COPENHAGEN BUSINESS SCHOOL 2017

CAND.MERC.(MAT)

MASTER THESIS

Parental effects on children's high school educational attainment

---

Author: Diana Hornshøj JENSEN

Supervisors: Birthe Larsen and Lisbeth la Cour

Date of submission: October 16 2017

Characters: 166, 126 / Number of pages: 80

Preface

I would like to say a big thank you to my supervisors Birthe Larsen and Lisbeth la Cour, for an excellent academic sparring and for their encouraging approach to the subject and the whole process.

Furthermore, I owe the Department of Economics at Copenhagen Business School a big thank you for giving me the opportunity to work with register data from Statistics Denmark. As register data is very detailed, the data is not attached in the thesis.

The thesis is written in 11pt and with 2 cm margin in the sides and 3 cm in the top and bottom.
Data processing is in SAS 9.4 and the code is not attached, as all work is done from a server from Statistic Denmark.

Resumé

I Danmark er den økonomiske ulighed de senere år vokset og flere har relativt færre midler at gøre godt med. Heriblandt er særligt indvandrere. På trods af, at flere forskningsprojekter peger mod, at Danmark er et af de lande med de bedste muligheder for uddannelsesmobilitet, peger ny forskning mod, at med bestemte uddannelses- og indkomstsmål er unge danske ikke mere uddannelsesmobile end unge amerikanere. Ydermere er andelen af underklassens børn i Danmark, som selv ender i underklassen steget i perioden fra 2003 til 2013. Herudover er den stigende debat, om hvordan integrationen i Danmark går, samt de gentagende resultater, der peger på, at indvandrere klarer sig dårligere end indfødte i folkeskolen med til at øge debatten. Derfor er information på området af stor vigtighed og det har derfor været det interessant at undersøge, hvorvidt der er en sammenhæng i forældres uddannelses- og indkomstniveau og sandsynligheden for, om deres børn får en gymnasial uddannelse. Herudover at undersøge om der er forskelle i disse effekter givet, om forældrene er indfødte, om de er af anden etnisk herkomst, om det at man bor i en bestemt kommune har betydning på, hvor stor forældreindflydelsen er og slutteligt, om der er forskel i forældrenes effekt alt efter om vi ser på drenge eller piger, men også hvordan forældreeffekten varierer med oprindelsesland. Jeg finder i denne afhandling, at børn af indfødte i højere grad end børn af indvandrere har en gymnasial uddannelse. Herudover finder jeg et stærkt forholdet mellem forældres uddannelse og sandsynligheden for om indfødte piger og især drenge får en gymnasial uddannelse, samt at forholdet mellem forældrenes indkomst og sandsynligheden for uddannelse er mindre stærkt. For børn med indvandre baggrund er forholdet mellem forældres indkomst- og uddannelsesniveau og sandsynligheden væsentligt svagere uanset børnenes køn, men denne påvirkning er blevet kraftigere i perioden fra 2006 til 2016. Herudover konkluderer jeg, at forældreeffekten også afhænger af, om forældre bor sammen eller ej. På kommunalt niveau finder jeg, uventet, at andelen af indvandrere med en gymnasial uddannelse er højere, end andelen af børn med indfødte forældre i Brøndby og Ishøj.

Contents

Chapter 1

Introduction

1.1 Motivation

There were three main motivations behind this thesis. Firstly the findings in Landersø and Heckman (2017), which show that in essential areas, social heritage in Denmark is about the same as in the US. Secondly the increased debate about immigrants and their migration to Denmark and lastly the fact that the inequality in Denmark is increasing.

Education is an important and sensibly tool, if the goal is to reduce economic inequality. The reasons are many, and some of them are listed here: education increases an individual's competencies and thus productivity, so therefore there is a known connection between income and education. Moreover education will increase the flexibility of individuals and in some sence is a help to self-help.[1] Most Danes believe that everyone should have a fair chance in achieving an education. Denmark is among the countries which spend the most on education and are well known for its high level of welfare state (like the other Scandinavian countries). Therefore many turns their heads to these countries, when discussing a model for reducing inequality and promoting intergenerational mobility (see Bailey et al. (2011)). And from an economical perspective Denmark is one of the most equal countries in the world (thanks to the education system, according to Thomas Piketty (2014)). But even though these factors, it is a fact, that over the last two decades inequality in Denmark has been changing for the worse: the inequality in Denmark is growing and more Danes live in poverty.[2] According to *De Økonomiske Råd* (The Economic Advide in Denmark) the Gini-coefficient in Denmark has increased from 20 in 1990 to 27 in 2014, and in this period the group of individuals with the absolute highest incomes, has increased their proportion of the income from 0.5% in 1990 to 1.6% in 2014. The higher inequality is *not* only caused by an increase in the top of

---

[1] [49] Lighed gennem uddannelle

[2] https://www.ae.dk/publikationer/danmark-paa-fattigdomskurs
https://www.ae.dk/analyser/de-rigeste-omraader-slaar-rekord-mens-de-fattiges-indkomst-falder

the income distribution. The lower part of the income distribution has increased as well. Comparing the income median with the income of the 10% with the highest income, the relationship was 1.48 in 1990, while it was 1.7 in 2014 and comparing the income median to the income of 10% with the lowest income the relationship was 1.6 in 1990 and 1.78 in 2014.
Investigating the group with the lowest income in Denmark, one finds that the proportion of immigrants (and descendants) is 28% greater than the proportion in the population which is 11%.[3] Not only is the proportion of immigrants in the low income group higher than the national proportion of immigrant, but facts are also, that the average income before tax for immigrants is at a lower level than found for native origin.[4] Given the higher income inequality the lowest income-quantile of the Danish population will be at a lower level of income. This implies that a greater proportion of children faces a bigger challenge if they would like to take a higher education, as children from families with greater resources more often tend to get a higher education than children from less fortunate families (see Thomsen et al. (2016)). Given the relative high proportion of immigrants in the low income group, a great part of immigrants faces that challenge, as a basic sociological fact is the permanency of social inheritance. Evidence also strongly implies that social origins impact on life chances just as much today as seen in the past when it comes to educational attainment.[5]

The immigration into Denmark increases[6] and since facts are, that students with both 1st and 2nd generation immigrant status, have a lower score in the PISA-tests in math,[7] it becomes even more relevant to improve knowledge on immigrants' educational structure in Denmark. Especially as the Danish government platform from November 2016, stated as a goal, that at least 50 percent of 30-year-olds have completed tertiary education, and as many as possible complete their education in the prescribed time[8]. Furthermore a new High School reform (Gymnasiereformen) was developed in 2016. This new reform requires that all students who want a high school education (Gymnasie/HHX/HTX/HF), needs to have above a 5.0 average in proficiency marks from primary school.[9] This leaves the proportion of children from families with low income levels, with a probably bigger challenge yearly on in their lives, as they have to meet up these requirements if they want the opportunity of a higher education.

My primary motivation for this thesis is to investigate the parental effect on the probability of education for children in Denmark. Furthermore I find it very interesting to study whether differences in parental effects might appear between children of immigrant origin and native origin. Secondly, I have a strong desire to increase my knowledge of parental impact on the probability of education, how the

---

[3] [49], Indkomstfordelingen i Denmark

[4] See Rosdahl et al (2013) and Table 10.10 in [53]

[5] [21]

[6] Figure 1.5 in [36]

[7] [1]

[8] http://ufm.dk/minister-og-ministerium/regeringsgrundlag-vision-og-strategier/regeringsgrundlag-november-2016

[9] Aftale mellem regeringen, Socialdemokraterne, Dansk Folkeparti, Liberal Alliance, Det Radikale Venstre, Socialistisk Folkeparti og Det Konservative Folkeparti om styrkede gymnasiale uddannelser

migration in Denmark is doing in form of education and how time has changed the impact. Therefore the research question will be:

1.2 Research question

To what extent do parental income and level of education affect the probability of whether boys and girls respectively take a high school education in Denmark and are there similarities between children of immigrants and natives, and municipality of residence - can this relationship be explained? And how has the parental effect changed over time?

The research question will be answered by setting up a logistic regression and the use of registered data from Statistic Denmark. The details will be explained later in the thesis.

1.3 Delimitation of the thesis

The thesis is delimited in some forms of the analysis as it only looks at data within two different years. Furthermore the thesis will be an empirical study on parental effects on social, mostly educational, mobility and the main focus is to report the empirical relationships. Secondly focus will be on the causal. The thesis does not use parental income (including social services) as income over the life cycle of children, which has showed to explain most of the variation in the connection between parental income and children's schooling (see Carneiro and Heckman (2002)).

1.4 Structure of the thesis

Thesis outlines: In chapter 2 there will be a review of the childcare and school system in the Danish welfare state and a very brief review of the municipalities in Denmark, the thesis will works with. Finally, parts of the motivation section will be elaborated followed and supported by some of the existing literature in the field of education. In chapter 3 an introduction to the empirical work, with a review of data is presented and there the stereotype families, who are used to present how children's probability of a high school education depends on the parents' affects will be presented. In chapter 4 theory of the logistic regression and the Maximum Likelihood Estimation method is presented, the model is set up in the benchmark scenario and analyzed by different robustness analysis and chapter 5 concludes.

Chapter 2

Explanation of the key concepts and Literature

This section will provide the reader with some knowledge about the welfare system in Denmark to help the frame and the understanding the analysis and lastly it will give a review of some of the literature in the field.

2.1 The welfare system in Denmark

The danish welfare state is by definition redistributive as it simultaneously allocates and transfers money to and from citizens, but it does not mean that it automatically reduces income inequalities. It is, as mentioned earlier, internationally prominent because of its relatively high level of expense and taxation, its degree of universality and its equalizing effect. In other words: the welfare state are playing a double role, both as piggy bank and as Robin Hood.[1] In Denmark individuals can be financially supported by the government trough the welfare state and receive social services, which is not for everyone and the rules are very strict and specific. Therefore they are not presented in the thesis. But this is mentioned, as many low income families are getting their main income from the social services.

2.1.1 The Day Care System in Denmark

In Denmark most children are taken care of outside their homes from they are one year old and when the child is younger than 3 years old they in daycare, which can be a private daycare or a public daycare. When children are older than 3 years old continue in kindergarten until schools starts at age 5 or 6. The daycare and education system is a big part of the welfare system in Denmark, and daycare is offered to all children in Denmark. The price of the daycare vary from municipality to municipality and from family to family. In every municipality the cost of a daycare depends on the level of parental income and

---

[1] [3]

for example if a parent is a single parent they will receive discount. For example in the municipality of Copenhagen, daycare cost approximately 3, 200 DKR, and if the total household income is below 528, 200 DKR the cost will be reduced.

Among Danish 1 year old children, it is 85%, who are in daycare and when they are 3 years old it is almost every child, who are in daycare (99%). But among children from first and second generation immigrants origin homes, the proportion differ from the majority. For families with immigrant origin it is 6% of 3 years old children who are are taken care of at home. And for mothers whose highest level of education is primary. And if one is looking at all mothers, who has primary school as highest level of education it is 6% of the 3 years old children whom are taken care of at home. Furthermore 7% of the children with parents outside the labor force who are not in daycare. Corresponding to this, some of these children probably are facing a greater challenge when it comes social mobility as, Heckman and Lochner (2000) with a great number of evaluations have found that, the prerequisites for learning and acquiring skills lie very early in childhood. I will return to this in chapter 3.

2.1.2 School System in Denmark

In 1903 the Danish education system changed to a single school system, with the law of public schools. This law was epoch-making, as it gave girls access to the public upper secondary school/high school (Gymnasieskolen), which they previously only had access to from private education, but also because it created a connection between school forms. High school was until then eight years long and then a six years education, and it was a school you started and ended without touching primary school. In 1903, high school became a three-year superstructure at the middle school and whose foundation was primary school. It was a democratization of access to high school, based on the wish that all those who had the skills should have access to higher education for the benefit of themselves and society. But in order to ensure that only the best ones achieved the target, there were a number of access tests that acted as a sorting mechanism. From 2019 one must have more than 5 on average in year grade and at least have an average of 3 in primary school graduation to get directly to high school, furthermore one must fulfill some social and personal prerequisites.[2]

The education system in Denmark is as following: At age 5 or 6 the children starts in 0th grade which is a kindergarten grade. Kindergarten grade is the first part of the compulsory primary school (Folkeskolen). Then follows nine years in primary school from 1th – 9th grade. It is the grades obtained in the final year, 9th grade, (Folkeskolen afgangseksamen) which serves the distinctive criterion for the acceptance in the further school system, unless the child chooses to continue primary schooling in 10th. In Denmark the first 10 years of schooling is obligatory, i.e. after 9th grade it is optional to continue in primary school in 10th grade, or to go directly to high school or another preferred path.

---

[2] For HF the average year grade is at least 4.
https://www.ug.dk/6til10klasse/optagelse-til-de-gymnasiale-uddannelser

After finishing 9th or 10th grade, it is possible to start in high school (Gymnasium/HHX)[3], which in Denmark takes 3 years. High school provides admission to Universities. Typically one will be accepted in university based on high school grades, but in a smaller proportion of cases it is possible to be admitted via quota 2., which is a mix of grades from high school, experience and motivated application. Therefore high school attendence is important for further during in life. If a child starts at age 6 in primary school it is finished at age 15 and then continue in boarding school before high school the child will be 19 years old, when finishing high school. Therefore individuals who are $20 - 30$ years old are chosen to work with. According to Thomsen and Andrade (2016), 5% of 35 years old Danes had not completed primary school in 2013.

2.1.3 Municipalities

Denmark has 98 municipalities and as in many other countries, there are big differences across the country within the municipalities. In this thesis I have chosen to work with four different municipalities. Brøndby and Ishøj are the first two, as these are the to municipalities with the highest proportion of immigrants. Kolding and Vejle are the last two, these two are chosen to represent "average" municipalities, as their proportion of immigrants are approximately the proportion of immigrant at the national level. Individuals in Brøndby and Ishøj are put together in one dataset, and individuals in Kolding and Vejle are put together in another, as individuals and the municipalities are assumed more or less alike.

2.2 Literature

This section contains a review of some of the literature in the field to help frame the analysis. As mentioned in the introducing section, Denmark is one of the most equal countries in the world seen from an economical perspective. But as in many other countries, the income inequality in Denmark is an increasing factor. Not as remarkable as in some of the countries that Thomas Piketty presents in the *Capital in the twenty-first century*, but facts shows, that over the last two decades inequality in Denmark has been changing for the worse. The Gini-coefficient has increased 25% points since the 1990's.[4] Regardless of whether you look before or after tax.[5] The higher income inequality implies that the lowest income-quantile of the Danish population will be at a lower level of income, i.e. the resources for these families will be at a lower level and there will be more of these families. Even though Esping-Andersen (2015) found, that parental background, from low resources families, is a little less important in Denmark and the other Scandinavian countries, than in the rest of the world, when it comes to intergenerational mobility, studies have found that the risk of children from low class families ending up in a low class family is 4.6 timer higher than children from higher class families ending up as a lower class family today.

---

[3] Or HF if one has taken 10th grade, only takes 2 years

[4] The Gini-coefficient is the most common measure for income inequality, if it is zero, there is complete equality of income.

[5] [49], Lighed og Skat

This factor was 2.8 in 2003.[6] Further more have Landersø and Heckmand (2017) showed that in some measurements Denmark is not more social mobile than USA. The increased inequality implies, that a greater proportion of children faces a bigger challenge if they would like to take a higher education, as children from families with greater resources more often tend to get a higher education than children from less fortunate families (see Thomsen et al. (2016)). And other studies (like Jantti et al. (2006)) showed, that one out of four sons will end up in a low income family, if their father belonged to the low income group. The Danish daycare- and school system is as mentioned earlier very expensive. But it is not for nothing. Many studies point to the fact, that attending high-quality institutions (as the Danish) contribute very well to children's later capacity in school and even through adulthood (See Corak (2005), Waldfogel (2005) and Dännrich and Esping-Andersen (2017)) and for low income families who sends their children to daycare, the bright side is the facts, that if the welfare state invests in childhood, as the Danish welfare state does, this is the best way to improve equal opportunities (see Esping-Andersen (2016) and Heckman and Krueger 2004).

Another reason why the field is important to have as much information about social mobility is, that a study (Fischer (2009)) showed that actual social mobility in society, measured by intergenerational earnings, elasticity and intergenerational dependence of the student's attainment, is positively correlated with social well-being, both for the well-being of society as a whole but also for individuals' social well-being, which might leave the reader wondering if there are a correlation within these facts and the fact that Denmark no longer is the happiest country in the world[7].

Denmark is one of the countries who spends the most on the education system many have found the "payoff" and the expensive school system interesting to study. Historically the education level among the Danish population has increased significantly since the 1980's[8], and young people from non-academic homes, ie. children having unskilled (non- or low educated) parents have received further education, which is great. Especially in the Danish society, which increasingly is based on knowledge and high technology, and therefore the level of education among the population plays a key role in ensuring high employment and productivity. In highly specialized societies, highly skilled workers are required and benefit the individuals self and the community, that everyone are given the opportunity of education. As education itself is costless in Denmark, hence the costs of acquiring education will be alternative costs, like what could have been earned. Therefore one could argue, that in some sense, all Danish children are born with equal opportunities for achieving education, but given it is usually found that parental education is very important to the fact of children are education them self, both because the young person learn about education from the parents, but also because educated parents usually are more well-off parents and therefore may be able to help economically while the young person is studying. See Erola et al. (2016), Esping-Andersen

---

[6] [37]

[7] https://www.thelocal.dk/20170320/denmark-no-longer-worlds-happiest-country-report

[8] jacobsen2004befolkningens, Figure 4.3 (2004)

(2004) and Bukodi and Goldthorpe (2012). Further more several studies through decades have shown that family income is highly correlated with education level of the family see Black and Devereux (2010) and Jæger and Holm (2007). These acknowledgments and the increasing gab in the income inequality might result in greater challenges for Danish the welfare system and the high quality it is known for.

Esping-Andersen (2015) studied and found that the Nordic welfare states, and the costless school system, have set lower barriers to educational attainment and social mobility for the working-class kids. Others have studied intergenerational mobility in Denmark and compared to what it is in other counties. Among those are the study, which motivated me to this thesis: Landersøe and Heckman (2017). In the paper Landersø and Heckmand studied whether there are major disparities in intergenerational- and educational mobility in United States and in Denmark. They examined the sources of differences in the social mobility and found that Denmark is a more mobile society than the U.S, if the measurement is not by educational mobility. Furthermore they found, that beside the more generous educational and childcare policy in Denmark, that young Danes educational pattern are similar to the pattern of young Americans: they found that the average educational mobility is remarkable similar in these two countries. Among their findings was that only 7% of Danes between 20-34, without parents with very low or non education, are enrolled in or have completed a tertiary education, which is just one percent lower than what it is in the United States. They found clear evidence that in both countries, the inclination to get an education depends on parental income and education and that the educational mobility in the two countries do not differ from each other. This started the thoughts of how parental effects, as income and educational influenced the educational choices in Denmark. And how these might differ across country of origin.

Even though the PISA 2015 report presents Denmark as one of the countries that achieve high levels of performance and equity in education outcomes see Gurria (2016). It is found in the PISA Ethnic report based on data from 2015[9], that there are large differences in test scores among the young Danes when it comes to country of origin. Beside all the above-mentioned educational factors, the thesis are working with differences which might appear between individuals of immigrant and native origin. This is motivated by the major focus on immigration and how integration in Denmark develops, this is (still) a very discussed subject in Denmark. Both in politics but also by citizens. And now, in this thesis.

The immigration into Denmark increases[10], and the immigrants are mostly from non-Western countries, and since it is a fact that students with both 1st and 2nd generation immigrant status, have a lower score in the PISA-tests in math,[11] it becomes even more relevant to improve knowledge on immigrants' educational structure and how it depends on parental factors in Denmark. Also since new studies looking at second-generation immigrants from different nationalities find that those whose parents come from high-scoring countries in the PISA test do better in school performance than their peers. Furthermore

---

[9] [14] and [28]

[10] [36]

[11] [1, (PISA 2015)]

they find a larger gap among students whose parents have poor education and have recently emigrated, suggesting the importance of country-specific cultural traits that parents progressively lose as they integrate in the new host countries (see De Philippis et al. (2016)). Therefore, if the knowledge of educational mobility (social mobility) is very specific, it will enlighten where to focus to improve, realizing these results.

Moreover as the test scores in primary school show that in average boys score higher than girls in math no matter the parental background of the students, differences across boys/girls become interesting to study. As it shows that immigrants in general are performing lower than natives. This might be correlated with the years since parents' migrated, since a positive correlation between parents' migration and children's academic achievement exists (see Skyt Nielsen and Schindler Rangvid (2011)). Furthermore it is found in the PISA data, that speaking Danish at home, makes a quite big difference to the better in test scores.

Somehow the tables are turned in high school, as girls after high school graduation perform better than the boys, in average. Research from DEA based on data from 2013 shows girls and boys with a primary school grade average at 8.0 (12 scale) will have a lower average after graduation high school and that, girls graduating from a 3 years long high school (STX/HHX/HTX) have a higher average than boys graduation. But the case is different, when looking at students graduating from 2 years high school (HF). Here boys are performing a tiny bit better than girls, but at the same time, both boys and girls are doing better than they where in primary school[12]. Students can only start in HF if they have attended 10th grade or have been working or doing something else for a few years. Besides these findings their calculations also show, that the differences between in grades, ie. the difference between the performance of girls and boys in high school has increased since 2006. So in general girls perform better in high school than in primary school, and boys performs at lower levels.

Furthermore, working studies from Denmark[13] shows, that in Denmark looking at high school attendance among immigrants, there are reasons looking at different municipalities, as a study points to that there are little evidence of correlation between negative attitudes and proportion of immigrants attending high school. But also finds, that network, ie. fraction of citizens in once municipality from with equal home country, has some influence on high school attendance. An increasing productivity is very important for the future of the welfare state in the Denmark (and the other Scandinavian countries) as there is a direct link with the development of labor productivity and the welfare state. As mentioned earlier, Danes government has as a goal that a great part of the young Danes get an education, and since a direct link between human capital, in terms of educational levels, and labor productivity[14] is found, it does not seem as an indifferent goal.

---

[12] [19]

[13] [5]

[14] [44]

Chapter 3

Introduction to the empirical work

In this section a detailed description of the register data and the stereotype families, which will be used in the analysis in the following chapter are presented.

3.1 Data

The data is Danish administrative register data from Statistics Denmark and is covering 100% of the population. Danish Register Data in Denmark is very rich and the level of details are of excellent quality compared to many other countries. The data is used by the Danish government, and researchers from around the world often prefer working with Danish register data, given the degree of details available, which can result in very detailed and unique projects and findings. This paper uses registers annual information and the level of details provides a good opportunity to examine what influence parental level of education and income have on whether young Danes choose to attend and complete high school or not. Furthermore it provides the possibility of investigating which impact, if any, municipality of residence has on these effects, i.e. if parental impacts differ across a few municipalities. Since the administrative register data contains clear information on individual's nationality, i.e. one can examen and differ individuals based on country of origin and further whether individuals are 1st or 2nd generation immigrant it makes the purpose of this paper possible to examine.

Statistics Denmark has anonymous individual identification numbers for each Dane, through family links parents are connected to their children with a mother/father anonymous identification number. If parents are registered of course (not all immigrants have both, if any, parents registered, furthermore other reasons why not all individuals have registered both parents could of course also be the case). As mentioned earlier, the level of information in the danish register data is very high, which imply that it is possible to find exactly the specific education an individual has as highest completed, and every education has a number. Therefore I will be setting up seven categorize of level of education, which will be explained later.

A handful of socio-economic variables from the register data are considered: age and gender of the young individuals, furthermore the individual's level of education (highest completed education). As the individuals easily can be linked with their parents, also parents' level of educational are used. Besides those variables, the socio-economic-variables containing nationality is used to specify if the individuals are natives or immigrants, in this paper. I simply distinguish between native and immigrants. This is done, as 90% of immigrants in Denmark are non-Western. Lastly in the thesis I will have a short review of the differences, which might appear when looking closer whether country of origin matter. Furthermore a variable listing individuals municipality of residence and a variable telling if an individuals parents live together are used. The variable used for controlling whether parents are together, is regardless of if they are married or if they are just living together, i.e. if parents have the same address they are noted as living together.

At first an income variable listing sum of salary for all appointments doing the year was used as parental income information. But as the analysis started I ran into troubles, as more than half of the immigrated children' parents, and approximately half of the native children' parents had no income from a job. Therefore I ended up using an variable including social services (kontanthjælp) but excluding property- and wealth income and before deduction of labor market contributions and special pension contributions, as the variable of parental income.

As I have access to the full population, it is possible to connect the young Danes to their parent(s) through family links, i.e. link to identification of children and parents of the population in the given years, 2006 and 2016. These two years are chosen as 2016 being the latest data and as I want to include 10 years of individuals in the analysis, 2006 become the other year. After finding the identification numbers of all the mothers, I created new variables containing her income and level of education. Afterwards the same was done for all identified fathers. Information on the level of parental income and education was found though the relevant variables and new variables containing parental income and level of educations was created. There are many different ways of working with the aspect of parental income, one could be working with it as log transformed data, but here I choose to set up groups containing specific levels of income, given the respective income distribution.

Before looking into the income distributions a list of details about the dataset are presented. The full dataset sample size is $5,707,251$ individuals in 2016 and $5,423,347$ individuals in 2006. As I am only interested in young Danes, who are old enough to have completed high school, I choose to keep every individual born between in the years $1986-1996$ in the dataset from 2016 and $1976-1986$ in the dataset from 2006. The large dataset sample are chosen on purpose, as the proportion of immigrants are relatively low (8.402% in 2006 and 12.33% in 2016). As I will be working with few municipality levels, it is necessary to have a relative large sample size if any significant story are to bee told. By removing all Danes too young to have completed a high school education and all of those "too old" of interest, the datasets now contains $748,409$ individuals in the sample from 2016 and $622,541$ individuals in 2006. In this way, we end up with a smaller dataset containing information about a group of young Danes whom just have or within a 10 years have finished high school education. After that, individuals with no registered parents in Denmark or with parents with negative income are deleted from the datasets too. After this, I end up having a datasets of $646,065$ and $569,017$ individuals in respectively the dataset from 2016 and 2006 with a parent-child match. The proportion of immigrants in the chosen sample dataset from 2016 are 10.43% and in 2006 it is only 6.36%.

Later on in the paper, there will be some robustness tests, among others to check, if children with only one registered parent significant differ from the results based on the datasets found above.

Parental income

After the reduction of datasets from the two chosen years, so that they only contains the individuals of interest, it is time to have a closer look to parental income. As mentioned earlier on in this paper, the work with parental income will be based on groups of income. To have the opportunity of working with groups new variables created. One variable for motherly groups and one for fatherly groups based on income. Since men still earns more that women in Denmark, it is necessary to divide the income distributions, so that we will be working with one for mothers and one for fathers.[1]. These created variables will tell which income group mother'/father' with a certain level of income will belong to. As these groups are based on income distributions, it is necessary to split each of the datasets up into two new datasets, given that income level of natives tends to be greater than level of income for immigrants[2]. I chose to set up one dataset containing natives and one containing immigrants to make it more simple doing the different analysis. When these datasets are set up, the income distributions for mother' and father' for both immigrants and natives are found. These distributions are the basis of all the groups of income.

These groups of income are based on the respective income distributions and are inspired by the work Niels Plaug has done in *Økonomisk Ulighed i Danmark*. The income groups in the paper are chosen so there are a group of mothers and fathers with low, low-middle, high-middel, high and the one percent's richest income compared to other parents with same gender and same history. These group' of income provides an opportunity to have some specification though the analysis. In every case, I have chosen to work with these 5 groups of income.

In all of the datasets the set up for groups of income, are as following: income group 1 is created so it will contain all parents (i.e. mothers or fathers) in the lowest income-fraction, as it contains parents with an income in the $0-25\%$ fraction of the income distribution. Income group 2 contains parents with an income within $25-50\%$ fraction of their gender. Income group 3 will contain the fraction of $50-75\%$, group 4 $75-99\%$ and then group 5 will contain the one percent's with the highest income. Income group

---

[1] http://www.dst.dk/da/Statistik/nyt/NytHtml?cid=23274: "Selvom forskellen mellem mænds og kvinders indkomster er blevet mindre de seneste årtier, så er mændenes indkomst fortsat 21 pct. højere end kvindernes i 2015. Højere erhvervsindkomst bidrager med 23 procentpoint til denne forskel. Det skyldes blandt andet, at mænd har en højere beskæftigelsesfrekvens, længere arbejdstid, tager mindre barsel og får en højere timeløn"

[2] http://www.business.dk/arbejdsmarked/indvandrere-faar-mindre-i-loen-end-danskere-med-samme-uddannelse

5 is mostly in this paper for investigating if there are differences in parental effect if your parents are rich (group 4) or very, very rich (group 5). Children of this group might act differently than the others, as their parents are so rich they probably do not have to worry about their own or their future family financial future.

After changing the income variable from only total salary, to also containing financial social services from the government, there still where some parents with zero or negative income. Al those with negative income are deleted from the data samples.

(Parental) Education

The level of parental education is as mentioned earlier also divided into groups. Again this is done to provide some sort of simplification, but still being able to find if parental level of education have some significant effects. These groups are based on parents highest completed education and the information are transformed in to new variables, i.e. level of education motherly/fatherly and children's.

The level of education (both parent's and children's) are split into 6 group: The first group, group number one, contains those who has primary school (mandatory in Denmark) as the highest completed education, the second group contains all craftsmen' educations ("Trade"-group), i.e. these two categories contains those who have not completed any kind of a high school education. Furthermore follows group number three, which contains individuals/ parents with high school (Gymnasium, HF, HHX or HTX) as the highest completed education. Group four contains people with a 2 year university degree, and then the two groups, five contains parents with respectively bachelor and then master and PhD in group six, as highest completed education.

It is known that for immigrants there can be some measurement insecurity within level of education obtained outside Denmark .

Assumptions: If an individual or parents have not got any registered level of education, I assume, that their level of education is primary school is the highest completed. All children with above high school education, are assumed to have completed high school.

Variable used from Statistics Denmark

Times: PNR, far_id, mor_id, IE_TYPE, OPR_LAND, KOM, FOED_DAG, EFALLE, AUDD, HFAUDD, fsp1e, PERSONINDK and (LONIND)
Created variables: "UddannelseVar" (Specific education) then "gym"/"highschool" to tell if the individuals have graduated high school or not. "i_lon", income from salary, "i_ind", income including social services, "i_udd" highest education completed, where "i" is mother and father. From these variables the grouping variables "i_longrp" and "i_indkgrp". Furthermore is a lot of dummi variables created to make the analysis easier to interpret to all levels of parental income and education, gender and country of origin.

3.1.1 The stereotype families

To evaluate the effect of parental level of income and education on whether young Danish children gets a high school education or not, the thesis will be working with stereotype families. This is done, as in the work with logistic regression the estimated coefficients are not easy to interpret and because all of the explanatory variables are categorical stereotype families make sense. In the following tables, the stereotype families are explained in details.

Table 3.1: Detailed information on stereotype families 1-6

Stereotype Family	The changing parental effect
1 (low)	All parental levels of income and education equal to 1
2 (low)	All parental levels of income and education equal to 2
3 (average)	All parental levels of income and education equal to 3
4 (average)	All parental levels of income and education equal to 4
5 (high)	All parental levels of income and education equal to 5
6 (high)	All parental levels of income equal to 5 and all levels of education equal to 6

Table 3.2: Detailed information on stereotype families 7-12

Stereotype Family	The changing parental effect	Notes
7	Fatherly education	Level of fatherly education equal to 3, all other parental effects equal to 1
8	Motherly education	Level of motherly education equal to 3, all other parental effects equal to 1
9	Motherly education	Level of mother education equal to 6, all other parental effects equal to 1
10	Fatherly education	Level of fatherly education equal to 6, all other parental effects equal to 1
11	Fatherly income group	Level of fatherly income equal to 4, all other parental effects equal to 1
12	Motherly income group	Level of fatherly education equal to 4 all other parental effects equal to 1

Chapter 4

Empirical work

This chapter contains all the empirical work. Firstly the logistic regression model be presented. Then the benchmark case, including al individuals are presented, followed by the case without individuals with only one parent registered in Denmark. Then the cases where parents do/do not live together is presented. The next scenario is a closer look at a few munitipalities and how parental effects effect the probability of education followed by, the analysis of how the effects differ with country of origin. Al the estimated model are documented in the appendix. Appendix A contains the benchmark case. Appendix B the scenario with only one parent registered in Denmark. Appendix C and D contains the models estimated for one/two households, appendix E and F for municipalities and finally the appendix G document the Western/non-Western case.

4.1 Model

The model for whether a child chose high school education or not, is done by the assumption that all individuals are rational agents, who search to maximize their lifetime utility. Whether they choose to attain high school depend on what maximize their lifetime utility. This is inspired by the Roy Model of education decision-making, originally suggested by Willis and Rosen (1979). The set up of the equation will mainly be a similar model for boys and girls, but with some differences for natives and immigrants. The aim of the equation is to examine how parental factors influence whether young Danes decide to get an education or not.

If young people achieve a decent education, this will increase the utility in form of higher productivity and expected wages, furthermore there will be a lower expected unemployment rate in the future (see. Becker and Tomes (1979)). The model may differ for immigrants and natives, in the way, that other factors will also have an impact on the immigrants' choice. A factor such as having a job put them in a better position when applying for citizenship, but also some differences in cultures as women are not suppose to work and there might be differences in norms.

To study intergeneration mobility in Denmark I am going to estimate a logistical regression model for the binary response variable. The depended variable (the response variable) is binary as it only takes two values, zero or one.

4.1.1 The logistic regression model

The empirical equation in this paper, is to be set up so the response variable so it will be the probability that young Danes achieve a high school education. The explanatory variables as well as the parents' income and the level of education. Inspired by Milhøj (1998) and Verbeek (2004) the section is set up. From this follows, that the goal is to find the characteristics stories that the explanatory variables $x_1, ..., x_p$ tells. If the explanatory variables are categorical, they are assumed presented as dummy variables. When a response variable, y_i, is binary it takes the value zero or one as it follows.

$$y_i = \begin{cases} 0 & \text{Have not graduated from high school "Failure"} \\ 1 & \text{Have graduated from high school - "Success"} \end{cases}$$

In this case, y_i is a realization of, in theory, a random variable Y_i that takes the value one with probability π_i and zero with probability $(1 - \pi_i)$ and then one observation in the data takes the form

$$(y_i, x_{i1}, ..., x_{ip})$$

which implies that data consists n sets of observations of $p+1$ variables, where each observation represents an individual, i.e. a young Dane. Furthermore it is assumed that $y_1, ..., y_n$ are observations of independent stochastic variables. From the details above it follows:

$$P\{Y_i = y_i\} = (\underbrace{\pi_i}_{p(y_i)=1})^{y_i}(\underbrace{1 - \pi_i}_{p(y_i)=0})^{1-y_i}, \quad y_i = 0, 1$$

As probabilities can not assume negative values or values larger than one, a linear function for this parameterization, as the following, can not be used. A linear function is the simplest idea and are set up as a regression model to explain y_i from the explanatory variables in the vector $\mathbf{x}_{ip}$, and are given as

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_p x_{ip} + u_i = \mathbf{x}_i^T \boldsymbol{\beta} + \mu_i \tag{4.1}$$

where $\mathbf{x}_{ip}^T = (x_{i1}, x_{i2}, ..., x_{ip})$. As y is binary, i.e. takes the value zero or one, β_j can not be interpreted as the change in y given a one-unit increase in x_j, ceteris paribus. Either y is unchanged or it changes from zero to one, or from one to zero, which could be a problem. This does not mean that, β_j does not still have some useful interpretation. Some further explanations are made before moving on.

Assuming that the error term has an expected value of zero given any values of the independent variables,

$E(u_i|\mathbf{x_i}) = 0$ such that the equation for the expected value of y_i, given $\mathbf{x}_i$ as following

$$E(y_i|\mathbf{x}_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_p x_{ip} = \mathbf{x}_i^T \boldsymbol{\beta}$$

where $\mathbf{x}$ contains all the explanatory variables.
Then the probability of "success" $y = 1$ (high school graduated) is the same as the expected value of y, and $P(y_i = 1|\mathbf{x}_i) = E(y|\mathbf{x}_i)$ will always be true. Therefore the following equation, also called a *binary response model*, can be set up

$$E[y_i] = P(y_i = 0) \cdot 0 + P(y_i = 1) \cdot 1 = P(y_i = 1|\mathbf{x_i}) = \mathbf{x}_i^T \boldsymbol{\beta} \tag{4.2}$$

where $P(y_i = 1|\mathbf{x_i})$ is the *response probability* and the aim to predict. Equation 4.2 says that the probability of success $p(x_i) = P(y_i = 1|\mathbf{x}_i)$ is a linear function of the x_j and as probabilities must always sum to one, it implies that $P(y_i = 0|\mathbf{x_i}) = 1 - P(y_i = 1|\mathbf{x}_i)$ is also a linear function of the x_j.
Beside the fact that $\mathbf{x}_i\boldsymbol{\beta}$ should lie between 0 and 1, there is another fundamental problem: the error term in equation (4.1) has a highly non-normal distribution and therefore suffers from heteroskedasticity.[1] As the response variable only takes two values zero and one, the error term, for a given value of x_i also only has two possible outcomes and the distribution of u_i can be summarized as

$$P\{u_i = -\mathbf{x}_i^T \boldsymbol{\beta}|\mathbf{x_i}\} = P\{y_i = 0|\mathbf{x_i}\} = 1 - \mathbf{x}_i^T \boldsymbol{\beta}$$
$$P\{u_i = 1 - \mathbf{x}_i^T \boldsymbol{\beta}|\mathbf{x_i}\} = P\{y_i = 1|\mathbf{x_i}\} = \mathbf{x}_i^T \boldsymbol{\beta}$$

which implies that the variance of the error term depends upon the model parameter $\boldsymbol{\beta}$ and upon explanatory variables according to $\mathrm{Var}[u_i|\mathbf{x}_i] = \mathbf{x}_i^T \boldsymbol{\beta}(1 - \mathbf{x}_i^T \boldsymbol{\beta})$ and therefore can not be constant.

As the aim of this paper is to find the connection and correlations between the explanatory variables and the response variable, and as this is not necessarily the linear connection, a transformation of the linear expression must be used, so the result can be perceived as a probability and to avoid linear probability model limitations. I consider a class of binary response models with the form

$$\pi_i \equiv P(y_i = 1|\mathbf{x}_i) = G(\underbrace{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_p x_{ip}}_{z}) = G(\beta_0 + \mathbf{x}_i^T \boldsymbol{\beta}) \tag{4.3}$$

where G is a function taking on the values between zero and one, so that $0 < G(z) < 1$ for all real numbers of z and to ensure that $0 \leq \pi \leq 1$ it is natural to assume that $G(\cdot)$ is a cumulative distribution function (cdf). When $G(\cdot)$ is a cdf, this cdf is only used for modeling the parameter, π and will not denote the cdf of y itself.
Furthermore $\mathbf{x}_i$ is the vector of covariates and $\boldsymbol{\beta}$ is a vector of regression coefficients. This defines the

---

[1] Verbeek [55]

systematic structure of the model. As there are a few different ways[2] in working with this non-linear function, G, the *Logit Model* approach is used in this paper. The other more common approach is the *Probit Model*, which specify the conditional probability

$$\pi \equiv \Phi(\mathbf{x}^T \boldsymbol{\beta}) = \int_{-\infty}^{\mathbf{x}^T \boldsymbol{\beta}} \phi(z) dz$$

where $\Phi(\cdot)$ is the standard normal cdf, with the standard normal density function $\phi(z) = 1/\sqrt{2\pi} \exp(-z^2/2)$. Why the Logit approach over Probit approach, one might think. Many has studied this field (see Cox (1966), Chambers and Cox (1967) and Chen and Tsurumi (2010)), and to my acknowledgment there still are no clear answer to which approach is the best. The expected value of a standard normal and a standard logistic random variable are zero, but their variances differ. A standard normal variable has variance 1, while a standard logistic random variable has variance $\pi^2/3$. Therefore the two distribution functions are very much similar, when corrected for the difference in scaling, but the logistic model (Logit) has slightly heavier tails, than the Probit model whose curves approaches the axes more quickly. Therefore it is common that in empirical work with those two models yield very similar results. Using Chen and Tsurumi (2010) one can argue that as long as the binary data, which is being modeled, are "balanced", meaning that it is roughly a $50 - 50$ split between the zero and one, then the information criteria does discriminating properly between Logit and Probit models, i.e. one can use both. Later in the thesis, I will show that there the binary response variable, is roughly "balanced".

Supposing that the *Logit* of the underlaying probability π_i is a linear function of the predictors, which is found from taking the logarithm of the odds

$$\text{logit}(\pi_i) = \log\Big(\frac{\pi_i}{1 - \pi_i}\Big) = \mathbf{x}_i^T \boldsymbol{\beta} \tag{4.4}$$

remember that $\pi_i = P(y_i = 1|\mathbf{x}_i)$ and again with $\mathbf{x}$ as a vector of covariates and $\boldsymbol{\beta}$ is a vector containing regression coefficients, this has the effect of removing the lower restriction, since the probability goes down to zero the odds approaches zero and the logit approaches $-\infty$. And as the probability goes to one, the other extreme, the odds approach $+\infty$ and likewise for the *logit*. Thus we have, the map of probability given by the *logits*, which range $(0, 1)$. If the value of the *logit* is less than zero, the probability represents less than a half, and all the values of *logit* above zero represents the probability higher than a half.

Then one can move from probability to odds by taking the exponential of equation (4.4)

$$\text{odds}_i = \frac{\pi_i}{1 - \pi_i} = e^{\mathbf{x_i}^T \boldsymbol{\beta}}$$

---

[2] Probit model, Logit model or a third choice could be Uniform Distribution over the interval [0,1]

which defines the ratio of the probability to its complements, or the ratio of favorable to unfavorable cases. The odds have no ceiling restrictions, as they may take any positive value. Thus, β_j represents the change in the logit of the probability associated with a one unit change in the j-th predictor holding all other predictors constant. If one solves for the probability

$$\pi_i = P(y_i = 1) = \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta}}}{1 + e^{\mathbf{x}_i^T \boldsymbol{\beta}}} = G(\mathbf{x}_i^T \boldsymbol{\beta}) \tag{4.5}$$

where $\mathbf{x}_i^T = (x_{i1}, ..., x_{ip})$. One can see that the left-hand-side is in familiar probability scale, and the right-hand-side is a non-linear function of predictors. Furthermore it becomes clear, that there are no simple way to express the effect on the probability of increasing a predictor by one unit while ceteris paribus. $G(\cdot)$ is the logistic cumulative distribution function (cdf).

An approximately answer can be obtained by taking derivatives with respect to x_j, but this would only make sense for continuous variables, which will not be relevant in this thesis, as explanatory variables are categorical.

One way to interpret the parameters is to consider the marginal effects of changes in the explanatory variables. If x_{ik} is a continuous explanatory variable, the marginal effect will be defined as

$$\frac{\partial G(\mathbf{x}_i^T \boldsymbol{\beta})}{\partial x_{ik}} = \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta}}}{(1 + e^{\mathbf{x}_i^T \boldsymbol{\beta}})^2} \beta_k$$

i.e. as the partial derivative of the probabiity that y_i equals one. Empirically, the marginal effect is typically defined at the "average" observation, then replacing x_i in the previous expression with the sample average.

For discrete explanatory variables, the effect of a change can be determined computing the implied probabilities for the two different outcomes, fixing the values of all the other explanatory variables. As the model is non-linear in parameters which have to be estimated, they can be estimated by the Maximum Likelihood method (derived in the following subsection). This provides an approximated covariance matrix for the estimators found so it becomes possible to test the hypotheses about the parameter values.

4.1.2 Parameter estimation - The Maximum Likelihood Estimator

Likelihood is a tool used for unknown parameters by summarizing the evidence of a dataset. Very often these unknown parameters of a distribution generally are denoted by θ. One of the most used technique for estimating parameters is the method of maximum likelihood, which estimates the values of the parameters that maximizes the likelihood (the joint probability function or joint density function) of an observed sample.

The basic principle of this estimation method is to establish a probability function, which is a common density function for the entire dataset, and to maximize it with respect to the parameters that are estimated. In other words, the maximum likelihood estimator, MLE, maximizes the likelihood for which the observed data will measure the estimated value of the unknown the best it can. The MLE is indispensable for nonlinear models, as in the present case of estimation of limited dependent variable models. The maximum likelihood estimation is nested on the distribution of y given $\mathbf{x}$, therefore the heteroskedasticity in $\text{Var}(y|\mathbf{x})$ is automatically accounted for.

It can be assumed, that in the general case of an i.i.d. dataset the likelihood function can be written as the common density function by multiplying the density functions of n number of observations as follows[3]

$$f(Y_1, Y_2, ..., Y_n) = f(Y_1|\boldsymbol{\theta}) \cdot f(Y_2|\boldsymbol{\theta}) \cdot ... \cdot f(Y_n|\boldsymbol{\theta})$$

where all the individual density functions depend on the vector of parameters, $\boldsymbol{\theta}$, and are the parameters, which will be estimated.

If the likelihood depends on p parameters, $\theta_1, \theta_2, ..., \theta_p$, those parameters are chosen to maximize the likelihood function

$$\mathcal{L}(\boldsymbol{\theta}; Y_1, Y_2, ..., Y_n) = f_{\mathbf{Y}}(\mathbf{Y}; \boldsymbol{\theta}) = \prod_{i=1}^{n} f(\mathbf{Y}_i; \boldsymbol{\theta}) = \prod_{i=1}^{n} \mathcal{L}(\boldsymbol{\theta}; \mathbf{Y}_i)$$

The Maximum Likelihood Estimator $\hat{\boldsymbol{\theta}}$ indicates the values of θ, which maximizes $\mathcal{L}(\boldsymbol{\theta}; \mathbf{Y})$ or equivalent maximizes the log-likelihoods function $\ell(\boldsymbol{\theta}; \mathbf{Y}) = \log \mathcal{L}(\boldsymbol{\theta}; \mathbf{Y})$. The function summarizes the information about $\boldsymbol{\theta}$ contained when $Y = y_i$. If the value of $\mathcal{L}(\boldsymbol{\theta}|y_i)$ is high for values of $\boldsymbol{\theta}$ this makes $Y = y_i$ more likely, and opposite are the values of $\mathcal{L}(\boldsymbol{\theta}|y_i)$ low, it makes $Y = y_i$ more unlikely.

In order to simplify the calculations, the logarithm is taken of the likelihood function. This has no significant effect on the output of the method, as well as the likelihood function and the log-likelihood function will have the identical value of the unknown parameters in $\boldsymbol{\theta}$. Also by taking the logarithm of the likelihood function, one makes sure, that the function always is positive which makes it possible to add the parts together instead of multiplying them. Taking the logarithm the function will be

$$\log \mathcal{L}(\boldsymbol{\theta}|Y_1, Y_2, ..., Y_n) = \log \prod_{i=1}^{n} f(\mathbf{Y}_i; \boldsymbol{\theta}) = \sum_{i=1}^{n} \log f(\mathbf{Y}_i|\boldsymbol{\theta})$$

thus obtaining an expression of the log-likelihood function in the general case, where the current density function is simply to be inserted. To apply the MLE method to a logit function it requires further comments.

In a distribution, the parameter of interest is $\boldsymbol{\beta}$, as n typically is fixed and/or known. If we let n be a random sample size and $f(y_i|\mathbf{x_i}, \boldsymbol{\beta})$ denote the density function for a random draw y_i from the dataset, conditional on $\mathbf{x_i} = \mathbf{x}$. The density of y_i given $\mathbf{x}_i$ is needed to obtain the maximum likelihood estimator,

---

[3] [56]

conditional on the explanatory variables and following is applicable. The maximum likelihood estimation of $\boldsymbol{\beta}$ maximizes the log-likelihood function

$$\max_{\beta}\sum_{i=1}^{n} \log f(y_i|\mathbf{x_i},\boldsymbol{\beta})$$

where $\boldsymbol{\beta}$ is a vector and the dummy argument in the maximization problem. In most cases, MLE of $\hat{\boldsymbol{\beta}}$, is consistent and has an approximate normal distribution in large samples. Since each y_i represents a binomial count in the i^{th} population, the joint probability density function for $\mathbf{Y}$ is

$$f(\mathbf{y};\boldsymbol{\beta}) = \prod_{i=1}^{n} \pi_i^{y_i}(1-\pi_i)^{1-y_i}, \qquad i = 1,2,...,n \tag{4.6}$$

so the probability of a success for any one of the n_i trails is π_i, the probability of y_i successes is $\prod \pi_i^{y_i}$ and likewise the probability of $1-y_i$ failures is $\prod(1-\pi_i)^{1-y_i}$. In equation (4.6) the joint probability expresses the values of $\mathbf{y}$ as a function of the known, fixed values for $\boldsymbol{\beta}$, which relates to π. Then the likelihood function can be set up, as it has the same form as the probability density function, except that the parameters of the function are reversed, i.e. the likelihood function expresses the values for $\boldsymbol{\beta}$ in terms of known, fixed values for $\mathbf{y}$ and $\mathbf{x}$.

The set up for maximum likelihood function for the distribution is to be estimated as follows

$$\mathcal{L}(\boldsymbol{\beta};\mathbf{y}) = \prod_{i=1}^{n} \pi_i^{y_i}(1-\pi_i)^{1-y_i} \tag{4.7}$$

Then taking the log to this expression ends up with

$$\ell(\boldsymbol{\beta};\mathbf{y}) = \sum_{i=1}^{N} y_i \log(\pi_i) + (1-y_i)\log(1-\pi_i) \tag{4.8}$$

And if one takes a look at the case of this thesis and the binary response variable, one can from equation (4.3) set up the conditional density is determined by two values

$$\pi_i = f(1;\mathbf{x},\boldsymbol{\beta}) = P(y_i = 1|1-\pi_i = \mathbf{x_i}) = G(\mathbf{x_i}\boldsymbol{\beta})$$
$$f(0;\mathbf{x},\boldsymbol{\beta}) = P(y_i = 0|\mathbf{x_i}) = 1 - G(\mathbf{x_i}\boldsymbol{\beta})$$

where the density can be written as $f(y;\mathbf{x},\boldsymbol{\beta}) = [1-G(\mathbf{x},\boldsymbol{\beta})]^{(1-y)} + [G(\mathbf{x},\boldsymbol{\beta})]^{y}$ for y equal to 0 and 1. Thus following equation can be found from taking the logs of the equation and one see consistency with equation (4.8)

$$\ell(\boldsymbol{\beta};\mathbf{y}) = \max_{\boldsymbol{\beta}} \sum_{i=1}^{n} \{ y_i \log \left[G(\mathbf{x_i}\boldsymbol{\beta}) \right] + (1-y_i) \log \left[1 - G(\mathbf{x_i}\boldsymbol{\beta}) \right] \} \tag{4.9}$$

As the log-likelihood function for observation i is a function of the parameters and the data $(\mathbf{x}_i, y_i)$ and is obtained by taking the log function of the following

$$\mathcal{L}(\mathbf{x};\boldsymbol{\beta}|y) = f_Y(y_i|\mathbf{x_i};\boldsymbol{\beta}) = [\underbrace{G(\mathbf{x_i}^T\boldsymbol{\beta})}_{y_i=1,\text{ "Success"}}]^{y_i} \cdot [\underbrace{1 - G(\mathbf{x_i}^T\boldsymbol{\beta})}_{y_i=0,\text{ "Failures"}}]^{1-y_i}$$

if $G(\cdot)$ is replaced by the model from equation (4.5)

$$f(y_i|\mathbf{x_i};\boldsymbol{\beta}) = \left[\frac{e^{\mathbf{x_i}^T\boldsymbol{\beta}}}{1+e^{\mathbf{x_i}^T\boldsymbol{\beta}}} \right]^{y_i} \cdot \left[1 - \frac{e^{\mathbf{x_i}^T\boldsymbol{\beta}}}{1+e^{\mathbf{x_i}^T\boldsymbol{\beta}}} \right]^{1-y_i} \tag{4.10}$$

$$= \left[\frac{e^{\mathbf{x_i}^T\boldsymbol{\beta}}}{1+e^{\mathbf{x_i}^T\boldsymbol{\beta}}} \right]^{y_i} \cdot \left[\frac{1}{1+e^{\mathbf{x_i}^T\boldsymbol{\beta}}} \right]^{1-y_i} = \frac{(e^{\mathbf{x_i}^T\boldsymbol{\beta}})^{y_i}}{1+e^{\mathbf{x_i}^T\boldsymbol{\beta}}} = \ell_i(\boldsymbol{\beta}) \tag{4.11}$$

where the vector x_i includes the intercepts. As $G(\cdot)$ is strictly between zero and one for the logit, $\ell_i(\boldsymbol{\beta})$ is well defined for all values of $\boldsymbol{\beta}$. Then the log-likelihood for the sample size n is obtained from equation (4.2), by summing equation (4.11) across all the observations:

$$\ell(\boldsymbol{\beta}) = \sum_{i=1}^{n} \ell_i(\boldsymbol{\beta}) = \sum_{i=1}^{n} \log \left[\frac{(e^{\mathbf{x_i}^T\boldsymbol{\beta}})^{y_i}}{1+e^{\mathbf{x_i}^T\boldsymbol{\beta}}} \right] = \sum_{i=1}^{n} y_i \mathbf{x_i}^T\boldsymbol{\beta} - \sum_{i=1}^{n} \log(1+e^{\mathbf{x_i}^T\boldsymbol{\beta}}) \tag{4.12}$$

The MLE of $\boldsymbol{\beta}$, denoted by $\widehat{\boldsymbol{\beta}}$ maximizes this log-likelihood. And as $G(\cdot)$ is the standard logit cumulative distribution function, the $\widehat{\boldsymbol{\beta}}$ is called the *logit estimator*. Differentiating with respect to $\boldsymbol{\beta}$, the MLE $\widehat{\boldsymbol{\beta}}_{MLE}$ solves And in specific case, i.e. be using the *logit model* from equation (4.5), and the fact that $G(\mathbf{x_i}^T\boldsymbol{\beta}) = e^{\mathbf{x_i}^T\boldsymbol{\beta}}/(1+e^{\mathbf{x_i}^T\boldsymbol{\beta}}) \Leftrightarrow 1/(1+e^{-\mathbf{x_i}^T\boldsymbol{\beta}})$ the logit MLE first-order condition of the logistic model is found from equation(4.12)

$$\frac{\partial \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{n} \left[y_i - \frac{e^{\mathbf{x}_i^T\boldsymbol{\beta}}}{1+e^{\mathbf{x}_i^T\boldsymbol{\beta}}} \right] x_i = 0 \tag{4.13}$$

the solution to of equation (4.13) is the maximum likelihood estimator $\hat{\beta}$ In general there are no explicit solutions for the $\widehat{\boldsymbol{\beta}}_{MLE}$. But then the Fisher's scoring method (also the Newton-Raphson method, see Calvin (1998) for discussion of differences) iterative procedure usually converges quickly, as logit models log-likelihood is globally concave.

From the estimate in equation (4.13) one can estimate the the probability that $y_i = 1$ for a given $\mathbf{x}_i$ as

$$\hat{\pi}_i = \frac{e^{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}}}{1+e^{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}}}.$$

and the first-order condition for the logit model found from equation (4.13) imply that

$$\sum_{i=1}^{n} \hat{\pi}_i \mathbf{x}_i = \sum_{i=1}^{n} y_i \mathbf{x}_i.$$

The predicted frequency is equal to the actual frequency, as if $\mathbf{x}_i$ contains a constant, then the estimated probabilities is equal to $\sum_{i=1}^{n} y_i$ or the number of observations in the dataset for which $y_i = 1$.

Classification table

It is always nice to know, whether an estimated model is a good or not that good, which is one of the reasons why classification tables are preferable. When one is classifying a set of binary data, a way of checking your model, could be to set up a training dataset and a test dataset. Meaning that if a dataset has 100 observations, one might use 90 observations to estimated model and the last 10 observations to check whether the model correctly classify those 10 observation or not. As the thesis is not focusing on machine learning techniques, the CTABLE function in SAS is used. The procedure provides a less expensive one-step approximation to the preceding parameter estimates, meaning that the procedure leaves one observation out, estimates the model and then predict the probability of of the left out observation.

4.2 Empirical Analysis and Results

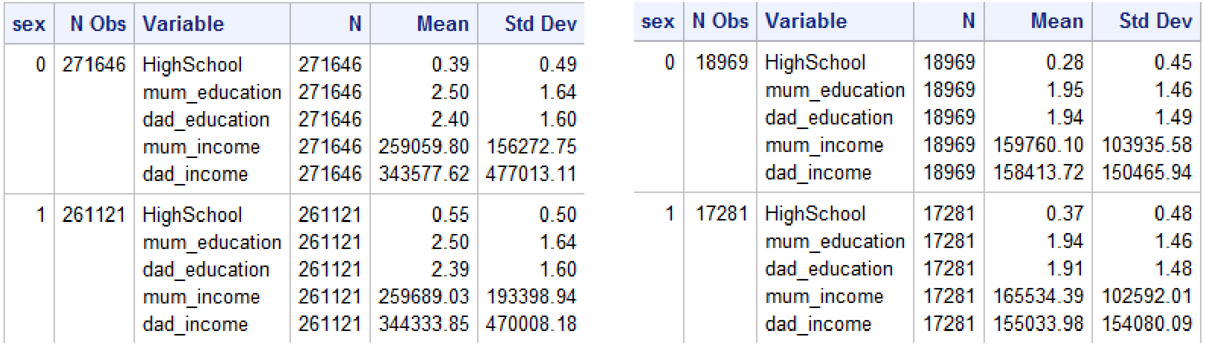
In this section, the models for boys and girls will be estimated. Furthermore the models will be estimated in different cases, which will be evolved through the section to conduct a in-depth study. The models will be estimated separately for immigrants and natives and then the analysis will be compared and commented on together with differences in the findings.

Firstly an overview of how many Danes in the datasets from the two different periods who, have graduated from high school and how many have not is given, including a look at parental average level of education and income. After this, the benchmark model with the explanatory variables parental level of education and income, will be estimated. In the following Robustness section different model estimations and analysis with the explanatory variables such as whether the young individuals mothers is living with the fathers.

Finally in the robustness section there will be a closer look on how models differ from municipalities and natives and immigrants and whether immigrants country of origin provides a different conclusion.

