university of groningen

faculty of spatial sciences

Doomed from Birth?

The associations of birthweight to obesity development in later life by looking for the importance of inherited and environmental factors among German twins

Lieke Heupink, S2759632

Email: l.f.heupink@student.rug.nl

Supervisor: Dr. Tobias C. Vogt

Second Assessor: Prof. Dr. Ir. Hinke Haisma

Master Thesis

MSc Population Studies 2017-2018

Population Research Centre

University of Groningen

Groningen: 2018, 29th of August

1 ABSTRACT

Background: Globally obesity affects over 650 million individuals. The mechanism underlying obesity is complex and no consensus on its workings prevails. Birthweight, genetic and environmental factors all have been theorized to influence obesity development. **Objective:** The aim of this study is twofold. First, it tests whether an association between birthweight and obesity in later life exists, after which the importance of genetics and social environment are determined. Data: Data from the TwinLife project is used. In total, 1,002 mono- and dizygotic twin pairs between the ages of 4 to 25 qualified for this study. **Methods:** The dependent variable is BMI, transformed to a z-score, with the main predictor being birthweight. Further, in the regression sex, mothers' age at birth and mothers' education are added. In the first regression, twins are treated as individuals. To look at the effect of twins a within-twin pair X-value is added and robust standard errors control for the paired structure of the data. **Findings:** Birthweight is positively associated with BMI when treating twins as individuals. When stratified the association is solely significant for twins with birthweight >2500 grams. The within twin pair X-value is mostly insignificant. In dizygotic twins the X-value is significant, which leads us to conclude that genetics and in-utero environments play a role in the association of birthweight and BMI. Furthermore, females have lower BMI compared to males. Next to that, high education and mothers' age >25 have a negative effect on BMI. **Conclusion:** Individuals are not doomed from birth, the mechanism of obesity is a complex interplay of birthweight, genetics and the social environment.

2 TABLE OF CONTENTS

3 TABLE OF FIGURES & TABLES

4 ACKNOWLEDGMENTS

The success of this thesis would not have been accomplished without the help and support of the following individuals. I would first like to thank my supervisor Tobias Vogt for giving me the opportunity to research twins from the TwinLife dataset, the coffee with milk during all our meetings and the very useful feedback. I want to thank Billie de Haas for helping me with the start-up process and giving feedback to scope the research to a thesis instead of a PhD project. Further, I want to thank my parents for unlimited streams of tea and food during the writing process. I am also proud on my brother for finally remembering my studies. Lastly, my thanks go out to my friends at the University of Groningen and Twente for their unlimited support, feedback and motivational quotes.

5 INTRODUCTION

Obesity is one of the major global health challenges. Estimations show that in 2016, 1.9 billion adults over 18 years old were overweight and of these adults 650 million were obese. In other words, meaning that 39% of the adult population was overweight (World Health Organization, 2018a). From 1975 to 2014, the Body Mass Index (BMI) has increased for both men and women from 21.7 kg/m2 to 24.2 kg/m2 and 22.1 kg/m2 to 24.4 kg/m2, respectively (NCD Risk Factor Collaboration, 2016). Differences between countries are evident, within the Organisation for Economic Co-operation (OECD), countries as Japan and South-Korea have less than 6% of the population over the age of 15 who are obese. While obesity rates in the United States indicate that over 38% of the population. In Europe, Germany had the second highest obesity rate (23.6%) which was a little below Finland (24.8%) (OECD, 2017, June). In high income countries, since the year 2000, the BMI has not increased as rapid as in the years before. Most likely this can be contributed to growing understanding that obesity is a serious public health threat. However, in other areas in the world the increase of BMI has not been slowing down (NCD Risk Factor Collaboration, 2016).

This current obesity pandemic has serious consequences for the wellbeing of populations. Obesity and overweight have many negative consequences for the health status of a population. Obesity and overweight are related to increased risk of death and many associated diseases. By way of illustration, some of these diseases include: hypertension[a], stroke[a], diabetes[a], certain forms of cancer[b], difficulties related to fertility[c], and increased risk on Alzheimer[d] (aDanaei et al., 2014; bCalle, Rodriguez, Walker-Thurmond, & Thun, 2003; cGesink Law, Maclehose, & Longnecker, 2006; cRamlau-Hansen et al., 2007; dKivipelto et al., 2005). Next to the physical consequences, even though obesity prevalence is high, obese person are facing discrimination and prejudice due to their body composition (Puhl & Heuer, 2009).Potentially, this explains the increased risk of depression and suicide, the lower quality of life, and declined career potentials among obese persons (Jesse, Xiaoling, Marbán, Henry, & Lowell, 2004).

Overweight and obesity are not merely health issues observed in the adult population. Globally, the prevalence of obesity and overweight among of children and adolescent has increased (Schienkiewitz, Brettschneider, Damerow, & Rosario, 2018). In 1975, 11 million children were obese which increased to 124 million in 2016. Additionally, in 2016, 213 million children were overweight which considering the obese children means that almost one in five (18.4%) children was overweight or obese (World Health Organization, 2018b). Luckily, in most high-income countries, including Germany, the rise in BMI has stabilized (Schienkiewitz et al., 2018). In Asia, the prevalence of obesity among children is lower than in the North-America, South-America and Europe. in various countries in Asia child obesity is still increasing rapidly (Lobstein, Baur, & Uauy, 2004). Although the trend of stabilizing obesity rates in high income countries is promising, in most countries the current prevalence are still serious issues. For example, in Germany, the prevalence of obesity among children from 3 to 17 years is 5.9% and a larger proportion, 15.4%, of these children is overweight (Schienkiewitz et al., 2018). Next to the previously mentioned consequences of overweight and obesity, children are even more vulnerable. In 2007, Sighn et al. showed that overweight and obese adolescents were between 20% and 58% likely to become overweight adults and between 24% and 90% likely to become obese adults, respectively.

Underlying this current obesity epidemic, despite being extensively researched, is that the mechanism of obesity is still debated Essentially, the development of overweight and obesity depends on a long-term imbalance in the energy consumption and expenditure. However, this is a huge simplification due to consumption and expenditure not being independent from each other (Naukkarinen, Rissanen, Kaprio, &

Pietiläinen, 2012). Further, the health status of a person is a lot more complex and depending on many factors (Halfon & Hochstein, 2002). In the case of obesity, research suggest that in-utero environment where babies develop and grow, are essential for obesity development. A study in 2010 by Morris et al. showed that children born in a household with low-income have higher birthweights and more chance to develop obesity later in life.

How birthweight influences the development of obesity is still under debate. Potential explanations include environmental, as well as, genetic factors. In 1990, Barker was one of the first persons to write and attract attention for the importance of the in-utero environment on the health status later in life (Almond & Currie, 2011). Barker observed that birthweight and development were influenced by maternal nutrition. During periods of undernutrition of mothers, babies were more likely to suffer from obesity and related diseases (Barker, 1990). Nowadays this importance of the in-utero environment is referred to as the "*fetal origins*" hypothesis. Another explanation describes the importance of certain genes which influence low birthweight and influence the weight and height of a person later in life (The, Adair, & Gordon-Larsen, 2010). Where previously the scientific debates were about if either genetics, nature, or the environment, nurture, determines the health outcomes in a person, the debate is shifting towards the shared understanding that genetic and environmental both influence health outcomes. Neither of the two are independent, nor are they static (Poulton et al., 2002).

In this thesis we use data from the TwinLife project to study the association of birthweight, social origin, genetic origin to obesity development. The aim of this thesis is to add to the understanding of the mechanism of obesity by studying twins. Twins are unique in that they share genetics, as well as, their environment in-utero and after. Hence, twin studies offer the unique opportunity to study the effects of birthweight, genes, environment, and interactions and fill the current knowledge gaps or explore scientific inconsistencies.

5.1 ACADEMIC AND SOCIAL RELEVANCE

5.1.1 Academic relevance & Uniqueness of twins

The TwinLife is the first dataset that allows to study the association between birthweight and obesity in German twins. There was no public twin study in Germany before that could have been used to study this relationship. Firstly, twins are often born with a different birthweight due to dissimilar intake of nutrition and unequal blood supply from the placenta (Baird et al., 2001). Twins share a large part from their genetic material due to the fertilization process. A dizygotic twin is born when the egg cell, ova, is fertilized with two separate sperm cells. Therefore, the proportion of genetic material shared by dizygotic twins is comparable to those of siblings around 50%. Monozygotic twins are born when an egg cell separates after separates after fertilization by a sperm cell. This leads to almost completely identical genetic material. Slight changes in the genetic make-up can still occur due to mutations and modifications which take place after the separation (Naukkarinen et al., 2012). By way of illustration, if within-twin pair differences are observed in the BMI among monozygotic twins, it cannot depend on the genetic origin due to being identical. A more likely explanation would be the difference in the environment of twins. The environment for twins is growing up is the same. Therefore, comparing within-twin pairs automatically controls for shared factors as smoking parents, household income and built environment. Besides comparing within-twin pairs, in this thesis also the twin pairs will be compared. These between pair difference can be caused by genes or the environment. All in all, twin studies offer the rare opportunity of making distinctions between genetic and environmental contributions on the associations of birthweight and obesity development.

Twin studies are a specific niche of statistic and epidemiological research. Most studies use different methods or do not have clear definitions for how they treat the paired structure of twin data methods used in reviews (Carlin, Gurrin, Sterne, Morley, & Dwyer, 2005). Nevertheless, these studies are important due to genetic factors being often ignored. Making and updating twin datasets is expensive, as well as, other genetic research (Mackenbach, 2005). To my knowledge there have not yet been any published articles on the topic of health outcomes and health inequalities using the data from the TwinLife project (TwinLife Bibliography, 2018). Additionally, there have only been a handful of studies using data from person being born during the global obesity epidemic.

5.1.2 Social relevance

The social relevance of this study is reflected in the size of the current obesity challenge the world is facing. As expressed, in the introduction the prevalence of obesity is very high. Further, we discussed the negative impact obesity can have on the health and psychological wellbeing of individuals. The issues, however, does not stop at the individual but negatively impacts entire countries. In Germany, the problem of obesity and overweight can be observed in high health expenditures. An estimation from 2008 proposed that obesity cost 16,797 million euros which is increase of 70% compared to 2002. These obesity cost account for 3.27% of the total German health expenditures (Lehnert, Streltchenia, Konnopka, Riedel-Heller, & König, 2015). Hence, it is not surprising that previous studies have suggested obesity among children and adolescent should be a public health priority to which requires successful preventive strategies (Singh, Mulder, Twisk, et al., 2007).

With the approach in this thesis, we allow to see what factors influences the development of obesity and how we potentially could control for this risk factors. The results of this study can therefore potentially give insight in how to lower the health expenditures by understanding the underlying mechanism. We can also inform policymakers or other persons of interest in the development of obesity and the negative aspects for individuals and society. Potentially, this thesis might proof that the social environment appears to have a moderating effect on the development of obesity. Policy makers could invest more into policies regarding improving the social economic status and decreasing social inequalities to prevent the amount of overweight and obese people.

5.2 STRUCTURE OF THESIS

In the previous chapters, we provided the background and the relevance of my research. The rest of the thesis consists five chapters. Firstly, chapter six provides the theoretical framework which is separated in four sub-chapters: The fetal environment, programming, birthweight and obesity; Genes, epigenetics, birthweight and obesity; The (social) environment; Synthesis of theories. All sub-chapters have the aim to provide theories and explanations for the relation of birthweight to obesity with focus on genetics and, shortly discussed, the environment. These theories form the base for the research and the conceptual model summarizes them. The chapter concludes with the research question and hypotheses. Further, chapter seven gives information on the dataset, participants and statistical methods. In chapter eight we show the results, describe the most important findings and compare findings of twin 1 and 2. The results are further discussed in chapter nine by comparing the results with other research, pointing out limitations and suggestion for future research. In conclusion chapter ten summarizes the findings and answers the research question (TwinLife, 2018)

6 THEORETICAL FRAMEWORK & CONCEPTUAL MODEL

The development of obesity is not only influenced by individual choices. The complexity of the mechanism is expressed of many overlapping theories which in this section I will try to distinct from each other. In the coming sections, the aim is to provide theories which try to explain the association between birthweight and obesity. A major determinant for excessive weight gain in childhood and after lays already in our environment before birth, the in-utero environment. The first set of theories will, therefore, focus on fetal programming and the fetal environment. Nevertheless, referring to the title of this thesis, also our genes are theorized to play a role. The second set of theories will focus on the genetic evidence for (higher susceptibility of) obesity development. These risks that stem from the in-utero environment and genetic make-up may be moderated by influences after we are born by factors from the obesogenic environment. This background will help to later determine the importance or intertwined complexity of in-utero, genetic and the environmental effects leading to obesity in later life. To summarize the theories will be presented in the conceptual model.

6.1 THEORETICAL FRAMEWORK

Previous research indicates that birthweight is related to obesity development. Studies show direct relations of high birthweight to elevated BMI in later life and that this relation has been stable in the past 48 years (Oken & Gillman, 2003; Rugholm et al., 2005; Zhao, Wang, Mu, & Sheng, 2012). By way of illustration, a twin study from East Flanders found that a one kilo increase of birthweight increases a bodyweight with 4.2 kilo and raises the BMI with 0.49 kg/m^2 within males (Loos, Beunen, Fagard, Derom, & Vlietinck, 2001). However, not all studies into birthweight found clear relations. In a study among British twins from Birmingham, the research did not find a relation between birthweight and blood pressure or glucose tolerance. Neither did their results indicate that shared genetic factors were present which determine fetal growth (Baird et al., 2001).

Multiple mechanisms have been indicated to underly this complex relation of birth weight and obesity development in later life. These mechanisms can roughly be separated in to two explanations. The first explanation focusses on fetal programming and experiences of the foetus before being born. The second group of theories explain the associations with one or more genes which are present in the foetus. These genes influence the development of the foetus and potentially have effect during growth or development in later life (Oken & Gillman, 2003).

6.1.1 The fetal environment, programming, birthweight and obesity

Firstly, the fetal environment and programming has been theorized to be of great importance in influencing birthweight and later BMI. The fetal environment covers the environment experimented by the twins before being born. Fetal programming is the change in the function or structure of the foetus which will persist during the development in-utero and even after. It is often caused by an external disadvantageous influence during the development of the foetus (Baird et al., 2001). The important underlying theories for fetal programming are "*the fetal origin hypothesis*" and "*thrifty genotype*". The fetal origin hypothesis was proposed first by D.J. Barker. It entailed that a disadvantageous in-utero environment permanently changes the physiology of a person. These permanent changes cause metabolic diseases and obesity (Barker, 1990; Morgan et al., 2010).

