction.

Education can teach students many skills. Maralani (2014) explains makes the distinction between analytical skills and “soft skills”. With analytical skills she means that it makes a student smarter. Examples of analytical skills are literacy, reasoning, numeracy, reading critically and writing persuasively. They can be taught at school, but may also be useful with regards to smoking decisions. She tested differences between 9<sup>th</sup> grade students taking algebra 1 and those with lower

math then algebra 1, but found no significant difference in the likelihood of smoking. She tested the same for 9<sup>th</sup> grade science, comparing to earth science, but found no significant results.

Another type of skills is soft skills, which is what is being focused on in this thesis. The definition of soft skills can depend on the context (Schulz, 2008). Maralani (2004) defines soft skills as non-cognitive skills like the skills to function in a group, time management skills and self-efficacy. Schulz (2008) explains that hard skills are technical requirements for a job, while soft skills complement them. It is therefore difficult to define them (Schulz, 2008). Also the soft skills could also be related to quitting smoking (Maralani, 2014).

In the past, employers often complained that students lack soft skills. A student could enter the business world with all the knowledge needed to fulfill the job, but lacked certain skills such as communication skills (Schulz, 2008). Nowadays students can get these skills. Group assignments and participating in student government are examples of how to improve social skills through education (Maralani, 2014). Soft skills can be acquired at educational institutions, but also be trained in courses or by socializing (Schulz, 2008).

There are several types of soft skills. The most important from a job perspective are communication skills. These can involve language skills, but also rhetoric skills. Having these skills at a decent level will be reflected in body language, discussion skills, and presentation. Critical thinking and structured thinking are also important, as they are related with the ability to solve problems. These skills are needed to properly make informed decisions. A final important skill is creativity, as it can result in thinking outside the box (Schulz, 2008).

Skills are not only related to education, but may also be related to smoking behavior (Maralani, 2004).

Personal identity and self-efficacy are important elements in changing smoking behavior (Crone & Van Kesteren, 2007). There are large differences between socioeconomic groups in their probabilities to quit smoking (Wetter et al., 2004). Socioeconomic characteristics can determine how a person identifies himself and how a person should behave. People can reevaluate their own identity. People who identify themselves as a quitter have a larger likelihood to quit smoking, and this regaining a new identity is a form of motivation. The capability to reevaluate behavior, accept a new identity, in combination with self-monitoring, can lead to a higher probability of quitting smoking (Crone & Van Kesteren, 2007).

Self-efficacy is an important skill relating to behavior, as is stated by the Social Cognition Theory of Bandura (1977). Any form of psychological behavior changes the amount and intensity of self-

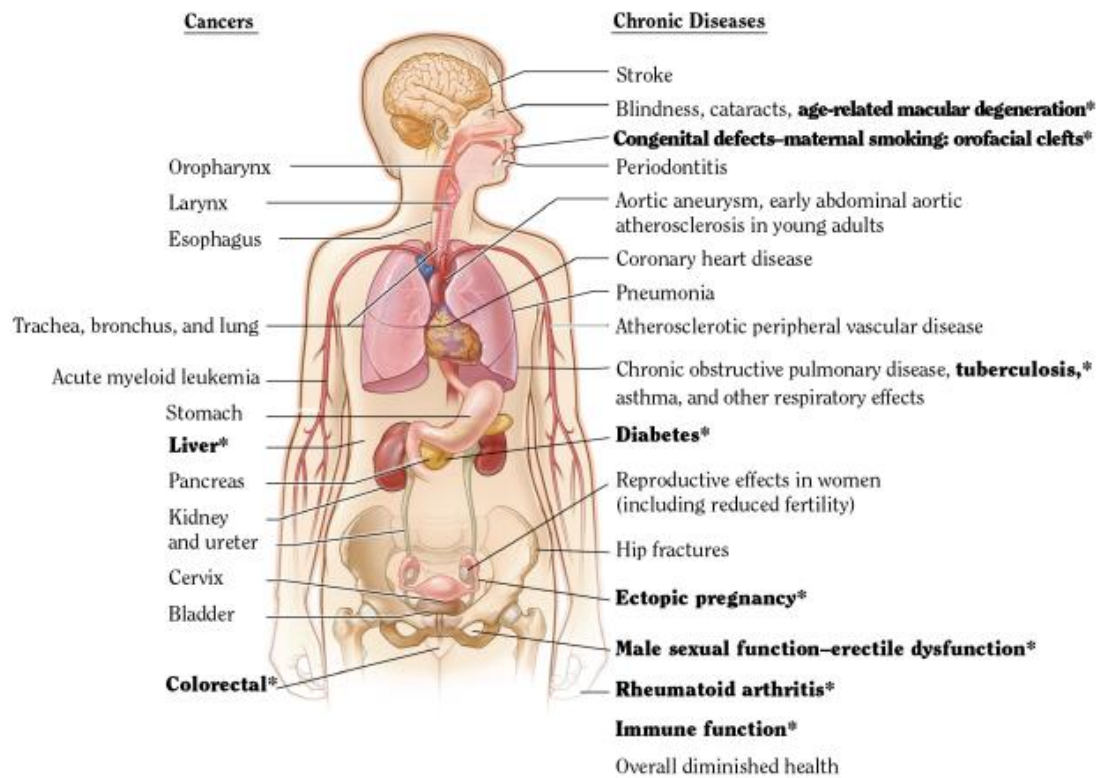
efficacy. An individual's expectations determine which behavior will be initiated, the amount of effort used, and how long this effort lasts.

Self-efficacy expectations are important for overcoming barriers (Crone & Van Kesteren, 2007). It is also one of the more important aspects of quitting smoking, as explained by Dijkstra & de Vries (2000). There are several types of self-efficacy that were measured, namely social self-efficacy, emotional self-efficacy, relapse-self efficacy, skill self-efficacy, and try self-efficacy. All of these were self-assessed. Of these only the variable relapse self-efficacy is associated with the number of past quit attempts, and only skill self-efficacy is a significant predictor of quitting smoking between two points in time. The point prevalence of quitting smoking had skill self-efficacy and relapse self-efficacy as predictors.

However, self-efficacy is not only related to smoking behavior, but also to education. There are large differences between the high and low educated. The low educated experience more social pressure to quit smoking, but also don't believe that they are capable to quit smoking, compared to highly educated. The latter is often defined as self-efficacy expectation (Willemsen, 2006).

## Smoking and Health

There has been evidence linking smoking and cancer since the 1950s (Sasco et al., 2004). Over time studies have shown causal links between smoking and cancer, but also many other illnesses (USDHHS, 2014). See the picture below from the 2014s Surgeon General’s report.



Source: USDHHS 2004, 2006, 2012.

Note: Each condition presented in bold text and followed by an asterisk (\*) is a new disease that has been causally linked to smoking

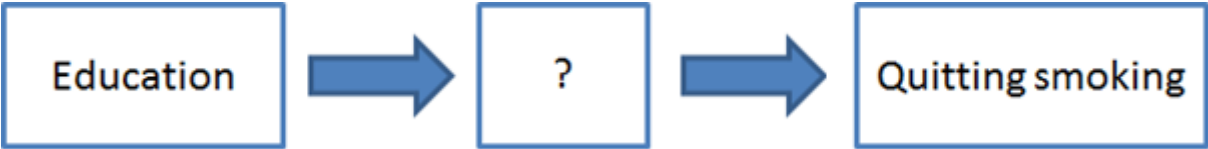
Source: USDHHS, 2014

The link between smoking and health is final part of the framework. It concludes the link between education, skills, smoking, and health.

