Bachelor Thesis

The effect of a change in the compulsory education system on the number of early school leavers

This research measured the effect of a change in the Dutch law of compulsory education in 2007 on the number of early school leavers. To tackle the problem of early school leavers the Dutch government introduced a change in the law of compulsory education. Schooling is important for individuals and society. Compulsory education helps more students get a basic qualification. This basic qualification is important for their future and their chances on the labor market. The main findings of this research are that the change in the law of compulsory education does not have a specific influence on the number of early school leavers in the Netherlands. It is important to note that not only this change caused a decrease in the number of early school leavers. Earlier projects, programs and investments done by the Dutch government had a significant positive influence on the number of early school leavers. Factors like, gender, household composition, paid work, self-reported happiness and self-reported health has a significant negative influence on the change of achieving a basic qualification for students with an age between 18 and 24 years. The analyses are conducted with the research methods, difference- in- difference analysis and multiple linear regression analysis. The outcomes are checked on robustness.

Professor: Dr. J.L.W. (Hans) van Kippersluis

Department of Applied Economics

Nishika Janki

374157

28th May 2015

Table of Contents

1. Introduction

Too many young people under the age of 18 experience difficulties in their learning career. They quit too early without any form of qualification. It is important to achieve a basic qualification because it gives a better perspective on the labour market and it gives you an own place in society. Apart from your individual achievement it is good for the whole society and economy if young people have a basic qualification. In order to tackle the problem of early school leaving without a basic qualification, the European Commission decided in 2000 that every country in Europe needs to reduce the percentage of early school leavers between the age of 18 and 24 before 2010. This agreement was a part of the popular Lisbon strategy (Commission, 2000). This strategy had to make the European economy the strongest. In addition to this strategy the European Commission came in 2009 with a new strategy named, Europe 2020. The seventeen European Countries made an agreement according the problem of early school leavers. This agreement implies that the number of early school leavers has to be under the 10% of the total students between the age of 18 and 24 years for each country (European-Commission, 2015).

In order to achieve the 10% the Dutch government decided in august 2007 to change the law of compulsory education. Before august 2007 there was partial compulsory education. Pupils who followed a full time education were of school age till 16. Pupils who followed a part time education were of school age till 18. After August 2007 all students have a school age till 16 years but they must attend school until they turn 18 or until they obtain a basic qualification. In the Netherlands a basic qualification is a VWO diploma, HAVO diploma or a MBO 2 diploma (Storimans, 2010).

The main purpose of the change in the law of compulsory education is to have more students with a basic qualification and to have less early school leavers. The Central Agency of Statistics measures the number of early school leavers every month and every year the Dutch Government presents a report about it. The report gives a wide understanding of the development of the problem. The number of early school leavers decreased since 2004. Next to this report the government studied much more. For example they found a link between the labour market and the career path of early school leavers. They also analyzed the link between age (18-23 year) and the number of early school leavers (Onderwijs, 2015). There has been a lot of case studies conducted on the effect of compulsory education, the effect of a change in compulsory education and why compulsory education matters. In one study the compelling effect of compulsory education is examined. The most important finding was that mandating education substantially increased adult income and substantially decreased the likelihood of being below the poverty line, unemployment and in a manual occupation (Oreopoulos, 2004). Another study has been conducted in order to measure the effect of a change in compulsory education on the earning of early school leavers. They found that a change in a compulsory education system significantly affect educational attainment, especially among individuals belonging to the lowest quintiles of the distribution of ability (Brunello, Fort, & Weber, 2012). More important for this research are the studies that looked at the importance of a achieving a basic qualification. A study showed that the probability of having a job with a basic qualification is considerably higher than without (Psacharopoulos & Layard, 1979). Early school leavers are of lesser health (Groot & Maassen van den Brink, 2007) and their children have a lower education level (Bowles, 1972).

The purpose of this research is to measure the effect of the change in the law of compulsory education in August 2007 on the number of early school leavers. The ideal way to measure this effect would be to measure the number of early school leavers with the treatment and compare this with the same group without the treatment. This is not possible because one country cannot be in the treatment group and control group at the same time. Therefore another method of

research is used. The effect will be measured in two parts. The first part will measure the average treatment effect. This part will show whether the change in compulsory education has worked. In other words it shows whether there are less early school leavers due to the policy change. To measure this difference-in-difference is conducted with The Netherlands as treatment group and several countries as control groups. The second part of this study will look at a change in the number of achieved basic qualification between the age 16 and 18 years and the age 18 and 24. This part will investigate whether the change in the law of compulsory education has led to an increase in pupils with a basic qualification between the age of 16 and 18 and the age of 18 and 24. Also this part will investigate which factors influence the change of achieving a basic qualification. These two parts together give an answer to the following research question:

What is the effect of the change in the law of compulsory education in August 2007 on the number of early school leavers?

The first part of this research will focus on the nation level. The data for this part is coming from the Central Agency of Statistics and Eurostat. The Central Agency of Statistics is a Dutch governmental institution that gathers statistical information about the Netherlands (CBS, Voortijdig Schoolverlaters, 2015). Eurostat is a Directorate-General of the European Commission located in Luxemburg. Its main responsibility is to provide statistical information of the European Union. The data of the number of early school leavers for several countries is collected from 2001 till 2012 because early data was not available and the number of early school leavers in 2013 is still uncertain (European-Commission, Eurostat, 2015). Both resources are reliable. The second part of this research will focus on individual level with the aid of the POLS data (CBS, Permanent Onderzoek LeefSituatie (POLS), 2015). This data is based on annual questionnaires filled in by random people in the Netherlands. The research sample consists of different parts that measure the living conditions. From this dataset the development in the number of basic qualifications can be investigated. The POLS data is from 2003 till 2009. After 2009 the Central Agency of Statistics stopped with the POLS.

For the first part of this research the average treatment effect is measured using difference-in-differences. With this research method the effect of a treatment on an outcome is calculated by comparing the average change over time in the outcome variable for the treatment group to the average change over time for the control group. The treatment group in this study is the Netherlands with control groups, France, Estonia and Latvia. In the second part of this research a multiple regression analysis will be conducted with a large number of control variables to measure the effect of different factors on achieving a basic qualification. From there, differences in factors influencing the chance of achieving a basic qualification before and after the treatment is measured.

In the remainder of this paper background information is presented. This background information consist of a literature review, the history of compulsory education in the Netherlands, a timeline of action taken by the government to reduce the number of early school leavers, explanation of the change in the law of compulsory education and an overview of compulsory education in Europe. From there the methodology and the descent of the data are presented. The methodology consists of the steps how this research is conducted and an explanation of the research methods. This will be followed by the results of this research and will lead to the conclusion. Before the conclusion a robustness check of the results is carried out. The last part will capture the limitations of this research and recommendations for further research.

2. Background information

This part of the paper presents some background information. This information is important for understanding the research and the outcomes. First an overview of the literature is presented. From there the history and the change in the Dutch law of compulsory education are explained. Further this will be followed by a timeline of action taken by the government to reduce the number of early school leavers. The last part consists of an overview of the compulsory education in Europe.

2.1 Literature overview

There have been a lot of case studies conducted to measure the effect of compulsory education. The main focus of those studies was to evaluate policy changes or important decisions. In this part of the research several papers will be discussed to have a wide overview of the literature. They are sorted in three parts. The first part will give an overview of the effect of early school living followed by the explanatory factors of early school leaving. In the last part the focus is on the effect of compulsory education.

2.1.1 The importance of schooling

In order to do this research it is necessary to look at the importance of going to school for students. The most favorable researches were in order to measure this by looking at the return of schooling. They measured the effect of schooling on future earnings or the effect of schooling on happiness. The main findings were that schooling is an important investment for someone's future. Although money is not everything they found that people with schooling have a healthier life-style (SH, DJ, & J, 1980) and a lower mortality rate (Groot & Maassen van den Brink, 2007). Also students with a basic qualification tend to see themselves happier and healthier. Schooling also affects the degree which one enjoys work and the likelihood of being unemployed. It leads individuals to make better decisions about health, marriage, and parenting. Schooling also improves patience, making individuals more goal-oriented and less likely to engage in risky behavior. It improves trust and social interaction, and may offer substantial consumption value to some students (Oreopoulos, How large are returns to schooling? : hint : money isn't everything , 2009).

Not only the importance of schooling itself is important but also looking at the importance of achieving a basic qualification is crucial. Statistics shows that people with only a basic qualification have almost twice the income of someone without a basic qualification. Throughout the years achieving a basic qualification became more and more important. Without a basic qualification the opportunities on the labor market are small and the chance of earning only the youth income becomes larger.[1]

The early school leavers do not only affect themselves but also society. When the number of early school leavers increase the societies costs will increase due to a bigger chance of potential crime activities (Lochner & Moretti, 2003). Also the economic growth will be strained because the labor force is less qualified (Hanushek & Wößmann, 2007). There are tax consequences because of lower tax revenue, higher costs on unemployment benefits and higher health care costs (Belfield, 2008).

[1] Appendix A - I

2.1.2 Explanatory factors

There are factors that might explain early school leaving. The importance of measuring those factors lies in policy making. Policymakers may have benefits if they know what is causing early school leaving. Knowing the reason behind it gives policymakers the opportunity to match their policy to the problem. Several papers suggest that gender has no significant influence or a small influence on early school leaving. The main reason male students leave school early is to work or because they see no benefits of schooling. Also the change in crime activities followed by early school leaving is considerably higher for male students. Female students leave school early due to family situations or lack of motivation (Stearns & Glennie, 2006). The factor, age, is also important to account for. They observe that, more than motivation, the first year of secondary education is crucial in shaping the dropping-out decision. In the first year of education the average age is 12 (Witte & Rogge, 2011). Due to policy the dropout rate between the ages of 18 and 24 is considerable higher than between the ages 12 and 16 (Commission, Early school leaving in Europe, 2012). The influence of parents is an important part of the decision of leaving school early. Research shows that parents with a higher education level have more influence on the education of their children. Also the children of parents with a higher education level also tend to have a higher level of education (McNeal, 1999). Social class and welfare has an influence on school leaving. In some cultures education it is more important than in other cultures. Immigrated students have a higher chance of leaving school early than native students in several countries (Byrne & Smyth, 2010). One research measured the effect of student mobility on the academic achievement of students. They found that students who move between schools often have more social, psychological, and academic problems than their classmates whose enrollment is stable. Those problems can be the cause of leaving school without a basic qualification (Knox, 2011). The quality of the school, class size and construction of the school has also a significant influence on the decision of early school leaving. Research suggests that schools with larger classes have respectively more dropouts than schools with small classes. This can be the case of students that need more attention (Rumberger, 1995).

2.1.3 Compulsory education

Compulsory education refers to a period of educational attendance required of all students. The period of compulsory education is often determined by the age of students. The working of the compulsory education is imposed by law (USLegal, 2015). The importance of education is explained in section 3.1.2. In order to guarantee this importance all countries have an education law. A part of this education law is the law of compulsory education. The question is whether compulsory education works. Research showed that compulsory education substantially increased adult income and substantially decreased the likelihood of being below the low-income cut-off unemployed, and in a manual occupation. These findings suggest significant gains from the law of compulsory education (Oreopoulos, The Compelling Effects of Compulsory Schooling, 2004). A research in the United Kingdom showed the effect of compulsory education. The change to the minimum school-leaving age in the United Kingdom from 14 to 15 had a powerful and immediate effect that redirected almost half the population of 14-year-olds in the mid-twentieth century to stay in school for one more year. One more year of education has one more year of returning (OREOPOULOS, 2006). The law of compulsory education has a significant effect on future incomes. Research shows that students compelled to take an extra year of schooling experienced an average increase of 10 to 14 percent in wages. They also find significant gains from education to health measures, employment and poverty status. In order to measure this they used the changes in the law of compulsory education (Oreopoulos', 2003). Another research measured the difference in the effect of a change in the law of compulsory education on the mortality rate of a male and female. They found that men benefit from

compulsory education both in the shorter and longer run. In contrast, compulsory schooling reforms have little or no effect on mortality for women. They also found that compulsory schooling reforms increased educational attainment by around 4 months for both men and women (Gathmann, Jürges, & Reinhold, 2012). One research used data from 12 European countries and the variation across countries and over time in the changes of minimum school leaving age, they studied the effects of the quantity of education on the distribution of earnings. They found that compulsory school reforms significantly affect educational attainment, especially among individuals belonging to the lowest quintiles of the distribution of ability. There is also evidence that additional education reduces conditional wage inequality, and that education and ability are substitutes in the earnings function (Brunello, Fort, & Weber, 2012).

2.2 The Dutch law of compulsory education

In the 8th century BC introduced Lycurgus in Sparta a school system. In this system the children had to go to school even when the parents did not want this. The great philosopher, Plato, writes about this in one of his dialogues. He wrote about the assent of the law. In a couple regions of the Republic of the Seven United Netherlands there was since the middle of the 17th century an indirect compulsory education system due to levying of school fees. Everyone needed to pay these fees whether their children were going to school or not (Storimans, 2010).

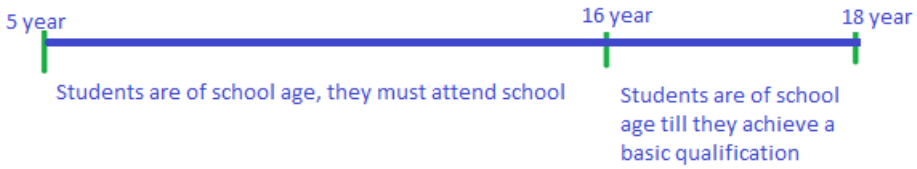
The first legislation on compulsory education was introduced on July 7th of 1900. The main driver behind the advent of the law on compulsory education was done in the footsteps of Van Houten. He was the first man that came out publicly against child labour. In order to prevent child labour they designed the law of compulsory education. The law said: *"Parents that make their children work on the land or in the factory are guilty. They are the cause of neglect of their children because they take the right of learning away. "*In that time there was no one that checked whether children did go to school or just worked at home (Storimans, 2010).

In 1969, almost 70 years later, the first law on compulsory education changed. They changed it into the current education law, the law of compulsory education 1969. This law said: Everyone between the age of 4 and 12 needs to have education. With this law there was the introduction of the local attendance officer. This officer was responsible for law enforcement. The officer had a commission and together they investigated violations of the compulsory education law 1969. They were also responsible for motivating students that violated the law (Storimans, 2010).

The next step in the law of compulsory education was the introduction of the partial compulsory education in 1971. The partial compulsory education is one year at the end of the full-time education period where the student could choose whether he wants to work or want to study. The student also could choose to work three days in the week and go to school for two days in the week. Since 1979 everyone in the Netherlands had a schooling age of 16. Till the 16th birthday of the student they needed to be held on the law of partial compulsory education (Storimans, 2010).