An example of a disadvantageous in-utero environments is expressed by the obese mother (Kleiser, Rosario, Mensink, Prinz-Langenohl, & Kurth, 2009; Ong, 2006). Multiple studies from different regions in the world showed that mothers with diabetes often have babies with relatively higher birthweights and their children also have higher BMI (Oken & Gillman, 2003). For mothers with gestational diabetes, compared to mothers without diabetes, babies are born with more fat mass and have higher likeliness of getting type-2-diabetes in later life (The et al., 2010). Another mechanism proposed to illustrate to influence of unfavourable fetal environment on obesity development in later life is the importance of gestational maternal hypertension. The high blood pressure leads to restrain growth in-utero. Proposed is that this restraint growth in-utero causes catch-up growth after birth which again related to high BMI in later life (Ong, 2006; The et al., 2010; Zhao et al., 2012). A study confirmed this relationship by associating high blood pressure of the foetus in different life stages to low birth weight and increased growth-rates after birth (Huxley, Shiell, & Law, 2000). Hence, obese mothers create disadvantageous fetal environments restricting the development of a healthy foetus and influencing birthweight.

Furthermore, the thrifty genotype hypothesis broadens the fetal origin hypothesis by adding details on how the foetus is programmed. The thrifty genotype hypothesis can therefore explain the potential mechanism developing to metabolic disease or type-2-diabetes, two diseases related to obesity. Most importantly underlying this hypothesis is the observation of undernutrition during the in-utero period. After birth, however, there is an abundance of food available. The hypothesis states that due to the shortages of food during the pregnancy, the foetus body is not programmed to deal with continuous streams of food (Hales & Barker, 2001; Morgan et al., 2010). A famous study conducted during the Hunger Winter in the Netherlands was only able to partly support the hypothesis. During the second world war between the end of the German Siege and the liberation in 1945, the Northern part of the Netherlands had large shortages of food. The proposed effect of developing obesity later in life due to malnutrition during the pregnancy was found for foetus who suffered malnutrition in their first trimester, but not for foetus who suffered malnutrition after this phase (Ribeiro, de Carvalho Lima, de Lira, Pedro Israel Cabral, & da Silva, Giselia Alves Pontes, 2015).

In conclusion, the mechanism for the association between birthweight and obesity development in later life is complex and evidence is not in agreement with each other. However, most studies do indicate that high birthweight is related to obesity, in the form of elevated BMI, in later life. Low birth weight has been associated with obesity related diseases and fast catch-up growth rates after birth. The fetal origin hypothesis and thrifty genotype hypothesis offer a part of the explanation, but are not yet able to explain the entire mechanism of obesity in later life.

6.1.2 Genes, epigenetics, birthweight and obesity

Besides the pathways indicate above, (non-syndromic) obesity has been estimated to have a high heritability between 40 and 70% (Herrera, Keildson, & Lindgren, 2011). Twin studies can help estimate the heritability of diseases. A previous twin study by Loos et al. in 2001 provides evidence for the influence of fetoplacental environment on obesity development, because of difference in birthweight and body composition within monozygotic twin pairs. By comparing the results from monozygotic and dizygotic twins, they propose that birthweight and obesity development are potentially influenced by a person’s genes, next to the influences of fetoplacental environment. The differences for dizygotic twins also appeared to be larger, indicating the contribution of genes. Hence, the second group of explanations focuses on the for presence of certain genes in the foetus which can cause obesity in later life.

Studies specific on the human genome have been able to connect certain genes or genetic variants which are associated with fat distributions or obesity. Estimates, which predict the heritability of birthweight, have been proposed to be between 10% and 40% (Yaghootkar & Freathy, 2012). A potential underlying theory which combines genes and birth weight to obesity is the "*fetal insulin hypothesis*". The hypothesis states that the foetus inherits insulin resistance principally through genes, affecting the birthweight, body composition, and risk on diabetes later in life. Already in the in-utero environment insulin-mediated growth is affected due to genetic factors in the foetus which regulate insulin secretion which is crucial in the growth of a foetus. By also potentially affecting the tissues which are sensitive for insulin, foetus is born smaller and low birthweight becomes a measure for insulin resistance in life (Hattersley & Tooke, 1999; Yaghootkar & Freathy, 2012). In 2001, a twin-study was not able to confirm this hypothesis. The within twin-pair difference for monozygotic and dizygotic twins were observed in the birthweight. Nevertheless, these birthweights did not correlate with blood pressure and glucose tolerance in later life. Importantly, in later life the blood pressure and glucose tolerance of monozygotic as higher correlation in monozygotic twins than dizygotic twins. All in all, this indicates that blood pressure and glucose tolerance are genetically transferred, but there is no significant correlation with birthweight (Baird, Osmond, MacGregor, 2001). Hence, the scientific evidence, regarding the fetal insulin hypothesis, is not consistent evidence.

Further research has been able to find, next to the AGT-gene, genes which are associated with obesity or BMI. One of the first genes to be related was the Fat Mass and Obesity (FTO) gene was related to BMI. More research is still required to understand the role of the gene. Sadly, most studies have not been able to be replicated (Herrera et al., 2011; Naukkarinen et al., 2012). Another review focused on showing the found associations between genetics variants and birthweight. One of the genes presented was the GCK-gene. If the mother carried the mutations and foetus did not, the birthweight was 600 grams higher compared to no-carriers. Vice versa if the foetus was carrying the mutation and the mother was not the birthweight was 500 grams lighter. When the mother and foetus contrastingly where carrier and/or not carrier, no effect compared to mothers without mutation was found. Additionally, other findings reported even more potential candidate genes. Differently from the GCK-gene, some of these genes only must be present in the foetus to affect the birthweight. Hence, studies have been able to connect genes to obesity and birthweight (Yaghootkar & Freathy, 2012). More research is needed to replicate data and find out the underlying mechanisms.

However, genes are only able to partly explain the heritability of obesity. Proposed is the importance the epigenetics of obesity. In epigenetics the DNA sequence in the cell nucleus is not changed, but it affects the expression of certain genes. Epigenetics is thought to already take place during the critical development periods of the foetus, namely exposure to certain environments can affect epigenetic marks which cause obesity (Herrera et al., 2011). The GHRL-gene is a proposed example of a gene which is influenced by environmental factors. The gene codes for the hormone ghrelin which plays an important role in regulating food intake. Within a twin pair one both twins can be carriers, but have their genes respond differently to the environment (Naukkarinen et al., 2012).

6.1.3 The obesogenic environment

In the current pandemic, the environment is proposed to play an important role. This important role of the environment observed in the past decade where a shift in psychical activity and food patterns of populations occurred. Diets have shifted to a high consumption of "*empty*" calories (Giskes et al., 2007). Examples of these foods containing "*empty*" calories are fast-food and soft drinks which contain little nutrients and lots of fats or sugars. These changes are accompanied by reduction in physical activity and further accelerated by changes in the environment (Dubois et al., 2012; Giskes et al., 2007). However, an important issue for finding the mechanism underlying obesity is that there is no agreement on a definition for the obesogenic environment (Kirk, Penney, & McHugh, 2010). One of the first to define the "*obesogenicity"* of environment was Swinburg he defined it as: *"the sum of influences that surroundings, opportunities, or conditions of live have on promoting obesity in individuals or populations*" (Swinburn, Egger, & Raza, 1999). Shortly, the obesogenic environment can be interpreted as all external factors and influences which promote unhealthy choices among individuals leading to obesity. Examples of how these influences or factors of the external environment are: household income (Kinge & Morris, 2010), social economic status or the built environment.

To express the social origin or environment of subjects, researchers often use the Social Economic Status (SES). Studies have been able to associate low SES to obesity development (Oken & Gillman, 2003; Zhao et al., 2012). Low social economic status during the childhood has been inversely related to BMI and waist-hip ratio, another measurement for fat distribution (Poulton et al., 2002). Additionally, a study among German children and adolescent showed that mothers from lower social economic status are positively associated with children with obesity development. In this study to categorize the SES of the participants is a combination of education, occupation and income of parents (Kleiser et al., 2009). Studies looking at the SES are very important, because they can guide policymakers into how to create more equal health outcomes.

Besides the SES, studies have also found other factors in the external environment which influence the association of birthweight and obesity. Smoking mothers have babies which are smaller and, thus, have babies with lower birthweight. However, the writers point out that smoking often runs in a family and person who smoke are more likely to have lower BMI's (Oken & Gillman, 2003). Therefore, if maternal smoking has the potential of working through the same mechanisms, as discussed in the previous sub-chapters, is unclear. Nevertheless, a study among German children and adolescent showed that overweight and obesity were associated positively with maternal smoking. Interestingly, this study also found positive relations between obesity and the following outcomes: maternal weight gain during pregnancy; high birthweight, high media consumption, migration background and low SES (Kleiser et al., 2009). Further, in a review from 2010, the scientist proposes nutritional status, lifestyle, low level of education, parity and gestational age to increase the risk for obesity and influence the birthweight (Zhao et al., 2012).

All in all, the obesogenic environment is a broad concept describing all external influences on individuals or populations. Previous studies and reviews have shown that a wide variety of external factors are related with obesity and/or low birthweight. Therefore, it will be important to also consider the effect of the environment in this thesis.

6.1.4 Synthesis of theories

Previous studies have been able to proof the association of deviating birthweight to either obesity or obesity related diseases. High birthweight is most often associated with higher BMI, while low birthweight is often associated with obesity related disease. The mechanisms underlying this association are still unclear, but there are theories with potential explanations. In this framework, we separated the theories in two groups.

The first group of explanations concerning fetal programming and in-utero experiences. The underlying theories for fetal programming are "the fetal origin hypothesis" and "thrifty genotype". Both focus on disadvantageous experience in-utero which have permanent effects during the life time, examples of these experience are obese mothers, mothers with diabetes, exposure to certain hormones and malnutrition and/or undernutrition.

The second group of explanations focus on the genetic information in the foetus and how these genes interact with the environment. One of the underlying theories is "fetal insulin hypothesis" which focus on the heritability of genes which are important for insulin. Furthermore, research has been able to find different genes connected to birthweight and or obesity. Next to these theories, also was pointed out that after birth babies with low birthweight can experience catch-up growth. It is unclear if low birthweight or the catch-up growth leads to unfavourable health outcomes later in life.

The theoretical framework was concluded by looking at the external environment which is experienced in-utero, as well as, during the life time and influences the development of obesity in later life. The external environment can be described as the obesogenic environment. The social economic status experienced during childhood and later in life is an important predicter for obesity and other health outcomes. Low social status is connected to higher BMI and obesity. Other factors in the external environment which might have a role in predicting obesity in later life are: education, lifestyle, parity, and migration background.

All in all, there are many theories which add or broaden each other. Both genetics, in-utero experiences, and the obesogenic environment seem to be important in determining birthweight and obesity in later life. The mechanism underlying obesity is most likely a combination of genetics and the external environment in-utero and during the lifetime. A metaphor used in the paper of Lobstein et al. (2004) describes the relation: the genetic origin *"loads the gun"* and the environment *"pulls the trigger"* (pg. 5). Nevertheless, the entire mechanism of birth weight and obesity was not able to be discovered.

6.2 CONCEPTUAL MODEL

The conceptual model presented by TwinLife performed as the starting point for a conceptual model (TwinLife, 2018). The model of TwinLife showed how the genetic origin and social origin influence areas of social inequality. Within this master thesis, we focus on the relation between birthweight and obesity with special attention for the importance of genetics and environmental effects. The theories discussed in the theoretical framework fits the conceptual model to the current research.

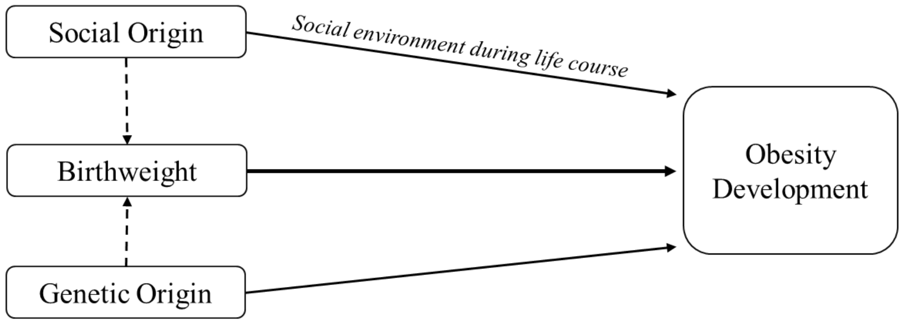

Figure 1: Conceptual model of the relation between birthweight, genetic origin and social origin to development of obesity

In figure 1, the conceptual model shows the influence of birthweight, social origin (environment), and genetic origin on obesity development. Additionally, the line for genetic origin to birthweight is dashed due to it becoming difficult to differentiate between genetic effects and in-utero effects. In-utero effects as explained in the theoretical framework can also come from the social environment and behaviours of the mother, as for instance smoking.

For this thesis, we decided not to look at all the different potential explaining factors which have been researched. Obesity and overweight are highly interdisciplinary and well researched topics. For this thesis, the variable of interest is birthweight, but the importance of inherited factors or the social environments are assessed by adding controlling variable. The social origin is expressed in the variables: social economic status, measured in education of mother, and the age of mother at birth. Hence, for the (social) environment we decided to focus on the family as main source of influence Lastly, for genetic origin an important variable is sex of the twins. Further operationalization of variables can be found in chapter 7. Figure two illustrates the result of adding the variables in the conceptual model.

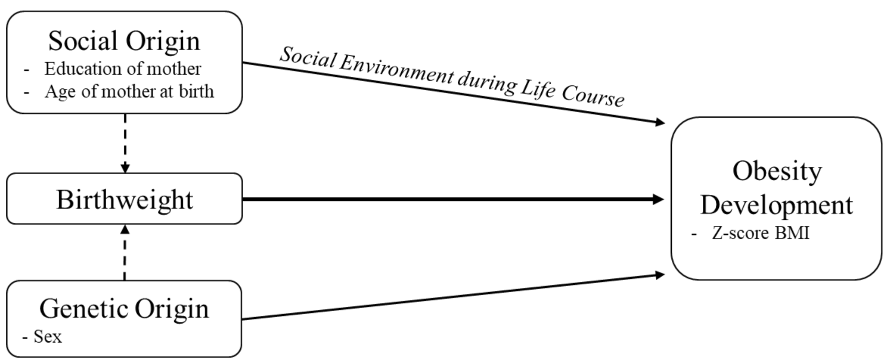

Figure 2: Conceptual model with variables for regressions

6.3 OBJECTIVE AND RESEARCH QUESTION

The objective of this study is to test if there is an association between birthweight and obesity development later in life among German Twins from the TwinLife dataset. The aim is to gain insight into the importance of shared genetic factors and the shared social origin of obesity development in later life. The objective can be re-formulated in the following research question:

> *What is the importance of inherited factors and shared factors from the environment in the association between birthweight and the development of (childhood)obesity among German twins born between 1990 and 2010?*

To answer the research question, four sub-research question with accompanying hypothesis are formulated. The sub-questions are:

- Is there an association between deviating birthweight to higher BMI?
  - *Accompanying assumptions:* Twins with higher birthweight will have higher BMI

- Is there a difference between monozygotic and dizygotic twins and the importance of inherited or social factors in the association of birthweight and obesity?
  - *Accompanying assumptions:* Dizygotic twins can act as comparison for monozygotic twins and if genetics play an important role. Birthweight difference will be smaller among monozygotic twins and so will the difference between twins expressed by genetics (and shared environmental) factors in birthweight strongly related to BMI.