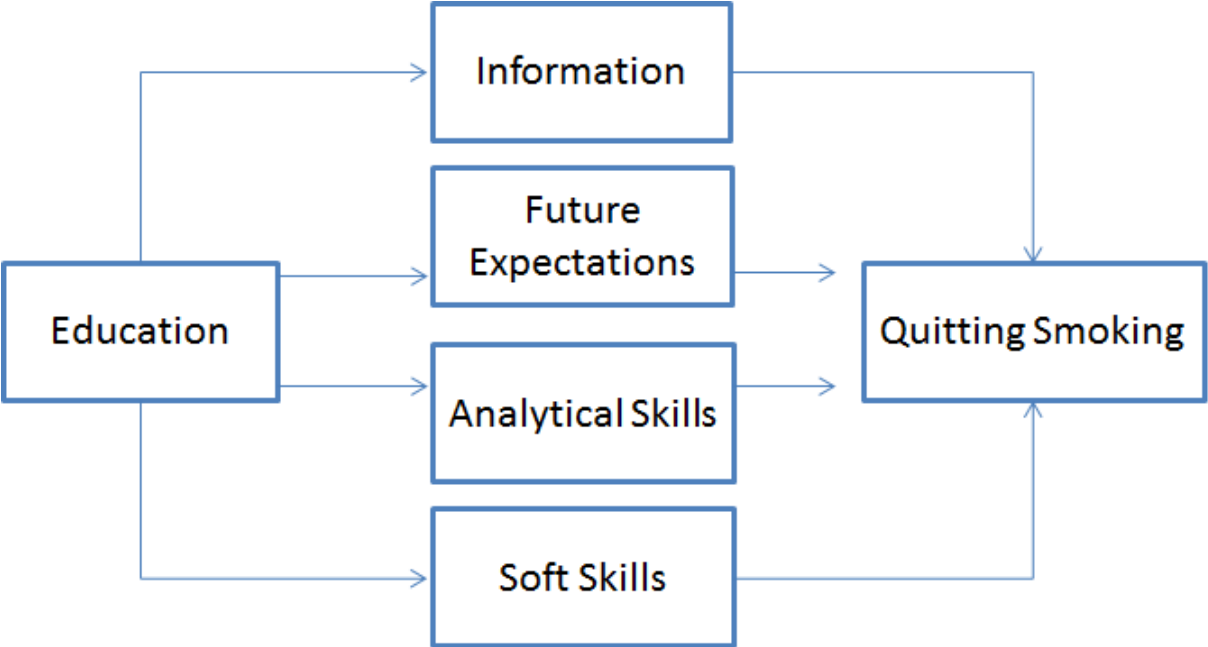


# Hypothesis development

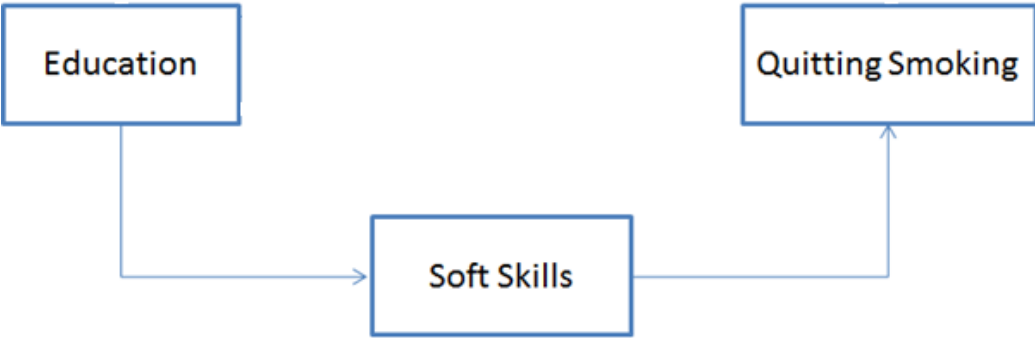
The literature concludes that education and quitting smoking seem causally connected, and so do smoking and health. In this thesis I focus on the relationship between education and quitting smoking and what could possibly connect the two.



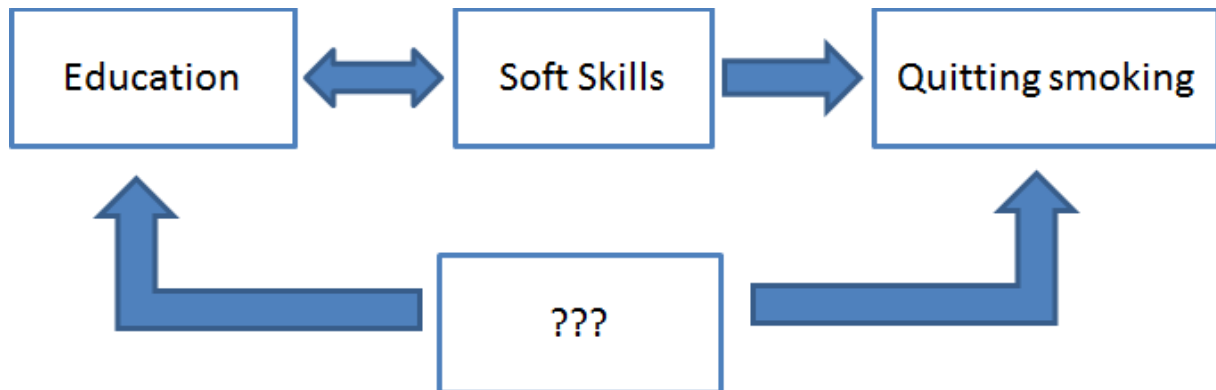
It is necessary to interpret the missing variable. Maralani (2014) explains 4 possible interpretations for this.



The hypothesis answering the research question is centered on the soft skills, which is the main focus of this thesis.



This must be done with reverse causality and joint determination of the variables in mind, as explained by Maralani (2014).



## Hypotheses

At the basis of this research is the relationship between education and quitting smoking, thus will confirm that education is causing a person to quit smoking. De Walque (2007) showed that this was true in his research, and similar effects are expected for this research. As shown in the research by Koning, Webbink & Martin (2010) there may be different outcomes for men and women. This means the relationship must be tested per gender. Most literature seems to find effects for both however, which I expect for this research to be the case as well.

H1a: a higher education has a positive effect on the likelihood of individuals to quit smoking

H1b: a higher education has a positive effect on the likelihood of men to quit smoking

H1c: a higher education has a positive effect on the likelihood of women to quit smoking

The relationship has to be explained, as it is unclear how education could cause people to quit smoking. The theory of this thesis is that education teaches student “soft” skills, and that these skills are used to quit smoking. Breaking up the relationship means first showing that education causes students to gain skills. First of all, people learn something at school according to the theory of Maralani (2014). Secondly, this is also shown in literature. For example, the research of Willemsen (2006) shows that education and self-efficacy are related variables. All skills will be tested separately, but will for now be summarized as “skills”.

H2a: a higher education has a positive effect on skills of individuals

H2b: a higher education has a positive effect on skills of men

H2c: a higher education has a positive effect on skills of women

The next part is to figure out if skills have a positive effect on the likelihood of quitting smoking.

Maralani (2014) suggested this link in her framework. More importantly, Dijkstra & de Vries (2000) have shown this relationship to be true with regards to the skill “self-efficacy”.

As shown in the literature women have more difficulty to quit smoking, thus should be tested separately (Bjornson et al., 1995).

H3a: a high skill value has a positive effect on the likelihood of individuals to quit smoking.

H3b: a high skill value has a positive effect on the likelihood of men to quit smoking.

H3c: a high skill value has a positive effect on the likelihood of women to quit smoking.

The final part is to combine everything into one model, answering the research question, combining everything. Education leads to quitting smoking through skills. This hasn't been shown in literature directly, but was part of the theory of Maralani (2014).

H4a: education has a positive effect on the likelihood of individuals to quit smoking through skills.

H4b: education has a positive effect on the likelihood of men to quit smoking through skills.

H4c: education has a positive effect on the likelihood of women to quit smoking through skills.

## Chapter 3: Data & Methods

### Data

In this study, I use the LISS data set<sup>1</sup>. This is a panel data set that contains information from randomly selected addresses in the Netherlands using the population register and claims to be a representative sample of the Dutch population. It contains over 5000 households and 8000 individuals. The panel members participate in a survey on the internet every month, but most variables in this study have only been measured once per year. LISS has a function that allows users to combine panel data in order to create new datasets.

In this study data from the health, personality and education will be used. Only people aged 16 and higher are surveyed. In 2014 the minimum age at which people are allowed to smoke went up from 16 to 18. This policy measure should show its results in the dataset, if working as intended.

The survey consists of 7 waves. Most variables were measured 7 times. People were asked their background variables once, but there are a few exceptions. Data for skills were available up to the 6<sup>th</sup> wave, and people who filled in the dataset in wave 2, weren't asked for them in wave 3. Waves 4 & 5 also interact in this way.

I will split the dataset in 2, and remove people under the age of 25 from my dataset for most of the research questions. Sander (1995) mentions that they are often missing a completed education, biasing the results. On the other hand I use the dataset with people below age 25 to check if skills increase when people are still in school.

---

<sup>1</sup> By its term & conditions on their website, the following has to be mentioned; in this thesis I make use of the data of the LISS (Longitudinal Internet Studies for the Social sciences) panel administered by CentERdata (Tilburg University, The Netherlands).

## Variables

Quickly scanning the dataset allows me to highlight variables that could be useful. The LISS dataset has information on smoking and education. William Sander (1995) used the years of for his education variable. I will be using the highest completed education at the moment the surveyed fill in the questionnaire for the first time. Due to the 25 year cutoff, people rarely change education from this point. The education variable will use dummies for primary school, vmbo, havo & vwo (combined) mbo, hbo and university, which are Dutch educational levels. There is also a dummy called schoolother, for people who changed education levels despite being over 25.

Primary school is the most basic form of education in the Netherlands. Completing primary school at the age of 12 allows a pupil to go to vmbo, havo or vwo (ranked from low to high). Mbo is senior secondary vocational education. Hbo and university are the highest forms of education.

The interviewees were also asked if they are currently pupils or students. This variable hasn't been consistent across all 7 waves however, as in the first 3 waves this explicitly mentioned interns, but in wave 4 and higher the survey only asked for pupils and students. As there are no people interviewed under the age of 16, the size of this group is limited.

Also available are the variables income, and the self-assessed "I have confidence in my capabilities", rated on a 1-7 scale. These variables should measure skills to a degree. It should be a measure of skills.

Income is generally corrected for, but theoretically speaking income should be related to skills. Income captures skill insofar this is valued by the market, and its inclusion is worth investigating. A possible interpretation would be that education leads to better thought processing or discipline, which leads to both income and quitting smoking. Income could have an effect on all variables. It could be used to purchase more years of education, or to buy more cigarettes. The relationship was also part of the model of Grossman (1972). On the other hand, education may increase income, which in turn could lead to less smoking (Vedøy, 2014). Even when this relationship is not caused by skills, finding this relation it is still relevant for policy reasons directly related to this topic, thus researching this is valuable.

The survey also asked people if they had ever smoked, and are smoking now. I define quitting smoking as not smoking now, but having done so in the past. People who never smoked weren't

asked if they have quit. Maralani (2013) stated that this was important. This means that the sample of people only contains former and current smokers when this variable is used.

The following variables were used to represent skills;

Skill	Range
Get chores done right away	1 to 5
Have excellent ideas	1 to 5
Like order	1 to 5
Shirk my duties	1 to 5
Follow a schedule	1 to 5
Spend time reflecting on things	1 to 5
Am exacting in my work	1 to 7
I feel that I'm a person of worth	1 to 7
I feel that I have a number of good qualities, at least on equal plane with others	1 to 7
I have confidence in my capabilities	1 to 7
I like to have the responsibility of handling a situation that requires a lot of thinking	1 to 7

Table 1

The questions being asked regarding the skills are the same across all waves.