From 1969 till 2006 the law of compulsory education changed a little bit at a time but the fundamental law never changed. But in 2007 the government decided to change this. From August 2007 all children between the ages of 5 and 16 living in the Netherlands are of school age. Parents or guardians of young people between the 6 and 16 years old must enroll their children in school. According to the law they must attend school. Students without a basic qualification must follow education till they turn 18. A basic qualification in the Netherlands is a pre-university diploma, HAVO diploma or MBO diploma at level 2 or higher. The basic qualification is one of the measures announced by the government to counter school dropout among the youth. The new law should increase the chances of young starters on the labor

market. The students must fulfill the qualification requirement by full-time education (Rijksoverheid, 2007).



2.3 Timeline

Since 1990 it became clear that early school leaving was a big issue. There was no favorable system that measured the number of early school leavers. Not only that, the control over the early school leavers was weak and unorganized. In order to solve this problem the Dutch government introduced in 1994 the RMC function. The RMC function was responsible for designing a good system where early school leavers were traced and registered. The RMC legislation implies that each school is responsible for indicating the number of early school leavers within a month. This law had to make sure that a decrease in the number of early school leavers was insured (OCW, 1994).

In 2000 the European Commission came with the Lisbon strategy. This strategy was a long term strategy that had to make the economy of the European Union the strongest economy of the world. The name Lisbon was allied with the place where the agreement happened between the European countries. In this agreement the Dutch government needed to make sure that the number of early school leavers between the age of 18 and 24 years decreased with 21000 in 2006 and with 35000 in 2010. Due to the Lisbon strategy the Dutch government put a target on the number of early school leavers. The number needed to decrease with 30 percent relative to 2002 before 2006. That means that the number had to decrease with more than 21000 (EuropaNU, 2012). To make sure the number of early school leavers decreased, the Dutch government introduced in 2004 operation JONG. This operation had the target to motivate young people to achieve a basic qualification. The Dutch government invested 22 million euro's in this project for 10 year. In cooperation with the RMC law the operation JONG succeed. The number of early school leavers decreased (Ministerie-OCW, 2007)[2].

Although the number of early school leavers decreased the Dutch government was not satisfied. They decided to change the law of compulsory education 1969. In august 2007 the change was a fact. The target of the new law was to have more students achieve basic qualifications. The government wanted to have more young people under the age of 23 years with a basic qualification (Rijksoverheid, 2007). Since 2007, the budget to discourage early school leaving raised up to 80 million per year (CPB, 2006). The effect of the change in the new law is experienced as positive. The number of early school leavers continued to decline and the number of achieved basic qualification between the age of 18 and 23 also continued to increase[3].

In 2012 the Dutch government decided to come up with a plan for 2012 till 2015. They made agreements about control, registration and regulations of early school leavers. The government invested 110 million euro each year. All the agreements are included in the performance covenants (Rijksoverheid-OCW, 2013). For 2016 the government has a target to reduce the number of early school leavers between the age 18 and 24 years to 25.000.

[2] Appendix A - II
[3] Appendix A - III

2.4 Compulsory education in Europe

In Europe, education systems are deeply rooted in national traditions and are characterized by specific national features. Today, economic, social, and cultural change strengthens the need for policy makers, business leaders, and scholars to learn more about the characteristics of national education systems (Hörner & Döbert, 2012). Education is an important investment for a valuable and stable economy. Every European country has their own education system. They differ in many ways. Some countries chose for an 8 year term of primary education while other countries chose for a 6 year term. The schooling age also differs for each country. The most countries use a school age from 5 till 16 years. Some countries like Belgium use a school age from 6 till 18 years. Only the Netherlands uses a basic qualification as requirement for not being school aged between the age of 16 and 18 (Commission-, 2014). The problem of early school leavers is not only a problem in the Netherlands. The problem occurs in every country. The European Commission made an agreement in 2000 to reduce the number of early school leavers in all European countries. Not all countries succeeded. A lot of countries still had a high percentage of early school leavers. In 2010 the European Commission made another agreement to reduce the number of early school leavers. For the future the European Commission introduced a project named European Youth in 2015. In this project the European Commission has an extra budget for countries to invest in their education system (Commission--, 2015). Spain had the highest percentage of early school leavers in 2013, followed by Malta. Slovenia has the lowest percentage with only 3.9 percent, followed by Czech Republic with 5.4 percent in 2013. Half of the European countries have achieved an early school leaving percentage below ten percent since 2010. The target for the European Commission is to have a percentage below ten percent in all European countries (Commission, Early school leaving in Europe, 2012).

2.5 New Research

There have been a lot of case studies conducted in the context of compulsory education. Previous studies showed what the effect of compulsory education is on individual and national level. The importance of schooling is also investigated in a lot of studies. Although there have been studies on the effect of a change in the compulsory education system in other countries, a clear evaluation of the change in the compulsory education in the Netherlands is missing. In 2007 the Dutch government presented a report about the number of early school leavers and whether the measures to prevent early school leaving works. This study will investigate whether the change in law of compulsory education had a significant positive effect on the number of early school leavers. Because the trend in the number of early school leaver in the Netherlands shows a decrease since 2004 it is interesting the know whether the further decrease in the number of early school leavers is caused due to the change in the law of compulsory education or due to earlier actions taken by the government. Previous studies never studied the real effect of the change in the law of compulsory education in the Netherlands.

3. Data

In this part of the paper the descent of the data is carefully explained. To conduct a good study it is important to have matching datasets. In order to conduct the first part of this study, data from Eurostat is used. Eurostat is a Directorate-General of the European Commission located in Luxemburg. Its main responsibilities are to provide statistical information to the institutions of the European Union. With the data from Eurostat the average treatment effect will be measured using the research method, difference-in-difference. With this research method the effect of a treatment on an outcome is calculated by comparing the average change over time in the outcome variable for the treatment group to the average change over time for the control group. For this study the Netherlands is the treatment group because they changed the law of compulsory education. The control groups are France, Latvia and Estonia. The reason for using those countries as control groups is as follows. The law in all four countries before the change is the same and they have almost the same trend in the number of early school leavers. The data coming from Eurostat consist of a percentage of early school leavers between the age of 18 and 24 years. For this research only data of the Netherlands, Latvia, France and Estonia is used. The data is from 2001 till 2012. Eurostat measures the number of early school leavers in cooperation with the European countries. They are together responsible for a clear view on the number of early school leavers.

To conduct the second part of this research the data from the Central Agency of Statistics used. The Central Agency of Statistic is a Dutch governmental institution that gathers statistical information about the Netherlands. They don't only provide macro statistics but also micro statistics based on surveys. To measure the effect of the change in the law of compulsory education on individual level a micro level database is used called POLS. This dataset is based on annual questionnaires filled in by random people in the Netherlands. The questionnaire consists of different parts that measure the living conditions. The questionnaires are held in the year from 2003 till 2009. After 2009 the Central Agency of Statistics stopped with this questionnaire. From the data some variables like age, gender, household composition, paid work, self-reported health and self-reported happiness are used. Those variables are used to measure the change in the effect of those variables on the chance of achieving a basic qualification due to the change in the law of compulsory education. The variable completed education level gave an indication whether someone achieved a basic qualification. Everyone between the age of 16 and 18 and between the age of 18 and 24 is filtered from the dataset and used.

Both resources are a part of government agencies. They have a structured method of collecting data and process them into useful datasets. Both provide their information and finding to large companies and governmental agencies. Important decision makers make use of data coming from both resources. A number of policies are based on the data and findings provided by Eurostat and The Central Agency of Statistics.

4. Methodology

This part of the paper will focus on the methodology of the conducted research. There will be careful explanation of how the data is used and which research methods are applied. To give a clear explanation of the methodology this part is divided in three subparts. The first subpart consists of an explanation of the experimental design. This is followed by steps of data sorting. The last part consists of an explanation of the robustness checks.

4.1 Experimental Design

To give an answer to the research question this study is conducted on two levels. To measure whether the change in the law of compulsory education in order to reduce the number of early school leavers worked the research method, difference-in-difference is used. With this method the average treatment effect can be measured. In order to do this the research will focus on the percentage of early school leavers between the age of 18 and 24. Next to the national level of this research there is a study conducted on individual level. The effect of factors on achieving a basic qualification is measured on this level. The outcomes before the change and after the change are compared with each other. With this part an eventual change in the influence of the factors due to policy change is measured. This analysis is done with a linear regression model. All analyses are conducted with a statistical program named SPSS. After those analyses, the outcomes will be checked with robustness checks. Those checks will give a clear view of whether the outcomes are significant right and not based on other factors. Using robustness checks it will be clear whether a causal effect is measured.

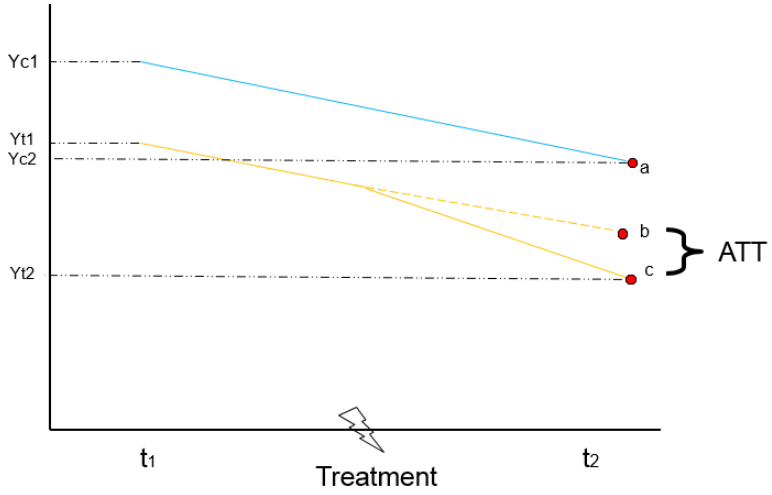
4.1.1 Research Methods

This part will explain the two research methods in more dept. While explaining the research methods the variables are also presented with the models. Difference-in-differences will be first explained followed by an explanation of a linear regression analysis.

4.1.1.1 Difference-in- Differences

To measure the average treatment effect on the treatment the research method difference-in-difference is used. With this research method the effect of a treatment on an outcome is calculated by comparing the average change over time in the outcome variable for the treatment group to the average change over time for the control group. This is a popular research method in applied research work nowadays. Difference-in-difference works well when evaluating policies or important decisions. It attempts to mimic a random assignment with a treatment group and comparison groups. To apply a difference-in-difference you need one treatment group. This group experiences a treatment. There is also a minimum of one control group needed. This group(s) does not experience the treatment. To make sure the outcomes can be compared data is needed before and after the treatment (Kippersluis, 2014). For this study The Netherlands is the treatment group. The Netherlands experienced the treatment, the change in the law of compulsory education. This research uses three control groups to increase the validity of the study. The control groups are France, Latvia and Estonia. It is important to point out that difference-in-differences can measure the effect of the treatment at one point. It will be unclear whether the effect is caused due to earlier treatments.

The key concept of difference-in-difference can be best explained using the following graph. The graph is based on fictional numbers to make the explanation simple and easy to understand. The y axis represents the number of early school leavers in percentages. The years are placed on the x axis. There are two countries used for this example. The orange line represents The Netherlands, the treated country. The blue line represents the control group, France. To make the explanation clearer, we assume that there are only two points in time. The first point, t_1 is before the treatment and t_2 is after the treatment. Difference-in-difference has one important assumption named the **parallel trend assumption**. This assumption implies that in absence of the treatment, the treatment group and control group have the same trend over time. The development in the number of school leavers needs to be almost equal. With two points in time it is impossible to check whether the parallel trend assumption holds. In order to tackle that problem this research makes use of nine time points (Kippersluis, 2014). Yc1 indicates the number of early school leavers in percentage for France before the treatment, the same goes for The Netherlands for Yt1. Yc2 stands for the number of early school leavers after the treatment for France. The same goes for The Netherlands for Yt2. The difference between the two countries, Yc1 – Yt1, is not a problem because difference-in-difference controls for that. In order to calculate the double differences a counterfactual is needed. The counterfactual is indicated by the dotted orange line. The counterfactual indicates what would happen to the treatment group if it was not treated. Due to the parallel trend assumption it is clear that line must have a direction like the line of the untreated group. The standard difference estimator is equal to A - C. The counterfactual 'normal' difference is calculated by B - A. Now the difference-in-difference can be estimated by C – B. This is called the average treatment effect. The following table will show how the difference-in-difference is calculated (Evans, 2008).



	Before treatment	After treatment	Difference
The Netherlands (treatment group)	Yt1	Yt2	ΔYt = Yt2 – Yt1
France (Control groups	Yc1	Yc2	Δ Yc = Yc2 – Yc1
Difference			$\Delta\Delta$ Y = ΔYt - ΔYc

The example above is based on two groups and two points in time. This research has one treatment group, The Netherlands, and three control groups, France, Latvia and Estonia. To hold on the parallel trend assumption, the research sample consists of data from the years 2001 to 2012. So this research is based on eleven points in time.

To measure the average treatment effect on the treated the following regression is used.

$$Y_{it} = \alpha + \rho * Countrydummy + \gamma_t + \beta * Treatmentdummy + \epsilon_{it}$$

Equation 1

The outcome of the regression, Y_{it} , is the number of early school leavers in percentage. The first variable, α, is the number of early school leavers in percentage for the control group(s). It is important to note that in this part of the research it is assumed that the control groups have the same starting point. Therefore the country dummy is equal to 1 if the concerned country is treated. So for this study the country dummy equals 1 for the Netherland and equals 0 for all other control countries. This dummy controls for the initial differences between the two groups, treatment group and control group, before the treatment. There will also be a control for the differences between the control groups with a robustness check. The variable γ_t stands for the time fixed effects. It controls for factors that might change over time. The treatment dummy equals 1 if the treatment occurs in a year. In this study it is equal to 1 for the Netherlands in the period between 2007 and 2012 because only than the treatment, change in the law of compulsory education, occurred. The average treatment effect on the treated is equal to the β in the first equation. If the β is positive for this study it implies that the treatment had a negative effect on the number of early school leavers. So this implies that the change in the law of compulsory education increased the number of early school leavers. If β is negative it implies that the treatment had a positive effect on the number of early school leavers, the number of early school leavers decreased. The difference-in-difference is conducted with a univariate general linear model.