- What is the importance of inherited and shared factors from the environment acting through birthweight among twins related to obesity development?
  - *Accompanying assumptions:* If inherited and shared factors from the environment are not important, there will be larger differences in birthweight and it is associated with higher BMI.

- What is the role of the social environment on to obesity development between twin-pairs?
  - *Accompanying assumptions:* The social environment will have a moderating effect on obesity development.

7 DATA & METHODS

In the following chapter the dataset and its quality, the participants and the methods of statistical analyses of this twin study are outlined. The study contributes to the study areas of public health, obesity development, twin studies and health inequalities.

7.1 DATASET AND THE QUALITY

In this thesis, we use the twin data from the German TwinLife project or the full name of the project: TwinLife – A genetically informatie, longitudinal study about the development of social inequality (link to data, see reference: (GESIS TwinLife, 2018)). The dataset follows a longitudinal and cross-sequential design with not solely information on twins, but also on family members. In total there are over 4000 families with twin pairs included in the dataset. The TwinLife is a DFG-funded long-term project with the aim to study genetic and social causes of chances in life and social inequalities.

The dataset is based on the answers from home- and telephone interviews. Therefore, the data is based on answers of the participants themselves. For children under the age of 13 parents were asked to fill in the questionnaires. Over the age of 13, the participants filled in the questionnaire themselves. In the dataset no information was available for BMI or within-twin pair difference which are therefore calculated by me. Furthermore, crucial to twin studies is the correct classification of the zygosity of twins. All twins over the age of 16 years were asked to fill in a questionnaire were characteristics were asked to determine the zygosity. If twins were younger than 16 years parental supervision was required. During the home interviews, interviews asked 328 twin pairs to participate in a saliva test to collect DNA. Comparison of the saliva test and questionnaires found that the questionnaires scored already high on the correct classification. For parental-reports 97% was correct and for self-reports 92% was correct (Lenau et al., 2017). In the current study we used the variable which combines the questionnaire and saliva test results to limit the number of wrongly classified twins.

7.2 PARTICIPANTS

Specifically, the dataset named: *ZA6701_Family_Wave_v1* is used. In 2014, the first wave of interviews took place. The data in this dataset is consist out of two home interviews and one telephone interview. The wave includes twins from four birth cohorts: 1990-93, 1997/98, 2003/04, and 2009/10 (GESIS TwinLife, 2018). The birth cohorts contain: 188, 291, 281 and 273 number of twin pairs. The age of the twins participating ranges between 4 till 25, for more information see the descriptive statistic table on the next page (table 1). For twins to be included into this research data on BMI (height and weight) and weight at birth must be available for both twins from a twin pair. Sadly, this led to the exclusion of almost half of the twin pairs due to missing data (n = 990). Further, twins who grew up without biological mother (n = 6) were excluded. Lastly, data is missing for the mothers age at birth (n = 1) and education of mother (n = 29). In total this led to the inclusion of 1,002 twin pairs, exclusion of twin pairs was executed by hand.

Table 1: Descriptive Statistics

	Complete sample (N = 1,002)			
	Cohort 90-93	Cohort 97/98	Cohort 03/04	Cohort 09/10
Twin pairs	181	285	270	266
Zygosity				
Monozygotic	105	138	105	111
Dizygotic	76	147	165	155
Sex				
Female	103	142	138	131
Male	78	143	132	135
Twin pair age (years)	23.50 (0.56)	16.98 (0.28)	11.03 (0.33)	5.06 (0.38)
Twin 1: Birthweight (grams)	2476.40 (534.56)	2435.31 (509.15)	2446.33 (576.14)	2447.55 (516.60)
Twin 2: Birthweight (grams)	2422.82 (555.26)	2359.13 (514.32)	2414.28 (554.03)	2398.43 (521.95)
Twin 1: BMI (kg/m^2)	22.75 (3.65)	21.52 (3.08)	17.34 (3.39)	14.89 (1.73)
Twin 2: BMI (kg/m^2)	22.73 (3.81)	21.37 (3.48)	17.08 (2.88)	14.90 (1.87)
Twin 1: z-score BMI	0.0174 (1.0089)	-0.0008 (1.0063)	-0.0166 (0.9819)	-0.0229 (0.9688)
Twin 2: z-score BMI	0.0123 (1.0148)	-0.0042 (1.0065)	-0.0277 (0.9724)	-0.0259 (0.9741)
Education Mother [a]				
Low	40	31	35	30
Medium	71	114	87	76
High	70	140	148	160
Age of mother at birth (years)	29.97 (4.11)	30.54 (4.31)	31.98 (4.89)	32.06 (4.92)

[a] The education of mothers refers to the following categories: Low = left school without qualification and leaving school with primary/lower school qualification (until grade 8); Medium = leaving school with secondary/intermediate qualification (until grade 10); High = university of applied sciences diploma (technical school qualification) and upper secondary diploma/university (abitur) entrance diploma.

7.3 METHODS: STATISTICAL ANALYSES

To find the association interested in this quantitative study we will explore the associations between birthweight, the genetic origin and social origin on obesity development. For the quantitative analysis the program Stata/ES 15 was used. My analysis is based on the standard twin regression techniques as discussed by Carlin, et al. (2005) in the paper *"Regression models for twin studies: a critical review"*. The method used in the current thesis consist out of three steps: standard OLS with normal standard errors, OLS with co-twin X-value and robust standard errors and the addition of strata. The robust standard errors are added in the regression to account for the clustering in the observations. OLS normally assumes that all observations are independent, however, in the case of twins the observations of a twin pair are clustered. The three steps will be further described in detail below.

7.3.1 Operationalization: independent & dependent variables

The main predictor is birthweight, because of the interest in the effect of birthweight on later BMI. In general birthweight will be treated as a continuous variable. However, for stratifications there are three different categories based on the definitions of the World Health Organization (WHO). The WHO differentiates between six low and high birth weight categories. There are four low birthweight categories, respectively: Extremely Low Birth Weight (ELBW) from 499 (or less) till 999 grams; Very Low Birth

Weight (VLBW) from 1000 till 1499 grams; and Low Birth Weight (LBW) from 1500 till 2499 grams. High birthweight is separated in two categories: Large Newborn from 4000 till 4499 grams and Exceptionally Large Newborn more than 4500 grams[1]. No twins with a high birthweight are included in the sample due to the heaviest twin being 3910 grams. The three categories in this thesis are ranging from: 499-1499 grams, 1500-2499, and 2499-3910.

Furthermore, the associations are controlled for the influence of the following variables: sex (male/female), social economic status (SES) (expressed as education of mother), age of mother at birth (in years), and birth cohort. The education of mother in the dataset consist out of six categories. To have more observations in one category, the variable is re-categorized in three groups: no/low education which also included six participants which filled out other education; middle education; lastly, high education; The continuous variable of age of mother is re-categorized in young mothers aged 18-24 years; mothers aged 25-34 years; and mothers older than 35 years. Mothers whom have babies after the age of 35 are often considered to be at risk pregnancies. Not all independent variables are readily available in the dataset and are calculated by hand.

The dependent variable of the research is the body mass index (BMI). The BMI is defined as the body mass divided by the square of the body weight (kg/m^2) and had to be calculated. For adults to be considered overweight the BMI must be higher than 25 kg/m^2 and to be considered obese BMI must be higher than 30 kg/m^2. For children between the ages 5 and 19 years often an adapted BMI measure is used to determine obesity and overweight. The WHO defines overweight in children as a standard deviation of 1 above the WHO growth reference median (World Health Organization, 2018). In this thesis a simplistic version of this formula is used to calculated z-scores for BMI per cohort. The z-scores are calculated as followed:

$$z - Score\ BMI_{ij} = \frac{X_{ij} - \mu_{cohortA}}{\sigma_{cohortA}} \qquad (f1)$$

In this formula X_{ij} is the BMI of a twin (i) from a specific twin pair (j). The $\mu_{cohortA}$ is the mean of one of the four cohorts (cohort a). Lastly, the $\sigma_{cohortA}$ represent the standard deviation of one of the specific cohorts. To be considered obese in the two youngest cohorts the found z-score must be ≥ 2.0 and to be considered overweight z-score must ≥ 1. In the two older cohorts, the z-scores come close to the normally used cut-off points for obesity and overweight. The z-score of 2.0 is equal to BMI ± 29.5 and a z-score of 1 is equal to BMI ± 25.5. By using z-score it becomes possible to use the entire sample at the same time in the regression.

7.3.2 OLS: Twins as individuals

The first analysis starts with the most straightforward approach which considers the twins as normal individuals within a sample. This approach ignores the fact that the nature of the data is paired, due to two twins being in one twin pair. The error term in this approach is, therefore, normal and not robust. The aim of the approach is to find if within the sample there is an association between the dependent variable, z-score BMI, and the independent variable, birthweight. By performing an Ordinary Least Squares (OLS) regression, the β_0 and β_c are estimated. The formula used in this approach is:

[1] World Health Organization: ICD-11. (2018). ICD-11 for mortality and morbidity statistics. Retrieved from https://icd.who.int/browse11/l-m/en - corresponding codes in ICD-11: KA21 and KA22

$$z - Score\ BMI\ (Y_{ij}) = \beta_0 + \beta_c X_{ij} + \varepsilon \qquad (f2)$$

In this formula the dependent variable is Y_{ij} is z-score BMI. The independent variable is β_c presents the birthweight. If β_c describes the increase of Y_{ij} if X_{ij} increases with one unit. In the Y and X values ij describes the twin pair (i) and the individual within the twin pair (j). Hence, to determine the role of birthweight on BMI, the first model estimates (1a) with dependent variable z-score BMI and independent variable birthweight is added. Nevertheless, the results from this global approach are useful due to the ability to show if there is a relation between BMI and birthweight.

To potentially further extend and improve the understanding of the mechanism of obesity, we control the association of birthweight and obesity with the other variables in our dataset. The formula for this approach is the same and the results are often referred to as the elaborated model (1b). In the (1b) second model the following independent variables are added: birthweight (in grams), sex (male/female), social economic status (SES) (= education of mother), age of mother at birth (years). By adding these variables to the simple regression, insight in between twin pair difference and the importance of the other variables raises.

7.3.3 OLS: Including the co-twin X-value

The second approach is more complex and has the potential to distinct between environmental, as well, as genetic effects. The paired nature of twins is no longer ignored and considered by including a co-twin X-value. In the formula this looks as followed:

$$BMI\ (Y_{ij}) = \beta_0 + \beta_w (X_{ij} - \bar{X}_i) + \beta_B \bar{X}_i + \varepsilon_r \qquad (f3)$$

$$BMI\ (Y_{twinpair10;twin1}) = \beta_0 + \beta_w (X_{twinpair10;twin1} - \bar{X}_{twinpair10}) + \beta_B \bar{X}_{twinpair10} + \varepsilon_r$$

In this formula the dependent variable is again Y_{ij}, z-score BMI and X_{ij} describes the independent variables. The co-twin X-value is described by $\beta_w (X_{ij} - \bar{X}_i)$. The X_{ij} describes the birthweight, X-value, of a specific twin (j) within a twin pair (i) and the $\bar{X}_i$ describes the average birthweight, X-value, of the twin pair (i). An example is given in formula ($f3$). The interpretation follows by having an increase of one-unit increase for the within twin pair difference ($X_{ij} - \bar{X}_i$), the Y_{ij} is expected to increase with the value of β_w. Hence, $\beta_w (X_{ij} - \bar{X}_i)$ reflects the “within effect” or in other words the differences within a specific twin pair.

On the other hand, the $\beta_B \bar{X}_i$ describes the “between effect”. It is an indication for what one twin pair shares and is different from other twin pairs. The interpretation of this variables is that when ($\bar{X}_i$) increases with one-unit, the Y_{ij} is expected to increase with the value of β_B. Lastly, the β_0 is the constant and ε_r is the robust standard error. In this paper, the $\overline{(X_i)}$ is expressed by the control variables of sex, age of mother at birth and SES (= education of mother). For twins in a twin pair, except for sex, these variables are shared and have the same value. Therefore, the average did not have to be calculated separately Between twin pairs the values of these variables are different.

In the first (2a) model, only includes the within pair birthweight difference (β_w). No other potential explaining variables are included, because it is assumed that the environment for twin pairs is the same during their development. Therefore, the environment would equal out between twins if it was included.

The second model (2b) includes the within pair birthweight difference (β_W) and between pair (β_B). The between pair differences are expressed by the following independent variables: sex (male/female), social economic status (SES) (= education of mother), age of mother at birth (years).

Additionally, to be able to check for the influences of differences between twin pairs a stratified approach was taken. As discussed in the theoretical framework, previous research shows that the obesity and birthweight relation might differ by age of mother at birth, sex and social economic status (= education of mother). Furthermore, the two regressions are executed for each birth cohort separately. Previous research has also compared outcomes from mono- and dizygotic twin. Although it is not possible with this method to compare results into detail, finding significant results for one of the two groups will add to explaining the association of birthweight and obesity in later life.

7.4 ETHICAL CONSIDERATIONS

The data used in this thesis are not collected by me personally, but by the TwinLife project (GESIS TwinLife, 2018). The TwinLife project ensured that the questionnaires for children under the age of 13 were filled by a parent. Furthermore, no names or living address are part of the dataset. Making the data as anonymous as possible. My supervisor Tobias Vogt had access to the TwinLife data and shared this with me. I promised and ensured that I use the data collected with care. I will not share the data and only use it for the current thesis. After this thesis, I will delete any data from the dataset.

8 RESULTS

The regression analysis finds when treating the twins as individuals that birthweight is significantly associated with the z-score BMI. When this association is controlled for other independent variables (sex, mothers' age at birth and education of mother), this association stays significant. Further, stratifying the regression for birthweight finds that individuals with a birthweight between 2499 – 3910 grams the associations between birthweight and z-score BMI is significant and not in the other strata of low birthweight. Hence, birthweight can give an indication of obesity development in later life. Specifically, for in babies born with a birthweight between 2499 – 3910 grams (table 1 and 2).