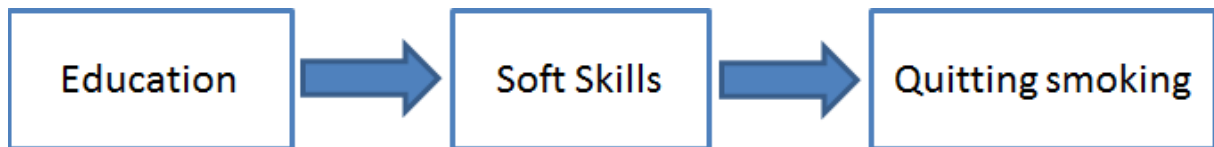
Other variables worth noting are urban, age and ethnic group, which are generally being corrected for. Also age squared is included to check for a nonlinear relationship between age and quitting smoking, as people of old age may value their health differently from the young.

## Methods

This is the research question;

**To what extent does an increase in skills explain the relationship between education and quitting smoking?**

To answer the research question, I must use several hypotheses. Combining these hypotheses will be sufficient to answer the research question. Different methods and techniques will be used for each part.



Above schedule is the simplified form of the main hypothesis. I need to divide this in separate hypotheses and test each part independently.

The first hypothesis is the following; education is positively associated with quitting smoking. The goal is to investigate whether my own data show a relationship between education and quitting smoking to confirm this.

A regular OLS regression would be able to show outputs below 0% or above 100%, and the relationship is unlikely to be linear. Therefore I will use a logit model, as I am trying to find likelihood to quit smoking. The model contains the logit of quitting smoking, explained by a constant, the beta of several levels of education, corrected for age, urban and ethnic group, and an error term. This will be done using the statistical program Stata, as is the rest of this research. I will be using primary school as the reference level for education, and a logarithmic function of income. I will use the dataset with people of at least 25 years old. The observations in the panel are not independent, as individuals are being surveyed multiple times, causing serial correlation.

The function is the following;

$$\text{Logit } QS = \beta_0 + \beta_1 * \text{vmbo} + \beta_2 * \text{havovwo} + \beta_4 * \text{mbo} + \beta_5 * \text{hbo} + \beta_6 * \text{university} + \beta_7 * \text{schoolother} + \beta_8 * \text{gender} + \beta_9 * \text{age} + \beta_{10} * \text{LNinc} + \beta_{11} * \text{ethnic} + e$$

Only the 6<sup>th</sup> wave was used rather than the entire dataset. Using the entire panel would mean using the same information more than once, and most people don't change their education in the panel. Using the entire panel, even when clustering, could lead to incorrect significant results. Because various researchers suggested that using young people in the dataset would cause bias these were excluded, but that also excluded a fixed effects regression as an option as education rarely changes.

The average marginal effects are then calculated. A joint significance (Wald) test was performed on both the various education levels and on the combination of age and age squared.

However, the coefficients of a logit's average marginal effects may be biased in case of heteroscedasticity. To check the validity of the model, a normal OLS regression is tested, tested on heteroscedasticity and on a Ramsey reset test. If this is heteroskedastic, heteroscedasticity is assumed for the logistic regression as well. An OLS regression and a logit a being tested with robust errors, both tested on joint significance of education. Only if the coefficients are similar the logit will

be interpreted. Otherwise the robust OLS is the best model. This is because the coefficients in a logit model can be biased in the presence of heteroscedasticity, but heteroscedasticity cannot cause the coefficients of a (robust) OLS to become biased. As it is impossible to test for heteroscedasticity for a logit model, the heteroscedasticity is tested for an OLS and assumed to apply to the logit as well. The very same thing is then done for an only woman group and an only men group.

Before doing the second hypothesis the Spearman correlation between all the variables will be checked. The reason for this is to prevent multicollinearity later in the research, and decisions to drop a variable must be made in case there will be serial correlation. The skills variable will be introduced in regressions, under the assumption the surveyed self-assess these skills well. For the readability of the paper, normal OLS regressions are preferred over ordered regressions. This does however add the assumption the difference between skill levels 1 and 2 is the same and the difference between 4 and 5.

The second hypothesis is; a high skill value has a positive effect on the likelihood of individuals to quit smoking. The idea is to investigate the first part of a potential causal relationship, where education leads to skills, and said skills lead to quitting smoking. Different regressions are used per skill, and separate regressions per gender. A linear regression is used for most skill variables. Income uses a logarithm however. In contrast with the first hypothesis, the dependent variable changes over time. Therefore the entire panel is used, rather than just the sixth wave. The personal identifiers were clustered, which makes them robust. It is known that the amount of smokers isn't constant over time, and policy measures can change, thus a dummy is used for each wave to correct for wave specific effects. Afterward the joint significance of both education and age is tested.

It must be noted that education could lead to skills, but skills could also lead to education. Due to limitations in my dataset there is no IV to correct for this. To add evidence that the relationship above is truly causal, another effect is measured. An assumption of this research is that people learn their skills while being at school. This will be tested. A proxy for education, namely the interaction effect between age and being in school is being used to see its effect on skills. If students learned their skills at school, these skill variables should get higher over time, especially at the ages that they are still in school. The entire dataset is used this time, rather than using the usual group of people above 25, and all 7 waves are included. For the sake of consistency both genders are also tested separately. An OLS regression is made to test if the proxy has a significant effect on skills, while correcting for ethnicity and education level, and age. This will be tested for heteroscedasticity and tested again using the cluster command, which also makes the errors robust.



It could be possible that skills simply increase with age, but aren't necessarily learned at school. The latter isn't a causal effect of education leading to skills, thus doesn't belong in this research.

This is followed by answering the question; are skills related to quitting smoking? The idea is to investigate the second part of a potential causal relationship, where skills lead to quitting smoking. This time a fixed effects panel regression model is used, attempting to explain the likelihood to quit smoking. A logit would be preferred but cannot be used if heteroscedasticity could be in the model for the reason that the coefficients can be biased. This method has been chosen over an OLS as it is important to hold personal (genetic) qualities constant whenever possible. For example, a person's IQ can be the driver of education as well as quitting smoking. The added value of the fixed effects panel model is that such variables, which aren't the result of education, no longer cause issues. Different regressions per gender are used, but it is again only tested among those who are 25 years and older. Redundant skills may be left out of the equation if needed.

The thought process behind this part of the thesis is that skills change over time, while individual characteristics stay the same, which can be used to measure the effect on smoking behavior.

Finally I will answer the research question; to what extent does education cause quitting smoking through skills? A logit regression is used, and another one per gender. The variable to explain is quitting smoking, and the predictors are the skills that have been significant in all previous hypothesis, except in hypothesis 1 and 4. If none of the skills has a proven causal effect then the association is measured using all skills. The other variables are education dummies, (log) income and ethnic. The same regression is made without the skills. If two different regressions are made, one with skills, and one without skills, the coefficients can be different. The effect of the skills is the difference between the coefficients of the educational dummies in the regressions.

Also tested in this section is the hypothesis that income leads to quitting smoking. As mentioned before the income variable may contain information regarding skills and will be tested the same way as other skills as much as possible. The effect cannot be measured when using the panel data however as income has only been measured once. In this section it is however possible to give an indication for the effect, but the method doesn't as good as ones used in hypothesis 3.

In short, the main question suggests a causal relationship where education (A) leads to skills (B) leads to smoking (C). The sub-questions simply tell us what B is, and tell the relationship between each part of the equation.

## Chapter 4: Results

### Education and Quitting Smoking

At first the effect of education on smoking is being tested. As explained in the variables section, the sample group only contains current and former smokers. Three different hypotheses will be tested. The first test will involve all people of 25 years and older. The second involves only men, and the third involves only women.

Before being able to say anything about the logistic regressions, an OLS was made yielding the following results;

Quit	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]	
<b>vmbo</b>	.0713635	.0311259	2.29	0.022	.0103337	.1323933
<b>havovwo</b>	.1198707	.038701	3.10	0.002	.043988	.1957534
<b>mbo</b>	.1111587	.0324524	3.43	0.001	.047528	.1747894
<b>hbo</b>	.1611224	.0327585	4.92	0.000	.0968914	.2253534
<b>university</b>	.1880902	.0415797	4.52	0.000	.1065631	.2696173
<b>schoolother</b>	.6536254	.4497609	1.45	0.146	-.2282392	1.53549
<b>gender</b>	-.040708	.0167264	-2.43	0.015	-.0735048	-.0079127
<b>age</b>	.0016966	.0039461	0.43	0.667	-.0060407	.0094338
<b>age2</b>	.0000618	.000035	1.76	0.078	-6.87e-06	.0001305
<b>LNinc</b>	.0071557	.0022	3.25	0.001	.0028419	.0114694
<b>foreign</b>	.0210975	.0258075	0.82	0.414	-.0295042	.0716993
<b>_cons</b>	.2554892	.1115319	2.29	0.022	.036804	.4741743

Table 2

However, the model tested positive on heteroscedasticity. This is in the appendix. The Ramsey RESET test was also significant however. This will be discussed later.

The regression was then made with robust standard errors.

Quit	Robust Coef.	Std. Err.	t	P>t
<b>vmbo</b>	.0713635	.0319409	2.23	0.026
<b>havovwo</b>	.1198707	.0394226	3.04	0.002
<b>mbo</b>	.1111587	.033477	3.32	0.001
<b>hbo</b>	.1611224	.0330513	4.87	0.000
<b>university</b>	.1880902	.0412264	4.56	0.000
<b>schoolother</b>	.6536254	.0435967	14.99	0.000
<b>gender</b>	-.0407087	.0168603	-2.41	0.016
<b>age</b>	.0016966	.0038086	0.45	0.656
<b>age2</b>	.0000618	.0000327	1.89	0.059
<b>LNinc</b>	.0071557	.0022108	3.24	0.001
<b>foreign</b>	.0210975	.0253043	0.83	0.404
<b>_cons</b>	.2554892	.1116784	2.29	0.022

Table 3

The Ramsey RESET test wasn't significant (see appendix), which means the test could not show misspecifications in the model, although it can still be there. All education dummies combined tested jointly significant (see appendix).

Vmbo has a significant positive effect on the likelihood of quitting smoking ( $P=0.026$ ). People having vmbo as their highest level of education have a 7.13 percentage points higher likelihood to quit smoking compared to people who have primary school as their highest level of education, *ceteris paribus*.