Parental average levels of income are found to be higher for natives than seen for immigrants, not surprisingly and further average education levels among native parents are found much higher than for immigrant parents. Some of low education average might due to the measurement insecurity, known in the field

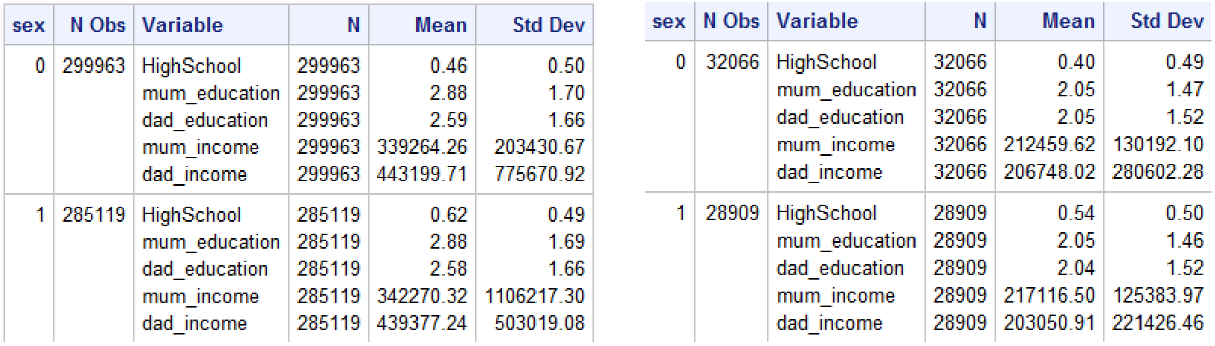
Figure 4.1: Left Table: Native origin, Right Table: Immigrant origin. Parental information 2006



sex	N Obs	Variable	N	Mean	Std Dev
0	271646	HighSchool	271646	0.39	0.49
		mum_education	271646	2.50	1.64
		dad_education	271646	2.40	1.60
		mum_income	271646	259059.80	156272.75
		dad_income	271646	343577.62	477013.11
1	261121	HighSchool	261121	0.55	0.50
		mum_education	261121	2.50	1.64
		dad_education	261121	2.39	1.60
		mum_income	261121	259689.03	193398.94
		dad_income	261121	344333.85	470008.18

sex	N Obs	Variable	N	Mean	Std Dev
0	18969	HighSchool	18969	0.28	0.45
		mum_education	18969	1.95	1.46
		dad_education	18969	1.94	1.49
		mum_income	18969	159760.10	103935.58
		dad_income	18969	158413.72	150465.94
1	17281	HighSchool	17281	0.37	0.48
		mum_education	17281	1.94	1.46
		dad_education	17281	1.91	1.48
		mum_income	17281	165534.39	102592.01
		dad_income	17281	155033.98	154080.09

Figure 4.2: Left Table: Native origin, Right Table: Immigrant origin. Parental information 2016



sex	N Obs	Variable	N	Mean	Std Dev
0	299963	HighSchool	299963	0.46	0.50
		mum_education	299963	2.88	1.70
		dad_education	299963	2.59	1.66
		mum_income	299963	339264.26	203430.67
		dad_income	299963	443199.71	775670.92
1	285119	HighSchool	285119	0.62	0.49
		mum_education	285119	2.88	1.69
		dad_education	285119	2.58	1.66
		mum_income	285119	342270.32	1106217.30
		dad_income	285119	439377.24	503019.08

sex	N Obs	Variable	N	Mean	Std Dev
0	32066	HighSchool	32066	0.40	0.49
		mum_education	32066	2.05	1.47
		dad_education	32066	2.05	1.52
		mum_income	32066	212459.62	130192.10
		dad_income	32066	206748.02	280602.28
1	28909	HighSchool	28909	0.54	0.50
		mum_education	28909	2.05	1.46
		dad_education	28909	2.04	1.52
		mum_income	28909	217116.50	125383.97
		dad_income	28909	203050.91	221426.46

The average education among native parents and income have increased relative more over time, compared to the development in parental income and education among immigrants. For native parents, fatherly income in average is highest, while motherly income in average is higher for immigrants. Among native parents, average motherly level of education is higher than for fathers (and the average for motherly education has increased more over time), while the average level of education for immigrant parents is approximate the same in both years.

Fraction of high school graduated young Danes

2006

To provide an overview of the proportion of young Danes who have completed a high school education, tables from the two datasets are set up. Firstly the dataset from 2006 is investigated. Data from 2006 contains all individuals born between 1976 – 1986 and who have either a mother or father registered in Denmark with non-negative income. The following table shows how many individuals born in this period, who have a high school education. The table also provides distinguishes on the differences there might be in natives and immigrants:

Table 4.1: Data from 2006

	Immigrant origin		Native origin		Total
High School	Individuals	Cumulative Freq.	Individuals	Cumulative Freq.	Percentage
No (0)	24,610	67.89%	281,986	52.93%	53.88%
Yes (1)	11,640	32.11%	250,781	47.07%	46.11%
Total	36,250		532,767		569,017

The table shows, that the larger fraction of individuals born between $1976 - 1986$ has not graduated from high school, regardless of their origin and gender. Beside that, the table shows that only 32.11% of the individuals with an immigrant origin of either 1^{st} or 2^{nd} generation born between $1976 - 1986$ had a high school education in 2006 and furthermore it shows that the fraction of immigrant individuals with a high school degree is relatively low compared to the fraction of native with an high school degree, which in 2006 was approximately 47%. Overall the table implies that clearly the greater fraction of individuals with immigrant origin did not graduate form high school in the mid-00's, and almost half of individuals with native origin did. These findings corresponds to the fact, that natives preform better in primary school and therefore probably are more likely to continue in high school.

As this thesis is interested in the differences there might be between gender, let's see how the fractions differ between genders. Besides that let's see if there might be any differences between gender and country of origin. The individuals born between $1976 - 86$ with an immigrant origin listed by having an high school education or not and by gender. From the 2006 data is given as

Table 4.2: Immigrant origin. Data 2006

Gender	Graduated High School	Individuals	Percentage	Gender Percentage
Boys	No (0)	13,725	37.86%	72.35%
	Yes (1)	5,244	14.47%	27.65%
Girls	No (0)	10,885	30.03%	62.99%
	Yes (1)	6,396	17.64%	37.01%

Table (4.2), is based only on individuals with an immigrant origin (either 1^{st} or 2^{nd} generation) shows that 37% of the girls have graduated from high school, while it is only 27.7% of the boys. Many people have studied this, and tried to explain and figure out why and why this is the case[4] - I will get back to that later in the thesis. Given the fraction of boys with a high school education there might be some difficulties in the section with model estimation, especially when working on a significantly level.

An identical table, again by gender, based on individuals with a native origin from the data in 2006 follows here:

---

[4] https://www.b.dk/nationalt/fire-ud-af-ti-unge-indvandrere-har-hverken-job-eller-uddannelse

Table 4.3: Native origin. Data 2006

Gender	Graduated High School	Individuals	Percent	Gender Percent
Boys	No (0)	164,516	30.88%	60.56%
	Yes (1)	107,130	20.11%	39.44%
Girls	No (0)	117,470	22.05%	44.99%
	Yes (1)	143,651	26.96%	55.01%

Table (4.3) shows that the greater fraction of the girls born between 1976 – 1986 have graduated from high school and that almost 40% of the boys with native origin have. In both scenarios, the fraction of natives with a high school education is larger than the fraction of immigrants by same gender, moreover one can notice that the fraction of native boys graduated (39.44%) is greater than the fraction of immigrated girls graduated (37.01%).

2016

This subsection provides the overview of the data from 2016, which contains all individuals born between 1986 – 1996, with either a mother or father registered in Denmark with a non-negative income.

Table 4.4: Data from 2016

	Immigrant origin		Native origin		Total
High School	Individuals	Cumulative Freq.	Individuals	Cumulative Freq.	Percentage
No (0)	32,693	53.62%	271,343	46.38%	47.06%
Yes (1)	28,282	46.38%	313,747	52.62%	52.94%
Total	60,975		585,090		646,065

Table (4.4) shows not only that the amount of individuals has increased, but also that in total a greater fraction of individuals, within the aging group the thesis is working with, has graduated high school in Denmark.

If one is familiar with the proportion of Danes who have graduated from high school the resent years, like presented in ”Danskernes Uddannelse” by Mie D. Pihl (2017), then the table above might seem wrong, or at least too low in the proportion of graduated young Danes. Mie D. Pihl presents that in 2015 more than 60% of young Danes had a high school education 10 years after graduating from primary school. The reason why the fraction of young Danes with a high school degree in the data from 2016 is lower, is that this group consists such a large gab of age, and that the oldest in the group, the smaller fraction has an education. If one looks closer at the younger generations in the dataset, a higher fraction graduated from high school is found.

Furthermore table (4.4) shows that compared to the data from 2006 the proportion of immigrants with a high school degree has increased from 32.1% to 46.4% in 2016, which is an increase of 14% points within the ten years. This can be compared to the increase in the fraction of native with a high school education, which only increases almost 6% points from 47% in 2006 to 52.6% in 2016. The proportion of individuals with a high school education has therefore increased relatively more for individuals with immigrant origin, than for individuals with native origin.

Now a closer look at the gender differences there might be. Firstly a table with individuals with immigrant origin are set up

Table 4.5: Immigrant origin. Data 2016

Gender	Graduated High School	Individuals	Percentage	Gender Percentage
Boys	No (0)	19,275	31.61%	60.11%
	Yes (1)	12,791	22.01 %	39.89 %
Girls	No (0)	13,418	20.98 %	46.41 %
	Yes (1)	15,491	25.4 %	53.59 %

Table (4.5) shows that more than half of the girls with immigrant background born between $1986-96$, with a mother or father with non-negative income has graduated from high school. The fraction of girls with immigrant origin with a high school education has increased by almost 17% points from 37% in 2016 to 53.6% in 2016. The fraction of boys, born in the same period and under the same parental assumptions, with a high school education has increased a little less than the girls, as it has increased by 12% points from 27.7% in 2006 to 39.9% 2016.

Information on natives born between $1986-96$ with a high school education data from 2016, provides the following table by gender

Table 4.6: Native origin. Data 2016

Gender	Graduated High School	Individuals	Percentage	Gender Percentage
Boys	No (0)	161,619	27.62%	53.88%
	Yes (1)	109,724	18.75%	46.12%
Girls	No (0)	138,346	23.65%	38.48%
	Yes (1)	175,401	29.98%	61.52%

Table (4.6) shows that more than 6 out of 10 girls with native origin, born between $1986-96$, have graduated from high school and that almost half of the boys (46%) have. In both scenarios, the fraction of natives with a high school education is larger than the fraction of immigrants by same gender. Compared to the results found for individuals with native origin based on data from 2006, one finds that the increased fraction is not as large as the increasing found within the immigrant origin-group. Furthermore the fraction of native boys graduated in 2016 is now lower than the fraction of girls with immigrant origin having an education.

So this short review of the two datasets provides the reader with the knowledge that in 2006, less than half of the individuals born between $1976 - 1986$ within the requirements, where at least one parent is registered in Denmark and has non-negative income including social services. Furthermore distinguish between origin is found that more than 60% of girls and 70% of boys with immigrant origin, did not have a high school education at the age of 30, which for natives was less than 50% and more than 60% for respectively girls and boys.
Investigating the differences between fraction of girls and boys with a high school education, clearly differences appear and are found from

$$\text{Diff}\ _i^{\text{Year}} = \text{Fraction Girls} - \text{Fraction Boys}$$

where $i = I, N$ respectively Immigrant/Native Background.

$$\begin{aligned}\text{Diff}\ _I^{2006} &= 37.01\% - 27.65\% = 9.36\% \text{ points}\\ \text{Diff}\ _I^{2016} &= 53.59\% - 39.89\% = 13.7\% \text{ points}\\ \text{Diff}\ _N^{2006} &= 55.01\% - 39.44\% = 15.57\% \text{ points}\\ \text{Diff}\ _N^{2016} &= 61.52\% - 46.12\% = 15.4\% \text{ points}\end{aligned}$$

Which imply that the proportion high school educated boys with native origin has increased relative more from 2006 to 2016, compared to the native girls. While from 2006 to 2016 the proportion of educated girls with immigrant origin have increased relative much, compared to the boys with immigrant origin. The findings that immigrants girls in higher frequency educate themselves than immigrant boys fit my prior expectations.

Correlations

A brief review of correlation coefficient, based on Wackerly et al. (2007), before estimating the models of interest. The most used way of looking at the "relationship" between to variables, is the bivariate Pearson Correlation, which provides a sample correlation coefficient, ρ, which measures the strength of a linear relationship between pairs of continuous variables. The Pearson Correlation in a two dimensional distribution of (X,Y) is a parametric measure measured by

$$\rho_{\text{XY}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{\text{cov}(X,Y)}{\sqrt{\text{var}(X)} \cdot \sqrt{\text{var}(Y)}}$$

and will always take a value between -1 and 1. A positive correlation implies that there is a positive connection between two variables (X, Y), i.e. if the correlation between those two variables is equal to one, that tells the reader that if X increases by one, then Y also increases by one. The correlations are found easily in SAS and for the "basic" explanatory variables the following table of correlations are found as

Table 4.7: Correlations between response variable and the explanatory variables

	Origin, Year	Mother_edu	Father_edu	Mother_in	Father_in
High School	Immigrant, 2006	0.1629	0.1299	0.1367	0.1041
High School	Native, 2006	0.3130	0.3141	0.2065	0.2031
High School	Immigrant, 2016	0.16	0.1391	0.1266	0.1386
High School	Native, 2016	0.3039	0.29355	0.24	0.2283

As expected, and corresponding to Erola and Lehti (2016), the correlation between high school education among individuals and their parents level of income and education is found to be positive. Implies, that parental higher education/income the greater is the probability that their child will get a high school education.

4.2.1 Estimation of models in Benchmark

The logistic regression is set up in SAS by *proc logistic* and is based on the four datasets, which are described previously in the thesis. The binary response variable, which is created from different education variables based on Denmark Statistic data, describes whether a given individual has graduated from high school or not, and the explanatory variables are parents' level of education and income, which is created in dummy variables. The model in SAS is set up as

$$y_i = \beta_0 + \boldsymbol{\beta}_k \mathbf{x}_{\text{i,mum-eduk}} + \boldsymbol{\beta}_k \mathbf{x}_{\text{i,dad-eduk}} + \boldsymbol{\beta}_l \mathbf{x}_{\text{i,mum-incgrpl}} + \boldsymbol{\beta}_l \mathbf{x}_{\text{i,dad-incgrpl}}$$

where $k \in [1:5]$ and $l \in [1:4]$. The reason why, k and l is 1 level lower than the amount of classes in each explanatory variable is, that one group of each explanatory variable is left out to avoid multicollinearity. In all the analyses the lowest group of income and the lowest level of education are left out. Their effect will appear in the intercept.

2006[5]

How were the parental effects in 2006? Firstly the effects from parental level of income and educations are found for the native origin. The effects are found by estimating a model for girls and a model for boys. The model estimated for boys, are modeled based on $271,646$ individuals, and the model estimated for girls are modeled based $261,121$ individuals. In both cases, individuals with only one registered parent in Denmark are included. Both models are found to be significant, as the p-value of the Wald Chi-Square

---

[5] For documentation of models and plots see appendix A

Test (and the other two tests as SAS prints) is found to be 0.00001. The tests test that at least one of the predictors' regression coefficient is not equal to zero in the model. The binary logit regression estimates for the parameters $\hat{\beta}$, found in the Maximum Likelihood estimates as the "Estimates", documented in appendix A. In all models estimated in the thesis, the intercept includes the lowest level of parental income and education. Not very surprisingly it is found, that all, but one, of the estimated coefficients are significant. This is not a surprise, as the dataset contains a lot of individuals. Furthermore it is not surprisingly the overall probability of an individual getting an education increases with a higher level of parental education for girls and boys. But for boys, it is surprising to find that if level of income changes from 1 to 2 among for both of their parents, this will reduce the probability of them getting an education. This implies, that if parental levels are at the absolute lowest level, the probability of education is higher. As one can not interpret the effects directly given the log transformation in the equation and the relationship, the estimates are hard to interpret. The following equation could be set up, given the boys' parents have a level education and income group equal to 2.

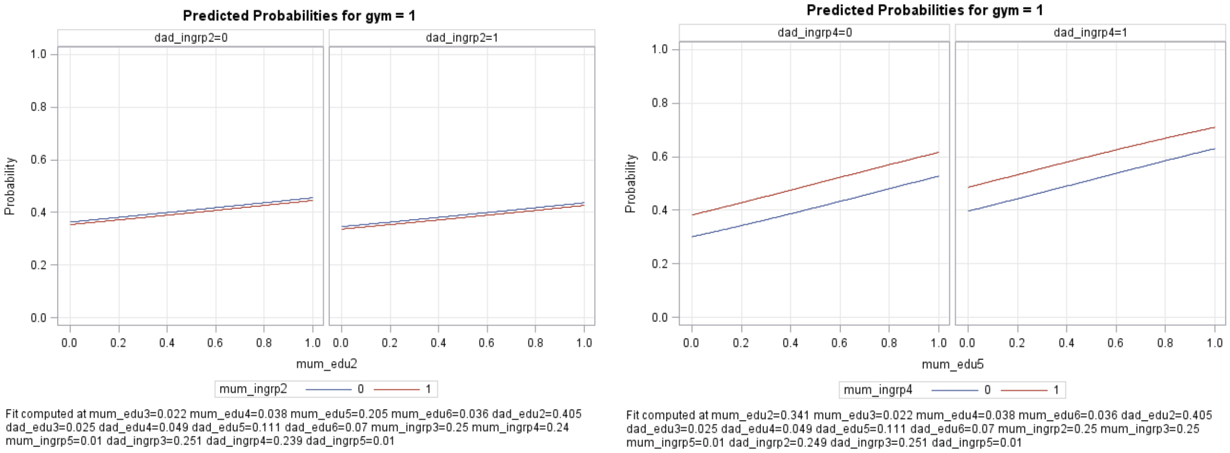
$$\log \frac{\pi}{1-\pi} = -1.5211 + 0.3819 x_{\text{mum\_edu2}} + 0.2584 x_{\text{dad\_edu2}} - 0.0417 x_{\text{mum\_inc2}} - 0.0772 x_{\text{dad\_inc2}}$$

One could use the interpretation that for a one unit change in a explanatory variable, the difference in log-odds for a positive outcome is expected to change by the respective β coefficient, given the other variables in the model are kept constant, i.e. the difference in log-odds is expected to be 0.3819 units higher for a level of motherly education equal to 2, ceteris paribus.

The odds ratios could also be found, but neither is directly in the interpretation. Which is why the stereotype families are chosen as the best way to interpret the parent effects. Another effect worth commenting on, is the effects of the models on discriminating between individuals having graduated from high school and who have not. The effect is summarized in the "c", the Concordance Statistic, which is the area under the ROC-curve. The value of "c" is found to be 73.51% and 72.1% for respectively boys and girls, this implies at what rate the model correctly predicts an observation, i.e. the rate, when the model correctly predict whether an individual has a high school education or not, based on parental level of education and level of income. One can control if the model predicts wrongly in one direction, meaning that the model is very good at predicting individuals with a high school education, but it preforms poorly when classifying individuals without a high school education. This is done by setting up a classification table, here set with the prior probability of 50%.

The *Sensitivity* is for boys found to 44.4%, provide the information ability of the models is to predict an individual with an high school education correctly, i.e. the proportion of $y_i = 1$ responses that were predicted to be $y_i = 1$ (remember $y_i = 1$ is high school graduated), and *Specificity* is the proportion of $y_i = 0$ responses that were predicted to be $y_i = 0$, which for boys is relatively high, 87.0% (but also what is most of). Furthermore the rate *False POS* 31%, which is the proportion of predicted $y_i = 1$ responses that were observed as $y_i = 0$. The opposite is the case for *False NEG* 29.4% where the proportion of predicted $y_i = 0$ responses actually is observed as $y_i = 1$. The model estimated for boys is not skew in its mis-classification, whereas the model estimated for girls is, as it falsely mis-classify a higher rate af individuals as, which are classified as not having an education, while they actually do have one. This might imply that some girls are more social mobile and therefore the model can not predict them correctly. The effect plots shows the differences, which appears when one looks at the probability of boys with native parents gets an education. It is found in two scenarios for a boy: when he comes from a low resource family and when he is from a family with greater resources.

Figure 4.3: Predicted probability plots for boys with native origin, data form 2006



The two plots to the left shows the probability of a boy getting an education, given his parents have low resources, while the two plots to the right shows the probability of education if he comes from a family with higher resources.

The set up in both plots: The blue line is the prior probability given the motherly level of income unknown. The red line, is when it is known. In the plot to the left (of the two associated), the fatherly level of income is not known, while it is in the plot to the right. To the left side in each plot the motherly level of education is unknown, while it is known to the right in each plot.

In the case where the boy has parents with fewer resources, one sees that the prior probability is always higher when parental level of income and education is known. Furthermore, one sees that change from the prior probability is very small, i.e. the increase in the probability of a boy getting an education is quite small whether he comes from a family where motherly and fatherly level of income and education is the lowest (the intercept) or if he comes form a family with a higher degree level of education and income.

If one instead looks at the probability of education for the boy, which comes from a family with more recourses, one see that the probability of getting education increases a lot, and that it is almost 15% higher than when his parents has no education. The same development is found for native girls, just at higher levels, which does not surprise as the proportion of girls getting an education is higher. For the girls with native origin a larger gab between the prior probability and motherly income and education is found at a higher level, and thereby the probability of education increases with more than what was the case for boys. Furthermore one higher level in fatherly income, does not change the probability much, which implies that motherly income and level of education has greater effect on the probability of girls getting an education or not. This fits the findings in Esping-Andersens (2004) and Bukodi and Goldthorpe (2012), who found that motherly effect had higher influence among girls.

Looking at a girls' probability of education if she comes from a high resource family, one sees that the gab between the prior probability of motherly level of income and also a high level of education provide a large difference as the slopes are very steep. Parental effect seems to have a great effect on probability of girls getting an education.

Testing whether parental effects differ between girls and boys with native origin based on the data from 2006, clear differences are found.[6] The null hypotheses is, that the parental effects at a specific level of education or income is equal for boys and girls, one can reject it if $p < 0.05$. Unless mothers have a bachelor degree, the motherly educational effect on the probability of education is alike between native girls and boys. For fathers' level of education almost every level differs significantly among girls and boys. (Except if he has short (2 years) university education.) But if parents belong to the riches percentage, the effects are equal for boys and girls, while it for all other levels of parental income provide us with a clear difference between boys and girls. More specific one can tell that if the maximum likelihood estimates of the tests are positive (as du3, dad education level 3, high school) then, that effect has a greater positive effect on girls, than on boys. And the opposite the fact that if the tests maximum likelihood estimator is negative, then the parental effect has a greater effect on boys. Therefore one can conclude, that almost every of the significant effects is of greater impact on the probability for boys, than for the girls.

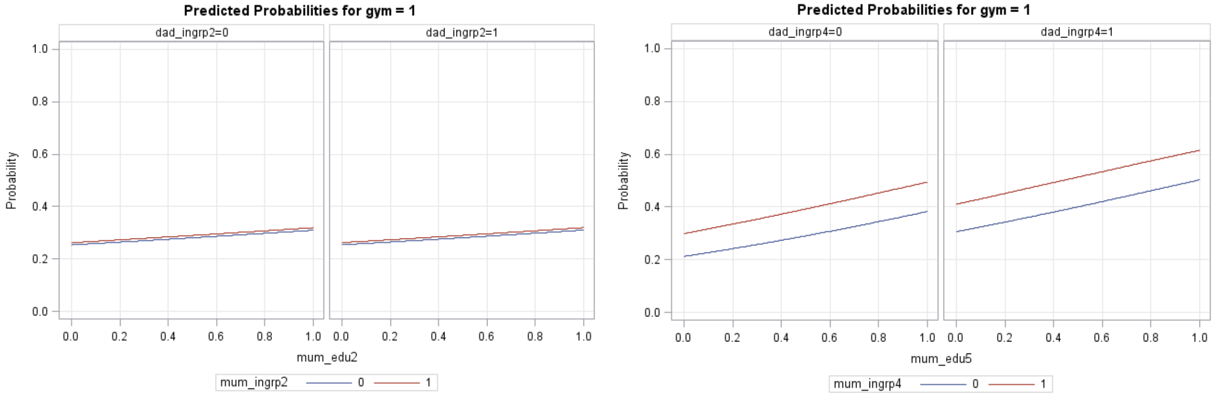
To investigate the development among the individuals with immigrant origin, models based on the data from 2006 are set up and the effect of parental level of income and educations on the probability of getting high school education is found. The model estimated for boys is modeled from $18,969$ individuals, while the model estimated for girls is based on $17,281$ individuals. Again it is found that both models are significant. The binary logit regression estimation for the parameters $\hat{\beta}$, is found and one sees that an increase in level of parental education increases the probability of education among the individuals, at a significant level of 5%. The low average of mothers with immigrant origin have a higher education, but the model seems not to have any uncertainty of the measurement at the area. The low amount of parents with higher education among immigrant origins parents, might due to the insecurity of measuring of educations among immigrants. Further the insecurity increases, when their home country becomes less developed.

---

[6] To test this, the dummy "sex" (indicating gender) was created as a zero-one variable. Zero for boys, and one for girls. After creating the dummi variable, before testing the hypotheses of equal effects, the maximum likelihoods estimators found by running the code including interactions are compared to the once found previously as they match. As they are found to match, it is known that the dummi variable and the interactions are setup correctly.

The models classify respectively 64.1% and 64% of the boys and girls correctly. Looking at the classification tables, it is found that the model for boys with immigrant origin estimate poor 10.6% with a high school education correct, while it predicts 97.2% without an education correctly. Remember that the greater part of boys with immigrant origin in the data from 2006 have not graduated from high school. It is also found, that the model more often classify a greater fraction as having an education $y_i = 1$, when they actually do not have an education, than it does the other way around (False POS >> False NEF). The probability slopes for high school education given a motherly level of education at the lowest, seem to be approximately alike, but with a prior probability at a point lower between the boys. The two plots to the left shows the predicted probability of boys getting an education, when their parents are of lower groups, i.e. low resource family. Motherly level of education and group of income are set to 2, and likewise for the fatherly income group. The blue line indicates the prior probability, before knowing the level of motherly income, the red line indicates when motherly level of income is known (here 2). The left plot is the prior probability before knowing fatherly level of income, while the right plot is when fatherly level of income is known as group 2. In both plots the left side in the plot indicate the prior probability, before knowing motherly level of education, while the right side of each plot indicates the probability after motherly level of education is known (here level 2). One see, that the prior probability of boys gets a high school education is lower, than after knowing the level of parental education and income, which is relatively low.

Figure 4.4: Predicted probability plots for boys with immigrant origin, data form 2006



Fit computed at mum_edu3=0.071 mum_edu4=0.035 mum_edu5=0.081 mum_edu6=0.037 dad_edu2=0.194 dad_edu3=0.053 dad_edu4=0.042 dad_edu5=0.065 dad_edu6=0.05 mum_ingrp3=0.245 mum_ingrp4=0.235 mum_ingrp5=0.009 dad_ingrp3=0.251 dad_ingrp4=0.243 dad_ingrp5=0.01

Fit computed at mum_edu2=0.19 mum_edu3=0.071 mum_edu4=0.035 mum_edu6=0.037 dad_edu2=0.194 dad_edu3=0.053 dad_edu4=0.042 dad_edu5=0.065 dad_edu6=0.05 mum_ingrp2=0.247 mum_ingrp3=0.245 mum_ingrp5=0.009 dad_ingrp2=0.254 dad_ingrp3=0.251 dad_ingrp5=0.01

The two plots to the right, the blue line again is prior probability before knowing motherly level of income group, the plot to the left is before knowing fatherly level of income and the left side of the plots are prior probability before knowing motherly level of education. One see a clear dependents on al the parental factors, as the gab between the lines are relatively big, the slopes are high and the difference in the to plots are clear.

Comparing the two stereotypes of families, it is found, that given the prior probability of education, it is in the left case, that families with lower resources makes the probability of education for a boy with immigrant origin to around 25%, furthermore the change in parental income does not seem to have any effect, the two lines are extremely close. While in the other case, the predicted probability of education is above 60% Here have in mind, that the fraction of boys with immigrant origin getting a high school education in 2006 was low. If one looks at the same two stereotype of families, but instead look at the predicted probabilities of girls getting an education, some of the patterns are they same the stereotype family with lower resources makes the probability of education among girls lower, while the family with higher resources makes the probability of education increase, as before. But for girls, a motherly income increase, higher the probability of education more than we saw among the boys. Further we see that the predicted probability of a girl getting an education, if se comes from a family with lower resources is around 40% and if she is from. Which is clearly higher than what we saw was the case with the boys.[7] Testing whether the parental effects are more or less alike between boys and girls with immigrant origin, it is found that only one parental effect provides different effect on the probability of girls vs. boys getting an education. The only effect which significantly differ between girls and boys is the highest level of education among fathers. This is the only test (testdu6) where the null hypotheses is rejected, as $p < 0.05$.

The four models based respectively native origin and immigrant origin are different in the results, but the overall effects are quite much alike, as it is found that higher level of education and income among parents implies higher probability of education. The differences in the parental affects will be further commented, after the models based on 2016 data are presented.

2016

The models based on the data from 2016 are built on $285,119$ girls and $299,963$ boys with native origin. From the SAS output it is found that the Model Fit Statistics based on the native sample, in the test of $\beta = 0$ one find that Likelihood Ratio Test and the Wald test both with a p-value < 0.0001 i.e. highly significant, which tells that the model as a whole fits significantly better than an "empty" model, i.e. the model explains a significant portion of variance in the data. The Maximum Likelihood Estimators are found and compared to what we saw in 2006, there is no negative effects on the probability of education for parental level of income 2. Therefore one can say, that higher level of education or higher level of income among parents implies higher probability for high school education among girls and boys. A closer look at the table summarizes the models effect on discriminating between individuals having/not having graduated from high school. The rate is found to 73.3% and 72.5% (for respectively boys and girls), meaning that the model is correct that rate of the time. This implies that when the models are predicting whether an individual are getting a high school education based on parental level of education and level of income, the model for the boys are doing a tiny bit better. The classification tables, tells

---

[7] See the plots in appendix A

that the model estimated for girls is a little skew, i.e. it predict more girls as not having an education, when they actually (than the other way around). This implies, that more girls are "breaking" out of the systematic the model has found, i.e. they might be breaking the social heritage and therefore are more social mobile to higher groups, than girls based on parental factors are predicted to get an education, but to not. Given this mis-classification, it might indicate that the are more social mobile, or at least that the parental effects are less similar among girls, than the boys.
The predicted probability for boys shows clear positive dependents of parental level of income an education. One see that if motherly level of education is 3 years at the university the probability of a boy getting a high school education, increases with more than 20%, while the much higher motherly level of income, here level 4, increases the probability with approximate 10%.

A closer look toward the girls with native origin based on data from 2016 one sees that the case are much alike the case with the boys. But found at higher levels and with greater dependents on motherly level of education and income. Testing whether the differences are significantly or not, provides the information that given a null hypotheses, that the effects are equal, it is found that most of the parental effects clearly differ between boys and girls within the sample of native origin. Among motherly effects that does not differ at a significant level between boys and girls, if her level of education is high school or master/ph.d. as highest completed education, or if she is among the one percentage with the highest income. Fatherly effects only differ between boys and girls, if he is in the two highest income groups. These results differ from 2006.

Looking at the datasets with immigrant origin, the models are built on respectively $32,066$ and $28,909$ for boys and girls. Again it is found that both models are significant, i.e. the βs do differ from zero, as their Wald Test Score as p-value clearly are less then 5%. The in the Maximum Likelihood Estimators shoes almost every parental effect is significant and compared the model estimation based on the 2006 dataset, more explanatory have become significant. This might be caused by more observations or the fact that influence of parental effects have become more similar among immigrants in 2016, than they were in 2006. All significant effects also provides us with the information, that higher level of parental education or income implies higher probability of high school education among the individuals. No surprises here. The models still classify more than half of the data correctly classify 64.6% for the boys and 63.9% for the girls with immigrant origin. Both the classification tables shows the models in 2016 are mis-classifying at a higher rate than in 2006. The model estimated for boys with immigrant origin are mis-classifying the most as not having an high school education, when they actually has one, while the model estimated for girls are mis-classifying a greater fraction as having a high school education, when they do not have one. So some how, the models estimated on immigrants origin has become worse in it predictions on whether girls or boys are getting a high school education. This implies that the patterns parental effects are more unclear. This might due to the higher rate of individuals with immigrant origin getting an education. The effect plots shows, that the overall probability of boys getting a high school education has increased a lot (approximately 40%), compared to what's seen based on data from 2006. Furthermore one sees that

an increase, even though it is just one level, in motherly income or education, the boys are more likely in getting an education, as the slopes in the to plots to the left have a larger gab between the lines and have a steeper slope. While the effect of a one level increase in fatherly income, does not seem to change much, in the case with a family of fewer resources, while if fatherly income level is known as high, it has a larger effect.
The prior probability of education among the girls are higher no matter what type of family she comes from, compared to 2006 and to the boys. The change in an increase in motherly income in a low resource family has not chanced as much, as one saw was the case for the boys. Further, if motherly level of education is at level 4 has relatively less effect, as the slopes are much flatter than what we saw in the model based on data from 2006, likewise with a motherly high income group, as the gab between the lines has decreased. Testing whether the parental effects differ between boys and girls. The results imply that significantly differences appear between boys and girls in few levels of fatherly education.

So based on data from 2006 and 2016 the models are estimated and it was found that the models estimated from data in 2006 are doing a better job predicting the probability of whether an individual are getting an education or not. This implies, that the individuals in the dataset in 2006 are more alike and the parental influence has a clear pattern, which indirectly implies that the similarities in parental effects are harder to find based on the individuals in 2016. With statistical glasses this is bad, but with a societal perspective this implies the whether a young Dane are getting an education or not depends a little less on parental effects, which could imply more social mobility.

4.2.2 The models used for stereotype families to predict probabilities of high school education - how do parental influence differ?

In this section a review of the differences within the model predictions is presented. Here the differences in parents' effect are shown and commented on, the developments within the parental effects among the origins are presented, as well as the differences between the origins. Based on the estimated models, the probability of education are presented, given the stereotype families presented in section 3.1.1. This implies the section will present the probability of education for 24 fictive individuals (12 girls and 12 boys). As all the explanatory variables are categorical it makes sense to evaluate the differences of parental effects on the probability of high school, given some stereotype families.

Girls

By the use of the four estimated models, the predicted probabilities for the girls from the stereotype families are presented in the following table. The first column shows the parental level of income and education, followed by second column which is the stereotype families, the third column tells the predicted probability of education for a girl with native origin (NA) based on the 2006 data, fourth column is the predicted probability of a girl with native origin based on the 2016 data. The two following columns represent the predicted probability of a girl with immigrant origin (IM) based on respectively data from 2006 and 2016. Finally the last two column represent the development from 2006 to 2016 in the probability of education for a girl within her origin.

Table 4.8: Predicted probabilities of high school education for girls based on stereotype families

Parental effects	"Stereotype Fam."	NA 06	NA16	IM06	IM16	ΔNA	ΔIM
Levels: 1	1 (prior prob.)	27.78%	27.35%	22.56%	36.11%	-1.52%	60.07%
Levels: 2	2	46.51%	54.05%	35.75%	55.33%	16.22%	54.76%
Levels: 3	3	79.57%	82.68%	44.39%	61.16%	3.90%	37.77%
Levels: 4	4	82.48%	86.74%	65.78%	77.86%	5.16%	18.36%
Levels: 5	5	92.31%	93.67%	88.55%	86.76%	1.48%	-2.02%
Levels: 5/6	6	96.53%	95.81%	85.01%	89.11%	-0.75%	4.82%
Dad Edu: 3	7	48.27%	46.91%	28.67%	41.44%	-2.82%	44.54%
Mum Edu: 3	8	50.30%	47.89%	30.99%	47.03%	-4.78%	51.76%
Mum Edu: 6	9	63.28%	57.11%	35.51%	56.86%	-9.75%	60.12%
Dad Edu: 6	10	61.37%	52.46%	30.50%	47.86%	-14.52%	56.92%
Dad Inc: 4	11	38.01%	40.53%	31.56%	50.39%	6.63%	59.66%
Mum Inc: 4	12	38.39%	42.42%	33.64%	48.08%	10.52%	42.93%

One see, that given girls comes from families with the lowest possible resources, the probability of girls with immigrant origin getting an education are higher, than if girls are of native origin in 2016. This could implies that immigrant girls are more mobile or it could also imply that the parental pattern for the probability of education is less alike.

The development in the predicted probability of education for girls with native origin shows, that girl "1", i.e. girls from families with very low resources, have reduced probability of education by 1.52% points from 2006 to 2016, while girls from stereotype family 2 have an increased probability of education with 16.22%. This indicates that girls from absolute lowest resource family are more likely to not get an education, while if her parents just have some education and a little higher income, her probabilities of education increases a lot.

Given the "last seven families", where a change in one parental level of income or eduction, it is found that the effect of increase in parental level of education has a lower influence on the probability of education, while the influence of parental income on probability of education is higher among native origin girls, when 2006 is compared to 2016. For girls with immigrant origin, the effects of an increase in the parents level education raises the probability for education a lot. This indicates, that girls of native origin are less influenced by parental level of education (but it is still the most powerful parental effect) and relative more effected by parental income, today than in 2006. For girls with immigrant origin, it seems like parental effect have increased a lot over time, but the relative change in the probability of education is more os less the same in 2006 and 2016.

Table 4.9: Relative changes in the predicted probability of education among girls when one parental level changes, and all other parental levels are set to 1

"Stereotype Family"	Changing effect	Δ NA 06	Δ NA16	ΔIM 06	ΔIM16
7	Dad Edu: 3	73.76%	71.47%	27.10%	14.77%
8	Mum Edu: 3	81.07%	75.07%	37.38%	30.26%
9	Mum Edu: 6	127.83%	108.78%	57.42%	57.48%
10	Dad Edu: 6	120.94%	91.77%	35.21%	32.55%
11	Dad Inc: 4	36.83%	48.15%	39.91%	39.56%
12	Mum Inc: 4	38.19%	55.09%	49.13%	33.16%

Reading this table, remember to have in mind, that the prior predicted probability of a native girl getting an education, is lower than the the predicted probability of a girl with immigrant native in 2016. One sees that an increase in the motherly level of education has a relatively greater effect on the probability of education for girls, regardless of country origin, than an increase in fatherly level of education.

The level of education among parents seem to be more important than their level of income in the case for native. These results corresponds to the findings in Bukodi and Goldthorpe (2012), who finds parental level of education are more import, than parental class or parental status and Erola et al. (2016) who found parental education more important than parents income when it comes to children's educational attainment. Furthermore it is found for girls with immigrant origin that fatherly level of income has a greater positive effect than level of his education, but also greater effect than motherly income. This is unlike the findings in Esping-Andersen (2004) and Lillard and Willis (1994) who both have found that mothers matter more for girls. Whether the high income among fathers with immigrant origin is correlated with a higher interface with the Danish society, and therefore knowledge about the importance of education is the reason that fatherly income matter more than motherly, further analysis would have to be made.

The conclusion in the benchmark scenario for girls will be: the parental influence is higher for native girls, than for immigrant girls and level of education seem more important than income on the probability of education. This implies that there might be other, more important, factors like network or culture for immigrants girls, when it comes to the facts of choosing high school attainment or not.

Boys

A closer look at the differences in parent' effects that appears for boys. The same approach with same stereotype families is used, then the following table is set up based on the four models estimated for boys.

Table 4.10: Predicted probabilities of high school education for boys based on stereotype families

Parental effects	"Stereotype Fam."	NA 06	NA 16	IM 06	IM16	Δ NA	Δ IM
Levels: 1	1 (prior prob.)	17.93%	17.07%	16.33%	23.60%	-4.80%	44.52%
Levels: 2	2	26.90%	32.75%	23.54%	42.69%	21.75%	81.35%
Levels: 3	3	66.01%	72.77%	34.99%	43.44%	10.24%	24.15%
Levels: 4	4	69.15%	73.37%	56.22%	68.37%	6.10%	21.62%
Levels: 5	5	87.67%	90.11%	80.05%	84.31%	2.78%	5.32%
Levels: 6	6	95.26%	94.48%	79.94%	85.40%	-0.81%	6.83%
Dad Edu:	7	38.47%	36.23%	22.69%	29.17%	-5.83%	28.54%
Mum Edu: 3	8	35.38%	36.24%	23.17%	30.29%	2.41%	30.72%
Mum Edu: 6	9	50.86%	42.64%	27.65%	38.52%	-16.16%	39.34%
Dad Edu: 6	10	51.71%	43.05%	27.65%	38.55%	-16.75%	39.45%
Dad Inc: 4	11	24.99%	26.59%	24.17%	34.73%	6.41%	43.70%
Mum Inc: 4	12	23.88%	27.42%	23.55%	32.76%	14.86%	39.13%

For boys one finds the same development over time, as seen for girls. For boys from stereotype families $1-6$ with native origin, it is found that all, but one, predicted probability have developed more, than we saw for the girls. The "but one" predicted probability, is the one for boys with parents from all the lowest groups. His probability of high school education has been reduced by almost 5% points from 2006 to 2016, implying boys from the lowest resource families are worse of today. That the development has increased more, might be corresponding with the fact, that the higher proportion of boys with an education. The predicted probabilities for boys with immigrant origin show, that boys from families with lowest resources have increased the probability of education a lot, but not as much as it has for girls. If an immigrant by comes from stereotype family 2 the predicted probability of education has increased, and it has increased more than it has for girls from the same stereotype family.

Looking at parental effects, the motherly level of income have had an increasing effect on the probability of education among boys. Whereas the influence of fatherly level of income on probability of boys with native origin getting a high school education has increased less over time. Hence motherly income effect are found more important in 2016.

Comparing increases in parental level of income and education among boys and girls with immigrant origin, it is found that increases in parental education and income have greater effect in the development of the probability for girls, than it has had for boys. Meaning that parental effects have become relative more important for girls over the years. Even though the importance of increased parental level of education and income is considered to be way more important for boys, than it is for girls, as the predicted probabilities of education increases way more for boys, than for girls.

Table 4.11: Relative changes in the predicted probability of education among boys when one parental level changes, and all other parental levels are set to 1

"Stereotype Family"	Changing effect	Δ NA 06	Δ NA16	ΔIM 06	Δ IM16
7	Dad Edu: 3	114.58%	112.24%	38.97%	23.60%
8	Mum Edu: 3	97.33%	112.27%	41.89%	28.35%
9	Mum Edu: 6	183.63%	149.79%	69.29%	63.22%
10	Dad Edu: 6	188.37%	152.17%	69.29%	63.35%
11	Dad Inc: 4	39.36%	55.76%	48.00%	47.16%
12	Mum Inc: 4	33.16%	60.65%	44.19%	38.81%

Comparing what changes in parental level of income and education do to the predicted probability of high school education for boys and girls, the increases in parental levels in general seems to be much more important for boys than for the girls, implying boys being more sensitive to parental effects, regardless of country of origin.

For boys with native origin, an increase in fatherly level of education 3 has become less important (the chance in the predicted probability has increased with less in 2016, than it did in 2016), while an increase in motherly level of education to level 3 has become more important in 2016, as the change in the predicted probability has increased from 97% to 112%. One can see that for native origin boys, an increase in parental level of education increases the probability for a high school education much more than seen for the girls with native origin. Moreover it is found for native boys, that increases in parental income also increases the probability of high school education more, than it did for the girls, and the effect has increased more for the boys, than seen for the girls. Furthermore one can conclude that for boys parental effect are important, while motherly effects for girls are found most influential.

Looking at increases in parental level of income and education for boys with immigrant origin, it is found that an increase in level of education from 1 to 3 (high school as highest completed education), motherly education has a higher effect on the probability, than an identical increase for fatherly level of education. But both effects have increased the probability of education with less in 2016, than it did in 2006. Actually all the relative changes in parental effect, increases the probability of an education among the boys less than it did in 2006, but here one should have in mind that the prior probability, which was when the parents have really few resources, was much lower in 2006 than in 2016. Comparing the results with the results for girls, the story is the same, but one see that a fatherly increase in education increases the probability of education more for boys, than it does for girls. Among girls with immigrant origin one saw, that an increase in motherly level of education increased the probability of education more than a corresponding increase in education for the fathers. For boys with immigrant origin, the large increase in parental education increases the predicted probability of education with corresponding values, which is unlike the findings in Esping-Andersen (2004) and Lilliard and Willis (1994), who found that motherly education is more important for girls, and fatherly education is more important for boys.

Based on the analysis I conclude, that natives are more sensitive to parental effects than immigrants, furthermore that boys are more sensitive than girls. For natives we see that the parental educational effects on the probability of education have decreased over time and the parental income has increased its influence over the years, but that the parental educational factor is the most important. While the parental effects for immigrants are unchanged or increased over time. What have caused the more unclear pattern in parental effects on the probability of education among immigrants is unknown, but it might indicate that other factors, such as network or culture are more important.

For girls I found that motherly level of education is most important, where the parental level of education are more equal in its effect on the probability of education for boys, but with fatherly level of education to matter most. The results of this analysis will later on be referred to as "benchmark".

4.3 Robustness Tests

This section will go through some robustness test of the results, to see how sensitive the findings are to whether parents live together or not. It will analyze how differences might appear within a few municipalities and lastly examen if the results for immigrants depends on country of origin, when they are split up into Wester origin and non-Western origin.

4.3.1 Removing individuals with only one parent registered in DK

This section very briefly presents a review of, whether parental effect on the probability of individuals getting a high school education changes, if al individuals with only one parent registered in Denmark are removed. For a full review of data, model estimations and results see appendix B.
It is interesting to research, as Weiss (1978) based on samples from US, find that the vast majority of one-parent families hold a disadvantageous position in society relatively to other family groups. Furthermore danish researches (as Arbejderbevægelsens Ervervsråd, AE) finds that single-parent children are doing worse in primary school.[8]

Removing the individuals with one parent registered in Denmark in 2006 has almost no effect on the proportion of individuals with a high school education within the native origin dataset (0.07%), while it increased the proportion of individuals with immigrant origin with 1.59% point, corresponding to 4.98%. A review of the "new" datasets gender specifications finds, no evolving change in native origin, while the individuals with immigrant origin has increased the proportion of as well boys as girls with a high school education. In 2016 it is found that the native origin data sample has increased the proportion of high school educated individuals by 1.09% point, corresponding to a increase in 2.07% the immigrant proportion with an high school education has increased 2.1% corresponding to 4.53%. The changes in the proportions within the natives are very low, while the proportion of boys as well as girls with immigrant origin who have an education, is raised. Even though the relatively small changes it gave in the sample composition for the native origin, the finding are corresponding to the findings in Weiss (1978).

The effect plots for native origin shows differences across parental effects on boys and girls, for respectively stereotype families of low-resources and high-resources. The effect plots of predicted probability among the boys, shows no clear changes in effect of parental effects on the predicted probability of education when one compare to benchmark, which included single parents and neither is the case for girls. Testing if parental effects significantly differ among boys and girls without single parents it is found that no significant chances appear. The effect plots for immigrant origin shows differences across parental effects on boys and girls, for respectively stereotype families of low-resources and high-resources the parental effects look much alike for the boys. Which tells, that motherly education has a (very) little higher effect on boys getting an education, if the mother is a single parent registered parent in Denmark. Among the girls a more clear sign appear: the prior probability of education has increased, but nothing applies that

---

[8] [46]

the parental effects on the probability of education have changed much. Whether the parental effects differ between gender are again tested. When testing at a 5% level, no parental effects are found to differ among boys and girls.

For girls with native origin, the changes compared to benchmark are to small for even commenting on. Among girls with immigrant origin few, but no dramatic changes have been found. One sees that girls from every stereotype families have higher probability of getting a high school education. Furthermore it is found that the development in the probability of education over time, has increased less than in the benchmark scenario, which implies that the parental effects on the probability of education for girls with immigrant origin and both parents registered in Denmark have increased more. Moreover, it is found for that parental income is just as important, as parental level of education for girls with immigrant origin. Which is unlike the findings in Bukodi and Goldthorpe (2012) who found, that parental level of education was the most important on educational attainment of their child.

As the predicted probability for a girl from stereotype family 1 , i.e. a girl from a family with absolute lowest level of income and education, are increased relatively more than the probability of education in the other stereotype families this implies that the changes in parental effects, seems to have less impact on the probability of education.

The predicted probabilities of education among boys allows an identical analysis for boys with native origin, as in benchmark: highly educated parents (level 6) does not raise the probability of education as much in 2016, as it did in 2006, i.e. parental level of educational as less important in the probability for education among boys, but the effect is still quite large.

For boys with immigrant origin, it is found that the probabilities of education are increased no matter what stereotype family he comes from, but especially if a boy comes from stereotype family 1. The relative effect on the probability given a change in one parental effect are reduced, but this is probably caused by the increase in the probability of education for a boy from stereotype family 1. If this is caused by, the model doing worse or because more boys are social mobile and parental effects therefore are less important/similar among the boys, could be interesting to study further. Comparing the parental effects on the probability of education among boys and girls with immigrant origin, one finds that, when removing all individuals without both parents registered in Denmark, girls over the years have become more sensitive to parental levels of education and income, but remember they in general are less sensitive to parental effects.

From these comments one may conclude that the estimated models are therefore quite robust to whether both parents are registered in Denmark or not which leads to the next robustness test: do the parental effects on the probability of their children getting an education depend on whether parents live together or not?

4.3.2 Are parents living together - one household?

The aim of this section is to investigated, if the findings of parental impact on the probability of young individuals gets an education or not and if it in some way are connected with whether parents are living together, when the individual are 17 years old +/− 2 years. This topic is found interesting to analyze, as studies have shown, that parents divorce is associated with lower educational attainment among their children (see Keith and Finlay (1988) and McLanahan and Sandefur (1994) and Weiss (1994)) The variable *efalle*, from Statistic Denmark provides information about the partner that an individual live with. Whether they are couples either in the form of marriage, registered partnership, cohabiting couple or cohabiting couple. Two individuals who are living together have their partner social security number (*PNR*) as *efalle*. By creating a variable listing mother's partner (her *efalle*), it is possible to identify whether her registered partner is the father to a child, as fatherly identification number is known and equal to her *efalle* if they are living together.

A review of average parental income and level of education shows, that parental levels of education and income are at higher levels when parents do *live* together, and when they do *not* live together, their average level of education and income is lower, compared to the benchmark case. And the changes in average are not small.[9] This is corresponding to Ploug (red.) (2017), as he point to parents with higher resources more often stay together, than parents with low resources.[10]

2006

It is found that 58.4% of the individuals with native origin in 2006 had parents who live together. Among those, 52.7% of the individuals have a high school education (compared to 47%), while it is only 39.14% of the individuals with parents who do *not* live together, who have a high school education. A review of the proportion of boys and girls with an educations shows, based on whether parents live together or not is set up.

Table 4.12: Native origin data from 2006, proportion with education

Gender	Graduated HS	Parents live together		Parents do *not* live together	
	High School	#	Gender %	#	Gender %
Boys	No (0)	87,509	55.0%	76,030	67.45%
	Yes (1)	71,585	45.0 %	36,695	32.55%
Girls	No (0)	59,713	39.23%	58.954	54.05%
	Yes (1)	92,504	60.77%	50.116	45.95%

Here one notes the clear differences on the proportion of individuals with an education, not only are

---

[9] See appendix C and D for summary of parental income and education

[10] see *Lighed gennem uddannelse*

girls sensitive to whether the parents live together or not, relatively it seems to be more important for boys with native origin. The very low proportion of individuals with an education among individuals with parents who do *not* live togehter, might due to the fact, that it is more often parents of low resource families who are divorced.[11] These findings corresponds to Keith and Finlay (1988) who finds that parental divorce has negative consequences for children's educational attainment and McLanahan and Sandefur (1994), how found that children whose parents are do not live together are twice as likely to drop out of high school as those whose parents live together.
Among individuals with immigrant origin based on data from 2006 it is 62.02% of the individuals who have parents registered at the same address. Among these, 33.12% of them have an education, while it is only 28.76% of the individuals without parents live together who have a high school education, compared to the average 32%. So again, are parents live together then the fraction of educated individuals is higher. Furthermore one sees that a lit is slightly more important for boys.

Table 4.13: Immigrant origin data from 2006, proportion with education

Gender	Graduated HS	Parents live together		Parents do *not* live together	
	High School	#	Gender %	#	Gender %
Boys	No (0)	8,609	70.68%	5,803	75.43%
	Yes (1)	3,571	29.32%	1,890	24.57 %
Girls	No (0)	6,826	62.63%	4,696	66.66%
	Yes (1)	4,073	37.37%	2,349	33.34%

2016

A closer look at the data from 2016 and the native origin, a clear differences in whether the parents live together or not. In the datasets 54.1% of the parents live together, which is more than 4% points less than in 2006. In the native sample, the fraction of individuals with a high school education is 45.42% among the individuals, whose parents are not registered together, while the fraction is 60.91% when the parents are registered with the same address. The increase in the proportion with an education, is larger among children whose parents live together.

Table 4.14: Native origin data from 2016, proportion with education

Gender	Graduated HS	Parents live together		Parents do *not* live together	
	High School	#	Gender %	#	Gender %
Boys	No (0)	74,276	46.81%	86,091	61.4%
	Yes (1)	85,721	53.58%	54,112	38.6 %
Girls	No (0)	47.140	31.29%	63,943	47.48%
	Yes (1)	103,506	68.71%	70,732	52.52%

---

[11] See Ploug (red.) (2017), Lighed gennem uddannelse

Compared to 2006 we see, that especially the proportion of educated boys with parents live together has increased relatively most, implying that boys have become more sensitive to the fact of one parents households over the years.

From the dataset based on the immigrant origin in 2016, it is found that 59.4% of individuals have parents who live together. The chances in the proportion of individuals with an education also has changed, but not as much as just seen among the natives. In the dataset where the parents live together, the fraction of individuals with a high school education is 49.95%, which is a small increase compared to the simple case with the full dataset. Among individuals with parents who have not the same registered address 40.13% has an education.

Table 4.15: Immigrant origin data from 2016, proportion with education

Gender	Graduated HS	Parents live together		Parents do *not* live together	
	High School	#	Gender %	#	Gender %
Boys	No (0)	10,945	56.01%	8,840	65.44%
	Yes (1)	8,595	43.99%	4,688	34.56 %
Girls	No (0)	7,519	43.33%	6,568	53.70%
	Yes (1)	9,833	56.67%	5,662	46.3%

Estimation of models[12]

2006

Setting up the models for native origin individuals with parents*living* together, almost all levels of parental effects are found significant. For boys if motherly income level is 2, her effects on the education is not significant, when it for girls is a fatherly level of income equal to 2, which is insignificant. The classification tables shows, that the model for boys is really good at estimating the boys without an education correctly, implying that a clear pattern in parental effects are found. In the model for girls, the mis-classification rate of girls classified as not having an education, when they actually do, is really high. This might imply that many girls differ from the pattern in parental effects, meaning that more girls are social mobile or at least for follow the parental effect pattern, that the model have found. The effect plots shows, that if a girl comes from a family with just a little higher resources (i.e. parental level equal to two) her probability of education increases more, than seen for the boys. Mutual findings within the model estimations are found, when parents do *not* live together. The classification tables shows, that the model estimated for boys have become even better in classifying boys without an education. This tells, that the patter in parental effects for boys not getting an education is more clear. The effect plots shows, that signs of a higher sensitivity of parental levels of education and income among the boys.

---

[12] See appendix C for documentation of the models and effect plots for parents who live together and appendix D for parents do not live together

Setting up the models for individuals with immigrant origin whose parents *do* live together, all levels except the lower levels of fatherly income seems to have positive significant effect on the probability of education. The models are classifying approximately the same proportion correct. The effect plots for boys, shows that parental effects on the probability have increased quite a lot compared to the benchmark case. For girls the effect plots do not show any clear changes, other than all over higher probability of education. The test of whether the parental effects differ between boys and girls shows, that no differences are significantly. For parents who do *not* live together some of the the maximum likelihood estimates, i.e. the parental effects, are not found to be significant and the models classification rates have decreased just a little. The model estimated for boys has a very poor rate at classifying boys with an actual education, furthermore it is mis-classifying a high rate of boys, as not having an education, while the actually do. The effect plots for boys imply, that parental effects are more important then in benchmark. Among the girls, the parental effects also seems to have a greater impact on the probability of education, than seen in benchmark case. The test shows no sign of different parental effects on boys and girls at a significant level.

2016

Estimating models based on individuals with native origin and parents who *live* together, the models are found significant and the maximum likelihood estimators are found with the same indications as we have seen before: increase in parental level of income and education implies an increased probability of education. Moreover it is found that the models mis-classify at a little higher rate. Especially the model for girls are classifying relatively more girls as *False NEG*, i.e. a higher proportion of girls are predicted as not having an education, when they actually has one. This could be an indicator of more girls with parents who live together are social mobile. The effect plots implicates lower parental impact on the probability of education. For girls from a stereotype family with lower resources, the predicted probabilities of education are found to be higher, but not much has changed. The predicted probabilities of education among girls given a higher resource family, it looks like the differences in fatherly level of education has less effect on the probability of education.

The models estimated, based on native parents who do *not* live together, are again found significantly, meaning that explanatory tells something about response variable (high school education). The classification of the model for boys mis-classify at a higher rate Implying, that the boys parental effects are more alike, when parents live together. The model for girls has reduced rate at classifying girls with education correctly, but instead the model is mis-classifying fewer girls as not having an education, when they actually has one. Implying either fewer girls are being social mobile, or that parental effects have become more similar to whether the girls are getting an education or not. The effect plots for boys indicates some changes in parental effects. In both stereotype family cases al the level of parental effects seems to have a greater impact on the probability of a boy getting an education, as the slope, the intercept and the gabs between the lines have increased. The parental effects, when parents are *not* living together, especially the motherly effects, becomes even more clear in the effect plots based on the model estimated for girls. Here the gabs between the lines, i.e. mother income level, have increased in both stereotype families and slopes has increased. The differences between the two associated plots, are not that clear, implying that fatherly effect are not as important as the motherly effects on the probability of education among the girls.

Testing whether parental effects differ between boys and girls, when parents live together, provides the same significant differences as the simple case including single parents.

Performing the test, a positive values of the maximum likelihood estimates in the test results implies that the parental effect is larger for girls than for boys (given all the maximum likelihood estimates are positive in the regression). The maximum likelihood estimates which are negative and significant, shows that the parental effects are higher among boys than among girls - therefore motherly income levels are of greater impact for the boys. Testing the parental effects, when parents are not living together it is found that more motherly levels of education are significantly different between boys and girls, while fewer of fatherly levels of education have become insignificant. Implying mothers have greater effect on the probability of education. When testing if parental effects differ between the models estimated for girls and boys, it is found that fatherly education has higher effects on increased probability of education among boys and girls, when parents are living together, while motherly level of education have significantly higher effect on the probability of education when parents are *not* living together.

Estimating the models for immigrant origin based on data from 2016, with parents *living* together no significant changes in the maximum likelihood estimators are found compared to benchmark. Both models are found to classify approximately at the same level, but the model for the girls seem to be a little more skew, as the rate of mis-classifying girls as not having an education, has increased. This might imply that girls with parents living together are more social mobile. At the effect plots for boys, it looks like motherly level of income in a high resource stereotype family has more effect on the probability of education, as the gab between the lines has increased a relative much, compared to benchmark. The effect plots based on the model estimated for girls tells an identical story. Motherly income effect on the probability of education has increased within the high resource stereotype family, while the opposite seem to be the case for motherly income effect for low resource stereotype families, as the gab between the lines has decreased.

Testing whether the parental effects differ between boys and girls. It is still found, that there are almost no significant difference in the parental effects on the probability of education among boys and girls with immigrant origin. The only parental effects with a significant difference on the probability of education among boys and girls are if fatherly level of education is higher, i.e. if he has a 2 years university education or if he has at least a master (minimum 5 years at university, then the effect increases the probability of getting education relative more for girls, than for boys. This might imply that girls are raised with the knowledge to educated them-self when fathers are educated, while boys might be raised with more free frames.

Bases on individuals with parents who do *not* live together, it is among the immigrant origin in 2016 found, that the maximum likelihood estimates tells the same story as seen and told before. The models are mis-classifying at a higher rate compared to both the models estimated in the benchmark case and the models estimated based on parents *live* together. The classification tables shows that the predicting the probability of high school for boys, have become more skew, as it wrongly predicts relative many boys as having an education, when they do not, than the other way around. On the other hand the model estimated to predict the probability of high school for girls is doing better at its mis-classifications. The effect plots for boys tells that motherly level of income has less effect on the probability of education, while motherly education has larger effect on the probability of education, when compared to the case where parents are living together. The effects plot for girls indicates the same story: Increase in motherly level of education increases the probability of education far more than an increase in motherly level of income, like the findings in Esping-Andersen (2004). When testing if the parental effects differ with gender, no significant differences between parental effects on the probability of education among boys and girls are found. For being able to compare the models with earlier findings, and the benchmark scenario, the estimated models are used to predict the probability of education among the 12 fictive stereotype families.

The models used for stereotype families to predict probabilities of education - how do parental effect seem to differ?

When parents *live* togehter

For native girls, one can conclude that the "prior probability" of education, (when a girl comes from stereotype family 1, i.e. parental levels of income and education are at the lowest level) is at a higher level than in the benchmark case and furthermore that the level in 2006 was very much higher, than in 2016. This might imply that in 2006 girls were more sensitive to the fact that parents live together, than in 2016. In 2016 it is not found to increase the prior probability of education that much, meaning that over the years whether parents live together or not, has less effect for the prior probability of education among native girls. In general it is found, that girls of lower resource families $1-3$ are more sensitive to parents living together as their predicted probability of education is higher than seen in benchmark, than girls with high resource parents, when it comes to high school education.