Secondly, to observe if associations are influenced by shared genetic factors (or the environment) the paired structure of the data is used. The paired structure of the data, or in other words the within twin difference, is expressed in the between X-value. Overall, for twin 1 and twin 2, there are no significant association found between X-value and z-score BMI (table 3). However, making strata's finds significant associations between the within twin X-value and dependent variable. In short, for twin 1 the association between twin X-value and z-score BMI is significant in the strata when babies are born with extremely/very low birthweight (only elaborated model, table 5), mothers' education is medium (simple and elaborated model, table 7) and when twin 1 is born from young mothers (only elaborated model, table 8). For twin 2 the associations found to be significant when: mothers' education is high (simple and elaborated model, table 7), twin 2 is born in older mothers 35+ (simple and elaborated model, table 8) and among dizygotic twins (simple and elaborated model, table 9).

In the following paragraphs these results will be explained, shortly, accompanied by the tables. The tables contain the information for twin 1 and accompanied with a textual comparison, the results of twin 2 are presented.

8.1 TWINS AS INDIVIDUALS

When treating twins as individuals, ignoring the paired structure of the data, both the simple (model 1a)

Table 2: OLS Regression treating twins as individuals, simple & elaborated model

	Complete Sample			
	Model 1a		Model 1b	
Birthweight (grams)	0.00025***	(0.00004)	0.00025***	(0.00004)
Sex *(Reference: male)*				
Female			-0.20464***	(0.04351)
Mothers' age at birth				
(Reference: Mother aged 18 - 24)				
Mother aged 25 – 34			-0.29158***	(0.08271)
Mother aged 35 +			-0.26677**	(0.09165)
Education of mother				
(*Reference: no/low education)*				
Middle education			-0.09947	(0.06964)
High education			-0.20416**	(0.06749)
Constant	-0.60681***	(0.10161)	-0.10120	(0.13271)
R^2	0.0177		0.0416	
N	2,004		2,004	

[a] The table includes: Coefficients (Standard Error, SE) [b] Significance is indicated by an asterisk. The asterisks have the following meaning * p < 0.05; ** p < 0.01; *** p < 0.001

The association in both models can be interpreted as followed: if the birthweight increases with 1 gram than the z-score BMI will increase with 0.00025. If the associations are controlled for variables (model 1b, referred to as elaborated model) which can potentially influence the found association, the association of birthweight and z-score BMI stays significant (table 1).

Furthermore, the R^2 gives an indication of how well the independent variables can explain the variance in the dependent variable, here: BMI z-score. Although the R^2 in both models is quite low, respectively 1.77% in the simple model and 4.16% in the elaborated model. It is important to note that adding the controlling variables the predictability of variance in the dependent variable increases.

Other significant findings are observed in the elaborated model. Being female compared to being male has a negative effect on z-score BMI. Additionally, mothers aged 25-34 and 35+ compared to young mothers have a negative effect on z-score BMI. Lastly, high education compared to no/low education finds a negative significant effect. The accompanying interpretation is that mothers with a high education compared to no/low education are associated with children with a lower z-score BMI and, thus, lower obesity. It is important to note that all significant found associations indicate a correlation and not causation. There is no control group when treating twins as individuals, making the regression not possible to indicate causation.

8.1.1 Twins as individuals, stratified for birthweight categories

Stratifying the regression for birthweight, merely shows in the category of a normal birthweight (2500-3910 grams) a positive significant association between birthweight and z-score BMI. This significant association is both observed in the simple and elaborated model.

Table 3: OLS Regression treating twins as individuals, stratified by birthweight

	Complete Sample		
	Extremely/Very Low birthweight (630-1499 grams)	Low birthweight (1500-2499 grams)	Normal birthweight (2500-3910 grams)
Birthweight (grams)	0.00030 (0.00035)	0.00016 (0.00013)	0.00032** (0.00011)
Constant	-0.699882 (0.42158)	-0.40571 (0.27488)	-0.82477* (0.32761)
R^2	0.0069	0.0015	0.0086
Birthweight (grams)	0.00008 (0.00035)	0.00014 (0.00013)	0.00315** (0.00011)
Sex *(Reference: male)*			
Female	-0.20177 (0.16944)	-2.67352*** (0.06449)	-0.13801* (0.06280)
Mothers' age at birth *(Reference: Mother aged 18 - 24)*			
Mother aged 25 – 34	-0.23535 (0.31366)	-0.46610*** (0.11273)	-0.04256 (0.13300)
Mother aged 35 +	-0.15896 (0.35270)	-0.46168*** (0.12755)	-0.00053 (0.14441)
Education of mother (R*eference: no/low education)*			
Middle education	0.27495 (0.28167)	-0.06399 (0.09710)	-0.21194* (0.10838)
High education	0.57670* (0.28478)	-0.15223 (0.09418)	-0.37847***(0.10418)
Constant	-0.51390 (0.53042)	0.29090 (0.29622)	
R^2	0.0679	0.0421	0.0313
N	111	986	925

Significance is indicated by an asterisk. The asterisks have the following meaning * p < 0.051; ** p < 0.01; *** p < 0.001

8.1.1.1 Comparing to Twin 2

Furthermore, comparing the elaborated models shows that per strata different independent variables have a significant effect on the dependent variable. In the extremely/very low birthweight strata a higher education compared to no/low education is significantly negatively associated with z-score BMI. In the low birthweight strata being female compared to being male has a strong negative impact. In addition, mothers aged 25-34 and 35+ compared to young mothers are negatively associated with a high BMI. Lastly, next to birthweight in the strata of normal birthweight, female compared to being male and middle and high education compared to no/low education are negatively associated with z-score BMI.

8.2 TWINS WITH BETWEEN TWIN X-VALUE

In the second part of the analysis the within twin pair X-value is added to give insight into the importance of inherited factors, as well as, the environment. The between twin X-value describes the difference of one twin minus the average of a twin pair. This subtraction makes the X-value free from confounding effects. In the simplistic model, 2a, no significant finding is found between z-score BMI and the between twin X-value. This can be interpreted as that the difference between the twin within twin pair is not significantly associated with explaining the z-score BMI of a twin. Hence, meaning that shared inherited factors and/or experiences between twins from a twin pair do not explain z-score BMI. Furthermore, when expanding the model by adding the independent variables which influence the association of z-score BMI. They are all found to be significant.

Neither running the model for the between-pair difference calculated for the twin 2, the simple model (p = 0.081) and elaborated model (p=0.086) are not found to be significant. In the model for twin 2 including the control variables, only the female variable was found to be significant (p=0.000). This could be interpreted as being female compared to being male will decrease the z-BMI score with -0.23694 (SE=0.062876). The constant in this model was also significant (p=0.030).

Table 4: OLS regression with between X-value and robust standard errors (twin 1)

	Model 2a		Model 2b	
Difference in birthweight (grams)	0.00028	(0.00017)	0.00029	(0.00017)
Sex *(Reference: male)*				
Female			-0.21321***	(0.06201)
Mothers' age at birth *(Reference: Mother aged 18 - 24)*				
Mother aged 25 – 34			-0.30073*	(0.14017)
Mother aged 35 +			-0.30364*	(0.14846)
Education of mother (*Reference: no/low education)*				
Middle education			-0.18585	(0.11520)
High education			-0.27740**	(0.10500)
Constant	-0.01500	(0.03094)	0.57973***	(0.17134)
R^2	0.0025		0.0319	
N	1,002		1,002	

Significance is indicated by an asterisk. The asterisks have the following meaning * p < 0.05; ** p < 0.01; *** p < 0.001

8.2.1 Synthesis of twins with between twin X-value stratified

Overall, the regression with between twin X-value is stratified for: birthweight, cohort, education of mother, mother's age at birth and zygosity. Overall the independent variable of interest, the between twin X-value, does not find to be significantly associated with z-score BMI (table 4). This means that the factors which are driving the difference within birthweight in a twin pair are not driving the obesity development in later life, expressed in z-score BMI. However, some strata did find significance between twin X-value and z-score BMI is observed. Specifically, for twin 1 the following strata's find this significant association: extremely/very low birthweight (limited observations 55), mothers with medium education, mothers aged 18-24 (limited observations 79). For twin 2, the strata mothers with high education, mothers aged 35+ and dizygotic twins shows a significant association between the X-value and z-score BMI.

8.2.2 Stratified by birthweight

The variable of interest here is the between X-value of birthweight. In the simple model no associations are found, but in the elaborated model this is observed for the strata of extremely/very low birthweight. Contrastingly, when treating the twins as individual's birthweight was only significant in the strata of normal birthweight. The X-value indicates that factors underlying this difference in birthweight also underly the z-score BMI. The coefficient, however, is negative and means when the difference in birthweight between twins increases the z-score BMI decreases. This could be expected, due to if the difference with extremely low birthweight is high, most likely it will mean that the other twin is heavier. As the literature and treating twins as individual regression showed, normal or higher birthweight are often associated with higher z-score BMI.

Table 5: OLS regression with between X-value and robust standard errors (twin 1), stratified by birthweight

	Complete Sample		
	Extremely/Very Low birthweight (630-1499 grams)	Low birthweight (1500-2499 grams)	Normal birthweight (2500-3910 grams)
Difference in birthweight (grams)	-0.00095 (0.00075)	0.00018 (0.00026)	0.00026 (0.00025)
Constant	-0.32917** (0.11611)	-0.04790 (0.04963)	0.04448 (0.04118)
R^2	0.0301	0.0008	0.0023
Difference in birthweight (grams)	-0.00137* (0.00064)	0.00022 (0.00026)	0.00022 (0.00025)
Sex *(Reference: male)*			
Female	-0.43911 (0.24673)	-0.22204* (0.09858)	-0.17028* (0.08439)
Mothers' age at birth *(Reference: Mother aged 18 - 24)*			
Mother aged 25 – 34	0.39370 (0.27930)	-0.44962* (0.19656)	-0.15668 (0.21849)
Mother aged 35 +	0.42806 (0.32023)	-0.40511 (0.21884)	-0.20908 (0.22351)
Education of mother (R*eference: little education)*			
Middle education	0.30095 (0.33396)	-0.14197 (0.15684)	-0.30608* (0.18772)
High education	0.76676* (0.28978)	-0.28446* (0.13729)	-0.42967** (0.17612)
Constant	-0.95654** (0.29820)	0.65475* (0.26086)	
R^2	0.2181	0.0409	0.0361
N	55	460	487

Significance is indicated by an asterisk. The asterisks have the following meaning * p < 0.05; ** p < 0.01; *** p < 0.001

Furthermore, the elaborated model shows that for extremely/very low and normal birthweight a high education compared to no/low education has a mediating effect on the z-score BMI. In extremely/very low the coefficient is positive leading to the interpretation that high education is associated with a higher z-score BMI. On the other hand, in normal birthweight a higher education has a negative association with z-score BMI.

8.2.2.1 Comparing to Twin 2

Birthweight is not necessarily a factor shared by twins. Therefore, the birthweight of twins can deviate making the strata of different sizes. For twin 2 there are 56 twins with Extremely/Very Low birthweight, 508 twins with low birthweight and 438 with normal birthweight. Running the regression with these strata find no significant findings in the simplistic model. In the elaborated model significance was found in the strata of low and normal birthweight. In low birthweight being female compared to being male finds a negative association to z-score BMI (-0.31657, SE=0.08723, p=0.000). Further, mothers aged 25-34 and 35+ compared to young mothers also find a significant association. The coefficients are both negative -0.47247 (SE=0.19945, p=0.018) for mothers aged 25-34 and -0.51058 (SE=0.21684, p=0.019) for mothers aged 35+ Hence, mothers of older age decrease among their offspring the development of obesity in later life. Lastly, the constant is significant in this stratum (0.53092, SE=0.25569, p=0.038). Although many variables are constant the R^2 indicates that only 5.09% of the variance in the dependent variable. Lastly, in twins with a normal birth weight only the oldest age category shows to have a significant influence. This can be interpreted as mothers of 35+ decrease the z-score BMI with -.031920 (SE-0.16132, p=0.48).

8.2.3 Stratified by cohort

The descriptive statics showed that the mothers' age of birth slowly increased. Additionally, the BMI of the younger cohorts has smaller standard deviations, meaning that the weight of the children is closer together. For these reasons, the birth cohorts are used as strata. Nevertheless, no significant effect is observed between the within twin pair X-value and the dependent variable. Furthermore, in the elaborated model the significant associations to the dependent variable differ per cohort. In the oldest cohort middle education compared to low/no education seems to influence the z-score BMI. Being female compared to being male seems to be associated the development of obesity in later life among the cohorts from 97/98 and 09/10. Lastly, the age of mothers at birth only seems to be significantly associated in the youngest cohort. This could be interpreted as twins born in mothers over the age of 25 years compared to young mothers are associated with a decrease of the z-score BMI.

Table 6: OLS regression with between X-value and robust standard errors (twin 1), stratified by cohort

	Complete Sample			
	Cohort 90-93	Cohort 97/98	Cohort 03/04	Cohort 09/10
Difference in birthweight (grams)	-0.00002 (0.00031)	0.00028 (0.00035)	0.00045 (0.00040)	0.00040 (0.00032)
Constant	0.01782 (0.07510)	-0.01153 (0.05792)	-0.02382 (0.05786)	-0.03266 (0.06125)
R^2	0.0000	0.0021	0.0062	0.0049
Difference in birthweight (grams)	0.00016 (0.00033)	0.00024 (0.00034)	0.00045 (0.00040)	0.00036 (0.00033)
Sex *(Reference: male)*				
Female	-0.27366 (0.14930)	-0.35121** (0.11729)	-0.03261 (0.12383)	-0.22684* (0.11667)
Mothers' age at birth *(Reference: Mother aged 18 - 24)*				
Mother aged 25 – 34	-0.23111 (0.32174)	-0.01124 (0.20870)	-0.37149 (0.29732)	-0.83920** (0.32006)
Mother aged 35 +	-0.05669 (0.38491)	-0.07784 (0.23231)	-0.43781 (0.29640)	-0.78894* (0.33725)
Education of mother (R*eference: little education)*				
Middle education	-0.49582* (0.24354)	-0.01371 (0.21014)	0.00130 (0.21762)	-0.16469 (0.24582)
High education	-0.30434 (0.23903)	-0.32135 (0.18838)	-0.13033 (0.18992)	-0.18609 (0.21900)
Constant	0.66791 (0.42151)	0.34801 (0.27568)	0.43029 (0.26600)	1.00970* (0.41551)
R^2	0.0584	0.0516	0.0249	0.0703
N	181	285	270	266

Significance is indicated by an asterisk. The asterisks have the following meaning * p < 0.053; ** p < 0.01; *** p < 0.001

8.2.3.1 Comparing to Twin 2

Running the model for the second twin, again difference in between X-value for birthweight was not found to be significant. Close to significance came the cohort from twins born in 2009/2010 where in the simplistic model the p-value is 0.061 and in the model with control variables the p-value is 0.055 Additionally, the model including control variables showed that being female compared to male was significant associated with z-score of BMI in the birth cohort of 1990-93 (p=0.023) and the birth cohort of 1997/1998. These regression models did not find other significant results.