For this research vwo and havo are combined. The likelihood of people with vwo or havo as their highest completed education quitting smoking is 11.99 percentage points higher compared to people who have primary school as their highest completed education, *ceteris paribus*. Also this was significant when using a 5% significance interval ( $P=0.002$ ).

Similarly, mbo ( $P<0.001$ ), hbo ( $P<0.001$ ), and university ( $P<0.001$ ) have significant positive effects on the likelihood to quit smoking compared to those with primary school as education. Their likelihood to quit smoking is respectively 10.88, 15.98 and 18.62 percentage points higher, *ceteris paribus*.

Age and age squared seem to have a positive effect, but don't turn up as being significant at a 5% significance level. Using the Wald test shows that they are jointly significant (see appendix). It must also be noted that age squared is significant at a 10% significance level.

Having 1% more income increases the likelihood of quitting smoking by 0.72 percentage points. This effect is significant at 5%. More details of this effect are shown later in this thesis. Being a foreigner has no significant effect.

Most important for this research is the likelihood of men quitting smoking is 4.1 percentage points lower than the quit probabilities of being a woman, *ceteris paribus*. Differences between genders will play an important role in this thesis and more details of gender differences will be shown later.

A logistic regression was also done (see appendix) as well as a robust version (see appendix). Its marginal effects regression showed the following results;

Average marginal effects		Number of obs = 3065				
Model VCE : Robust						
	dy/dx	Std. Err.	z	P>z	95% Conf.	Interval
<b>vmbo</b>	.0733627	.0299208	2.45	0.014	.014719	.1320064
<b>havovwo</b>	.123793	.0386749	3.20	0.001	.0479917	.1995943
<b>mbo</b>	.1147778	.0313427	3.66	0.000	.0533472	.1762084
<b>hbo</b>	.1651749	.0321739	5.13	0.000	.1021151	.2282346
<b>university</b>	.1956251	.0434012	4.51	0.000	.1105602	.2806899
<b>schoolother</b>	0	(omitted)				
<b>gender</b>	-.039432	.0166463	-2.37	0.018	-.0720583	-.006805
<b>age</b>	-.0059614	.0042744	-1.39	0.163	-.014339	.0024163
<b>age2</b>	.0001337	.0000398	3.36	0.001	.0000557	.0002117
<b>LNinc</b>	.0073164	.0021877	3.34	0.001	.0030286	.0116042
<b>foreign</b>	.0238563	.0255972	0.93	0.351	-.0263134	.074026

Table 4

However, some of the results are too different from the robust OLS regression, meaning that heteroscedasticity in this model could be problematic. This is why the model earlier is interpreted instead of this one.

## Education and Quitting Smoking; the effect for men

The same procedure as before is being followed, the OLS regression tested significantly positive for heteroscedasticity (see appendix). A regular logit regression was made, as well as a logit with robust errors (see appendix). However, for the coefficients to be unbiased only a linear model with robust errors will be interpreted.

Linear regression		Number of obs	1568	
		F( 9, 1558)	21.64	
		Prob > F	0.0000	
		R-squared	0.0903	
		Root MSE	44251	
Quit	Coef.	Robust St. Err.	t	P>t
<b>vmbo</b>	.0697248	.0430238	1.62	0.105
<b>havovwo</b>	.0959365	.0549563	1.75	0.081
<b>mbo</b>	.1183018	.0435925	2.71	0.007
<b>hbo</b>	.1554335	.0433169	3.59	0.000
<b>university</b>	.153607	.0527167	2.91	0.004
<b>age</b>	-.0059477	.0052345	-1.14	0.256
<b>age2</b>	.0001334	.0000437	3.05	0.002
<b>LNinc</b>	.0089115	.0029839	2.99	0.003
<b>foreign</b>	.0024415	.0347777	0.07	0.944
<b>_cons</b>	.4011018	.1552033	2.58	0.010

Table 5

The results for men are clearly different from the mixed group. Men who have completed vmbo, have and vwo no longer have a significantly higher likelihood to quit smoking compared to people who only completed primary school. These schools start at the age of 12. Mbo, hbo and university do have a significant positive effect on the likelihood to quit smoking.

Yet again being foreign has no significant influence on the probabilities to quit smoking, and neither does age itself. Age squared does have a significant influence, and so does the joint effect of age and age squared (see appendix). Income matters yet again and its coefficient is higher than for the mixed group. Most importantly, the education dummies have a jointly significant effect on the likelihood to quit smoking. Also noteworthy is the low R square. The model doesn't explain that much regarding quitting smoking.

## Education and Quitting Smoking; the effect for women

Same procedure as before is used. The OLS model heteroskedastic (see appendix). The logit model were tested but due to heteroscedasticity not preferred. Another linear model with robust errors will be interpreted, these were the results;

Quit	Coef.	Robust SE	t	P>t
<b>vmbo</b>	.0739278	.047821	1.55	0.122
<b>havovwo</b>	.1378509	.0575446	2.40	0.017
<b>mbo</b>	.0983669	.0522304	1.88	0.060
<b>hbo</b>	.1649434	.0511717	3.22	0.001
<b>university</b>	.2465165	.0665273	3.71	0.000
<b>schoolother</b>	.6954341	.0633011	10.99	0.000
<b>age</b>	.0118547	.0057709	2.05	0.040
<b>age2</b>	-.0000362	.0000514	-0.70	0.481
<b>LNinc</b>	.0058442	.0033162	1.76	0.078
<b>foreign</b>	.0393911	.0369341	1.07	0.286
<b>_cons</b>	.0108951	.1638631	0.07	0.947

Table 6

Age has a significant effect on the likelihood of quitting smoking. It is also jointly significant with age square as an effect on quitting smoking (see appendix). In the model of men only age squared was significant and had a positive effect. In both cases age has an effect of smoking cessation, but in case of women the relationship seems more linear.

Vmbo has no significant effect, but the other education levels do. Most importantly, the education levels have a jointly significant effect on quitting smoking. Also noteworthy is that the coefficients for the educational dummies for men were never higher than 0.16. The effect of education on the likelihood on quitting smoking is bigger for women than for men, the only exception being vmbo.

Income does not, in contrast with men, have a significant effect on quitting smoking. Also in contrast with the literature, being foreign has no significant effect on quitting smoking.

## Skills

Before the relationship between education and skills is tested a spearman test is performed;

	Chores	HaveEx	LikeOrder	ShirkD	FollowS	Reflecting	AmExac
ChoresDone	1.000						
HaveExcellent	0.1044	1.000					
LikeOrder	0.3733	0.0837	1.000				
ShirkDuties	0.2983	0.1193	0.2767	1.000			
FollowSchedule	0.2976	0.1687	0.3776	0.2161	1.000		
Reflecting	0.1163	0.2238	0.2041	0.1946	0.2440	1.000	
AmExacting	0.1690	0.2517	0.2448	0.2336	0.2885	0.2141	1.000
PersonOfWorth	0.0881	0.2140	0.1087	0.2629	0.0973	0.1504	0.1446
GoodQualities	0.0839	0.3216	0.1288	0.2595	0.1424	0.1975	0.2337
Confidence	0.1605	0.3195	0.1272	0.2744	0.1695	0.1859	0.2321
Responsibility	0.0451	0.3126	0.0294	0.1301	0.1955	0.2229	0.2963

	PersonW	GoodQ	Confid	Responsibility
PersonOfWorth	1.000			
GoodQualities	0.7207	1.000		
Confidence	0.5235	0.5728	1.000	
Responsibility	0.1595	0.2291	0.2935	1.000

Table 7

As seen above none of the 11 skills has a problematic correlation with another skill. The strongest correlation has a coefficient of 0.7207, between having good qualities and doing chores. This makes multicollinearity between skills unlikely.

## Education and Income

The next hypothesis to test is; education has a significant effect on skills. As income can be related to market related skills it is also tested in this section as if it were a skill.

Regression	OLS	Number of obs:		66323
		F( 15, 4682)		74.47
		Prob > F		0.0000
		R-squared		0.0925
		Root MSE		3.691
LNinc	Coef.	Robust SE	t	P>t
<b>vmbo</b>	0.4482	0.1498	2.99	0.003
<b>havovwo</b>	1.3318	0.1852	7.19	0.000
<b>mbo</b>	1.3439	0.1554	8.65	0.000
<b>hbo</b>	2.3655	0.1541	15.35	0.000
<b>university</b>	2.9315	0.1806	16.23	0.000
<b>schoolother</b>	0.2151	1.0964	0.20	0.844
<b>gender</b>	1.1476	0.0762	15.05	0.000
<b>age</b>	0.0981	0.0151	6.50	0.000
<b>age2</b>	-0.0010	0.0001	-7.15	0.000
<b>foreign</b>	-0.4032	0.1032	-3.91	0.000
<b>dwave2</b>	-0.0049	0.0055	-0.88	0.378
<b>dwave3</b>	0.0036	0.0083	0.43	0.665
<b>dwave4</b>	0.0070	0.0108	0.65	0.518
<b>dwave5</b>	-0.0079	0.0131	-0.61	0.544
<b>dwave6</b>	-0.0190	0.0151	-1.26	0.207
<b>dwave7</b>	-0.0467	0.0170	-2.74	0.006
<b>_cons</b>	0.2007	0.3893	0.52	0.606

Table 8

All educational levels have a significantly positive effect on income. This was as expected. Having a high education on the curriculum vitae is generally thought of to be a sign of skills, and associates with a high salary.

Being a man is positively associated with having a high salary, and being a foreigner is negatively associated with a lower salary. This could indicate a discrimination effect.



Also age and age squared are positively associated with a high salary, which makes sense when careers take off.

The OLS regressions containing only men or women show the very same effects. The only exception is that vmbo doesn't lead to a significantly higher salary than primary school when the dataset contains only men, *ceteris paribus*. This is surprising as completing this education is considered to have some value. A possible explanation would be that the minimum wage allows people who only completed primary school to earn a similar salary as people with vmbo.