Table 4.16: Predicted probabilities of high school education for girls based on stereotype families, given the parents *live* together

Parental levels	"Stereotype Fam."	NA 06	Na16	IM 06	IM16	Δ NA	Δ IM
Levels: 1	1	38.92%	31.22%	26.74%	38.92%	-19.78%	45.55%
Levels: 2	2	54.41%	59.10%	36.37%	54.50%	8.62%	49.85%
Levels: 3	3	66.29%	84.05%	45.12%	66.29%	26.79%	46.92%
Levels: 4	4	78.72%	86.81%	67.92%	78.72%	10.28%	15.90%
Levels: 5	5	86.28%	93.89%	87.91%	86.28%	8.82%	-1.85%
Levels: 5/6	6	88.47%	95.12%	80.08%	88.47%	7.52%	10.48%
Dad Edu: 3	7	43.84%	52.52%	33.56%	43.84%	19.80%	30.63%
Mum Edu: 3	8	51.48%	50.88%	37.29%	51.47%	-1.17%	38.03%
Mum Edu: 6	9	57.29%	57.77%	31.69%	57.29%	0.84%	80.78%
Dad Edu: 6	10	50.75%	56.89%	36.89%	50.75%	12.10%	37.57%
Dad Inc: 4	11	50.91%	43.77%	30.66%	50.91%	-14.02%	66.05%
Mum Inc: 4	12	54.63%	44.77%	41.52%	54.63%	-18.05%	31.58%

Unlike benchmark, the parental effects on the probability of education seem to have increased a little over time for native girls, but here one should have the prior probability of education in mind.

For girls with immigrant origin, the prior probability of eduction, i.e. the probability of education for a girl from stereotype family 1, has increased compared to benchmark. In 2006 the increase was higher than seen in 2016, but unlike the native girls, the probability of education for a girl of stereotype family 2-6, has not changed much compared to benchmark. This implies that only immigrant origin girls, from families with very low resource family have a higher probability of getting a high school education, when her parents live together, the increased prior probability of education leads to lower relative effects when one parental level is changed.

One note, that motherly level of education is found to have greater impact on the probability of education among the native and immigrant origin girls, than fatherly level of education, which correspond to the findings in Esping-Andersen (2004).

Table 4.17: Relative changes in the predicted probability of education among girls whose parents *live* together, one parental effect change and all other parental levels are set to 1

"Stereotype Fam"	Changing effect	Δ NA 06	Δ NA16	Δ IM 06	Δ IM16
7	Dad Edu: 3	12.64%	68.23%	25.50%	12.64%
8	Mum Edu: 3	32.27%	62.97%	39.45%	32.25%
9	Mum Edu: 6	47.20%	85.04%	18.51%	47.20%
10	Dad Edu: 6	30.40%	82.22%	37.96%	30.40%
11	Dad Inc: 4	30.81%	40.20%	14.66%	30.81%
12	Mum Inc: 4	40.36%	43.40%	55.27%	40.36%

A review of the finding for girls, when parents are live together compared to the benchmark case: native origin girls from lower stereotype families (families 1-3) are more likely to get a high school education, than in benchmark, while it is only immigrant origin girls from stereotype family 1 who have increased probability of education. This implies that girls are more social mobile, when parents live together. When parents live together, I can conclude that parental effects are relative less important (but the fact that they live together is important) than in benchmark, but that their influence have become more important over time. This might correspond to the fact that the prior probability of education was lower in 2016 and that the pattern in parental effect has become more clear.

Furthermore it is once again found, that motherly effect, especially education, is the most important factor among girls' probability of education. But in the scenario of parents living together, I notes that the parental effects have become more equal for native girls, meaning that fatherly effects increases when parents live together.

For native origin boys, it is also found, that the prior probability of education is higher, especially in 2006. Comparing benchmark to when parents *live* together in 2016 the probability of education does not increase much. This implies, that boys from families with very low resources, over the years have become less sensitive to whether parents live together or not. In general boys of native origin from families with fewer resources (stereotype family 1-3) in 2006 were very sensitive to the fact that parents live together or not.

Table 4.18: Predicted probabilities of high school education for boys based on stereotype families, given parents live together

	"Boy"	NA 06	Na16	IM 06	IM16	Δ NA	Δ IM
Levels: 1	1	25.37%	19.67%	18.12%	25.37%	-22.47%	40.01%
Levels: 2	2	42.47%	37.57%	25.60%	42.47%	-11.54%	65.90%
Levels: 3	3	49.76%	74.73%	37.32%	49.77%	50.18%	33.36%
Levels: 4	4	73.41%	73.49%	55.87%	73.41%	0.11%	31.39%
Levels: 5	5	79.01%	90.55%	75.82%	79.01%	14.61%	4.21%
Levels: 5/6	6	81.51%	94.12%	74.40%	81.51%	15.47%	9.56%
Dad Edu: 3	7	30.78%	42.24%	24.26%	30.78%	37.23%	26.88%
Mum Edu: 3	8	32.65%	35.70%	26.21%	32.65%	9.34%	24.57%
Mum Edu: 6	9	36.33%	44.23%	24.40%	36.33%	21.75%	48.89%
Dad Edu: 6	10	43.84%	48.71%	30.70%	43.84%	11.11%	42.80%
Dad Inc: 4	11	36.29%	28.86%	24.46%	36.28%	-20.47%	48.32%
Mum Inc: 4	12	39.11%	29.11%	27.56%	39.11%	-25.57%	41.91%

For immigrant boys, the prior probability also have increased, but just as for girls with immigrant origin, it only seems like the fact that parents living together has an impact on the probability for boys from families with very low resources. But unlike for the immigrant origin girls, the over time development in the stereotype families $3-6$ is higher for the boys when parents live together. But even though I did not find any signs of significant difference in parental effect among boys and girls with immigrant origin, it seems like, the parental impact is higher on the probability of education for boys, than for girls in 2016, as a change in one parental level, increased the relative probability of education more for boys, than for girls.

Table 4.19: Relative changes in the predicted probability of education among boys when parents live together, when one parental level change and all other parental levels are set to 1

"Boy"	Changing effect	Δ NA 06	Δ NA16	Δ IM 06	Δ IM16
7	Dad Edu: 3	21.32%	114.74%	33.89%	21.32%
8	Mum Edu: 3	28.70%	81.49%	44.65%	28.70%
9	Mum Edu: 6	43.20%	124.86%	34.66%	43.20%
10	Dad Edu: 6	72.80%	147.64%	69.43%	72.80%
11	Dad Inc: 4	43.04%	46.72%	34.99%	43.00%
12	Mum Inc: 4	54.16%	47.99%	52.10%	54.16%

The finding for boys, when parents live together compared to the benchmark case: fatherly level of education, seems to be very important on the probability of education, again in correspond to the findings of Esping-Andersen (2004). Furthermore the probability of education for native origin boys from lower stereotype families (1-2) is higher than in benchmark, while it is only immigrant origin boys from stereotype family 1 who have increased probability of education. This implies, that boys of low resource families are more likely to be social mobile, if parents live together, which corresponts to the general finding in McLanahan and Sandefur (1994), where children of single-parents are found to meet greater challenges in life. Moreover the development in parental effects, seem to have increased over time, regardless of boys' country of origin. Implying the pattern in parental effect on the probability of education has become more clear over the years. Furthermore one might notes that for native boys the effect of an increased parental income has reduced a lot over time. Why parental level of income have become less important when parents live together, could be interesting to do further analysis on. These results seem to fit the effect plots.

When parents do *not* live together

The predicted probabilities are found for the girls and boys from the stereotype families, when parents do *not* live together. For girls of native origin the prior probability of education, i.e. the probability of education for a girl from stereotype family 1, is found to be lower in 2006, but much higher in 2016, compared to the benchmark case. This could be caused by an increasing development of more social mobile girls from families among girls whose parents do not live together, or caused by the fact that the pattern in parental effect have become less clear, and other factors could be more important. Factors as stepparental level of education and income or network. In 2016 the native girls are just as sensitive to parental income as to parental education, which was not the case in benchmark. Furthermore the pattern in the parental influence have decreased a lot, this might imply that other factors such as stepparents or network have become more important.

Table 4.20: Predicted probabilities of high school education for girls based on stereotype families, given parents do *not* live together

Parental level	"Stereotype Fam."	NA 06	Na16	IM 06	IM16	Δ NA	Δ IM
Levels: 1	1 (prior prob.)	27.00%	38.92%	20.35%	31.83%	44.15%	56.41%
Levels: 2	2	34.91%	54.41%	28.91%	47.26%	55.86%	63.47%
Levels: 3	3	75.93%	66.29%	40.25%	52.65%	-12.70%	30.81%
Levels: 4	4	79.11%	78.72%	58.11%	71.01%	-0.49%	22.20%
Levels: 5	5	90.13%	86.28%	90.38%	87.35%	-4.27%	-3.35%
Levels: 5/6	6	96.06%	88.46%	90.13%	87.59%	-7.91%	-2.82%
Dad Edu: 3	7	46.75%	43.84%	24.18%	37.19%	-6.22%	53.80%
Mum Edu: 3	8	51.84%	51.47%	28.18%	41.58%	-0.71%	47.55%
Mum Edu: 6	9	64.19%	57.30%	43.29%	56.96%	-10.73%	31.58%
Dad Edu: 6	10	57.90%	50.75%	27.38%	38.89%	-12.35%	42.04%
Dad Inc: 4	11	33.29%	50.91%	25.86%	41.22%	52.93%	59.40%
Mum Inc: 4	12	38.91%	54.63%	29.70%	42.61%	40.40%	43.47%

For the probability of education among girls of immigrant origin the story is different. Here the prior probability of education is lower than in the benchmark case. Actually in most stereotype families, the predicted probability of education is lower i 2016 compared to benchmark, furthermore the girls of immigrant origin are found to be relative more sensitive to motherly income and education than fatherly, compared to benchmark. Even though the fatherly impact on the probability of education has increased most over time. This implies that girls with immigrant origin are much more sensitive to whether parents live together or not, when it comes to the probability of education.

Table 4.21: Relative changes in the predicted probability of education among girls with parents *not* living together, when one parental level changes and all other parental levels are set to 1

"Girl"	Changing effect"	Δ NA 06	Δ NA16	Δ IM 06	Δ IM16
7	Dad Edu: 3	73.15%	12.64%	18.82%	16.84%
8	Mum Edu: 3	92.00%	32.25%	38.48%	30.63%
9	Mum Edu: 6	137.74%	47.23%	112.73%	78.95%
10	Dad Edu: 6	114.44%	30.40%	34.55%	22.18%
11	Dad Inc: 4	23.30%	30.81%	27.08%	29.50%
12	Mum Inc: 4	44.11%	40.36%	45.95%	33.87%

When parents do *not* live together and are stereotype families of lower levels, the findings for girls are very depending on origin, as immigrant girls are found to be relative more sensitive to the parental effects when parents do not live together. Regardless of origin, mothers level of education and level of income are found to have greater impact on the probability of education than fatherly effects. For girls with immigrant origin the fatherly impact is lower on the probability of education in 2016 than in 2006, and motherly level of education are found to have much higher influence than in benchmark and when parents live together. This might imply, that if immigrant girls live with a single parent, it is more often to be their mother.

Comparing these findings to benchmark and to when parents live together, then it is found that girls with immigrant origin are relatively more sensitive to motherly levels of education and income, while native girls are relatively less sensitive to parental changes in 2016, when parents do not live together. Implying that when parents do not live together maybe other factors have become more important on the probability of education among native girls. These factors could be stepparent's level of education and income or maybe network.

Among boys with native origin, the prior probability of education also have increased, and like for native girls the probability of education has increased most among the lower levels of stereotype families (1-2). Furthermore it is found, that parental level of education have an even larger influence on the probability of education among boys, than seen in benchmark or when parents live together in 2006. The prior probability of education have increased in 2016 which provides the information, that parental effects have less effect in 2016 than in 2006. This implies, that the pattern in the influence of parental effect on the probability of education among native boys have become less clear. Which could be caused by the fact, that other factors such as stepparents, or network could by an important factors, when it comes to native boys high school attainment.

Table 4.22: Predicted probabilities of high school education for boys based on stereotype families, given parents do *not* live together

	"Boy"	NA 06	Na16	IM 06	IM16	Δ NA	Δ IM
Levels: 1	1 (prior prob.)	18.38%	25.37%	16.59%	21.77%	38.03%	31.22%
Levels: 2	2	19.81%	42.47%	16.61%	40.77%	114.39%	145.45%
Levels: 3	3	59.75%	49.73%	30.52%	37.98%	-16.77%	24.44%
Levels: 4	4	64.18%	73.42%	45.66%	60.65%	14.40%	32.83%
Levels: 5	5	86.27%	79.01%	87.29%	88.52%	-8.42%	1.41%
Levels: 5/6	6	95.41%	81.51%	82.55%	89.06%	-14.57%	7.89%
Dad Edu: 3	7	37.37%	30.78%	21.46%	25.71%	-17.63%	19.80%
Mum Edu: 3	8	35.73%	32.66%	22.37%	29.54%	-8.59%	32.05%
Mum Edu: 6	9	53.78%	36.33%	35.69%	45.34%	-32.45%	27.04%
Dad Edu: 6	10	50.49%	43.85%	20.69%	28.02%	-13.15%	35.43%
Dad Inc: 4	11	22.27%	36.28%	19.69%	29.85%	62.91%	51.60%
Mum Ind: 4	12	24.88%	39.11%	22.54%	28.47%	57.19%	26.31%

The results for natives in 2006 and for immigrant boys seem to add up with the fact that divorced mothers currently have child custody more than 90% of the time (Glick (1979)) and the fact that boys somehow are expected to have greater postdivorce problems than girls. Since it appears for boys with divorced parents and if mother has custody, they have a greater risk at being troubled, which seems to ad up to the fact, that fewer boys with parents not living together get a high school education. The fact that motherly level of education has decreased over the years for native boys, might imply that fathers in higher levels gets custody or that factors like network or stepparents are more important.

For boys of immigrant origin it is found, that in 2006 the prior probability is relatively the same as the benchmark case, while in 2016 the prior probability of education is lower, meaning that immigrant boys from families with very low resources are set worse, if parents do *not* live together. The historical development in the predicted probability of education does not change much, compared to the benchmark case. For immigrant boys, with parents not living together, motherly levels of education tends to have a much higher impact on the probability of education. This might be due to the fact that mothers more often tend to have the custody Glick (1979) and therefore are more responsible and a figure to look up to. Or to the fact that more mothers than fathers are registered in Denmark.

Table 4.23: Relative changes in the predicted probability of education among boys with parents do *not* live together, when one parental level changes and all other parental levels are set to 1

"Boys"	Changing effect"	Δ NA 06	Δ NA16	Δ IM 06	Δ IM16
7	Dad Edu: 3	103.32%	21.32%	29.36%	18.10%
8	Mum Edu: 3	94.40%	28.73%	34.84%	35.69%
9	Mum Edu: 6	192.60%	43.20%	115.13%	108.27%
10	Dad Edu: 6	174.70%	72.84%	24.71%	28.71%
11	Dad Inc: 4	21.16%	43.00%	18.69%	37.12%
12	Mum Inc: 4	35.36%	54.16%	35.86%	30.78%

Therefore one can conclude: that parental level of education seems to have a very large impact on young Danes probability of education. Especially the ones with native origin, when parents do not live together, but historically the parental effect are less important, or at least not as powerful as in the past, which could imply other factors could be explaining the relationship better.

The fact that parents getting divorced/no longer live together is hard on both boys and girls. And in this thesis it is presented, that the proportion of educated individuals are much lower, when parents are split up, which corresponds to the fact that individuals living with single parents or stepparents are found to receive less parental encouragement and attention with respect to education (see Keith and Finlay (1988)).

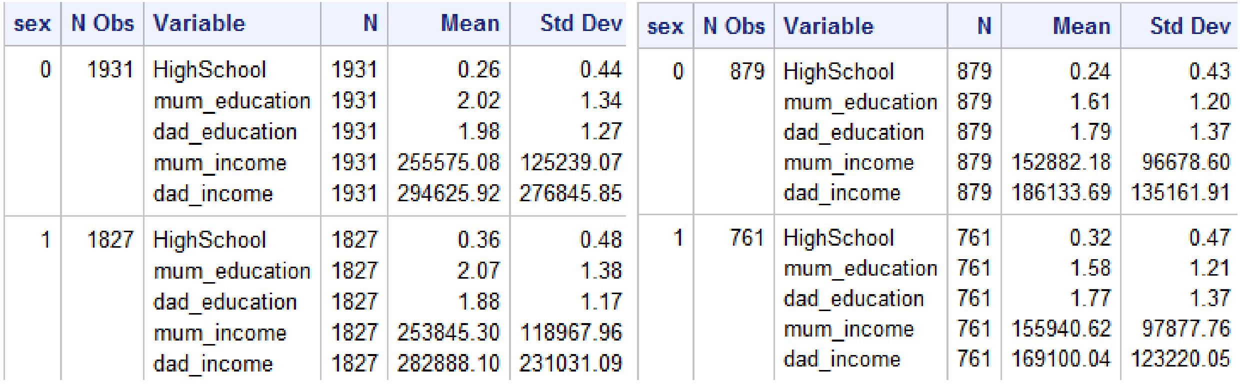
Therefore it is found that the models are not very robust, when it comes to if parents are living together or not. An interesting extension to these studies could be looking at stepparents and their affects on the probability of education among children.

4.3.3 Does the results depend on municipality of residence?

As many factors, as average income, average level of education and political conviction etc. differ across municipalities and furthermore as Bennett et al. (2017) finds a positive correlation between negative attitude towards immigrants on municipality levels, measured as votes for Dansk Folkeparti, and proportion of boys with immigrant origin attending high school, it motivated me to investigate whether parental effect might differ across a municipalities. I have chosen to take a few municipalities into account. These municipalities are chosen mostly based on interest, but also on proportion of immigrants in municipalities.[13] Ishøj is the municipality in Denmark with the highest proportion of immigrants, here 37.5% of the inhabitants are immigrants or descendants, second highest on the list of municipalities with the highest proportion of immigrants origin is Brøndby, where 31.2% inhabitants are immigrants or descendants. As the municipalities are quite close geographically but also quite similar in income and level of education, they will be treated as one. Another set of data is created to investigate based on municipalities with a lower proportion of immigrants. The dataset includes the individuals from Kolding and Vejle. In the two municipalities the proportion of immigrants is lower, as they were respectively 11.6% and 11.1% in 2016.

Setting up the datasets to work with, the data for the two municipalities with the highest proportion of immigrant origin citizens, Ishøj and Brøndby based on the data for 2016, involves $6,307$ individuals, where $3,740$ have native origin and $2,567$ with immigrant origin born between $1986-1996$. While in 2006 the municipalities together had $3,758$ individuals with native origin and $1,640$ individuals with immigrants born between $1976-1986$. A review of the parental information within the two municipalities

Figure 4.5: Summary of parental information. Left Table: Native origin, Right Table: Immigrants 2006 Brøndby and Ishøj



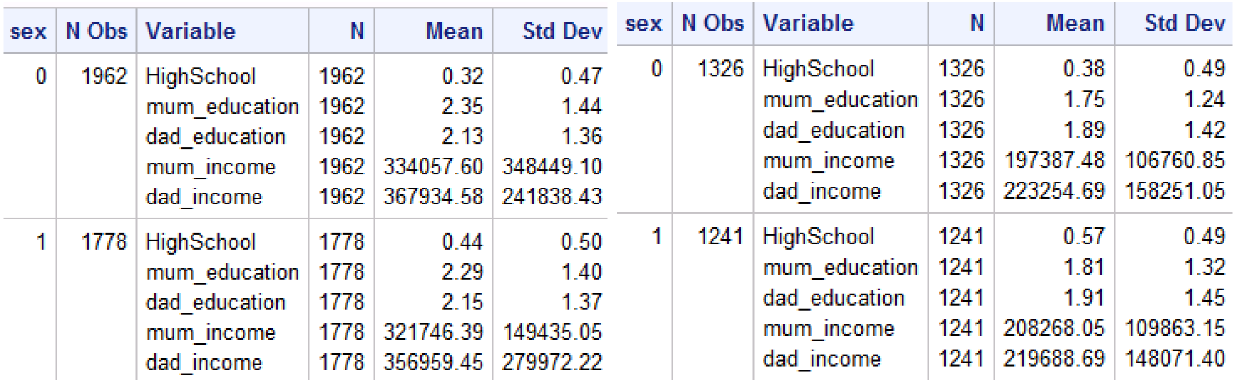
sex	N Obs	Variable	N	Mean	Std Dev
0	1931	HighSchool	1931	0.26	0.44
		mum_education	1931	2.02	1.34
		dad_education	1931	1.98	1.27
		mum_income	1931	255575.08	125239.07
		dad_income	1931	294625.92	276845.85
1	1827	HighSchool	1827	0.36	0.48
		mum_education	1827	2.07	1.38
		dad_education	1827	1.88	1.17
		mum_income	1827	253845.30	118967.96
		dad_income	1827	282888.10	231031.09

sex	N Obs	Variable	N	Mean	Std Dev
0	879	HighSchool	879	0.24	0.43
		mum_education	879	1.61	1.20
		dad_education	879	1.79	1.37
		mum_income	879	152882.18	96678.60
		dad_income	879	186133.69	135161.91
1	761	HighSchool	761	0.32	0.47
		mum_education	761	1.58	1.21
		dad_education	761	1.77	1.37
		mum_income	761	155940.62	97877.76
		dad_income	761	169100.04	123220.05

---

[13] Proportions of immigrants are found in [52] table 1.5.

Figure 4.6: Summary of parental information. Left Table: Native origin, Right Table: Immigrants 2016 Brøndby and Ishøj



sex	N Obs	Variable	N	Mean	Std Dev
0	1962	HighSchool	1962	0.32	0.47
		mum_education	1962	2.35	1.44
		dad_education	1962	2.13	1.36
		mum_income	1962	334057.60	348449.10
		dad_income	1962	367934.58	241838.43
1	1778	HighSchool	1778	0.44	0.50
		mum_education	1778	2.29	1.40
		dad_education	1778	2.15	1.37
		mum_income	1778	321746.39	149435.05
		dad_income	1778	356959.45	279972.22

sex	N Obs	Variable	N	Mean	Std Dev
0	1326	HighSchool	1326	0.38	0.49
		mum_education	1326	1.75	1.24
		dad_education	1326	1.89	1.42
		mum_income	1326	197387.48	106760.85
		dad_income	1326	223254.69	158251.05
1	1241	HighSchool	1241	0.57	0.49
		mum_education	1241	1.81	1.32
		dad_education	1241	1.91	1.45
		mum_income	1241	208268.05	109863.15
		dad_income	1241	219688.69	148071.40

tells, that parental level of income is lower than in the benchmark case, and relative lower for the native parents. The dataset for the municipalities in Jytland, includes 18,103 young individuals with native origin and 2,016 individuals with immigrant origin. Based on data from 2006 the sample includes 12,640 with native origin and 921 with immigrant origin. Then parental levels of income are created, based on these individuals. For both origins the levels are at a much lower level in Brøndby and Ishøj, implying, that average income among parents in Ishøj and Brøndby are much lower, compared to the average incomes in the nation, while the levels of income in Kolding and Vejle are approximately equal to the national levels.

Brøndby and Ishøj[14]

2006

In the two municipalities with the highest proportion of immigrants in Denmark, 30.97% of the individuals with native origin have a high school education. This is a much lower fraction, than the overall proportion in Denmark, which was 46%. For individual with immigrant origin it is 27.62% with a high school education, which is five percentage lower compared to the overall proportion of immigrants with an education in Denmark. This implies, that native origin are worse of living in these municipalities than immigrant origins, measured in proportion with a high school education.

---

[14] For documentation of models see appendix E

Table 4.24: Data from Ishøj and Brøndby 2006

Gender	Graduated HS	Immigrant origin		Native origin	
	High School	#	Gender %	#	Gender %
Boys	No (0)	667	75.88%	1,426	73.85%
	Yes (1)	212	24.12%	505	26.15%
Girls	No (0)	520	68.33%	1,168	63.93%
	Yes (1)	241	31.67%	659	36.07%

Compared to the national proportion of individuals with a high school education, it is found that among the immigrant origin, the proportion of boys with an education is 3% points lower in the municipalities, while the fraction of girls is (6%) points lower than national proportion. For individuals of native origin it is found that 14% points fewer boys and 19% points fewer girls have a high school education in the municipalities, compared the a national proportions.

2016

In the two municipalities with the highest proportion of immigrants in Denmark, 37.67% of the individuals with native origin have a high school education. This is a much lower fraction, than seen as the average proportion for natives for Denmark. Among individuals with immigrant origin 47.52% have a high school education. This is higher than the average fraction of individuals with immigrant origin in Denmark, where 46.38% have an education.

Table 4.25: Data from Ishøj and Brøndby 2016

Gender	Graduated HS	Immigrant origin		Native origin	
	High School	#	Gender %	#	Gender %
Boys	No (0)	816	61.54%	1,338	68.2%
	Yes (1)	510	38.46%	624	31.8%
Girls	No (0)	531	42.79%	993	55.85%
	Yes (1)	710	57.21%	785	44.15%

On a national level the proportion of girls and boys with native origin, who have an education increased by 6% points from 2006 to 2016 and the proportion of girls with immigrant origin having an education increased by almost 17% points and 11% points for boys. In Brøndby and Ishøj, the development for boys with immigrant origin was higher than national level, as the proportion increased with 14% point. The proportion of girls with immigrant origin with a high school education increased even further, in total 26% points. While the individuals with immigrant origin have a development more impressive than what is seen as average in Denmark. The story for natives who live in Brøndby and Ishøj are less stunning. Here the development is much lower than seen as an average in the country. Exactly why children with immigrant origin are doing so much better in these municipalities this study can not tell. Whether it is factors in primary school, personal factors, the fact that proportion of immigrants living so close is very high providing them with a closer network, or that they do not meet that many cultural differences, or if is the fact, that native children are not much better in schools, and that immigrants therefore feel that they have a fair chance of "being the best". Or what it is, could be very interesting make further study on. For now, we will have to see, if the parental factors pay a different role.

Estimation of the models within the municipalities

2006

Estimating the models based on data for 2006, they are all found to be significant. For natives, the parental levels of income do not tell us much significantly about the probability of a high school education for the individuals (motherly level of income 4, and fatherly level of income 3 and 4 does), while parental levels of educations does in all categories. An increase in parental level of education and income is again found to have positive effect on the probability of education. Furthermore, it is found that the models classify approximately at the same levels as seen earlier on, and that non of them are skew. For individuals of immigrant origin the models are worse of, as few parental effects are significant. For boys, significant variables are motherly level of education $1-3$, her level of income if it is 2 or 4, or fatherly level of income, if it level 4, as these are the only significant at a 5% level. Implying no clear pattern in parental effects on the probability of education. For girls with immigrant origin, the same levels of motherly educations are found to be significant, furthermore are fatherly levels of education also significant. Only motherly level of income 4 is significant. This might be do to the small sample, but could also be caused by a week relationship between parental level of education and income and the probability of an individual with immigrant origin getting an education. At all levels of parental education and income, an increase is found to have positive effect on the probability of education.

2016

Estimating models based on the samples from 2016, they are all found to be significant. Furthermore it is found for the native origin individuals, that almost every level of parental education and income have a positive and significant effect on the probability of education. For both genders, a parental level of income at level 5 is not significant, and for girls neither are parental level of income equal to 2. The models both classify more or less as in benchmark. For boys with immigrant origin, all motherly level of education has significant and positive effect. Fathers with "trade" or high school as highest completed education has no significant effect, while all other levels of education among fathers has a positive significant effect. High levels of fatherly income increases the probability of education significantly, while mothers belonging in income group 2 or 4 increases the probability of education significantly. The probability of education among girls with immigrant origin increases significantly with an increase in: motherly level of education, a fatherly level of education equal to 3 or 5, a mother within the income groups 2 or 4. The model estimated for girls with immigrant origin has troubles with estimating girls who has education, as it mis-classify a high rate of girls as not having an education, when they actually has one, implying, that no absolute pattern with parental level of education and income and the probability of education is found.

The models used for stereotype families to predict probabilities of education - how do they differ?

The introductory studies in this robustness test showed, a very low proportion of high school educated native girls. The predicted probabilities are created to tell how parental effects play a role in that story. Investigating the predicted probability of education for girls, who live in Brøndby and Ishøj one find the prior probability of education for a girl with native origin is low. Very low compared to the benchmark case. Even though parental effects are modeled to have a great effect, as increase in especially motherly level of education has a very large effect on the probability of education. Again parental level of income is not as important as education, but fatherly level of income is found to be more important for the probability of education, than motherly income. Historically it is seen that motherly effects have greater impact on daughters' educational attainment, and fatherly effects are seen to be higher for sons (See Esping-Andersen (2004)), which is why it is a little surprising to find that fatherly level of education seem to have increased influence on the probability of education for native girls in 2016.

Table 4.26: Predicted probabilities of high school education for girls based on stereotype families living in Brøndby or Ishøj

Parental levels	"Stereotype Fam."	NA 06	NA 16	IM 06	IM16	Δ NA	Δ IM
Levels: 1	1 (prior prob.)	17.64%	22.22%	19.12%	41.12%	25.96%	115.06%
Levels: 2	2	30.25%	38.07%	44.78%	63.16%	25.85%	41.05%
Levels: 3	3	73.02%	76.51%	53.66%	65.63%	4.78%	22.31%
Levels: 4	4	79.49%	76.30%	66.31%	75.40%	-4.01%	13.71%
Levels: 5	5	80.26%	80.88%	52.61%	96.48%	0.77%	83.39%
Levels: 5/6	6	93.24%	91.17%	37.22%	96.47%	-2.22%	159.19%
Dad Edu: 3	7	30.63%	38.74%	35.81%	54.25%	26.48%	51.49%
Mum Edu: 3	8	45.09%	42.38%	30.18%	45.27%	-6.01%	50.00%
Mum Edu: 6	9	56.27%	46.23%	30.23%	58.62%	-17.84%	93.91%
Dad Edu: 6	10	37.22%	53.78%	23.59%	48.21%	44.49%	104.37%
Dad Inc: 4	11	28.66%	32.14%	19.70%	47.81%	12.14%	142.69%
Mum Inc: 4	12	24.82%	26.25%	33.79%	55.97%	5.76%	65.64%

For girls with immigrant origin, the development within the predicted probability is very high, not only in the prior probability, but at all stereotype families, the probability of education has increased from 2006 to 2016, since the proportion of girls with immigrant origin having an education had increased this is not that surprising. The high development in the prior probability tells us, that not only are more girls get an education, but the proportion of girls from families with very low resources has doubled, implying that more girls are more social mobile. Or it could indicate that the increased proportion with an education might not be explained with parental effects, but need to be explained by other factors. The relative changes for girls with immigrant origin are small compared to the relative changes in the probability for native girls. This is caused by the high prior probability for immigrant girls, which implies that girls are less sensitive to parental effects than in benchmark and that other factors could be more explaining when it comes to the probability of education.

Table 4.27: Relative changes in the predicted probability of education among girls living in Brøndby or Ishøj, when one parental level changes and all other parental levels are set to 1

"Stereotype Fam."	Changing effect"	Δ NA 06	Δ NA16	Δ IM 06	Δ IM16
7	Dad Edu: 3	73.64%	74.35%	87.29%	31.93%
8	Mum Edu: 3	155.61%	90.73%	57.85%	10.09%
9	Mum Edu: 6	218.99%	108.06%	58.11%	42.56%
10	Dad Edu: 6	111.00%	142.03%	23.38%	17.24%
11	Dad Inc: 4	62.47%	44.64%	3.03%	16.27%
12	Mum Inc: 4	40.70%	18.14%	76.73%	36.11%

Given the model estimations, the native girls are found to be very sensitive to the parental levels of education, when they do live in Brøndby or Ishøj. The exact reason why, are not told by these findings but need further studying. It may be explained by the facts, that their parents seem to parents with low resources, as average income and level of education for parents in these municipalities is much lower than the average income and education in benchmark. In low resource famlies Korupp et al. (2002) found that mothers mattered the most, which is not clear among natives here.

For native boys one finds a very low level of prior probability of education, in both years. In 2006 the parental effects in stereotype familiies $3-4$ are found to increase the probability more in 2006 than in 2016. The probability of education among boys with native origin, seem to be very sensitive to parents with very high resources in 2016, as this increases the probability of education much. It is found that parental levels of education is very important, while parental income seem to be just as important as in benchmark. Especially if parents are very well education.

Table 4.28: Predicted probabilities of high school education for boys based on stereotype families living in Brøndby or Ishøj

Parental levels	"Stereotype Fam."	NA 06	NA 16	IM 06	IM16	Δ NA	Δ IM
Levels: 1	1 (prior prob.)	12.39%	12.15%	10.65%	19.82%	-1.94%	86.10%
Levels: 2	2	15.64%	26.12%	34.64%	42.82%	67.01%	23.61%
Levels: 3	3	63.51%	55.17%	43.28%	48.21%	-13.13%	11.39%
Levels: 4	4	62.02%	61.95%	50.71%	75.15%	-0.11%	48.20%
Levels: 5	5	66.08%	84.85%	36.11%	90.28%	28.40%	150.01%
Levels: 5/6	6	80.31%	95.58%	39.21%	94.84%	19.01%	141.88%
Dad Edu: 3	7	26.84%	19.51%	17.74%	27.85%	-27.31%	56.99%
Mum Edu: 3	8	29.96%	27.63%	18.10%	30.16%	-7.78%	66.63%
Mum Edu: 6	9	43.95%	42.62%	14.29%	35.10%	-3.03%	145.63%
Dad Edu: 6	10	31.06%	31.25%	19.39%	43.87%	0.61%	126.25%
Dad Inc: 4	11	17.20%	19.16%	17.04%	33.36%	11.40%	95.77%
Mum Inc: 4	12	19.71%	19.52%	18.28%	31.33%	-0.96%	71.39%

For immigrant boys the impact of parental income and education has grown a lot over time. Both income and educational effects increase the probability of education. And as found in Esping-Andersen (2004) the fatherly effects have a greater impact on the probability of education among boys with immigrant origin. The finding that parental education is more important to probability of education among children than parental income, correspond to the findings in Erola et al. (2016) and in Bukodi and Goldthorpe (2012), who found that parental class and social status mattered, also found parental education as the strongest predictor of children's education in the UK.

Table 4.29: Relative changes in the predicted probability of education among boys living in Brøndby or Ishøj, when one parental level changes and all other parental levels are set to 1

"Stereotype Fam."	Changing effect	Δ NA 06	Δ NA16	Δ IM 06	Δ IM16
7	Dad Edu: 3	116.63%	60.58%	66.57%	40.51%
8	Mum Edu: 3	141.81%	127.41%	69.95%	52.17%
9	Mum Edu: 6	254.72%	250.78%	34.18%	77.09%
10	Dad Edu: 6	150.69%	157.20%	82.07%	121.34%
11	Dad Inc: 4	38.82%	57.70%	60.00%	68.31%
12	Mum Inc: 4	59.08%	60.66%	71.64%	58.07%

These findings of parental effect for native girls in 2006 and immigrant girls in 2016 correspond to the conclusion in Korupp et al. (2002), since they find that for parents with low resources, it was often motherly resources, which mattered for the children's educational attainment, with the use of data from the Netherlands, West Germany and the U.S.. Likewise findings are found by the use of Dutch surveys by

Buis (2013). He found that mother's education mattered more for children's attainment than the father's. Somehow fatherly influence on the probability of education for native girls has increased. Hence fathers' education and income have become more important for the girls.

One can conclude that natives are much more sensitive to parental effects when they live in Brøndby and Ishøj, meaning that a clear pattern in parental effects and the probability of education is fund. And is found to be very important. For immigrants I did not find a clear and strong relationship between parental level of education and income. Therefore there might be other factors explaining the relationship of educational attainment, this could for example be network.

Kolding and Vejle

2006

In the two municipalities the proportion of individuals with native origin who have a high school education is 42.20%, which is a lower fraction compared to the 47% as the overall proportion for native origin in 2006 in Denmark. For individual with immigrant origin it was 25.62%, compared to the overall proportion of immigrants with a high school education (32%) in 2006 in Denmark. This implies, proportion of individuals with a high school education is lower in the two municipalities.

Table 4.30: Data from Kolding and Vejle 2006

Gender	Graduated HS	Immigrant origin		Native origin	
	High School	#	Gender %	#	Gender %
Boys	No (0)	398	79.82%	4,245	67.27%
	Yes (1)	100	20.08%	2,065	32.73 %
Girls	No (0)	287	67.85%	3,065	48.36%
	Yes (1)	136	32.15%	3,273	51.64%

Compared to the national proportion of individuals with a high school education, is found to be at a lower level, especially the boys regardless of origin.

2016

In 2016 it were 42.93% of the individuals with native origin in the municipalities, who had a high school education. This is a much lower fraction, than seen as the overall proportion with an education in Denmark. Among individuals with immigrant origin 40.57% had a high school education, which is lower than what is found for individuals with immigrant origin on a national level, where 46.38% have an education. This implies that the development of the proportion of high school educated immigrants within the municipalities is greater than the development within the individuals with native origin.

Table 4.31: Data from Kolding and Vejle 2016

Gender	Graduated HS	Immigrant origin		Native origin	
	High School	#	Gender %	#	Gender %
Boys	No (0)	671	63.24%	6,105	65.77%
	Yes (1)	390	36.76%	3,178	34.23%
Girls	No (0)	527	55.18%	4,241	47.94%
	Yes (1)	428	44.82%	4,606	52.06%

On a national level the proportion of girls and boys with native origin who have an education increased by 6% points, from 2006 to 2016 and the proportion of girls with immigrant origin who have an education increased by almost 17% points and 11% points for boys. In Vejle and Kolding the development the proportion of girls with immigrant origin with a high school education only increased with 12% points, but the development for boys with immigrant origin was higher than national level, as the proportion increased with 16% point, i.e. a development higher than seen in Brøndby and Ishøj. The development for native origins are not impressive, as the proportion of boys and girls with an education only is increased by 1% for girls and 2% for boys.

A look at average parental income and education implies, that in 2006 the average education and income are found at a little lower in Kolding and Vejle compared to average for the native origin, while the average for parents with immigrant origin are found to have higher level of education in the municipalities, also motherly level of income are found at higher levels than in Denmark. Implying that the average family of immigrant origin have higher motherly income and higher average parental level of education in Kolding and Vejle. In 2016 the average parental level of education and income have not increased as much as the average in Denmark, regardless of origin. This implies that the average development in the municipalities lower, is the average development in Denmark, which the proportion of educated boys and girls also showed.

Estimation of the models[15]

2006

The models based on 2006 data are found for native to have parental education levels as significant factors among the explanatory variables. The models for natives classify at levels approximately what is seen earlier in the thesis. The models estimated for individuals with immigrant origin are found to be significant, but most parental levels are insignificant. Furthermore the models are very skew, as both of them wrongly predict a very high rate of individuals as having an high school education, when they actually has non. This implies, that even though a clear pattern in parental effects are found in the

---

[15] For documentation of findings see appendix F

data, some boys does not follow this pattern, indirectly implying that other factors than parental level of education and income might be very or more important.[16]

2016

Estimating the models based on individuals in Kolding and Vejle in 2016 the models are found to be significant. The models estimated on the native sample has significant parental levels of income and education, and increased parental levels of education or income increases the probability of education. The models estimated based on the immigrant origin are found to have most of the motherly levels of education as significant explanatory variables, but parental effects are found insignificant.[17] Furthermore is the skewness in the classification tables gone. Implying that there might be a pattern in the educational behavior among girls, but remember it is insignificant.

In the predicted probabilities of education for the girls based on their corresponding stereotype families, it is found that girls from families with low resources (stereotype family 1-2) her probability of education where higher in 2006, than in 2016, implying higher level of social mobility among native girls in Kolding and Vejle, than seen on national level in 2006. Moreover the findings for native girls are similar to the benchmark case, higher level of resources, higher probability of education. The girls from Kolding and Vejle, differ from the benchmark case in the dependency of fatherly levels of education and income are less important to the probability of education and motherly effects in 2016 are found to be more important.

---

[16] I tested, whether the parental effects became equal, if I instead estimated one model for boys and girls, this did not change anything.

[17] I tested, if estimating one joint model for boys and girls, instead of estimating two models, one for girls and one for boys, would make more parental effects significant. It would not.

Table 4.32: Predicted probabilities of high school education for girls based on stereotype families living in Kolding or Vejle

Parental levels	”Stereotype Fam.”	NA 06	NA 16	IM 06	IM16	Δ NA	Δ IM
Levels: 1	1 (prior prob.)	30.58%	24.05%	19.56%	25.32%	-21.35%	29.45%
Levels: 2	2	49.39%	45.24%	22.67%	51.59%	-8.40%	127.57%
Levels: 3	3	63.66%	71.51%	38.97%	59.47%	12.33%	52.60%
Levels: 4	4	79.06%	83.41%	61.38%	71.96%	5.50%	17.24%
Levels: 5	5	83.60%	91.73%	90.98%	52.95%	9.72%	-41.80%
Levels: 5/6	6	93.31%	95.93%	91.24%	70.86%	2.81%	-22.34%
Dad Edu: 3	7	40.71%	35.79%	15.64%	32.94%	-12.09%	110.61%
Mum Edu: 3	8	44.01%	41.05%	37.25%	42.98%	-6.73%	15.38%
Mum Edu: 6	9	71.04%	58.46%	29.82%	57.44%	-17.71%	92.62%
Dad Edu: 6	10	50.88%	41.08%	31.49%	37.06%	-19.26%	17.69%
Dad Inc: 4	11	39.75%	37.13%	22.44%	39.19%	-6.59%	74.64%
Mum Inc: 4	12	35.66%	38.25%	31.12%	28.15%	7.26%	-9.54%

As the proportion of immigrants in the municipalities is very low and the pattern effect of parental level of education and income on the probability of education are not very alike, most of the parental are found insignificant. But with that in mind, one can still conclude that motherly level of education seems to be the greatest effect on the probability of education among the girls. The reasons why this might be the scenario is commented on earlier on in the thesis.

Table 4.33: Relative changes in the predicted probability of education among girls living in Kolding or Vejle, when one parental level changes and all other parental levels are set to 1

”Girl”	Changing effect”	Δ NA 06	Δ NA16	Δ IM 06	Δ IM16
7	Dad Edu: 3	33.13%	48.81%	-20.04%	30.09%
8	Mum Edu: 3	43.92%	70.69%	90.44%	69.75%
9	Mum Edu: 6	132.31%	143.08%	52.45%	126.86%
10	Dad Edu: 6	66.38%	70.81%	60.99%	46.37%
11	Dad Inc: 4	29.99%	54.39%	14.72%	54.78%
12	Mum Inc: 4	16.61%	59.04%	59.10%	11.18%

In general for the boys in the two municipalities in Jutland the prior probability of education is lower. In 2006 fatherly levels seemed to be the most important factors in the probability of education among the native boys. Similar to the findings that father' education are more important for sons, see Esping-Andersen (2004). The development over the years only changed that with the fact, that a mother with five or more years of education has become more important on the probability of education among boys. Implying that boys have become more sensitive. It could be interesting to study if divorces in the municipalities have anything to do with this.

Table 4.34: Predicted probabilities of high school education for boys based on stereotype families living in Kolding or Vejle

	"Stereotype Fam."	NA 06	NA 16	IM 06	IM16	Δ NA	Δ IM
Levels: 1	1	15.28%	13.19%	17.05%	22.36%	-13.68%	31.14%
Levels: 2	2	26.47%	25.10%	15.22%	37.08%	-5.18%	143.63%
Levels: 3	3	52.99%	57.81%	18.30%	39.36%	9.10%	115.08%
Levels: 4	4	69.10%	55.10%	41.83%	73.37%	-20.26%	75.40%
Levels: 5	5	84.52%	81.42%	91.50%	87.82%	-3.67%	-4.02%
Levels: 5/6	6	92.73%	88.17%	91.00%	91.33%	-4.92%	0.36%
Dad Edu: 3	7	35.71%	28.59%	14.20%	23.95%	-19.94%	68.66%
Mum Edu: 3	8	22.06%	23.22%	25.05%	28.89%	5.26%	15.33%
Mum Edu: 6	9	36.36%	34.64%	29.24%	34.96%	-4.73%	19.56%
Dad Edu: 6	10	39.97%	31.06%	12.37%	30.16%	-22.29%	143.82%
Dad Inc: 4	11	22.30%	20.20%	17.31%	32.45%	-9.42%	87.46%
Mum Inc: 4	12	20.68%	20.94%	21.97%	26.22%	1.26%	19.34%

Among boys with immigrant origin, the results imply that fathers with a high school education or a father with 5+ years of education decreases the probability of education in 2006 ,and in 2016 increases the probability of education a lot, but as the factors are not significant the results are not to comment on.

Table 4.35: Relative changes in the predicted probability of education among boys living in Kolding or Vejle, when one parental level changes and all other parental levels are set to 1

"Boy"	Changing effect"	Δ NA 06	Δ NA16	Δ IM 06	Δ IM16
7	Dad Edu: 3	133.70%	116.76%	-16.72%	7.11%
8	Mum Edu: 3	44.37%	76.04%	46.92%	29.20%
9	Mum Edu: 6	137.96%	162.62%	71.50%	56.35%
10	Dad Edu:6	161.58%	135.48%	-27.45%	34.88%
11	Dad Inc: 4	45.94%	53.15%	1.52%	45.13%
12	Mum Inc: 4	35.34%	58.76%	28.86%	17.26%

Nothing dramatic is found on a significant level in the two municipalities in Jutland. The parental effects within the immigrant origin, we saw for girls that a well educated mother increases the probability of education a lot. Further comments on parental effects among immigrants are omitted, as most of the results are insignificant. Among natives it is found, the development of parental effects on the probability of education is positive, meaning that parental factors have become more important when it comes to high school attainment for their children. This is unlike the findings for native in the national level.

Therefore the model is not very robust to the levels in specific municipalities. But the overall conclusions such as natives being more sensitive to parental level of education and income, than immigrants and further that boys are more sensitive than girls, seem to be more or less the same.

4.3.4 Distinguish between immigrants home country in Western /non-Western home country

This section will provide a review of the differences in parental impact on the probability of education, given the immigrants are split into two groups (Western/non-Western) based on country of origin. Given the great degree of detains that Statistic Denmark offer, it is also possible to investigate how parents 'effects may differ from immigrants' country of origin. Immigrants' country of origin, will in this section be distinguished as non-western and western[18] and a closer look at how, if any, differences' in parental level of income and education might appear on the probability of high school education.

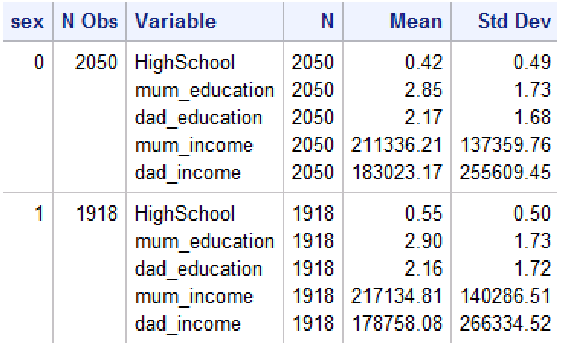
In 2006 it is found that 10.92% of the individuals with immigrant origin came from a Western country, while it was 10.13% percentage in 2016, i.e. the composition of Western and non-Western immigrants is almost unchanged over the years. Setting up the parental levels of income in both years, it is found that more than 25% of the individuals with Western origin have fathers belonging in income group one has a zero income. This might be due to the high proportion of individuals with western origin, who have no father registered in Denmark. The parental levels of income and education are also found to differ from each other. The income and education levels among parents with Western origin are higher, than among non-Western parents, which corresponds well to the findings in the report *Indvandrere i Danmark 2016*[19].

---

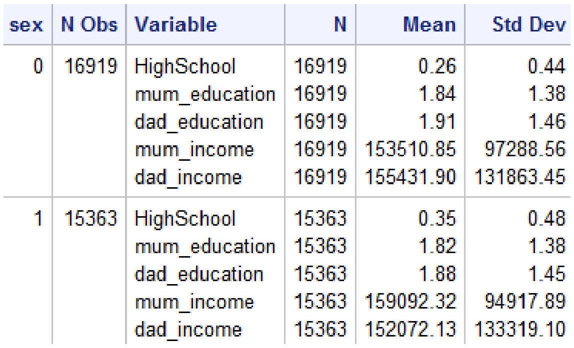
[18] Inspired by the composition of data variable IELANDG3, which is Immigrants/descendants broken down by country of origin by categories Western and Non-Western countries, from Statistic Denmarks. A variable is created to provide information on Western/Non-Western based on OPR_Land. Source: http://www.dst.dk/extranet/staticsites/TIMES3/html/2dfe91de-bd15-4fc9-b18c-b6166d20e745.htm

[19] [52]

Figure 4.7: Left Table: Western origin, Righ Table: non-Western origin, summary from 2006. Gender: Boys 0, Girls 1.

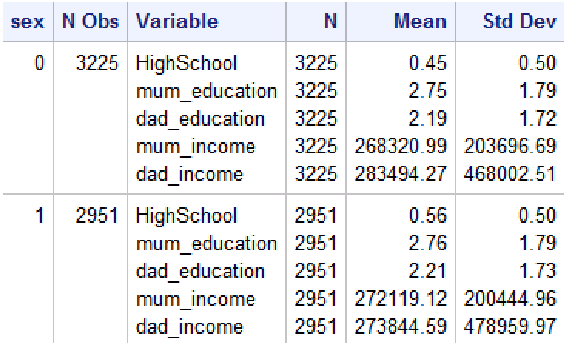


sex	N Obs	Variable	N	Mean	Std Dev
0	2050	HighSchool	2050	0.42	0.49
		mum_education	2050	2.85	1.73
		dad_education	2050	2.17	1.68
		mum_income	2050	211336.21	137359.76
		dad_income	2050	183023.17	255609.45
1	1918	HighSchool	1918	0.55	0.50
		mum_education	1918	2.90	1.73
		dad_education	1918	2.16	1.72
		mum_income	1918	217134.81	140286.51
		dad_income	1918	178758.08	266334.52

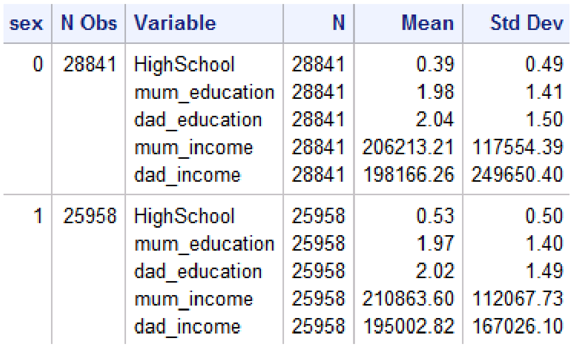


sex	N Obs	Variable	N	Mean	Std Dev
0	16919	HighSchool	16919	0.26	0.44
		mum_education	16919	1.84	1.38
		dad_education	16919	1.91	1.46
		mum_income	16919	153510.85	97288.56
		dad_income	16919	155431.90	131863.45
1	15363	HighSchool	15363	0.35	0.48
		mum_education	15363	1.82	1.38
		dad_education	15363	1.88	1.45
		mum_income	15363	159092.32	94917.89
		dad_income	15363	152072.13	133319.10

Figure 4.8: Left Table: Western origin, Righ Table: non-Western origin, summary from 2016. Gender: Boys 0, Girls 1.



sex	N Obs	Variable	N	Mean	Std Dev
0	3225	HighSchool	3225	0.45	0.50
		mum_education	3225	2.75	1.79
		dad_education	3225	2.19	1.72
		mum_income	3225	268320.99	203696.69
		dad_income	3225	283494.27	468002.51
1	2951	HighSchool	2951	0.56	0.50
		mum_education	2951	2.76	1.79
		dad_education	2951	2.21	1.73
		mum_income	2951	272119.12	200444.96
		dad_income	2951	273844.59	478959.97



sex	N Obs	Variable	N	Mean	Std Dev
0	28841	HighSchool	28841	0.39	0.49
		mum_education	28841	1.98	1.41
		dad_education	28841	2.04	1.50
		mum_income	28841	206213.21	117554.39
		dad_income	28841	198166.26	249650.40
1	25958	HighSchool	25958	0.53	0.50
		mum_education	25958	1.97	1.40
		dad_education	25958	2.02	1.49
		mum_income	25958	210863.60	112067.73
		dad_income	25958	195002.82	167026.10

It is found that among immigrants of non-Western origins the fatherly level of education in average are highest, while it is opposite for Western origins. Which is comparable to the findings that there are many more native women than native men with short-, medium- or higher education, whereas there are more men than women with such education among immigrants with non-Western origin (Bonke and Schultz-Nielsen (2012)). When distinguish between whether immigrants country of origin is Western or non-Western a review of the proportion of individuals with a high school education shows clear differences as 48.23% of individuals immigrated from a Western country has a high school education, while it is only 30.22% of individuals immigrated from a non-Western country who have an education. A closer look that the differences in gender

Table 4.36: High School educated individuals. Western/Non-Western origin

		2006				2016			
	Graduated	Western		Non-Western		Western		Non-Western	
	HS	#	Gender %	#	Gender %	#	Gender %	#	Gender %
Boys	No (0)	1,182	57.66%	12,543	74.14%	1,780	55.19%	17,454	60.66%
	Yes (1)	868	42.34%	4376	25.86 %	1,182	44.81%	11,346	39.34%
Girls	No (0)	872	45.46%	10,013	65.18%	1,294	43.85%	12,124	46.71%
	Yes (1)	1,046	54.54%	5,350	35.82%	1,657	56.15%	13,834	53.29%

shows that a higher proportion of boys with western origin is getting a high school education compared to the native origin boys in the full dataset in 2006. In 2016 it is found that 50.23% of individuals with a western origin have an education while it is 45.95% of the individuals with non-western origin. The development among the proportion of individuals with western origin, seems to be much lower, than the very large increased proportion of individuals with non-western origin.

Estimating the models for individuals with Western/non-Western origin

Estimating the models for predicting the probability of education among individuals from the stereotype families, all the models are found to be significant. Most of parental levels of education and income are found to have positive significant effects on the probability of education, as we have seen before and furthermore are the classification rates of the models approximately without skewness in its classifications and what is seen earlier. The predicted probabilities of education among girls with Western origin, provide the information that the prior probabilities of education are found at higher levels in in 2006 than for immigrants in the benchmark case, which might be implying that girls with Western origin are more social mobile in 2006 than they are found to be in 2016. Furthermore it is found that are parental of the groups with highest resources, the probability of education has decreased over time but the overall predicted probability is higher than found for immigrants girls, which also fits the fact that the proportion of immigrants with western origin and a high school education is higher, than the overall proportion of immigrants with an education. Again motherly effects seems to have a greater impact on probability of education most, and moreover all parental levels seem to be at higher effects compared to the relative changes we saw among the overall immigrant girls, in the benchmark scenario, like Esping-Andersen (2004).

Table 4.37: Predicted probabilities of high school education for girls based on stereotype families

Parental Lev.	"Stereotype Fam."	WE 06	WE 16	non-WE06	non-WE16	ΔWE	Δnon-WE
Levels: 1	1 (prior prob.)	31.31%	32.51%	21.25%	35.63%	3.83%	67.67%
Levels: 2	2	40.70%	41.30%	35.83%	55.82%	1.47%	55.79%
Levels: 3	3	73.78%	69.43%	43.23%	57.77%	-5.90%	33.63%
Levels: 4	4	75.78%	81.01%	62.99%	77.25%	6.90%	22.64%
Levels: 5	5	97.37%	88.45%	82.87%	89.98%	-9.16%	8.58%
Levels: 5/6	6	98.20%	91.60%	73.67%	91.14%	-6.72%	23.71%
Dad Edu: 3	7	49.81%	37.13%	26.33%	40.56%	-25.46%	54.04%
Mum Edu: 3	8	41.52%	53.64%	29.34%	45.92%	29.19%	56.51%
Mum Edu: 6	9	53.50%	62.71%	29.52%	54.35%	17.21%	84.11%
Dad Edu: 6	10	53.55%	59.55%	25.97%	45.61%	11.20%	75.63%
Dad Inc: 4	11	45.96%	42.54%	30.26%	50.61%	-7.44%	67.25%
Mum Inc: 4	12	44.42%	48.96%	31.29%	47.00%	10.22%	50.21%

Note surprisingly, the predicted probabilities for immigrants with non-Western are more or less equal to the findings for the overall immigrants, but there still are some remarkable notes to make. The development of parental effect over time has increased, but the parental effects are relative less important to the probability of education, an increase in fatherly level of income from 1 to 4 (significant levels) has increased effect on the probability of education among the girls with non-Western origin, hence fatherly level of income seems to be more important than motherly level of income.

Table 4.38: Relative changes in the predicted probability of education among girls with immigrant origin when all, but one, parental levels are set to 1

"Stereotype Fam."	Changing effect	Δ WE 06	Δ WE 16	ΔNon-WE 06	Δ Non-WE16
7	Dad Edu: 3	59.09%	14.21%	23.91%	13.84%
8	Mum Edu: 3	32.61%	65.00%	38.07%	28.88%
9	Mum Edu: 6	70.87%	92.89%	38.92%	52.54%
10	Dad Edu: 6	71.03%	83.17%	22.21%	28.01%
11	Dad Inc: 4	46.79%	30.85%	42.40%	42.04%
12	Mum Inc: 4	41.87%	50.60%	47.25%	31.91%

In the predicted probabilities of education among boys from Western origin, the development of probability of education over time, are lower compared to the overall development for boys of immigrant origin. This implies, that both boys and girls with Western origin have increased the probability of education over time less than the development of the probability of education among immigrants regardless of stereotype family. This indirectly implies that immigrants of non-Western origin have increased over time, as we see in the following table and have seen earlier. Furthermore one sees that, especially mother's education and income have development in the stereotype families 7-12 implies. And the relative changes tells that parental effect are very important to immigrant boys with Western origin, and fatherly education and income effects on the probability of education were greatest in 2006, but this has changed in 2016 where motherly effects changes the probability of education relative more than fatherly effects, which is the opposite of Esping-Andersen (2004). This should not correspond to the fact that many of the Western origin has no father registered in Denmark.

Table 4.39: Predicted probabilities of high school education for boys based on stereotype families

Parental Levels	"Stereotype Fam."	WE 06	WE 16	non-WE06	non-WE16	ΔWE	Δnon-WE
Levels: 1	1 (prior prob.)	18.30%	22.19%	15.85%	23.20%	21.26%	46.37%
Levels: 2	2	34.16%	37.26%	23.61%	42.44%	9.07%	79.75%
Levels: 3	3	59.65%	57.58%	33.80%	40.40%	-3.47%	19.53%
Levels: 4	4	66.09%	69.73%	52.34%	67.29%	5.51%	28.56%
Levels: 5	5	66.51%	70.22%	81.61%	85.40%	5.58%	4.64%
Levels: 5/6	6	66.94%	85.03%	77.43%	84.41%	27.02%	9.01%
Dad Edu: 3	7	22.16%	34.63%	22.24%	28.26%	56.27%	27.07%
Mum Edu: 3	8	30.40%	34.53%	21.97%	29.52%	13.59%	34.37%
Mum Edu: 6	9	39.96%	51.42%	22.12%	35.55%	28.68%	60.71%
Dad Edu: 6	10	33.40%	40.21%	25.16%	37.03%	20.39%	47.18%
Dad Edu: 4	11	31.67%	31.66%	23.46%	33.77%	-0.03%	43.95%
Mum Inc: 4	12	32.37%	32.42%	21.41%	32.34%	0.15%	51.05%

For immigrant boys in Denmark of non-Western origin, the historical delvelopment of the predicted probability found very much higher in the low and average income stereotype families, when comparing to the benchmark case. This implies, that immigrant boys with non-Western origin in the families have increased probability of getting an education. Especially a boy from the lowest stereotype family has increased the probability of education over time. From the stereotype families $7 - 12$, the development over time also exceeds what is seen for immigrants in the benchmark scenario, implying that for non-Western immigrants, the impact of parental effects are higher for boys of non-Western origin, than for the overall immigrant population. For boys with non-Western origin, fatherly effects are found to have greater effect on the probability of education.