4.1.1.2 Multiple linear regression analysis

A multiple linear regression attempts to measure the relationship between variables by fitting a linear equation to the observed data. It models the relationship between one outcome variable and multiple explanatory variables. The outcome in a linear regression is called the dependent variable. The factors influencing the outcome are called independent variables or explanatory variables (Moore, McCabe, Alwan, Craig, & Duckworth, 2011). Formally the model for a multiple linear regression, given n observations is:

$$Yi = \beta 0 + \beta 1 * X1 + \beta 2 * X2 + \beta 3 * X3 + \beta 4 * X4 + \cdots + \beta n * Xn$$

Equation 2

The correlation between the factors and the outcome variable can be measured on the basis of equation 2. There are important assumptions that need to hold for the multiple linear regression analysis. If these assumptions do not hold the outcomes of the model could be biased. The first assumption implies that the relationship between the independent and dependent variables need to be linear. It is also important to check for outliers since multiple linear regressions are sensitive to outlier effects. The linearity assumption can best be tested with a scatter plot (Moore, McCabe, Alwan, Craig, & Duckworth, 2011). The second assumption implies that a multiple linear regression assumes that there is little or no multicollinearity in the data. Multicollinearity occurs when the independent variables are not independent form each other. This will be checked with a correlation matrix. In this matrix the Pearson Bivariate Correlation among all independent variables the correlation coefficient must be smaller than 0.08. The third assumption checks for exogeneity in the model. This implies that factor may not influence the relationship between the independent variable and the dependent variable. When this is not the case there will be a lack of validity in the outcomes of the regression models. Having a third factor influencing the relationship between the independent variable and dependent variable will cause a non-causal relationship. This assumption can be checked with a correlation test. The

independent variable and the error term should not be correlated with each other. Also the mean of the error term, E[ε], must be almost equal to zero (Sloutions, 2015).

After checking the assumptions of a linear regression model, the regressions can be conducted. This analysis is divided in six parts. The first analysis measures the influence of factors on achieving a basic qualification for the group with an age between 16 and 18. The first analysis is based on the sample from 2003 till 2009. The second analysis measures the influence of factors on achieving a basic qualification for the group with an age between 16 and 18 years but now only before the change in the law of compulsory education, so 2003 till 2006. The third analysis does the same but after the change in the law of compulsory education, so 2007 till 2009. The fourth analysis measures the influences of factors on achieving a basic qualification for the group with an age between 18 and 24 years. This analysis is based on a sample from 2003 till 2009. The fifth analysis measure the influence of the factors on achieving a basic qualification for the same group but now before the change in the law of compulsory education. The last research does the same but with a research sample after the change, so a research sample from 2007 till 2009. For all six analyzes the following regression is used:

$$Yi = \alpha + \beta 1i * Gender + \beta 2i * household\ composition + \beta 3i * paid\ work + \beta 4i * self\ reported\ health + \beta 5i * self\ reported\ happiness + \varepsilon i$$

Equation 3

The outcome of this regression is defined as a dummy variable. It equals 1 when someone achieved a basic qualification and it equals 0 when someone did not. With this regression the relationship between the explanatory factors and the outcome variable can be measured. All the explanatory variables were included in the questionnaire from 2003 till 2009 in the same design. Only the variables, self-reported health and self-reported happiness were not included in the POLS questionnaire of 2009. The betas in the equation represent the strength of the relationship between the factor and the outcome variable. When the beta equals a positive value it means that the factor and the outcome variable have a positive relationship. The factors increase in that case the chance of achieving a basic qualification. When the beta equals a negative value it implies that the relationship between the factor and the outcome variables is negative. In that case the factor will decrease the chance of achieving a basic qualification (Moore, McCabe, Alwan, Craig, & Duckworth, 2011).

4.1.2 Significance

The outcome of the analysis will come with a p-value. The significance level in this research is 0.05 (5%). This implies that repetition of the research will give the same results in 95% of the cases. When the p-value is greater than 0.05 the outcome is seen as not significant. This implies that there is no statistical evidence that the difference in the groups is not due to the chance or factors (Moore, McCabe, Alwan, Craig, & Duckworth, 2011). When the p-value is below the 0.05 it means that the outcome is significant.

4.2 Data Sorting

Most researches are based on large datasets with a lot of information. Not all information are useful for any research. To make sure you have the right information it is important to sort the datasets carefully. The first step is finding the right data. To make sure that the right research method could be conducted the number of early school leavers for four European countries had to be in one file. This research only focuses on students between the age of 18 and 24 years. To measure which factors might affect achieving a basic qualification, the POLS data is used. The research sample is large across the years and there are also a large number of questions. In total there are 132 variables included in the datasets. Only a small part the dataset is used for this

research. Also only people between the age of 16 and 18 years and between the age of 18 and 24 years are used. Sorting the data was an important part in conducting this study.

4.3 Robustness checks

The results of the analyses will be tested on robustness by performing robustness checks; this is done in order to test the outcomes on reliability. It has been assumed that the control groups have the same starting point on the number of early school leavers. With a robustness check will be checked whether the outcomes of the difference-in-differences are the same when the control groups do not have the same starting point on the number of early school leavers. In general difference-in-differences the country dummy is equal to 1 if the concerned country is treated. So for this study the country dummy equals 1 for the Netherlands and equals 0 for all other control countries. In the robustness check the county dummy will have a value of 1 for the Netherlands, 2 for France, 3 for Latvia and 4 for Estonia. This check controls for the initial differences ***between*** the control groups. The second check will check the validity of the outcomes of the difference-in-difference analyses with a robustness check named placebo treatment. This check implies that there will be assumed that the treatment was not in 2007 but one year earlier in 2006. If the outcome is not significant it means that the real effect of the change in the law of compulsory education in 2007 on the number of early school leavers is measured. If the outcome for 2006 is still significant it means that other developments had an effect on the reduction of early school leavers. In that case it will be assumed that the treatment was in 2005 and not in 2007. If the outcome is still significant the treatment year will be 2004 (Gertler, Martinez, Premand, Rawlings, & Vermeersch, 2010).

The outcomes of the multiple linear regressions will be checked on the basis of a binary logistic regression and a univariate general linear model. The outcome of the binary logistic regression is binary, it is 1 if someone achieved a basic qualification and it equals 0 when someone did not achieve a basic qualification. A binary logistic regression is used to predict a binary response based on one or more predictor variables. This regression deals with a situation where the outcome variable, achieving a basic qualification, can have only two possible types. In this case it is 0 or 1. It predicts the probability that an observation falls into one of those types of the outcome variable (Laerd, 2005). If the outcomes of the binary logistic regression are the same as the outcomes of the multiple linear regression it can be assumed that the model is robust and the outcomes are reliable (Moore, McCabe, Alwan, Craig, & Duckworth, 2011). The univariate general linear model will check the differences within the variables. For the variables self-reported happiness and self-reported health, people had to choose between five possible numbers. When someone chose 1 it implied that the health or happiness is very bad. When someone chose 5 it implied that the health or happiness was very good. There was an order between the 1 and 5 of the measurement of the health and happiness. In this robustness check it will be assumed that those two variables are not covariates but fixed effects. From this robustness check it can be concluded whether the factor has a significant effect on the change of achieving a basic qualification.

5. Results

The results of the analyses will be presented in this part. The results are divided in two parts. The first part will give an overview of the results of the difference-in-differences analyzes. The second part will focus on the results of the multiple regression analyses.

5.1 Difference-in-differences

The Netherlands was the treatment group with France, Latvia and Estonia as control groups in the first analysis. The trends in the lines are not perfectly parallel but those three countries, of all European countries with the same law before the treatment as the law of the Netherlands, came the closest to the parallel trend assumption[4]. After conducting the difference-in-differences it can be concluded that the change in the law of compulsory education had a significant positive effect on the number of early school leavers. From the analysis with the Netherlands as treatment group and France, Latvia and Estonia as control groups the average treatment effect is -2.4%. This outcome implies that due to the change in compulsory education the number of early school leavers in the Netherlands decreased with 2.4%. Because the p-value of the outcome is smaller than the significant level, $0.014<0.05$, it can be assumed that the outcome of the first analysis is significant[5]. From now on the countries are individually compared in the analysis. In the second analysis France was the control group and the Netherlands the treatment group. The two countries have almost the same trend in the number of early school leavers[6]. From the results it can be concluded that the average treatment effect is -3.25%. The outcome implies that if France is the only control group the change in compulsory education has a positive effect on the number of early school leavers in the Netherlands. The outcome is significant and therefore it can be assumed that the treatment decreased the number of school leavers with 3.25%[7]. When Latvia is the control group and the Netherlands the treatment group the outcome, -0.983%, is not significant[8]. The p-value of 0.097 is greater than the significant level of 0.05. Although the two countries have a similar trend in the number of early school leavers the outcome of the test is small and not significant[9]. Estonia was the control group in the last analysis. Both countries have not quite the same trend from 2001 till 2006. From 2006 the trend is more parallel[10]. The average treatment effect is -2.97% if Estonia is the only control group. The outcome is significant and that implies that the change in the law of compulsory education has a positive influence on the number of early school leavers in the Netherlands[11]. From all this analysis it can be concluded that the change in the law of compulsory education had a significant positive influence on the number of early school leavers in the Netherlands.

5.2 Multiple regression analysis

Before presenting the results of the multiple regression analysis the assumptions of a multiple linear regression analysis are checked. The assumptions are first checked for the data of the group with an age between 16 and 18. The first assumption implies that the relationship between the independent and dependent variables needs to be linear. This assumption is checked with a scatterplot. From the plot can be concluded that the assumption of linearity is

[4] Appendix B - I
[5] Appendix B - II
[6] Appendix C - I
[7] Appendix C - II
[8] Appendix D - I
[9] Appendix D- II
[10] Appendix E - I
[11] Appendix E - II

met[12]. The second assumption implies that in a multiple linear regression there is little or no multicollinearity in the data. This is checked with a correlation matrix. From the table it can be concluded that there is a little multicollinearity in the data. This is for the variables self-reported happiness and self-reported health[13]. The last assumption checks for exogeneity in the model. This implies that factors may not influence the relationship between the independent variable and the dependent variable. When this is the case there will be a lack of validity in the outcomes of the regression models. Having a third factor influencing the relationship between the independent variable and dependent variable will cause a non-causal relationship. This assumption can be checked with a correlation test. The independent variable and the error term should not be correlated with each other. Also the mean of the error term, $E[\varepsilon]$, must be almost equal to zero. Because this cannot be tested this assumption will be discussed in the interpretation part of this paper.

The assumptions for the multiple linear regression model are also checked for the data with people between the age of 18 and 24 years. The first assumption of linearity is met[14]. The second assumption[15] of the multiple linear regression model are also met. The last assumption of exogeneity will be discussed in the interpretation part of this paper.

Now that the assumptions of a multiple linear regression model are checked the results of the analyses can be presented. The multiple regression analysis is conducted six times, so the results are divided in six parts. The first part investigated the effect of the explanatory factors on the chance of achieving a basic qualification for the group with an age between 16 and 18 years. The research sample for the first analysis consists of data from 2003 and 2009. From the results it can be concluded that four factors, gender, household composition, self-reported happiness and self-reported health do not influence the chance of achieving a basic qualification. The factor paid work does influence the chance of achieving a basic qualification. This implies that students that have paid work, more than 12 hours in a week, have a lower chance of achieving a basic qualification[16]. If only the years before the change are included in the data the results shows that only paid work have a significant influence on the chance of achieving a basic qualification. Paid work decreases the chance of achieving a basic qualification[17]. This result does not hold if only the years after the change are included in the data. None of the factors have a significant influence on the change of achieving a basic qualification[18].

The above presented results were of the group with people between the age of 16 and 18 years. Now the results of the same regression model are presented but for the group with people between the age of 18 and 24 years are presented. If the research sample consists of the years 2003 till 2009 the results shows that all explanatory factors have a significant influence on the chance of achieving a basic qualification. Gender has a small positive influence on the chance of achieving a basic qualification. All the other factors have a negative influence on the chance of achieving a basic qualification. The factor self-reported happiness has the highest negative coefficient. This implies that the lower the self-reported happiness the lower the chance of achieving a basic qualification. The same goes for self-reported health. Paid work also has a negative influence on achieving a basic qualification. The more a student works, the less the chance of achieving a basic qualification. Household compensation has a really small negative

[12] Appendix F - I
[13] Appendix F - II
[14] Appendix G - I
[15] Appendix G II
[16] Appendix H - I
[17] Appendix H - II
[18] Appendix H - III

influence on achieving a basic qualification. This implies that the more members in a household the lower the chance[19]. When the years before the change in the law of compulsory education are included in the data the results shows that all factors have a significant influence on the chance of achieving a basic qualification. The factors have almost the same relationship with the outcome variable as with all years included[20]. When only the years after the change in the law of compulsory education are included in the data the results are different than the other analyses for this group. Gender has a larger influence on the chance of achieving a basic qualification. Self-reported happiness has no significant influence anymore. Self-reported health has a small negative effect on achieving a basic qualification. This is the same case for the variables paid work and household composition. When the outcomes of the regression after the change are compared with the outcomes of the regression before the change it can be concluded that the influences of the factor on the change of achieving a basic qualification has decreased.

5.3 Robustness checks

The results of the robustness checks are presented below. The first part will check whether the outcomes from the difference-in-difference are reliable with a placebo treatment. The second part will check whether the outcomes from the regression are reliable with a binary logistic regression.

5.3.1 Differences in control groups

In the general analyses it was assumed that the control group has the same starting point of the number of early school leavers. Because this was only an assumption this robustness check will check whether the results change after this assumption does not apply. For each control group there is a different dummy. Due to the different dummies the difference-in-difference analyses controls for the differences **between** the control groups. From the results of this robustness check it can be concluded that change in the law of compulsory education still has a significant positive influence on the number of early school leavers[21]. This implies that the differences in the control groups do not have an impact on the outcome of the analyses.

5.3.2 Placebo treatment

To know whether the results of the difference-in-difference are reliable a placebo treatment is introduced in the dataset. It is assumed that the treatment began in 2006 and not in 2007. In the regression the treatment has dummy a value of 1 for 2006 till 2012. The research sample consists of the Netherlands as treatment group and France, Latvia and Estonia as control groups. The results show that the outcome of -2.4 is still significant. This implies that not only the change in the law of compulsory education is responsible for a reduction in the number early school leavers[22]. Suppose that the treatment started in 2005 and not in 2007 than the results implies that the outcome of the difference-in-difference is still significant. Not only in 2006 but also in 2005 there were factors that caused a decrease in the number of early school leavers[23]. To know whether this also hold in 2004 it is assumed that the treatment began in 2004 and not in 2007. The results of this test are not significant[24].

[19] Appendix I - I
[20] Appendix I - II
[21] Appendix L - I
[22] Appendix J - I
[23] Appendix J - II
[24] Appendix J - III

5.3.3 Binary Logistic regression

A binary logistic regression is used to check whether the outcomes of the linear regression are robust. The robustness check is conducted two times. The first time it is conducted with a sample of people with an age between 16 and 18 years. The second time it is conducted with a sample of people with an age between 18 and 24 years. The results for the first analysis show that the outcome is the same as for the linear regression. Only paid work has a significant influence on the chance of achieving a basic qualification. The strength of the relationship between paid work and the outcome variable has decreased. The influence became greater. A student that works has a lower chance of achieving a basic qualification[25]. For the group with an age between the 18 and 24 years the results are also the same as those from the linear regression. All factors have a significant influence on achieving a basic qualification. The relationship between the factors and the outcome variable are the same[26]. The significance level of the outcomes is almost the same as the outcome of the linear regression model.