8.2.4 Stratified by education

The difference between twins, expressed in the X-value, in birthweight is found to be significantly associated among the group of twins with a mother with a medium education. This association shows in the simple and elaborated model. This finding expresses that the factors which drive the difference in birthweight between twins in a pair may also drive the difference in z-score BMI. Hence, this shows that among mothers with medium education, genetic and environmental factors (as discussed in the theoretical framework) are positively significantly associated with z-score BMI. The interpretation of the coefficient shows that when the difference in birthweight increases, the z-score BMI increases. In the elaborated model for mothers with a medium education, next the between X-value, no independent variables are significantly associated. This could indicate that the effects of sex and mothers' age at birth are mediated by the X-value. Furthermore, in the simple model of the strata low education and high education only the constant is significantly associated. The elaborated model shows that twins born from mothers with a low education and mothers aged 25-34 compared to young mothers, the z-score BMI decreases.

Table 7: OLS regression with between X-value and robust standard errors (twin 1), stratified by education

	Complete Sample		
	Low education	Medium education	High education
Difference in birthweight (grams)	0.00024 (0.00048)	0.00099** (0.00035)	-0.00014 (0.00019)
Constant	0.22279* (0.10123)	0.00423 (0.05775)	-0.08559* (0.03660)
R^2	0.0014	0.0246	0.0009
Difference in birthweight (grams)	0.00021 (0.00047)	0.00104** (0.00035)	-0.00020 (0.00019)
Sex *(Reference: male)*			
Female	-0.19565 (0.20320)	-0.16917 (0.11615)	-0.27701*** (0.07178)
Mothers' age at birth *(Reference: Mother aged 18 - 24)*			
Mother aged 25 – 34	-0.66182* (0.33814)	-0.15173 (0.19456)	-0.21783 (0.16804)
Mother aged 35 +	-0.40680 (0.39947)	-0.21661 (0.22706)	-0.23471 (0.17541)
Constant	0.83339** (0.36346)	0.23522 (0.18327)	0.27592 (0.16933)
R^2	0.0517	0.0327	0.0302
N	136	348	518

Significance is indicated by an asterisk. The asterisks have the following meaning * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

8.2.4.1 Comparing to Twin 2

When running the simplistic regression for the second twin the within-twin pair difference is not significant for low education (p=0.111). Contrastingly to the results from twin one, also medium education is not significantly associated with z-score BMI (p=0.390). However, high education is significantly associated with the dependent variable (p=0.006). The association can be interpreted when the difference in birthweight within a twin pair increases with 1 gram, the z-score BMI increases with 0.00054 (SE=0.00019). In the second model with the controlling variables added, the same pattern can be observed for the within-twin pair difference. Meaning that low education (p=0.111) and medium education (p=0.249) do not find a significant association, but high education is significantly associated (0.00059; SE=0.00019; p=0.002). Furthermore, in both the medium (p=0.025) and high education (p=0.000), being female compared to being male have a negative association with the z-score BMI. The corresponding coefficients are respectively: -0.27198 (SE=0.12109) and -0.25992 (SE=0.07254). No other, significant results were observed.

8.2.5 Stratified by mothers’ age at birth

The regression by strata of mother’s age for twin 1 shows two significant associations in the elaborated model. Firstly, for the between X-value in birthweight among low educated mothers. However, this association must be interpreted with causation, because the finding was not significant in the simplistic model and there are only 79 observations in the strata. Otherwise the interpretation is that the shared factors, genetic or environmental, which drive the difference in birthweight are associated with a higher z-score BMI.

Table 8: OLS regression with between X-value and robust standard errors (twin 1), stratified by mothers' age at birth

	Complete Sample		
	Mother aged 18 – 24	Mother aged 25 - 34	Mother aged 35 +
Difference in birthweight (grams)	0.00108 (0.00063)	0.00038 (0.00021)	-0.00029 (0.00031)
Constant	0.28714* (0.13942)	-0.03866 (0.03597)	-0.045189 (0.06364)
R^2	0.0244	0.0048	0.0033
Difference in birthweight (grams)	0.00119* (0.00058)	0.00041 (0.00021)	-0.00033 (0.00031)
Sex *(Reference: male)*			
Female	-0.31726 (0.28027)	-0.23076** (0.07355)	-0.17657 (0.12055)
Mothers' age at birth *(Reference: Mother aged 18 - 24)*			
Mother aged 25 – 34	-0.56437 (0.37195)	-0.05861 (0.13088)	-0.35332 (0.27629)
Mother aged 35 +	-0.63173 (0.34718)	-0.17006 (0.11931)	-0.41655 (0.24942)
Constant	0.87764* (0.39978)	0.18475 (0.11556)	0.41099 (0.24405)
R^2	0.0919	0.0237	0.0271
N	79	691	232

Significance is indicated by an asterisk. The asterisks have the following meaning * p < 0.05; ** p < 0.01; *** p < 0.001

8.2.5.1 Comparing to Twin 2

Running the regression for the second twin finds in the simple and elaborated model both significant association between difference in birthweight and z-score BMI, solely, for mothers aged 35+. In the simple model this association is 0.00116 (SE=0.00034; p=0.001) and in the elaborated model this associations are 0.00121 (SE=0.00035, p=0.001). This contrast the non-significant findings in twin 1. Other significant findings are observed in the elaborated model for mothers aged 18-24 and 25-34. Firstly, being female compared to being male finds negative significant association to z-score BMI in both categories, respectively: -0.65226 (SE=0.28931; p=0.027) and -0.23581 (SE=0.07286; p=0.001). Lastly, the constant is significant in the category of mothers aged 18-24 (0.81541; SE=0.39999; p=0.045).

8.2.6 Stratified by zygosity

As discussed before monozygotic twins share more genetic material than dizygotic twins. However, in both the simple and elaborated model the between X-value does not show to be significant in any of the two groups. Indicating that the difference in birthweight is not originating from the same factors which determine obesity in later life. In both elaborated models both being female compared to male shows to have a negative significant impact. The association is stronger for monozygotic twins than dizygotic twins potentially with the underlying reason that dizygotic twins are also mixed-sex couples. Lastly, in dizygotic twins the mothers' age at birth and education are all negatively significantly associated.

Table 9: OLS regression with between X-value and robust standard errors (twin 1), stratified by zygosity

	Complete Sample	
	Monozygotic	Dizygotic
Difference in birthweight (grams)	0.00015 (0.00027)	0.00038 (0.00022)
Constant	-0.00654 (0.04595)	-0.02296 (0.04200)
R^2	0.0007	0.0049
Difference in birthweight (grams)	0.00019 (0.00028)	0.00037 (0.00022)
Sex *(Reference: male)*		
Female	-0.25083** (0.09616)	-0.18017* (0.08318)
Mothers' age at birth *(Reference: Mother aged 18 - 24)*		
Mother aged 25 – 34	-0.21699 (0.18391)	-0.44066* (0.22210)
Mother aged 35 +	-0.18551 (0.20288)	-0.46727* (0.23092)
Education of mother (R*eference: no/low education)*		
Middle education	-0.13903 (0.17584)	-0.22163 (0.15234)
High education	-0.19882 (0.15847)	-0.33776* (0.14092)
Constant	0.46425 (0.26221)	0.74285*** (0.22109)
R^2	0.0266	0.0414
N	459	543

Significance is indicated by an asterisk. The asterisks have the following meaning * p < 0.05; ** p < 0.01; *** p < 0.001

8.2.6.1 Comparing to Twin 2

In monozygotic twins the regressions of twin 2 no significant associations to the dependent variable in the simple model. In the elaborate model only being female compared to being male was significant (p=0.003). The results can be interpreted as being female compared to male decreases the z-score BMI with -0.28529 (SE=0.09395).

Running the simplistic model for twin 2 found that among dizygotic twins within-twin pair difference was significant associated with the z-score BMI (p=0.048). When controlling for other variables in the elaborated model still finds within-twin pair difference in birthweight to be significant associated with the dependent variable (p=0.044). This could be interpreted as when the birthweight of twin 2 increases with one gram compared to the average birthweight of the twin pair, the z-score BMI increase with 0.00046 (SE = 0.00023). Furthermore, also being female compared to male was found to be significantly associated with z-score BMI (-0.19990 (0.08522), p=0.019). This contrast the findings of twin 1 where all other variables were significant except from the between X-value.

9 DISCUSSION

From the results, a positive association between birthweight and BMI can be observed. The association is more apparent when treating twins as individuals, than looking at the within twin pair differences. However, association does not mean causation, and, in this case, the casual mechanism of birthweight and BMI cannot be confirmed with the current data or approach. Neither is it possible from this data to perfectly differentiate between genetics and in-utero effects, or, in-utero effects and postnatal circumstance. Birthweight, however, might still be an indicator of genetics and in-utero effects.

In this section of the thesis, we discuss the results starting by the regression of twins as individuals, followed by a discussion on why many results in the section with the stratified regression were stratified not significant and why it is hard to differentiate between in-utero or genetic effects. We decided to start with the regression of twins as individuals, because this gives insight into if there is an association and to follow the structure of the results. Furthermore, we give an overview of the strengths and limitations of this study. Lastly, we supply ideas for future research with attention on how to separate between genetic and in-utero effects.

9.1.1 Discussion: Twins as individuals

The first analysis shows that birthweight is positively significantly associated with BMI in later life, especially for twins born with a normal birthweight. These findings are in line with previous studies which indicated a relation of high birthweight and BMI (Loos, Beunen, Fagard, Derom, & Vlietinck, 2001; Oken & Gillman, 2003; Ong, 2006). Looking with more detail to our sample, the relation when stratified is only significant in individuals with the highest birthweight. Our results corroborate with the findings of Rugholm et al. (2015) who observed that the relative risk on developing overweight during childhood increased significantly for children with birthweights from 3.5 to 4.0 kilogram compared to children with birthweight ranging from 3.0 to 3.5 kg's. Our results can, however, not confirm findings from the same study which describe that children born with a birthweight from 4.0 to 4.5 kg and 4.5 to 5.5 kg, the relative risk increased. For the last group, compared to the reference group the risk of overweight even doubled (Rugholm et al., 2005). These findings cannot be confirmed, because our dataset does not have any cases of twins with such a high birthweight life. Further, a study by The et al. (2010) also showed that individuals born with a high birthweight (>4.0 kg), were more likely to develop obesity than individuals with a birthweight between 2.5 and 4.0 kg if the mother was not obese (The et al., 2010). Again, supporting our findings that high birthweight is associated with increased BMI in childhood and adolescence.

Next to this evidence, it is important to note that effect of birthweight on BMI might be limited, as the R^2 is only 1.77%. A study by Loos et al. in 2001 showed birthweight had a more influential relation with weight and height in later life than with BMI. The paper concludes that albeit BMI was showed to generally increase with increased birthweight, the weight was the driving factor instead of height (Loos et al., 2001). If we would have used weight as dependent variable instead of BMI, a stronger correlation might be observed. Furthermore, the variance expressed in the R^2 would most likely be higher. Weight is more often used for malnutrition by comparing the weight at a specific age to a reference, weight-for-height would already be more descriptive. Nevertheless, BMI is a more standard approach to gain insight in relative adiposity in children and therefore more fitting for the current research (Lobstein, Baur, & Uauy, 2004)

From the simplistic model, where we treated the twins as individuals, we are not sure of the found association between birthweight and BMI is confounded by other factors we know from the second twin. Therefore, we do a second regression with additional variables. With an R^2 of 4.17%, the more elaborated model is better at explaining the variance in the z-score BMI, indicating that sex, age of mother and education of mother may play an important role in the association with z-score BMI. This confirms that birthweight might be a piece of a bigger puzzle explaining obesity, these other independent variables are necessary to complete the puzzle. Hence, we will shortly discuss the potential mechanisms explaining the association between these three other variables included in the elaborated model.

9.1.2 Discussion: Twins and the within twin pair X-value

The within twin pair X-value describes the difference of twin 1 or twin 2 from the average birthweight of the corresponding twin pair. The newly created variables express an association which is free of confounding due to the shared environment of te twins. The X-value in birthweight will not be affected constant confounders, thus which are the same for both twins, but these confounders can still influence the between twin pair effects. Confounders which vary within a twin pair can affect both the X-value and the other variables in the regression. This explains, in the case when the X-value to be significant and not zero, that the X-value reflects an association which is free from confounding from shared factors of twins from a pair. However, the individual-twin specific factors, as in-utero experiences or genetics in dizygotic twins, underly the differences in the X-value. The same individual-twin specific factors which cause the difference in the X-value also are than quite certain to drive the difference in the Y, in this case BMI reviews (Carlin, Gurrin, Sterne, Morley, & Dwyer, 2005).

To make the previous paragraph less cryptic, we discuss our own results. Overall, the X-value in the simple (only X-value) and in the elaborated model the between X-value was not significant. This indicates that we cannot conclude that in-utero or genetic effects specific to either twin 1 or twin 2 explain obesity development in later life. Nevertheless, for specific groups from the sample we conclude that twin specific characteristics which differ from the other twin do explain obesity development. For twin 1, the results show that in-utero experiences, genetics, or other twin specific characteristic decrease obesity development when the mother has a medium education. This finding does not conclude that in a casual form having a higher education leads to less obesity, it only shows that within twins who have mothers with a medium education there are twin-specific factors which decrease the development of obesity. The other two strata, extremely/very low birthweight and young mothers, which find the X-value to be significant should be interpreted with lots of caution due to the limited observations. There is a high likeliness these findings are significant because of coincidence or selection of twins.

For twin 2, none of the strata which were significant in twin 1 found the same results. For twin to the following strata the X-value is significant: mothers with high education, mothers aged 35+ and dizygotic twins shows a significant association between the X-value and z-score BMI. Especially, the result of dizygotic twins to be significant in twin 2 shows that most likely genetics do play a role in the birthweight and (z-score) BMI association. More specific details on the potential underlying theories can be found in the following paragraphs.

9.1.3 Obesity development and the importance of sex

Firstly, our data shows that often being female compared to male is significantly associated to z-score BMI when treating twins as individuals (table 2 and 3), as well as, when using the between X-value. Overall, the interpretation follows females compared to the male reference group have a lower z-score BMI. This negative association is observed in all the regressions. This negative effect is most likely partly strengthened by the already existing difference in BMI between males and females in the twin sample. The female BMI is 18.57 while the male BMI is higher with 19.04. From this data it is however, not possible to conclude how sex influences the relation of birthweight and BMI. Additionally, it stays complicated to explain this found difference between males and females, because the knowledge on sex-difference in heritability around body compositions and BMI from birth till in late life is still limited (Dubois et al., 2012).