Table 4.40: Relative changes in the predicted probability of education among boys with immigrant origin when all, but one, parental levels are set to 1

"Family"	Changing effect	Δ WE 06	Δ WE 16	Δnon-WE 06	Δ non-WE16
7	Dad Edu: 3	21.09%	56.06%	40.32%	21.81%
8	Mum Edu: 3	66.12%	55.61%	38.61%	27.24%
9	Mum Edu: 6	118.36%	131.73%	39.56%	53.23%
10	Dad Edu: 6	82.51%	81.21%	58.74%	59.61%
11	Dad Inc: 4	73.06%	42.68%	48.01%	45.56%
12	Mum Inc: 4	76.89%	46.10%	35.08%	39.40%

The probability of education for immigrants of Western origin motherly characteristics are found the most important factor when it comes to the fact of predicting the probability of education, as her level of income and education leads to a higher increase in probability of education.

Comparing parental effect across country of origin, for boys and girls with Western origin the connection between the probability and parental factors are found stronger, meaning that they are more sensitive to parental changes. Indirectly this also leads to the fact that for individuals in Denmark of non-Western origin, the connection of parental factors and the probability of education are less strong, meaning that one can not make conclusions about the probability of education given parental levels of education and income with the same certainty, as for individuals of Western origin.

4.4 Perspective

The results of all these data leave me with many more topics that could be exciting to examine in this field of studying. A natural extension of this thesis would be to dig deeper into why the influence of immigrant parents in general is lower when it comes to the probability of education, compared to the influence of native parents. Is the reason a difference in culture, network or norms? It could be interesting to analyze: the stepparents' influence on the probability of education among their stepchildren or if parental joint income would tell a different story, or to see if the results of a robustness test testing the connection between an average parents' income and the probability of their children getting a high school education. Another interesting topic that comes up is when looking at the results from two municipalities with a high number of immigrants, Brøndby and Ishøj, is why children of immigrant origin here do relatively much better measured by high school education, than children with native origin. And finally it could be interesting to see, if the results would change if the analysis are based on just 2 or 5 years instead of ten years?

Chapter 5

Conclusion

In this thesis I have studied to what extent parental income and level of education effect the probability of education among their children. Furthermore, I have studied the differences appearing in the parental effects on the probability of education among children of native and immigrant origin in Denmark based on data from 2006 and 2016. In the educational data from 2006, the children are born between $1976-1986$ and in the education data from 2016, the children are born between $1986-1996$. When the children were connected with their parents through family links and the acknowledgment of parental income and level of education, too, the data was ready for analysis. In all the analysis made in this thesis, I found significant relationship between parental impacts and the probability of education among their children.

In the benchmark scenario, which is the full datasets only excluding children with parents having a negative income the proportion of native children with a high school education is higher than the proportion of immigrant children having a high school education. For children with native origin, I found a very strong relationship between parental level of education and the probability of education and a less strong relationship between parental income and the probability of education. But for children with immigrant origin the influence of parental income is almost as high, as the influence of parental education, and therefore differs from the conclusion among natives. Moreover I found that influence of parents' income and level of education on the probability of education are higher among boys than girls and also much higher among children of native origin than among children of immigrant origin. This might imply that other factors, such as norm, culture or network could be more important for immigrants, when it comes to education attainment. Furthermore I concluded, that the development of parental influence on the probability of a high school education has increased for girls and boys with immigrant origin. As the influence of parent's education on the probability of education has decreased over time, the influence of parental income on the probability of education among natives has increased over time, regardless of gender.

Removing individuals without both parents registered in Denmark, I found no clear changes in the parental effects on the probability of education among the children of native origin, but for children with immigrant origin it increased the probability of education. And for girls with immigrant origin, the parental effect becomes relatively less important than in benchmark.

When the analysis distinguishes between whether or not parents live together, clear differences in the proportion of children with a high school education were found: the proportion of children with a high school education is much higher when parents live together. This fits the literature stating that parental divorce has negative consequences for the attainment of children's education and the facts that parents who live together have higher levels of education and income.[1] Moreover, I found that the parent effect was much more important among children whose parents did *not* live together in 2006, but in 2016 the relationship among parental effects and the probability of education became less strong. Whether the less strong relationship between parental effects and the probability of education became less clear in 2016, is caused because native girls have become more social mobile or the fact that other factors such as stepparents' effects influence the probability of educational attainment or simply just because the pattern in the parental effects have become less clear, could be interesting to do further studies on. Girls with immigrant origin, seem to be more sensitive to motherly effects than in benchmark. For boys with native origin the story is comparable to the story for native girls: when parents live apart they are found to be very sensitive to parental effects in 2006 and less sensitive to parental effects in 2016. Whether this is caused by an unclear pattern of parental effects on the probability of education, or the fact that children with native origin with parents who do not live together have become more social mobile is not clear. Beside the lower prior probability of education and the higher effect of motherly level of education, the story for boys with immigrant origin did not change much compared to benchmark. The fact is that parents getting divorced/no longer living together especially effects boys. Girls are also effected by this though at a lower level.

When parents live together, the influence of parent's effect has increased over time and for natives the father's influence becomes more important regardless of gender, pointing to the fact, that when parents live together the father becomes more influential. Furthermore when parents live together native girls from low resource families are found to have higher probability of education, compared to benchmark. For native boys from low resource families I found, that over time they have become less sensitive to the fact that parents live together. In 2006 parents living together increased the probability of education among boys much more, as when he came from a low resource family. The fact that parents living together was only found to be important for boys with immigrant origin from low resource families. Even though I did not find any significant differences in parental effects among boys and girls of immigrant origin. The insignificant findings point to that boys being more sensitive to the fact that parents live together.

The parental effects also seem to differ within municipalities. In Kolding and Vejle, the parental level of education is found to be very important and gender specific: motherly effects clearly have higher influence on the probability of education among girls while fathers' effect are most important for boys. In the analysis from Brøndby and Ishøj, the two municipalities with the highest proportion of immigrants in Denmark, I found interesting results: the fraction of educated natives was lower than the fraction of educated immigrants in 2016. In these municipalities the parental income and education is found to be

---

[1] Keith and Finlay (1988)

even more influential, and for natives it is found that especially motherly level of education is important, regardless of gender. For girls with immigrant origin living in Brøndby or Ishøj, parental effects do not seem to be very influential though boys are more influenced on parental effects than in the other scenarios. This implies that for girls with immigrant origin in Brøndby and Ishøj there might be other factors, such as network, culture or norms influencing the educational attainment, as the results point to that they get a high school education regardless of parental income and education. Furthermore no signs in the average of parental income or level of education are to imply why immigrants do so much better in these municipalities than at a national level, so the reasons are still unknown.

The differences in the investigations made in Western origin/non-Western origin, the Western origin are found to be influenced by parents effect approximately as much as native origins are. Higher motherly level of education increases the probability of education for both girls and boys more than seen in benchmark. The findings among non-Western origin are similar to the benchmark case for immigrants, intuitively this makes sense as almost 90% of the immigrants in Denmark are non-Western. For non-Western origin, the effect of motherly level of education has increased mostly over time.

In general, most of the findings in this thesis seem to fit the literature: the attainment of children's education is highly correlated to parental level of education. Motherly effects are most important among girls, while the influence of parents seems to be more equal among boys, however, the father's effects have a slightly higher influence and furthermore I have found clear signs of higher fraction of children with a high school education when parents live together, than when they do not. I have found that a higher proportion of natives has a high school education compared to immigrants, and furthermore that natives are more influenced by parental effects than immigrants. Implying that other factors than parental level of education and income might be more important for immigrants when it comes to the fact of choosing high school or not.

References

[1] Arnbak, E. and Bremholm, J. (2016). Pisa 2015: Danske unge i en international sammenligning.

[2] Bailey, M. J. and Dynarski, S. M. (2011). Gains and gaps: Changing inequality in us college entry and completion. Technical report, National Bureau of Economic Research.

[3] Barr, N. A. (2001). *The welfare state as piggy bank: information, risk, uncertainty, and the role of the state*. Oxford University Press.

[4] Becker, G. S. and Tomes, N. (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of political Economy*, 87(6):1153–1189.

[5] Bennett, P., la Cour, L., Larsen, B., and Waisman, G. (2017). Negative attitudes, network and education.

[6] Black, S. E. and Devereux, P. J. (2010). Recent developments in intergenerational mobility. Technical report, National Bureau of Economic Research.

[7] Bonke, J. and Schultz-Nielsen, M. L. (2013). *Integration blandt ikke-vestlige indvandrere-Arbejde, familie, netværk og forbrug*. Syddansk Universitet. Institut for Grnseregionsforskning. Syddansk Universitetsforlag.

[8] Buis, M. L. (2012). The composition of family background: The influence of the economic and cultural resources of both parents on the offspring's educational attainment in the netherlands between 1939 and 1991. *European Sociological Review*, 29(3):593–602.

[9] Bukodi, E. and Goldthorpe, J. H. (2012). Decomposing 'social origins': The effects of parents' class, status, and education on the educational attainment of their children. *European Sociological Review*, 29(5):1024–1039.

[10] Calvin, J. A. (1998). Regression models for categorical and limited dependent variables.

[11] Carneiro, P. and Heckman, J. J. (2002). The evidence on credit constraints in post-secondary schooling. *The Economic Journal*, 112(482):705–734.

[12] Chambers, E. A. and Cox, D. R. (1967). Discrimination between alternative binary response models. *Biometrika*, 54(3-4):573–578.

[13] Chen, G. and Tsurumi, H. (2010). Probit and logit model selection. *Communications in Statistics—Theory and Methods*, 40(1):159–175.

[14] Christensen, V. T., Egelund, N., Fredslund, E. K., and Jensen, T. P. (2014). Pisa etnisk 2012: Pisa 2012 med fokus på unge med indvandrerbaggrund.

[15] Corak, M. (2003). Income mobility between generations.

[16] Cox, D. R. (1966). Some procedures associated with the logistic qualitative response curve.

[17] Dämmrich, J. and Esping-Andersen, G. (2017). 7. preschool and reading competencies–a cross-national analysis. *Childcare, Early Education and Social Inequality: An International Perspective*, page 133.

[18] De Philippis, M. and Rossi, F. (2016). Parents, schools and human capital differences across countries.

[19] Elias, S. V. and Ramsløv, K. F. (2014). Gymnasiet taber drengene.

[20] Emery, R. E. (1982). Interparental conflict and the children of discord and divorce. *Psychological bulletin*, 92(2):310.

[21] Erikson, R. and Goldthorpe, J. H. (1992). *The constant flux: A study of class mobility in industrial societies*. Oxford University Press, USA.

[22] Erola, J., Jalonen, S., and Lehti, H. (2016). Parental education, class and income over early life course and children's achievement. *Research in Social Stratification and Mobility*, 44:33–43.

[23] Esping-Andersen, G. (2004). Untying the gordian knot of social inheritance. *Research in social stratification and mobility*, 21:115–138.

[24] Esping-Andersen, G. (2015). Welfare regimes and social stratification. *Journal of European Social Policy*, 25(1):124–134.

[25] Esping-Andersen, G. (2016). Families in the 21st century.

[26] Fischer, J. A. (2009). The welfare effects of social mobility: An analysis for oecd countries.

[27] Glick, P. C. and Norton, A. J. (1977). Marrying, divorcing, and living together in the us today. *Population Bulletin*, 32(5):1.

[28] Greve, J. and Krassel, K. F. (2017). Pisa etnisk 2015.

[29] Gurria, A. (2016). Pisa 2015 results in focus. *PISA in Focus*, (67):1.

[30] Heckman, J. J., Krueger, A. B., and Friedman, B. (2004). Inequality in america. *Cambridge (Mass.).*

[31] Heckman, J. J. and Lochner, L. (2000). Rethinking education and training policy: Understanding the sources of skill formation in a modern economy. *Securing the future: Investing in children from birth to college*, pages 47–83.

[32] Holm, A., Jæger, M. M., Karlson, K. B., and Reimer, D. (2013). Incomplete equalization: The effect of tracking in secondary education on educational inequality. *Social science research*, 42(6):1431–1442.

[33] Jacobsen, K. F. (2004). *Befolkningens uddannelsesniveau.* Danmarks Statistik.

[34] Jæger, M. M. and Holm, A. (2007). Does parents' economic, cultural, and social capital explain the social class effect on educational attainment in the scandinavian mobility regime? *Social Science Research*, 36(2):719–744.

[35] Jantti, M., Bratsberg, B., Roed, K., Raaum, O., Naylor, R., Osterbacka, E., Bjorklund, A., and Eriksson, T. (2006). American exceptionalism in a new light: a comparison of intergenerational earnings mobility in the nordic countries, the united kingdom and the united states.

[36] Jespersen, H. (2016). Integration: Status og udvikling 2016.

[37] Juul, J. S. (2016). Den sociale arv er blevet stærkere i danmark.

[38] Keith, V. M. and Finlay, B. (1988). The impact of parental divorce on children's educational attainment, marital timing, and likelihood of divorce. *Journal of Marriage and the Family*, pages 797–809.

[39] Korupp, S. E., Ganzeboom, H. B., and Van Der Lippe, T. (2002). Do mothers matter? a comparison of models of the influence of mothers' and fathers' educational and occupational status on children's educational attainment. *Quality & Quantity*, 36(1):17–42.

[40] Landersø, R. and Heckman, J. J. (2017). The scandinavian fantasy: The sources of intergenerational mobility in denmark and the us. *The Scandinavian Journal of Economics*, 119(1):178–230.

[41] Lillard, L. A. and Willis, R. J. (1994). Intergenerational educational mobility: Effects of family and state in malaysia. *Journal of Human Resources*, pages 1126–1166.

[42] McLanahan, S. and Sandefur, G. (1994). *Growing Up with a Single Parent. What Hurts, What Helps.* ERIC.

[43] Milhøj, A. (1998). Maksimum likelihood estimation. In *Den Store Danske Encyklopædi.* Gyldendal.

[44] Nelson, R. R. and Phelps, E. S. (1966). Investment in humans, technological diffusion, and economic growth. *The American economic review*, 56(1/2):69–75.

[45] Ottesen, M. H., Andersen, D., Dahl, K. M., Hansen, A. T., and Œstergaard, S. V. (2014). Boern og unge i danmark - velfærd og trivsel 2014.

[46] Pihl, M. D. (2017a). Børn af enlige på kontanthjælp klarer sig dårligere i skolen.

[47] Pihl, M. D. (2017b). Danskernes uddannelse: Flere får en uddannelse, men faglærte taber terræn. *AE*.

[48] Piketty, T. (2014). Capital in the 21st century.

[49] Ploug, N. and (red.) (2017). *Økonomisk ulighed i Danmark*. Jurist- og Økonomforbundets Forlag.

[50] Rosdahl, A., Fridberg, T., Jakobsen, V., and Jørgensen, M. (2013). *Færdigheder i læsning, regning og problemløsning med IT i Danmark*. SFI-Det Nationale Forskningscenter for Velfærd.

[51] Skyt Nielsen, H. and Schindler Rangvid, B. (2011). The impact of parents' years since migration on children's academic achievement. Technical report, Discussion Paper series, Forschungsinstitut zur Zukunft der Arbeit.

[52] Statistik, D. Indvandrere i danmark 2016. 2016.

[53] Statistik, Danmark, y. p. Indkomster 2014. tema: Formuer.

[54] Thomsen, J.-P. and Andrade, S. B. (2016). Uddannelsesmobilitet i danmark.

[55] Verbeek, M. (2004). A guide to modern econometrics. southern gate, chichester, west sussex, england hoboken.

[56] Wackerly, D., Mendenhall, W., and Scheaffer, R. (2007). *Mathematical statistics with applications*. Nelson Education.

[57] Waldfogel, J. (2006). What do children need? *Juncture*, 13(1):26–34.

[58] Weiss, R. S. (1979). Growing up a little faster: The experience of growing up in a single-parent household. *Journal of Social Issues*, 35(4):97–111.

[59] Willis, R. J. and Rosen, S. (1979). Education and self-selection. *Journal of political Economy*, 87(5, Part 2):S7–S36.

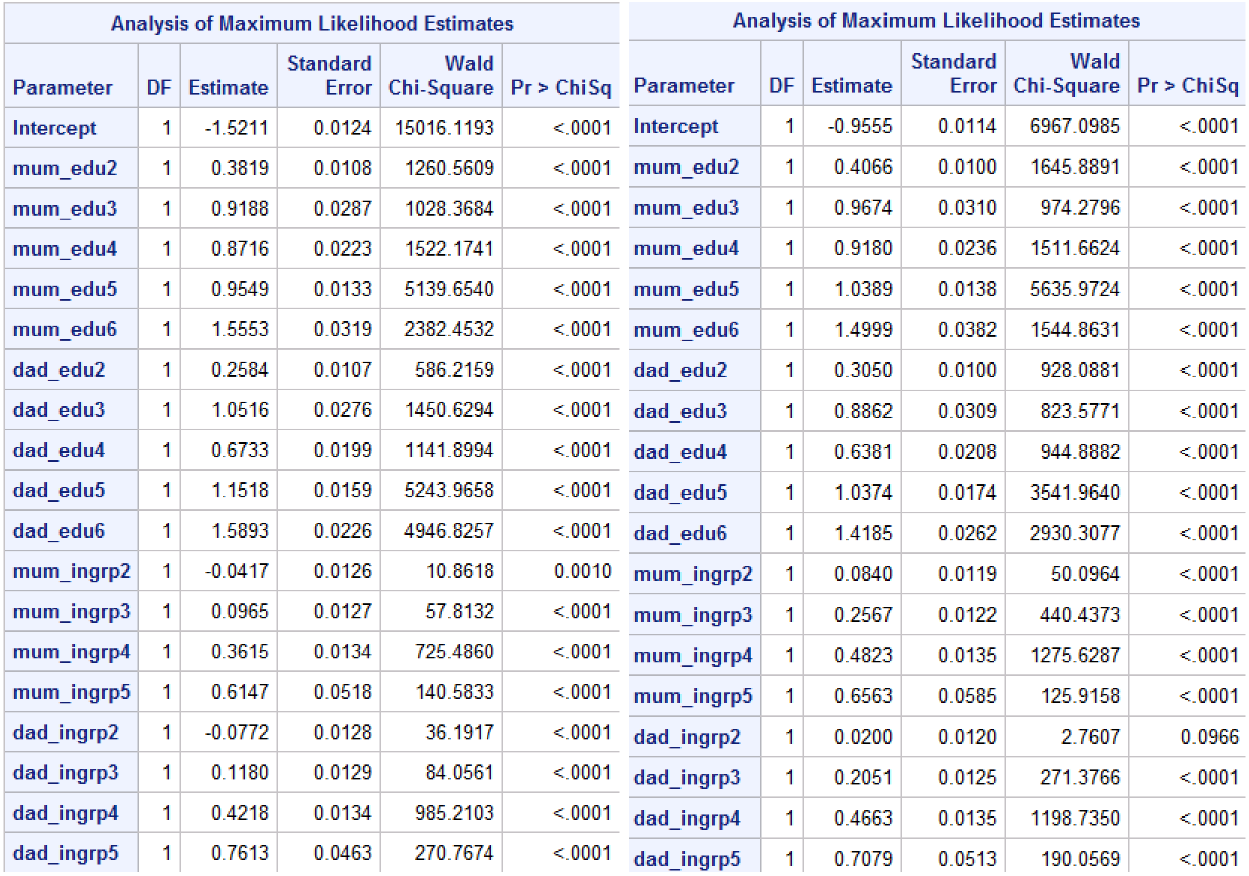
[60] Økonomiske Råd, D. (2016). Dansk økonomi efteråret 2016.

Appendix A

Benchmark

2006

Figure A.1: Left Table: Boys, Right Table: Girls. Native origin, data form 2006



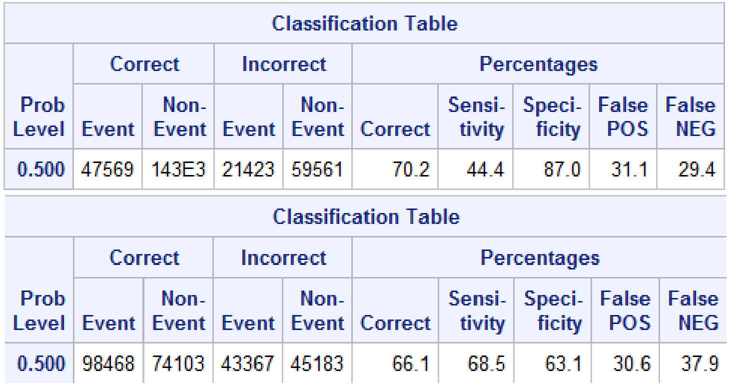
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.5211	0.0124	15016.1193	<.0001
mum_edu2	1	0.3819	0.0108	1260.5609	<.0001
mum_edu3	1	0.9188	0.0287	1028.3684	<.0001
mum_edu4	1	0.8716	0.0223	1522.1741	<.0001
mum_edu5	1	0.9549	0.0133	5139.6540	<.0001
mum_edu6	1	1.5553	0.0319	2382.4532	<.0001
dad_edu2	1	0.2584	0.0107	586.2159	<.0001
dad_edu3	1	1.0516	0.0276	1450.6294	<.0001
dad_edu4	1	0.6733	0.0199	1141.8994	<.0001
dad_edu5	1	1.1518	0.0159	5243.9658	<.0001
dad_edu6	1	1.5893	0.0226	4946.8257	<.0001
mum_ingrp2	1	-0.0417	0.0126	10.8618	0.0010
mum_ingrp3	1	0.0965	0.0127	57.8132	<.0001
mum_ingrp4	1	0.3615	0.0134	725.4860	<.0001
mum_ingrp5	1	0.6147	0.0518	140.5833	<.0001
dad_ingrp2	1	-0.0772	0.0128	36.1917	<.0001
dad_ingrp3	1	0.1180	0.0129	84.0561	<.0001
dad_ingrp4	1	0.4218	0.0134	985.2103	<.0001
dad_ingrp5	1	0.7613	0.0463	270.7674	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.9555	0.0114	6967.0985	<.0001
mum_edu2	1	0.4066	0.0100	1645.8891	<.0001
mum_edu3	1	0.9674	0.0310	974.2796	<.0001
mum_edu4	1	0.9180	0.0236	1511.6624	<.0001
mum_edu5	1	1.0389	0.0138	5635.9724	<.0001
mum_edu6	1	1.4999	0.0382	1544.8631	<.0001
dad_edu2	1	0.3050	0.0100	928.0881	<.0001
dad_edu3	1	0.8862	0.0309	823.5771	<.0001
dad_edu4	1	0.6381	0.0208	944.8882	<.0001
dad_edu5	1	1.0374	0.0174	3541.9640	<.0001
dad_edu6	1	1.4185	0.0262	2930.3077	<.0001
mum_ingrp2	1	0.0840	0.0119	50.0964	<.0001
mum_ingrp3	1	0.2567	0.0122	440.4373	<.0001
mum_ingrp4	1	0.4823	0.0135	1275.6287	<.0001
mum_ingrp5	1	0.6563	0.0585	125.9158	<.0001
dad_ingrp2	1	0.0200	0.0120	2.7607	0.0966
dad_ingrp3	1	0.2051	0.0125	271.3766	<.0001
dad_ingrp4	1	0.4663	0.0135	1198.7350	<.0001
dad_ingrp5	1	0.7079	0.0513	190.0569	<.0001

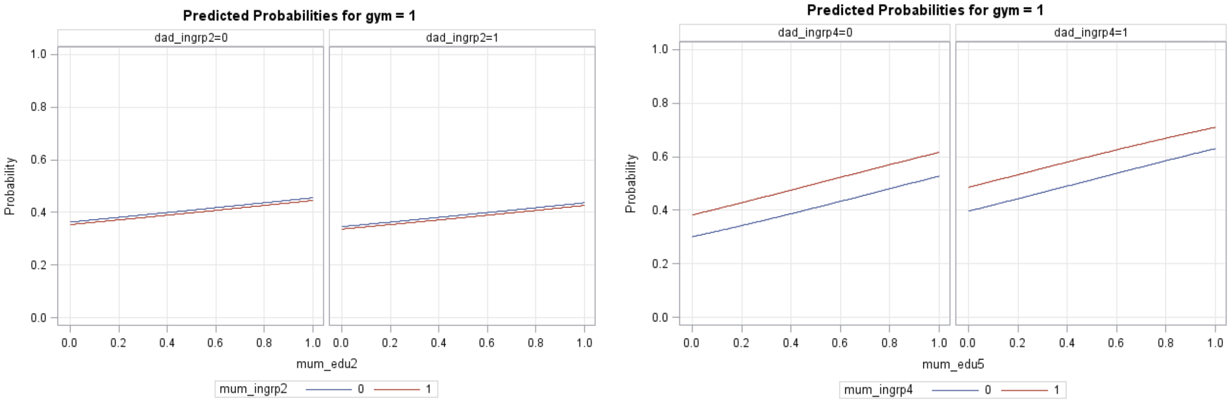
Figure A.2: Left Plot: Boys, Right Plot: Girls. Native origin, data form 2006



Classification Table									
	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	47569	143E3	21423	59561	70.2	44.4	87.0	31.1	29.4

Classification Table									
	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	98468	74103	43367	45183	66.1	68.5	63.1	30.6	37.9

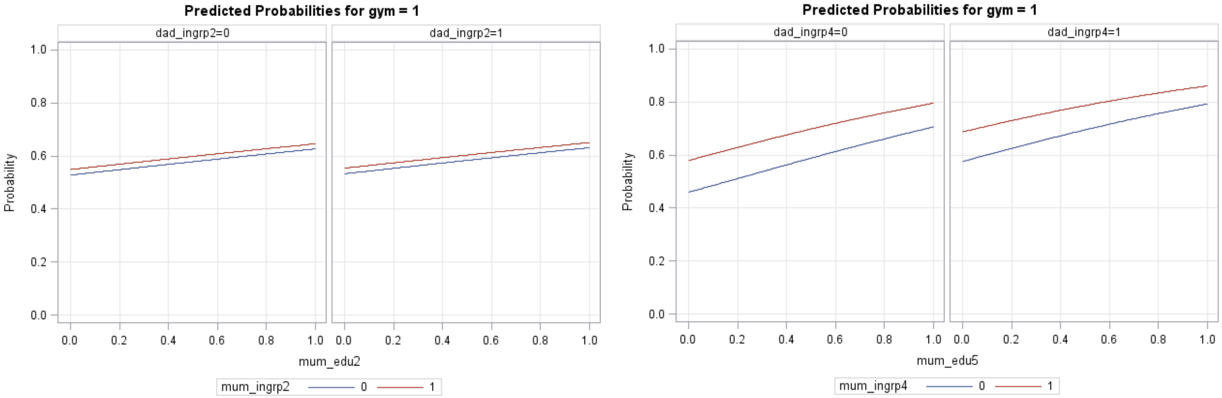
Figure A.3: Predicted probability plots for boys with native origin, data form 2006



Fit computed at mum_edu3=0.022 mum_edu4=0.038 mum_edu5=0.205 mum_edu6=0.036 dad_edu2=0.405 dad_edu3=0.025 dad_edu4=0.049 dad_edu5=0.111 dad_edu6=0.07 mum_ingrp3=0.25 mum_ingrp4=0.24 mum_ingrp5=0.01 dad_ingrp3=0.251 dad_ingrp4=0.239 dad_ingrp5=0.01

Fit computed at mum_edu2=0.341 mum_edu3=0.022 mum_edu4=0.038 mum_edu6=0.036 dad_edu2=0.405 dad_edu3=0.025 dad_edu4=0.049 dad_edu5=0.111 dad_edu6=0.07 mum_ingrp2=0.25 mum_ingrp3=0.25 mum_ingrp5=0.01 dad_ingrp2=0.249 dad_ingrp3=0.251 dad_ingrp5=0.01

Figure A.4: Predicted probability plots for girls with native origin, data form 2006



Fit computed at mum_edu3=0.022 mum_edu4=0.039 mum_edu5=0.204 mum_edu6=0.035 dad_edu2=0.404 dad_edu3=0.024 dad_edu4=0.049 dad_edu5=0.11 dad_edu6=0.071 mum_ingrp3=0.25 mum_ingrp4=0.24 mum_ingrp5=0.01 dad_ingrp3=0.249 dad_ingrp4=0.241 dad_ingrp5=0.01

Fit computed at mum_edu2=0.342 mum_edu3=0.022 mum_edu4=0.039 mum_edu6=0.035 dad_edu2=0.404 dad_edu3=0.024 dad_edu4=0.049 dad_edu5=0.11 dad_edu6=0.071 mum_ingrp2=0.25 mum_ingrp3=0.25 mum_ingrp5=0.01 dad_ingrp2=0.251 dad_ingrp3=0.249 dad_ingrp5=0.01

The test results are set up as the following the label *testXYZ* indicates which effects are tested. **X** indicates respectively M/D for mum/dad, Y indicates respectively U/I for education/Income and finally Z indicates the level of education or income, recall the 1 is the lowest level.

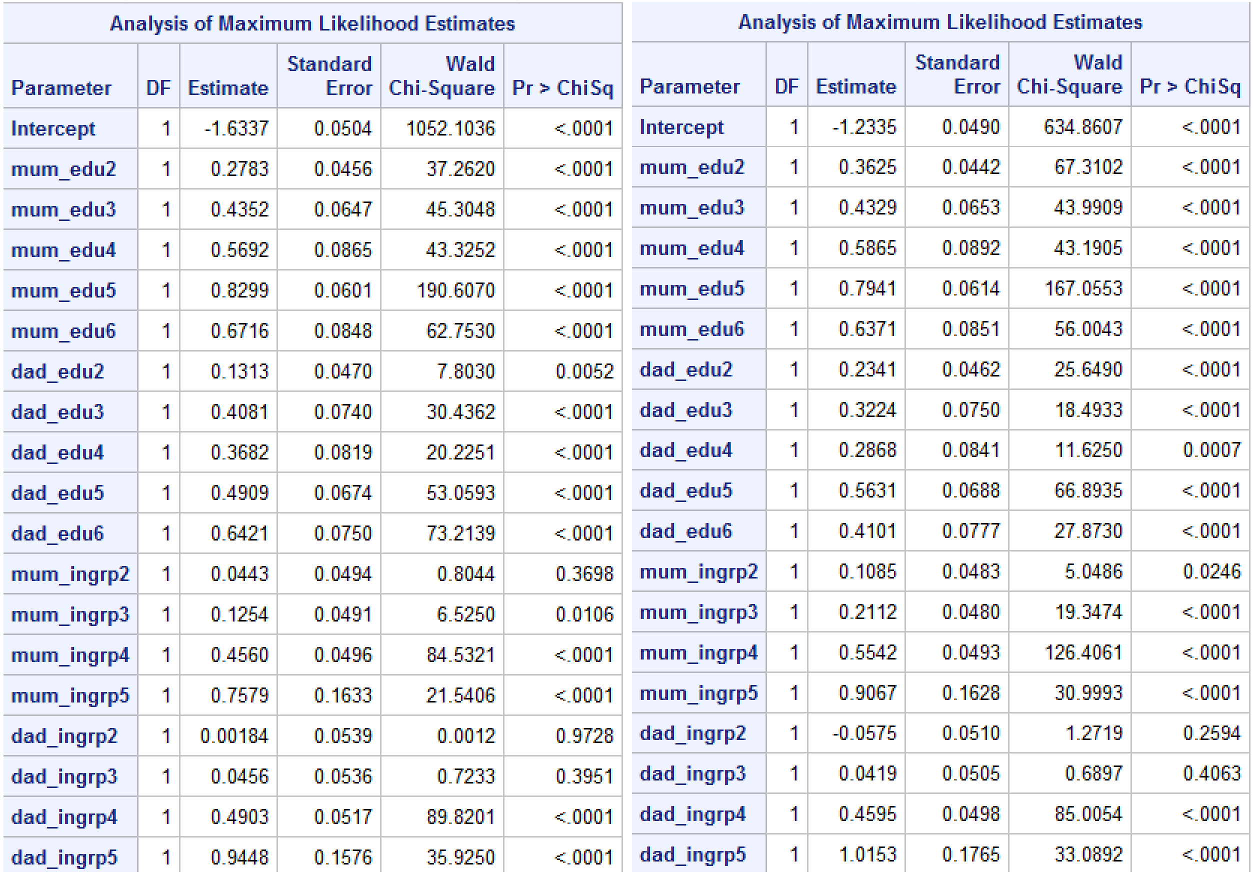
Figure A.5: Left Table: DF, Maximum likelihood estimates, std error, Wald Chi-square, P-value of $\beta_i^{\text{Boys}} = \beta_i^{\text{Girls}}$. Right Table: Test $\beta_i^{\text{Boys}} = \beta_i^{\text{Girls}}$, where i is a parental level of education or income. Native origin, data form 2006

mu2	1	-0.0248	0.0147	2.8404	0.0919
mu3	1	-0.0486	0.0422	1.3271	0.2493
mu4	1	-0.0464	0.0325	2.0406	0.1531
mu5	1	-0.0839	0.0192	19.1016	<.0001
mu6	1	0.0558	0.0497	1.2618	0.2613
du2	1	-0.0466	0.0146	10.1235	0.0015
du3	1	0.1655	0.0414	15.9551	<.0001
du4	1	0.0352	0.0288	1.4938	0.2216
du5	1	0.1144	0.0236	23.5025	<.0001
du6	1	0.1710	0.0346	24.4136	<.0001
mi2	1	-0.1257	0.0173	52.5190	<.0001
mi3	1	-0.1602	0.0176	82.5322	<.0001
mi4	1	-0.1208	0.0190	40.2481	<.0001
mi5	1	-0.0414	0.0782	0.2800	0.5967
di2	1	-0.0972	0.0176	30.4967	<.0001
di3	1	-0.0871	0.0179	23.6839	<.0001
di4	1	-0.0445	0.0190	5.4803	0.0192
di5	1	0.0534	0.0691	0.5975	0.4395

Linear Hypotheses Testing Results

Label	Wald Chi-Square	DF	Pr > ChiSq
testmu2	2.8404	1	0.0919
testmu3	1.3271	1	0.2493
testmu4	2.0406	1	0.1531
testmu5	19.1016	1	<.0001
testmu6	1.2618	1	0.2613
testdu2	10.1235	1	0.0015
testdu3	15.9551	1	<.0001
testdu4	1.4938	1	0.2216
testdu5	23.5025	1	<.0001
testdu6	24.4136	1	<.0001
testmi2	52.5190	1	<.0001
testmi3	82.5322	1	<.0001
testmi4	40.2481	1	<.0001
testmi5	0.2800	1	0.5967
testdi2	30.4967	1	<.0001
testdi3	23.6839	1	<.0001
testdi4	5.4803	1	0.0192
testdi5	0.5975	1	0.4395

Figure A.6: Left Table: Boys, Right Table: Girls. Immigrant origin, data form 2006



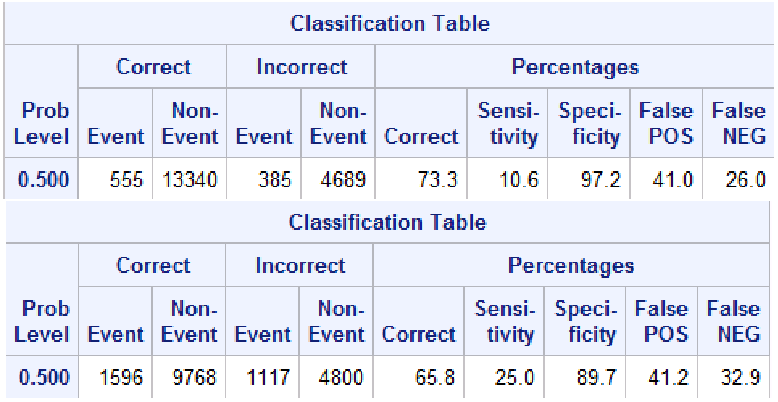
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.6337	0.0504	1052.1036	<.0001
mum_edu2	1	0.2783	0.0456	37.2620	<.0001
mum_edu3	1	0.4352	0.0647	45.3048	<.0001
mum_edu4	1	0.5692	0.0865	43.3252	<.0001
mum_edu5	1	0.8299	0.0601	190.6070	<.0001
mum_edu6	1	0.6716	0.0848	62.7530	<.0001
dad_edu2	1	0.1313	0.0470	7.8030	0.0052
dad_edu3	1	0.4081	0.0740	30.4362	<.0001
dad_edu4	1	0.3682	0.0819	20.2251	<.0001
dad_edu5	1	0.4909	0.0674	53.0593	<.0001
dad_edu6	1	0.6421	0.0750	73.2139	<.0001
mum_ingrp2	1	0.0443	0.0494	0.8044	0.3698
mum_ingrp3	1	0.1254	0.0491	6.5250	0.0106
mum_ingrp4	1	0.4560	0.0496	84.5321	<.0001
mum_ingrp5	1	0.7579	0.1633	21.5406	<.0001
dad_ingrp2	1	0.00184	0.0539	0.0012	0.9728
dad_ingrp3	1	0.0456	0.0536	0.7233	0.3951
dad_ingrp4	1	0.4903	0.0517	89.8201	<.0001
dad_ingrp5	1	0.9448	0.1576	35.9250	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.2335	0.0490	634.8607	<.0001
mum_edu2	1	0.3625	0.0442	67.3102	<.0001
mum_edu3	1	0.4329	0.0653	43.9909	<.0001
mum_edu4	1	0.5865	0.0892	43.1905	<.0001
mum_edu5	1	0.7941	0.0614	167.0553	<.0001
mum_edu6	1	0.6371	0.0851	56.0043	<.0001
dad_edu2	1	0.2341	0.0462	25.6490	<.0001
dad_edu3	1	0.3224	0.0750	18.4933	<.0001
dad_edu4	1	0.2868	0.0841	11.6250	0.0007
dad_edu5	1	0.5631	0.0688	66.8935	<.0001
dad_edu6	1	0.4101	0.0777	27.8730	<.0001
mum_ingrp2	1	0.1085	0.0483	5.0486	0.0246
mum_ingrp3	1	0.2112	0.0480	19.3474	<.0001
mum_ingrp4	1	0.5542	0.0493	126.4061	<.0001
mum_ingrp5	1	0.9067	0.1628	30.9993	<.0001
dad_ingrp2	1	-0.0575	0.0510	1.2719	0.2594
dad_ingrp3	1	0.0419	0.0505	0.6897	0.4063
dad_ingrp4	1	0.4595	0.0498	85.0054	<.0001
dad_ingrp5	1	1.0153	0.1765	33.0892	<.0001

Figure A.7: Top Table: Boys, Bottom Table: Girls. Immigrant origin, data form 2006



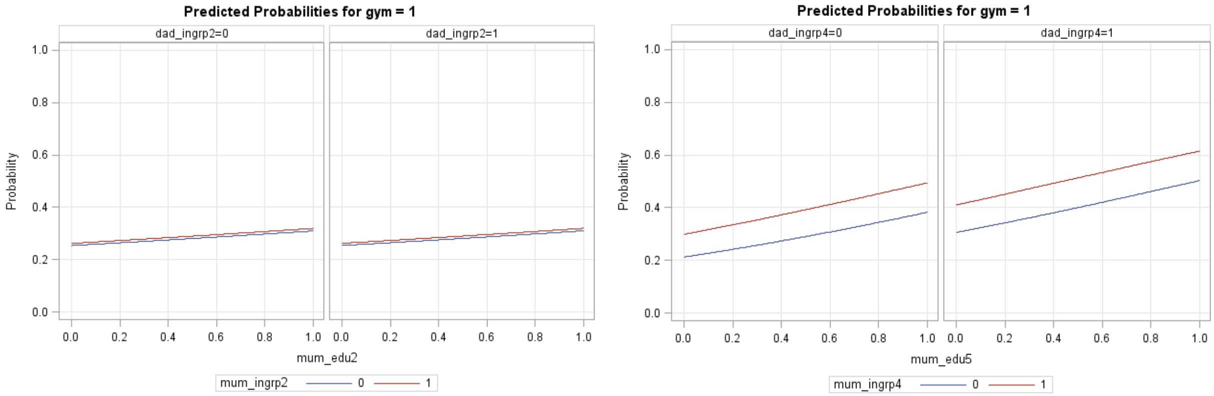
Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	555	13340	385	4689	73.3	10.6	97.2	41.0	26.0

Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	1596	9768	1117	4800	65.8	25.0	89.7	41.2	32.9

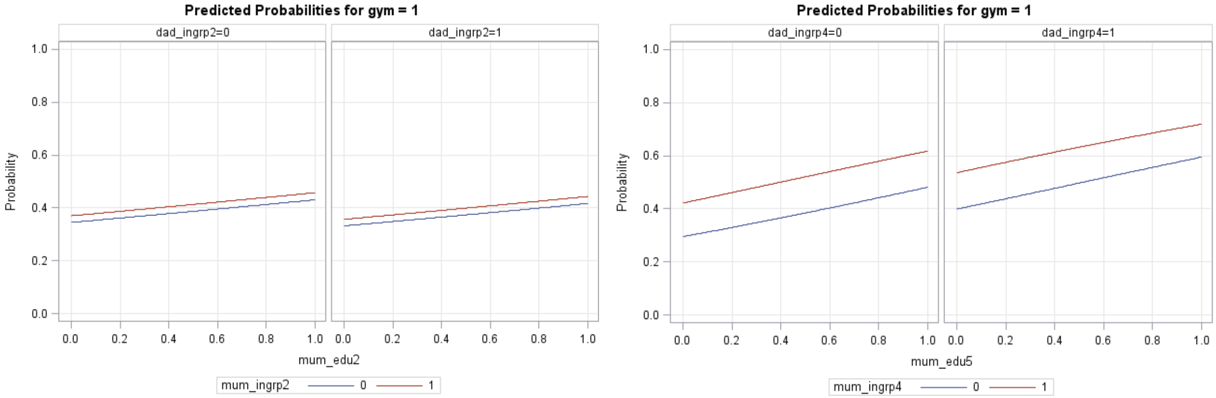
Figure A.8: Predicted probability plots for boys with immigrant origin, data form 2006



Fit computed at mum_edu3=0.071 mum_edu4=0.035 mum_edu5=0.081 mum_edu6=0.037 dad_edu2=0.194 dad_edu3=0.053 dad_edu4=0.042 dad_edu5=0.065 dad_edu6=0.05 mum_ingrp3=0.245 mum_ingrp4=0.235 mum_ingrp5=0.009 dad_ingrp3=0.251 dad_ingrp4=0.243 dad_ingrp5=0.01

Fit computed at mum_edu2=0.19 mum_edu3=0.071 mum_edu4=0.035 mum_edu6=0.037 dad_edu2=0.194 dad_edu3=0.053 dad_edu4=0.042 dad_edu5=0.065 dad_edu6=0.05 mum_ingrp2=0.247 mum_ingrp3=0.245 mum_ingrp5=0.009 dad_ingrp2=0.254 dad_ingrp3=0.251 dad_ingrp5=0.01

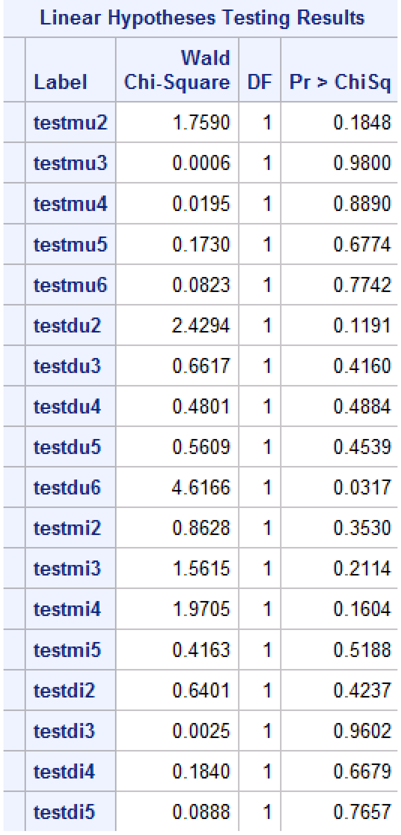
Figure A.9: Predicted probability plots for girls with immigrant origin, data form 2006



Fit computed at mum_edu3=0.067 mum_edu4=0.033 mum_edu5=0.081 mum_edu6=0.038 dad_edu2=0.181 dad_edu3=0.051 dad_edu4=0.04 dad_edu5=0.063 dad_edu6=0.05 mum_ingrp3=0.255 mum_ingrp4=0.246 mum_ingrp5=0.011 dad_ingrp3=0.249 dad_ingrp4=0.237 dad_ingrp5=0.01

Fit computed at mum_edu2=0.189 mum_edu3=0.067 mum_edu4=0.033 mum_edu6=0.038 dad_edu2=0.181 dad_edu3=0.051 dad_edu4=0.04 dad_edu5=0.063 dad_edu6=0.05 mum_ingrp2=0.253 mum_ingrp3=0.255 mum_ingrp5=0.011 dad_ingrp2=0.246 dad_ingrp3=0.249 dad_ingrp5=0.01

Figure A.10: Test $\beta_i^{\text{Boys}} = \beta_i^{\text{Girls}}$, where i is a parental level of education or income. Immigrant origin, data form 2006

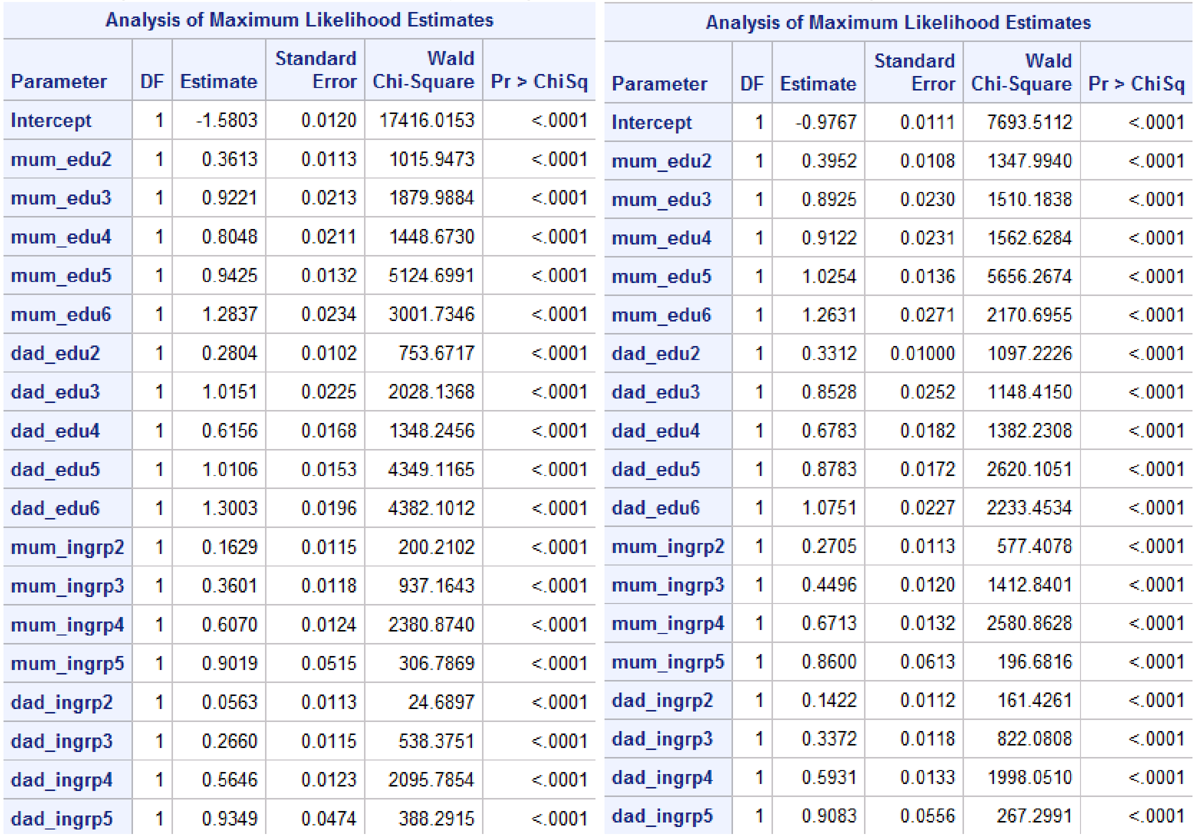


Linear Hypotheses Testing Results

Label	Wald Chi-Square	DF	Pr > ChiSq
testmu2	1.7590	1	0.1848
testmu3	0.0006	1	0.9800
testmu4	0.0195	1	0.8890
testmu5	0.1730	1	0.6774
testmu6	0.0823	1	0.7742
testdu2	2.4294	1	0.1191
testdu3	0.6617	1	0.4160
testdu4	0.4801	1	0.4884
testdu5	0.5609	1	0.4539
testdu6	4.6166	1	0.0317
testmi2	0.8628	1	0.3530
testmi3	1.5615	1	0.2114
testmi4	1.9705	1	0.1604
testmi5	0.4163	1	0.5188
testdi2	0.6401	1	0.4237
testdi3	0.0025	1	0.9602
testdi4	0.1840	1	0.6679
testdi5	0.0888	1	0.7657

2016

Figure A.11: Left Table: Boys, Right Table: Girls, native origin data form 2016



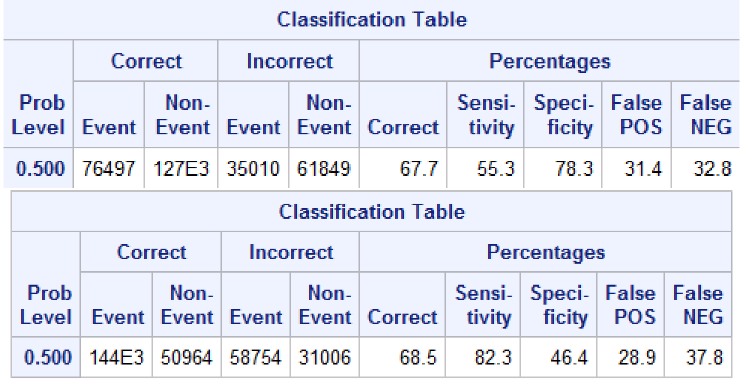
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.5803	0.0120	17416.0153	<.0001
mum_edu2	1	0.3613	0.0113	1015.9473	<.0001
mum_edu3	1	0.9221	0.0213	1879.9884	<.0001
mum_edu4	1	0.8048	0.0211	1448.6730	<.0001
mum_edu5	1	0.9425	0.0132	5124.6991	<.0001
mum_edu6	1	1.2837	0.0234	3001.7346	<.0001
dad_edu2	1	0.2804	0.0102	753.6717	<.0001
dad_edu3	1	1.0151	0.0225	2028.1368	<.0001
dad_edu4	1	0.6156	0.0168	1348.2456	<.0001
dad_edu5	1	1.0106	0.0153	4349.1165	<.0001
dad_edu6	1	1.3003	0.0196	4382.1012	<.0001
mum_ingrp2	1	0.1629	0.0115	200.2102	<.0001
mum_ingrp3	1	0.3601	0.0118	937.1643	<.0001
mum_ingrp4	1	0.6070	0.0124	2380.8740	<.0001
mum_ingrp5	1	0.9019	0.0515	306.7869	<.0001
dad_ingrp2	1	0.0563	0.0113	24.6897	<.0001
dad_ingrp3	1	0.2660	0.0115	538.3751	<.0001
dad_ingrp4	1	0.5646	0.0123	2095.7854	<.0001
dad_ingrp5	1	0.9349	0.0474	388.2915	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.9767	0.0111	7693.5112	<.0001
mum_edu2	1	0.3952	0.0108	1347.9940	<.0001
mum_edu3	1	0.8925	0.0230	1510.1838	<.0001
mum_edu4	1	0.9122	0.0231	1562.6284	<.0001
mum_edu5	1	1.0254	0.0136	5656.2674	<.0001
mum_edu6	1	1.2631	0.0271	2170.6955	<.0001
dad_edu2	1	0.3312	0.01000	1097.2226	<.0001
dad_edu3	1	0.8528	0.0252	1148.4150	<.0001
dad_edu4	1	0.6783	0.0182	1382.2308	<.0001
dad_edu5	1	0.8783	0.0172	2620.1051	<.0001
dad_edu6	1	1.0751	0.0227	2233.4534	<.0001
mum_ingrp2	1	0.2705	0.0113	577.4078	<.0001
mum_ingrp3	1	0.4496	0.0120	1412.8401	<.0001
mum_ingrp4	1	0.6713	0.0132	2580.8628	<.0001
mum_ingrp5	1	0.8600	0.0613	196.6816	<.0001
dad_ingrp2	1	0.1422	0.0112	161.4261	<.0001
dad_ingrp3	1	0.3372	0.0118	822.0808	<.0001
dad_ingrp4	1	0.5931	0.0133	1998.0510	<.0001
dad_ingrp5	1	0.9083	0.0556	267.2991	<.0001

Figure A.12: Top Table: Boys, Bottom Table: Girls, native origin data form 2016



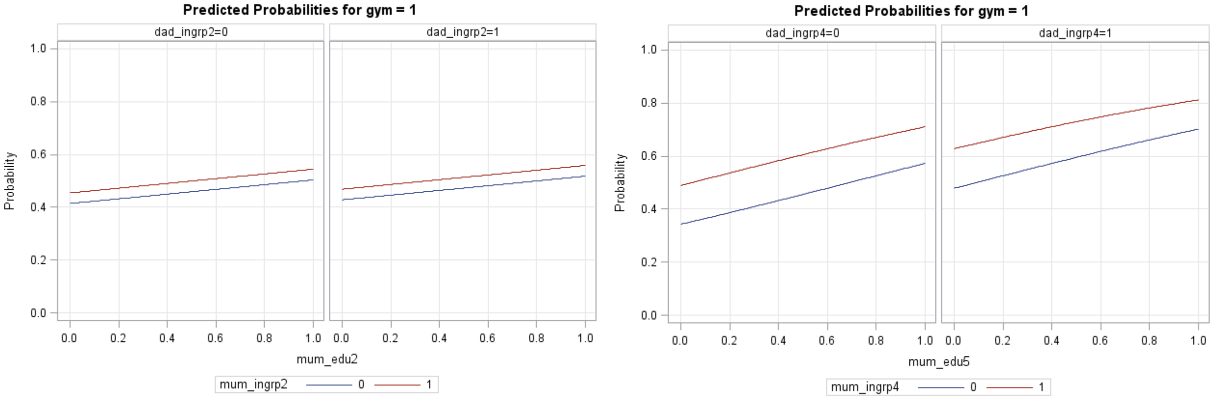
Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	76497	127E3	35010	61849	67.7	55.3	78.3	31.4	32.8

Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	144E3	50964	58754	31006	68.5	82.3	46.4	28.9	37.8

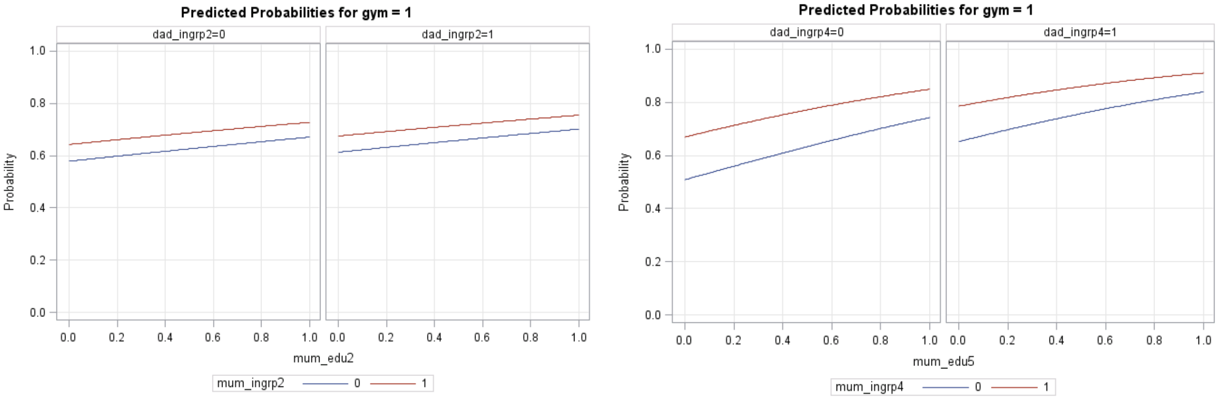
Figure A.13: Predicted probability plots for boys with native origin, data form 2016



Fit computed at mum_edu3=0.042 mum_edu4=0.042 mum_edu5=0.242 mum_edu6=0.062 dad_edu2=0.401 dad_edu3=0.037 dad_edu4=0.07 dad_edu5=0.114 dad_edu6=0.087 mum_ingrp3=0.241 mum_ingrp4=0.234 mum_ingrp5=0.01 dad_ingrp3=0.229 dad_ingrp4=0.222 dad_ingrp5=0.009

Fit computed at mum_edu2=0.386 mum_edu3=0.042 mum_edu4=0.042 mum_edu6=0.062 dad_edu2=0.401 dad_edu3=0.037 dad_edu4=0.07 dad_edu5=0.114 dad_edu6=0.087 mum_ingrp2=0.241 mum_ingrp3=0.241 mum_ingrp5=0.01 dad_ingrp2=0.228 dad_ingrp3=0.229 dad_ingrp5=0.009

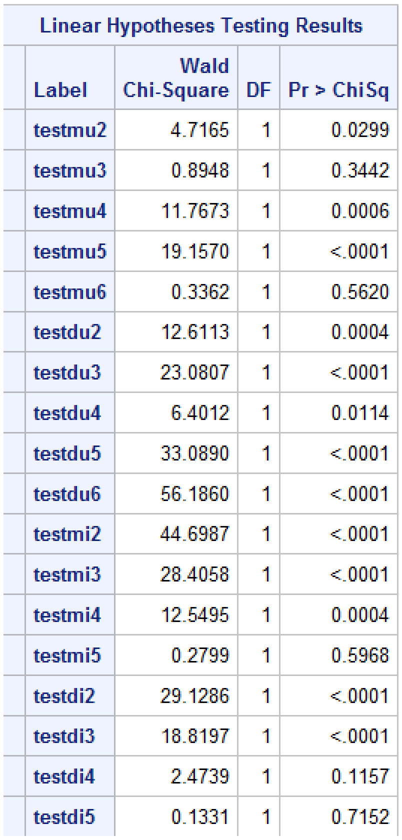
Figure A.14: Predicted probability plots for girls with native origin, data form 2016



Fit computed at mum_edu3=0.041 mum_edu4=0.043 mum_edu5=0.242 mum_edu6=0.061 dad_edu2=0.404 dad_edu3=0.037 dad_edu4=0.07 dad_edu5=0.113 dad_edu6=0.087 mum_ingrp3=0.242 mum_ingrp4=0.231 mum_ingrp5=0.009 dad_ingrp3=0.229 dad_ingrp4=0.221 dad_ingrp5=0.009

Fit computed at mum_edu2=0.386 mum_edu3=0.041 mum_edu4=0.043 mum_edu6=0.061 dad_edu2=0.404 dad_edu3=0.037 dad_edu4=0.07 dad_edu5=0.113 dad_edu6=0.087 mum_ingrp2=0.243 mum_ingrp3=0.242 mum_ingrp5=0.009 dad_ingrp2=0.229 dad_ingrp3=0.229 dad_ingrp5=0.009