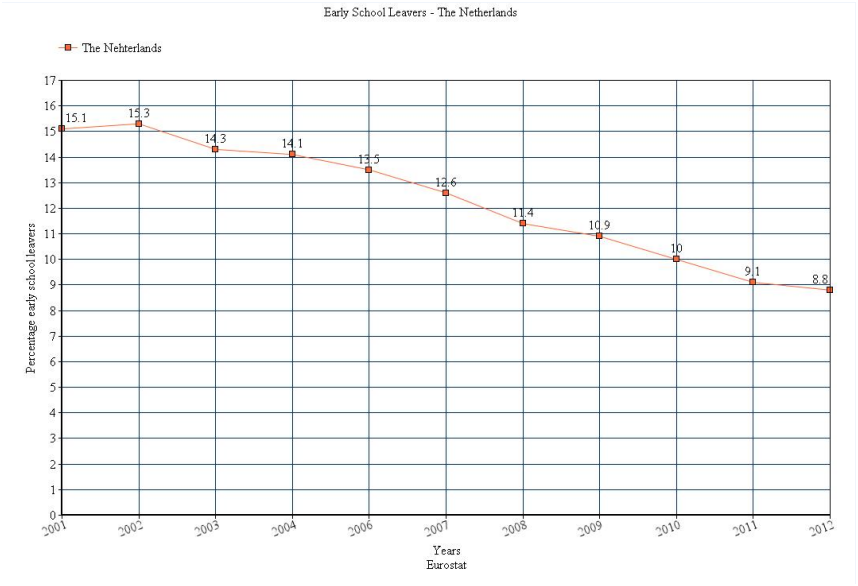
5.3.4 Differences within variables

Within the variables self-reported health and self-reported happiness there can be differences in the degree that influences the change of achieving a basic qualification. The difference between good health and really good health is smaller than the difference between bad health and really bad health. The same goes for self-reported happiness. To check whether the differences within variables influence the outcome of the linear regression model a robustness check has been conducted. From the results of the robustness check there can be concluded that for the group with an age between 16 and 18 years, self-reported happiness and self-reported happiness have no significant influence on the chance of achieving a basic qualification. The variables gender and household composition also do not have an influence on the chance of achieving a basic qualification[27]. For the group with an age between the 18 and 24 years to following can be concluded. The factor self-reported happiness has a significant influence on the chance of achieving a basic qualification. Meaning that the happier the student, the higher the chance of achieving a basic qualification. This does not apply for the factor self-reported health. This factor does not have a significant influence on the chance of achieving a basic qualification[28].

[25] Appendix K - I
[26] Appendix K - II
[27] Appendix M - I
[28] Appendix M- II

6. Interpretation

From the results of the difference-in-difference analysis it can be concluded that the change in the law of compulsory education did not have a specific influence on the number of early school leavers in the Netherlands. The robustness checks showed that earlier placebo treatments had also a significant influence on the number of early school leavers. This implies that not only the change in the law of compulsory education has a positive influence. There are other factors that might have caused a decrease in the number of early school leavers. Looking to the development of the number of early school leavers in the Netherlands it can be concluded that it decreased since 2004 with more than 6%. The actions taken by the Dutch government in order to decrease the number of early school leavers are explained in section 3.3. From this timeline it is clear that the government invested a lot to make sure that the number of early school leavers decreased over the years. The main motivation to invest in preventing early school leaving was due to the rules of the European Commission. In 2000 they decided that the number of early school leavers needed to decrease in all European countries in 2006. In order to achieve this target the Dutch government invested in programs and projects to decrease the number of early school leavers. In 2004 they introduced the project JONG. This might explain the decrease in the number of early school leavers. Earlier studies, see section 3.2, showed that a change in compulsory education systems had a positive influence on the number of early school leavers. Students who had an additional year of school benefit from more income and better life conditions. The difference-in-difference in this research also showed that the change in the law of compulsory education did not have a specific influence on the number of early school leavers. The decrease in the number of early school leavers is caused by a downward trend. Other investments, programs and projects had an earlier positive effect in the change in the law of compulsory education and that caused a decrease in the number of early school leavers.



When outcomes of the linear regression are compared the following can be concluded. Factors like, gender, household composition, self-reported happiness and self-reported health do not influence the chance of achieving a basic qualification for students with an age between 16 and 18 years. They are young and have less or no responsibilities in life. Due to this they might not feel the consequences of the factors. They have fewer responsibilities in life and might not be aware of the importance of the factors. This might explain why the factors do not have a significant influence. The only factor that direct influence the chance of achieving a basic qualification for the group with an age between the 16 and 18 years is paid work. When students work more than 12 hours in a week the chance of achieving a basic qualification decreases. For the group with an age between 18 and 24 years all factors have a significant influence on the chance of achieving a basic qualification. Students in that age group stand different in life. Some students with that age have responsibilities toward themselves and others. The factors influence them significantly because those factors became more important over the years. When the years before the change in the law of compulsory education are compared to the years after the

change in the law of compulsory education the following can be concluded. For the group of with people with an age between the 16 and 18 years applies that all factors including paid work do not have a significant influence on the chance of achieving a basic qualification after the change in the law of compulsory education. This implies that the factors do not have an influence anymore. This can be a good result because due to the change students that work now have an equal chance of achieving a basic qualification. The change in the law of compulsory education has ensured that factors do not influence the chance of achieving a basic qualification anymore. For the group of with people with an age between the 18 and 24 years applies that all factors have a significant influence in the chance of achieving a basic qualification before the change. After the change in the law of compulsory education the significant influences of the factors was still the same. In section 2.1.2 explanatory factor of early school leaving are explained. This study showed that some findings are in line with the literature. Gender has a small significant influence on the chance of achieving a basic qualification for people with an age between the 18 and 24 years. Male students have a bigger chance of achieving a basic qualification in that group. This might be due to responsibility or culture. Also the dropping out rate is significantly higher in the group with an age between 18 and 24 years than in the group with an age between 16 and 18. Studies showed that environmental factors like household composition have an influence on early school leaving. This study found that the influence only occurs in the group with an age between the 18 and 24 years and is small. From the robustness checks it can be concluded that the influences of the factors for both groups remained the same as the outcomes of the linear regression model. The significance of the outcomes is equal in both analyses, linear regression model and binary logistic model. The outcome of the robustness check that measured the differences within the variables self-reported happiness and self-reported health can be concluded that only self-reported happiness has a significant influence on students with an age between the 18 and 24 years. The reason behind this finding might be that most people in that age group has a strong attitude that their happiness is causing motivation. When they are not happy their motivation of achieving a basic qualification is slight. This study found that this does not apply for self-reported health.

Although a relationship has been found between the factors and the chance of achieving a basic qualification, it is questionable whether this relationship is causal. Other factors might influence on one hand the chance of achieving a basic qualification but also on the other hand there might be other factors that influence the independent variables in the regression models. The factor paid work has a significant influence on the chance of achieving a basic qualification in both groups. When the relationship between paid work and the chance of achieving a basic qualification is measured, the assumptions of a regression model tells us that paid work can only affect the change of achieving a basic qualification not vice versa. In reality it possible that having a basic qualification increases the chance of having paid work in both groups. When school is easy for students with an age between 16 and 18 they might have a job next to their school work. Having a basic qualification will increase the chance of having paid work for people with an age between 18 and 24 because their chances on the labor market became better with a basic qualification. Due to a lack on inverse causality it is important to note that this study only found correlation between the independent variables and dependent variable.

7. Conclusion

This study measured the effect of the change in the law of compulsory education on the number of early school leavers. Early school leaving is a problem for the future of students and society. To tackle this problem the European Commission introduced an agreement in 2000 in order to decrease the number of early school leavers in all European countries. In addition to this agreement the European Commission came in 2009 with a new strategy named, Europe 2020. The seventeen European Countries made an agreement according the problem of early school leavers. This agreement implies that the number of early school leavers has to be under the 10% of the total students between the age of 18 and 24 for each country. In order to achieve the 10% the Dutch government decided in august 2007 to change the law of compulsory education. Before august 2007 there was partial compulsory education. Pupils who followed a full time education were of school age till 16. Pupils who followed a part time education were of school age till 18. After August 2007 all students have a school age till 16 years but they must attend school until they turn 18 or until they obtain a basic qualification. In the Netherlands a basic qualification is a VWO diploma, HAVO diploma or a MBO 2 diploma. The main purpose of the change in the law of compulsory education is to get more students with a basic qualification and to have less early school leavers. There has been a lot of case studies conducted on the effect of compulsory education, the effect of a change in compulsory education and why compulsory education matters. Schooling is important, not only for students themselves but also for society and economy. Compulsory education ensures that more students have a basic qualification and a better change on the labor market. On the basis of statistics from the Central Agency of Statistics and Eurostat there has been several analyses conducted to answer the following research question:

What is the effect of the change in the law of compulsory education in August 2007 on the number of early school leavers?

From several analyses it can be concluded that the change in the law of compulsory education does not have a specific influence on the number of early school leavers. The outcomes of the difference-in-differences analyses showed that not only did this change cause a decrease in the number of early school leavers. Other actions taken by the Dutch government in order to decrease the number of early school leavers has a significant positive influence on the number of early school leavers. It is important to highlight that the change in the law of compulsory education had a positive outcome but not only that change caused the decrease in the number of early school leavers. It is also important to know what factors influence the chance of achieving a basic qualification. The factors, gender, household composition, paid work, self-reported happiness and self-reported health have a significant influence on students with an age between 18 and 24 years. All factors, expect gender, are negative correlated with the change of achieving a basic qualification. The factors influence the chance of achieving a basic qualification significantly before and after the change in the law of compulsory education. The factors have no significant effect on the chance of achieving a basic qualification for students with an age between 16 and 18 years. Only the factors paid work significantly influence the chance of achieving a basic qualification in a negative way. The more a student works the lower the chance of achieving a basic qualification. Due to a lack on inverse causality it is important to note that this study only found correlation between the independent variables and dependent variable.

From the results there can be concluded that the change in the law of compulsory education in august 2007 does not have a specific influence on the number of early school leavers with an age between 18 and 24 years. The reduction in the number of early school leavers is caused due to a

downward trend that began in 2004. Factors such as gender, household composition, paid work, self-reported happiness and self-reported health have a significant influence on the chance of achieving a basic qualification for them. Before and after the change the influences remain the same.

8. Limitations and recommendations

This study knows a couple of limitations. The POLS data was not complete because the Central Agency of Statistics decided to change the design of the questionnaire in 2009. Due to the change of the questionnaire the variables, self-reported happiness and self-reported health are missing. To measure a causal effect of the factors on the change of achieving a basic qualification there is more data needed. The dataset is from 2003 till 2009. There was change in the law of compulsory education in 2007. The change might influence people later still, maybe later than 2009. A limitation of this research is the number of years in the dataset. Another important limitation of this research has to do with the Parallel Trend Assumption. This assumption implies that the treatment group and control groups need to have the same trend in the outcome variable in absence of the treatment. The Netherlands was the treatment group and France, Latvia and Estonia was the control groups. The trends of the number of early school leavers were not exactly the same for the four countries. The control groups have the same law of compulsory education as the Netherlands before the change and were the closest in having the same trend as the Netherlands. Although several analyses are conducted there was no causal effect found of the change in the law of compulsory education on the number of early school leavers. As the results show that the change significantly influences the number of early school leavers positively it is not clear whether it is only because of the change in the law or also due to the other project, programs and investment in order to reduce the number of early school leavers.

It is interesting to know what would happened if the law of compulsory education changed. But now into a law where everyone has to achieve a basic qualification irrespectively someone's age. For further research is it important to know whether the quality of the school has a significant influence on the change of achieving a basic qualification. Countries outside the European Union can also be studied. For example how does India tackle the problem of early school leavers or the USA. The percentage early school leavers has to be under the 10% in all European countries. Not all countries have achieved that yet and it might be interesting to investigate if they are able to achieve it.

9. References

Belfield, C. (2008). *The Cost of Early School-leaving and School Failure .* New York: City University of New York.

Bowles, S. (1972). *Schooling and Inequality from Generation to Generation.* Chicago: The University of Chicago Press.

Brunello, G., Fort, M., & Weber, G. (2012). *Changes in Compulsory Schooling, Education and the Distribution of Wages in Europe .* The Economic Journal.

Byrne, D., & Smyth, E. (2010). *The Dynamics of Early School Leaving.* Dublin: The Liffey Press .

CBS. (2015, Maart 26). Permanent Onderzoek LeefSituatie (POLS). Den Haag, Zuid-Holland, Nederland.

CBS. (2015, Maart 16). *Voortijdig Schoolverlaters.* Den Haag: CBS.

Commission, E. (2000, Maart). *Europa Nu*. Opgehaald van Lissabon strategy: http://www.europa-nu.nl/id/vh7dotpnvfze/lissabonstrategie

Commission, E. (2012). *Early school leaving in Europe.* Brussell : European Commission .

Commission-, E. (2014). *The structure of the European education systems .* Brussell: European Commission.

Commission--, E. (2015). *Early school leaving.* Brussell: Eupopean Commission.

CPB. (2006). *Voortijdig schoolverlaten in Nederland: omvang, beleid en resultaten .* Den Haag: Centraal Planbureau .

EuropaNU. (2012). *Lissabonstrategy.* Leiden: University Leiden.

European-Commission. (2015, 4 14). *Education and Training*. Opgehaald van European Commission: http://ec.europa.eu/education/policy/school/early-school-leavers_en.htm

European-Commission. (2015, April 14). *Eurostat*. Opgehaald van Eurostat: http://ec.europa.eu/eurostat

Evans, B. (2008). *Difference-in-Difference.*

Gathmann, C., Jürges, H., & Reinhold, S. (2012). *Compulsory Schooling Reforms, Education and Mortality in Twentieth Century Europe .* CES.

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. (2010). *Impact Evaluation in Practice.* Washington : The World Bank.

Groot, W., & Maassen van den Brink, H. (2007). *The health effects of education.* Oxford: Oxford University Press.

Hanushek, E. A., & Wößmann, L. (2007). *The Role of Education Quality in Economic Growth.* World Bank Policy Research Working Paper.

Hörner, W., & Döbert, H. (2012). *The Education Systems of Europe – Introductory Methodological Remarks .*

Kippersluis, H. v. (2014, May 25). Difference-in-differences. Rotterdam, Zuid-Holland, The Netherlands.