## Education and Skills

There are 11 skills as seen before. Using these 11 as dependent variables, then testing for men, women and the whole dataset, means 33 OLS regressions are made and included in the appendix. Each time a Wald test is performed to test the significance of the joint effect of education and included in the tables below. Also the r squared of the regression and the values of the university coefficients are shown. The expectation given the theory is that people learn to make schedules, learn to do chores, gain qualities and gain all sorts of skills.

Skill	University	R-squared	F-test	Prob>F
<b>Chores done</b>	-0.269	0.0403	13.25	0.0000
<b>Excellent ideas</b>	0.232	0.0331	20.48	0.0000
<b>Like order</b>	-0.1598	0.0205	3.64	0.0001
<b>Shirk duties (5=low)</b>	0.1938	0.0518	6.73	0.0000
<b>Follow a schedule</b>	0.2956	0.0153	17.22	0.0000
<b>Reflecting</b>	0.3004	0.0236	28.44	0.0000
<b>Am Exacting</b>	0.4303	0.0459	41.35	0.0000
<b>Person of worth</b>	0.198	0.0149	6.40	0.0000
<b>Good Qualities</b>	0.378	0.0359	26.72	0.0000
<b>Confidence</b>	0.2714	0.0260	10.66	0.0000
<b>Responsibility</b>	1.404	0.1204	134.33	0.0000

Table 9

The education dummies are jointly significant in their effect on every skill, even at a 1% significance value, *ceteris paribus*. Having a university degree was however negatively associated with the skills "chores done" and "like order". Its largest effect was that on the skill "I like to have the responsibility

of handling a situation that requires a lot of thinking”. People with a university degree assessed this value on average 1.404 points higher than people with primary school as their highest completed education, ceteris paribus. It must also be noted that the values of R-squares are low, indicating the independent variables as education cannot predict the values of skills well.

The following comparison is for men;

<b>Skill</b>	<b>University</b>	<b>R-squared</b>	<b>F-test</b>	<b>Prob&gt;F</b>
<b>Chores done</b>	-0.2241	0.0417	83.20	0.0000
<b>Excellent ideas</b>	0.2782	0.0381	19.10	0.0000
<b>Like order</b>	-0.0231	0.0160	9.52	0.0000
<b>Shirk duties (5=low)</b>	0.2586	0.0454	8.96	0.0000
<b>Follow a schedule</b>	0.2757	0.0189	48.25	0.0000
<b>Reflecting</b>	0.3252	0.0256	29.98	0.0000
<b>Am Exacting</b>	0.4717	0.0357	19.10	0.0000
<b>Person of worth</b>	0.1136	0.0144	5.37	0.0000
<b>Good Qualities</b>	0.3538	0.0422	12.23	0.0000
<b>Confidence</b>	0.2245	0.0170	11.32	0.0000
<b>Responsibility</b>	1.356	0.1119	114.37	0.0000

Table 10

This is the table for woman;

<b>Skill</b>	<b>University</b>	<b>R-squared</b>	<b>F-test</b>	<b>Prob&gt;F</b>
<b>Chores done</b>	-0.3245	0.0241	9.40	0.0000
<b>Excellent ideas</b>	0.2240	0.0210	11.46	0.0000
<b>Like order</b>	-0.290	0.0133	6.44	0.0000
<b>Shirk duties (5=low)</b>	0.1172	0.0478	1.89	0.0780
<b>Follow a schedule</b>	0.3191	0.0185	9.20	0.0000
<b>Reflecting</b>	0.2787	0.0233	13.03	0.0000
<b>Am Exacting</b>	0.4183	0.0565	24.79	0.0000
<b>Person of worth</b>	0.4378	0.0165	4.14	0.0004
<b>Good Qualities</b>	0.4645	0.0332	14.27	0.0000
<b>Confidence</b>	0.3547	0.0178	5.97	0.0000
<b>Responsibility</b>	1.540	0.0807	63.42	0.0000

Table 11

The results are remarkably similar for both men and women, and are similar to the combined table. “Chores done” and “like order” are negatively associated with education, and jointly significant at 5%. The details are yet again in the appendix. The notable difference with the earlier results is that for women shirking duties is no longer associated with education.

The results for this hypothesis were mostly as expected. Most of this skills are positively and significantly associated with (the joint effect of) education. Two skills were negatively associated with education. Perhaps people that never follow an education are forced into doing their chores. Liking order is more an attitude then a skill, but perhaps not having a boss around and doing homework in the spare time is something students grow to like.

## Age and Skills

The previous regressions attempted to show the association between skills and education. A regression can show that if one variable is high, the other is also high. It could be the case that education leads to skills, which is the theory. But being skilled may also lead to a better education. This section attempts to show that, according to theory, education leads to more skills. If the effects of hypothesis 2 are causal, skills should increase with age as long as people are in school.

Skill	Breusch-Pagan	Age*Student	P	age	R-squared
Chores done	0.0072	-0.0053	0.001	0.0081	0.0406
Excellent ideas	0.0000	0.0002	0.829	-0.0031	0.0226
Like order	0.0000	-0.0034	0.029	0.0055	0.0165
Shirk duties (5=low)	0.0000	-0.0020	0.146	0.0119	0.0728
Follow a schedule	0.0000	-0.005	0.001	0.0042	0.0275
Reflecting	0.0000	0.0044	0.000	0.0054	0.0242
Am Exacting	0.0000	-0.0029	0.055	-0.0001	0.0430
Person of worth	0.0000	-0.0030	0.142	0.0062	0.0171
Good Qualities	0.0000	-0.0044	0.004	-0.0005	0.0249
Confidence	0.0000	-0.0001	0.949	0.0042	0.0160
Responsibility	0.0000	0.0101	0.000	0.0024	0.0940

Table 12

All OLS regression had heteroscedasticity. The first column in the table above shows that using a Breusch-Pagan test for heteroscedasticity. Clustering was used as described in the methods for the further regressions.

Most skills increase with age, having excellent ideas, exacting at work and having good qualities being the exceptions. The student variable includes both students and pupils. The combination of age and the interaction effect of age and student shows how much these skills change during the time people are students or pupils. The variable of interest is the interaction effect however, as it shows the difference compared to the trend. The variables “chores done” and “like order” increase with age during the time that people are students or pupils, but the effect is lower than the positive trend when they are not in school. , The variables “follow a schedule”, “I am exacting” and “good qualities” decrease with age when people are still in school, and the effect of schooling is a negative portion within the trend. This means these last 5 variables are unlikely be the causal effect of education, unless the effect of education is negative. The same goes for having excellent ideas, shirking duties, confidence, and being a person of worth. The reason for this is that the interaction effect of age and being a student/pupil is insignificant, even at a 10% significance level.

Skills that could be the causal positive result of education are “responsibility” and “reflecting” (see the table above).

The same test is done for both men and women. The results for men are these;

<b>Skill</b>	<b>Age*Student</b>	<b>P</b>	<b>age</b>	<b>R-squared</b>
<b>Chores done</b>	-0.0027	0.272	0.0105	0.0541
<b>Excellent ideas</b>	-0.0008	0.662	-0.0052	0.0310
<b>Like order</b>	-0.0031	0.245	0.0073	0.0276
<b>Shirk duties (5=low)</b>	-0.0006	0.795	0.0125	0.0774
<b>Follow a schedule</b>	-0.0083	0.001	0.0065	0.0427
<b>Reflecting</b>	0.0029	0.179	0.0049	0.0262
<b>Am Exacting</b>	-0.0101	0.000	0.0010	0.0407
<b>Person of worth</b>	-0.0103	0.003	0.0032	0.0146
<b>Good Qualities</b>	-0.0096	0.000	-0.0041	0.0281
<b>Confidence</b>	-0.0044	0.136	0.0002	0.0118
<b>Responsibility</b>	0.0031	0.412	-0.0001	0.1025

Table 13

The only significant results were those of “follow a schedule”, “Am exacting at my work”, “I am a person of worth” and “I have good qualities”. The coefficients are negative meaning that education is unlikely to have a positive causal effect on skills for men. This contradicts the theory that is being tested, and this remarkable result will be discussed in the discussion section.

These were the results for women;

Skill	Age*Student	P	age	R-squared
Chores done	-0.0064	0.001	0.0070	0.0373
Excellent ideas	0.0007	0.656	-0.0016	0.0185
Like order	-0.0033	0.088	0.0045	0.0147
Shirk duties (5=low)	-0.0021	0.236	0.0122	0.0817
Follow a schedule	-0.0031	0.104	0.0025	0.0243
Reflecting	0.0057	0.000	0.0059	0.0239
Am Exacting	0.0014	0.471	-0.0009	0.0518
Person of worth	0.0019	0.469	0.0089	0.0206
Good Qualities	-0.0005	0.784	0.0030	0.0278
Confidence	0.0023	0.325	0.0071	0.0182
Responsibility	0.0136	0.000	0.0026	0.0763

Table 14

When testing education on the proxy for education only the skills of reflecting and liking responsibility tested significant and positive. Chores done had a significant negative effect compared to the positive trend of age, ceteris paribus. The same holds true for the “skill” of liking order. The latter is only significant at a 10% significance level.

The most important finding of the relationship between age and most skills doesn’t look positive, thus their causality is put into question. On the other hand, measurement errors due to people self-assessing their skills incorrectly may have influenced these results (see discussion).

It must yet again be noted however that the crucial variable “Student” was differently defined in different waves in the survey. In the first two waves it included internships and in the rest it didn’t. Although it isn’t likely to have a major effect, it may still cause bias. Dummies per wave have been included in the models to counteract this problem as much as possible. The amount of people being a student across the waves do need seem to be larger in the first 2 waves (see appendix).