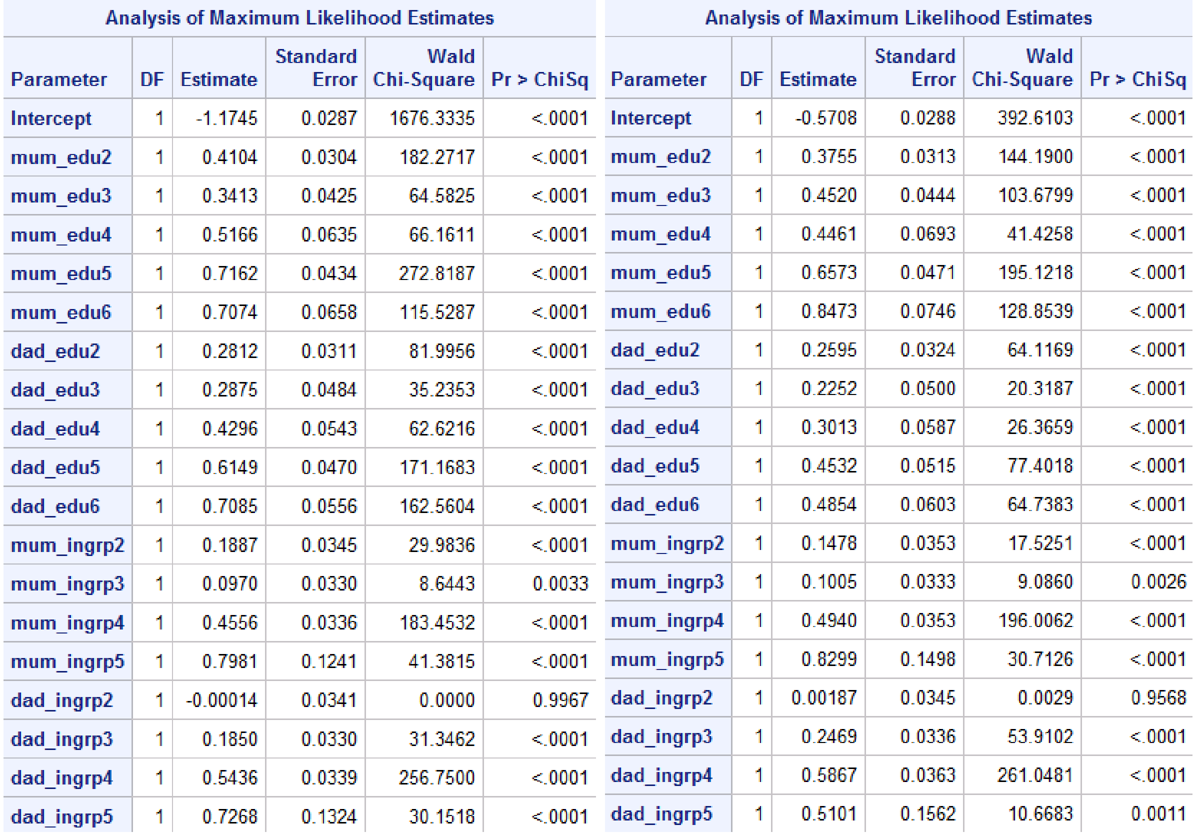
Figure A.15: Test $\beta_i^{\text{Boys}} = \beta_i^{\text{Girls}}$, where i is a parental level of education or income. Native origin, data form 2016



Linear Hypotheses Testing Results

Label	Wald Chi-Square	DF	Pr > ChiSq
testmu2	4.7165	1	0.0299
testmu3	0.8948	1	0.3442
testmu4	11.7673	1	0.0006
testmu5	19.1570	1	<.0001
testmu6	0.3362	1	0.5620
testdu2	12.6113	1	0.0004
testdu3	23.0807	1	<.0001
testdu4	6.4012	1	0.0114
testdu5	33.0890	1	<.0001
testdu6	56.1860	1	<.0001
testmi2	44.6987	1	<.0001
testmi3	28.4058	1	<.0001
testmi4	12.5495	1	0.0004
testmi5	0.2799	1	0.5968
testdi2	29.1286	1	<.0001
testdi3	18.8197	1	<.0001
testdi4	2.4739	1	0.1157
testdi5	0.1331	1	0.7152

Figure A.16: Left Table: Boys, Right Table: Girls, immigrant origin data form 2016



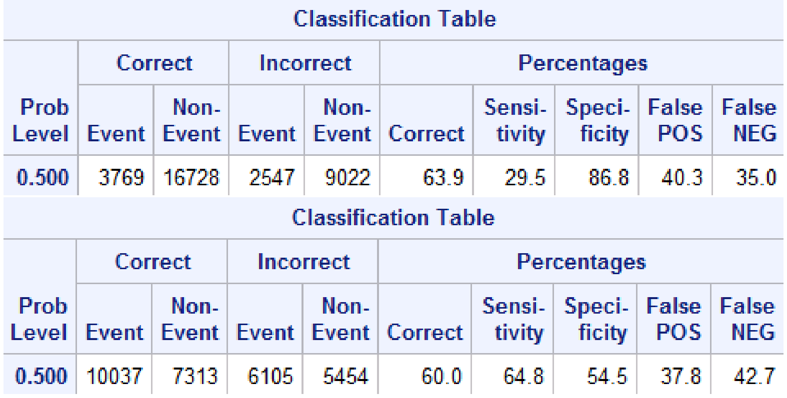
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.1745	0.0287	1676.3335	<.0001
mum_edu2	1	0.4104	0.0304	182.2717	<.0001
mum_edu3	1	0.3413	0.0425	64.5825	<.0001
mum_edu4	1	0.5166	0.0635	66.1611	<.0001
mum_edu5	1	0.7162	0.0434	272.8187	<.0001
mum_edu6	1	0.7074	0.0658	115.5287	<.0001
dad_edu2	1	0.2812	0.0311	81.9956	<.0001
dad_edu3	1	0.2875	0.0484	35.2353	<.0001
dad_edu4	1	0.4296	0.0543	62.6216	<.0001
dad_edu5	1	0.6149	0.0470	171.1683	<.0001
dad_edu6	1	0.7085	0.0556	162.5604	<.0001
mum_ingrp2	1	0.1887	0.0345	29.9836	<.0001
mum_ingrp3	1	0.0970	0.0330	8.6443	0.0033
mum_ingrp4	1	0.4556	0.0336	183.4532	<.0001
mum_ingrp5	1	0.7981	0.1241	41.3815	<.0001
dad_ingrp2	1	-0.00014	0.0341	0.0000	0.9967
dad_ingrp3	1	0.1850	0.0330	31.3462	<.0001
dad_ingrp4	1	0.5436	0.0339	256.7500	<.0001
dad_ingrp5	1	0.7268	0.1324	30.1518	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.5708	0.0288	392.6103	<.0001
mum_edu2	1	0.3755	0.0313	144.1900	<.0001
mum_edu3	1	0.4520	0.0444	103.6799	<.0001
mum_edu4	1	0.4461	0.0693	41.4258	<.0001
mum_edu5	1	0.6573	0.0471	195.1218	<.0001
mum_edu6	1	0.8473	0.0746	128.8539	<.0001
dad_edu2	1	0.2595	0.0324	64.1169	<.0001
dad_edu3	1	0.2252	0.0500	20.3187	<.0001
dad_edu4	1	0.3013	0.0587	26.3659	<.0001
dad_edu5	1	0.4532	0.0515	77.4018	<.0001
dad_edu6	1	0.4854	0.0603	64.7383	<.0001
mum_ingrp2	1	0.1478	0.0353	17.5251	<.0001
mum_ingrp3	1	0.1005	0.0333	9.0860	0.0026
mum_ingrp4	1	0.4940	0.0353	196.0062	<.0001
mum_ingrp5	1	0.8299	0.1498	30.7126	<.0001
dad_ingrp2	1	0.00187	0.0345	0.0029	0.9568
dad_ingrp3	1	0.2469	0.0336	53.9102	<.0001
dad_ingrp4	1	0.5867	0.0363	261.0481	<.0001
dad_ingrp5	1	0.5101	0.1562	10.6683	0.0011

Figure A.17: Left Table: Boys, Right Table: Girls. Immigrant origin, data form 2016



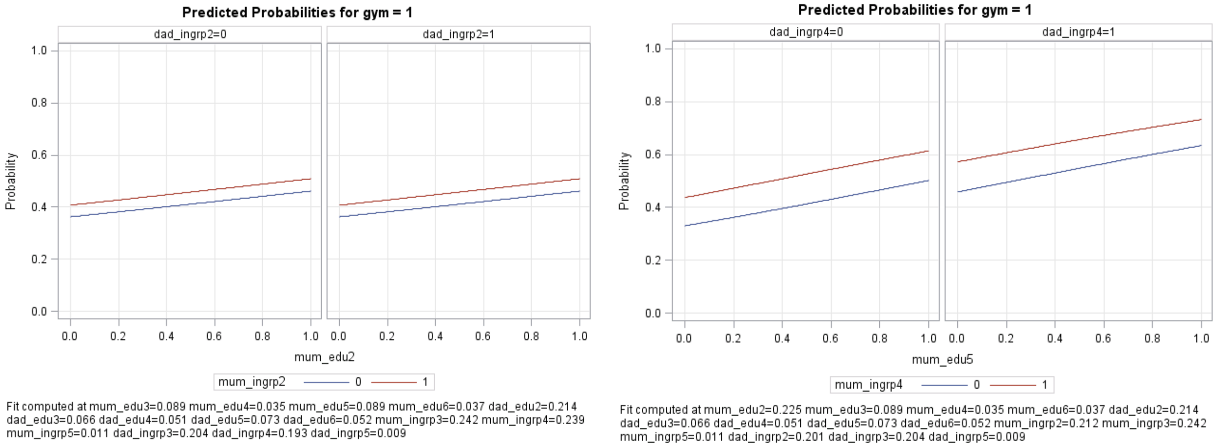
Classification Table

Prob Level	Correct		Incorrect		Percentages				
	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	3769	16728	2547	9022	63.9	29.5	86.8	40.3	35.0

Classification Table

Prob Level	Correct		Incorrect		Percentages				
	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	10037	7313	6105	5454	60.0	64.8	54.5	37.8	42.7

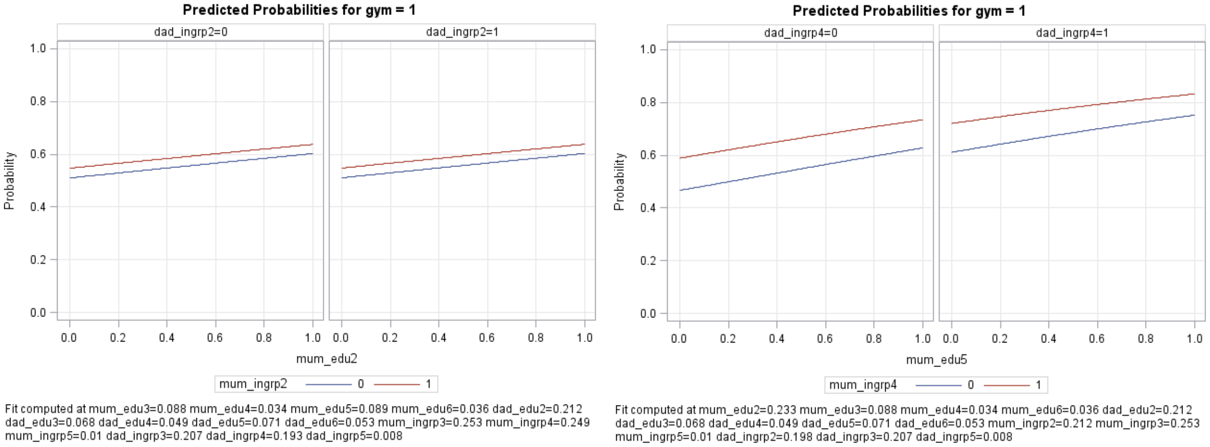
Figure A.18: Predicted probability plots for boys with immigrant origin, data form 2016



Fit computed at mum_edu3=0.089 mum_edu4=0.035 mum_edu5=0.089 mum_edu6=0.037 dad_edu2=0.214 dad_edu3=0.066 dad_edu4=0.051 dad_edu5=0.073 dad_edu6=0.052 mum_ingrp3=0.242 mum_ingrp4=0.239 mum_ingrp5=0.011 dad_ingrp3=0.204 dad_ingrp4=0.193 dad_ingrp5=0.009

Fit computed at mum_edu2=0.225 mum_edu3=0.089 mum_edu4=0.035 mum_edu6=0.037 dad_edu2=0.214 dad_edu3=0.066 dad_edu4=0.051 dad_edu5=0.073 dad_edu6=0.052 mum_ingrp2=0.212 mum_ingrp3=0.242 mum_ingrp5=0.011 dad_ingrp2=0.201 dad_ingrp3=0.204 dad_ingrp5=0.009

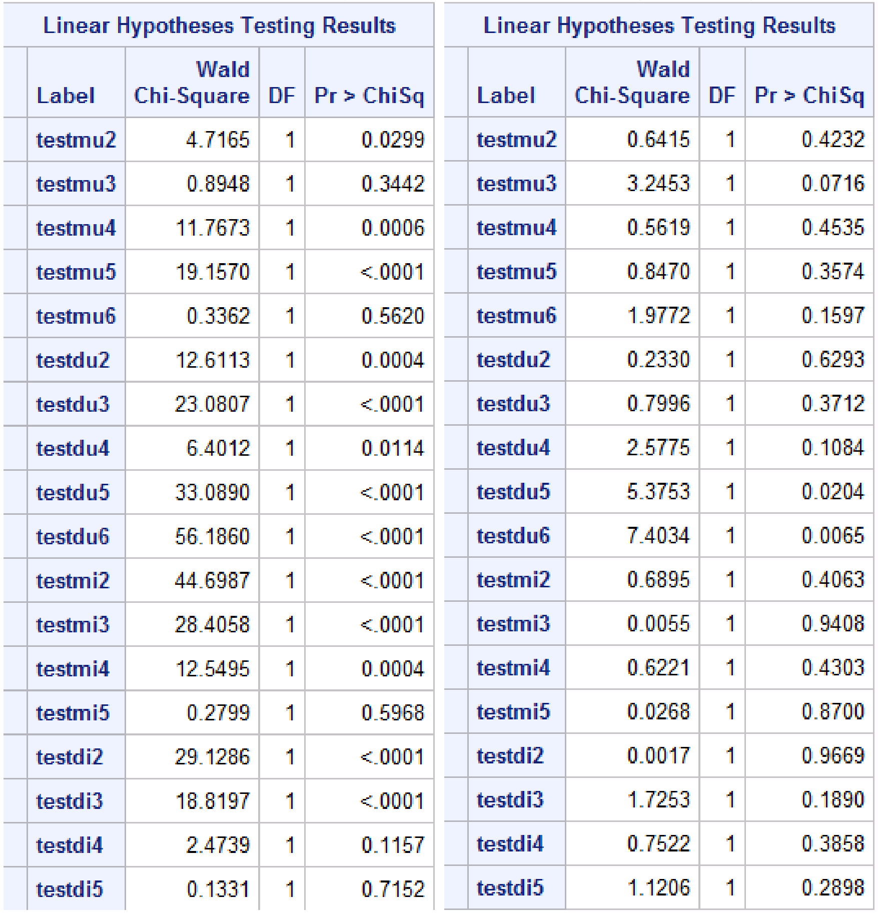
Figure A.19: Predicted probability plots for girls with immigrant origin, data form 2016



Fit computed at mum_edu3=0.088 mum_edu4=0.034 mum_edu5=0.089 mum_edu6=0.036 dad_edu2=0.212 dad_edu3=0.068 dad_edu4=0.049 dad_edu5=0.071 dad_edu6=0.053 mum_ingrp3=0.253 mum_ingrp4=0.249 mum_ingrp5=0.01 dad_ingrp3=0.207 dad_ingrp4=0.193 dad_ingrp5=0.008

Fit computed at mum_edu2=0.233 mum_edu3=0.088 mum_edu4=0.034 mum_edu6=0.036 dad_edu2=0.212 dad_edu3=0.068 dad_edu4=0.049 dad_edu5=0.071 dad_edu6=0.053 mum_ingrp2=0.212 mum_ingrp3=0.253 mum_ingrp5=0.01 dad_ingrp2=0.198 dad_ingrp3=0.207 dad_ingrp5=0.008

Figure A.20: Left Table: Native origin, Right Table: Immigrant origin, data form 2016



Linear Hypotheses Testing Results

Label	Wald Chi-Square	DF	Pr > ChiSq
testmu2	4.7165	1	0.0299
testmu3	0.8948	1	0.3442
testmu4	11.7673	1	0.0006
testmu5	19.1570	1	<.0001
testmu6	0.3362	1	0.5620
testdu2	12.6113	1	0.0004
testdu3	23.0807	1	<.0001
testdu4	6.4012	1	0.0114
testdu5	33.0890	1	<.0001
testdu6	56.1860	1	<.0001
testmi2	44.6987	1	<.0001
testmi3	28.4058	1	<.0001
testmi4	12.5495	1	0.0004
testmi5	0.2799	1	0.5968
testdi2	29.1286	1	<.0001
testdi3	18.8197	1	<.0001
testdi4	2.4739	1	0.1157
testdi5	0.1331	1	0.7152

Linear Hypotheses Testing Results

Label	Wald Chi-Square	DF	Pr > ChiSq
testmu2	0.6415	1	0.4232
testmu3	3.2453	1	0.0716
testmu4	0.5619	1	0.4535
testmu5	0.8470	1	0.3574
testmu6	1.9772	1	0.1597
testdu2	0.2330	1	0.6293
testdu3	0.7996	1	0.3712
testdu4	2.5775	1	0.1084
testdu5	5.3753	1	0.0204
testdu6	7.4034	1	0.0065
testmi2	0.6895	1	0.4063
testmi3	0.0055	1	0.9408
testmi4	0.6221	1	0.4303
testmi5	0.0268	1	0.8700
testdi2	0.0017	1	0.9669
testdi3	1.7253	1	0.1890
testdi4	0.7522	1	0.3858
testdi5	1.1206	1	0.2898

Appendix B

Removing single parents

2006

The original data samples contained respectively $532,767$ and $36,250$ individuals with native and immigrant origins. The sample size of the immigrant origin data reduces to $29,596$ after removing individuals without both parents registered in Denmark. This leads an small increase in proportion immigrants with a high school education to 33.70%. The sample size of the native origin data are reduced to $528,660$ individuals, where 47.14% have a high school education. The relative greatest change in sample size is seen among the immigrant origin, which also was expected. Removing the individuals with one parent registered in Denmark has almost no effect on the proportion of individuals with a high school education within the native origin dataset (0.07%), while it increased the proportion of individuals with immigrant origin with 1.59% point, corresponding to 4.98%. A review of the "new" datasets gender specifications are found in the following table

Table B.1: Data from 2006

Gender	Graduated HS	Immigrant origin			Native origin		
	High School	#	Cum. Freq.	Gender %	#	Cum. Freq.	Gender %
Boys	No (0)	11,013	37.21%	70.95%	163,118	30.85%	60.51%
	Yes (1)	4,510	15.24%	29.05%	106,454	20.14%	39.49%
Girls	No (0)	8,610	29.09%	61.18%	116,339	22.01%	44.9%
	Yes (1)	5,463	18.46%	38.82%	142,749	27.0%	55.10%
	Total	29,596			528,660		

and one finds, no evolving change in native origin, while the individuals with immigrant origin has increased the proportion of as well boys as girls with a high school education. The non-existing change in the proportion individuals with native origin still might provide differences in the effects of parental level of income and education.

2016

The sample from the benchmark case contained $585,090$ individuals with native origin. Removing very individual without both parents registered in Denmark reduces the sample size to $579,918$ individuals. The sample containing individuals with immigrant origin is reduced to a sample size with $52,197$ individuals after removing individuals without both parents registered in Denmark.
The new samples, provides the information that 53.71% of natives with both parents registered in Denmark, had a high school education in 2016, while 48.48% with immigrant origin did. Deleting individuals without both parents registered in Denmark from the datasets, provides a higher fraction of individuals who have graduated from high school, as the proportions in benchmark was 52.62% for natives and 46.38% for immigrants in 2016. Furthermore one sees, that while the native origin data sample has increased the proportion of high school educated individuals by 1.09% point, corresponding to a increase in 2.07% the immigrant proportion with an high school education has increased 2.1% corresponding to 4.53%. Then the fractions of high school graduated, by gender, are found as:

Table B.2: Data from 2016

Gender	Graduated HS	Immigrant origin			Native origin		
	High School	#	Cum. Freq.	Gender %	#	Cum. Freq.	Gender %
Boys	No (0)	15,959	30.57%	58.18%	159,940	27.58%	53.8%
	Yes (1)	10,931	21.98%	41.82%	137,368	23.69%	46.2%
Girls	No (0)	11,472	20.94%	26.51%	108,500	18.71%	38.39%
	Yes (1)	13,835	26.51 %	55.86%	174,110	30.02%	61.61%
	Total	52,197			579,918		

Here one finds, that the changes in the proportions within the natives are very low, while the proportion of boys as well as girls with immigrant origin who have an education, is raised. And given the relatively small changes it gave in the sample composition for the native origin, the finding are corresponding to the findings in Weiss (1978) . Looking at the correlation between high school graduated individuals and parental level of income and education, in both years, after removing all individuals with single parent, the changes are small.

Table B.3: Correlations without individuals with single-parents

	Origin, Year	Mother_edu	Father_edu	Mother_in	Father_in
High School	Immigrant, 2006	0.156	0.136	0.11385	0.1087
High School	Native, 2006	0.3134	0. 3159	0.2066	0.2035
High School	Immigrant, 2016	0.1469	0.1318	0.1225	0.1291
High School	Native, 2016	0.3038	0.2945	0.2394	0.2 288

Estimation of models

2006

Setting up the models for the native dataset based on data from 2006, significant models are found. Further on significant levels in the maximum likelihood estimates (see appendix B), it is found that an higher level of parental education and income increases the probability of education among boys and girls. Just as seen earlier in the models estimated based in the benchmark case. So at first, no clear change. A closer look at the models performance in predicting, i.e. the models performance in discriminating between individuals having/not having. graduated from high school. The effect is summarized in the "c", the concordance Statistic (the area under the ROC-curve) and is observed as 73.5% for boys and 72.1% for girls - approximately what is seen the benchmark. Checking if the models only predicts wrongly in one direction, is again done with setting up classification tables with the prior probability of 50% The classification table provides the same results as in the benchmark scenario, especially for the model estimated on native boys. Moreover one also sees, that the model for girls still have some skewness, as predicting girls with an education mis-classify 37.9% of them, as not having an education. Again, this might be explained with girls being more social mobile.

The following effect plots shows differences across parental effects on boys and girls, for respectively stereotype families of low-resources and high-resources. The effect plots of predicted probability among the boys, shows no clear changes in effect of parental effects on the predicted probability of education when one compare to the simple model, which included single parents and neither is the case for girls. Testing if parental effects significantly differ among boys and girls without single parents it is found that no significant chances appear. A closer look at the immigrant origin with both parents registered in Denmark, might bring some changes in the models estimating, as the dataset experienced a major change: Both in amount of individuals, but also in the proportion of individuals with a high school education. The maximum likelihoods estimates for boys shows, fewer levels of fatherly income being significant and further that these effects has a negative effect on probability of high school education. For girls the maximum likelihoods estimates have the same effect, and all fatherly effects are significant and the second income level has negative effect on the probability. The models performance for boys is a little lower, while the model for girls predicts approximately the same rate correctly. Both models are still skew in its prediction of high school education, as the both still classify more than 40% of the individuals in the dataset as having an education while they are observed as not having one, but it is "only" mis-classifying 27.1% of the boys as not having an education, while they actually do, and 34.2% of the girls. This tells that the models are not perfect, but I will take it. There are many reasons why the model is not super, among others a reason could be that there are not so many observations to base the models upon, but it is more likely that it could be due to the fact, that the explanatory variables are not very good as explanatory variables when modelling whether or not young immigrants are getting an education, as parental effects have few clear connections with high school education among their children. The effect plots shows differences across parental effects on boys and girls, for respectively stereotype families of low-resources and high-resources the parental effects look much alike for the boys, maybe with a little smaller effect across motherly level of education. Which tells, that motherly education has a (very) little higher effect on boys getting an education, if the mother is a single registered parent in Denmark. Among the girls a more clear sign appear: The prior probability of education has increased, but nothing applies that the parental effects on the probability of education have changed much. Whether the parental effects differ between gender are again tested. When testing at a 5% level, no parental effects are found to differ among boys and girls.

2016

Setting up the models the for native origin based on data from 2016 without single parents, provides maximum likelihood estimates which are much like those we saw when single parents where included in the dataset. Having in mind, that the lowest level of education and income are hidden in the intercept, all increases in parental level of education and income still imply higher probability of high school education. A closer look at the table which summarizes the models effect of discriminating between individuals having graduated from high school and who have notprovides the information that the models classify the individuals correctly 73.6% and 72.5% (for respectively boys and girls) of the time. And the model have improved a tiny bit for the boys, compared to what is seen in the case where single parents are included. The classification tables, which tells whether the model has any ”weak spots”, i.e. predicts wrongly in one direction, shows what we have seen before therefore the model estimated for girls still have some troubles classifying girls with an education correctly. Therefore no changes in the effect plots are not surprisingly. Remember that the relatively change in the native origin data from 2016 was very small. Further no changes in differences in parental effects between boys and girls are found.

Based on the estimated models, the changes in parental effects on the probability of education seems quite unchanged after removing all individuals without both parents registered in Denmark. A closer look at the stereotype families will provide a more detailed conclusion.
For girls the predicted probability of education are found as:

Table B.4: Predicted probabilities of high school education for girls based on stereotype families without single parents

Parents levels	"Stereotype Fam"	NA 06	NA16	IM06	IM16	Δ NA	Δ IM
Levels: 1	1	27.73%	27.18%	28.95%	41.21%	-1.98%	42.35%
Levels: 2	2	46.51%	54.17%	35.55%	53.79%	16.47%	51.31%
Levels: 3	3	79.45%	82.51%	43.84%	60.04%	3.85%	36.95%
Levels: 4	4	82.51%	86.71%	66.26%	77.63%	5.09%	17.16%
Levels: 5	5	92.27%	93.71%	87.70%	84.08%	1.56%	-4.13%
Levels: 5/6	6	96.49%	95.81%	82.38%	86.96%	-0.70%	5.56%
Dad Edu: 3	7	48.25%	46.83%	35.99%	45.27%	-2.94%	25.78%
Mum Edu: 3	8	50.05%	47.61%	37.85%	51.47%	-4.88%	35.98%
Mum Edu: 6	9	63.04%	56.63%	37.74%	60.03%	-10.17%	59.06%
Dad Edu: 6	10	61.36%	52.46%	38.89%	52.22%	-14.50%	34.28%
Dad Inc: 4	11	38.01%	40.42%	33.85%	53.37%	6.34%	57.67%
Mum Inc: 4	12	38.33%	42.15%	40.80%	52.94%	9.97%	29.75%

For girls with native origin, the changes are to small for event commenting on. Among girls with immigrant origin some few, but no dramatic changes have been found. One sees that the a girl from every "stereotype family" has a higher probability of getting a high school education. Furthermore one can see, that the development in the probability of education over time, has increased less than in the benchmark scenario, which implies that the parental effects on the probability of education for girls with immigrant origin and both parents registered in Denmark higher. Moreover, it is found for that parental income is just as important, as parental level of education for girls with immigrant origin.

Table B.5: Relative changes in the predicted probability of education among girls (only one parent registered in DK *not* included) when one parental level changes, and all other parental levels are set to 1

"Stereotype Family"	Changing effect	Δ NA 06	Δ NA16	ΔIM 06	Δ IM16
7	Dad Edu=3	74.00%	72.30%	24.32%	9.85%
8	Mum Edu=3	80.49%	75.17%	30.74%	24.90%
9	Mum Edu=6	127.34%	108.35%	30.36%	45.67%
10	Dad Edu=6	121.28%	93.01%	34.34%	26.72%
11	Dad Inc=4	37.07%	48.71%	16.93%	29.51%
12	Mum Inc=4	38.23%	55.08%	40.93%	28.46%

As the predicted probability for a girl from stereotype family 1 , i.e. a girl from a family with absolute lowest level of income and education, are increased relatively more than the probability of education in the other stereotype families this implies that the changes in parental effects, seems to have less impact on the probability of education.

The predicted probabilities of education among boys allows an identical analysis for boys with native origin, as in benchmark: Highly educated parents (level 6) does not raise the probability of education as much in 2016, as it did in 2006, but the effect is still quite large.

Table B.6: Predicted probabilities of high school education for boys based on stereotype families without single parents

Parental levels	"Stereotype Fam."	NA 06	NA16	IM06	IM16	Δ NA	Δ IM
Levels: 1	1	17.83%	16.94%	20.00%	26.58%	-4.99%	32.90%
Levels: 2	2	26.93%	32.81%	23.74%	41.31%	21.83%	74.01%
Levels: 3	3	65.85%	72.74%	34.78%	42.75%	10.46%	22.92%
Levels: 4	4	69.11%	73.31%	54.66%	67.81%	6.08%	24.06%
Levels: 5	5	87.74%	90.06%	78.21%	83.99%	2.64%	7.39%
Levels: 5/6	6	95.26%	94.43%	75.20%	84.06%	-0.87%	11.78%
Dad Edu: 3	7	38.41%	36.25%	27.33%	31.82%	-5.62%	16.43%
Mum Edu: 3	8	35.02%	33.77%	27.26%	32.77%	-3.57%	20.21%
Mum Edu: 6	9	50.49%	42.15%	28.67%	37.78%	-16.52%	31.78%
Dad Edu: 6	10	51.67%	43.08%	32.23%	42.12%	-16.62%	30.69%
Dad Inc: 4	11	24.96%	26.52%	25.58%	36.92%	6.25%	44.33%
Mum Inc: 4	12	23.74%	27.21%	27.53%	36.46%	14.62%	32.44%

For boys with immigrant origin, it is found that the probabilities of education are increased no matter what stereotype family he comes from, but especially if the boy comes from stereotype family 1. The relatively effect on the probability given a change in one parental effect are reduced, but this is probably caused by the increase in the probability of education for a boy from stereotype family 1

Table B.7: Relative changes in the predicted probability of education among boys (only one parent registered in DK *not* included) when one parental level changes, and all other parental levels are set to 1

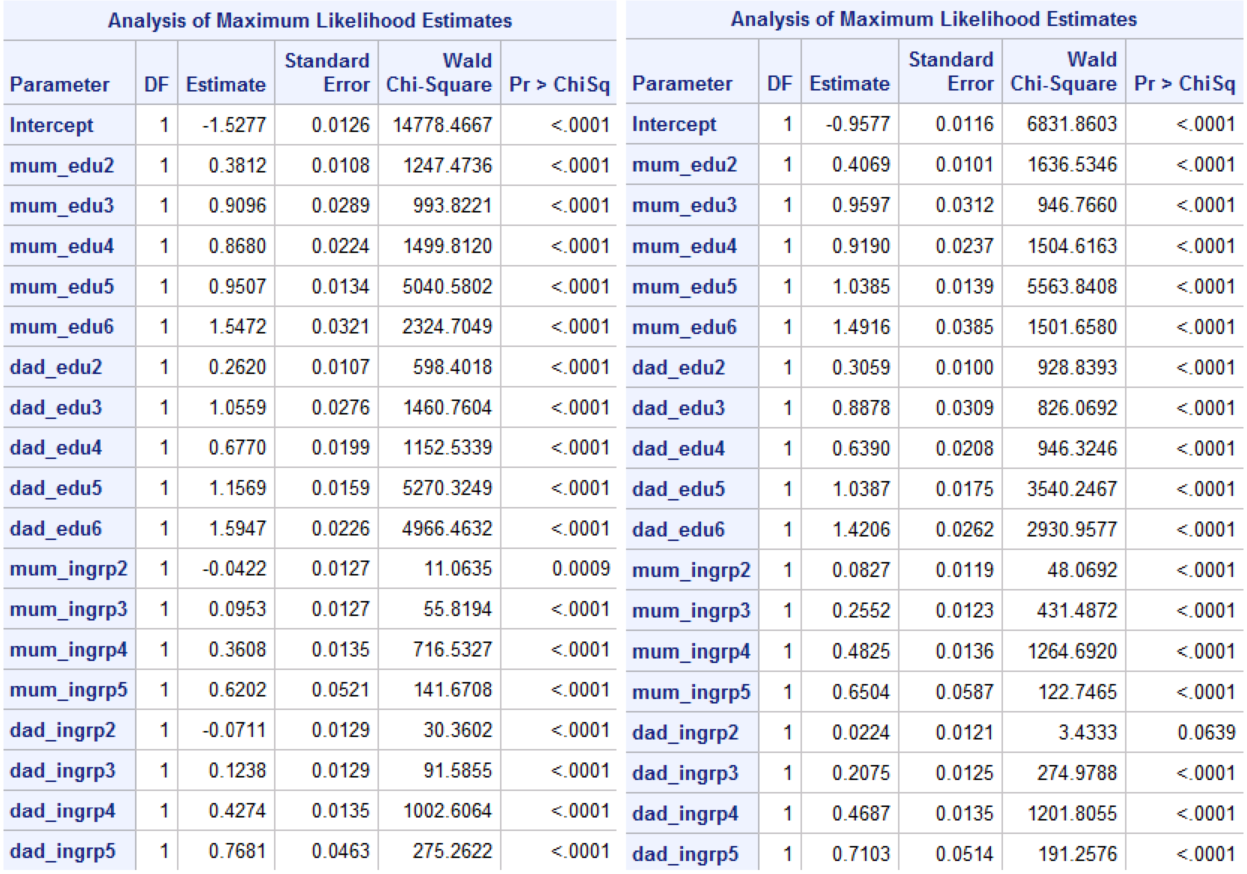
"Boy"	Changing effect	Δ NA 06	Δ NA16	ΔIM 06	Δ IM16
7	Dad Edu: 3	115.42%	113.99%	36.65%	19.71%
8	Mum Edu: 3	96.41%	99.35%	36.30%	23.29%
9	Mum Edu: 6	183.17%	148.82%	43.35%	42.14%
10	Dad Edu: 6	189.79%	154.31%	61.15%	58.47%
11	Dad Inc: 4	39.99%	56.55%	27.90%	38.90%
12	Mum Inc: 4	33.15%	60.63%	37.65%	37.17%

Comparing the parental effects on the probability of education among boys and girls with immigrant origin, one finds that, when removing all individuals without both parents registered in Denmark, girls over the years have become more sensitive to parental levels of education and income. The case is mutual to the findings in benchmark, but the relative changes with an increase in one level of parental effect, has less impact to the probability of education.

From these comments one can conclude that the estimated models are therefore quite robust to whether both parents are registered in Denmark or not. Which leads the the next robustness test: Does the parental effects on the probability of their children are getting an education depend on whether parents are living together or not?

2006

Figure B.1: Left Table: Boys, Right Table: Girls, native origin data form 2006 with both parents registered in DK



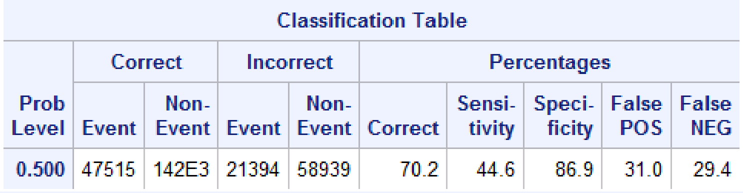
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.5277	0.0126	14778.4667	<.0001
mum_edu2	1	0.3812	0.0108	1247.4736	<.0001
mum_edu3	1	0.9096	0.0289	993.8221	<.0001
mum_edu4	1	0.8680	0.0224	1499.8120	<.0001
mum_edu5	1	0.9507	0.0134	5040.5802	<.0001
mum_edu6	1	1.5472	0.0321	2324.7049	<.0001
dad_edu2	1	0.2620	0.0107	598.4018	<.0001
dad_edu3	1	1.0559	0.0276	1460.7604	<.0001
dad_edu4	1	0.6770	0.0199	1152.5339	<.0001
dad_edu5	1	1.1569	0.0159	5270.3249	<.0001
dad_edu6	1	1.5947	0.0226	4966.4632	<.0001
mum_ingrp2	1	-0.0422	0.0127	11.0635	0.0009
mum_ingrp3	1	0.0953	0.0127	55.8194	<.0001
mum_ingrp4	1	0.3608	0.0135	716.5327	<.0001
mum_ingrp5	1	0.6202	0.0521	141.6708	<.0001
dad_ingrp2	1	-0.0711	0.0129	30.3602	<.0001
dad_ingrp3	1	0.1238	0.0129	91.5855	<.0001
dad_ingrp4	1	0.4274	0.0135	1002.6064	<.0001
dad_ingrp5	1	0.7681	0.0463	275.2622	<.0001

Analysis of Maximum Likelihood Estimates

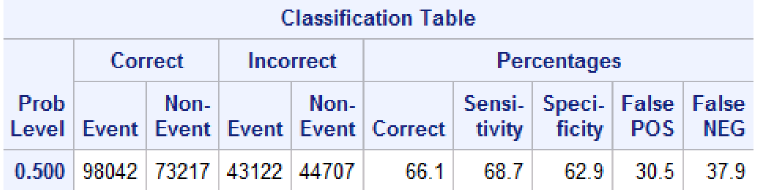
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.9577	0.0116	6831.8603	<.0001
mum_edu2	1	0.4069	0.0101	1636.5346	<.0001
mum_edu3	1	0.9597	0.0312	946.7660	<.0001
mum_edu4	1	0.9190	0.0237	1504.6163	<.0001
mum_edu5	1	1.0385	0.0139	5563.8408	<.0001
mum_edu6	1	1.4916	0.0385	1501.6580	<.0001
dad_edu2	1	0.3059	0.0100	928.8393	<.0001
dad_edu3	1	0.8878	0.0309	826.0692	<.0001
dad_edu4	1	0.6390	0.0208	946.3246	<.0001
dad_edu5	1	1.0387	0.0175	3540.2467	<.0001
dad_edu6	1	1.4206	0.0262	2930.9577	<.0001
mum_ingrp2	1	0.0827	0.0119	48.0692	<.0001
mum_ingrp3	1	0.2552	0.0123	431.4872	<.0001
mum_ingrp4	1	0.4825	0.0136	1264.6920	<.0001
mum_ingrp5	1	0.6504	0.0587	122.7465	<.0001
dad_ingrp2	1	0.0224	0.0121	3.4333	0.0639
dad_ingrp3	1	0.2075	0.0125	274.9788	<.0001
dad_ingrp4	1	0.4687	0.0135	1201.8055	<.0001
dad_ingrp5	1	0.7103	0.0514	191.2576	<.0001

Figure B.2: Left Plot: Boys, Right Plot: Girls. Native origin, data form 2006 with both parents registered in DK



Classification Table

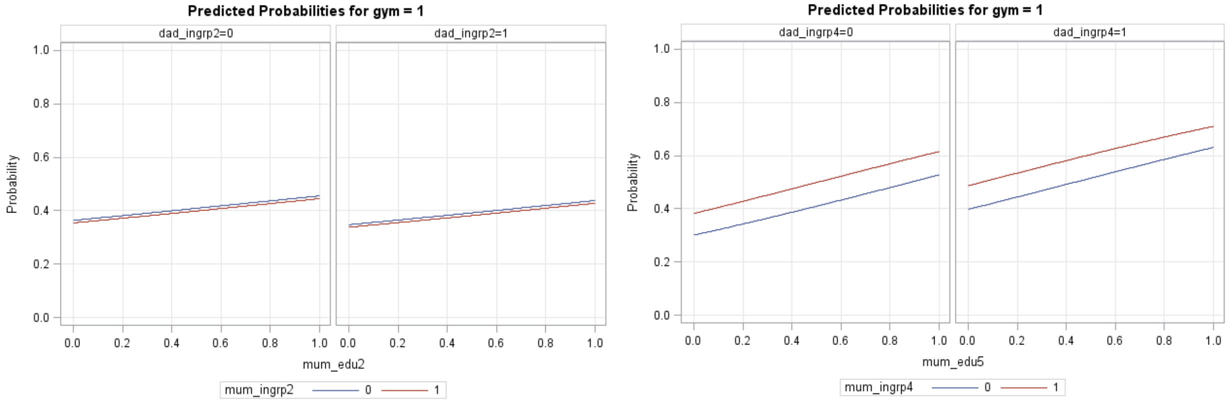
	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	47515	142E3	21394	58939	70.2	44.6	86.9	31.0	29.4



Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	98042	73217	43122	44707	66.1	68.7	62.9	30.5	37.9

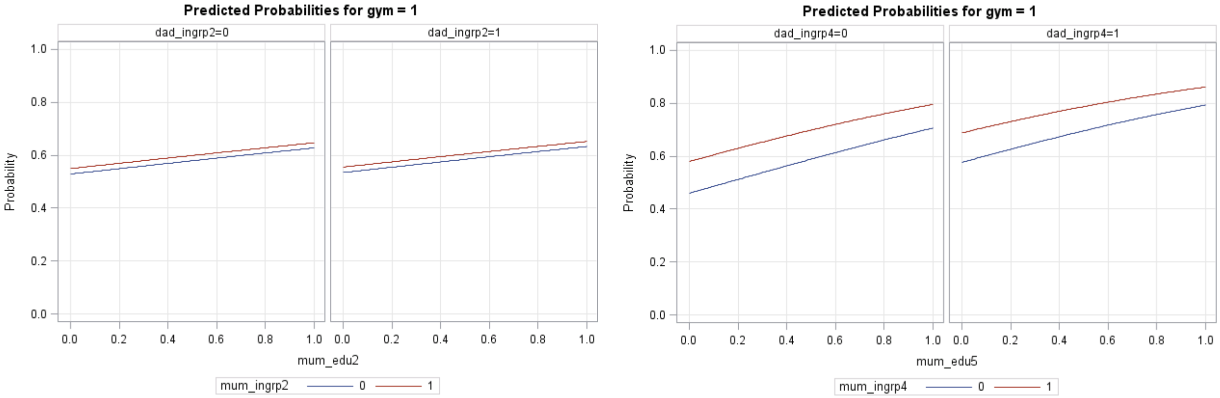
Figure B.3: Predicted probability plots for boys with native origin, data form 2016 with both parents registered in DK



Fit computed at mum_edu3=0.022 mum_edu4=0.038 mum_edu5=0.205 mum_edu6=0.035 dad_edu2=0.408 dad_edu3=0.025 dad_edu4=0.05 dad_edu5=0.111 dad_edu6=0.07 mum_ingrp3=0.25 mum_ingrp4=0.24 mum_ingrp5=0.01 dad_ingrp3=0.253 dad_ingrp4=0.241 dad_ingrp5=0.01

Fit computed at mum_edu2=0.342 mum_edu3=0.022 mum_edu4=0.038 mum_edu6=0.035 dad_edu2=0.408 dad_edu3=0.025 dad_edu4=0.05 dad_edu5=0.111 dad_edu6=0.07 mum_ingrp2=0.25 mum_ingrp3=0.25 mum_ingrp5=0.01 dad_ingrp2=0.251 dad_ingrp3=0.253 dad_ingrp5=0.01

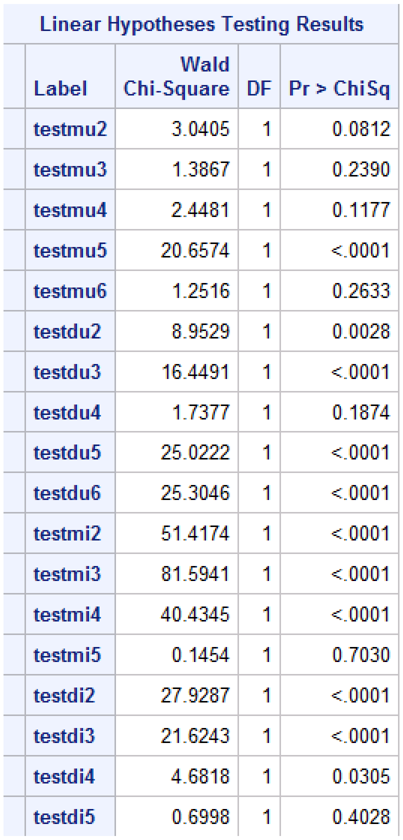
Figure B.4: Predicted probability plots for girls with native origin, data form 2016 with both parents registered in DK



Fit computed at mum_edu3=0.022 mum_edu4=0.039 mum_edu5=0.204 mum_edu6=0.035 dad_edu2=0.407 dad_edu3=0.024 dad_edu4=0.049 dad_edu5=0.111 dad_edu6=0.071 mum_ingrp3=0.25 mum_ingrp4=0.241 mum_ingrp5=0.01 dad_ingrp3=0.251 dad_ingrp4=0.242 dad_ingrp5=0.01

Fit computed at mum_edu2=0.342 mum_edu3=0.022 mum_edu4=0.039 mum_edu6=0.035 dad_edu2=0.407 dad_edu3=0.024 dad_edu4=0.049 dad_edu5=0.111 dad_edu6=0.071 mum_ingrp2=0.25 mum_ingrp3=0.25 mum_ingrp5=0.01 dad_ingrp2=0.253 dad_ingrp3=0.251 dad_ingrp5=0.01

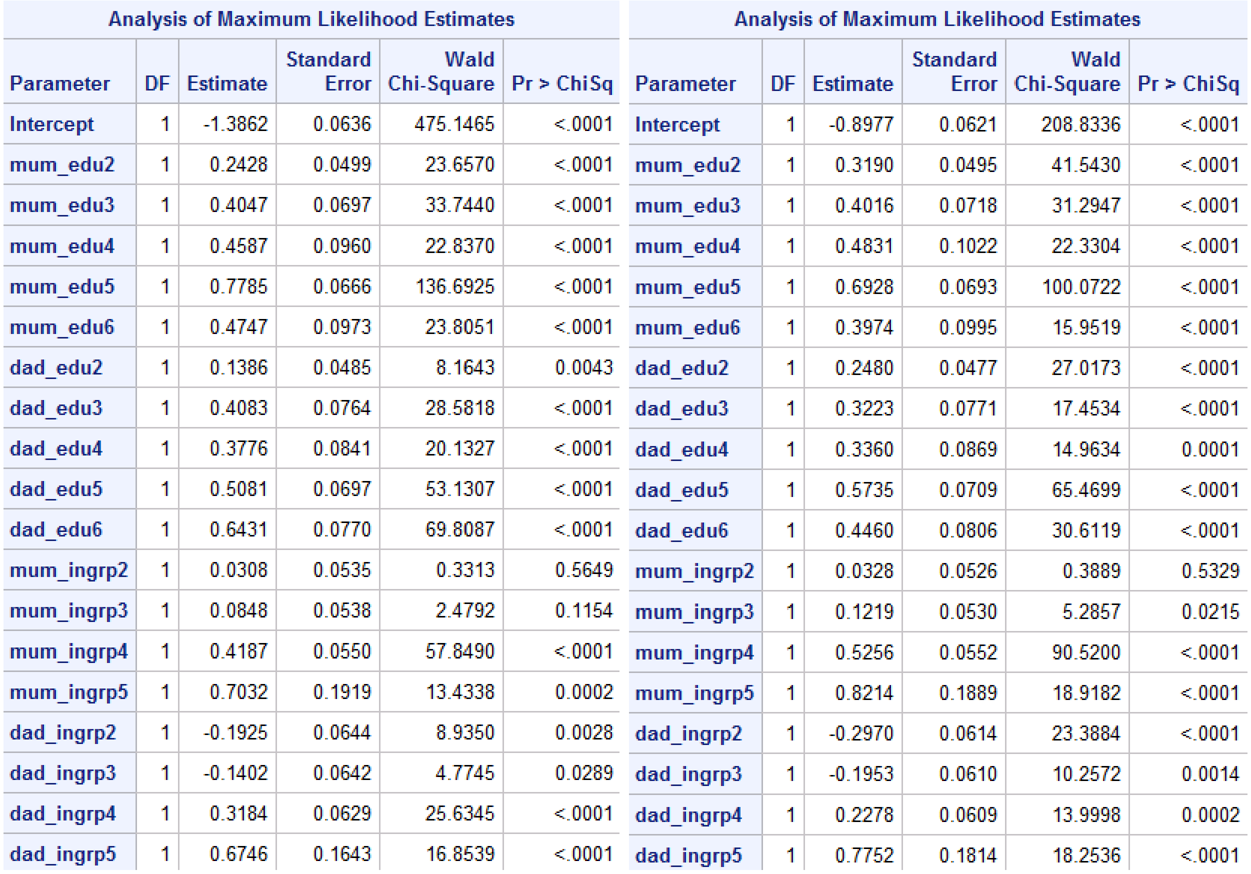
Figure B.5: Test $\beta_i^{\text{Boys}} = \beta_i^{\text{Girls}}$, where i is a parental level of education or income. Native origin, data form 2006 with both parents registered in DK



Linear Hypotheses Testing Results

Label	Wald Chi-Square	DF	Pr > ChiSq
testmu2	3.0405	1	0.0812
testmu3	1.3867	1	0.2390
testmu4	2.4481	1	0.1177
testmu5	20.6574	1	<.0001
testmu6	1.2516	1	0.2633
testdu2	8.9529	1	0.0028
testdu3	16.4491	1	<.0001
testdu4	1.7377	1	0.1874
testdu5	25.0222	1	<.0001
testdu6	25.3046	1	<.0001
testmi2	51.4174	1	<.0001
testmi3	81.5941	1	<.0001
testmi4	40.4345	1	<.0001
testmi5	0.1454	1	0.7030
testdi2	27.9287	1	<.0001
testdi3	21.6243	1	<.0001
testdi4	4.6818	1	0.0305
testdi5	0.6998	1	0.4028

Figure B.6: Left Table: Boys, Right Table: Girls. Immigrant origin, data form 2006 with both parents registered in DK



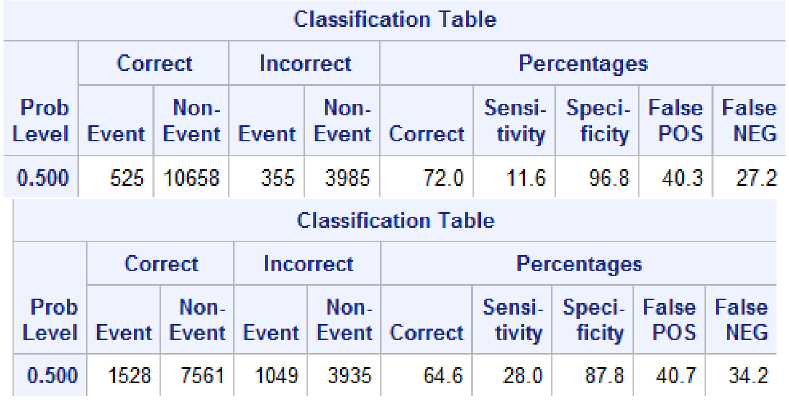
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.3862	0.0636	475.1465	<.0001
mum_edu2	1	0.2428	0.0499	23.6570	<.0001
mum_edu3	1	0.4047	0.0697	33.7440	<.0001
mum_edu4	1	0.4587	0.0960	22.8370	<.0001
mum_edu5	1	0.7785	0.0666	136.6925	<.0001
mum_edu6	1	0.4747	0.0973	23.8051	<.0001
dad_edu2	1	0.1386	0.0485	8.1643	0.0043
dad_edu3	1	0.4083	0.0764	28.5818	<.0001
dad_edu4	1	0.3776	0.0841	20.1327	<.0001
dad_edu5	1	0.5081	0.0697	53.1307	<.0001
dad_edu6	1	0.6431	0.0770	69.8087	<.0001
mum_ingrp2	1	0.0308	0.0535	0.3313	0.5649
mum_ingrp3	1	0.0848	0.0538	2.4792	0.1154
mum_ingrp4	1	0.4187	0.0550	57.8490	<.0001
mum_ingrp5	1	0.7032	0.1919	13.4338	0.0002
dad_ingrp2	1	-0.1925	0.0644	8.9350	0.0028
dad_ingrp3	1	-0.1402	0.0642	4.7745	0.0289
dad_ingrp4	1	0.3184	0.0629	25.6345	<.0001
dad_ingrp5	1	0.6746	0.1643	16.8539	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.8977	0.0621	208.8336	<.0001
mum_edu2	1	0.3190	0.0495	41.5430	<.0001
mum_edu3	1	0.4016	0.0718	31.2947	<.0001
mum_edu4	1	0.4831	0.1022	22.3304	<.0001
mum_edu5	1	0.6928	0.0693	100.0722	<.0001
mum_edu6	1	0.3974	0.0995	15.9519	<.0001
dad_edu2	1	0.2480	0.0477	27.0173	<.0001
dad_edu3	1	0.3223	0.0771	17.4534	<.0001
dad_edu4	1	0.3360	0.0869	14.9634	0.0001
dad_edu5	1	0.5735	0.0709	65.4699	<.0001
dad_edu6	1	0.4460	0.0806	30.6119	<.0001
mum_ingrp2	1	0.0328	0.0526	0.3889	0.5329
mum_ingrp3	1	0.1219	0.0530	5.2857	0.0215
mum_ingrp4	1	0.5256	0.0552	90.5200	<.0001
mum_ingrp5	1	0.8214	0.1889	18.9182	<.0001
dad_ingrp2	1	-0.2970	0.0614	23.3884	<.0001
dad_ingrp3	1	-0.1953	0.0610	10.2572	0.0014
dad_ingrp4	1	0.2278	0.0609	13.9998	0.0002
dad_ingrp5	1	0.7752	0.1814	18.2536	<.0001

Figure B.7: Top Table: Boys, Bottom Table: Girls. Immigrant origin, data form 2006 with both parents registered in DK



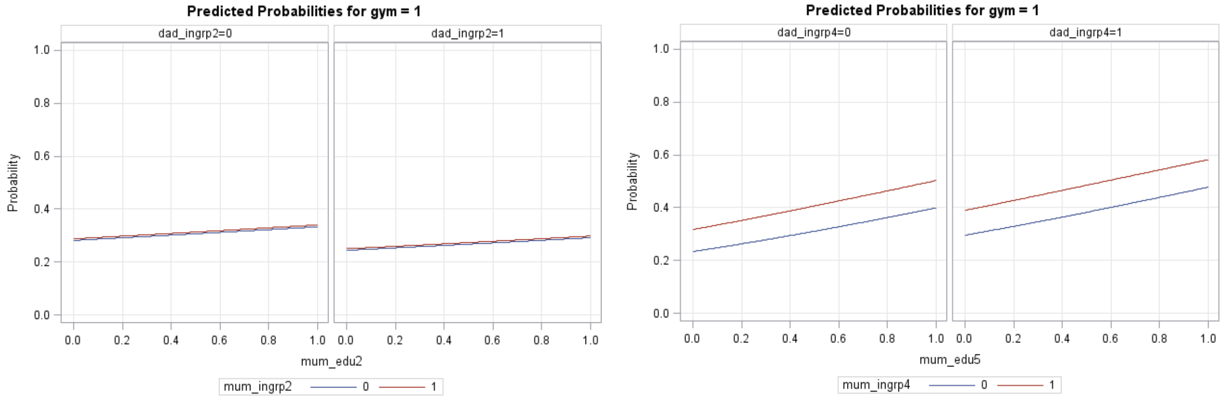
Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	525	10658	355	3985	72.0	11.6	96.8	40.3	27.2

Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	1528	7561	1049	3935	64.6	28.0	87.8	40.7	34.2

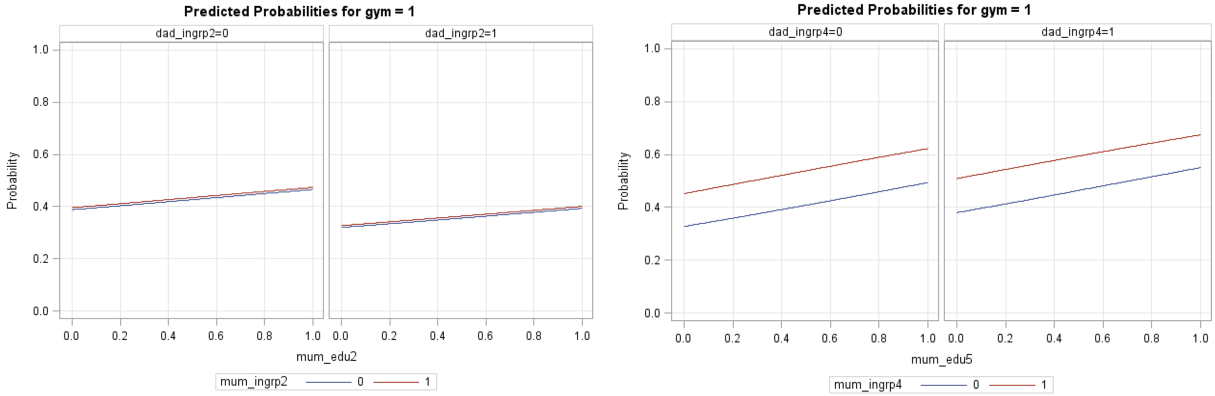
Figure B.8: Predicted probability plots for boys with immigrant origin, data form 2006 with both parents registered in DK



Fit computed at mum_edu3=0.072 mum_edu4=0.034 mum_edu5=0.078 mum_edu6=0.034 dad_edu2=0.221 dad_edu3=0.06 dad_edu4=0.048 dad_edu5=0.074 dad_edu6=0.059 mum_ingrp3=0.253 mum_ingrp4=0.227 mum_ingrp5=0.008 dad_ingrp3=0.286 dad_ingrp4=0.28 dad_ingrp5=0.012

Fit computed at mum_edu2=0.181 mum_edu3=0.072 mum_edu4=0.034 mum_edu6=0.034 dad_edu2=0.221 dad_edu3=0.06 dad_edu4=0.048 dad_edu5=0.074 dad_edu6=0.059 mum_ingrp2=0.271 mum_ingrp3=0.253 mum_ingrp5=0.008 dad_ingrp2=0.294 dad_ingrp3=0.286 dad_ingrp5=0.012

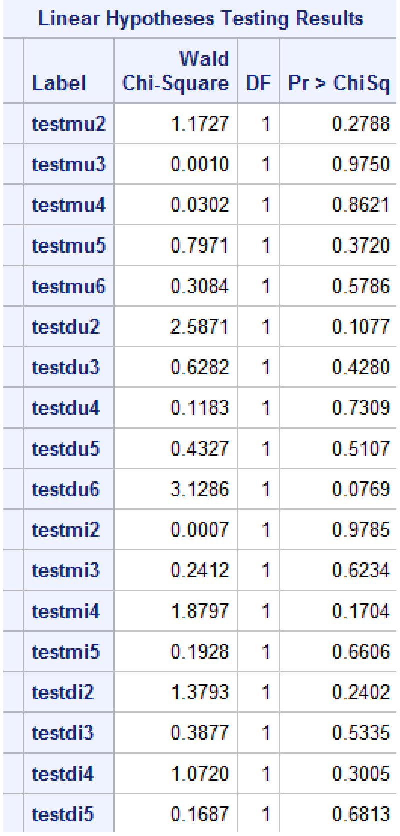
Figure B.9: Predicted probability plots for girls with immigrant origin, data form 2006, with both parents registered in DK



Fit computed at mum_edu3=0.067 mum_edu4=0.031 mum_edu5=0.078 mum_edu6=0.035 dad_edu2=0.21 dad_edu3=0.059 dad_edu4=0.045 dad_edu5=0.073 dad_edu6=0.058 mum_ingrp3=0.262 mum_ingrp4=0.233 mum_ingrp5=0.01 dad_ingrp3=0.29 dad_ingrp4=0.275 dad_ingrp5=0.012

Fit computed at mum_edu2=0.174 mum_edu3=0.067 mum_edu4=0.031 mum_edu6=0.035 dad_edu2=0.21 dad_edu3=0.059 dad_edu4=0.045 dad_edu5=0.073 dad_edu6=0.058 mum_ingrp2=0.276 mum_ingrp3=0.262 mum_ingrp5=0.01 dad_ingrp2=0.291 dad_ingrp3=0.29 dad_ingrp5=0.012

Figure B.10: Test $\beta_i^{\text{Boys}} = \beta_i^{\text{Girls}}$, where i is a parental level of education or income. Immigrant origin, data form 2006