Knox, R. M. (2011). *EFFECTS OF STUDENT MOBILITY ON THE ACADEMIC ACHIEVEMENT .* Louisiana State University.

Laerd. (2005). *Binomial Logistic Regression using SPSS*. Opgehaald van Laerd Statistics: https://statistics.laerd.com/spss-tutorials/binomial-logistic-regression-using-spss-statistics.php

Lochner, L., & Moretti, E. (2003). *The Effect of Education on Crime: Evidence from Prison Inmates ,Arrests, and Self-Reports.* University of Western Ontario.

McNeal. (1999). *Parent Involvement as Social Capital.* 2001: Journal of Socioeconomics.

Ministerie-OCW. (2007). *Operatie Jong en Jeugdzorgbrigade.* Den Haag: Ministerie van Binnenlandse Zaken en Koninkrijksrelaties .

Moore, McCabe, Alwan, Craig, & Duckworth. (2011). *The practice of Statistics for Business and Economics.* New York : W. H. Freeman and Company.

OCW. (1994). *RMC Functie.* Den Haag: Ministerie van onderwijs.

Onderwijs, M. v. (2015, Maart 3). *Beleid*. Opgehaald van Aanval op schooluitval: http://www.aanvalopschooluitval.nl/

Oreopoulos', P. (2003). *do dropsouts drop out too soon?* Cambrige: NATIONAL BUREAU OF ECONOMIC RESEARCH.

Oreopoulos, P. (2004). *The Compelling Effects of Compulsory Schooling.* Canada: University of Toronto.

OREOPOULOS, P. (2006). *Estimating Average and Local Average Treatment Effects of Education when Compulsory Schooling Laws Really Matter.* Canada: Canadian Journal of Economics.

Oreopoulos, P. (2009). *How large are returns to schooling? : hint : money isn't everything .* Cambridge : National Bureau of Economic Research.

Psacharopoulos, G., & Layard, R. (1979). *Human Capital and Earnings: British Evidence and a Critique.* Oxford: Oxford University Press.

Rijksoverheid. (2007). *Leerplicht.* Den Haag: Rijksoverheid.

Rijksoverheid-OCW. (2013). *Aanpak voortijdig schoolverlaten.* Den Haag: Rijksoverheid .

Rumberger, R. W. (1995). *Dropping Out of Middle School: A Multilevel Analysis of Students and Schools .* Santa Barbara : University of California.

SH, C., DJ, O., & J, L. (1980). *The effects of education on health.* Washington, D.C: World Bank.

Sloutions, S. (2015, May 25). *Assumptions of Multiple Linear Regression*. Opgehaald van StatisticsSloutions: http://www.statisticssolutions.com/assumptions-of-multiple-linear-regression/

Stearns, E., & Glennie, E. J. (2006). *When and Why Dropouts Leave High School.* Sage Publications.

Storimans, T. (2010). *Geschiedenis en achtergronden van de wettelijke leerplicht.* Amsterdam: Nederlandse Vereniging voor Onderwijsrecht (NVOR).

Trouw. (2006, Februarie 8). Premie moet schoolverlaten tegengaan. *Trouw*, p. 1.

USLegal. (2015, May 2015). Compulsory education law.

Witte, K. D., & Rogge, N. (2011). *Dropout from secondary education: All's well that begins well.* Leuven : Universiteit van Amsterdam.

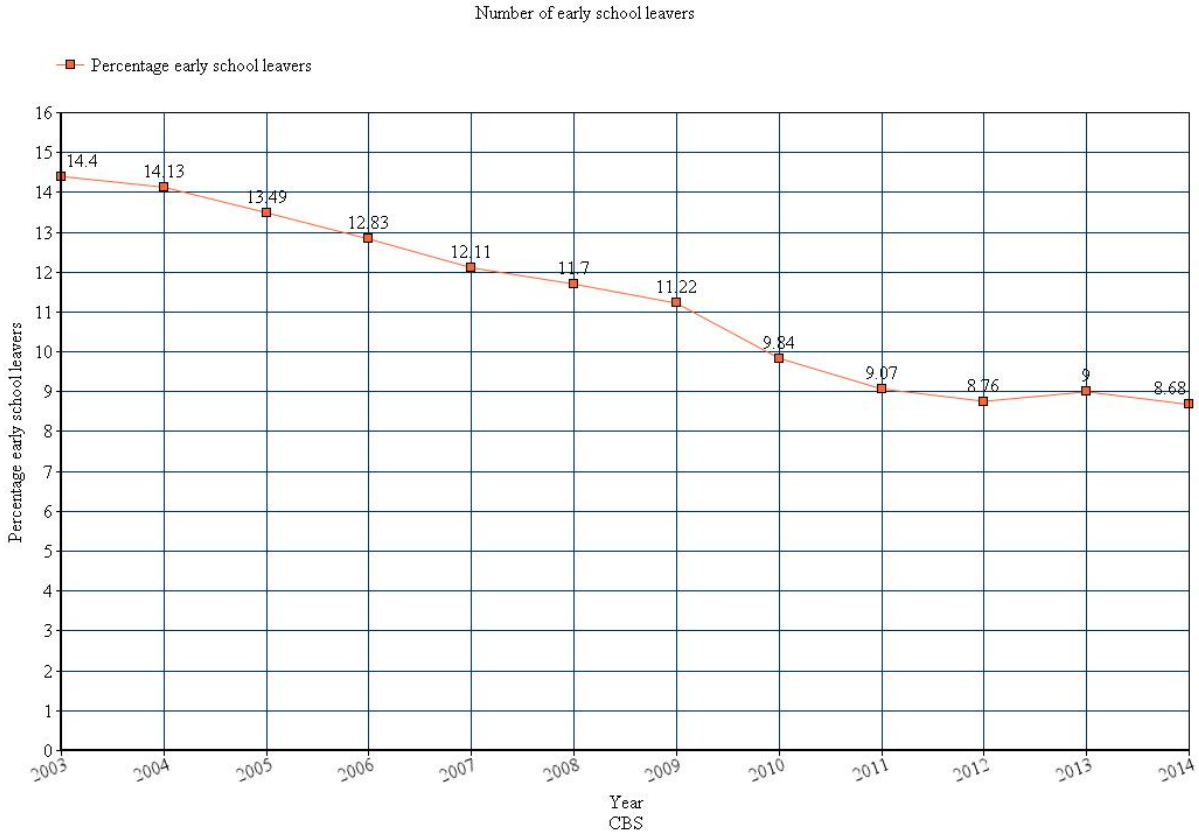
10. Appendix

Appendix A – I

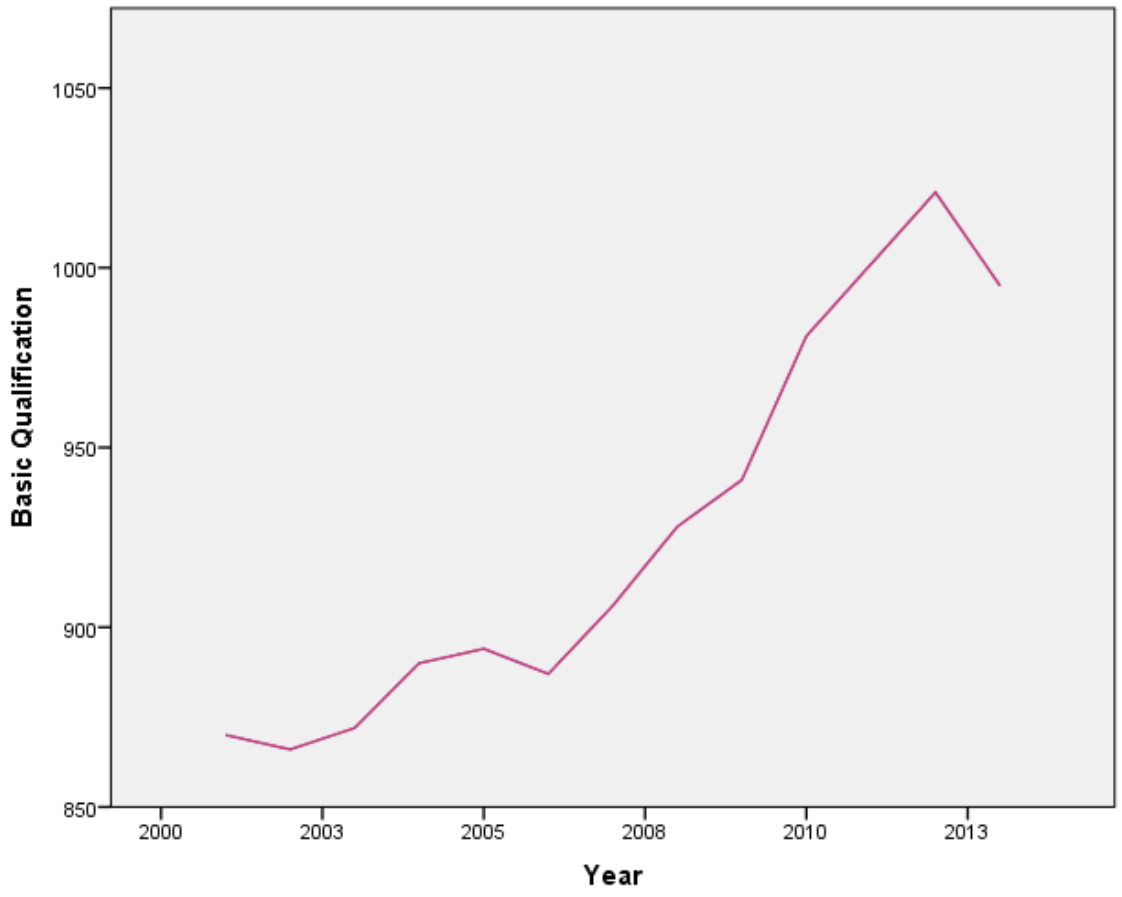
	2005	2005	2006	2006	2007	2007	2008	2008
School leavers without income	73,3%	26,7%	69,8%	30,2%	68,2%	31,8%	64,5%	35,5%
School leavers with income	43,9%	56,1%	38,8%	61,2%	37,2%	62,8%	34,6%	65,4%
0,5 tot 0,9 keer (yought wage)	49,2%	50,8%	45,1%	54,9%	43,5%	56,5%	39,6%	60,4%
0,9 tot 1,2 keer (yought wage)	44,1%	55,9%	39,9%	60,1%	38,8%	61,2%	36,7%	63,3%
1,2 tot 1,5 keer (yought wage)	40,9%	59,1%	33,9%	66,1%	33,0%	67,0%	30,8%	69,2%
>= 1,5 keer (yought wage)	37,2%	62,8%	32,2%	67,8%	30,6%	69,4%	29,4%	70,6%

	Early School leavers
	School leavers with a basic qualification

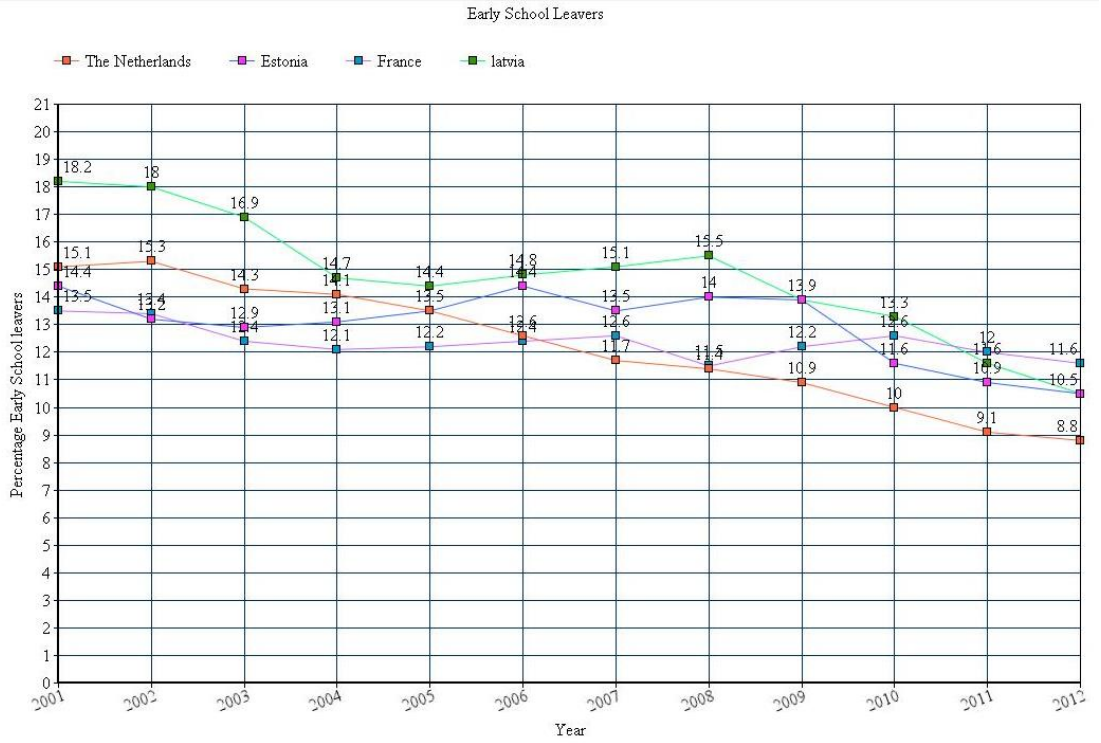
Appendix A - II



Appendix A – III



Appendix B – I



Appendix B – II

Treatment group = The Netherlands

Control group = France – Latvia - Estonia

Tests of Between-Subjects Effects

Dependent Variable: Percentageearlyschoolleavers

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	119,942[a]	13	9,226	4,758	,000
Intercept	3232,080	1	3232,080	1666,905	,000
CountryDummy	,020	1	,020	,010	,920
Year	64,298	11	5,845	3,015	,007
TreatmentDummy	12,960	1	12,960	6,684	,014
Error	65,925	34	1,939		
Total	8402,200	48			
Corrected Total	185,867	47			

a. R Squared = ,645 (Adjusted R Squared = ,510)