Literature explains that clear distinctions between sex and body compositions are already observed in the earliest stages of life. Previous research shows that girls are born with a lower birthweight than boys (Rugholm et al., 2005; Wilkin & Murphy, 2006). Additionally, in our sample girls have a birthweight of 2,381.78 grams (SE=537.17, N=1,028) while boys are heavier with a birthweight of 2,465.27 gram (SE=528.98, N=976). As our study and previous research finds that higher birthweight is related to higher BMI. This might point to that boys already from birth are more susceptible to develop obesity than girls. Nevertheless, the mechanism causing this difference is birthweight is not yet worked out. One hypothesis includes the presence of sex-specific genes (Wilkin & Murphy, 2006). Hence, birthweight might still be of interest in explaining the difference in BMI among males and females. In future studies, it might be interesting to stratify the results for sex. With the current methods this was not possible due to the inclusion of mixed-gender twin pairs. Finding the mechanism underlying sex difference in birthweight might help to understand sex-difference in the development of obesity.

9.1.4 Obesity development and the importance of mother's age at birth

Within our results we found that age of mothers plays an important role in obesity development in later life. Mothers aged 25 and older at birth compared to young mothers are shown to be negatively significant associated to BMI. This association is found when treating twins as individuals and when the between X-value is added. In the statistical analysis with the X-value, findings suggest significance of the X-value for twin 1 in young mothers and for twin 2 in mothers older than 35. In both cases this should be interpreted as an indication that the bigger the difference from the twin average, the higher the BMI.

To the author's knowledge, there is little research which focusses on the connection between maternal age, birthweight and BMI. However, in a research by Myrskylä and Fenelonthey in 2012 a relation was found signalling that mothers younger than 25 and older than 35 compared to mothers aged 25-34 have children with a worse health. Their definition of worse health includes obesity, but also mortality, self-rated health, height and diagnosed conditions. They also made a plot specifically for the association between obesity and age of mothers, the same negative association was presented by a u-shape in the plot.

In the same research, they controlled the relation between mother's age at birth and obesity of the child for the education of the mother (and lifespan shared by mother and child). The result showed no effect on the probability of having obese children for young mothers, but in mothers over the age of 30 the probability of having a child with obesity decreased (Myrskylä & Fenelon, 2012). This indicates that for older mothers an education can have a moderating effect on the development of obesity in their offspring Hence, age of mother is an interesting variable in the relation of birthweight and obesity development. However, the association found is most likely influenced by other factors which are associated with older age such as education.

9.1.5 Obesity development and the importance of mother's education

As mentioned before, education plays an important role in obesity development (Kleiser, Rosario, Mensink, Prinz-Langenohl, & Kurth, 2009; Oken & Gillman, 2003; Webbink, Martin, & Visscher, 2010; Zhao, Wang, Mu, & Sheng, 2012). However, according to our analysis, education is less often significant than the age of mothers. Firstly, in both elaborated models (table 2 & 4) only high education compared to low/no education is significantly associated to obesity development. A reason for the lack of significant results for medium education could be that the difference between the category of medium and high education is quite small, namely 2 school-grades. There is no consistency in medium, high or both categories of educations compared to the reference of low/no education to be significantly associated with BMI. A potential explanation could be that a part of the effect of education acts through the age of mothers and therefore is observed in the current regressions.

However, an interesting result is observed when considering statistical analysis, both when treating twins as individuals and with the X-value, for normal birthweight was that both medium a high education compared to no/low education are negatively significant associated with z-score BMI. This indicates that medium and higher education in our sample have a positive influence by decreasing the development of obesity, which agrees with the findings of Myrskylä and Fenelonthey in 2012. Additionally, in a study among German adolescents, the researchers formulated the social economic status of individuals on parents' income, employment, and education. Their results show children with a low SES, meaning children having parents with low education, are more often obese. This association is still stronger among children who behave in a way which would prevent obesity development and have low SES than for children from medium and high SES who behave unhealthy (Kleiser, Rosario, Mensink, Prinz-Langenohl, & Kurth, 2009)

In this study, we merely looked at education of mothers. This partly to limit the exclusion of twin pairs due to missing data. However, in 2008 the study from Webbink et al. (2010) explains previous studies in America on the effect of parental education on the child's education show that the father's education has a positive effect, but mother's education does not have an effect. In their own results they find that education positively influences the body composition of boys, but again no effect for girls. Contrastingly, in our own results for the statistical analysis stratified education (table 7), we observe that only in the strata of high education of mothers, twins being female compared to male has a negative effect on z-score BMI. Showing that being a female twin within the strata of high education of mothers leads to lower BMI compared to boys. Thus, showing that high education does likely influence the relation of being a female (or male) and BMI. This difference in our results and previous results might be due to the timing of the research and more females obtaining higher education which can also be observed in our own descriptive statistics (table 1). Hence, more research into the influence of parental education and BMI development could help policymakers to understand and optimize the effect of BMI decreasing policies in schools. Findings which suggest that low education mothers for instance or more likely to have obese children, might inform policy makers to help new-moms also outside schools. For future research, also involving data on father's education might be insightful.

9.1.6 Obesity development and the importance of cohorts

In the descriptive statistics, we observe changes in the variables used in this thesis between the four cohorts. The age of mothers at birth increased from 29.97 years to 32.06 years in the later cohort. Further, in the younger birth cohorts, from 2009/2010, more mothers obtain a higher education, the percentage of mothers with high education increases from 38%, 49%, 54% to 61%. Next to these observed changes, also medical innovations and reduction of smoking (Reitsma et al., 2017) leads us to expect the observation of differences between cohorts. However, in the statistical analyses for cohorts (table 6), we observe little noteworthy differences. Only in the youngest cohort the age of mothers seems to be significantly associated to the decrease of the z-score BMI.

Additionally, stronger cohort effects could be expected among the second twin. The younger twins of the twin pairs experience more stress during the birth and are on higher risk of death in-utero related to the process while giving birth (Smith, Fleming, & White, 2007). With these insights in this stress experienced by twin 2, we could expect over time more medical measures are taken to decrease the stress for this twin and decrease the risk of death or other unfavourable outcomes. However, in our dataset twin 2 is born with a lighter birthweight than twin 1 which most likely is a sign for the unequal and disadvantageous experience in-utero for twin 2. Further, our results show no cohort effects for twin 2 which indicate that the experience have in-utero or during birth have become more equal to twin 1 over time.

In 2016, Yokoyama et al. found that the birth order of twins can explain differences in BMI, in both boys and girls until the age of 12 if the results were adjusted for birthweight. This led to the conclusion that a part of the effect of difference between twins in BMI is due to unique experience in-utero and the birth order, next to a part of the association being explained by birthweight (Yokoyama et al., 2016).The methods used in our study are not designed to conclude the same results, nor are they able to rule out these conclusions. However, in our results in cohorts we do not observe large differences for either twin 1 or twin 2.

9.1.7 Obesity development and the importance zygosity

Stratifying the regression for zygosity makes it possible to control even better for genetics, especially for monozygotic twins. Monozygotic twins share most of their genetics and we assumed that they experience their environments growing up the same, meaning that when statistically analysing monozygotic twins, we can test the influence of different experience by monozygotic twins in-utero environment (Loos, Beunen, Fagard, Derom, & Vlietinck, 2001). Sadly, in our results the X-value is not significant among monozygotic. Suggesting that the factors driving the no-shared factors among monozygotic twins in a pair, thus the unique in-utero environments, are not driving the obesity development in later life. The in-utero environments are not experienced differently by the twins or the between X-value had to be significant. These results do not give additional information on the influence of genetics and birthweight o BMI and thus not excluding genetic factors to be of importance (The et al., 2010).

In a study from 2001 in male twins from East-Flanders, in monozygotic twins no association was found between an intra-twin for birthweight and BMI (p=0.33). Nevertheless, next to BMI this study also included other measurements of body composition, namely height and weight, which were significantly associated with birthweight (Loos, Beunen, Fagard, Derom, & Vlietinck, 2001). All in all, from merely looking at monozygotic twins we learn that in-utero effects, which can be experienced differently by twins from a twin pair, might only have limited effect on the association of birthweight and BMI.

Interestingly, as introduced early, in the dizygotic analyses for twin 2 the between X-value of birthweight is significant. However, at the same time, in the monozygotic twins the between X-value is never significant. Difference between mono- and dizygotic twins are observed in other research. In 2005, IJzerman et al. reported that within twin pair difference in birthweight were negatively significant associated to differences in blood pressure within dizygotic twins, but not within monozygotic twins. They concluded that within this association of birthweight and blood pressure, genetics play an important role (Ijzerman, Boomsma, & Stehouwer, 2005). From our results, we can conclude that genetics also may play a role in the difference between birthweights and BMI in later life. The X-value expresses the unique characteristics of twins from a dizygotic pair which are driving the difference in birthweight might also be driving the difference in BMI. However, it is hard to distinct between genetics and the in-utero effects, because the X-value only express the importance of non-shared characteristic which can be both genetics and in-utero effects. Crucial with this conclusion is that the assumption of twins experiencing their environment the same is true. However, caution should be taken due to twins still can be treated differently by society or have different experiences (Mackenbach, 2005).

Lastly, in the regression of dizygotic twin 1 all other variables besides the within twin pair difference of birthweight were significant. Suggesting, that next to genetics or in-utero effects, the social environment still has a strong influence. This indicates that twins with certain genes are maybe more susceptible to develop a lower BMI in certain environments. All in all, the findings within the dizygotic twin pairs do indicate that genetic a/or a difference in the experience while in-utero drive obesity development.

9.2 STRENGTHS & LIMITATIONS

The strengths of this study are expressed in the opportunity of studying over 2000 twins from the TwinLife dataset. The quality of the dataset from TwinLife is very high which is for instance shown by the determination process of zygosity. This process was supported by saliva test and researched separately which showed almost all twins were correctly assessed (Lenau et al., 2017). Further, the TwinLife dataset is relatively new and consist of twins which are born during the obesity epidemic. To the authors' knowledge, even though more research is being conducted in persons born after the year 2000, there are still little studies done in this new population.

Furthermore, the results of this study can have implications, as also expressed in the relevance, for policymakers. Our results point in the direction that both genetics and the social environment play a role in obesity. Therefore, these results inform that policymakers do have the potential to influence our social environment and decrease the impacts of genetics. Furthermore, with the right policies in place also the costs of obesity and the wellbeing of populations can increase.

Nevertheless, there also multiple limitations which decreased the quality of the approach, data and eventually the results. Cautionary notes for this study focus on quality of (self-reported) data, BMI as dependent variable, missing (confounding) variables, and the representativeness of twins. Firstly, the data in the dataset is based on self- or parental-reports which can lead to various report biases. One of the main reasons for exclusion was that birthweight data was not available. It is also possible that parents or individuals made an educated guess which may be different from their actual birthweight which is known as recall bias. Additionally, we checked for faulty or impossible answers on questions from the questionnaire during the preparation of the data. However, this does not guarantee for all wrong data to be excluded. For birthweight, height and length individuals could accidentally give the wrong number or make other mistakes. Furthermore, for BMI to be accurate it is advised to have measurements are done by a professional (Lobstein et al., 2004). Hence, we cannot assure that there are no wrongfully reported data in the dataset.

For the dependent variable we decided to use BMI, because of length and weight to be available in the TwinLife dataset. However, debates have discussed if BMI is a good measurement of body fat and, thus, obesity. BMI is not a sensitive measurement for body fat due to it ignoring body types. Short, tall or very muscular persons can have the same BMI's, even if they have different amounts of bodyfat (Lobstein et al., 2004). BMI is described as giving insight into the "heaviness" of a persons' body more than the presence and distribution of fat (Loos, Beunen, Fagard, Derom, & Vlietinck, 2001). Other measurements for obesity used in other research are waist-hip-ratio's or skinfold-thickness. These measures do have their own advantages and disadvantages, as being harder and more expensive to obtain

(Lobstein et al., 2004). Additionally, we transformed the BMI into a z-score to study all twins in one go and compare between cohorts. However, we took a simplistic approach which caused that the z-score do not have to be entirely correct and therefore classifying people as obese or overweight while they are not. To increase the quality of the z-score especially in the younger children for each age another z-score should be calculated.

For future research, it would be interesting to add new variables to the TwinLife dataset to potentially increase the quality of research on the topic of obesity and overweight. As explained above, more direct methods for measuring fat would increase understanding of which factors lead to obesity. Other factors which could be added to increase the quality of research on the current topic would be data on smoking behaviour of parents and information on body compositions of parents. For the current research, it is most important that this data is available at birth. As discussed in the theoretical framework both smoking and obese mothers can have a confounding effect on the relation of birthweight and obesity. Currently, I was not able to test for these factors which potentially could have led to different results.

Furthermore, for repetition of this research adjustments to the approach could also enhance the quality of the associations. Within the scope of this thesis, I decided to only compare twins from a twin pair and not siblings. Comparison of twins and siblings can give an extra insight into genetic and environmental difference. Additionally, I divided the group in twin 1, first born, and twin 2, second born. Nevertheless, another approach could have been to distinct between the heaviest and lighter twin from a twin pair. Especially, because we did not find cohort effects which led us to assume that stress experienced by the second twin during birth is most likely expressed in the birthweight. Thirdly, I decided to limit the control variables to sex, mothers age at birth, education of mother. The SES in this thesis was expressed by merely looking at the education of mothers, while for instance in the KIGGS study they formulate the SES by combing information on parental education, occupation and income (Kleiser, Rosario, Mensink, Prinz-Langenohl, & Kurth, 2009). Lastly, in this discussion we already stated it was still difficult to differentiate between the environment, genetics and in-utero effects. Hence, by selecting other variables and re-grouping strata it could increase the quality.

Lastly, causation should be taken regarding the representativeness of the twin sample. Webbink et al. (2008) discussed shortly that twins are quite representative for the rest of the population, but results from twin studies are not always transferable for the entire population. If we compare our sample of twins to the rest of the German population, the following difference stand out: birthweight, age of mothers, and education of mothers. In our sample we have no individuals with a weight more than 4000 grams, while in Germany in 2010 over 68,000 babies were born with a birthweight above 4000 grams Furthermore, for the age of mothers we did not control if more siblings were born before or after the birth of the twin pair. If twins were mostly born as first child mothers the average found would be lower (UN Data - Demographic Statistics Database, 2018). The sample could, for instance accidentally, consist of mostly younger mothers. If we compare the age of mothers from this sample in 2009 and 2010 to the rest of Germany, we find that the age is slightly lower to our sample. Overall the age at births, if combining the different averages of the age at birth order for German mothers, of mothers in Germany was in 2009 31.78 years and in 2010 31.83 (Statistisches Bundesamt (Destatis), 2018). In our sample most of the mother's attained a high education, which most likely is caused by a selection effect. In the birth cohort of 2009/2010 over 60% of the mothers had a high education, while for Germany overall the percentage ranged between 25.3% in 2009 and 25.8% in 2010 (Statistisches Bundesamt - Destatis, 2018). The national numbers of Germany are for both sexes, and most likely would have been a little bit higher if only people in reproductive ages (18-45) were considered. Also, because high education has a moderating effect on obesity among offspring (Webbink, Martin, & Visscher, 2010), the high percentage of educated mothers could have decreased the actual association between birthweight and obesity. All in all, the representativeness of the current sample for the rest of Germany can be discussed. Most likely, findings in twin give insight in the relation, but will have to be tested in the overall population too.