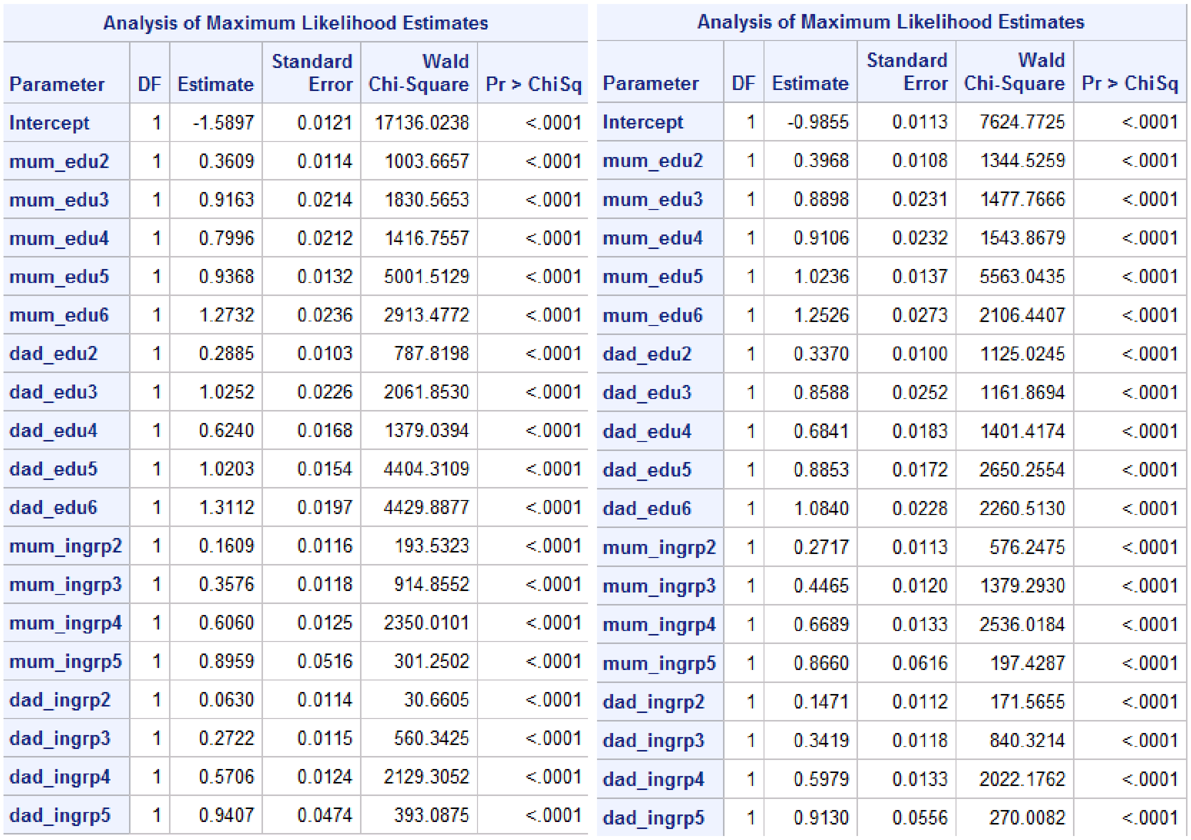


Linear Hypotheses Testing Results

Label	Wald Chi-Square	DF	Pr > ChiSq
testmu2	1.1727	1	0.2788
testmu3	0.0010	1	0.9750
testmu4	0.0302	1	0.8621
testmu5	0.7971	1	0.3720
testmu6	0.3084	1	0.5786
testdu2	2.5871	1	0.1077
testdu3	0.6282	1	0.4280
testdu4	0.1183	1	0.7309
testdu5	0.4327	1	0.5107
testdu6	3.1286	1	0.0769
testmi2	0.0007	1	0.9785
testmi3	0.2412	1	0.6234
testmi4	1.8797	1	0.1704
testmi5	0.1928	1	0.6606
testdi2	1.3793	1	0.2402
testdi3	0.3877	1	0.5335
testdi4	1.0720	1	0.3005
testdi5	0.1687	1	0.6813

2016

Figure B.11: Left Table: Boys, Right Table: Girls, native origin data form 2016 with both parents registered in DK



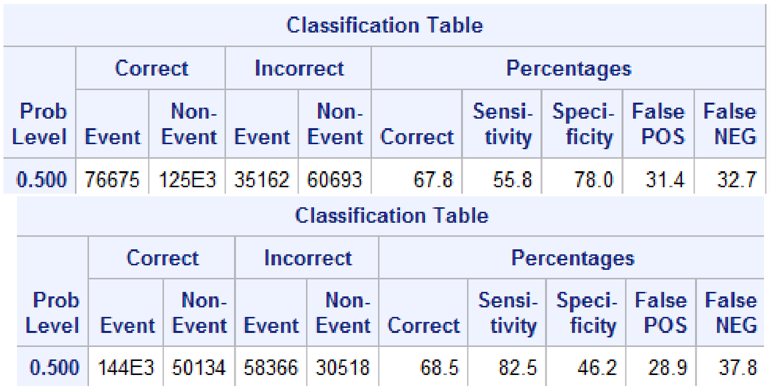
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.5897	0.0121	17136.0238	<.0001
mum_edu2	1	0.3609	0.0114	1003.6657	<.0001
mum_edu3	1	0.9163	0.0214	1830.5653	<.0001
mum_edu4	1	0.7996	0.0212	1416.7557	<.0001
mum_edu5	1	0.9368	0.0132	5001.5129	<.0001
mum_edu6	1	1.2732	0.0236	2913.4772	<.0001
dad_edu2	1	0.2885	0.0103	787.8198	<.0001
dad_edu3	1	1.0252	0.0226	2061.8530	<.0001
dad_edu4	1	0.6240	0.0168	1379.0394	<.0001
dad_edu5	1	1.0203	0.0154	4404.3109	<.0001
dad_edu6	1	1.3112	0.0197	4429.8877	<.0001
mum_ingrp2	1	0.1609	0.0116	193.5323	<.0001
mum_ingrp3	1	0.3576	0.0118	914.8552	<.0001
mum_ingrp4	1	0.6060	0.0125	2350.0101	<.0001
mum_ingrp5	1	0.8959	0.0516	301.2502	<.0001
dad_ingrp2	1	0.0630	0.0114	30.6605	<.0001
dad_ingrp3	1	0.2722	0.0115	560.3425	<.0001
dad_ingrp4	1	0.5706	0.0124	2129.3052	<.0001
dad_ingrp5	1	0.9407	0.0474	393.0875	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.9855	0.0113	7624.7725	<.0001
mum_edu2	1	0.3968	0.0108	1344.5259	<.0001
mum_edu3	1	0.8898	0.0231	1477.7666	<.0001
mum_edu4	1	0.9106	0.0232	1543.8679	<.0001
mum_edu5	1	1.0236	0.0137	5563.0435	<.0001
mum_edu6	1	1.2526	0.0273	2106.4407	<.0001
dad_edu2	1	0.3370	0.0100	1125.0245	<.0001
dad_edu3	1	0.8588	0.0252	1161.8694	<.0001
dad_edu4	1	0.6841	0.0183	1401.4174	<.0001
dad_edu5	1	0.8853	0.0172	2650.2554	<.0001
dad_edu6	1	1.0840	0.0228	2260.5130	<.0001
mum_ingrp2	1	0.2717	0.0113	576.2475	<.0001
mum_ingrp3	1	0.4465	0.0120	1379.2930	<.0001
mum_ingrp4	1	0.6689	0.0133	2536.0184	<.0001
mum_ingrp5	1	0.8660	0.0616	197.4287	<.0001
dad_ingrp2	1	0.1471	0.0112	171.5655	<.0001
dad_ingrp3	1	0.3419	0.0118	840.3214	<.0001
dad_ingrp4	1	0.5979	0.0133	2022.1762	<.0001
dad_ingrp5	1	0.9130	0.0556	270.0082	<.0001

Figure B.12: Top Table: Boys, Bottom Table: Girls, native origin data form 2016 with both parents registered in DK



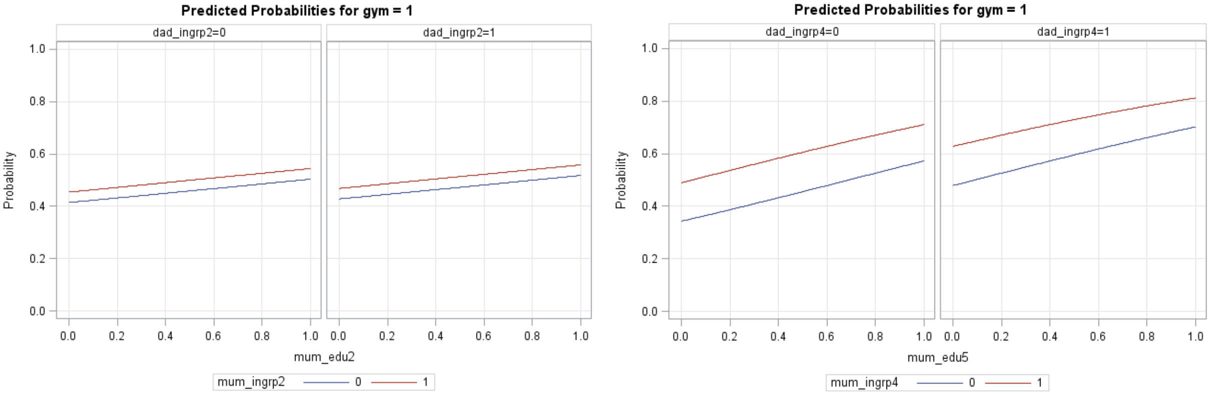
Classification Table

Prob Level	Correct		Incorrect		Percentages				
	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	76675	125E3	35162	60693	67.8	55.8	78.0	31.4	32.7

Classification Table

Prob Level	Correct		Incorrect		Percentages				
	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	144E3	50134	58366	30518	68.5	82.5	46.2	28.9	37.8

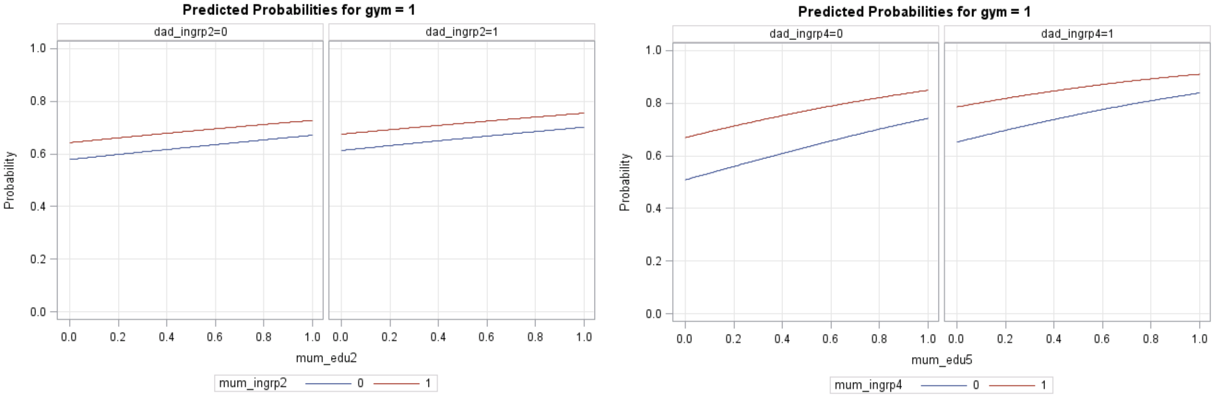
Figure B.13: Predicted probability plots for boys with native origin, data form 2016 with both parents registered in DK



Fit computed at mum_edu3=0.042 mum_edu4=0.042 mum_edu5=0.242 mum_edu6=0.062 dad_edu2=0.401 dad_edu3=0.037 dad_edu4=0.07 dad_edu5=0.114 dad_edu6=0.087 mum_ingrp3=0.241 mum_ingrp4=0.234 mum_ingrp5=0.01 dad_ingrp3=0.229 dad_ingrp4=0.222 dad_ingrp5=0.009

Fit computed at mum_edu2=0.386 mum_edu3=0.042 mum_edu4=0.042 mum_edu6=0.062 dad_edu2=0.401 dad_edu3=0.037 dad_edu4=0.07 dad_edu5=0.114 dad_edu6=0.087 mum_ingrp2=0.241 mum_ingrp3=0.241 mum_ingrp5=0.01 dad_ingrp2=0.228 dad_ingrp3=0.229 dad_ingrp5=0.009

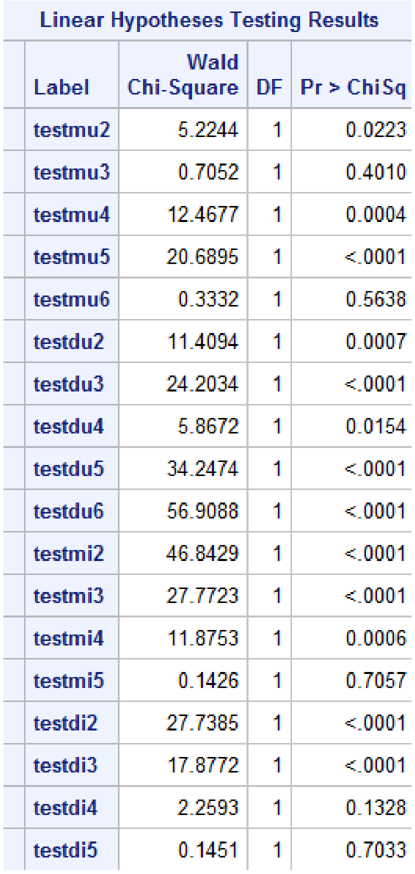
Figure B.14: Predicted probability plots for girls with native origin, data form 2016 with both parents registered in DK



Fit computed at mum_edu3=0.041 mum_edu4=0.043 mum_edu5=0.242 mum_edu6=0.061 dad_edu2=0.404 dad_edu3=0.037 dad_edu4=0.07 dad_edu5=0.113 dad_edu6=0.087 mum_ingrp3=0.242 mum_ingrp4=0.231 mum_ingrp5=0.009 dad_ingrp3=0.229 dad_ingrp4=0.221 dad_ingrp5=0.009

Fit computed at mum_edu2=0.386 mum_edu3=0.041 mum_edu4=0.043 mum_edu6=0.061 dad_edu2=0.404 dad_edu3=0.037 dad_edu4=0.07 dad_edu5=0.113 dad_edu6=0.087 mum_ingrp2=0.243 mum_ingrp3=0.242 mum_ingrp5=0.009 dad_ingrp2=0.229 dad_ingrp3=0.229 dad_ingrp5=0.009

Figure B.15: Test $\beta_i^{\text{Boys}} = \beta_i^{\text{Girls}}$, where i is a parental level of education or income. Native origin, data form 2016 with both parents registered in DK



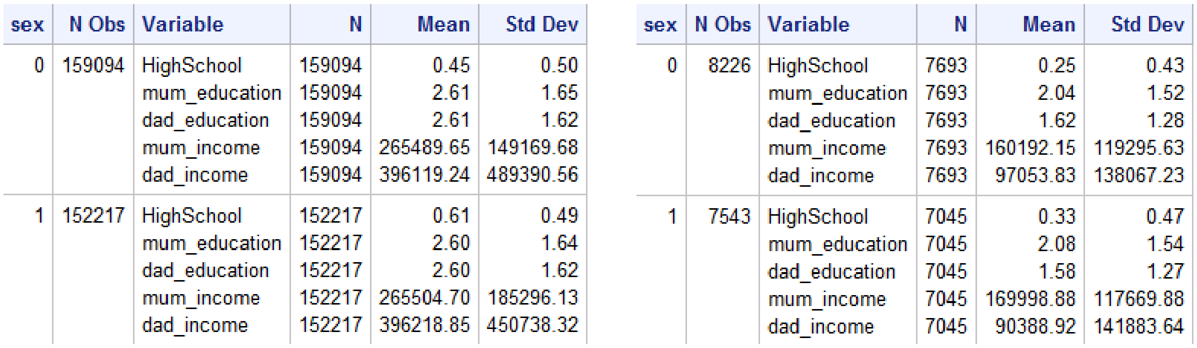
Linear Hypotheses Testing Results

Label	Wald Chi-Square	DF	Pr > ChiSq
testmu2	5.2244	1	0.0223
testmu3	0.7052	1	0.4010
testmu4	12.4677	1	0.0004
testmu5	20.6895	1	<.0001
testmu6	0.3332	1	0.5638
testdu2	11.4094	1	0.0007
testdu3	24.2034	1	<.0001
testdu4	5.8672	1	0.0154
testdu5	34.2474	1	<.0001
testdu6	56.9088	1	<.0001
testmi2	46.8429	1	<.0001
testmi3	27.7723	1	<.0001
testmi4	11.8753	1	0.0006
testmi5	0.1426	1	0.7057
testdi2	27.7385	1	<.0001
testdi3	17.8772	1	<.0001
testdi4	2.2593	1	0.1328
testdi5	0.1451	1	0.7033

Appendix C

When parents live together

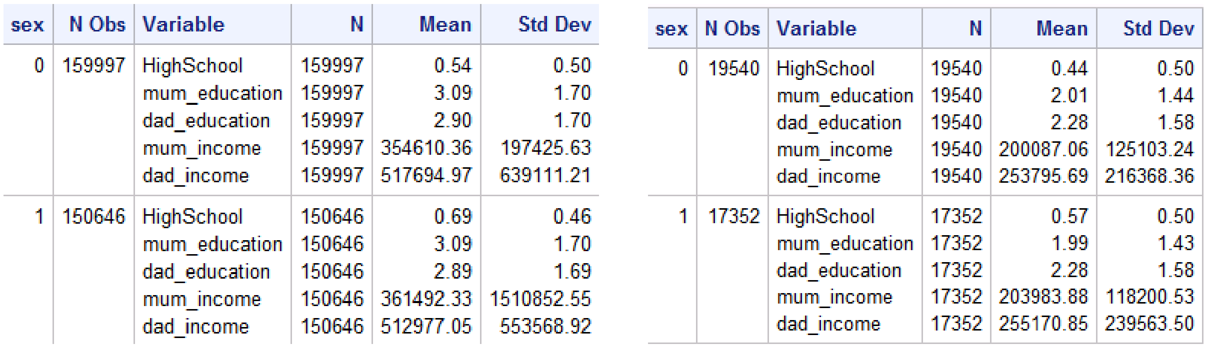
Figure C.1: Left Table: Native origin, Right Table: Immigrant origin. Summary parents living together 2006



sex	N Obs	Variable	N	Mean	Std Dev
0	159094	HighSchool	159094	0.45	0.50
		mum_education	159094	2.61	1.65
		dad_education	159094	2.61	1.62
		mum_income	159094	265489.65	149169.68
		dad_income	159094	396119.24	489390.56
1	152217	HighSchool	152217	0.61	0.49
		mum_education	152217	2.60	1.64
		dad_education	152217	2.60	1.62
		mum_income	152217	265504.70	185296.13
		dad_income	152217	396218.85	450738.32

sex	N Obs	Variable	N	Mean	Std Dev
0	8226	HighSchool	7693	0.25	0.43
		mum_education	7693	2.04	1.52
		dad_education	7693	1.62	1.28
		mum_income	7693	160192.15	119295.63
		dad_income	7693	97053.83	138067.23
1	7543	HighSchool	7045	0.33	0.47
		mum_education	7045	2.08	1.54
		dad_education	7045	1.58	1.27
		mum_income	7045	169998.88	117669.88
		dad_income	7045	90388.92	141883.64

Figure C.2: Left Table: Native origin, Right Table: Immigrant origin. Summary parents living together 2016

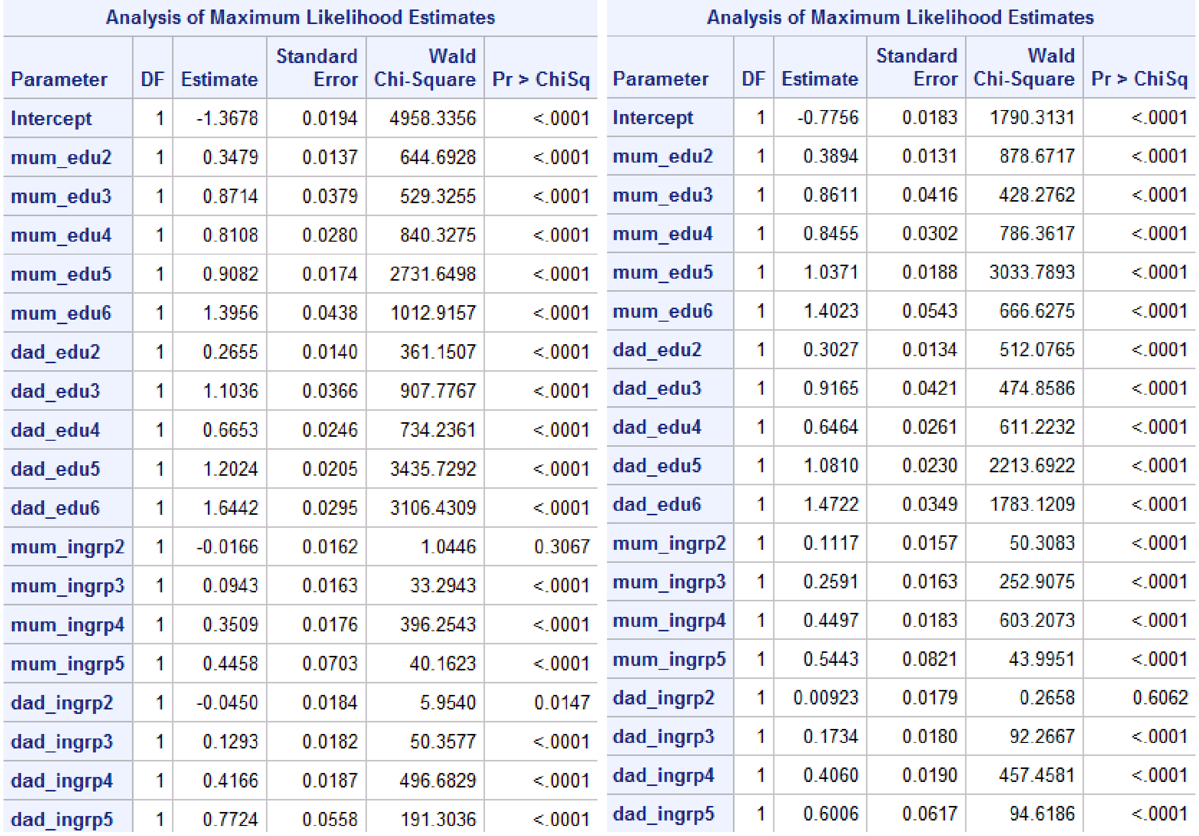


sex	N Obs	Variable	N	Mean	Std Dev
0	159997	HighSchool	159997	0.54	0.50
		mum_education	159997	3.09	1.70
		dad_education	159997	2.90	1.70
		mum_income	159997	354610.36	197425.63
		dad_income	159997	517694.97	639111.21
1	150646	HighSchool	150646	0.69	0.46
		mum_education	150646	3.09	1.70
		dad_education	150646	2.89	1.69
		mum_income	150646	361492.33	1510852.55
		dad_income	150646	512977.05	553568.92

sex	N Obs	Variable	N	Mean	Std Dev
0	19540	HighSchool	19540	0.44	0.50
		mum_education	19540	2.01	1.44
		dad_education	19540	2.28	1.58
		mum_income	19540	200087.06	125103.24
		dad_income	19540	253795.69	216368.36
1	17352	HighSchool	17352	0.57	0.50
		mum_education	17352	1.99	1.43
		dad_education	17352	2.28	1.58
		mum_income	17352	203983.88	118200.53
		dad_income	17352	255170.85	239563.50

2006

Figure C.3: Left Table: Boys, Right Table: Girls, native origin data form 2006 parents living together



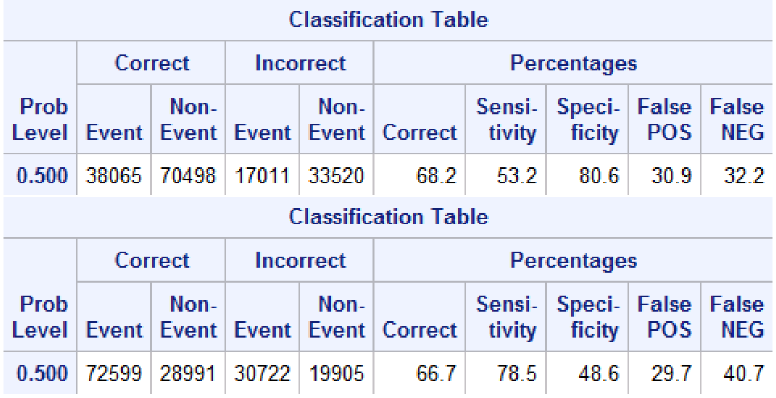
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.3678	0.0194	4958.3356	<.0001
mum_edu2	1	0.3479	0.0137	644.6928	<.0001
mum_edu3	1	0.8714	0.0379	529.3255	<.0001
mum_edu4	1	0.8108	0.0280	840.3275	<.0001
mum_edu5	1	0.9082	0.0174	2731.6498	<.0001
mum_edu6	1	1.3956	0.0438	1012.9157	<.0001
dad_edu2	1	0.2655	0.0140	361.1507	<.0001
dad_edu3	1	1.1036	0.0366	907.7767	<.0001
dad_edu4	1	0.6653	0.0246	734.2361	<.0001
dad_edu5	1	1.2024	0.0205	3435.7292	<.0001
dad_edu6	1	1.6442	0.0295	3106.4309	<.0001
mum_ingrp2	1	-0.0166	0.0162	1.0446	0.3067
mum_ingrp3	1	0.0943	0.0163	33.2943	<.0001
mum_ingrp4	1	0.3509	0.0176	396.2543	<.0001
mum_ingrp5	1	0.4458	0.0703	40.1623	<.0001
dad_ingrp2	1	-0.0450	0.0184	5.9540	0.0147
dad_ingrp3	1	0.1293	0.0182	50.3577	<.0001
dad_ingrp4	1	0.4166	0.0187	496.6829	<.0001
dad_ingrp5	1	0.7724	0.0558	191.3036	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.7756	0.0183	1790.3131	<.0001
mum_edu2	1	0.3894	0.0131	878.6717	<.0001
mum_edu3	1	0.8611	0.0416	428.2762	<.0001
mum_edu4	1	0.8455	0.0302	786.3617	<.0001
mum_edu5	1	1.0371	0.0188	3033.7893	<.0001
mum_edu6	1	1.4023	0.0543	666.6275	<.0001
dad_edu2	1	0.3027	0.0134	512.0765	<.0001
dad_edu3	1	0.9165	0.0421	474.8586	<.0001
dad_edu4	1	0.6464	0.0261	611.2232	<.0001
dad_edu5	1	1.0810	0.0230	2213.6922	<.0001
dad_edu6	1	1.4722	0.0349	1783.1209	<.0001
mum_ingrp2	1	0.1117	0.0157	50.3083	<.0001
mum_ingrp3	1	0.2591	0.0163	252.9075	<.0001
mum_ingrp4	1	0.4497	0.0183	603.2073	<.0001
mum_ingrp5	1	0.5443	0.0821	43.9951	<.0001
dad_ingrp2	1	0.00923	0.0179	0.2658	0.6062
dad_ingrp3	1	0.1734	0.0180	92.2667	<.0001
dad_ingrp4	1	0.4060	0.0190	457.4581	<.0001
dad_ingrp5	1	0.6006	0.0617	94.6186	<.0001

Figure C.4: Top Table: Boys, Bottom Table: Girls. Native origin, data form 2006 parents living together



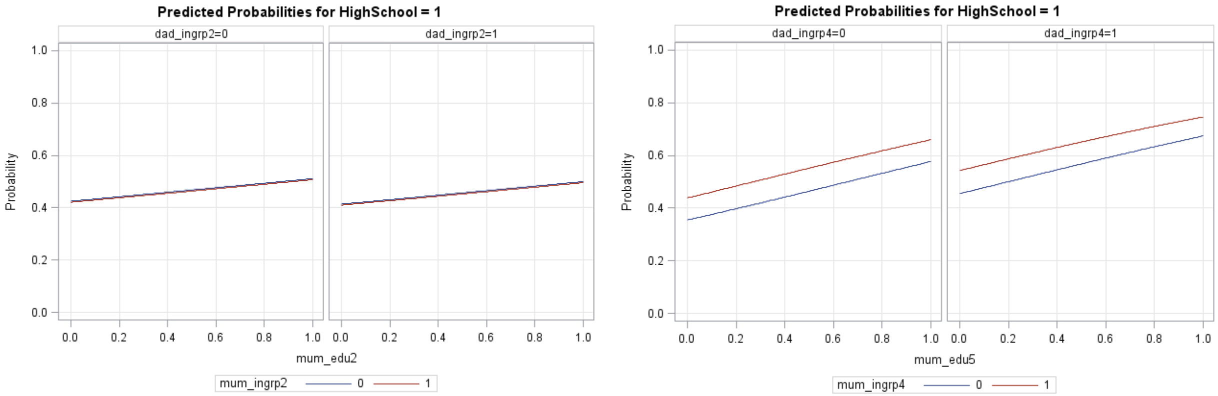
Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	38065	70498	17011	33520	68.2	53.2	80.6	30.9	32.2

Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	72599	28991	30722	19905	66.7	78.5	48.6	29.7	40.7

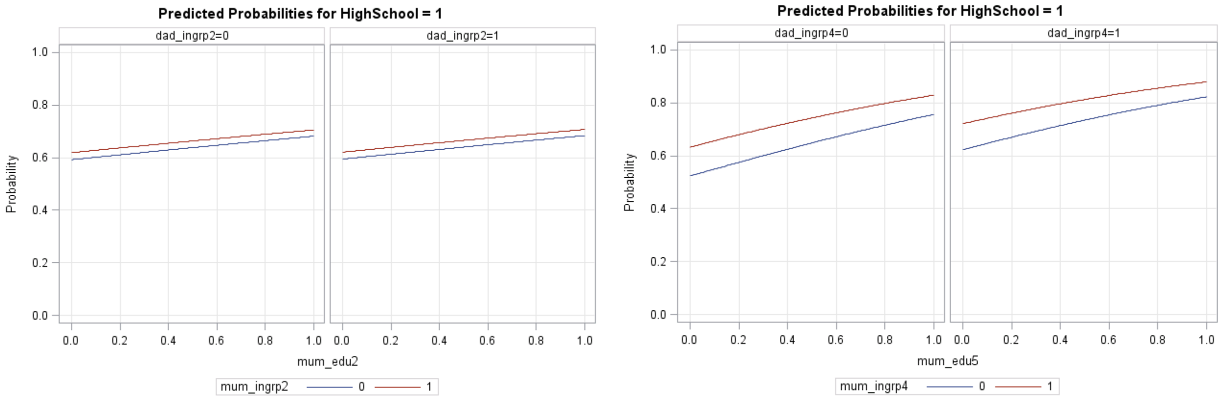
Figure C.5: Predicted probability plots for boys with native origin, data form 2006 with parents living together



Fit computed at mum_edu3=0.023 mum_edu4=0.043 mum_edu5=0.221 mum_edu6=0.037 dad_edu2=0.457 dad_edu3=0.025 dad_edu4=0.058 dad_edu5=0.129 dad_edu6=0.081 mum_ingrp3=0.264 mum_ingrp4=0.243 mum_ingrp5=0.01 dad_ingrp3=0.29 dad_ingrp4=0.29 dad_ingrp5=0.012

Fit computed at mum_edu2=0.369 mum_edu3=0.023 mum_edu4=0.043 mum_edu6=0.037 dad_edu2=0.457 dad_edu3=0.025 dad_edu4=0.058 dad_edu5=0.129 dad_edu6=0.081 mum_ingrp2=0.262 mum_ingrp3=0.264 mum_ingrp5=0.01 dad_ingrp2=0.264 dad_ingrp3=0.29 dad_ingrp5=0.012

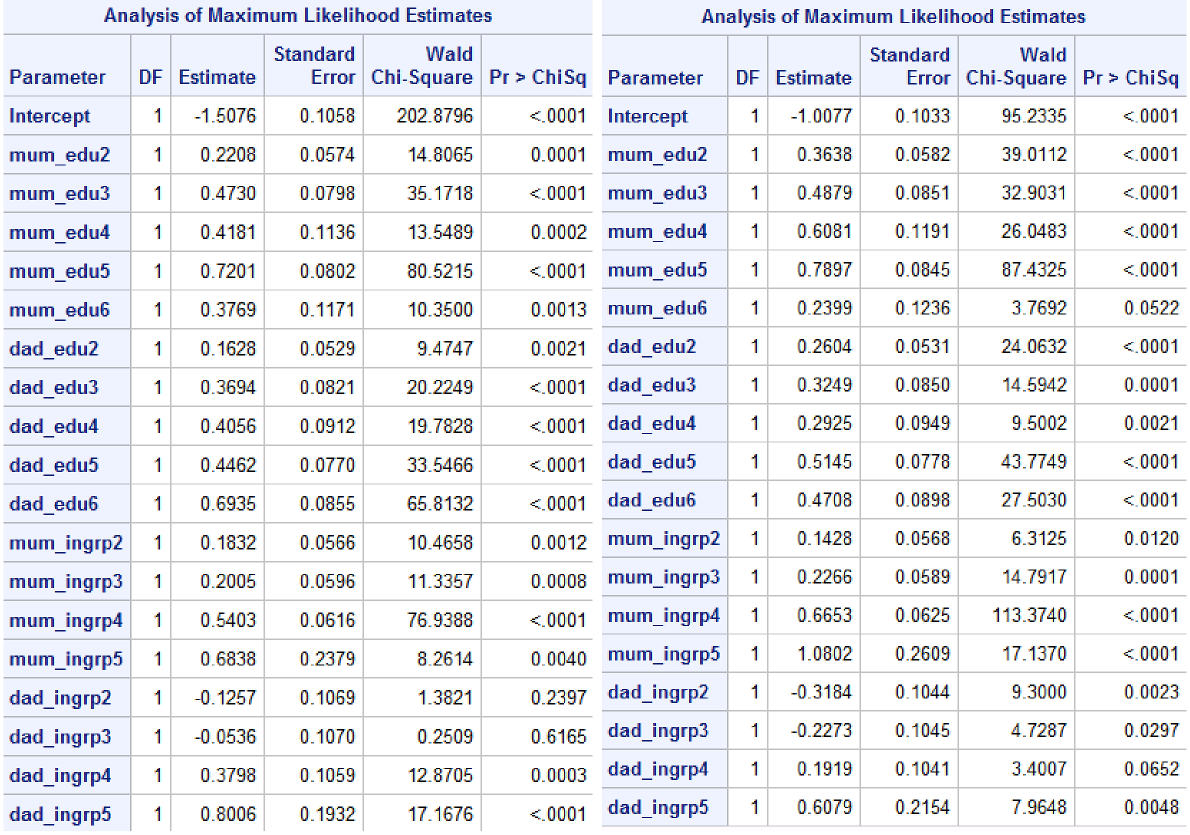
Figure C.6: Predicted probability plots for girls with native origin, data form 2006 with parents living together



Fit computed at mum_edu3=0.023 mum_edu4=0.044 mum_edu5=0.218 mum_edu6=0.036 dad_edu2=0.457 dad_edu3=0.025 dad_edu4=0.057 dad_edu5=0.127 dad_edu6=0.083 mum_ingrp3=0.262 mum_ingrp4=0.243 mum_ingrp5=0.01 dad_ingrp3=0.289 dad_ingrp4=0.291 dad_ingrp5=0.013

Fit computed at mum_edu2=0.369 mum_edu3=0.023 mum_edu4=0.044 mum_edu6=0.036 dad_edu2=0.457 dad_edu3=0.025 dad_edu4=0.057 dad_edu5=0.127 dad_edu6=0.083 mum_ingrp2=0.264 mum_ingrp3=0.262 mum_ingrp5=0.01 dad_ingrp2=0.265 dad_ingrp3=0.289 dad_ingrp5=0.013

Figure C.7: Left Table: Boys, Right Table: Girls, immigrant origin data form 2006 parents living together



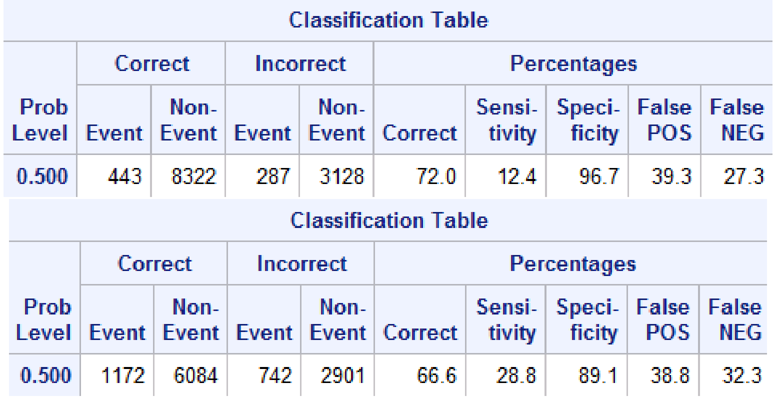
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.5076	0.1058	202.8796	<.0001
mum_edu2	1	0.2208	0.0574	14.8065	0.0001
mum_edu3	1	0.4730	0.0798	35.1718	<.0001
mum_edu4	1	0.4181	0.1136	13.5489	0.0002
mum_edu5	1	0.7201	0.0802	80.5215	<.0001
mum_edu6	1	0.3769	0.1171	10.3500	0.0013
dad_edu2	1	0.1628	0.0529	9.4747	0.0021
dad_edu3	1	0.3694	0.0821	20.2249	<.0001
dad_edu4	1	0.4056	0.0912	19.7828	<.0001
dad_edu5	1	0.4462	0.0770	33.5466	<.0001
dad_edu6	1	0.6935	0.0855	65.8132	<.0001
mum_ingrp2	1	0.1832	0.0566	10.4658	0.0012
mum_ingrp3	1	0.2005	0.0596	11.3357	0.0008
mum_ingrp4	1	0.5403	0.0616	76.9388	<.0001
mum_ingrp5	1	0.6838	0.2379	8.2614	0.0040
dad_ingrp2	1	-0.1257	0.1069	1.3821	0.2397
dad_ingrp3	1	-0.0536	0.1070	0.2509	0.6165
dad_ingrp4	1	0.3798	0.1059	12.8705	0.0003
dad_ingrp5	1	0.8006	0.1932	17.1676	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.0077	0.1033	95.2335	<.0001
mum_edu2	1	0.3638	0.0582	39.0112	<.0001
mum_edu3	1	0.4879	0.0851	32.9031	<.0001
mum_edu4	1	0.6081	0.1191	26.0483	<.0001
mum_edu5	1	0.7897	0.0845	87.4325	<.0001
mum_edu6	1	0.2399	0.1236	3.7692	0.0522
dad_edu2	1	0.2604	0.0531	24.0632	<.0001
dad_edu3	1	0.3249	0.0850	14.5942	0.0001
dad_edu4	1	0.2925	0.0949	9.5002	0.0021
dad_edu5	1	0.5145	0.0778	43.7749	<.0001
dad_edu6	1	0.4708	0.0898	27.5030	<.0001
mum_ingrp2	1	0.1428	0.0568	6.3125	0.0120
mum_ingrp3	1	0.2266	0.0589	14.7917	0.0001
mum_ingrp4	1	0.6653	0.0625	113.3740	<.0001
mum_ingrp5	1	1.0802	0.2609	17.1370	<.0001
dad_ingrp2	1	-0.3184	0.1044	9.3000	0.0023
dad_ingrp3	1	-0.2273	0.1045	4.7287	0.0297
dad_ingrp4	1	0.1919	0.1041	3.4007	0.0652
dad_ingrp5	1	0.6079	0.2154	7.9648	0.0048

Figure C.8: Top Table: Boys, Bottom Table: Girls. Immigrant origin, data form 2006 parents living together



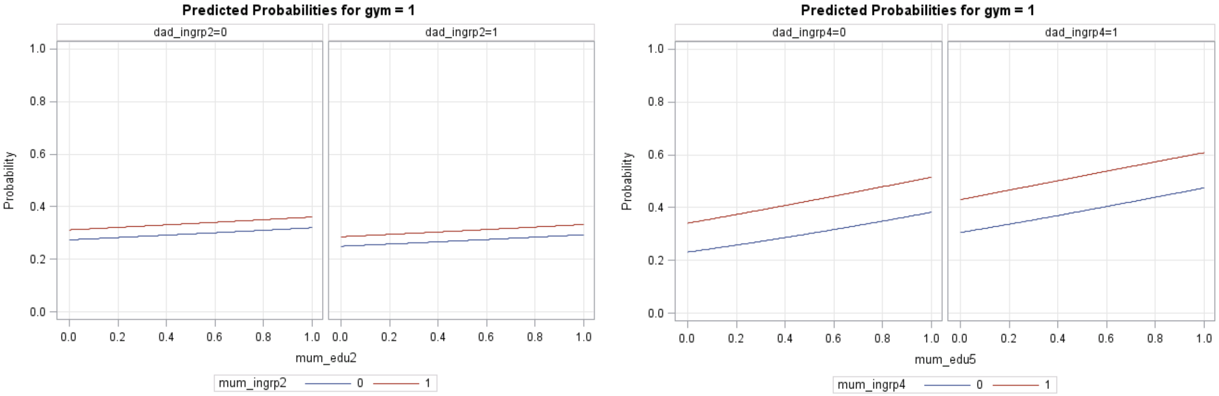
Classification Table

Prob Level	Correct		Incorrect		Percentages				
	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	443	8322	287	3128	72.0	12.4	96.7	39.3	27.3

Classification Table

Prob Level	Correct		Incorrect		Percentages				
	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	1172	6084	742	2901	66.6	28.8	89.1	38.8	32.3

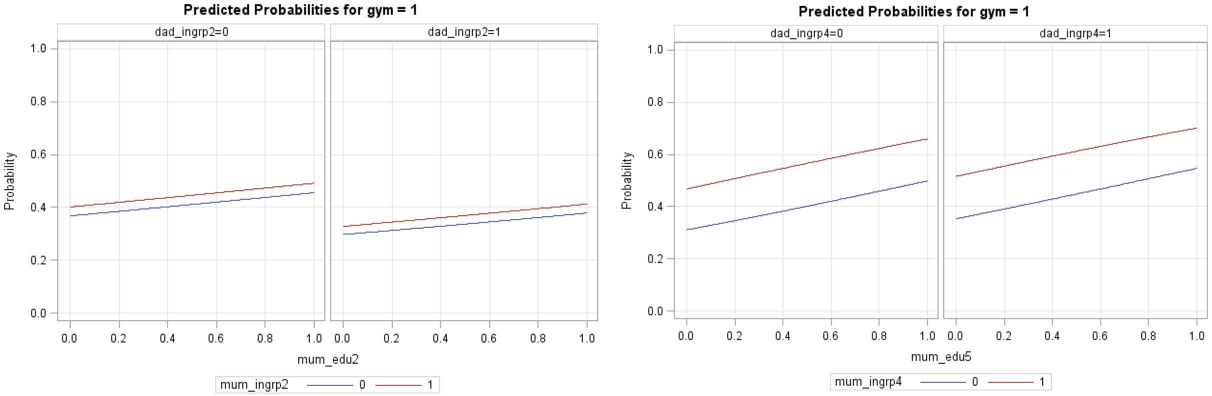
Figure C.9: Predicted probability plots for boys with immigrant origin, data form 2006 with parents living together



Fit computed at mum_edu3=0.068 mum_edu4=0.031 mum_edu5=0.067 mum_edu6=0.03 dad_edu2=0.242 dad_edu3=0.068 dad_edu4=0.052 dad_edu5=0.078 dad_edu6=0.062 mum_ingrp3=0.217 mum_ingrp4=0.194 mum_ingrp5=0.007 dad_ingrp3=0.302 dad_ingrp4=0.308 dad_ingrp5=0.013

Fit computed at mum_edu2=0.169 mum_edu3=0.068 mum_edu4=0.031 mum_edu6=0.03 dad_edu2=0.242 dad_edu3=0.068 dad_edu4=0.052 dad_edu5=0.078 dad_edu6=0.062 mum_ingrp2=0.274 mum_ingrp3=0.217 mum_ingrp5=0.007 dad_ingrp2=0.333 dad_ingrp3=0.302 dad_ingrp5=0.013

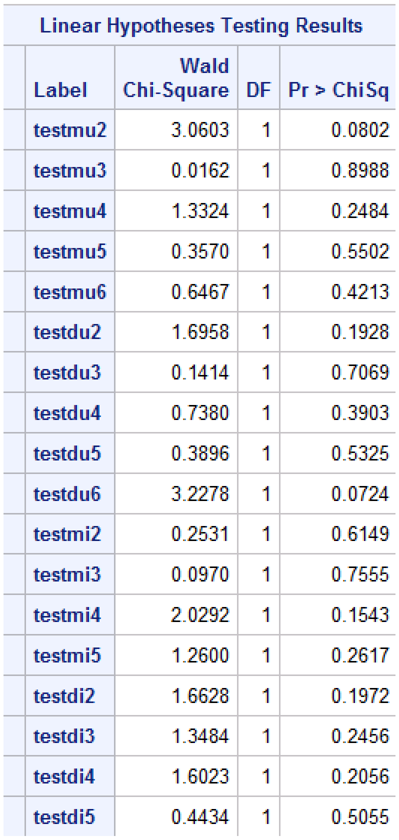
Figure C.10: Predicted probability plots for girls with immigrant origin, data form 2006 with parents living together



Fit computed at mum_edu3=0.061 mum_edu4=0.03 mum_edu5=0.068 mum_edu6=0.029 dad_edu2=0.231 dad_edu3=0.065 dad_edu4=0.051 dad_edu5=0.081 dad_edu6=0.061 mum_ingrp3=0.231 mum_ingrp4=0.201 mum_ingrp5=0.007 dad_ingrp3=0.304 dad_ingrp4=0.312 dad_ingrp5=0.012

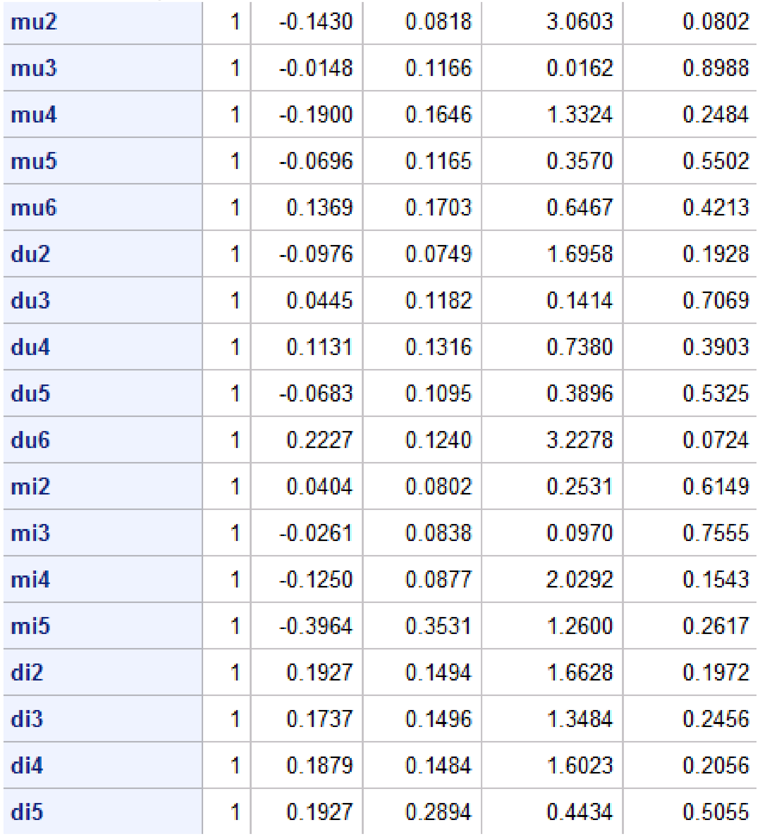
Fit computed at mum_edu2=0.162 mum_edu3=0.061 mum_edu4=0.03 mum_edu6=0.029 dad_edu2=0.231 dad_edu3=0.065 dad_edu4=0.051 dad_edu5=0.081 dad_edu6=0.061 mum_ingrp2=0.279 mum_ingrp3=0.231 mum_ingrp5=0.007 dad_ingrp2=0.329 dad_ingrp3=0.304 dad_ingrp5=0.012

Left plot: Test $\beta_i^{\text{Boys}} = \beta_i^{\text{Girls}}$, where i is a parental level of education or income.
Right plot: Observation, $\hat{\beta}$, Std.error. Wald. P-value. Immigrant origin, data form 2006 parents living together



Linear Hypotheses Testing Results

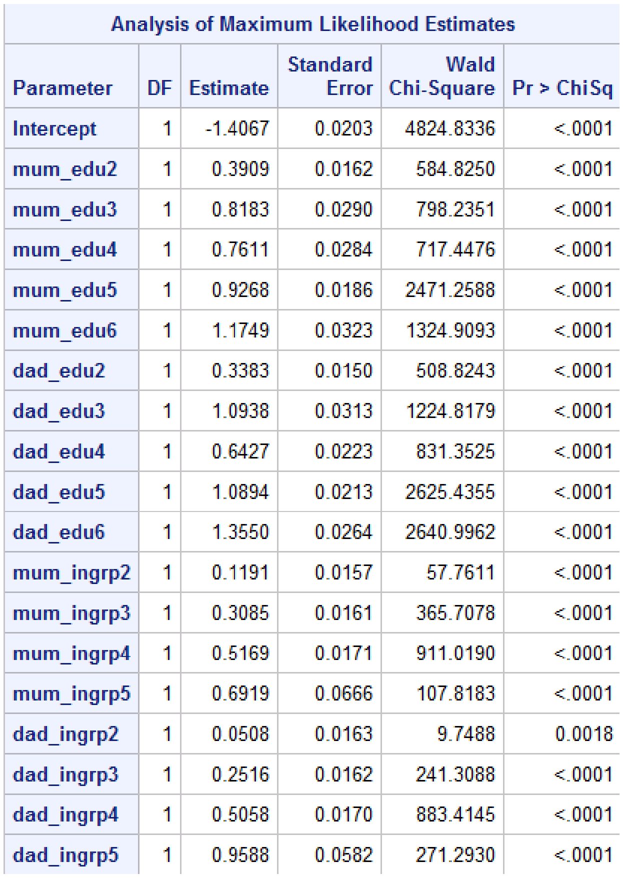
Label	Wald Chi-Square	DF	Pr > ChiSq
testmu2	3.0603	1	0.0802
testmu3	0.0162	1	0.8988
testmu4	1.3324	1	0.2484
testmu5	0.3570	1	0.5502
testmu6	0.6467	1	0.4213
testdu2	1.6958	1	0.1928
testdu3	0.1414	1	0.7069
testdu4	0.7380	1	0.3903
testdu5	0.3896	1	0.5325
testdu6	3.2278	1	0.0724
testmi2	0.2531	1	0.6149
testmi3	0.0970	1	0.7555
testmi4	2.0292	1	0.1543
testmi5	1.2600	1	0.2617
testdi2	1.6628	1	0.1972
testdi3	1.3484	1	0.2456
testdi4	1.6023	1	0.2056
testdi5	0.4434	1	0.5055



mu2	1	-0.1430	0.0818	3.0603	0.0802
mu3	1	-0.0148	0.1166	0.0162	0.8988
mu4	1	-0.1900	0.1646	1.3324	0.2484
mu5	1	-0.0696	0.1165	0.3570	0.5502
mu6	1	0.1369	0.1703	0.6467	0.4213
du2	1	-0.0976	0.0749	1.6958	0.1928
du3	1	0.0445	0.1182	0.1414	0.7069
du4	1	0.1131	0.1316	0.7380	0.3903
du5	1	-0.0683	0.1095	0.3896	0.5325
du6	1	0.2227	0.1240	3.2278	0.0724
mi2	1	0.0404	0.0802	0.2531	0.6149
mi3	1	-0.0261	0.0838	0.0970	0.7555
mi4	1	-0.1250	0.0877	2.0292	0.1543
mi5	1	-0.3964	0.3531	1.2600	0.2617
di2	1	0.1927	0.1494	1.6628	0.1972
di3	1	0.1737	0.1496	1.3484	0.2456
di4	1	0.1879	0.1484	1.6023	0.2056
di5	1	0.1927	0.2894	0.4434	0.5055

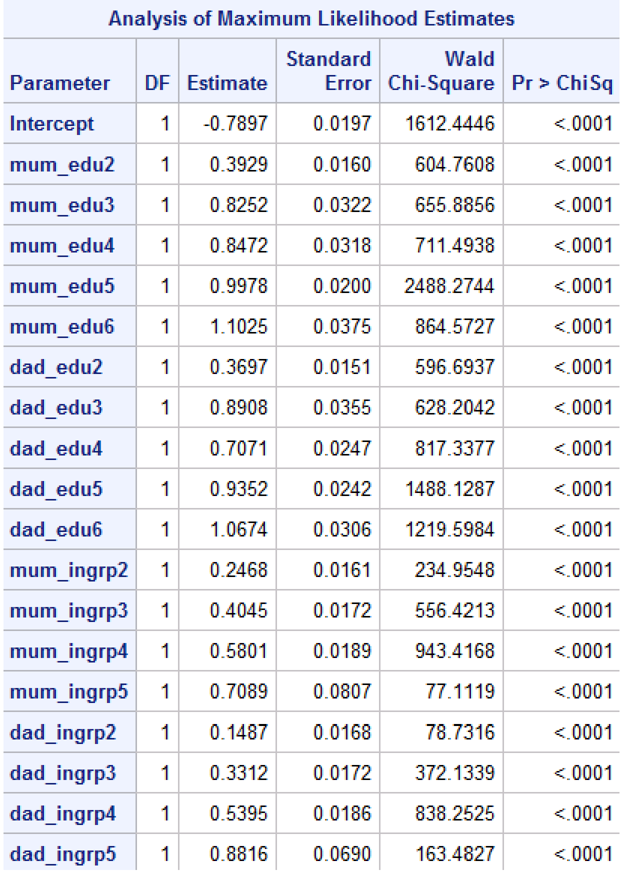
2016

Figure C.11: Left Table: Boys, Right Table: Girls, native origin data form 2016 with parents living together



Analysis of Maximum Likelihood Estimates

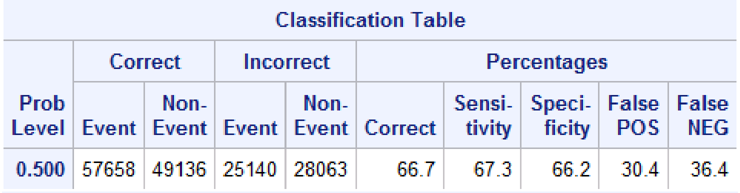
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.4067	0.0203	4824.8336	<.0001
mum_edu2	1	0.3909	0.0162	584.8250	<.0001
mum_edu3	1	0.8183	0.0290	798.2351	<.0001
mum_edu4	1	0.7611	0.0284	717.4476	<.0001
mum_edu5	1	0.9268	0.0186	2471.2588	<.0001
mum_edu6	1	1.1749	0.0323	1324.9093	<.0001
dad_edu2	1	0.3383	0.0150	508.8243	<.0001
dad_edu3	1	1.0938	0.0313	1224.8179	<.0001
dad_edu4	1	0.6427	0.0223	831.3525	<.0001
dad_edu5	1	1.0894	0.0213	2625.4355	<.0001
dad_edu6	1	1.3550	0.0264	2640.9962	<.0001
mum_ingrp2	1	0.1191	0.0157	57.7611	<.0001
mum_ingrp3	1	0.3085	0.0161	365.7078	<.0001
mum_ingrp4	1	0.5169	0.0171	911.0190	<.0001
mum_ingrp5	1	0.6919	0.0666	107.8183	<.0001
dad_ingrp2	1	0.0508	0.0163	9.7488	0.0018
dad_ingrp3	1	0.2516	0.0162	241.3088	<.0001
dad_ingrp4	1	0.5058	0.0170	883.4145	<.0001
dad_ingrp5	1	0.9588	0.0582	271.2930	<.0001



Analysis of Maximum Likelihood Estimates

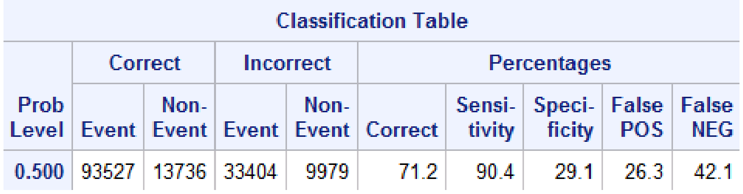
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.7897	0.0197	1612.4446	<.0001
mum_edu2	1	0.3929	0.0160	604.7608	<.0001
mum_edu3	1	0.8252	0.0322	655.8856	<.0001
mum_edu4	1	0.8472	0.0318	711.4938	<.0001
mum_edu5	1	0.9978	0.0200	2488.2744	<.0001
mum_edu6	1	1.1025	0.0375	864.5727	<.0001
dad_edu2	1	0.3697	0.0151	596.6937	<.0001
dad_edu3	1	0.8908	0.0355	628.2042	<.0001
dad_edu4	1	0.7071	0.0247	817.3377	<.0001
dad_edu5	1	0.9352	0.0242	1488.1287	<.0001
dad_edu6	1	1.0674	0.0306	1219.5984	<.0001
mum_ingrp2	1	0.2468	0.0161	234.9548	<.0001
mum_ingrp3	1	0.4045	0.0172	556.4213	<.0001
mum_ingrp4	1	0.5801	0.0189	943.4168	<.0001
mum_ingrp5	1	0.7089	0.0807	77.1119	<.0001
dad_ingrp2	1	0.1487	0.0168	78.7316	<.0001
dad_ingrp3	1	0.3312	0.0172	372.1339	<.0001
dad_ingrp4	1	0.5395	0.0186	838.2525	<.0001
dad_ingrp5	1	0.8816	0.0690	163.4827	<.0001

Figure C.12: Top Table: Boys, Bottom Table: Girls, native origin data form 2016 parents living together



Classification Table

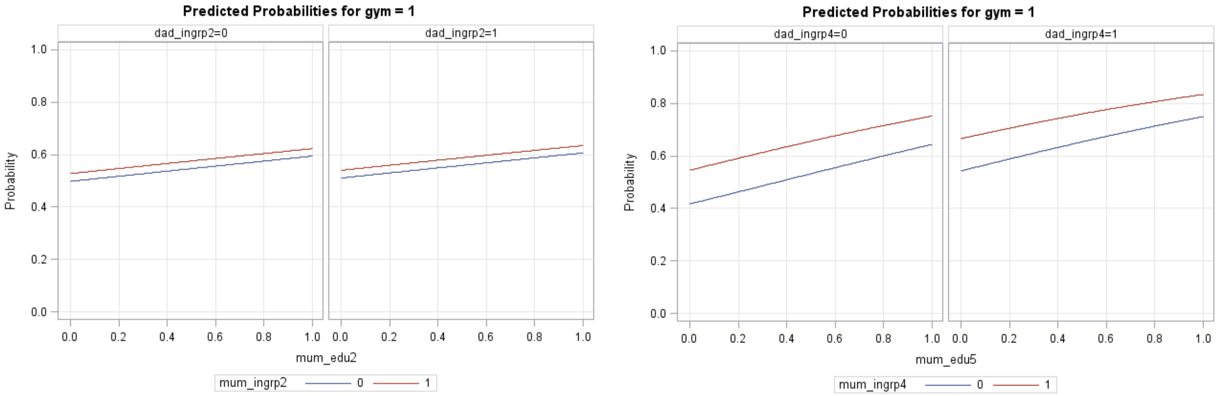
Prob Level	Correct		Incorrect		Percentages				
	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	57658	49136	25140	28063	66.7	67.3	66.2	30.4	36.4



Classification Table

Prob Level	Correct		Incorrect		Percentages				
	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	93527	13736	33404	9979	71.2	90.4	29.1	26.3	42.1

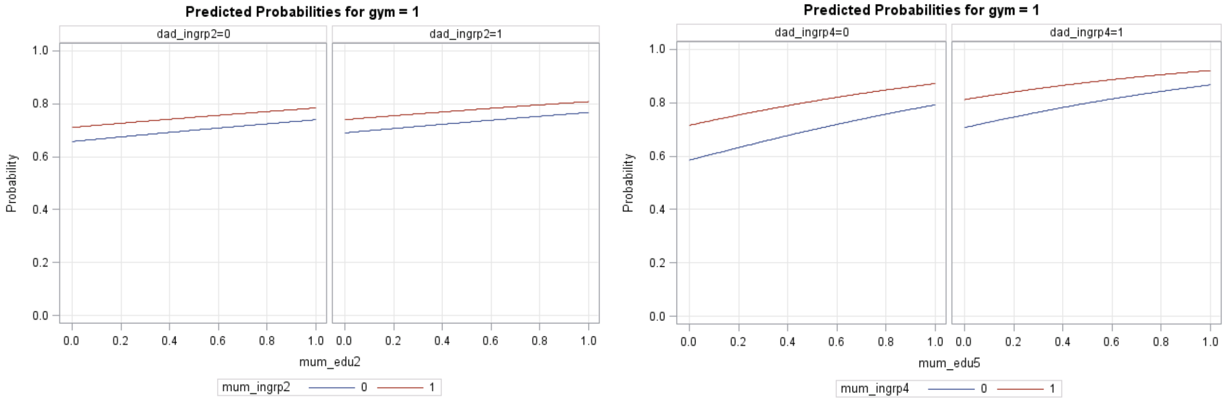
Figure C.13: Predicted probability plots for boys with native origin, data form 2016 and parents living together



Fit computed at mum_edu3=0.045 mum_edu4=0.047 mum_edu5=0.272 mum_edu6=0.073 dad_edu2=0.441 dad_edu3=0.04 dad_edu4=0.085 dad_edu5=0.138 dad_edu6=0.113 mum_ingrp3=0.256 mum_ingrp4=0.252 mum_ingrp5=0.012 dad_ingrp3=0.267 dad_ingrp4=0.279 dad_ingrp5=0.012

Fit computed at mum_edu2=0.396 mum_edu3=0.045 mum_edu4=0.047 mum_edu6=0.073 dad_edu2=0.441 dad_edu3=0.04 dad_edu4=0.085 dad_edu5=0.138 dad_edu6=0.113 mum_ingrp2=0.257 mum_ingrp3=0.256 mum_ingrp5=0.012 dad_ingrp2=0.246 dad_ingrp3=0.267 dad_ingrp5=0.012

Figure C.14: Predicted probability plots for girls with native origin, data form 2016 and parents living together

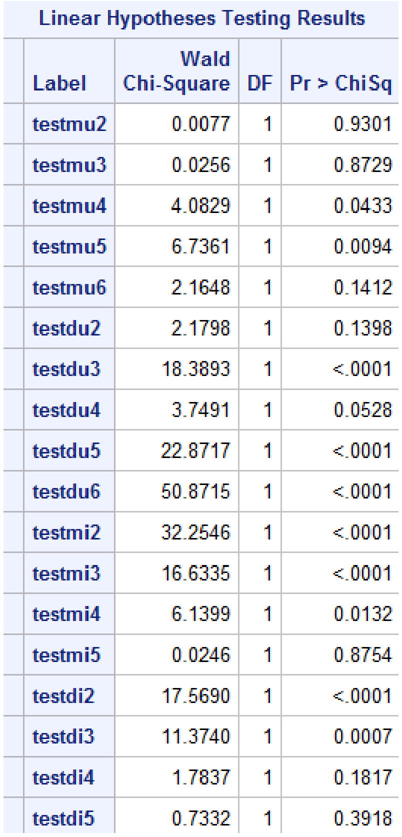


Fit computed at mum_edu3=0.045 mum_edu4=0.048 mum_edu5=0.272 mum_edu6=0.073 dad_edu2=0.444 dad_edu3=0.041 dad_edu4=0.085 dad_edu5=0.136 dad_edu6=0.111 mum_ingrp3=0.257 mum_ingrp4=0.25 mum_ingrp5=0.011 dad_ingrp3=0.267 dad_ingrp4=0.278 dad_ingrp5=0.012

Fit computed at mum_edu2=0.397 mum_edu3=0.045 mum_edu4=0.048 mum_edu6=0.073 dad_edu2=0.444 dad_edu3=0.041 dad_edu4=0.085 dad_edu5=0.136 dad_edu6=0.111 mum_ingrp2=0.258 mum_ingrp3=0.257 mum_ingrp5=0.011 dad_ingrp2=0.247 dad_ingrp3=0.267 dad_ingrp5=0.012

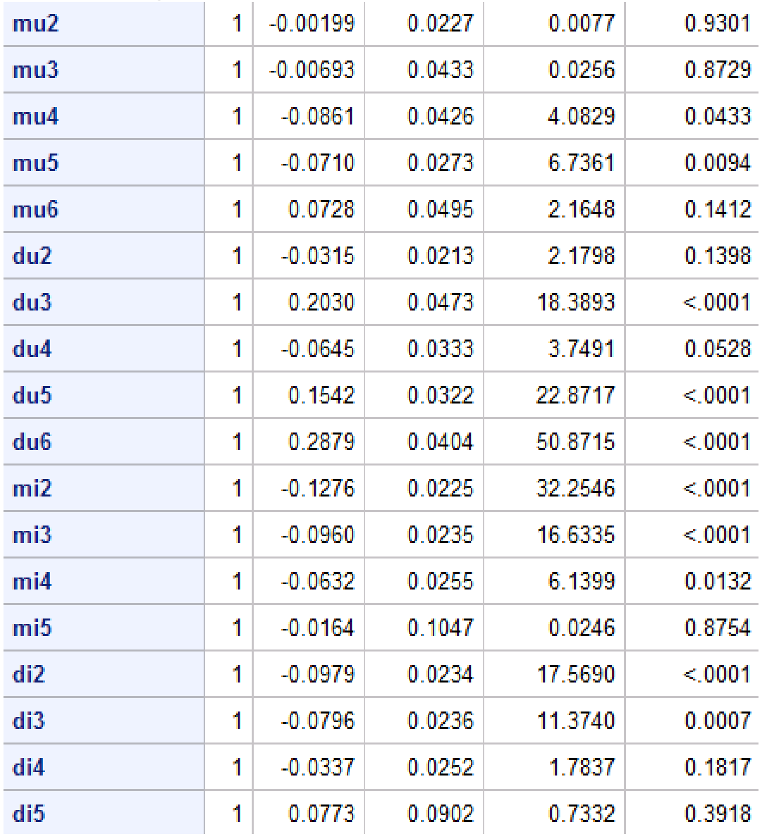
Left plot: Test $\beta_i^{\text{Boys}} = \beta_i^{\text{Girls}}$, where i is a parental level of education or income.