Parameter Estimates

Dependent Variable: Percentageearlyschoolleavers

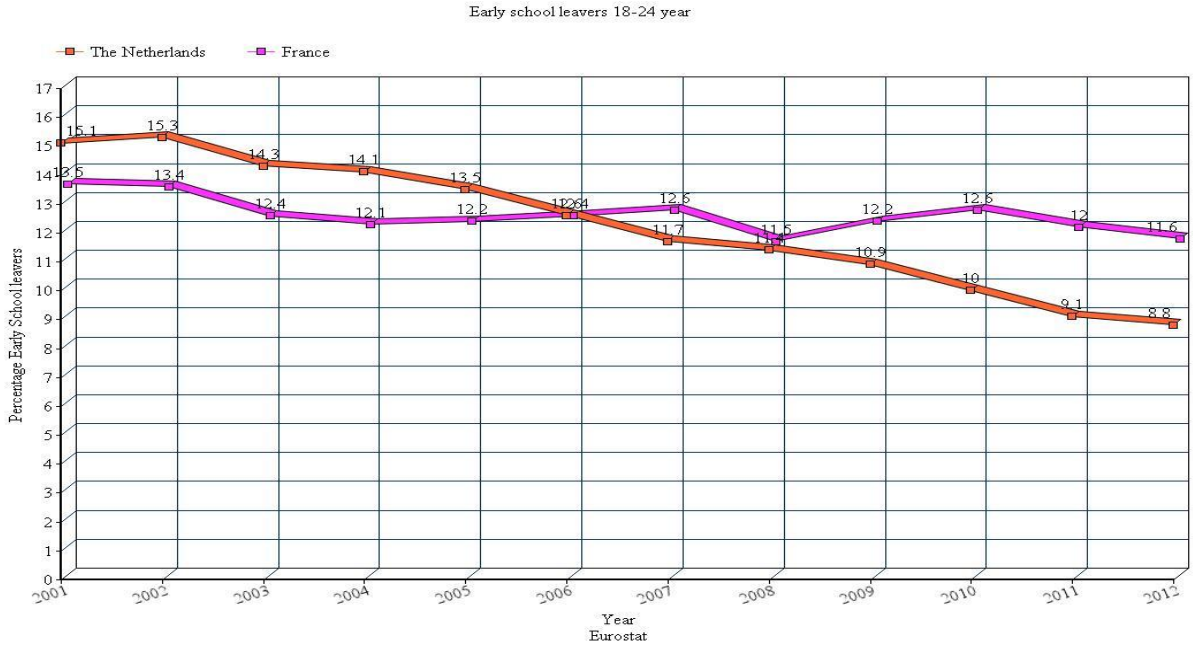
Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	11,000	,971	11,330	,000	9,027	12,973
[CountryDummy=0]	-,067	,656	-,102	,920	-1,401	1,267
[CountryDummy=1]	0[a]	.	.	.	.	.
[Year=2001]	4,350	1,012	4,300	,000	2,294	6,406
[Year=2002]	4,025	1,012	3,979	,000	1,969	6,081
[Year=2003]	3,175	1,012	3,139	,003	1,119	5,231
[Year=2004]	2,550	1,012	2,521	,017	,494	4,606
[Year=2005]	2,425	1,012	2,397	,022	,369	4,481
[Year=2006]	2,375	1,012	2,348	,025	,319	4,431
[Year=2007]	3,100	,985	3,148	,003	1,099	5,101
[Year=2008]	2,750	,985	2,793	,009	,749	4,751
[Year=2009]	2,375	,985	2,412	,021	,374	4,376
[Year=2010]	1,525	,985	1,549	,131	-,476	3,526
[Year=2011]	,550	,985	,559	,580	-1,451	2,551
[Year=2012]	0[a]	.	.	.	.	.
TreatmentDummy	-2,400	,928	-2,585	,014	-4,287	-,513

a. This parameter is set to zero because it is redundant.

Appendix C – I

Treatment group = The Netherlands

Control group = France



Appendix C – II

Tests of Between-Subjects Effects

Dependent Variable: Percentageearlyschoolleavers

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	56,099[a]	13	4,315	9,482	,001
Intercept	2064,563	1	2064,563	4536,671	,000
CountryDummy	6,601	1	6,601	14,505	,003
Year	11,895	11	1,081	2,376	,092
TreatmentDummy	15,844	1	15,844	34,815	,000
Error	4,551	10	,455		
Total	3694,070	24			
Corrected Total	60,650	23			

a. R Squared = ,925 (Adjusted R Squared = ,827)

Parameter Estimates

Dependent Variable: Percentageearlyschoolleavers

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	12,567	,646	19,457	,000	11,128	14,006
[CountryDummy=0]	-1,483	,389	-3,808	,003	-2,351	-,616
[CountryDummy=1]	0[a]	.	.	.	.	.
[Year=2001]	2,475	,729	3,397	,007	,851	4,099
[Year=2002]	2,525	,729	3,465	,006	,901	4,149
[Year=2003]	1,525	,729	2,093	,063	-,099	3,149
[Year=2004]	1,275	,729	1,750	,111	-,349	2,899
[Year=2005]	1,025	,729	1,407	,190	-,599	2,649
[Year=2006]	,675	,729	,926	,376	-,949	2,299
[Year=2007]	1,950	,675	2,891	,016	,447	3,453
[Year=2008]	1,250	,675	1,853	,094	-,253	2,753
[Year=2009]	1,350	,675	2,001	,073	-,153	2,853
[Year=2010]	1,100	,675	1,631	,134	-,403	2,603
[Year=2011]	,350	,675	,519	,615	-1,153	1,853
[Year=2012]	0[a]	.	.	.	.	.
TreatmentDummy	-3,250	,551	-5,900	,000	-4,477	-2,023

a. This parameter is set to zero because it is redundant.

Appendix D – I

Treatment group = The Netherlands & Control group = Lativia

Tests of Between-Subjects Effects

Dependent Variable: Percentageearlyschoolleavers

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	148,592[a]	13	11,430	26,372	,000
Intercept	2263,253	1	2263,253	5221,888	,000
CountryDummy	12,201	1	12,201	28,150	,000
Year	66,758	11	6,069	14,003	,000
TreatmentDummy	1,450	1	1,450	3,346	,097
Error	4,334	10	,433		
Total	4518,830	24			
Corrected Total	152,926	23			

a. R Squared = ,972 (Adjusted R Squared = ,935)

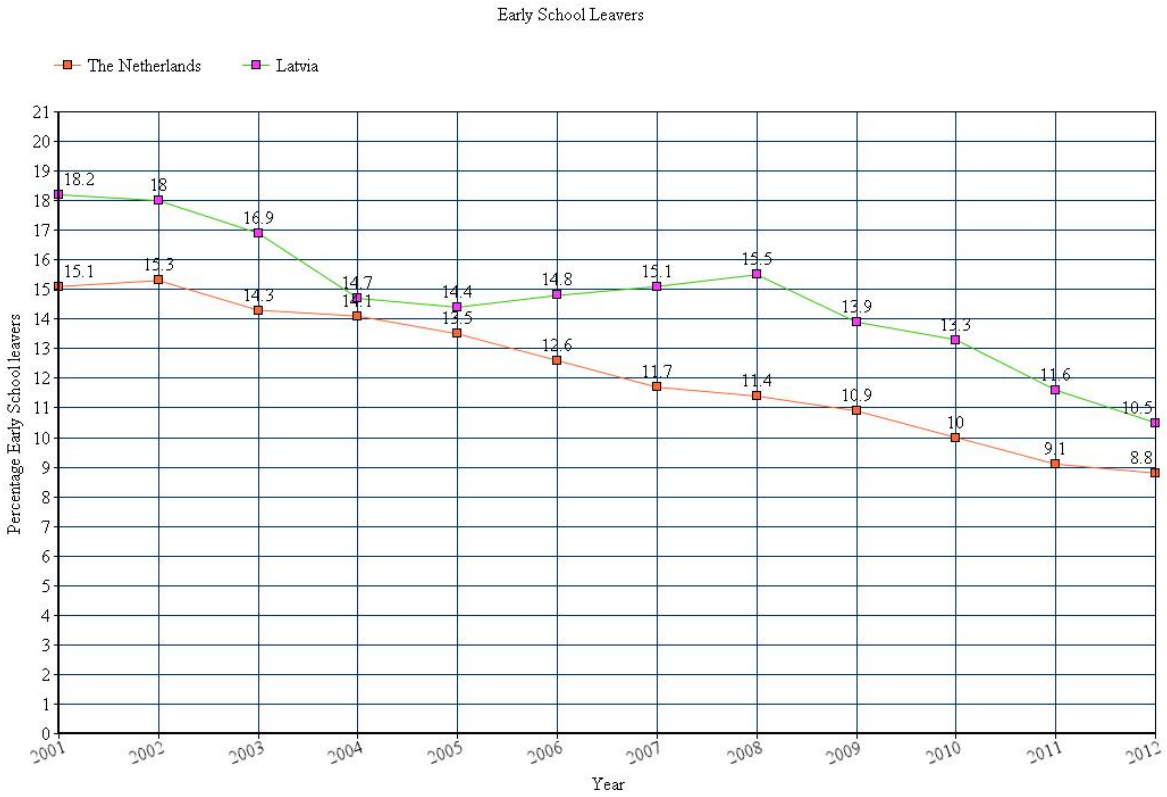
Parameter Estimates

Dependent Variable: Percentageearlyschoolleavers

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	9,133	,630	14,490	,000	7,729	10,538
[CountryDummy=0]	2,017	,380	5,306	,000	1,170	2,864
[CountryDummy=1]	0[a]	.	.	.	.	.
[Year=2001]	6,508	,711	9,153	,000	4,924	8,093
[Year=2002]	6,508	,711	9,153	,000	4,924	8,093
[Year=2003]	5,458	,711	7,676	,000	3,874	7,043
[Year=2004]	4,258	,711	5,988	,000	2,674	5,843
[Year=2005]	3,808	,711	5,356	,000	2,224	5,393
[Year=2006]	3,558	,711	5,004	,001	1,974	5,143
[Year=2007]	3,750	,658	5,696	,000	2,283	5,217
[Year=2008]	3,800	,658	5,772	,000	2,333	5,267
[Year=2009]	2,750	,658	4,177	,002	1,283	4,217
[Year=2010]	2,000	,658	3,038	,013	,533	3,467
[Year=2011]	,700	,658	1,063	,313	-,767	2,167
[Year=2012]	0[a]	.	.	.	.	.
TreatmentDummy	-,983	,538	-1,829	,097	-2,181	,214

a. This parameter is set to zero because it is redundant.

Appendix D – II

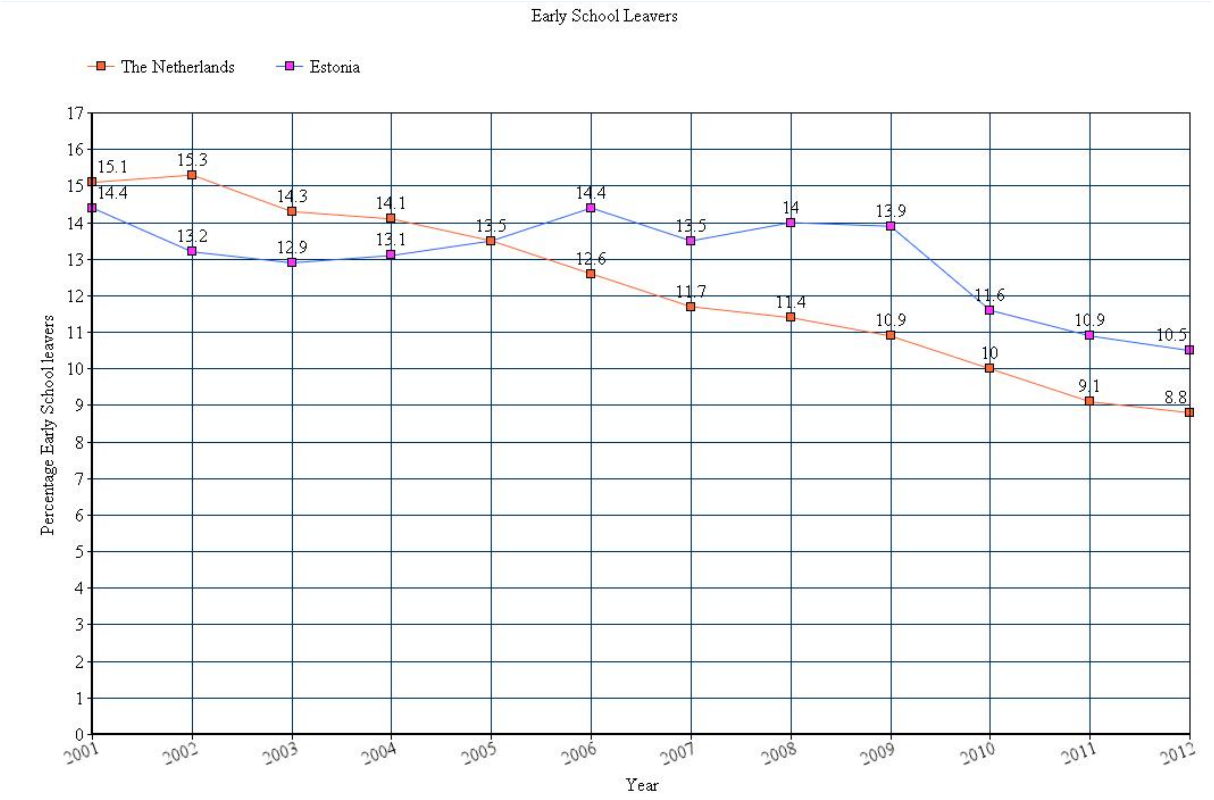


Appendix E – I

Treatment group = The Netherlands

Control group = Estonia

$$Yit = \alpha + \beta * countrydummy + \delta t + \gamma * Dit + \varepsilon it$$



Appendix E – II

Tests of Between-Subjects Effects

Dependent Variable: Percentageearlyschoolleavers

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	75,025[a]	13	5,771	15,884	,000
Intercept	2138,670	1	2138,670	5886,248	,000
CountryDummy	1,613	1	1,613	4,440	,061
Year	27,567	11	2,506	6,897	,002
TreatmentDummy	13,202	1	13,202	36,335	,000
Error	3,633	10	,363		
Total	3893,940	24			
Corrected Total	78,658	23			

a. R Squared = ,954 (Adjusted R Squared = ,894)