9.3 FUTURE RESEARCH

The current thesis is not able to conclusive indicate how birthweight and obesity are caused by either the environment, genetic or in-utero effects. More future research should be necessary on this topic to be able to describe the entire mechanism. From this research birthweight, sex, education of mother and the age of mother seem to be involved in the mechanism. However, more research is required on the sex difference in birthweight and obesity. As well as, social environment factors as education which potentially have different effects on sex. For policymakers this research could than indicate, for instance, if school-aimed policy work on both sexes or merely for boys or girls. Lastly, most of the comparable studies mentioned in this thesis were conducted among Caucasian samples from the America, Australia or Europe. It would be interesting to also do comparable twin research in other countries to get a better understanding between nationalities.

Additionally, the aim of this thesis was to differentiate between nature, genetics, and nurture, environmental effects. Although with the current approach we tried to reach this goal, it is still not possible to see difference between genetics and in-utero effects. To study this more clearly focus on comparative mono- and dizygotic twins could help, if assumed that monozygotic twins share most of their genetic information. Also, the more advanced methods explained by Carlin et al. could ensure a higher quality of the data and findings. Either with the more advance methods or the current methods, it would also be interesting Further building on the current structure, it would be interesting to follow the cohorts and see how their weight, length and BMI chances over the years. This will potentially give insight in a life perspective and development of overweight and obesity. In the current thesis we only look at two static moments in time, birth and moment of interview.

Lastly, the current results and previous discussed studies show that low birthweight and high birthweight show to effect obesity development. Here we only fitted a linear relation, it would be interesting to fit also test for nonlinear relation as u- and j-shaped relations.

10 CONCLUSION

The objective of this study was to test if there is an association between birthweight and obesity development later in life among German Twins from the TwinLife dataset. Secondly, we wanted to gain insight into the importance of shared genetic factors and the shared social origin of obesity development in later life.

Our results showed a positive association between birthweight and BMI in later life among German Twins aged 4 to 25. The association is more apparent when treating twins as individuals in the strata for twins born with a higher birthweight (> 2500 grams). However, association does not mean causation, and, in this case, the casual mechanism of birthweight and BMI cannot be confirmed with the current data or approach.

With the X-value for within pair birthweight differences we gained insight into the importance of the social environment, as well as genetics. From the current results, overall the X-value was not significant only for specific stratums. The finding which gave the most insight in the importance of genetic and partly answering our second sub-questions regards the results of the dizygotic twins. Within the regression for the second twin, the X-value was significant. This indicates that factors which are not shared by the twins from a specific twin pair explain the difference in birthweight. The same factors which explain this difference in birthweight are also causing the elevated BMI. Dizygotic twins are not genetically the same leading to the conclusion that genetics do play a role in the relation of birthweight and BMI. Sadly, in this research we were not able to distinct between the unique experience of the in-utero environment of twins. Hence, this finding of a significant X-value is not solely genetics, but can also refer to the differences in the in-utero environment.

Next to the interest in birthweight, we also gained insight in the importance of sex, education of mothers', age of mothers at birth, and cohorts. Almost in all regression, being female compared to male had a significant negative impact on the development of obesity. Potentially this relation is already influenced or determined from birth were girls are born lighter than boys. Future research is required to gain more insight in difference birthweight at birth and the development of body composition to explain if males and females develop obesity differently.

Further, the other three variables give insight in the importance of the social environment. These also answer the last two sub-questions on the importance of the social environment. Although expected due to innovations and changed descriptive statics, no differences were observed in cohorts. The age of mothers at birth was more often significant than education of mothers. Potentially, a part of the effect of education was already accounted for by the age of mothers. Having a higher education compared to lower educations more often was found to decrease the obesity development in the twins. Additionally, having older mothers compared to young mothers (< 25years) were found to again decrease the obesity development. Our results did not suggest that older mothers have a higher likeliness of obese children, this however, might be due to education of older mothers.

All in all, this research made it possible to observe a positive association between birthweight and BMI. Additionally, it found that both genetics, in-utero experiences and the social environment are of importance. Sadly, in this research we were not able to distinct between genetics and in-utero effects. Therefore, more research is still required to explain the mechanism of obesity development. Nevertheless, from this study the conclusion suggest that obesity is formed by an interplay of genetics, in-utero effects and the social environment. Hence, individuals are not necessarily doomed from birth.

11 REFERENCES

Almond, D., & Currie, J. (2011). Killing me softly: The fetal origins hypothesis. *Journal of Economic Perspectives, 25*(3), 153-172.

Baird, J., Osmond, C., MacGregor, A., Snieder, H., Hales, C., & Phillips, D. (2001). Testing the fetal origins hypothesis in twins: The birmingham twin study. *Diabetologia, 44*(1), 33-39.

Barker, D. J. (1990). The fetal and infant origins of adult disease. *BMJ (Clinical Research Ed.), 301*(6761), 1111.

Calle, E. E., Rodriguez, C., Walker-Thurmond, K., & Thun, M. J. (2003). Overweight, obesity, and mortality from cancer in a prospectively studied cohort of US adults. *New England Journal of Medicine, 348*(17), 1625-1638.

Carlin, J. B., Gurrin, L. C., Sterne, J. A., Morley, R., & Dwyer, T. (2005). Regression models for twin studies: A critical review. *International Journal of Epidemiology, 34*(5), 1089-1099.

Danaei, G., Lu, Y., Singh, G., Stevens, G., Cowan, M., Farzadfar, F., . . . Khang, Y. (2014). Cardiovascular disease, chronic kidney disease, and diabetes mortality burden of cardiometabolic risk factors from 1980 to 2010: A comparative risk assessment. *The Lancet, 2*(2), 634-647.

Dubois, L., Kyvik, K. O., Girard, M., Tatone-Tokuda, F., Pérusse, D., Hjelmborg, J., . . . Lichtenstein, P. (2012). Genetic and environmental contributions to weight, height, and BMI from birth to 19 years of age: An international study of over 12,000 twin pairs. *PLOS One, 7*(2), e30153.

Gesink Law, D., Maclehose, R. F., & Longnecker, M. P. (2006). Obesity and time to pregnancy. *Human Reproduction, 22*(2), 414-420.

GESIS TwinLife. (2018). TwinLife - data description and access. Retrieved from https://paneldata.org/twinlife

Giskes, K., Kamphuis, C. B., van Lenthe, F. J., Kremers, S., Droomers, M., & Brug, J. (2007). A systematic review of associations between environmental factors, energy and fat intakes among adults: Is there evidence for environments that encourage obesogenic dietary intakes? *Public Health Nutrition, 10*(10), 1005-1017.

Hales, C. N., & Barker, D. J. (2001). The thrifty phenotype hypothesis. *British Medical Bulletin, 60*(1)

Halfon, N., & Hochstein, M. (2002). Life course health development: An integrated framework for developing health, policy, and research. *The Milbank Quarterly, 80*(3), 433-479.

Hattersley, A. T., & Tooke, J. E. (1999). The fetal insulin hypothesis: An alternative explanation of the association of low bir thweight with diabetes and vascular disease. *The Lancet, 353*(9166), 1789-1792.

Herrera, B. M., Keildson, S., & Lindgren, C. M. (2011). Genetics and epigenetics of obesity. *Maturitas, 69*(1), 41-49.

Huxley, R. R., Shiell, A. W., & Law, C. M. (2000). The role of size at birth and postnatal catch-up growth in determining systolic blood pressure: A systematic review of the literature. *Journal of Hypertension, 18*(7), 815-831.

Ijzerman, R. G., Boomsma, D. I., & Stehouwer, C. D. (2005). Intrauterine environmental and genetic influences on the association between birthweight and cardiovascular risk factors: Studies in twins as a means of testing the fetal origins hypothesis. *Paediatric and Perinatal Epidemiology, 19*, 10-14.

Jesse, R., Xiaoling, Q., Marbán, S. L., Henry, R., & Lowell, B. C. (2004). The obesity pandemic: Where have we been and where are we going? *Obesity Research, 12*, 88S-101S. doi:10.1038/oby.2004.273

Kinge, J. M., & Morris, S. (2010). Socioeconomic variation in the impact of obesity on health-related quality of life. *Social Science & Medicine, 71*(10), 1864-1871.

Kirk, S. F., Penney, T. L., & McHugh, T. (2010). Characterizing the obesogenic environment: The state of the evidence with directions for future research. *Obesity Reviews, 11*(2), 109-117.

Kivipelto, M., Ngandu, T., Fratiglioni, L., Viitanen, M., Kåreholt, I., Winblad, B., . . . Nissinen, A. (2005). Obesity and vascular risk factors at midlife and the risk of dementia and alzheimer disease. *Archives of Neurology, 62*(10), 1556-1560.

Kleiser, C., Rosario, A. S., Mensink, G. B., Prinz-Langenohl, R., & Kurth, B. (2009). Potential determinants of obesity among children and adolescents in germany: Results from the cross-sectional KiGGS study. *BMC Public Health, 9*(1), 46.

Lehnert, T., Streltchenia, P., Konnopka, A., Riedel-Heller, S. G., & König, H. (2015). Health burden and costs of obesity and overweight in germany: An update. *The European Journal of Health Economics, 16*(9), 957-967.

Lenau, F., Hahn, E., Peters, A., Gottschling, J., Thiel, W., & Spinath, F. M. (2017). Zygosity determination in twin studies: A validation of zygosity questionnaires using DNA in the german TwinLife study.

Lobstein, T., Baur, L., & Uauy, R. (2004). Obesity in children and young people: A crisis in public health. *Obesity Reviews, 5,* 4-85.

Loos, R., Beunen, G., Fagard, R., Derom, C., & Vlietinck, R. (2001). Birth weight and body composition in young adult men—a prospective twin study. *International Journal of Obesity, 25*(10), 1537.

Mackenbach, J. P. (2005). Genetics and health inequalities: Hypotheses and controversies. *Journal of Epidemiology and Community Health, 59*(4), 268-273. doi:59/4/268 [pii]

Morgan, A. R., Thompson, J. M., Murphy, R., Black, P. N., Lam, W., Ferguson, L. R., & Mitchell, E. A. (2010). Obesity and diabetes genes are associated with being born small for gestational age: Results from the auckland birthweight collaborative study. *BMC Medical Genetics, 11*(1), 125.

Myrskylä, M., & Fenelon, A. (2012). Maternal age and offspring adult health: Evidence from the health and retirement study. *Demography, 49*(4), 1231-1257.

Naukkarinen, J., Rissanen, A., Kaprio, J., & Pietiläinen, K. (2012). Causes and consequences of obesity: The contribution of recent twin studies. *International Journal of Obesity, 36*(8), 1017.

NCD Risk Factor Collaboration. (2016). Trends in adult body-mass index in 200 countries from 1975 to 2014: A pooled analysis of 1698 population-based measurement studies with 19· 2 million participants. *The Lancet, 387*(10026), 1377-1396.

OECD. (2017, June). *Obesity update 2017.* ().OECD. Retrieved from OECD Health Statistics 2017

Oken, E., & Gillman, M. W. (2003). Fetal origins of obesity. *Obesity Research, 11*(4), 496-506.
Ong, K. K. (2006). Size at birth, postnatal growth and risk of obesity. *Hormone Research, 65 Suppl 3,* 65-69. doi:91508 [pii]

Poulton, R., Caspi, A., Milne, B. J., Thomson, W. M., Taylor, A., Sears, M. R., & Moffitt, T. E. (2002). Association between children's experience of socioeconomic disadvantage and adult health: A life-course study. *The Lancet, 360*(9346), 1640-1645.

Puhl, R. M., & Heuer, C. A. (2009). The stigma of obesity: A review and update. *Obesity, 17*(5), 941-964.

Ramlau-Hansen, C. H., Thulstrup, A. M., Nohr, E., Bonde, J. P., Sørensen, T., & Olsen, J. (2007). Subfecundity in overweight and obese couples. *Human Reproduction, 22*(6), 1634-1637.

Reitsma, M. B., Fullman, N., Ng, M., Salama, J. S., Abajobir, A., Abate, K. H., . . . Abyu, G. Y. (2017). Smoking prevalence and attributable disease burden in 195 countries and territories, 1990–2015: A systematic analysis from the global burden of disease study 2015. *The Lancet, 389*(10082), 1885-1906.

Ribeiro, A. M., de Carvalho Lima, M., de Lira, Pedro Israel Cabral, & da Silva, Giselia Alves Pontes. (2015). Low birth weight and obesity: Causal or casual casual association? *Revista Paulista De Pediatria (English Edition), 33*(3), 340-348.

Rugholm, S., Baker, J. L., Olsen, L. W., Schack-Nielsen, L., Bua, J., & Sørensen, T. I. A. (2005). Stability of the association between birth weight and childhood overweight during the development of the obesity epidemic. *Obesity Research, 13*(12), 2187-2194. doi:10.1038/oby.2005.271

Schienkiewitz, A., Brettschneider, A., Damerow, S., & Rosario, A. S. (2018). Übergewicht und adipositas im kindes-und jugendalter in Deutschland–Querschnittergebnisse aus KiGGS welle 2 und trends.

Smith, G. C., Fleming, K. M., & White, I. R. (2007). Birth order of twins and risk of perinatal death related to delivery in england, northern ireland, and wales, 1994-2003: Retrospective cohort study. *BMJ (Clinical Research Ed.), 334*(7593), 576. doi:bmj.39118.483819.55 [pii]

Statistisches Bundesamt - Destatis. (2018). Educational attainment - educational attainment of the population in germany. Retrieved from https://www.destatis.de/EN/FactsFigures/SocietyState/EducationResearchCulture/EducationalLevel/Tables/EducationalAttainmentPopulationGermany.html

Statistisches Bundesamt (Destatis). (2018). Genesis - Age of the mother:
difference between the child's year of birth and the mother's year of birth ('year-of-birth method'). Age of the mother:
difference between the child's year of birth and the mother's year of birth ('year-of-birth method'). .

Swinburn, B., Egger, G., & Raza, F. (1999). *Dissecting obesogenic environments: The development and application of a framework for identifying and prioritizing environmental interventions for obesity* doi:http://dx.doi.org/10.1006/pmed.1999.0585

The, N. S., Adair, L. S., & Gordon-Larsen, P. (2010). A study of the birth weight–obesity relation using a longitudinal cohort and sibling and twin pairs. *American Journal of Epidemiology, 172*(5), 549-557.

TwinLife. (2018). Study. Retrieved from http://www.twin-life.de/en/study

UN Data - Demographic Statistics Database. (2018). Live births by birth weight and sex of child. Retrieved from http://data.un.org/Data.aspx?q=birth+weight&d=POP&f=tableCode%3a60

Webbink, D., Martin, N. G., & Visscher, P. M. (2010). Does education reduce the probability of being overweight? *Journal of Health Economics, 29*(1), 29-38.