Right plot: Observation, $\hat{\beta}$, Std.error. Wald. P-value. Native origin, data form 2016 parents living together



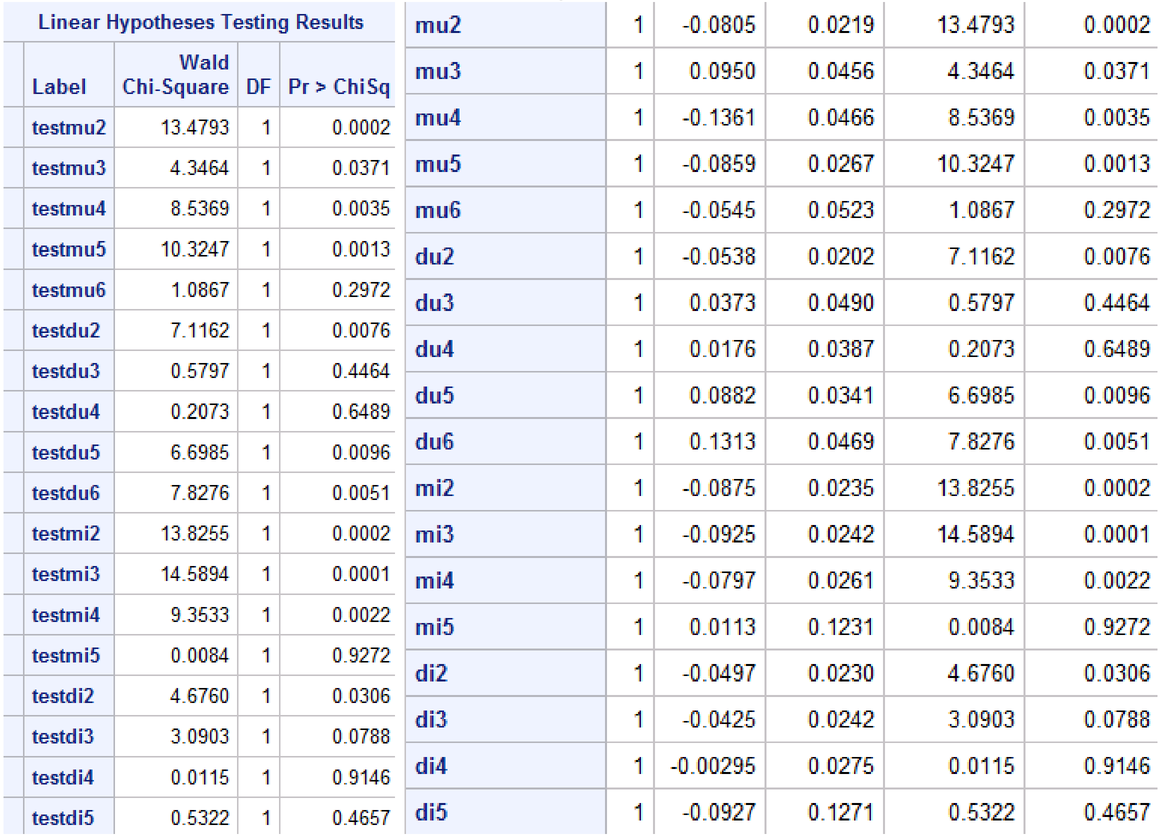
Linear Hypotheses Testing Results

Label	Wald Chi-Square	DF	Pr > ChiSq
testmu2	0.0077	1	0.9301
testmu3	0.0256	1	0.8729
testmu4	4.0829	1	0.0433
testmu5	6.7361	1	0.0094
testmu6	2.1648	1	0.1412
testdu2	2.1798	1	0.1398
testdu3	18.3893	1	<.0001
testdu4	3.7491	1	0.0528
testdu5	22.8717	1	<.0001
testdu6	50.8715	1	<.0001
testmi2	32.2546	1	<.0001
testmi3	16.6335	1	<.0001
testmi4	6.1399	1	0.0132
testmi5	0.0246	1	0.8754
testdi2	17.5690	1	<.0001
testdi3	11.3740	1	0.0007
testdi4	1.7837	1	0.1817
testdi5	0.7332	1	0.3918



mu2	1	-0.00199	0.0227	0.0077	0.9301
mu3	1	-0.00693	0.0433	0.0256	0.8729
mu4	1	-0.0861	0.0426	4.0829	0.0433
mu5	1	-0.0710	0.0273	6.7361	0.0094
mu6	1	0.0728	0.0495	2.1648	0.1412
du2	1	-0.0315	0.0213	2.1798	0.1398
du3	1	0.2030	0.0473	18.3893	<.0001
du4	1	-0.0645	0.0333	3.7491	0.0528
du5	1	0.1542	0.0322	22.8717	<.0001
du6	1	0.2879	0.0404	50.8715	<.0001
mi2	1	-0.1276	0.0225	32.2546	<.0001
mi3	1	-0.0960	0.0235	16.6335	<.0001
mi4	1	-0.0632	0.0255	6.1399	0.0132
mi5	1	-0.0164	0.1047	0.0246	0.8754
di2	1	-0.0979	0.0234	17.5690	<.0001
di3	1	-0.0796	0.0236	11.3740	0.0007
di4	1	-0.0337	0.0252	1.7837	0.1817
di5	1	0.0773	0.0902	0.7332	0.3918

Left plot: Test $\beta_i^{\text{Boys}} = \beta_i^{\text{Girls}}$, where i is a parental level of education or income.
Right plot: Observation, $\hat{\beta}$, Std.error. Wald. P-value. Native origin, data form 2016 parents *not* living together

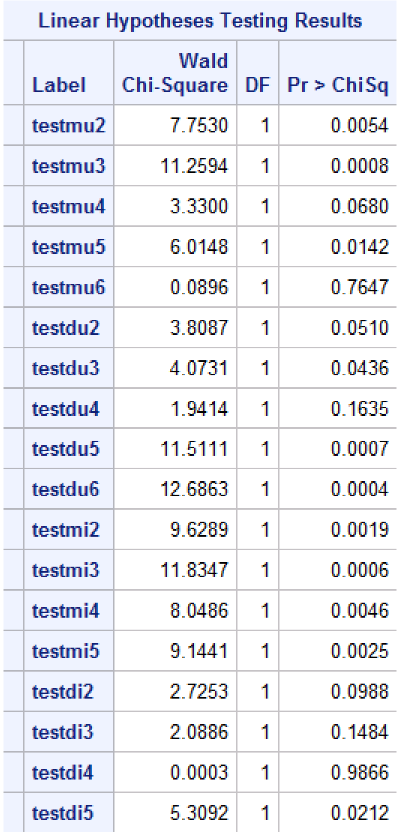


Linear Hypotheses Testing Results

Label	Wald Chi-Square	DF	Pr > ChiSq
testmu2	13.4793	1	0.0002
testmu3	4.3464	1	0.0371
testmu4	8.5369	1	0.0035
testmu5	10.3247	1	0.0013
testmu6	1.0867	1	0.2972
testdu2	7.1162	1	0.0076
testdu3	0.5797	1	0.4464
testdu4	0.2073	1	0.6489
testdu5	6.6985	1	0.0096
testdu6	7.8276	1	0.0051
testmi2	13.8255	1	0.0002
testmi3	14.5894	1	0.0001
testmi4	9.3533	1	0.0022
testmi5	0.0084	1	0.9272
testdi2	4.6760	1	0.0306
testdi3	3.0903	1	0.0788
testdi4	0.0115	1	0.9146
testdi5	0.5322	1	0.4657

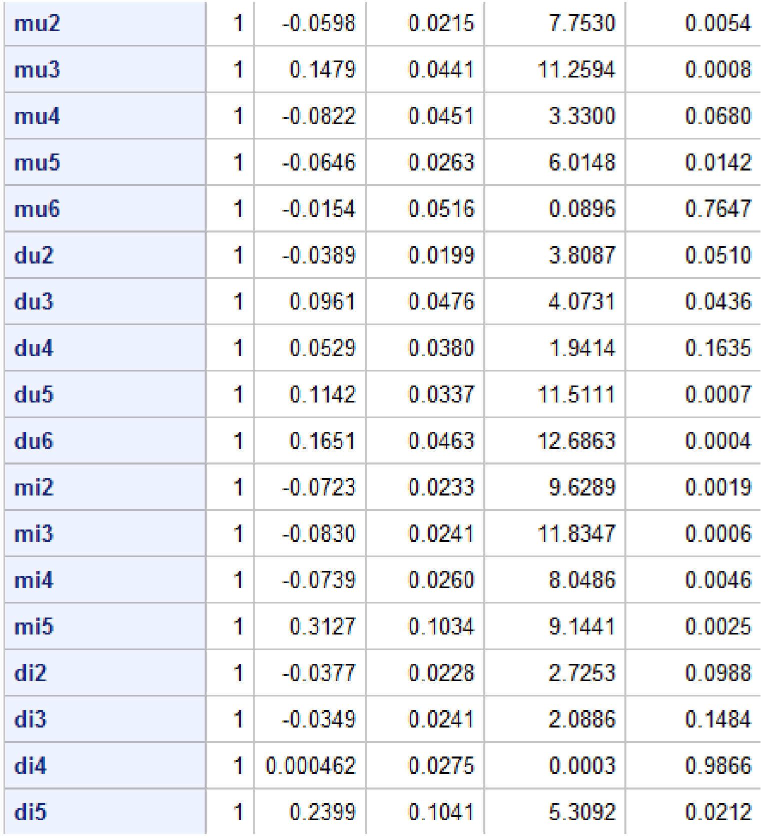
mu2	1	-0.0805	0.0219	13.4793	0.0002
mu3	1	0.0950	0.0456	4.3464	0.0371
mu4	1	-0.1361	0.0466	8.5369	0.0035
mu5	1	-0.0859	0.0267	10.3247	0.0013
mu6	1	-0.0545	0.0523	1.0867	0.2972
du2	1	-0.0538	0.0202	7.1162	0.0076
du3	1	0.0373	0.0490	0.5797	0.4464
du4	1	0.0176	0.0387	0.2073	0.6489
du5	1	0.0882	0.0341	6.6985	0.0096
du6	1	0.1313	0.0469	7.8276	0.0051
mi2	1	-0.0875	0.0235	13.8255	0.0002
mi3	1	-0.0925	0.0242	14.5894	0.0001
mi4	1	-0.0797	0.0261	9.3533	0.0022
mi5	1	0.0113	0.1231	0.0084	0.9272
di2	1	-0.0497	0.0230	4.6760	0.0306
di3	1	-0.0425	0.0242	3.0903	0.0788
di4	1	-0.00295	0.0275	0.0115	0.9146
di5	1	-0.0927	0.1271	0.5322	0.4657

Left plot: Test $\beta_i^{\text{Not together}} = \beta_i^{\text{Together}}$, where i is a parental level of education or income. Right plot: Observation, $\hat{\beta}$, Std.error. Wald. P-value. Native origin, data form 2016 parents live/*not* living together



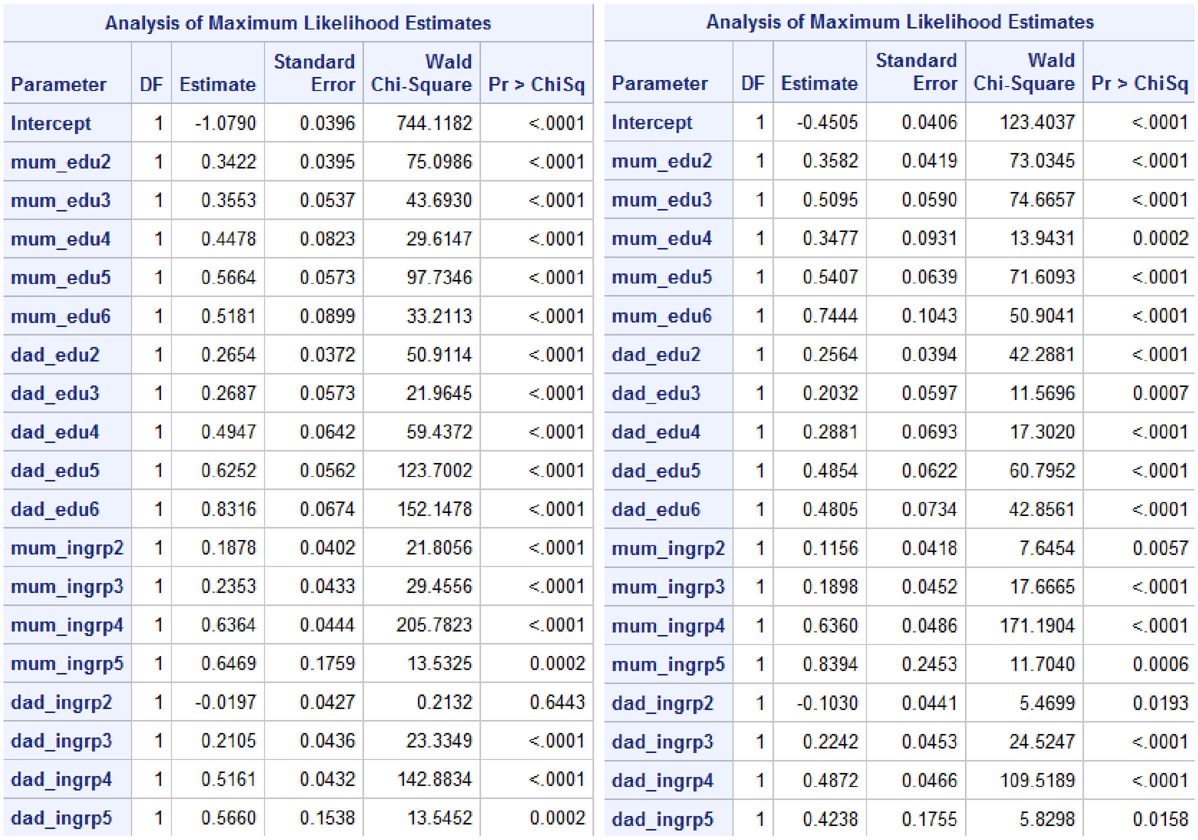
Linear Hypotheses Testing Results

Label	Wald Chi-Square	DF	Pr > ChiSq
testmu2	7.7530	1	0.0054
testmu3	11.2594	1	0.0008
testmu4	3.3300	1	0.0680
testmu5	6.0148	1	0.0142
testmu6	0.0896	1	0.7647
testdu2	3.8087	1	0.0510
testdu3	4.0731	1	0.0436
testdu4	1.9414	1	0.1635
testdu5	11.5111	1	0.0007
testdu6	12.6863	1	0.0004
testmi2	9.6289	1	0.0019
testmi3	11.8347	1	0.0006
testmi4	8.0486	1	0.0046
testmi5	9.1441	1	0.0025
testdi2	2.7253	1	0.0988
testdi3	2.0886	1	0.1484
testdi4	0.0003	1	0.9866
testdi5	5.3092	1	0.0212



mu2	1	-0.0598	0.0215	7.7530	0.0054
mu3	1	0.1479	0.0441	11.2594	0.0008
mu4	1	-0.0822	0.0451	3.3300	0.0680
mu5	1	-0.0646	0.0263	6.0148	0.0142
mu6	1	-0.0154	0.0516	0.0896	0.7647
du2	1	-0.0389	0.0199	3.8087	0.0510
du3	1	0.0961	0.0476	4.0731	0.0436
du4	1	0.0529	0.0380	1.9414	0.1635
du5	1	0.1142	0.0337	11.5111	0.0007
du6	1	0.1651	0.0463	12.6863	0.0004
mi2	1	-0.0723	0.0233	9.6289	0.0019
mi3	1	-0.0830	0.0241	11.8347	0.0006
mi4	1	-0.0739	0.0260	8.0486	0.0046
mi5	1	0.3127	0.1034	9.1441	0.0025
di2	1	-0.0377	0.0228	2.7253	0.0988
di3	1	-0.0349	0.0241	2.0886	0.1484
di4	1	0.000462	0.0275	0.0003	0.9866
di5	1	0.2399	0.1041	5.3092	0.0212

Figure C.15: Left Table: Boys, Right Table: Girls, immigrant origin data form 2016 parents living together



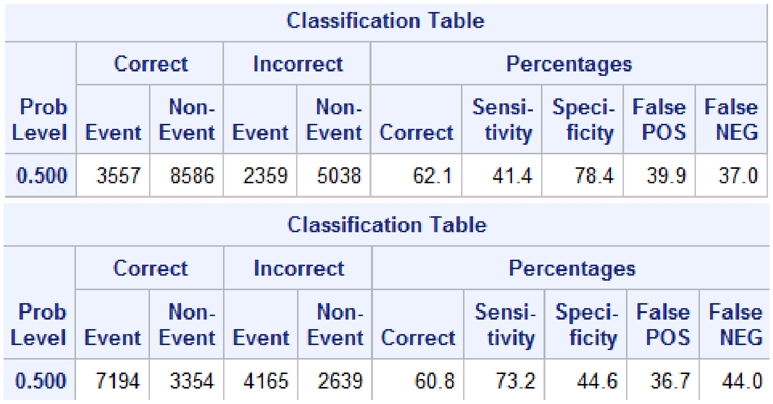
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.0790	0.0396	744.1182	<.0001
mum_edu2	1	0.3422	0.0395	75.0986	<.0001
mum_edu3	1	0.3553	0.0537	43.6930	<.0001
mum_edu4	1	0.4478	0.0823	29.6147	<.0001
mum_edu5	1	0.5664	0.0573	97.7346	<.0001
mum_edu6	1	0.5181	0.0899	33.2113	<.0001
dad_edu2	1	0.2654	0.0372	50.9114	<.0001
dad_edu3	1	0.2687	0.0573	21.9645	<.0001
dad_edu4	1	0.4947	0.0642	59.4372	<.0001
dad_edu5	1	0.6252	0.0562	123.7002	<.0001
dad_edu6	1	0.8316	0.0674	152.1478	<.0001
mum_ingrp2	1	0.1878	0.0402	21.8056	<.0001
mum_ingrp3	1	0.2353	0.0433	29.4556	<.0001
mum_ingrp4	1	0.6364	0.0444	205.7823	<.0001
mum_ingrp5	1	0.6469	0.1759	13.5325	0.0002
dad_ingrp2	1	-0.0197	0.0427	0.2132	0.6443
dad_ingrp3	1	0.2105	0.0436	23.3349	<.0001
dad_ingrp4	1	0.5161	0.0432	142.8834	<.0001
dad_ingrp5	1	0.5660	0.1538	13.5452	0.0002

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.4505	0.0406	123.4037	<.0001
mum_edu2	1	0.3582	0.0419	73.0345	<.0001
mum_edu3	1	0.5095	0.0590	74.6657	<.0001
mum_edu4	1	0.3477	0.0931	13.9431	0.0002
mum_edu5	1	0.5407	0.0639	71.6093	<.0001
mum_edu6	1	0.7444	0.1043	50.9041	<.0001
dad_edu2	1	0.2564	0.0394	42.2881	<.0001
dad_edu3	1	0.2032	0.0597	11.5696	0.0007
dad_edu4	1	0.2881	0.0693	17.3020	<.0001
dad_edu5	1	0.4854	0.0622	60.7952	<.0001
dad_edu6	1	0.4805	0.0734	42.8561	<.0001
mum_ingrp2	1	0.1156	0.0418	7.6454	0.0057
mum_ingrp3	1	0.1898	0.0452	17.6665	<.0001
mum_ingrp4	1	0.6360	0.0486	171.1904	<.0001
mum_ingrp5	1	0.8394	0.2453	11.7040	0.0006
dad_ingrp2	1	-0.1030	0.0441	5.4699	0.0193
dad_ingrp3	1	0.2242	0.0453	24.5247	<.0001
dad_ingrp4	1	0.4872	0.0466	109.5189	<.0001
dad_ingrp5	1	0.4238	0.1755	5.8298	0.0158

Figure C.16: Top Table: Boys, Bottom Table: Girls. Immigrant origin, data form 2016 parents living together



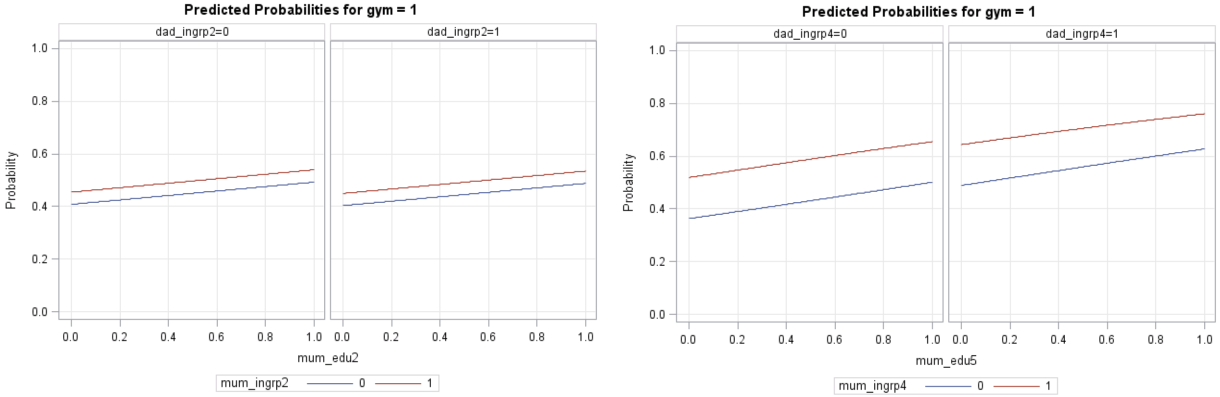
Classification Table

Prob Level	Correct		Incorrect		Percentages				
	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	3557	8586	2359	5038	62.1	41.4	78.4	39.9	37.0

Classification Table

Prob Level	Correct		Incorrect		Percentages				
	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	7194	3354	4165	2639	60.8	73.2	44.6	36.7	44.0

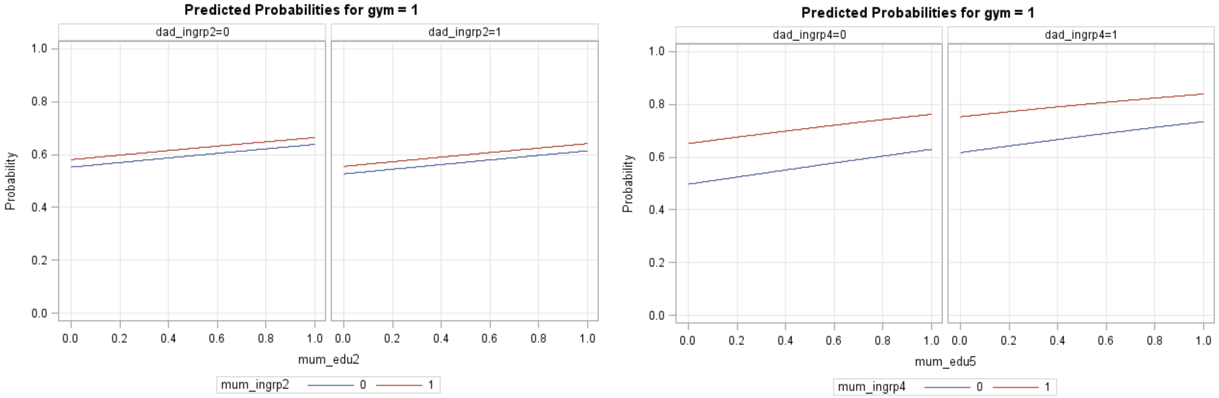
Figure C.17: Predicted probability plots for boys with immigrant origin, data form 2016 with parents living together



Fit computed at mum_edu3=0.089 mum_edu4=0.035 mum_edu5=0.085 mum_edu6=0.034 dad_edu2=0.27 dad_edu3=0.08 dad_edu4=0.061 dad_edu5=0.088 dad_edu6=0.062 mum_ingrp3=0.192 mum_ingrp4=0.207 mum_ingrp5=0.01 dad_ingrp3=0.222 dad_ingrp4=0.251 dad_ingrp5=0.011

Fit computed at mum_edu2=0.214 mum_edu3=0.089 mum_edu4=0.035 mum_edu6=0.034 dad_edu2=0.27 dad_edu3=0.08 dad_edu4=0.061 dad_edu5=0.088 dad_edu6=0.062 mum_ingrp2=0.27 mum_ingrp3=0.192 mum_ingrp5=0.01 dad_ingrp2=0.256 dad_ingrp3=0.222 dad_ingrp5=0.011

Figure C.18: Predicted probability plots for girls with immigrant origin, data form 2016 with parents living together



Fit computed at mum_edu3=0.087 mum_edu4=0.032 mum_edu5=0.084 mum_edu6=0.033 dad_edu2=0.268 dad_edu3=0.083 dad_edu4=0.061 dad_edu5=0.086 dad_edu6=0.064 mum_ingrp3=0.203 mum_ingrp4=0.215 mum_ingrp5=0.007 dad_ingrp3=0.229 dad_ingrp4=0.251 dad_ingrp5=0.011

Fit computed at mum_edu2=0.221 mum_edu3=0.087 mum_edu4=0.032 mum_edu6=0.033 dad_edu2=0.268 dad_edu3=0.083 dad_edu4=0.061 dad_edu5=0.086 dad_edu6=0.064 mum_ingrp2=0.274 mum_ingrp3=0.203 mum_ingrp5=0.007 dad_ingrp2=0.253 dad_ingrp3=0.229 dad_ingrp5=0.011

Figure C.19: Left plot: Test $\hat{\beta}_i^{\text{Girls}} = \hat{\beta}_i^{\text{Boys}}$, where i is a parental level of education or income. Right plot: Observation, $\hat{\beta}$, Std.error. Wald. P-value. Immigrant origin, data form 2016 with parents living together

Linear Hypotheses Testing Results

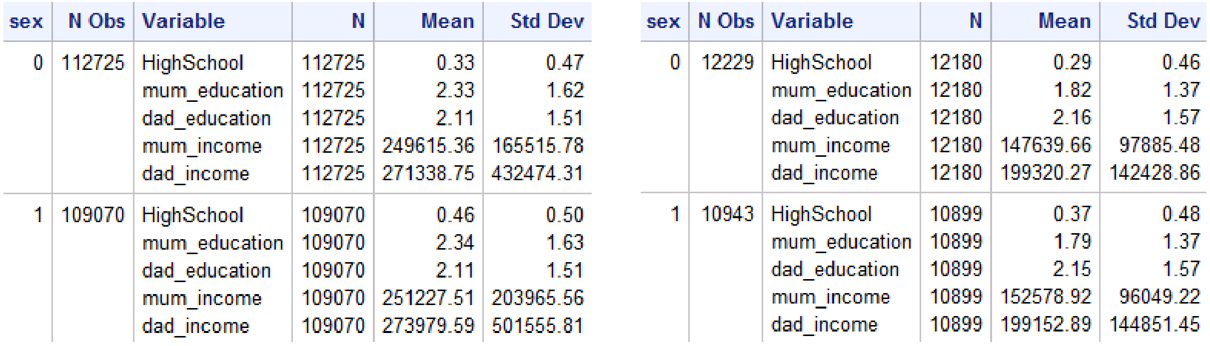
Label	Wald Chi-Square	DF	Pr > ChiSq
testmu2	0.0766	1	0.7820
testmu3	3.7384	1	0.0532
testmu4	0.6499	1	0.4201
testmu5	0.0899	1	0.7643
testmu6	2.6999	1	0.1004
testdu2	0.0274	1	0.8686
testdu3	0.6256	1	0.4290
testdu4	4.7873	1	0.0287
testdu5	2.7811	1	0.0954
testdu6	12.4122	1	0.0004
testmi2	1.5465	1	0.2137
testmi3	0.5264	1	0.4681
testmi4	0.0000	1	0.9961
testmi5	0.4061	1	0.5240
testdi2	1.8469	1	0.1741
testdi3	0.0475	1	0.8274
testdi4	0.2077	1	0.6486
testdi5	0.3713	1	0.5423

mu2	1	-0.0159	0.0576	0.0766	0.7820
mu3	1	-0.1543	0.0798	3.7384	0.0532
mu4	1	0.1002	0.1243	0.6499	0.4201
mu5	1	0.0257	0.0858	0.0899	0.7643
mu6	1	-0.2263	0.1377	2.6999	0.1004
du2	1	0.00897	0.0542	0.0274	0.8686
du3	1	0.0655	0.0828	0.6256	0.4290
du4	1	0.2066	0.0944	4.7873	0.0287
du5	1	0.1399	0.0839	2.7811	0.0954
du6	1	0.3511	0.0997	12.4122	0.0004
mi2	1	0.0722	0.0580	1.5465	0.2137
mi3	1	0.0454	0.0626	0.5264	0.4681
mi4	1	0.000322	0.0658	0.0000	0.9961
mi5	1	-0.1924	0.3019	0.4061	0.5240
di2	1	0.0833	0.0613	1.8469	0.1741
di3	1	-0.0137	0.0628	0.0475	0.8274
di4	1	0.0289	0.0635	0.2077	0.6486
di5	1	0.1422	0.2334	0.3713	0.5423

Appendix D

When parents do *not* live together

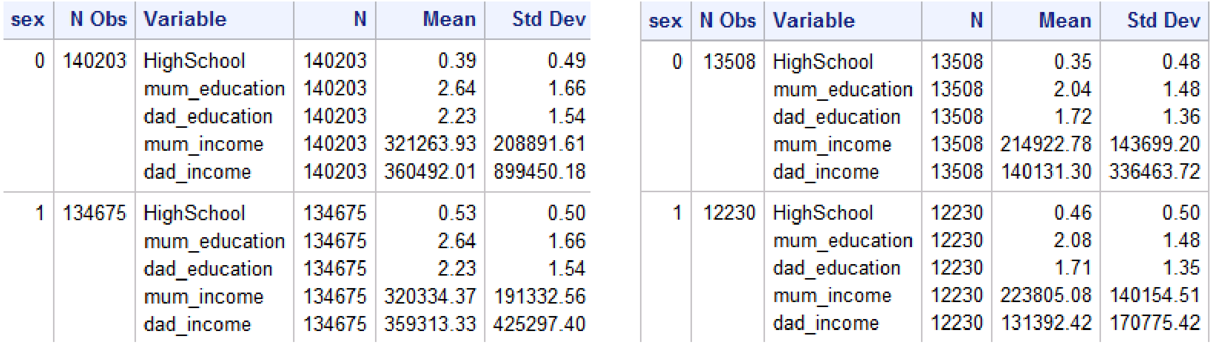
Figure D.1: Left Table: Native origin, Right Table: Immigrant origin. Summary when parents do *not* live together 2006



sex	N Obs	Variable	N	Mean	Std Dev
0	112725	HighSchool	112725	0.33	0.47
		mum_education	112725	2.33	1.62
		dad_education	112725	2.11	1.51
		mum_income	112725	249615.36	165515.78
		dad_income	112725	271338.75	432474.31
1	109070	HighSchool	109070	0.46	0.50
		mum_education	109070	2.34	1.63
		dad_education	109070	2.11	1.51
		mum_income	109070	251227.51	203965.56
		dad_income	109070	273979.59	501555.81

sex	N Obs	Variable	N	Mean	Std Dev
0	12229	HighSchool	12180	0.29	0.46
		mum_education	12180	1.82	1.37
		dad_education	12180	2.16	1.57
		mum_income	12180	147639.66	97885.48
		dad_income	12180	199320.27	142428.86
1	10943	HighSchool	10899	0.37	0.48
		mum_education	10899	1.79	1.37
		dad_education	10899	2.15	1.57
		mum_income	10899	152578.92	96049.22
		dad_income	10899	199152.89	144851.45

Figure D.2: Left Table: Native origin, Right Table: Immigrant origin. Summary when parents do *not* live together 2016

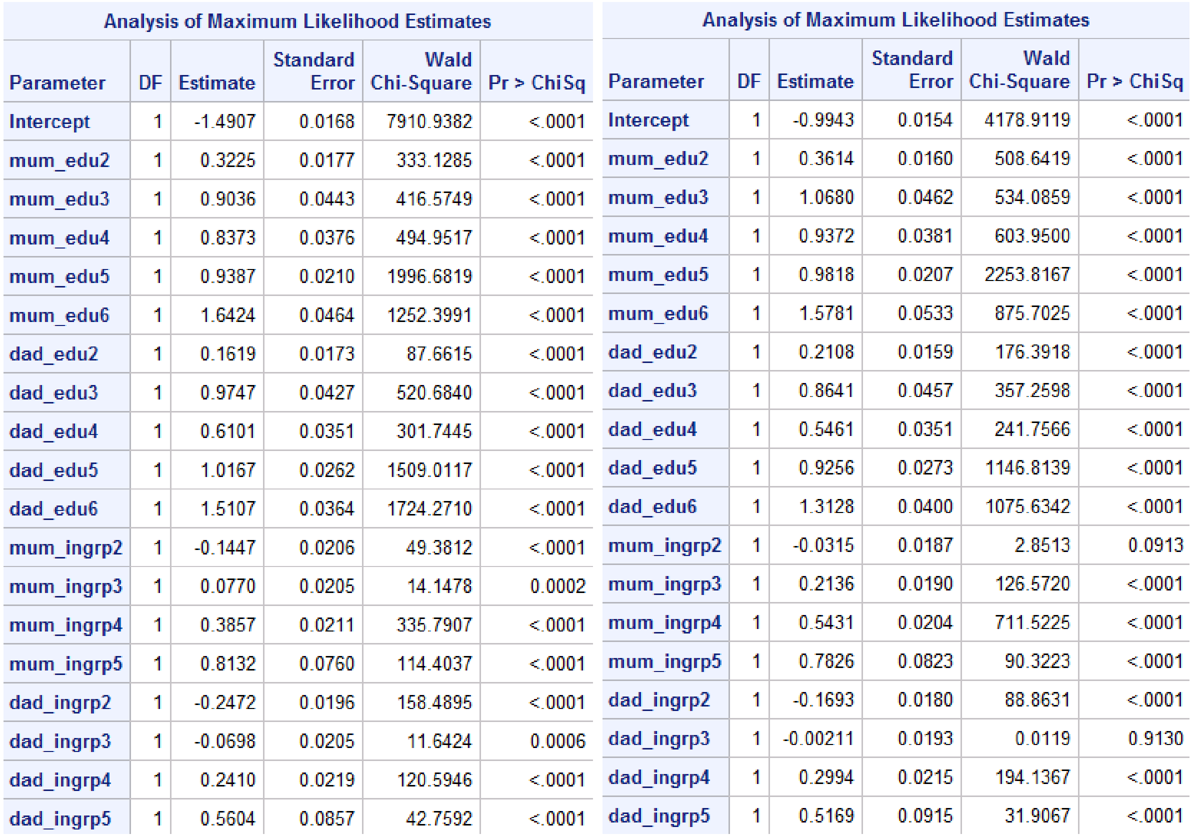


sex	N Obs	Variable	N	Mean	Std Dev
0	140203	HighSchool	140203	0.39	0.49
		mum_education	140203	2.64	1.66
		dad_education	140203	2.23	1.54
		mum_income	140203	321263.93	208891.61
		dad_income	140203	360492.01	899450.18
1	134675	HighSchool	134675	0.53	0.50
		mum_education	134675	2.64	1.66
		dad_education	134675	2.23	1.54
		mum_income	134675	320334.37	191332.56
		dad_income	134675	359313.33	425297.40

sex	N Obs	Variable	N	Mean	Std Dev
0	13508	HighSchool	13508	0.35	0.48
		mum_education	13508	2.04	1.48
		dad_education	13508	1.72	1.36
		mum_income	13508	214922.78	143699.20
		dad_income	13508	140131.30	336463.72
1	12230	HighSchool	12230	0.46	0.50
		mum_education	12230	2.08	1.48
		dad_education	12230	1.71	1.35
		mum_income	12230	223805.08	140154.51
		dad_income	12230	131392.42	170775.42

2006

Figure D.3: Left Table: Boys, Right Table: Girls, native origin data form 2006 when parents do *not* live together



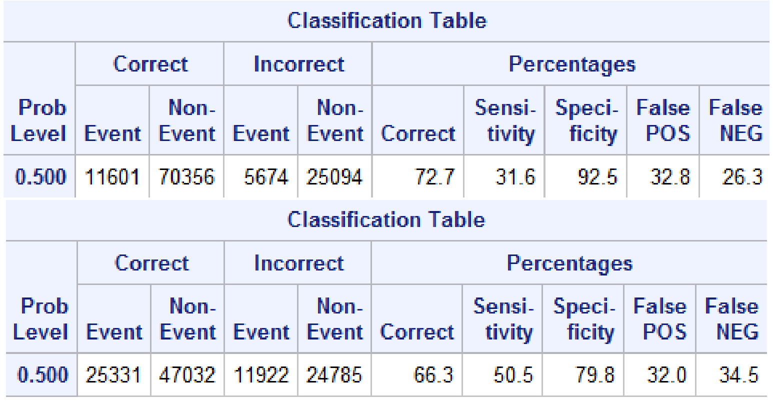
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.4907	0.0168	7910.9382	<.0001
mum_edu2	1	0.3225	0.0177	333.1285	<.0001
mum_edu3	1	0.9036	0.0443	416.5749	<.0001
mum_edu4	1	0.8373	0.0376	494.9517	<.0001
mum_edu5	1	0.9387	0.0210	1996.6819	<.0001
mum_edu6	1	1.6424	0.0464	1252.3991	<.0001
dad_edu2	1	0.1619	0.0173	87.6615	<.0001
dad_edu3	1	0.9747	0.0427	520.6840	<.0001
dad_edu4	1	0.6101	0.0351	301.7445	<.0001
dad_edu5	1	1.0167	0.0262	1509.0117	<.0001
dad_edu6	1	1.5107	0.0364	1724.2710	<.0001
mum_ingrp2	1	-0.1447	0.0206	49.3812	<.0001
mum_ingrp3	1	0.0770	0.0205	14.1478	0.0002
mum_ingrp4	1	0.3857	0.0211	335.7907	<.0001
mum_ingrp5	1	0.8132	0.0760	114.4037	<.0001
dad_ingrp2	1	-0.2472	0.0196	158.4895	<.0001
dad_ingrp3	1	-0.0698	0.0205	11.6424	0.0006
dad_ingrp4	1	0.2410	0.0219	120.5946	<.0001
dad_ingrp5	1	0.5604	0.0857	42.7592	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.9943	0.0154	4178.9119	<.0001
mum_edu2	1	0.3614	0.0160	508.6419	<.0001
mum_edu3	1	1.0680	0.0462	534.0859	<.0001
mum_edu4	1	0.9372	0.0381	603.9500	<.0001
mum_edu5	1	0.9818	0.0207	2253.8167	<.0001
mum_edu6	1	1.5781	0.0533	875.7025	<.0001
dad_edu2	1	0.2108	0.0159	176.3918	<.0001
dad_edu3	1	0.8641	0.0457	357.2598	<.0001
dad_edu4	1	0.5461	0.0351	241.7566	<.0001
dad_edu5	1	0.9256	0.0273	1146.8139	<.0001
dad_edu6	1	1.3128	0.0400	1075.6342	<.0001
mum_ingrp2	1	-0.0315	0.0187	2.8513	0.0913
mum_ingrp3	1	0.2136	0.0190	126.5720	<.0001
mum_ingrp4	1	0.5431	0.0204	711.5225	<.0001
mum_ingrp5	1	0.7826	0.0823	90.3223	<.0001
dad_ingrp2	1	-0.1693	0.0180	88.8631	<.0001
dad_ingrp3	1	-0.00211	0.0193	0.0119	0.9130
dad_ingrp4	1	0.2994	0.0215	194.1367	<.0001
dad_ingrp5	1	0.5169	0.0915	31.9067	<.0001

Figure D.4: Top Table: Boys, Bottom Table: Girls. Native origin, data form 2006 when parents do *not* live together



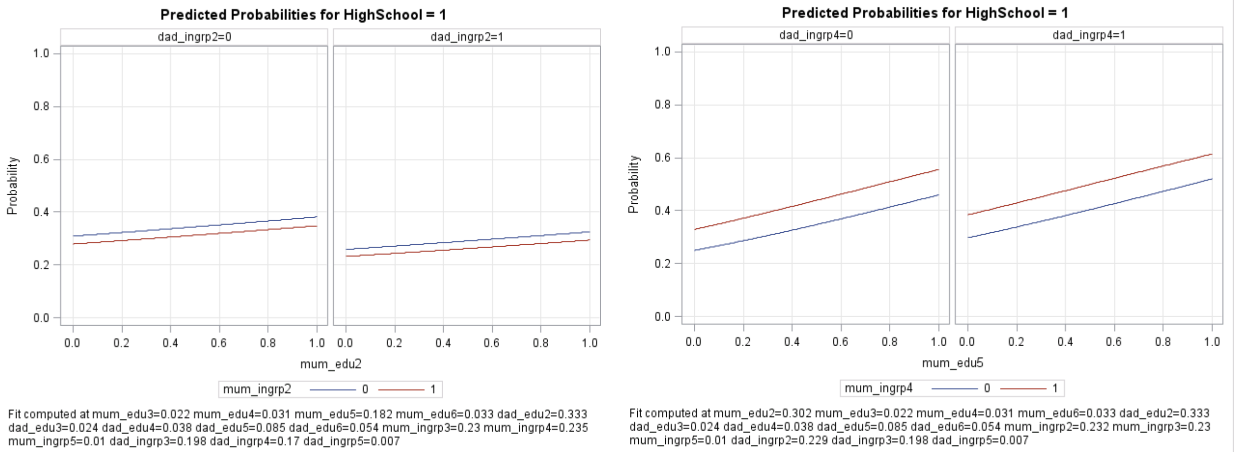
Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	11601	70356	5674	25094	72.7	31.6	92.5	32.8	26.3

Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	25331	47032	11922	24785	66.3	50.5	79.8	32.0	34.5

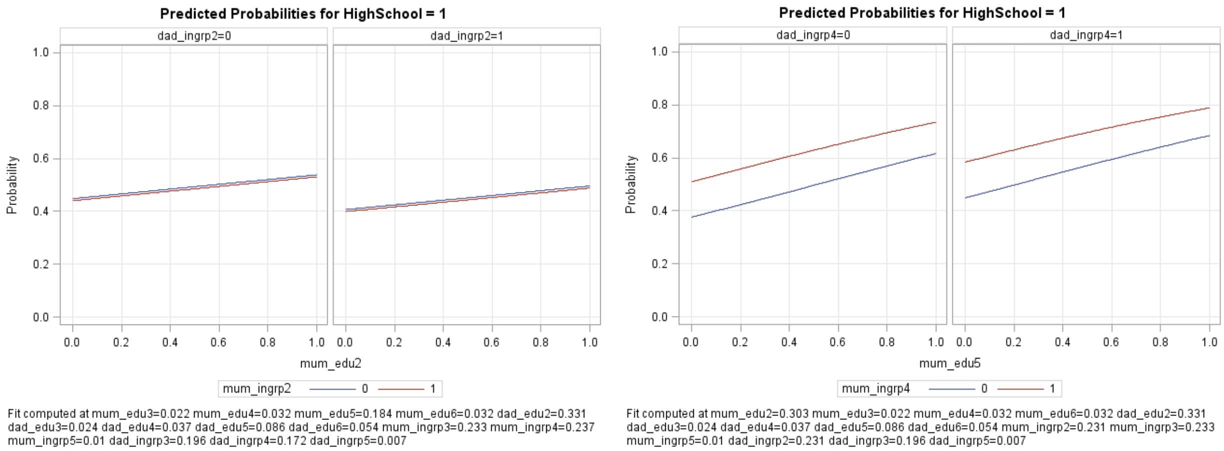
Figure D.5: Predicted probability plots for boys with native origin, data form 2006 when parents do *not* live together



Fit computed at mum_edu3=0.022 mum_edu4=0.031 mum_edu5=0.182 mum_edu6=0.033 dad_edu2=0.333 dad_edu3=0.024 dad_edu4=0.038 dad_edu5=0.085 dad_edu6=0.054 mum_ingrp3=0.23 mum_ingrp4=0.235 mum_ingrp5=0.01 dad_ingrp3=0.198 dad_ingrp4=0.17 dad_ingrp5=0.007

Fit computed at mum_edu2=0.302 mum_edu3=0.022 mum_edu4=0.031 mum_edu6=0.033 dad_edu2=0.333 dad_edu3=0.024 dad_edu4=0.038 dad_edu5=0.085 dad_edu6=0.054 mum_ingrp2=0.232 mum_ingrp3=0.23 mum_ingrp5=0.01 dad_ingrp2=0.229 dad_ingrp3=0.198 dad_ingrp5=0.007

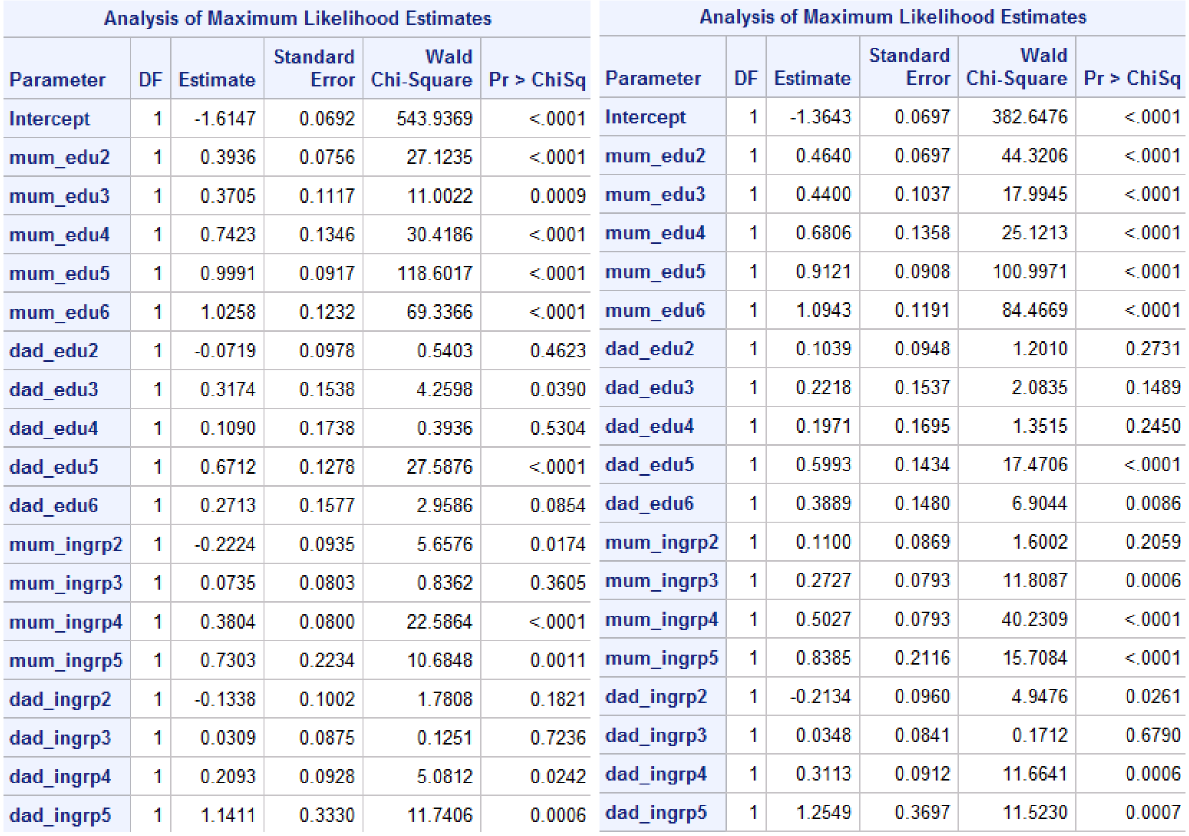
Figure D.6: Predicted probability plots for girls with native origin, data form 2006 when parents do *not* live together



Fit computed at mum_edu3=0.022 mum_edu4=0.032 mum_edu5=0.184 mum_edu6=0.032 dad_edu2=0.331 dad_edu3=0.024 dad_edu4=0.037 dad_edu5=0.086 dad_edu6=0.054 mum_ingrp3=0.233 mum_ingrp4=0.237 mum_ingrp5=0.01 dad_ingrp3=0.196 dad_ingrp4=0.172 dad_ingrp5=0.007

Fit computed at mum_edu2=0.303 mum_edu3=0.022 mum_edu4=0.032 mum_edu6=0.032 dad_edu2=0.331 dad_edu3=0.024 dad_edu4=0.037 dad_edu5=0.086 dad_edu6=0.054 mum_ingrp2=0.231 mum_ingrp3=0.233 mum_ingrp5=0.01 dad_ingrp2=0.231 dad_ingrp3=0.196 dad_ingrp5=0.007

Figure D.7: Left Table: Boys, Right Table: Girls, immigrant origin data form 2006 when parents do *not* live together



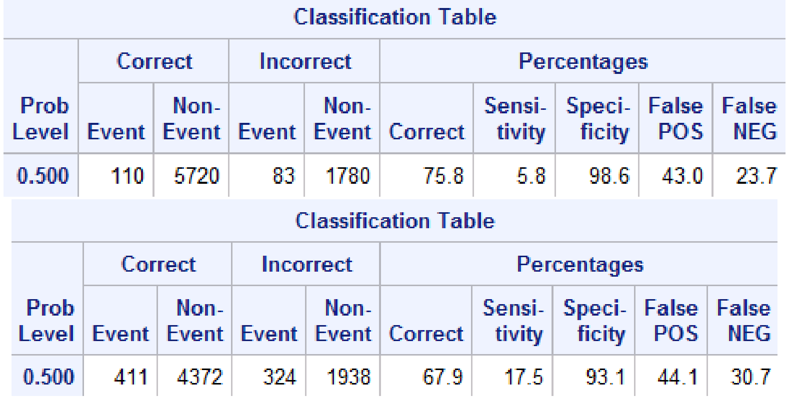
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.6147	0.0692	543.9369	<.0001
mum_edu2	1	0.3936	0.0756	27.1235	<.0001
mum_edu3	1	0.3705	0.1117	11.0022	0.0009
mum_edu4	1	0.7423	0.1346	30.4186	<.0001
mum_edu5	1	0.9991	0.0917	118.6017	<.0001
mum_edu6	1	1.0258	0.1232	69.3366	<.0001
dad_edu2	1	-0.0719	0.0978	0.5403	0.4623
dad_edu3	1	0.3174	0.1538	4.2598	0.0390
dad_edu4	1	0.1090	0.1738	0.3936	0.5304
dad_edu5	1	0.6712	0.1278	27.5876	<.0001
dad_edu6	1	0.2713	0.1577	2.9586	0.0854
mum_ingrp2	1	-0.2224	0.0935	5.6576	0.0174
mum_ingrp3	1	0.0735	0.0803	0.8362	0.3605
mum_ingrp4	1	0.3804	0.0800	22.5864	<.0001
mum_ingrp5	1	0.7303	0.2234	10.6848	0.0011
dad_ingrp2	1	-0.1338	0.1002	1.7808	0.1821
dad_ingrp3	1	0.0309	0.0875	0.1251	0.7236
dad_ingrp4	1	0.2093	0.0928	5.0812	0.0242
dad_ingrp5	1	1.1411	0.3330	11.7406	0.0006

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.3643	0.0697	382.6476	<.0001
mum_edu2	1	0.4640	0.0697	44.3206	<.0001
mum_edu3	1	0.4400	0.1037	17.9945	<.0001
mum_edu4	1	0.6806	0.1358	25.1213	<.0001
mum_edu5	1	0.9121	0.0908	100.9971	<.0001
mum_edu6	1	1.0943	0.1191	84.4669	<.0001
dad_edu2	1	0.1039	0.0948	1.2010	0.2731
dad_edu3	1	0.2218	0.1537	2.0835	0.1489
dad_edu4	1	0.1971	0.1695	1.3515	0.2450
dad_edu5	1	0.5993	0.1434	17.4706	<.0001
dad_edu6	1	0.3889	0.1480	6.9044	0.0086
mum_ingrp2	1	0.1100	0.0869	1.6002	0.2059
mum_ingrp3	1	0.2727	0.0793	11.8087	0.0006
mum_ingrp4	1	0.5027	0.0793	40.2309	<.0001
mum_ingrp5	1	0.8385	0.2116	15.7084	<.0001
dad_ingrp2	1	-0.2134	0.0960	4.9476	0.0261
dad_ingrp3	1	0.0348	0.0841	0.1712	0.6790
dad_ingrp4	1	0.3113	0.0912	11.6641	0.0006
dad_ingrp5	1	1.2549	0.3697	11.5230	0.0007

Figure D.8: Top Table: Boys, Bottom Table: Girls. Immigrant origin, data form 2006 when parents do *not* live together



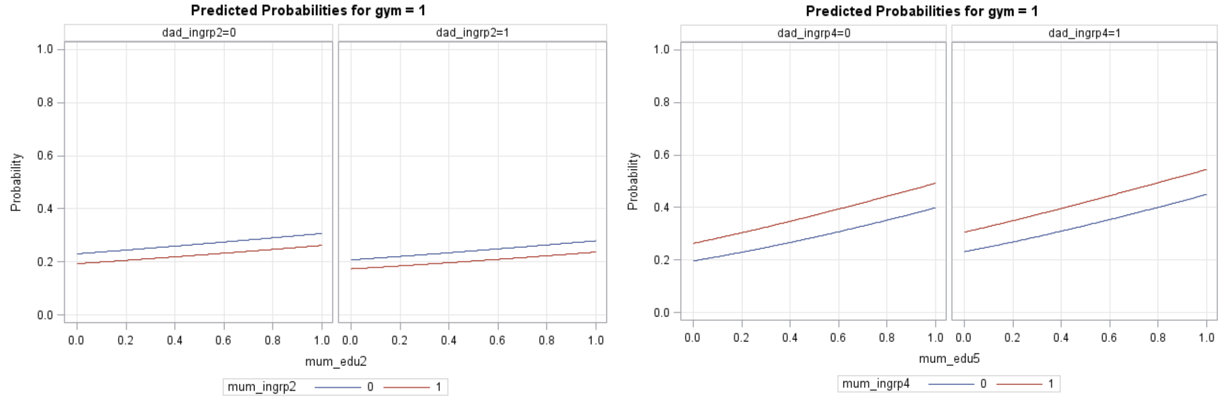
Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	110	5720	83	1780	75.8	5.8	98.6	43.0	23.7

Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	411	4372	324	1938	67.9	17.5	93.1	44.1	30.7

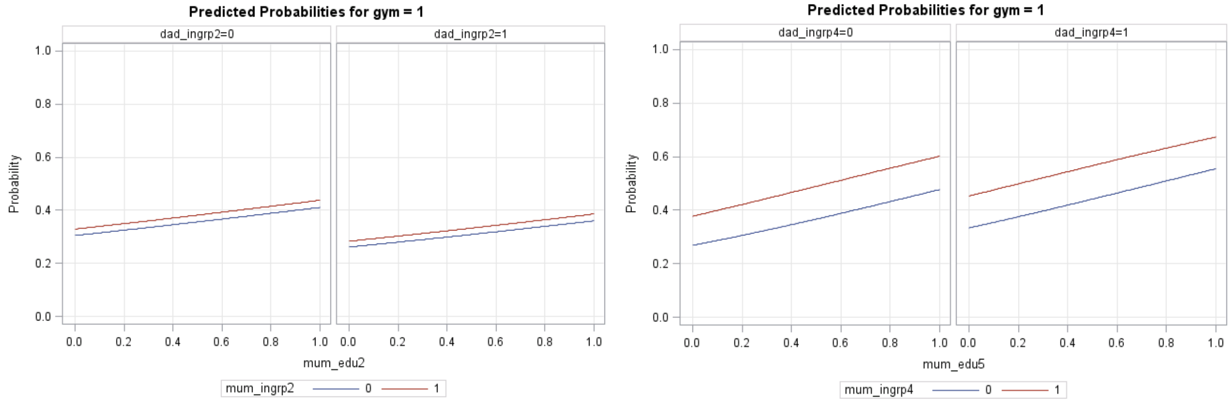
Figure D.9: Predicted probability plots for boys with immigrant origin, data form 2006 when parents do *not* live together