Another problem with the variable Student is that the amount of students in the sample is limited. In wave 6 for example, the sample only contained 444 students, and 5123 non-students. The small sample size and the use of robust errors may cause these relationships between skills and age not to be significant

### Skills and Smoking Cessation

The third hypothesis is; a change in skills leads to an increase in the likelihood to quit smoking. A fixed effects panel regression is created for men, women and the groups combined. To omit statistical issues as heteroscedasticity (see appendix) robust errors are used for all models.

	<b>Coefficient</b>	<b>Robust SE</b>	<b>t</b>	<b>P&gt;t</b>
<b>Chores done</b>	0.0018	0.0047	0.37	0.712
<b>Have excellent ideas</b>	0.0035	0.0057	0.61	0.542
<b>Like order</b>	0.0021	0.0048	0.44	0.660
<b>Shirk my duties (5=low)</b>	-0.0020	0.0050	-0.40	0.687
<b>Follow a schedule</b>	0.0006	0.0047	0.12	0.903
<b>Reflecting</b>	0.0011	0.0051	0.22	0.823
<b>Am exacting at my work</b>	-0.0043	0.0047	-0.91	0.365
<b>Person of worth</b>	-0.0037	0.0041	-0.89	0.373
<b>Good qualities</b>	-0.0051	0.0053	-0.96	0.338
<b>Confidence</b>	0.0004	0.0043	0.09	0.926
<b>Responsibility</b>	-0.0036	0.0029	-1.25	0.211
<b>age</b>	0.0329	0.0081	4.08	0.000
<b>age2</b>	-0.0002	0.0001	-2.56	0.011
<b>_cons</b>	-0.5433	0.2481	-2.26	0.024

Table 15

The model above shows that people with a higher age have an increased likelihood to quit smoking compared to those with a lower age, ceteris paribus. Every extra year increases the likelihood of quitting with 3.29 percentage points. This effect is even significant at 1%. The small effects of age squared have to be reduced from this effect however to see the true size of the effect of age, that effect is negative and significant at 5%.

---

**R-sq: within = 0.0184**

**between = 0.0438**

**overall = 0.0422**

---

*Table 16*

Again noteworthy is that the regression only explains a small portion of what is going on, as seen in the table above.

The most important conclusion from the above regression is that none of the skills had a significant effect on the likelihood to quit smoking. The regressions for men and women showed that same result (see appendix).

In earlier hypothesis the conclusion was that education and responsibility were the only soft skills that seem to be causality affected by education. The same test was done with only these 2 skills.

<b>Quit</b>	<b>Coefficient</b>	<b>Robust SE</b>	<b>t</b>	<b>P&gt;t</b>
<b>Reflecting</b>	0.0006	0.0049	0.12	0.901
<b>Responsibility</b>	-0.0039	0.0029	-1.37	0.172
<b>age</b>	0.0328	0.0081	4.06	0.000
<b>age2</b>	-0.0002	0.0001	-2.50	0.012
<b>_cons</b>	-0.5917	0.2390	-2.48	0.013

*Table 17*

The skill of reflecting and responsibility are not significantly associated with quitting smoking. A possible explanation is that the skills are self-assessed, and the surveyed not being able to do this well (see discussion).

These are the results for men and women respectively, only including the two aforementioned skills;

<b>Quit</b>	<b>Coefficient</b>	<b>Robust SE</b>	<b>t</b>	<b>P&gt;t</b>
<b>Reflecting</b>	0.0043	0.0065	0.66	0.506
<b>Responsibility</b>	-0.0089	0.0040	-2.23	0.026
<b>age</b>	0.0373	0.0119	3.12	0.002
<b>age2</b>	-0.0002	0.0001	-2.22	0.026
<b>_cons</b>	-0.7172	0.3688	-1.94	0.052

*Table 18*

Quit	Coefficient	Robust SE	t	P>t
<b>Reflecting</b>	-0.0032	0.0076	-0.42	0.673
<b>Responsibility</b>	0.0003	0.0040	0.07	0.942
<b>age</b>	0.0279	0.0114	2.44	0.015
<b>age2</b>	-0.0001	0.0001	-1.18	0.240
<b>_cons</b>	-0.4574	0.3212	-1.42	0.155

Table 19

Yet again reflecting and responsibility have no significantly positive effect on the likelihood of quitting smoking. This means that the most important part of the research, namely testing the theory that education leads to quitting smoking through skills, is unproven. Or from a different perspective; within the limitations of the model the conclusion is that education does not lead to a higher likelihood to quit smoking through skills.

## Education Skills and Smoking Cessation

An OLS regression was made including all variables except skills for the 3 types of sample groups that were used throughout this thesis (men, women, combined). All these regressions were heteroskedastic (see appendix). A robust OLS regression using only wave 6, as described in the methods. These regressions test the effect that education has on quitting smoking, with and without including skills. The differences in the coefficients of the education dummies tell the effect of skills. These were the results without skills for the combined group of men and women;

Quit	Coefficient	Robust SE	t	P>t
<b>LNinc</b>	.0071557	.0022108	3.24	0.001
<b>vmbo</b>	.0713635	.0319409	2.23	0.026
<b>mbo</b>	.1111587	.033477	3.32	0.001
<b>hbo</b>	.1611224	.0330513	4.87	0.000
<b>university</b>	.1880902	.0412264	4.56	0.000
<b>havovwo</b>	.1198707	.0394226	3.04	0.002
<b>schoolother</b>	.6536254	.0435967	14.99	0.000
<b>age</b>	.0016966	.0038086	0.45	0.656
<b>age2</b>	.0000618	.0000327	1.89	0.059
<b>foreign</b>	.0210975	.0253043	0.83	0.404



<b>geslacht</b>	-0.0407087	.0168603	-2.41	0.016
<b>_cons</b>	.2554892	.1116784	2.29	0.022

Table 20

These are the results when skills are included;

<b>Quit</b>	<b>Coefficient</b>	<b>Robust SE</b>	<b>t</b>	<b>P&gt;t</b>
<b>LNinc</b>	.0065107	.0024089	2.70	0.007
<b>ChoresDone</b>	.0240948	.0104094	2.31	0.021
<b>HaveExcellent</b>	.001593	.0134252	0.12	0.906
<b>LikeOrder</b>	-0.0076414	.0117665	-0.65	0.516
<b>ShirkDuties</b>	-0.0052348	.011689	-0.45	0.654
<b>FollowSchedule</b>	.0212006	.011568	1.83	0.067
<b>Reflecting</b>	-0.030028	.0124294	-2.42	0.016
<b>AmExacting</b>	.0107704	.010752	1.00	0.317
<b>PersonOfWorth</b>	.0087311	.0108774	0.80	0.422
<b>GoodQualities</b>	-0.0134022	.0137185	-0.98	0.329
<b>Confidence</b>	-0.0004083	.0111321	-0.04	0.971
<b>Responsibility</b>	-0.0181509	.0065056	-2.79	0.005
<b>vmbo</b>	.0662109	.0334915	1.98	0.048
<b>mbo</b>	.0961049	.0357992	2.68	0.007
<b>hbo</b>	.1692359	.0358148	4.73	0.000
<b>university</b>	.202336	.0452299	4.47	0.000
<b>havovwo</b>	.1149157	.0421058	2.73	0.006
<b>schoolother</b>	.6655676	.0576508	11.54	0.000
<b>age</b>	.004689	.0043399	1.08	0.280
<b>age2</b>	.0000385	.0000369	1.04	0.296
<b>foreign</b>	.0082429	.0279815	0.29	0.768
<b>geslacht</b>	-0.0284188	.0189013	-1.50	0.133
<b>_cons</b>	.2407294	.1469257	1.64	0.101

Table 21

In both regressions education and quitting smoking are positively associated. However, the coefficients of the education dummies are different. In table 20 the education dummies contain information regarding skills, but this is corrected for in table 21. The differences are shown in the table below.

Level	Coefficient (Table 20)	Coefficient (Table 21)	Table 20-21
vmbo	0.0713635	0.0662109	0.0051526
mbo	0.1111587	0.0961049	0.0150538
hbo	0.1611224	0.1692359	-0.0081135
university	0.1880902	0.202336	-0.0142458
havovwo	0.1198707	0.1149157	0.004955
schoolother	0.6536254	0.6655676	-0.0119422

Table 22

The effect of skills that explain the relationship between education and the likelihood of quitting smoking is marginal. The effect these 11 skills explain only account for an extra 0.51 percentage points to quit smoking when comparing vmbo to no education. Small effects were expected, but perhaps not this small. The role it plays in the effect on university and hbo seems negative. In earlier hypotheses certain skills were already negatively associated with quitting smoking. An explanation could be that the surveyed don't self-assess these skills well. Another explanation lies in the fact that only people that once smoked were included. Skills may be correlated with both starting smoking and education, causing a bias in the results. People that reflect well might not start smoking to begin with, and people that reflect well may also have a different education level.

In earlier hypotheses it was shown that the skills either don't have an effect on quitting smoking or that the effects weren't causal. The effects are therefore to be interpreted as a correlation. The effects for men and women don't show notable differences and are included in the appendix.

These regressions also show an effect that could not be tested using the panel regressions in hypothesis 3; the effect of income on quitting smoking. A logarithm was used to measure the effect of income. As seen in table 21 an increase of 1% income is associated with an extra 0.65 percentage points to quit smoking.

If income was only a budget constraint one would expect that those with more income could buy more cigarettes. However, higher socioeconomic groups of all sorts tend to behave differently than others. It must also be noted that the effect of income changes when skills are not included in the model. This shows that income, as predicted, is partly explained by the effect of skills. Unexpectedly the effect of skills was mostly negative however, and this is reflected in the effect of income in the 2 different models above.

## Chapter 5: Conclusions

This thesis attempts to improve the literature around the suspected causal effect where an increase educational level leads to improved smoking behavior. Maralani (2014) explained that there are several effects that can explain how education can lead to improved smoking behavior. She argues that one of the ways that this can happen is that students learn “soft skills”, like time management skills at school and use said skills to for smoking decisions. The most important finding of this thesis is that the 11 skills that were tested were not significantly positively associated with either quitting smoking or education, or failed to pass a test for causality.