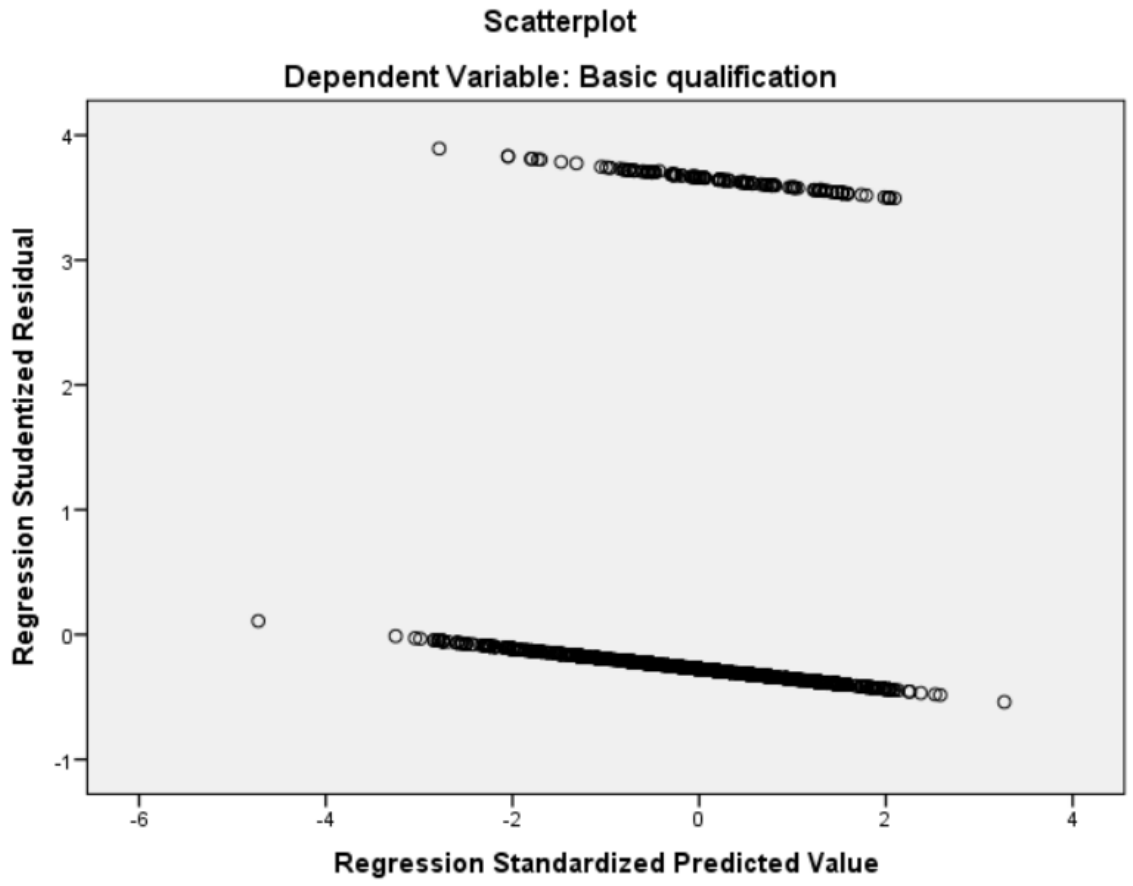
Parameter Estimates

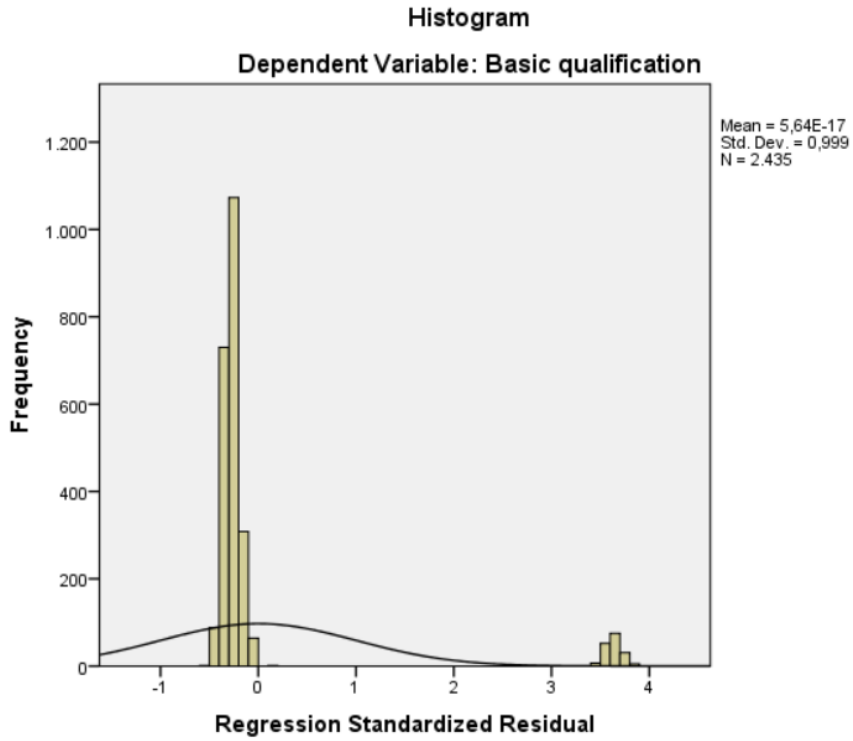
Dependent Variable: Percentageearlyschoolleavers

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	11,500	,577	19,927	,000	10,214	12,786
[CountryDummy=0]	-,733	,348	-2,107	,061	-1,509	,042
[CountryDummy=1]	0[a]	.	.	.	.	.
[Year=2001]	3,617	,651	5,555	,000	2,166	5,067
[Year=2002]	3,117	,651	4,787	,001	1,666	4,567
[Year=2003]	2,467	,651	3,789	,004	1,016	3,917
[Year=2004]	2,467	,651	3,789	,004	1,016	3,917
[Year=2005]	2,317	,651	3,558	,005	,866	3,767
[Year=2006]	1,917	,651	2,944	,015	,466	3,367
[Year=2007]	3,400	,603	5,641	,000	2,057	4,743
[Year=2008]	3,050	,603	5,060	,000	1,707	4,393
[Year=2009]	2,750	,603	4,562	,001	1,407	4,093
[Year=2010]	1,150	,603	1,908	,086	-,193	2,493
[Year=2011]	,350	,603	,581	,574	-,993	1,693
[Year=2012]	0[a]	.	.	.	.	.
TreatmentDummy	-2,967	,492	-6,028	,000	-4,063	-1,870

a. This parameter is set to zero because it is redundant.

Appendix F – I





Appendix F – II

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions					
				(Constant)	Gender	Household composition	Paid work	Self reported happiness	Self-reported health
1	1	5,209	1,000	,00	,00	,01	,00	,00	,00
	2	,535	3,121	,00	,00	,98	,00	,00	,00
	3	,095	7,414	,00	,70	,00	,00	,12	,18
	4	,078	8,151	,00	,04	,00	,08	,26	,71
	5	,065	8,977	,01	,11	,00	,40	,50	,00
	6	,018	17,072	,98	,14	,00	,52	,12	,10

a. Dependent Variable: Basic qualification

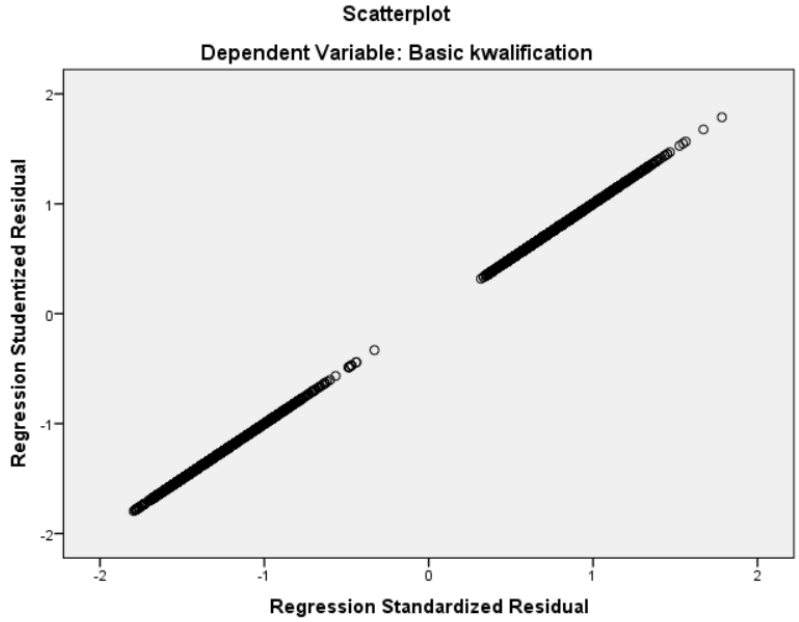
Model Summary[b]

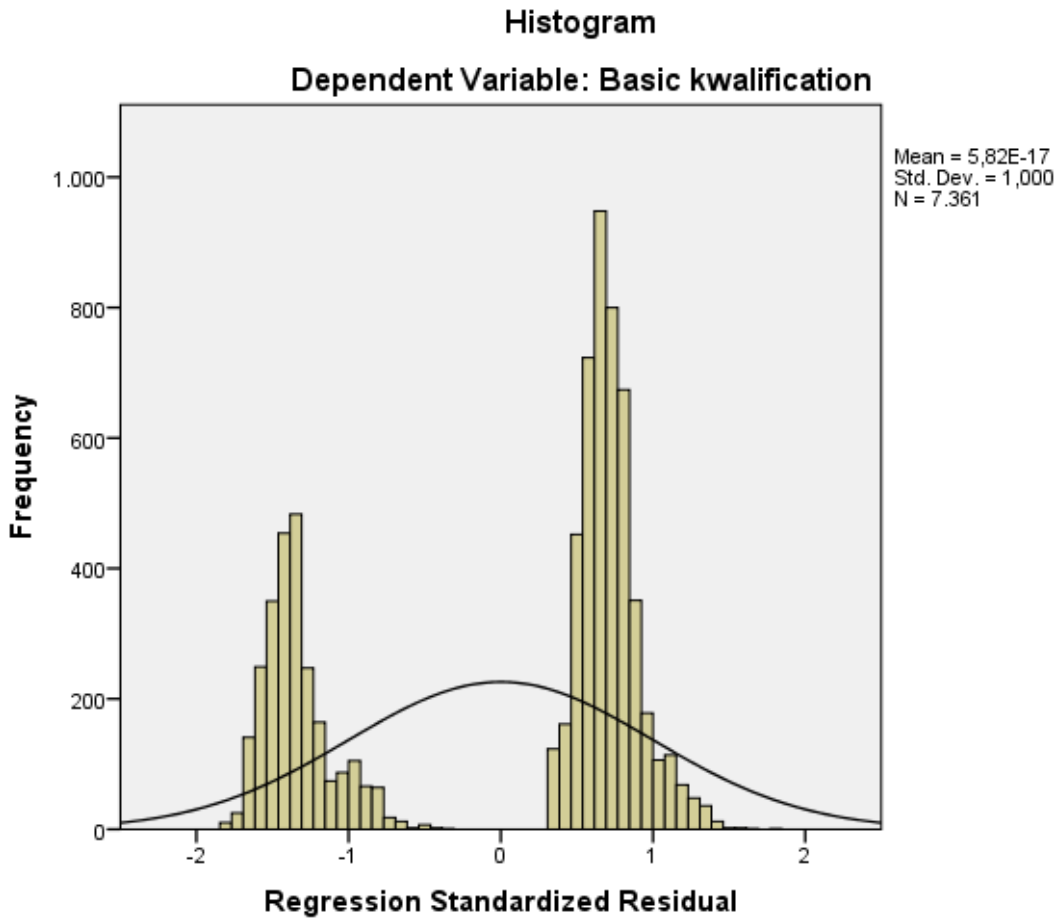
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	,080[a]	,006	,004	,254	2,008

a. Predictors: (Constant), Self-reported health, Paid work , Gender, Household composition, Self reported happiness

b. Dependent Variable: Basic qualification

Appendix G – I


Scatterplot
Dependent Variable: Basic kwalification
Regression Studentized Residual
Regression Standardized Residual
2
1
0
-1
-2
-2
-1
0
1
2


Histogram
Dependent Variable: Basic kwalification
Mean = 5,82E-17
Std. Dev. = 1,000
N = 7.361
Frequency
1.000
800
600
400
200
0
-2
-1
0
1
2
Regression Standardized Residual

Appendix G – II

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions					
				(Constant)	Gender	Household composition	Paid work	Self reported happiness	Self-reported health
1	1	5,088	1,000	,00	,00	,01	,00	,00	,00
	2	,619	2,867	,00	,00	,97	,00	,00	,00
	3	,102	7,050	,00	,45	,01	,04	,22	,21
	4	,094	7,350	,00	,23	,00	,51	,01	,22
	5	,071	8,438	,00	,06	,00	,14	,59	,49
	6	,025	14,209	,99	,25	,00	,31	,17	,06

a. Dependent Variable: Basic kwalification

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	,204[a]	,042	,041	,466	1,995

a. Predictors: (Constant), Self-reported health, Paid work , Household composition, Gender, Self reported happiness

b. Dependent Variable: Basic kwalification

Appendix H – I

POLS DATA – age 16 till 18 – 2003 t/m 2009

$$Yi = \alpha + \beta 1i * Gender + \beta 2i * household\ composition + \beta 3i * paid\ work + \beta 4i * self\ reported\ health + \beta 5i * self\ reported\ happiness + \varepsilon i$$

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,080[a]	,006	,004	,254

a. Predictors: (Constant), Self-reported health, Paid work , Gender, Household composition, Self reported happiness

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1,024	5	,205	3,168	,007[b]
	Residual	157,107	2429	,065		
	Total	158,131	2434			

a. Dependent Variable: Basic qualification

b. Predictors: (Constant), Self-reported health, Paid work , Gender, Household composition, Self reported happiness

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	,151	,034		4,487	,000			
	Gender	-,005	,010	-,010	-,491	,623	-,012	-,010	-,010
	Household composition	-,001	,001	-,040	-1,975	,048	-,042	-,040	-,040
	Paid work	-,021	,007	-,061	-3,003	,003	-,062	-,061	-,061
	Self reported happiness	,011	,009	,024	1,187	,235	,013	,024	,024
	Self-reported health	-,010	,009	-,025	-1,208	,227	-,024	-,025	-,024

a. Dependent Variable: Basic qualification

Appendix H – II

POLS DATA – age 16 till 18 – 2003 t/m 2006

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,090[a]	,008	,006	,258

a. Predictors: (Constant), Self-reported health, Paid work , Gender, Household composition, Self reported happiness

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1,058	5	,212	3,190	,007[b]
	Residual	128,886	1943	,066		
	Total	129,944	1948			

a. Dependent Variable: Basic qualification

b. Predictors: (Constant), Self-reported health, Paid work , Gender, Household composition, Self reported happiness

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	,178	,039		4,600	,000			
	Gender	-,007	,012	-,013	-,578	,563	-,014	-,013	-,013
	Household composition	-,001	,001	-,042	-1,827	,068	-,045	-,041	-,041
	Paid work	-,026	,008	-,074	-3,246	,001	-,075	-,073	-,073
	Self reported happiness	,007	,011	,015	,661	,508	,003	,015	,015
	Self-reported health	-,010	,010	-,023	-,995	,320	-,023	-,023	-,022

a. Dependent Variable: Basic qualification

Appendix H – III

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,071[a]	,005	-,005	,242

a. Predictors: (Constant), Self-reported health, Paid work , Household composition, Gender, Self reported happiness

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,143	5	,029	,492	,782[b]
	Residual	28,005	480	,058		
	Total	28,148	485			

a. Dependent Variable: Basic qualification

b. Predictors: (Constant), Self-reported health, Paid work , Household composition, Gender, Self reported happiness