Wilkin, T., & Murphy, M. (2006). The gender insulin hypothesis: Why girls are born lighter than boys, and the implications for insulin resistance. *International Journal of Obesity, 30*(7), 1056.

World Health Organization. (2018a). Obesity and overweight. Retrieved from http://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight

World Health Organization. (2018b). World health statistics 2018: Monitoring health for the SDGs, sustainable development goals.

World Health Organization: ICD-11. (2018). ICD-11 for mortality and morbidity statistics. Retrieved from https://icd.who.int/browse11/l-m/en

Yaghootkar, H., & Freathy, R. M. (2012). Genetic origins of low birth weight. *Current Opinion in Clinical Nutrition & Metabolic Care, 15*(3), 258-264.

Yokoyama, Y., Jelenkovic, A., Sund, R., Sung, J., Hopper, J. L., Ooki, S., . . . Tarnoki, D. L. (2016). Twin's birth-order differences in height and body mass index from birth to old age: A pooled study of 26 twin cohorts participating in the CODATwins project. *Twin Research and Human Genetics, 19*(2), 112-124.

Zhao, Y., Wang, S., Mu, M., & Sheng, J. (2012). Birth weight and overweight/obesity in adults: A meta-analysis. *European Journal of Pediatrics, 171*(12), 1737-1746.

12 APPENDIX

12.1 TO DO FILES USED IN STATA

12.1.1 Twins as individuals

```
///dataset for twin 1
        ///start by dropping and generating
        *drop cases without biological mother
        drop if fpr0107_m_1 <=0
        drop if fpr0107_m_1 == 23
                                    //1032 cases left
        *drop cases without age of mother
        drop if age0100_m_1 == -95
                                    //1031 cases left
        *drop cases without education of mother
        drop if eca0100_m_1 <= 0
                                    //1002 cases left

        *generate AGE OF MOTHER AT BIRTH OF CHILD
        generate age_m_atbirth_t = age0100_m_1-age0100_t_1

        *continious variables to be grouped : education mother & age of mother
        summarize age_m_atbirth_t eca0100_m_1

        egen cat_age_m_atbirth_t = cut(age_m_atbirth_t), at (0, 17, 25, 35, 50) icode
        table cat_age_m_atbirth_t, contents(freq min age_m_atbirth_t max age_m_atbirth_t)
                        //3 catgeories young 18-24, average 25-34, risk 35+
        tab cat_age_m_atbirth_t

        generate copy_eca0100_m_1 = eca0100_m_1
        recode eca0100_m_1 6=1
        egen cat_eca0100_m_1 = cut(eca0100_m_1), at (0, 1, 3, 4, 6) icode
        table cat_eca0100_m_1, contents (freq min eca0100_m_1 max eca0100_m_1)
        tab cat_eca0100_m_1
                //3 categories low edu (primary and other), medium (secondary), high (applied science & uni)

///dataset for twin 2
clear
cd "c:\Users\Lieke\Documents\MS Population\MSc Thesis\STATA\Stata"
use fid age0100_m_1 BMI_u zhBMI_u z_BMI_u_all z_BMI_u_3 nbi0115_u_1 nbi0114_u_1 bdy0100_u_1
bdy0100u_1 age0100_u_1 fpr0107_m_1 fpr0104_u_1 fpr0105_u_1 cgr_1 sex_u_1 tsex_1 zyg0102_1 nuts1_a_1_1
gkpol7_1_1 fpc_1 fam0100_m_1 fam0100_f_1 eca0100_m_1 eca0100_f_1 dia5400u_1 mig0100_m_1
mig0100_f_1 using "ZA6701_family_wave1_v1-0-0 incl all variables, z-scores, and 1038 cases"
        ///start by dropping and generating
        *drop cases without biological mother
        drop if fpr0107_m_1 <=0
        drop if fpr0107_m_1 == 23
                                    //1032 cases left
        *drop cases without age of mother
        drop if age0100_m_1 == -95
                                    //1031 cases left
        *drop cases without education of mother
        drop if eca0100_m_1 <= 0
                                    //1002 cases left
```

```
	*generate AGE OF MOTHER AT BIRTH OF CHILD
	generate age_m_atbirth_u = age0100_m_1-age0100_u_1

	*continious variables to be grouped : education mother & age of mother
	summarize age_m_atbirth_u eca0100_m_1

	egen cat_age_m_atbirth_u = cut(age_m_atbirth_u), at (0, 17, 25, 35, 50) icode
	table cat_age_m_atbirth_u, contents(min age_m_atbirth_u max age_m_atbirth_u)
				//3 catgeories young 18-24, average 25-34, risk 35+
	tab cat_age_m_atbirth_u

	generate copy_eca0100_m_1 = eca0100_m_1
	recode eca0100_m_1 6=1
	egen cat_eca0100_m_1 = cut(eca0100_m_1), at (0, 1, 3, 4, 6) icode
	table cat_eca0100_m_1, contents (min eca0100_m_1 max eca0100_m_1)
	tab cat_eca0100_m_1
		//3 categories low edu (primary and other), medium (secondary), high (applied science & uni)

///copy data from dataset twin 2 and add in dataset of twin 1 - gives following dataset
*dataset: ZA6701_family_wave1_v1-0-0 final dataset for twins as individuals
*regression do not have to account for pairing in data, because twins are observed as individuals
cd "c:\Users\Lieke\Documents\MS Population\MSc Thesis\STATA\Stata"
use "ZA6701_family_wave1_v1-0-0 final dataset for twins as individuals"
sum nbi0115_t_1 sex_t_1 cat_age_m_atbirth_t cat_eca0100_m_1

///birth weight categories
generate c_nbi0115_t_1 = nbi0115_t_1
egen cat3_nbi0115_t_1 = cut(c_nbi0115_t_1), at (0, 1, 1499, 2499, 3999) icode
table cat3_nbi0115_t_1, contents(freq min c_nbi0115_t_1 max c_nbi0115_t_1)

//run regressions
regress z_BMI_t_3 nbi0115_t_1
regress z_BMI_t_3 nbi0115_t_1 i.sex_t_1 i.cat_age_m_atbirth_t i.cat_eca0100_m_1
			///made into tabel in word draft
regress z_BMI_t_3 nbi0115_t_1, vce(cluster fid)
regress z_BMI_t_3 nbi0115_t_1 i.sex_t_1 i.cat_age_m_atbirth_t i.cat_eca0100_m_1, vce(cluster fid)

regress z_BMI_t_3 ib3.cat3_nbi0115_t_1
regress z_BMI_t_3 ib2.cat3_nbi0115_t_1 i.sex_t_1 i.cat_age_m_atbirth_t i.cat_eca0100_m_1

by cat3_nbi0115_t_1, sort : regress z_BMI_t_3 nbi0115_t_1
by cat3_nbi0115_t_1, sort : regress z_BMI_t_3 nbi0115_t_1 i.sex_t_1 i.cat_age_m_atbirth_t i.cat_eca0100_m_1
			//interesting results "healthy birthweights are related with z-bmi"
```

12.1.2 Twins stratified

```
//start by dropping and generating
*drop cases without biological mother
drop if fpr0107_m_1 <=0
drop if fpr0107_m_1 == 23
                                        //1032 cases left
*drop cases without age of mother
drop if age0100_m_1 == -95
                                        //1031 cases left
*drop cases without education of mother
drop if eca0100_m_1 <= 0
                                        //1002 cases left

*generate AGE OF MOTHER AT BIRTH OF CHILD
generate age_m_atbirth_t = age0100_m_1-age0100_t_1
generate age_m_atbirth_u = age0100_m_1-age0100_u_1

*continious variables to be grouped : education mother & age of mother
summarize age_m_atbirth_t age_m_atbirth_u eca0100_m_1

egen cat_age_m_atbirth_t = cut(age_m_atbirth_t), at (0, 17, 25, 35, 50) icode
table cat_age_m_atbirth_t, contents(freq min age_m_atbirth_t max age_m_atbirth_t)
egen cat_age_m_atbirth_u = cut(age_m_atbirth_u), at (0, 17, 25, 35, 50) icode
table cat_age_m_atbirth_u, contents(freq min age_m_atbirth_u max age_m_atbirth_u)
                            //3 catgeories young 18-24, average 25-34, risk 35+
tab cat_age_m_atbirth_t
tab cat_age_m_atbirth_u

generate copy_eca0100_m_1 = eca0100_m_1
recode eca0100_m_1 6=1
egen cat_eca0100_m_1 = cut(eca0100_m_1), at (0, 1, 3, 4, 6) icode
by cgr_1, sort : table cat_eca0100_m_1, contents (freq min eca0100_m_1 max eca0100_m_1)
tab cat_eca0100_m_1
codebook eca0100_m_1
              //3 categories low edu (primary and other), medium (secondary), high (applied science & uni)
        //(1) left school without edu and primary/lower secondary/other;
                            //(2) intermediate secondary (grade 10)
                            //(3) univeristy of applied schience, upper secondary - university entrance diploma)

generate c_nbi0115_t_1 = nbi0115_t_1
egen cat3_nbi0115_t_1 = cut(c_nbi0115_t_1), at (0, 1, 1499, 2499, 3999) icode
table cat3_nbi0115_t_1, contents(freq min c_nbi0115_t_1 max c_nbi0115_t_1)
generate c_nbi0115_u_1 = nbi0115_u_1
egen cat3_nbi0115_u_1 = cut(c_nbi0115_u_1), at (0, 1, 1499, 2499, 3999) icode
table cat3_nbi0115_u_1, contents(freq min c_nbi0115_u_1 max c_nbi0115_u_1)

generate A_birthw=(nbi0115_t_1+nbi0115_u_1)/2
generate DT_birthw=nbi0115_t_1-A_birthw
generate DT_birthw_u=nbi0115_u_1-A_birthw

///regression simple///twin 1/////////////////////////////////////////////////////////////////////////
regress z_BMI_t_3 DT_birthw, vce(cluster fid)
estimates store m1, title (Model 1a)
regress z_BMI_t_3 DT_birthw i.sex_t_1 i.cat_age_m_atbirth_t i.cat_eca0100_m_1, vce(cluster fid)
estimates store m2, title (Model 1b)
estout m1 m2 , cells(b(star fmt(5)) se(par fmt(5)))   ///
```

```
  legend label varlabels(_cons constant)              ///
  stats(r2 N, fmt(4 0) label(R-sqr N))

///strata/////twin 1///////////////////////////////////////////////////////////////////////////
*birthweight
by cat3_nbi0115_t_1, sort : regress z_BMI_t_3 DT_birthw, vce(cluster fid)
by cat3_nbi0115_t_1, sort : regress z_BMI_t_3 DT_birthw i.sex_t_1 i.cat_age_m_atbirth_t i.cat_eca0100_m_1 ,
vce(cluster fid)

*cohort
by cgr_1, sort : regress z_BMI_t_3 DT_birthw, vce(cluster fid)
by cgr_1, sort : regress z_BMI_t_3 DT_birthw i.sex_t_1 i.cat_age_m_atbirth_t i.cat_eca0100_m_1, vce(cluster fid)

*age mother twin 1
by cat_age_m_atbirth_t, sort : regress z_BMI_t_3 DT_birthw, vce(cluster fid)
            //no significance simple
by cat_age_m_atbirth_t, sort : regress z_BMI_t_3 DT_birthw i.sex_t_1 i.cat_age_m_atbirth_t i.cat_eca0100_m_1,
vce(cluster fid)
            //significance DT for young mothers (79 cases)

*education of mother
by cat_eca0100_m_1, sort : regress z_BMI_t_3 DT_birthw, vce(cluster fid)
by cat_eca0100_m_1, sort : regress z_BMI_t_3 DT_birthw i.sex_t_1 i.cat_age_m_atbirth_t i.cat_eca0100_m_1,
vce(cluster fid)

*mono vs di
by zyg0102_1, sort : regress z_BMI_t_3 DT_birthw, vce(cluster fid)
by zyg0102_1, sort : regress z_BMI_t_3 DT_birthw i.sex_t_1 i.cat_age_m_atbirth_t i.cat_eca0100_m_1, vce(cluster
fid)

///regression simple///twin 2//////////////////////////////////////////////////////////////////////////////
regress z_BMI_u_3 DT_birthw_u, vce(cluster fid)
regress z_BMI_u_3 DT_birthw_u i.sex_u_1 i.cat_age_m_atbirth_u i.cat_eca0100_m_1, vce(cluster fid)
                        //add to appendix

by cgr_1, sort : tab eca0100_f_1
///strata/////twin 2///////////////////////////////////////////////////////////////////////////
*birthweight
by cat3_nbi0115_u_1, sort : regress z_BMI_u_3 DT_birthw_u, vce(cluster fid)
by cat3_nbi0115_u_1, sort : regress z_BMI_u_3 DT_birthw_u i.sex_u_1 i.cat_age_m_atbirth_u i.cat_eca0100_m_1,
vce(cluster fid)

*cohort
by cgr_1, sort : regress z_BMI_u_3 DT_birthw_u, vce(cluster fid)
by cgr_1, sort : regress z_BMI_u_3 DT_birthw_u i.sex_u_1 i.cat_age_m_atbirth_u i.cat_eca0100_m_1, vce(cluster
fid)

*education
by cat_eca0100_m_1, sort : regress z_BMI_u_3 DT_birthw_u, vce(cluster fid)
by cat_eca0100_m_1, sort : regress z_BMI_u_3 DT_birthw_u i.sex_u_1 i.cat_age_m_atbirth_u i.cat_eca0100_m_1,
vce(cluster fid)

*age mother twin 2
by cat_age_m_atbirth_u, sort : regress z_BMI_u_3 DT_birthw_u, vce(cluster fid)
by cat_age_m_atbirth_u, sort : regress z_BMI_u_3 DT_birthw_u i.sex_u_1 i.cat_age_m_atbirth_u
i.cat_eca0100_m_1, vce(cluster fid)
```

*mono vs di
by zyg0102_1, sort : regress z_BMI_u_3 DT_birthw_u, vce(cluster fid)
by zyg0102_1, sort : regress z_BMI_u_3 DT_birthw_u i.sex_u_1 i.cat_age_m_atbirth_u i.cat_eca0100_m_1, vce(cluster fid)

*sex
by sex_u_1, sort : regress z_BMI_u_3 DT_birthw_u, vce(cluster fid)
by sex_u_1, sort : regress z_BMI_u_3 DT_birthw_u i.sex_u_1 i.cat_age_m_atbirth_u i.cat_eca0100_m_1, vce(cluster fid)Appendix An appendix or epilogue, if relevant, should appear at the end of the manuscript. In the appendix, large or less important tables, figures, maps, or other illustrations are displayed. If a questionnaire has been used in the research, include it in the appendix.

12.2 Regression Results – Twin 2

Regression results from twin two are available on request. When handing in the thesis they will be attached to the thesis.