The second most important finding is the explanation regards the effect of income. The theory, which is not based on earlier research, is that education leads to income, which leads to quitting smoking. Part of the theory is that income captures market related skills that are valued by the employer, thus partly depends on the same theory. The tests show significant associations between (logarithmic) income and quitting smoking. It was also shown that when regressing quitting smoking on income, the coefficients of income change when correcting for skills. This means that the effect that income has on quitting smoking is partly a skill related effect.

It was already known that education had a causal effect on the likelihood to be smoking (Grimard & Parent, 2007), and had shown the association with quitting smoking (Koning, Webbink & Martin, 2010). This research confirms the latter association is valid for this dataset. All educational levels have significant effects on quitting smoking. The same effect was found when testing for men and women separately.

The relationship between skills and education was being tested second. The education dummies were jointly significantly associated with every skill, but not always in a positive way. Education seemed to have a positive effect on the self-assessed “skills” of having excellent ideas, not shirking duties, following a schedule, reflecting, exacting at work, having good qualities, being a person of worth, having confidence in you capabilities and liking to take responsibility. Also shown was that income is significantly associated with a higher income. This was expected.

If the theory is correct, skills would increase over time when students are still in school. If skills increase naturally with age, it should increase even more while still being a student. Testing the effect of the interaction effect of age and student on all 11 skills showed that only 2 skills could possibly explain the effect of education on quitting smoking causally. For men this was responsibility, for women it was both reflecting and responsibility.

Several skills, including having good qualities, are positively and significantly associated to education, but the relationship doesn't run from education to skills in most cases. Having good qualities and certain skills may be a requirement to complete such an education. Another option is that another effect can increase both skills and education.

The relationship between education and skills was tested next. A fixed effects panel regression showed that none of the 11 skills had a significant positive effect on the likelihood to quit smoking. Combined with the low R-squared values of the models these facts show that other factors may explain the relationship better.

In the last part of the thesis it was shown that skills only explain a small portion of the education effect in quitting smoking. Income on the other hand did have a significant effect on the likelihood to quit smoking. The data provided cannot however show causality for this effect. A potential causal effect for education leading to improved smoking behavior through income deserves a recommendation for further research (see discussion).

The research question was answered using the 2<sup>nd</sup> and 4<sup>th</sup> hypothesis. The question was; to what extent does an increase in skills explain the relationship between education and quitting smoking? Summarized, education could only causally lead to reflecting and responsibility. And both these skills do not lead to a significant increased likelihood to quit smoking. The causal part of the theory has not been proven. Either it isn't true or the surveyed have difficulties assessing their skills (see discussion). Skills do however explain part of the relationship between education and smoking in a non-causal way. The small associations in table 22 showed that skills play a role depending on the educational level and can be associated with either a decrease or increase of the likelihood of quitting smoking of less than 2 percentage points.

## Chapter 6: Discussion

This thesis attempted to show how education leads to the skills required, like responsibility, to quit smoking. As seen in the conclusions, several hypotheses were rejected. Eventually the possibility of these skills interpreting a positive effect between education and smoking in a causal way was rejected for all skills. In this section the research itself will be discussed. Underlined are the most important issues.

### Survey and variables

The LISS dataset provided a lot of information. It had most of the variables that were needed, especially regarding education and smoking. Only people that ever smoked were asked if they had quit smoking. This forced the sample to only contain people who ever smoked. Combining this with only using people aged 25 and higher cut away a part of the dataset. The downside of this is that valuable information is missing. Maralani (2013) explained that the probability of smoking is mostly determined by the decision never to start smoking, and that the relationship between education and quitting smoking is only a small effect. This may explain in part some of the low R-squares in the regressions. The R-squares have been low for all regressions, meaning the models only explain a small portion of the effects on the dependent variables. The goal of this thesis is mostly to explain the relationship between education, smoking and skills. However, the relationship between education and skills has often been researched for a different reason. From a public policy perspective it can be desired to find an effective way to get people to stop smoking. This can be done by finding all the important predictors of quitting smoking and researching the cost effectiveness of the policy measures. This at the very least required the r-squares of this research to be high. The low r-squares leave room for omitted variables that may represent other variables (and different policy measures) with higher coefficients. This research is there for limited with regards to policy implications.

Also lacking was the variable of skills. It was better to have had IQ and problem solving capabilities as skills in the dataset. Persistence would also be good. “I have confidence in my capabilities” and “I have good qualities” and “I am a person of worth” were used as skills, but are more of a proxy of all skills rather than one in general. The skills are also subjective; especially “I have excellent ideas”. On the other hand the variable income was significant everywhere and isn’t subjective.

Probably the most important topic of discussion is the self-assessment of skills. Skills are most likely difficult to assess objectively yourself. This may cause a measurement error, which may lead to implausible results. The skill “reflecting” has an ironic bias as people who don’t reflect well may still self-assess this value as high as a result of poor reflecting.

A measurement error could be problematic. Obviously the value should have a meaning, but the self-assessed values might tell more about the confidence people have in their own skills than those skills actually being reliable variables. If certain education or smoking groups are overestimating their skills more than other groups, then the skills variable could be biased.

In the cross sectional data the measurement errors could have led to standard errors being higher. The issues for measurement error in the fixed effects model are different however. The model measures differences in responses of a person being interviewed. A measurement error in 1 wave therefor can lead to the outcome of a combination of waves being incorrect. An example; a person in wave 1 think he/she has good qualities, then figures this was incorrect and fills in a lower score in wave 2. The fixed effects model then assumes that the variable “good qualities” has good done, but in fact the measurement error has simply been corrected. The change was related to the error, not to the skill.

There was problem with the skills dataset. The interviewed were all asked what their personality was/ skills were in wave 1 and 2 of the personality section. People that already gave answers to these questions in wave 2 weren’t asked for them in wave 3. The same applies to wave 5. This means there is a lack of data for these waves, and is also one of the reasons why only wave 6 was used to test hypothesis 1.

There was another problem with the survey. The questions seem to be the same across the waves. There is 1 exception however. The variable student hasn’t been the same across the waves. In earlier versions student could mean student or pupil. In several later waves also interns were included. This means there are could be more students in those waves, and student has a different meaning. Wave dummies have been included but may not solve the problem. The variable student, more specifically the interaction effect of age\*student, was crucial with regards to testing for causality.

The website of LISS data tells that only people of age 16 and higher were surveyed, but the background variables showed that some people weren’t older than 3. This might be caused by measurement errors, or not following the rules. People below the age of 4 have been cut from the dataset. Another reason to do this was that they haven’t gone to school yet.

## Methods and results

The initial results confirmed that there is an association between education and quitting smoking. Obviously better than showing an association is proving causality. Sander (1995) used parental education as an instrumental variable to show that education leads to quitting smoking, rather than smoking behavior leading to more education. Such a variable wasn't in the dataset. This is a limitation of the research. Using a war draft, such as de Walque (2007), would have been another way of showing this, but there wasn't such a dataset available.

The amount of options available to research the relationship between education and quitting smoking is limited. Fixed effects regressions aren't possible as education doesn't change. This has to do with the fact that only people of age 25 and older were used to prevent bias. Another problem is the presence of heteroscedasticity. The coefficients of a logit may be biased in the presence of heteroscedasticity, thus another model than logit must be used. Another problem is that in addition to education not changing, smoking behavior isn't likely to change per person either, especially per wave. This is related to the surveyed being asked the quit smoking question a limited amount of times while only covering a small time frame, smokers having quit by the first wave, smokers not quitting until the last wave. Smoking is addictive, thus a person keeps smoking for quite a number of years before changing their smoking behavior. Using the entire panel and clustering could show significant results even if they aren't significant, simply because the same thing is measured several times. Measuring smoking decisions more than just 7 waves would have been better.

Another limitation of researching the relationship between education and smoking behavior is related to the problem of smoking at young ages. Children smoke before they gain extra years of education or before they complete a certain level of education. This makes it difficult to make statements regarding causality and the influence of group behavior has to be corrected for (Maralani, 2013). This is one of the reasons the variable quitting smoking has been researched. Quitting smoking is done at a later age, as a person has to start smoking before quitting.

The relationship between education and skills shows that, when people are still students/pupils, their "skill" in doing chores and liking order, significantly decreases. These results also seem to show causality. Previously only significant positive effects were discussed, as the theory is being tested. These results seem to imply that education makes people relatively lazy. Either that or the surveyed aren't self-assessing well. The latter makes most sense, although these results are questionable. An added value of a school is to teach people to become better at all sorts of things. Liking order is probably not a real skill however. The reason to see it as a skill was that liking order may be related

to being orderly. Perhaps the surveyed think order is just less important later in their studies, without being related to a skill. The chores done variable stands for “get chores done right away”. Perhaps the highly educated have more work to do and therefore cannot do their chores right away. Doing all your chores right away can be considered a skill, but having more chores makes this impossible. Another option is that education makes people lazy, and this is the part that would be questionable.

In hypothesis 2 it would have made more sense from a statistical point of view to use an ordered regression rather than an OLS regression. Skills rank from 1 to 5 or 7. The chosen method is only valid under the assumption that the difference between 1 and 2 is the same as the difference as the difference between 2 and 3. This assumption isn't likely to be completely true. On the other hand this makes the thesis much more readable then when showing and interpreting the results compared to ordered regressions, and the variable being ordinal or cardinal often makes very little difference (Ferrer & Frijters, 2004).

Since 2014 it isn't allowed to smoke before the age of 18. This could affect wave 7. This made it sometimes necessary to add dummies to correct for year specific effects. A remarkable effect is that of education and income has a significant dummy for wave 7. The dummy for wave 7 showed that the year 2014 had a significantly negative effect on income. The variable income is a background variable however. It was only measured once. Also there wasn't a recession in 2014. The most likely explanation is that more poor people joined the survey in that year, or rich people no longer answered the questions of the survey. The population has been somewhat different per wave.