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	,057	,068		,842	,400			
	Gender	-,001	,022	-,002	-,036	,971	-,002	-,002	-,002
	Household composition	-,001	,001	-,031	-,686	,493	-,028	-,031	-,031
	Paid work	-,004	,015	-,012	-,257	,797	-,010	-,012	-,012
	Self reported happiness	,025	,020	,060	1,302	,194	,053	,059	,059
	Self-reported health	-,012	,017	-,032	-,699	,485	-,029	-,032	-,032

a. Dependent Variable: Basic qualification

Appendix I – I

POLS DATA – Age 18 till 24 – 2003 t/m 2009

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,204[a]	,042	,041	,466

a. Predictors: (Constant), Self-reported health, Paid work , Household composition, Gender, Self reported happiness

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	69,795	5	13,959	64,160	,000[b]
	Residual	1600,195	7355	,218		
	Total	1669,990	7360			

a. Dependent Variable: Basic kwalification

b. Predictors: (Constant), Self-reported health, Paid work , Household composition, Gender, Self reported happiness

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	,845	,030		27,835	,000			
	Gender	,082	,011	,086	7,509	,000	,082	,087	,086
	Household composition	-,007	,001	-,128	-11,169	,000	-,145	-,129	-,127
	Paid work	-,031	,007	-,053	-4,632	,000	-,065	-,054	-,053
	Self reported happiness	-,052	,009	-,065	-5,493	,000	-,102	-,064	-,063
	Self-reported health	-,051	,009	-,070	-5,843	,000	-,085	-,068	-,067

a. Dependent Variable: Basic kwalification

Appendix I – II

POLS DATA – Age 18 till 24 – 2003 t/m 2006

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,197[a]	,039	,038	,468

a. Predictors: (Constant), Self-reported health, Paid work , Household composition, Gender, Self reported happiness

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	53,258	5	10,652	48,692	,000[b]
	Residual	1315,146	6012	,219		
	Total	1368,404	6017			

a. Dependent Variable: Basic kwalification

b. Predictors: (Constant), Self-reported health, Paid work , Household composition, Gender, Self reported happiness

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	,845	,034		24,971	,000			
	Gender	,078	,012	,082	6,419	,000	,077	,082	,081
	Household composition	-,007	,001	-,124	-9,759	,000	-,138	-,125	-,123
	Paid work	-,029	,007	-,051	-4,004	,000	-,063	-,052	-,051
	Self reported happiness	-,054	,011	-,068	-5,133	,000	-,101	-,066	-,065
	Self-reported health	-,049	,010	-,066	-5,029	,000	-,081	-,065	-,064

a. Dependent Variable: Basic kwalification

Appendix I – III

POLS DATA – Age 18 till 24 – 2007 t/m 2009

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,236[a]	,056	,052	,461

a. Predictors: (Constant), Self-reported health, Paid work , Household composition, Gender, Self reported happiness

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	16,845	5	3,369	15,825	,000[b]
	Residual	284,645	1337	,213		
	Total	301,491	1342			

a. Dependent Variable: Basic kwalification

b. Predictors: (Constant), Self-reported health, Paid work , Household composition, Gender, Self reported happiness

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	,842	,069		12,183	,000			
	Gender	,103	,026	,108	4,029	,000	,105	,110	,107
	Household composition	-,008	,001	-,147	-5,427	,000	-,173	-,147	-,144
	Paid work	-,039	,016	-,064	-2,379	,018	-,072	-,065	-,063
	Self reported happiness	-,041	,021	-,054	-1,937	,053	-,106	-,053	-,051
	Self-reported health	-,061	,020	-,083	-2,980	,003	-,105	-,081	-,079

a. Dependent Variable: Basic kwalification

Appendix J – I

Tests of Between-Subjects Effects

Dependent Variable: Percentageearlyschoolleavers

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	119,702[a]	13	9,208	4,732	,000
Intercept	2735,100	1	2735,100	1405,483	,000
CountryDummy	,280	1	,280	,144	,707
Year	65,644	11	5,968	3,067	,006
TreatmentDummy	12,720	1	12,720	6,537	,015
Error	66,165	34	1,946		
Total	8402,200	48			
Corrected Total	185,867	47			

a. R Squared = ,644 (Adjusted R Squared = ,508)

Parameter Estimates

Dependent Variable: Percentageearlyschoolleavers

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	11,158	1,014	11,002	,000	9,097	13,219
[CountryDummy=0]	-,273	,720	-,379	,707	-1,737	1,191
[CountryDummy=1]	0[a]	.	.	.	.	.
[Year=2001]	4,347	1,014	4,286	,000	2,286	6,408
[Year=2002]	4,022	1,014	3,966	,000	1,961	6,083
[Year=2003]	3,172	1,014	3,128	,004	1,111	5,233
[Year=2004]	2,547	1,014	2,511	,017	,486	4,608
[Year=2005]	2,422	1,014	2,388	,023	,361	4,483
[Year=2006]	2,975	,986	3,016	,005	,970	4,980
[Year=2007]	3,100	,986	3,143	,003	1,095	5,105
[Year=2008]	2,750	,986	2,788	,009	,745	4,755
[Year=2009]	2,375	,986	2,408	,022	,370	4,380
[Year=2010]	1,525	,986	1,546	,131	-,480	3,530
[Year=2011]	,550	,986	,558	,581	-1,455	2,555
[Year=2012]	0[a]	.	.	.	.	.
TreatmentDummy	-2,411	,943	-2,557	,015	-4,328	-,495

a. This parameter is set to zero because it is redundant.

Appendix J – II

Tests of Between-Subjects Effects

Dependent Variable: Percentageearlyschoolleavers

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	116,227[a]	13	8,941	4,365	,000
Intercept	2192,403	1	2192,403	1070,386	,000
CountryDummy	,270	1	,270	,132	,719
Year	68,160	11	6,196	3,025	,006
TreatmentDummy	9,245	1	9,245	4,514	,041
Error	69,640	34	2,048		
Total	8402,200	48			
Corrected Total	185,867	47			

a. R Squared = ,625 (Adjusted R Squared = ,482)

Parameter Estimates

Dependent Variable: Percentageearlyschoolleavers

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	11,113	1,103	10,077	,000	8,871	13,354
[CountryDummy=0]	-,300	,826	-,363	,719	-1,979	1,379
[CountryDummy=1]	0[a]	.	.	.	.	.
[Year=2001]	4,412	1,043	4,230	,000	2,293	6,532
[Year=2002]	4,087	1,043	3,918	,000	1,968	6,207
[Year=2003]	3,237	1,043	3,104	,004	1,118	5,357
[Year=2004]	2,612	1,043	2,504	,017	,493	4,732
[Year=2005]	3,025	1,012	2,989	,005	,968	5,082
[Year=2006]	2,975	1,012	2,940	,006	,918	5,032
[Year=2007]	3,100	1,012	3,063	,004	1,043	5,157
[Year=2008]	2,750	1,012	2,717	,010	,693	4,807
[Year=2009]	2,375	1,012	2,347	,025	,318	4,432
[Year=2010]	1,525	1,012	1,507	,141	-,532	3,582
[Year=2011]	,550	1,012	,543	,590	-1,507	2,607
[Year=2012]	0[a]	.	.	.	.	.
TreatmentDummy	-2,150	1,012	-2,125	,041	-4,207	-,093

a. This parameter is set to zero because it is redundant.

Appendix J – III

Tests of Between-Subjects Effects

Dependent Variable: Percentageearlyschoolleavers

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	111,795[a]	13	8,600	3,947	,001
Intercept	1624,090	1	1624,090	745,482	,000
CountryDummy	,040	1	,040	,018	,893
Year	71,791	11	6,526	2,996	,007
TreatmentDummy	4,813	1	4,813	2,209	,146
Error	74,072	34	2,179		
Total	8402,200	48			
Corrected Total	185,867	47			

a. R Squared = ,601 (Adjusted R Squared = ,449)

Parameter Estimates

Dependent Variable: Percentageearlyschoolleavers

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	10,872	1,238	8,781	,000	8,356	13,388
[CountryDummy=0]	-,133	,984	-,136	,893	-2,133	1,866
[CountryDummy=1]	0[a]	.	.	.	.	.
[Year=2001]	4,528	1,082	4,186	,000	2,330	6,726
[Year=2002]	4,203	1,082	3,886	,000	2,005	6,401
[Year=2003]	3,353	1,082	3,100	,004	1,155	5,551
[Year=2004]	3,150	1,044	3,018	,005	1,029	5,271
[Year=2005]	3,025	1,044	2,898	,007	,904	5,146
[Year=2006]	2,975	1,044	2,850	,007	,854	5,096
[Year=2007]	3,100	1,044	2,970	,005	,979	5,221
[Year=2008]	2,750	1,044	2,635	,013	,629	4,871
[Year=2009]	2,375	1,044	2,276	,029	,254	4,496
[Year=2010]	1,525	1,044	1,461	,153	-,596	3,646
[Year=2011]	,550	1,044	,527	,602	-1,571	2,671
[Year=2012]	0[a]	.	.	.	.	.
TreatmentDummy	-1,689	1,136	-1,486	,146	-3,998	,620

a. This parameter is set to zero because it is redundant.

Appendix K – I

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1216,501[a]	,007	,017

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

Classification Table[a]

			Predicted		
			Basic qualification		Percentage Correct
	Observed		0	1	
Step 1	Basic qualification	0	2265	0	100,0
		1	170	0	,0
	Overall Percentage				93,0

a. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Gender	-,075	,160	,221	1	,638	,927
	Householdcomposition	-,022	,011	3,931	1	,047	,978
	Paidwork	-,318	,106	8,960	1	,003	,728
	Selfreportedhappiness	,176	,146	1,448	1	,229	1,193
	Selfreportedhealth	-,165	,135	1,497	1	,221	,848
	Constant	-1,365	,508	7,222	1	,007	,255

a. Variable(s) entered on step 1: Gender, Householdcomposition, Paidwork, Selfreportedhappiness, Selfreportedhealth.

Appendix K – II

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	9208,248[a]	,040	,056

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than ,001.

Classification Table[a]

Observed			Predicted		
			Basic kwalification		Percentage Correct
			0	1	
Step 1	Basic kwalification	0	369	2192	14,4
		1	294	4506	93,9
	Overall Percentage				66,2

a. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Gender	,379	,051	56,206	1	,000	1,461
	Householdcomposition	-,030	,003	115,473	1	,000	,970
	Paidwork	-,139	,030	21,085	1	,000	,870
	Selfreportedhappiness	-,236	,043	29,501	1	,000	,790
	Selfreportedhealth	-,234	,040	33,614	1	,000	,791
	Constant	1,502	,140	115,852	1	,000	4,492

a. Variable(s) entered on step 1: Gender, Householdcomposition, Paidwork, Selfreportedhappiness, Selfreportedhealth.

Appendix L – I

Parameter Estimates

Dependent Variable: Percentageearlyschoolleavers

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	10,550	,544	19,384	,000	9,441	11,659
[Year=2001]	4,350	,700	6,218	,000	2,925	5,775
[Year=2002]	4,025	,700	5,754	,000	2,600	5,450
[Year=2003]	3,175	,700	4,538	,000	1,750	4,600
[Year=2004]	2,550	,700	3,645	,001	1,125	3,975
[Year=2005]	2,425	,700	3,466	,002	1,000	3,850
[Year=2006]	2,375	,700	3,395	,002	,950	3,800
[Year=2007]	3,100	,681	4,553	,000	1,713	4,487
[Year=2008]	2,750	,681	4,039	,000	1,363	4,137
[Year=2009]	2,375	,681	3,488	,001	,988	3,762
[Year=2010]	1,525	,681	2,240	,032	,138	2,912
[Year=2011]	,550	,681	,808	,425	-,837	1,937
[Year=2012]	0[a]	.	.	.	.	.
[CountryDummy=1]	,450	,508	,887	,382	-,584	1,484
[CountryDummy=2]	-,608	,393	-1,547	,132	-1,409	,192
[CountryDummy=3]	1,758	,393	4,473	,000	,958	2,559
[CountryDummy=4]	0[a]	.	.	.	.	.
TreatmentDummy	-2,400	,642	-3,738	,001	-3,708	-1,092

a. This parameter is set to zero because it is redundant.

Appendix M - I

Parameter Estimates

Dependent Variable: Basic qualification

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	,094	,164	,573	,567	-,227	,415
[Selfreportedhappiness=1]	,045	,148	,305	,760	-,244	,334
[Selfreportedhappiness=2]	,058	,147	,391	,696	-,231	,346
[Selfreportedhappiness=3]	,066	,149	,446	,655	-,225	,358
[Selfreportedhappiness=4]	,092	,157	,586	,558	-,216	,399
[Selfreportedhappiness=5]	0[a]	.	.	.	.	.
[Selfreportedhealth=1]	,014	,067	,209	,834	-,117	,145
[Selfreportedhealth=2]	,001	,066	,015	,988	-,129	,131
[Selfreportedhealth=3]	-,007	,068	-,103	,918	-,140	,126
[Selfreportedhealth=4]	0[a]	.	.	.	.	.
Gender	-,005	,010	-,503	,615	-,026	,015
Householdcomposition	-,001	,001	-1,959	,050	-,002	1,253E-006
Paidwork	-,021	,007	-3,005	,003	-,035	-,007

a. This parameter is set to zero because it is redundant.

Appendix M – II

Parameter Estimates

Dependent Variable: Basic kwalification

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	,131	,163	,805	,421	-,189	,451
[Selfreportedhappiness=0]	,051	,224	,227	,821	-,388	,489
[Selfreportedhappiness=1]	,340	,104	3,261	,001	,136	,544
[Selfreportedhappiness=2]	,297	,104	2,865	,004	,094	,500
[Selfreportedhappiness=3]	,268	,106	2,536	,011	,061	,474
[Selfreportedhappiness=4]	,095	,113	,845	,398	-,125	,316
[Selfreportedhappiness=5]	0[a]	.	.	.	.	.
[Selfreportedhealth=0]	,253	,370	,682	,495	-,473	,978
[Selfreportedhealth=1]	,254	,137	1,859	,063	-,014	,522
[Selfreportedhealth=2]	,216	,136	1,583	,113	-,051	,483
[Selfreportedhealth=3]	,137	,137	1,000	,317	-,131	,405
[Selfreportedhealth=4]	,150	,145	1,039	,299	-,133	,433
[Selfreportedhealth=5]	0[a]	.	.	.	.	.
Gender	,082	,011	7,490	,000	,061	,104
Householdcomposition	-,007	,001	-11,163	,000	-,008	-,006
Paidwork	-,030	,007	-4,425	,000	-,043	-,016

a. This parameter is set to zero because it is redundant.