Noteworthy is that nearly all regressions suffered from heteroscedasticity. This involves all hypotheses, with or without skills. Robust errors have been used in most regressions, but the downside is that effects that might exist may show up as insignificant due to this. On the other hand a lot of regressions were made, especially in hypothesis 2. When many variables are tested, there is a good chance something became significant even when no such effect exists. Accepting only results with a P value under 0.01 might have been better.

Another problem was the presence of significant Ramsey RESET tests. This could mean a lot of different things. In a lot of cases, regressions didn't test positive for this when age squared wasn't involved, but became significant when age squared was in the model, and sometimes it was the other way around. Consistency was preferred, testing the same model for men, women and everybody. After including age squared, this has led to doing nothing with significant Ramsey RESET tests.



## Looking ahead

The results have led, step by step, to a situation where none of the skills involved could explain the relationship between education and quitting smoking. Not being significant is also a result, although not being a practical one. Significant results could have led to schools being more of a tool in the cessation of smoking, as it is then known how to get people to stop smoking. To complete the framework of Maralani (2014), it would be necessary to have an improved panel survey with skills that can be objectively measured. This is however unlikely to happen. The results of skills such as “good qualities”, assuming limited measurement errors, do not add much value to the explanations of quitting smoking as seen by the R-squared of the regressions.

The most interesting result is the effect of income. Income is significantly associated with a high education, and also with a bigger likelihood to quit smoking. One would assume that if income is just a budget constraint a smoker could buy more cigarettes. Income is a background variable that has only been measured once in this dataset however. It would be interesting to see more research on this topic. Added value would be proving causality. With this recommendation for more research, I conclude this thesis.

## Literature

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, 84(2), 191.

Bjornson, W., Rand, C., Connett, J. E., Lindgren, P., Nides, M., Pope, F., ... & O'Hara, P. (1995). Gender differences in smoking cessation after 3 years in the Lung Health Study. *American Journal of Public Health*, 85(2), 223-230.

Blake, S. M., Klepp, K. I., Pechacek, T. F., Folsom, A. R., Luepker, R. V., Jacobs, D. R., & Mittelmark, M. B. (1989). Differences in smoking cessation strategies between men and women. *Addictive Behaviors*, 14(4), 409-418.

Chaloupka, F. J. (1990). *Rational addictive behavior and cigarette smoking* (No. w3268). National Bureau of Economic Research.

Crone, M. R., & van Kesteren, N. M. C. (2007). *De rol van de eigen identiteit en de eigen effectiviteitsverwachtingen bij het stoppen met roken: een literatuuronderzoek*. TNO Kwaliteit van Leven.

Cutler, D. M., & Lleras-Muney, A. (2010). Understanding differences in health behaviors by education. *Journal of health economics*, 29(1), 1-28.

Cutler, D. M., & Lleras-Muney, A. (2006). *Education and health: evaluating theories and evidence* (No. w12352). National Bureau of Economic Research.

De Walque, D. (2004). Education, information, and smoking decisions: evidence from smoking histories, 1940-2000. *World Bank policy research working paper*, (3362).

De Walque, D. (2007). Does education affect smoking behaviors?: Evidence using the Vietnam draft as an instrument for college education. *Journal of health economics*, 26(5), 877-895.

Dijkstra, A., & Vries, H. D. (2000). Self-efficacy expectations with regard to different tasks in smoking cessation. *Psychology and Health*, 15(4), 501-511.

Doll, R., & Hill, A. B. (1950). Smoking and carcinoma of the lung. *British medical journal*, 2(4682), 739.

Eisinger, R. A. (1972). Psychosocial predictors of smoking behavior change. *Social Science & Medicine* (1967), 6(1), 137-144.

Ferrer-i-Carbonell, A., & Frijters, P. (2004). How Important is Methodology for the estimates of the determinants of Happiness?\*. *The Economic Journal*, 114(497), 641-659.

Field, M., Mogg, K., & Bradley, B. P. (2005). Alcohol increases cognitive biases for smoking cues in smokers. *Psychopharmacology*, 180(1), 63-72.

Fiore, M. C., Novotny, T. E., Pierce, J. P., Giovino, G. A., Hatziandreu, E. J., Newcomb, P. A., ... & Davis, R. M. (1990). Methods used to quit smoking in the United States: Do cessation programs help?. *Jama*, 263(20), 2760-2765.

Grimard, F., & Parent, D. (2007). Education and smoking: Were Vietnam war draft avoiders also more likely to avoid smoking?. *Journal of Health Economics*, 26(5), 896-926.

Grossman, M. (1972). On the concept of health capital and the demand for health. *The Journal of Political Economy*, 223-255.

Grossman, M. (1976). The correlation between health and schooling. In *Household production and consumption* (pp. 147-224). NBER.

Haenszel, W., Miller, H. P., & Shimkin, M. B. (1956). Tobacco smoking patterns in the United States.

Hatziandreu, E. J., Pierce, J. P., Lefkopoulou, M., Fiore, M. C., Mills, S. L., Novotny, T. E., ... & Davis, R. M. (1990). Quitting smoking in the United States in 1986. *Journal of the National Cancer Institute*, 82(17), 1402-1406.

Hughes, J. R., Keely, J. P., Fagerstrom, K. O., & Callas, P. W. (2005). Intentions to quit smoking change over short periods of time. *Addictive behaviors*, 30(4), 653-662.

Kabat, G. C., & Wynder, E. L. (1987). Determinants of quitting smoking. *American Journal of Public Health*, 77(10), 1301-1305.

Kenkel, D. S. (1991). Health behavior, health knowledge, and schooling. *Journal of Political Economy*, 287-305.

Koning, P., Webbink, D., & Martin, N. G. (2010). The effect of education on smoking behavior: new evidence from smoking durations of a sample of twins. *Empirical Economics*, 1-19.

Lahelma, E., Martikainen, P., Laaksonen, M., & Aittomäki, A. (2004). Pathways between socioeconomic determinants of health. *Journal of Epidemiology and Community Health*, 58(4), 327-332.

- Maralani, V. (2013). Educational inequalities in smoking: the role of initiation versus quitting. *Social Science & Medicine*, *84*, 129-137.
- Maralani, V. (2014). Understanding the links between education and smoking. *Social science research*, *48*, 20-34.
- Meara, E. (2001). *Why is health related to socioeconomic status?* (No. w8231). National Bureau of Economic Research.
- Mermelstein, R., Cohen, S., Lichtenstein, E., Baer, J. S., & Kamarck, T. (1986). Social support and smoking cessation and maintenance. *Journal of consulting and clinical psychology*, *54*(4), 447.
- Monden, C. W., van Lenthe, F., De Graaf, N. D., & Kraaykamp, G. (2003). Partner's and own education: does who you live with matter for self-assessed health, smoking and excessive alcohol consumption?. *Social Science & Medicine*, *57*(10), 1901-1912.
- Novotny, T. E., Warner, K. E., Kendrick, J. S., & Remington, P. L. (1988). Smoking by blacks and whites: socioeconomic and demographic differences. *American journal of public health*, *78*(9), 1187-1189.
- Ockene, J. K., Benfari, R. C., Nuttall, R. L., Hurwitz, I., & Ockene, I. S. (1982). Relationship of psychosocial factors to smoking behavior change in an intervention program. *Preventive medicine*, *11*(1), 13-28.
- Pamuk, E. (Ed.). (1999). *Health United States 1998: With Socioeconomic Status and Health Chart Book*. DIANE Publishing.
- Roche, A. M., Eccleston, P., & Sanson-Fisher, R. (1996). Teaching smoking cessation skills to senior medical students: a block-randomized controlled trial of four different approaches. *Preventive medicine*, *25*(3), 251-258.
- Rogers, E. M. (2010). *Diffusion of innovations*. Simon and Schuster.
- Sacerdote, B. (2000). *Peer effects with random assignment: Results for Dartmouth roommates* (No. w7469). National bureau of economic research.
- Sander, W. (1995). Schooling and quitting smoking. *The Review of Economics and Statistics*, 191-199.
- Sasco, A. J., Secretan, M. B., & Straif, K. (2004). Tobacco smoking and cancer: a brief review of recent epidemiological evidence. *Lung cancer*, *45*, S3-S9.
- Schulz, B. (2008). The importance of soft skills: Education beyond academic knowledge.

Taubman, P., & Rosen, S. (1980). Healthiness, education, and marital status.

Tayade, M. C. (2014). Smoking: Historical aspects and Composition. *Asian J of Biosciences*, 1(01), 15-19.

US Public Health Service. (1980). *Health Consequences of Smoking For Women-a Report of the Surgeon General*.

US Department of Health and Human Services. (2014). The health consequences of smoking—50 years of progress. *A report of the Surgeon General*.

Vedøy, T. F. (2014). Tracing the cigarette epidemic: An age-period-cohort study of education, gender and smoking using a pseudo-panel approach. *Social science research*, 48, 35-47.

Viscusi, W. K. (1992). *Smoking: Making the risky decision*. Oxford University Press.

Wetter, D. W., Cofta-Gunn, L., Irvin, J. E., Fouladi, R. T., Wright, K., Daza, P., Mazas, C., Cinciripini, P.M. & Gritz, E. R. (2005). What accounts for the association of education and smoking cessation?. *Preventive Medicine*, 40(4), 452-460.

Willemsen, M. C. (2006). Rokers onder vuur?. *Inloed van de gewijzigde tabakswet op rokers, met speciale aandacht voor verschillen tussen sociaaleconomische klassen*. Den Haag: STIVORO.

Winkleby, M. A., Jatulis, D. E., Frank, E., & Fortmann, S. P. (1992). Socioeconomic status and health: how education, income, and occupation contribute to risk factors for cardiovascular disease. *American journal of public health*, 82(6), 816-820.

## Appendix

This section contains the output of the Stata do files that contain the regressions. Separate do files were used for creating variables and merging data sets. These were not included in the appendix. The part below contains all relevant outputs.