Fit computed at mum_edu3=0.068 mum_edu4=0.037 mum_edu5=0.092 mum_edu6=0.044 dad_edu2=0.128 dad_edu3=0.033 dad_edu4=0.027 dad_edu5=0.046 dad_edu6=0.031 mum_ingrp3=0.262 mum_ingrp4=0.273 mum_ingrp5=0.013 dad_ingrp3=0.183 dad_ingrp4=0.143 dad_ingrp5=0.006

Fit computed at mum_edu2=0.201 mum_edu3=0.068 mum_edu4=0.037 mum_edu6=0.044 dad_edu2=0.128 dad_edu3=0.033 dad_edu4=0.027 dad_edu5=0.046 dad_edu6=0.031 mum_ingrp2=0.174 mum_ingrp3=0.262 mum_ingrp5=0.013 dad_ingrp2=0.129 dad_ingrp3=0.183 dad_ingrp5=0.006

Figure D.10: Predicted probability plots for girls with immigrant origin, data form 2006 when when parents do *not* live together

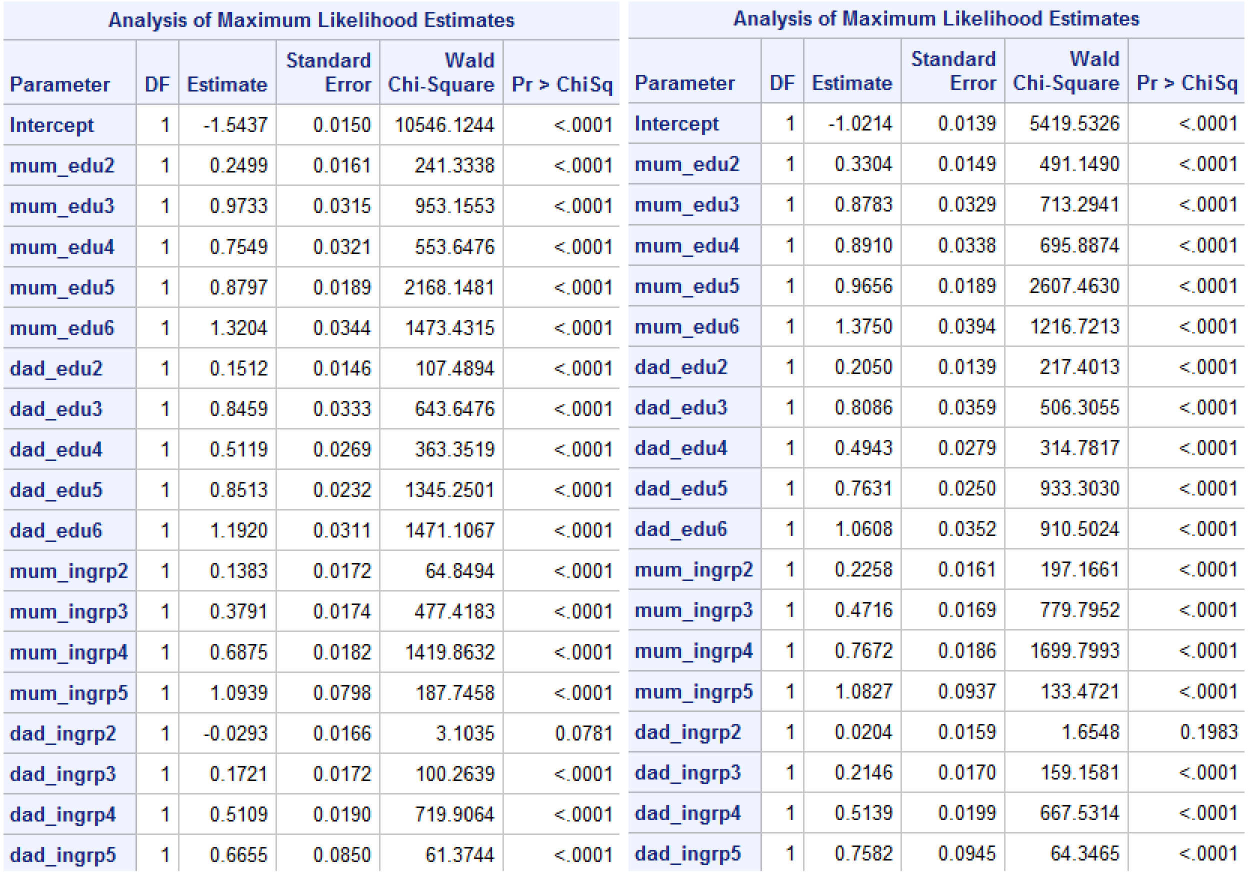


Fit computed at mum_edu3=0.07 mum_edu4=0.036 mum_edu5=0.094 mum_edu6=0.049 dad_edu2=0.112 dad_edu3=0.033 dad_edu4=0.026 dad_edu5=0.036 dad_edu6=0.036 mum_ingrp3=0.268 mum_ingrp4=0.292 mum_ingrp5=0.016 dad_ingrp3=0.171 dad_ingrp4=0.131 dad_ingrp5=0.006

Fit computed at mum_edu2=0.212 mum_edu3=0.07 mum_edu4=0.036 mum_edu6=0.049 dad_edu2=0.112 dad_edu3=0.033 dad_edu4=0.026 dad_edu5=0.036 dad_edu6=0.036 mum_ingrp2=0.189 mum_ingrp3=0.268 mum_ingrp5=0.016 dad_ingrp2=0.117 dad_ingrp3=0.171 dad_ingrp5=0.006

2016

Figure D.11: Left Table: Boys, Right Table: Girls, native origin data form 2016 with parents *not* living together



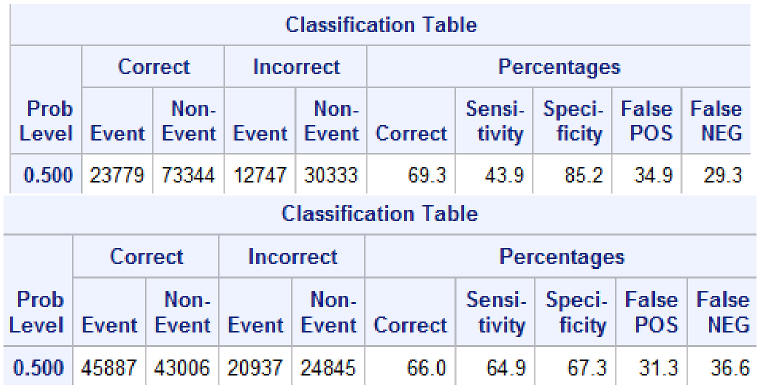
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.5437	0.0150	10546.1244	<.0001
mum_edu2	1	0.2499	0.0161	241.3338	<.0001
mum_edu3	1	0.9733	0.0315	953.1553	<.0001
mum_edu4	1	0.7549	0.0321	553.6476	<.0001
mum_edu5	1	0.8797	0.0189	2168.1481	<.0001
mum_edu6	1	1.3204	0.0344	1473.4315	<.0001
dad_edu2	1	0.1512	0.0146	107.4894	<.0001
dad_edu3	1	0.8459	0.0333	643.6476	<.0001
dad_edu4	1	0.5119	0.0269	363.3519	<.0001
dad_edu5	1	0.8513	0.0232	1345.2501	<.0001
dad_edu6	1	1.1920	0.0311	1471.1067	<.0001
mum_ingrp2	1	0.1383	0.0172	64.8494	<.0001
mum_ingrp3	1	0.3791	0.0174	477.4183	<.0001
mum_ingrp4	1	0.6875	0.0182	1419.8632	<.0001
mum_ingrp5	1	1.0939	0.0798	187.7458	<.0001
dad_ingrp2	1	-0.0293	0.0166	3.1035	0.0781
dad_ingrp3	1	0.1721	0.0172	100.2639	<.0001
dad_ingrp4	1	0.5109	0.0190	719.9064	<.0001
dad_ingrp5	1	0.6655	0.0850	61.3744	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.0214	0.0139	5419.5326	<.0001
mum_edu2	1	0.3304	0.0149	491.1490	<.0001
mum_edu3	1	0.8783	0.0329	713.2941	<.0001
mum_edu4	1	0.8910	0.0338	695.8874	<.0001
mum_edu5	1	0.9656	0.0189	2607.4630	<.0001
mum_edu6	1	1.3750	0.0394	1216.7213	<.0001
dad_edu2	1	0.2050	0.0139	217.4013	<.0001
dad_edu3	1	0.8086	0.0359	506.3055	<.0001
dad_edu4	1	0.4943	0.0279	314.7817	<.0001
dad_edu5	1	0.7631	0.0250	933.3030	<.0001
dad_edu6	1	1.0608	0.0352	910.5024	<.0001
mum_ingrp2	1	0.2258	0.0161	197.1661	<.0001
mum_ingrp3	1	0.4716	0.0169	779.7952	<.0001
mum_ingrp4	1	0.7672	0.0186	1699.7993	<.0001
mum_ingrp5	1	1.0827	0.0937	133.4721	<.0001
dad_ingrp2	1	0.0204	0.0159	1.6548	0.1983
dad_ingrp3	1	0.2146	0.0170	159.1581	<.0001
dad_ingrp4	1	0.5139	0.0199	667.5314	<.0001
dad_ingrp5	1	0.7582	0.0945	64.3465	<.0001

Figure D.12: Top Table: Boys, Bottom Table: Girls, native origin data form 2016 parents *not* living together



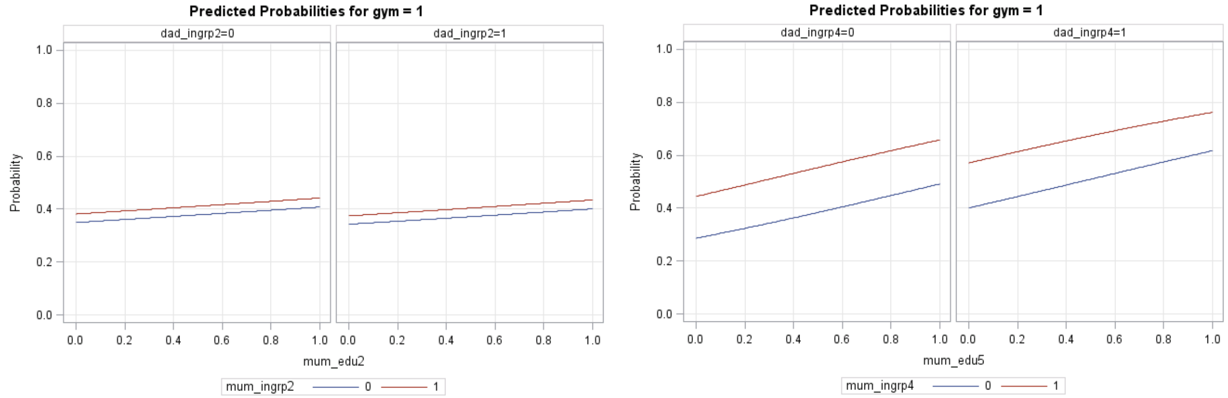
Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	23779	73344	12747	30333	69.3	43.9	85.2	34.9	29.3

Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	45887	43006	20937	24845	66.0	64.9	67.3	31.3	36.6

Figure D.13: Predicted probability plots for boys with native origin, data form 2016 and parents *not* living together



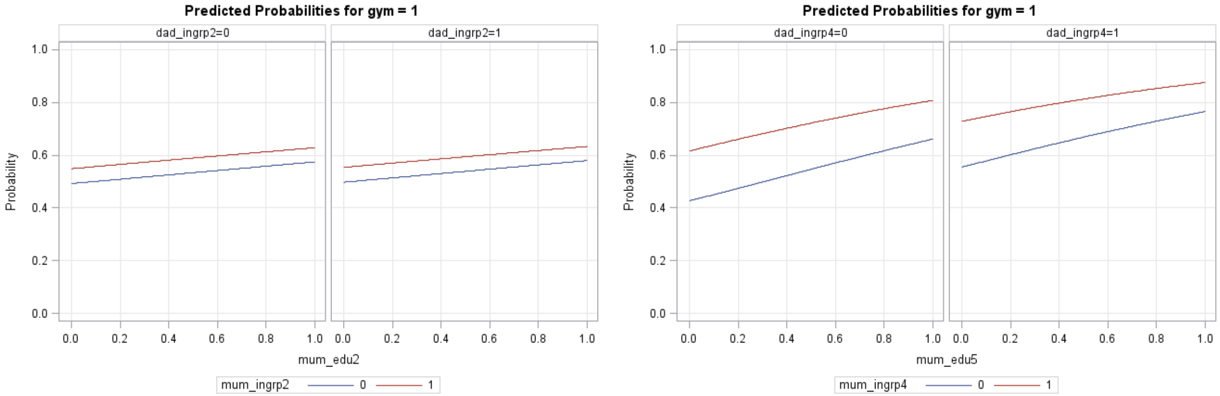
Predicted Probabilities for gym = 1

Predicted Probabilities for gym = 1

Fit computed at mum_edu3=0.038 mum_edu4=0.037 mum_edu5=0.207 mum_edu6=0.05 dad_edu2=0.358 dad_edu3=0.033 dad_edu4=0.053 dad_edu5=0.087 dad_edu6=0.059 mum_ingrp3=0.223 mum_ingrp4=0.212 mum_ingrp5=0.008 dad_ingrp3=0.188 dad_ingrp4=0.158 dad_ingrp5=0.006

Fit computed at mum_edu2=0.374 mum_edu3=0.038 mum_edu4=0.037 mum_edu6=0.05 dad_edu2=0.358 dad_edu3=0.033 dad_edu4=0.053 dad_edu5=0.087 dad_edu6=0.059 mum_ingrp2=0.223 mum_ingrp3=0.223 mum_ingrp5=0.008 dad_ingrp2=0.208 dad_ingrp3=0.188 dad_ingrp5=0.006

Figure D.14: Predicted probability plots for girls with native origin, data form 2016 and parents *not* living together



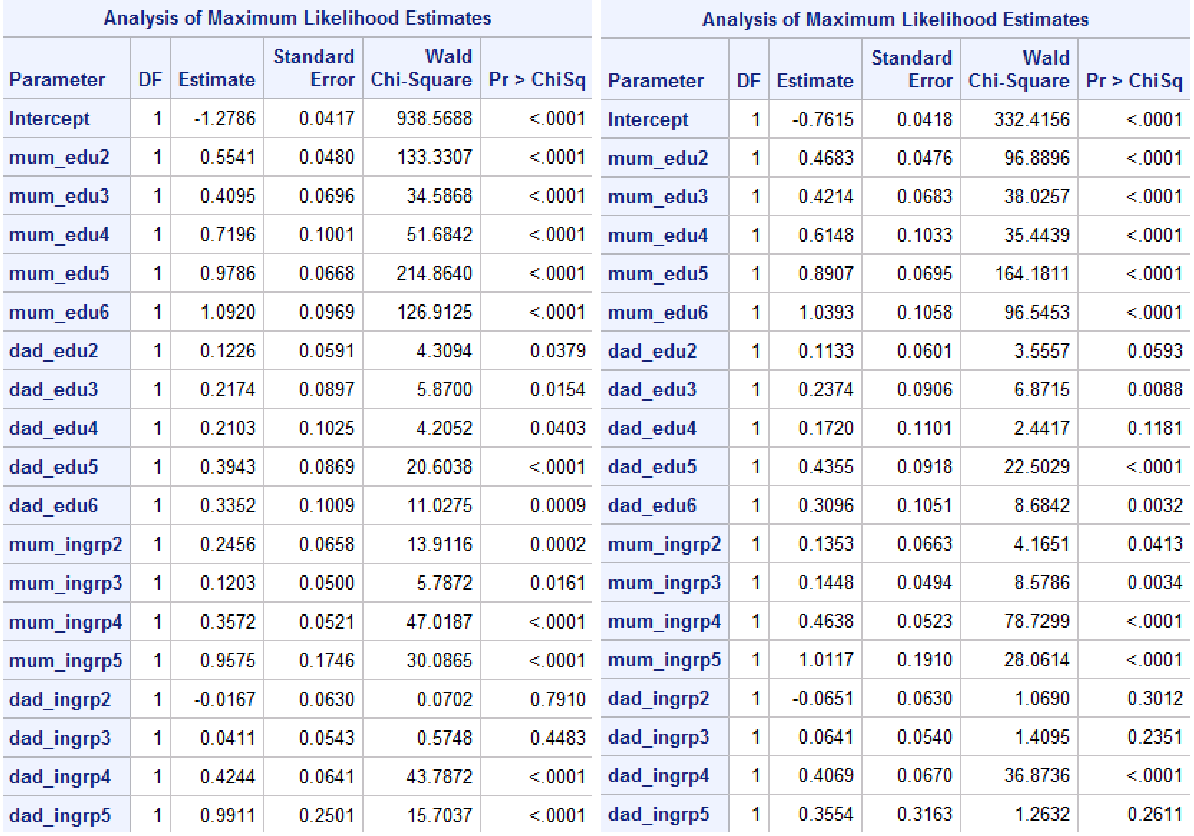
Predicted Probabilities for gym = 1

Predicted Probabilities for gym = 1

Fit computed at mum_edu3=0.038 mum_edu4=0.037 mum_edu5=0.208 mum_edu6=0.049 dad_edu2=0.359 dad_edu3=0.034 dad_edu4=0.053 dad_edu5=0.087 dad_edu6=0.059 mum_ingrp3=0.224 mum_ingrp4=0.211 mum_ingrp5=0.008 dad_ingrp3=0.188 dad_ingrp4=0.159 dad_ingrp5=0.006

Fit computed at mum_edu2=0.374 mum_edu3=0.038 mum_edu4=0.037 mum_edu6=0.049 dad_edu2=0.359 dad_edu3=0.034 dad_edu4=0.053 dad_edu5=0.087 dad_edu6=0.059 mum_ingrp2=0.225 mum_ingrp3=0.224 mum_ingrp5=0.008 dad_ingrp2=0.208 dad_ingrp3=0.188 dad_ingrp5=0.006

Figure D.15: Left Table: Boys, Right Table: Girls, immigrant origin data form 2016 with parents *not* living together



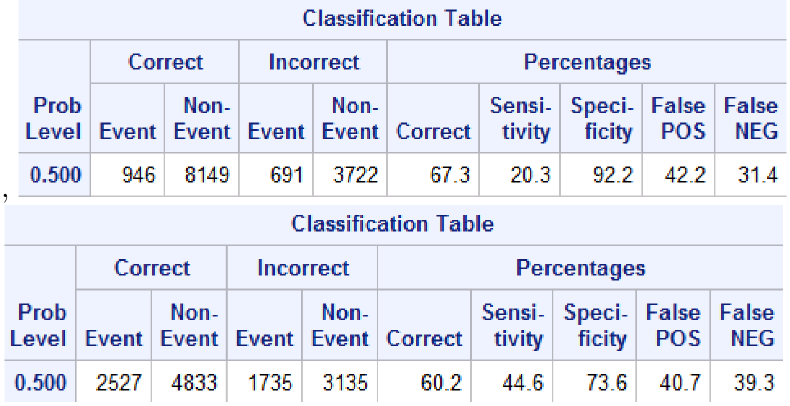
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.2786	0.0417	938.5688	<.0001
mum_edu2	1	0.5541	0.0480	133.3307	<.0001
mum_edu3	1	0.4095	0.0696	34.5868	<.0001
mum_edu4	1	0.7196	0.1001	51.6842	<.0001
mum_edu5	1	0.9786	0.0668	214.8640	<.0001
mum_edu6	1	1.0920	0.0969	126.9125	<.0001
dad_edu2	1	0.1226	0.0591	4.3094	0.0379
dad_edu3	1	0.2174	0.0897	5.8700	0.0154
dad_edu4	1	0.2103	0.1025	4.2052	0.0403
dad_edu5	1	0.3943	0.0869	20.6038	<.0001
dad_edu6	1	0.3352	0.1009	11.0275	0.0009
mum_ingrp2	1	0.2456	0.0658	13.9116	0.0002
mum_ingrp3	1	0.1203	0.0500	5.7872	0.0161
mum_ingrp4	1	0.3572	0.0521	47.0187	<.0001
mum_ingrp5	1	0.9575	0.1746	30.0865	<.0001
dad_ingrp2	1	-0.0167	0.0630	0.0702	0.7910
dad_ingrp3	1	0.0411	0.0543	0.5748	0.4483
dad_ingrp4	1	0.4244	0.0641	43.7872	<.0001
dad_ingrp5	1	0.9911	0.2501	15.7037	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.7615	0.0418	332.4156	<.0001
mum_edu2	1	0.4683	0.0476	96.8896	<.0001
mum_edu3	1	0.4214	0.0683	38.0257	<.0001
mum_edu4	1	0.6148	0.1033	35.4439	<.0001
mum_edu5	1	0.8907	0.0695	164.1811	<.0001
mum_edu6	1	1.0393	0.1058	96.5453	<.0001
dad_edu2	1	0.1133	0.0601	3.5557	0.0593
dad_edu3	1	0.2374	0.0906	6.8715	0.0088
dad_edu4	1	0.1720	0.1101	2.4417	0.1181
dad_edu5	1	0.4355	0.0918	22.5029	<.0001
dad_edu6	1	0.3096	0.1051	8.6842	0.0032
mum_ingrp2	1	0.1353	0.0663	4.1651	0.0413
mum_ingrp3	1	0.1448	0.0494	8.5786	0.0034
mum_ingrp4	1	0.4638	0.0523	78.7299	<.0001
mum_ingrp5	1	1.0117	0.1910	28.0614	<.0001
dad_ingrp2	1	-0.0651	0.0630	1.0690	0.3012
dad_ingrp3	1	0.0641	0.0540	1.4095	0.2351
dad_ingrp4	1	0.4069	0.0670	36.8736	<.0001
dad_ingrp5	1	0.3554	0.3163	1.2632	0.2611

Figure D.16: Top Table: Boys, Bottom Table: Girls. Immigrant origin, data form 2016 with parents *not* living together



Classification Table

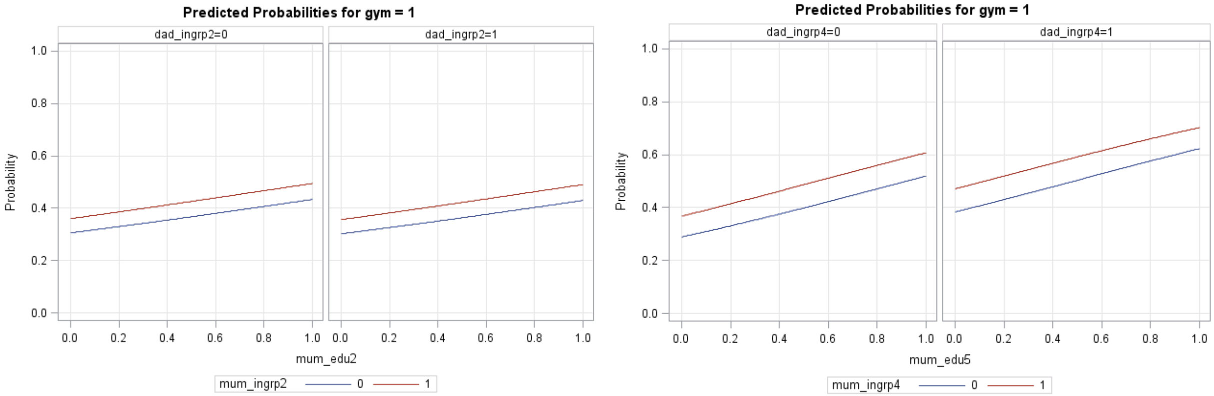
Prob Level	Correct		Incorrect		Percentages				
	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	946	8149	691	3722	67.3	20.3	92.2	42.2	31.4

,

Classification Table

Prob Level	Correct		Incorrect		Percentages				
	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	2527	4833	1735	3135	60.2	44.6	73.6	40.7	39.3

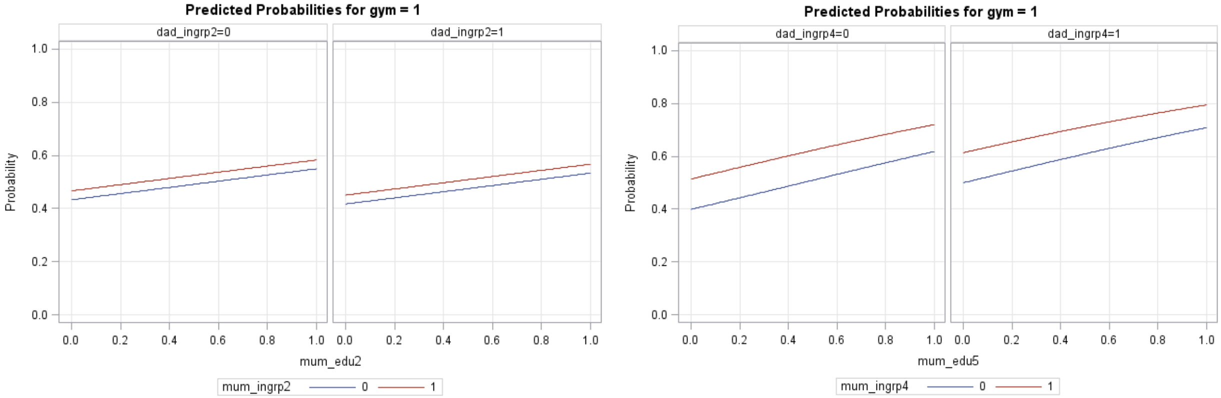
Figure D.17: Predicted probability plots for boys with immigrant origin, data form 2016 with parents *not* living together



Fit computed at mum_edu3=0.081 mum_edu4=0.034 mum_edu5=0.089 mum_edu6=0.04 dad_edu2=0.136 dad_edu3=0.047 dad_edu4=0.035 dad_edu5=0.05 dad_edu6=0.037 mum_ingrp3=0.297 mum_ingrp4=0.269 mum_ingrp5=0.013 dad_ingrp3=0.183 dad_ingrp4=0.111 dad_ingrp5=0.006

Fit computed at mum_edu2=0.226 mum_edu3=0.081 mum_edu4=0.034 mum_edu6=0.04 dad_edu2=0.136 dad_edu3=0.047 dad_edu4=0.035 dad_edu5=0.05 dad_edu6=0.037 mum_ingrp2=0.112 mum_ingrp3=0.297 mum_ingrp5=0.013 dad_ingrp2=0.119 dad_ingrp3=0.183 dad_ingrp5=0.006

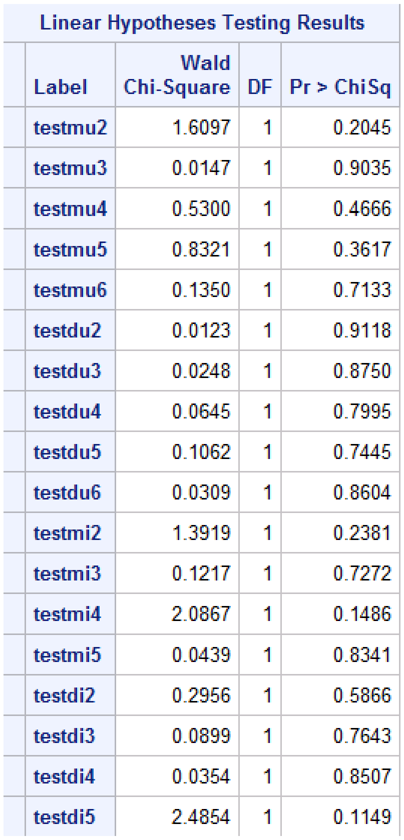
Figure D.18: Predicted probability plots for girls with immigrant origin, data form 2016 with parents *not* living together



Fit computed at mum_edu3=0.085 mum_edu4=0.035 mum_edu5=0.092 mum_edu6=0.04 dad_edu2=0.134 dad_edu3=0.048 dad_edu4=0.031 dad_edu5=0.049 dad_edu6=0.037 mum_ingrp3=0.31 mum_ingrp4=0.284 mum_ingrp5=0.013 dad_ingrp3=0.181 dad_ingrp4=0.11 dad_ingrp5=0.004

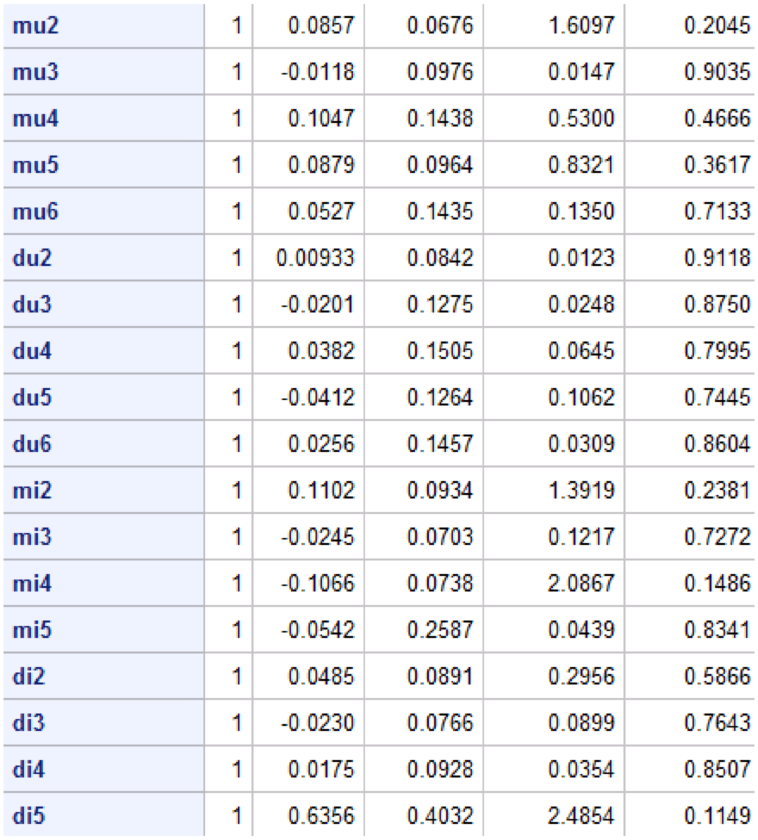
Fit computed at mum_edu2=0.238 mum_edu3=0.085 mum_edu4=0.035 mum_edu6=0.04 dad_edu2=0.134 dad_edu3=0.048 dad_edu4=0.031 dad_edu5=0.049 dad_edu6=0.037 mum_ingrp2=0.113 mum_ingrp3=0.31 mum_ingrp5=0.013 dad_ingrp2=0.117 dad_ingrp3=0.181 dad_ingrp5=0.004

Figure D.19: Left plot: Test $\hat{\beta}_i^{\text{Girls}} = \hat{\beta}_i^{\text{Boys}}$, where i is a parental level of education or income. Right plot: Observation, $\hat{\beta}$, Std.error. Wald. P-value. Immigrant origin, data form 2016 with parents *not* living together



Linear Hypotheses Testing Results

Label	Wald Chi-Square	DF	Pr > ChiSq
testmu2	1.6097	1	0.2045
testmu3	0.0147	1	0.9035
testmu4	0.5300	1	0.4666
testmu5	0.8321	1	0.3617
testmu6	0.1350	1	0.7133
testdu2	0.0123	1	0.9118
testdu3	0.0248	1	0.8750
testdu4	0.0645	1	0.7995
testdu5	0.1062	1	0.7445
testdu6	0.0309	1	0.8604
testmi2	1.3919	1	0.2381
testmi3	0.1217	1	0.7272
testmi4	2.0867	1	0.1486
testmi5	0.0439	1	0.8341
testdi2	0.2956	1	0.5866
testdi3	0.0899	1	0.7643
testdi4	0.0354	1	0.8507
testdi5	2.4854	1	0.1149

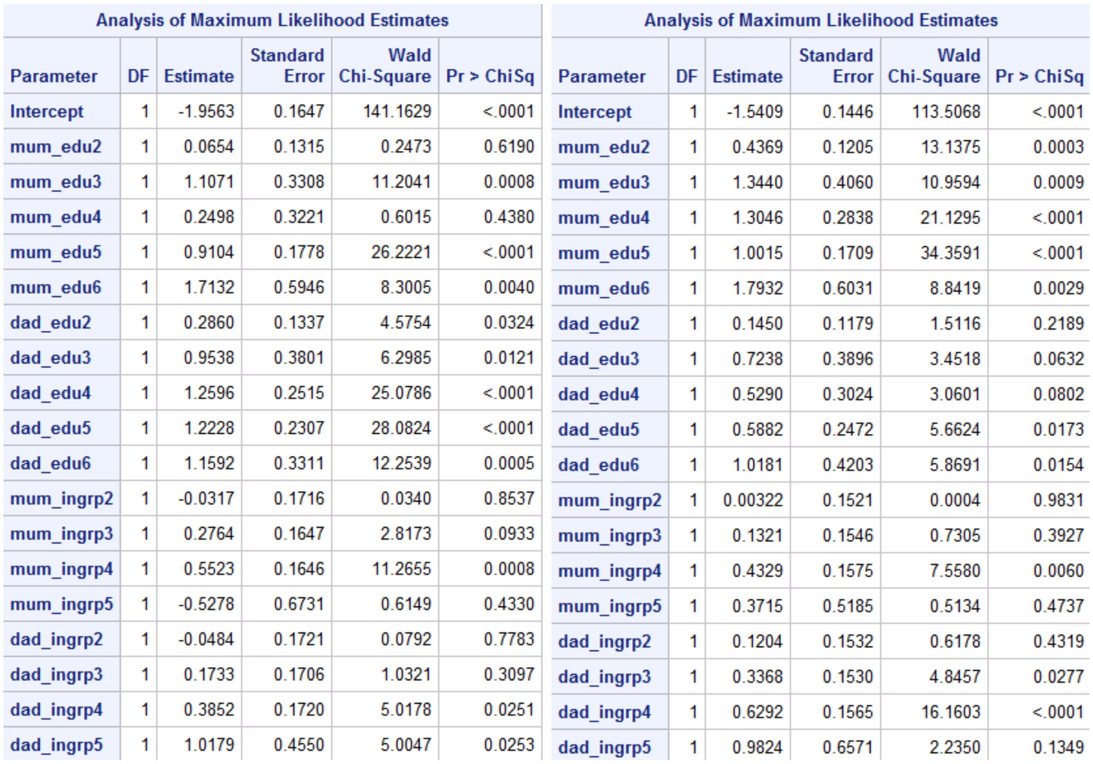


mu2	1	0.0857	0.0676	1.6097	0.2045
mu3	1	-0.0118	0.0976	0.0147	0.9035
mu4	1	0.1047	0.1438	0.5300	0.4666
mu5	1	0.0879	0.0964	0.8321	0.3617
mu6	1	0.0527	0.1435	0.1350	0.7133
du2	1	0.00933	0.0842	0.0123	0.9118
du3	1	-0.0201	0.1275	0.0248	0.8750
du4	1	0.0382	0.1505	0.0645	0.7995
du5	1	-0.0412	0.1264	0.1062	0.7445
du6	1	0.0256	0.1457	0.0309	0.8604
mi2	1	0.1102	0.0934	1.3919	0.2381
mi3	1	-0.0245	0.0703	0.1217	0.7272
mi4	1	-0.1066	0.0738	2.0867	0.1486
mi5	1	-0.0542	0.2587	0.0439	0.8341
di2	1	0.0485	0.0891	0.2956	0.5866
di3	1	-0.0230	0.0766	0.0899	0.7643
di4	1	0.0175	0.0928	0.0354	0.8507
di5	1	0.6356	0.4032	2.4854	0.1149

Appendix E

Brøndby and Ishøj

Figure E.1: Left Table: MLE Boys, Right MLE Table: Girls, native origin data form 2006 Brøndby and Ishøj



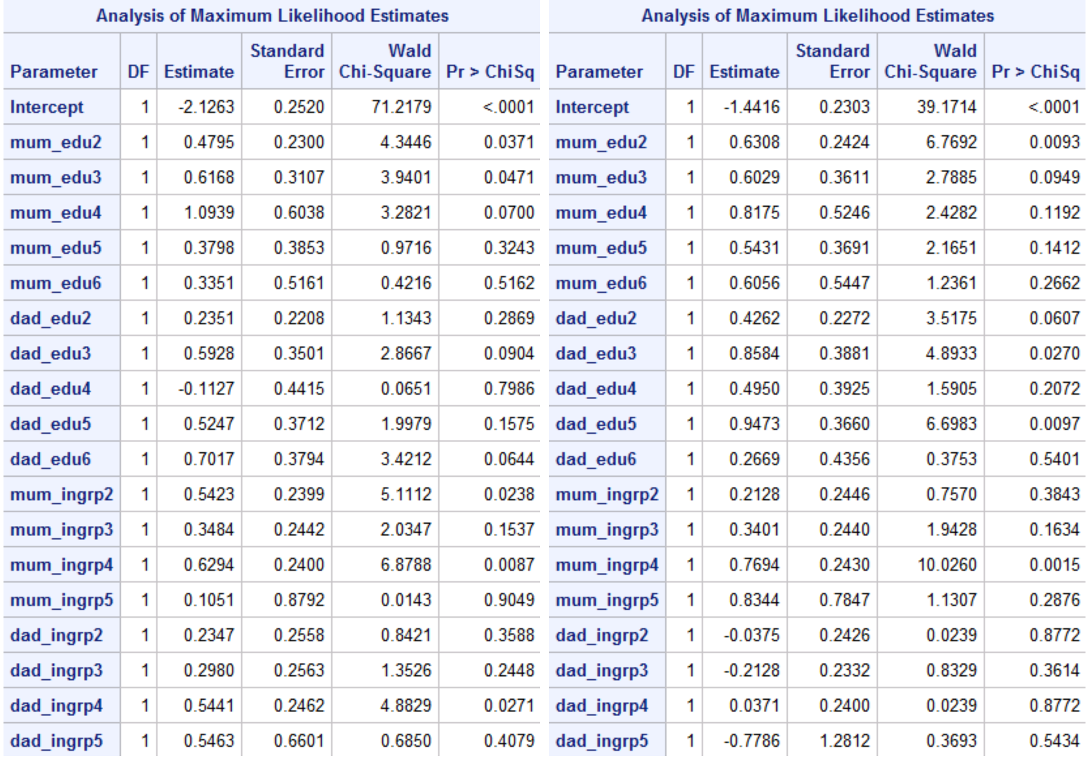
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.9563	0.1647	141.1629	<.0001
mum_edu2	1	0.0654	0.1315	0.2473	0.6190
mum_edu3	1	1.1071	0.3308	11.2041	0.0008
mum_edu4	1	0.2498	0.3221	0.6015	0.4380
mum_edu5	1	0.9104	0.1778	26.2221	<.0001
mum_edu6	1	1.7132	0.5946	8.3005	0.0040
dad_edu2	1	0.2860	0.1337	4.5754	0.0324
dad_edu3	1	0.9538	0.3801	6.2985	0.0121
dad_edu4	1	1.2596	0.2515	25.0786	<.0001
dad_edu5	1	1.2228	0.2307	28.0824	<.0001
dad_edu6	1	1.1592	0.3311	12.2539	0.0005
mum_ingrp2	1	-0.0317	0.1716	0.0340	0.8537
mum_ingrp3	1	0.2764	0.1647	2.8173	0.0933
mum_ingrp4	1	0.5523	0.1646	11.2655	0.0008
mum_ingrp5	1	-0.5278	0.6731	0.6149	0.4330
dad_ingrp2	1	-0.0484	0.1721	0.0792	0.7783
dad_ingrp3	1	0.1733	0.1706	1.0321	0.3097
dad_ingrp4	1	0.3852	0.1720	5.0178	0.0251
dad_ingrp5	1	1.0179	0.4550	5.0047	0.0253

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.5409	0.1446	113.5068	<.0001
mum_edu2	1	0.4369	0.1205	13.1375	0.0003
mum_edu3	1	1.3440	0.4060	10.9594	0.0009
mum_edu4	1	1.3046	0.2838	21.1295	<.0001
mum_edu5	1	1.0015	0.1709	34.3591	<.0001
mum_edu6	1	1.7932	0.6031	8.8419	0.0029
dad_edu2	1	0.1450	0.1179	1.5116	0.2189
dad_edu3	1	0.7238	0.3896	3.4518	0.0632
dad_edu4	1	0.5290	0.3024	3.0601	0.0802
dad_edu5	1	0.5882	0.2472	5.6624	0.0173
dad_edu6	1	1.0181	0.4203	5.8691	0.0154
mum_ingrp2	1	0.00322	0.1521	0.0004	0.9831
mum_ingrp3	1	0.1321	0.1546	0.7305	0.3927
mum_ingrp4	1	0.4329	0.1575	7.5580	0.0060
mum_ingrp5	1	0.3715	0.5185	0.5134	0.4737
dad_ingrp2	1	0.1204	0.1532	0.6178	0.4319
dad_ingrp3	1	0.3368	0.1530	4.8457	0.0277
dad_ingrp4	1	0.6292	0.1565	16.1603	<.0001
dad_ingrp5	1	0.9824	0.6571	2.2350	0.1349

Figure E.2: Left Table: MLE Boys, Right MLE Table: Girls, Immigrant origin data form 2006 Brøndby and Ishøj



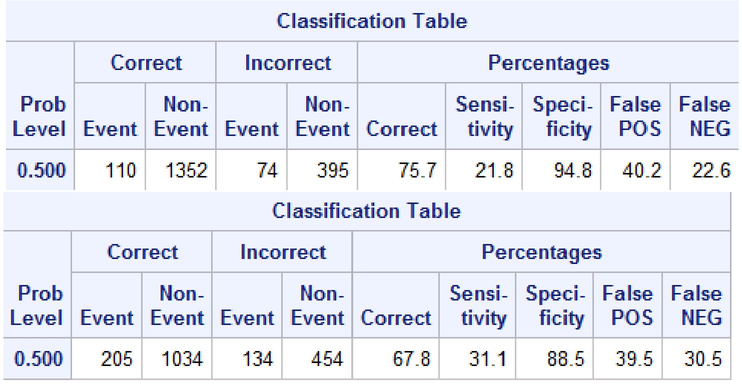
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.1263	0.2520	71.2179	<.0001
mum_edu2	1	0.4795	0.2300	4.3446	0.0371
mum_edu3	1	0.6168	0.3107	3.9401	0.0471
mum_edu4	1	1.0939	0.6038	3.2821	0.0700
mum_edu5	1	0.3798	0.3853	0.9716	0.3243
mum_edu6	1	0.3351	0.5161	0.4216	0.5162
dad_edu2	1	0.2351	0.2208	1.1343	0.2869
dad_edu3	1	0.5928	0.3501	2.8667	0.0904
dad_edu4	1	-0.1127	0.4415	0.0651	0.7986
dad_edu5	1	0.5247	0.3712	1.9979	0.1575
dad_edu6	1	0.7017	0.3794	3.4212	0.0644
mum_ingrp2	1	0.5423	0.2399	5.1112	0.0238
mum_ingrp3	1	0.3484	0.2442	2.0347	0.1537
mum_ingrp4	1	0.6294	0.2400	6.8788	0.0087
mum_ingrp5	1	0.1051	0.8792	0.0143	0.9049
dad_ingrp2	1	0.2347	0.2558	0.8421	0.3588
dad_ingrp3	1	0.2980	0.2563	1.3526	0.2448
dad_ingrp4	1	0.5441	0.2462	4.8829	0.0271
dad_ingrp5	1	0.5463	0.6601	0.6850	0.4079

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.4416	0.2303	39.1714	<.0001
mum_edu2	1	0.6308	0.2424	6.7692	0.0093
mum_edu3	1	0.6029	0.3611	2.7885	0.0949
mum_edu4	1	0.8175	0.5246	2.4282	0.1192
mum_edu5	1	0.5431	0.3691	2.1651	0.1412
mum_edu6	1	0.6056	0.5447	1.2361	0.2662
dad_edu2	1	0.4262	0.2272	3.5175	0.0607
dad_edu3	1	0.8584	0.3881	4.8933	0.0270
dad_edu4	1	0.4950	0.3925	1.5905	0.2072
dad_edu5	1	0.9473	0.3660	6.6983	0.0097
dad_edu6	1	0.2669	0.4356	0.3753	0.5401
mum_ingrp2	1	0.2128	0.2446	0.7570	0.3843
mum_ingrp3	1	0.3401	0.2440	1.9428	0.1634
mum_ingrp4	1	0.7694	0.2430	10.0260	0.0015
mum_ingrp5	1	0.8344	0.7847	1.1307	0.2876
dad_ingrp2	1	-0.0375	0.2426	0.0239	0.8772
dad_ingrp3	1	-0.2128	0.2332	0.8329	0.3614
dad_ingrp4	1	0.0371	0.2400	0.0239	0.8772
dad_ingrp5	1	-0.7786	1.2812	0.3693	0.5434

Figure E.3: Top Table: Boys, Bottom Table: Girls, Classification tables native origin data form 2006 Brøndby and Ishøj



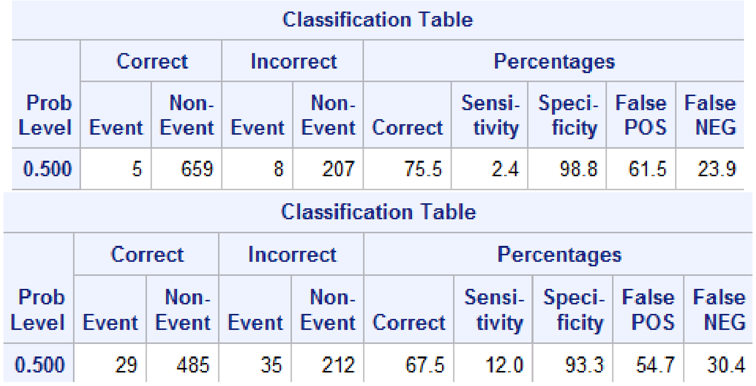
Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensitivity	Specificity	False POS	False NEG
0.500	110	1352	74	395	75.7	21.8	94.8	40.2	22.6

Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensitivity	Specificity	False POS	False NEG
0.500	205	1034	134	454	67.8	31.1	88.5	39.5	30.5

Figure E.4: Top Table: Boys, Bottom Table: Girls, Classification tables immigrant origin data form 2006 Brøndby and Ishøj



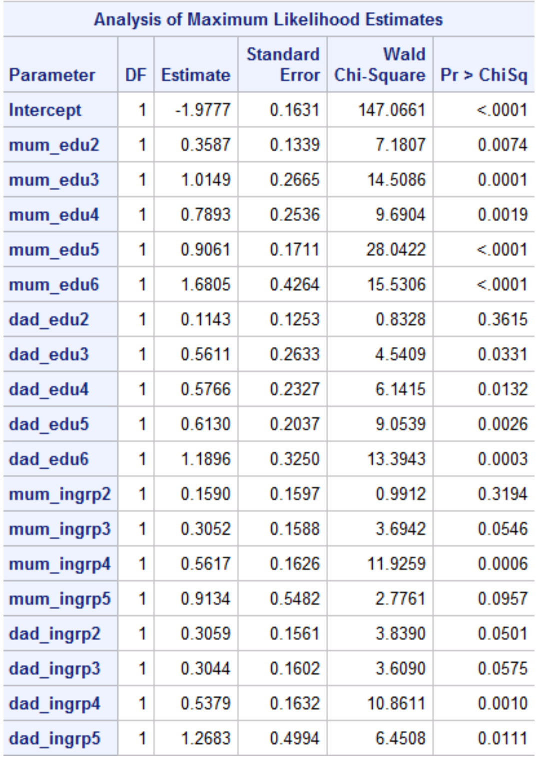
Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensitivity	Specificity	False POS	False NEG
0.500	5	659	8	207	75.5	2.4	98.8	61.5	23.9

Classification Table

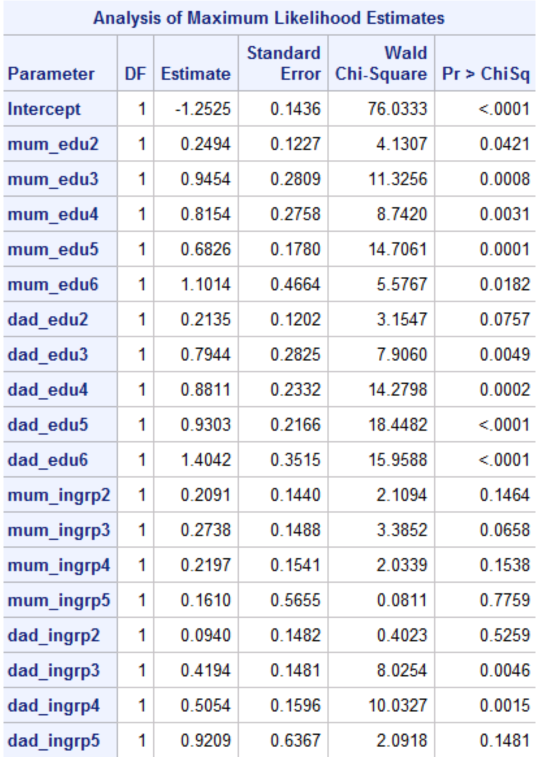
	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensitivity	Specificity	False POS	False NEG
0.500	29	485	35	212	67.5	12.0	93.3	54.7	30.4

Figure E.5: Left Table: MLE Boys, Right MLE Table: Girls, native origin data form 2016 Brøndby and Ishøj



Analysis of Maximum Likelihood Estimates

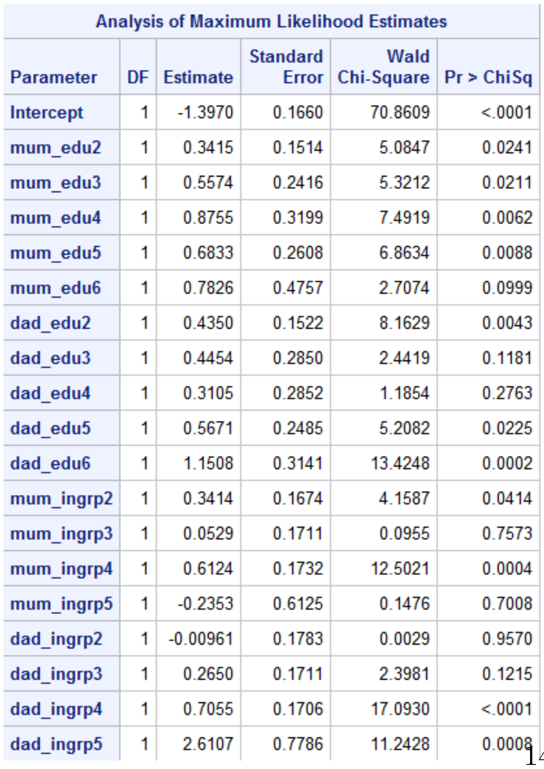
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.9777	0.1631	147.0661	<.0001
mum_edu2	1	0.3587	0.1339	7.1807	0.0074
mum_edu3	1	1.0149	0.2665	14.5086	0.0001
mum_edu4	1	0.7893	0.2536	9.6904	0.0019
mum_edu5	1	0.9061	0.1711	28.0422	<.0001
mum_edu6	1	1.6805	0.4264	15.5306	<.0001
dad_edu2	1	0.1143	0.1253	0.8328	0.3615
dad_edu3	1	0.5611	0.2633	4.5409	0.0331
dad_edu4	1	0.5766	0.2327	6.1415	0.0132
dad_edu5	1	0.6130	0.2037	9.0539	0.0026
dad_edu6	1	1.1896	0.3250	13.3943	0.0003
mum_ingrp2	1	0.1590	0.1597	0.9912	0.3194
mum_ingrp3	1	0.3052	0.1588	3.6942	0.0546
mum_ingrp4	1	0.5617	0.1626	11.9259	0.0006
mum_ingrp5	1	0.9134	0.5482	2.7761	0.0957
dad_ingrp2	1	0.3059	0.1561	3.8390	0.0501
dad_ingrp3	1	0.3044	0.1602	3.6090	0.0575
dad_ingrp4	1	0.5379	0.1632	10.8611	0.0010
dad_ingrp5	1	1.2683	0.4994	6.4508	0.0111



Analysis of Maximum Likelihood Estimates

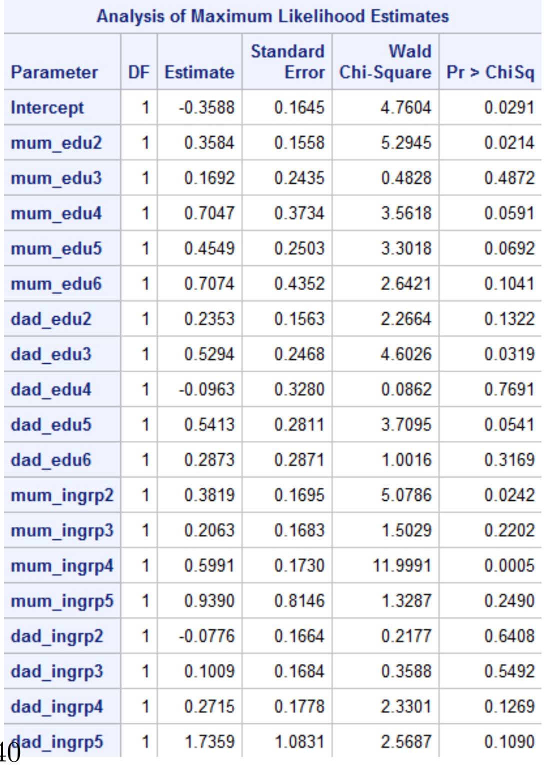
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.2525	0.1436	76.0333	<.0001
mum_edu2	1	0.2494	0.1227	4.1307	0.0421
mum_edu3	1	0.9454	0.2809	11.3256	0.0008
mum_edu4	1	0.8154	0.2758	8.7420	0.0031
mum_edu5	1	0.6826	0.1780	14.7061	0.0001
mum_edu6	1	1.1014	0.4664	5.5767	0.0182
dad_edu2	1	0.2135	0.1202	3.1547	0.0757
dad_edu3	1	0.7944	0.2825	7.9060	0.0049
dad_edu4	1	0.8811	0.2332	14.2798	0.0002
dad_edu5	1	0.9303	0.2166	18.4482	<.0001
dad_edu6	1	1.4042	0.3515	15.9588	<.0001
mum_ingrp2	1	0.2091	0.1440	2.1094	0.1464
mum_ingrp3	1	0.2738	0.1488	3.3852	0.0658
mum_ingrp4	1	0.2197	0.1541	2.0339	0.1538
mum_ingrp5	1	0.1610	0.5655	0.0811	0.7759
dad_ingrp2	1	0.0940	0.1482	0.4023	0.5259
dad_ingrp3	1	0.4194	0.1481	8.0254	0.0046
dad_ingrp4	1	0.5054	0.1596	10.0327	0.0015
dad_ingrp5	1	0.9209	0.6367	2.0918	0.1481

Figure E.6: Left Table: MLE Boys, Right MLE Table: Girls, immigrant origin data form 2016 Brøndby and Ishøj



Analysis of Maximum Likelihood Estimates

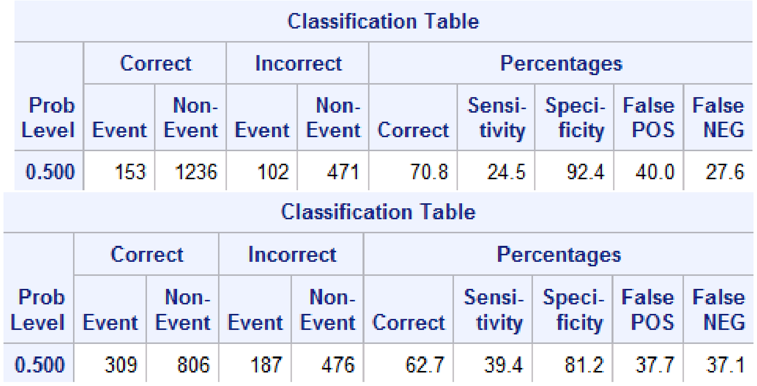
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.3970	0.1660	70.8609	<.0001
mum_edu2	1	0.3415	0.1514	5.0847	0.0241
mum_edu3	1	0.5574	0.2416	5.3212	0.0211
mum_edu4	1	0.8755	0.3199	7.4919	0.0062
mum_edu5	1	0.6833	0.2608	6.8634	0.0088
mum_edu6	1	0.7826	0.4757	2.7074	0.0999
dad_edu2	1	0.4350	0.1522	8.1629	0.0043
dad_edu3	1	0.4454	0.2850	2.4419	0.1181
dad_edu4	1	0.3105	0.2852	1.1854	0.2763
dad_edu5	1	0.5671	0.2485	5.2082	0.0225
dad_edu6	1	1.1508	0.3141	13.4248	0.0002
mum_ingrp2	1	0.3414	0.1674	4.1587	0.0414
mum_ingrp3	1	0.0529	0.1711	0.0955	0.7573
mum_ingrp4	1	0.6124	0.1732	12.5021	0.0004
mum_ingrp5	1	-0.2353	0.6125	0.1476	0.7008
dad_ingrp2	1	-0.00961	0.1783	0.0029	0.9570
dad_ingrp3	1	0.2650	0.1711	2.3981	0.1215
dad_ingrp4	1	0.7055	0.1706	17.0930	<.0001
dad_ingrp5	1	2.6107	0.7786	11.2428	0.0008



Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.3588	0.1645	4.7604	0.0291
mum_edu2	1	0.3584	0.1558	5.2945	0.0214
mum_edu3	1	0.1692	0.2435	0.4828	0.4872
mum_edu4	1	0.7047	0.3734	3.5618	0.0591
mum_edu5	1	0.4549	0.2503	3.3018	0.0692
mum_edu6	1	0.7074	0.4352	2.6421	0.1041
dad_edu2	1	0.2353	0.1563	2.2664	0.1322
dad_edu3	1	0.5294	0.2468	4.6026	0.0319
dad_edu4	1	-0.0963	0.3280	0.0862	0.7691
dad_edu5	1	0.5413	0.2811	3.7095	0.0541
dad_edu6	1	0.2873	0.2871	1.0016	0.3169
mum_ingrp2	1	0.3819	0.1695	5.0786	0.0242
mum_ingrp3	1	0.2063	0.1683	1.5029	0.2202
mum_ingrp4	1	0.5991	0.1730	11.9991	0.0005
mum_ingrp5	1	0.9390	0.8146	1.3287	0.2490
dad_ingrp2	1	-0.0776	0.1664	0.2177	0.6408
dad_ingrp3	1	0.1009	0.1684	0.3588	0.5492
dad_ingrp4	1	0.2715	0.1778	2.3301	0.1269
dad_ingrp5	1	1.7359	1.0831	2.5687	0.1090

Figure E.7: Top Table: Boys, Bottom Table: Girls, Classification tables native origin data form 2016 Brøndby and Ishøj



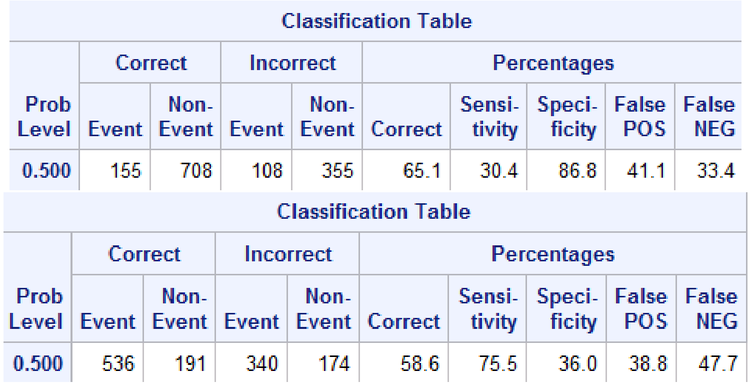
Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	153	1236	102	471	70.8	24.5	92.4	40.0	27.6

Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	309	806	187	476	62.7	39.4	81.2	37.7	37.1

Figure E.8: Top Table: Boys, Bottom Table: Girls, Classification tables immigrant origin data form 2006 Brøndby and Ishøj



Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	155	708	108	355	65.1	30.4	86.8	41.1	33.4

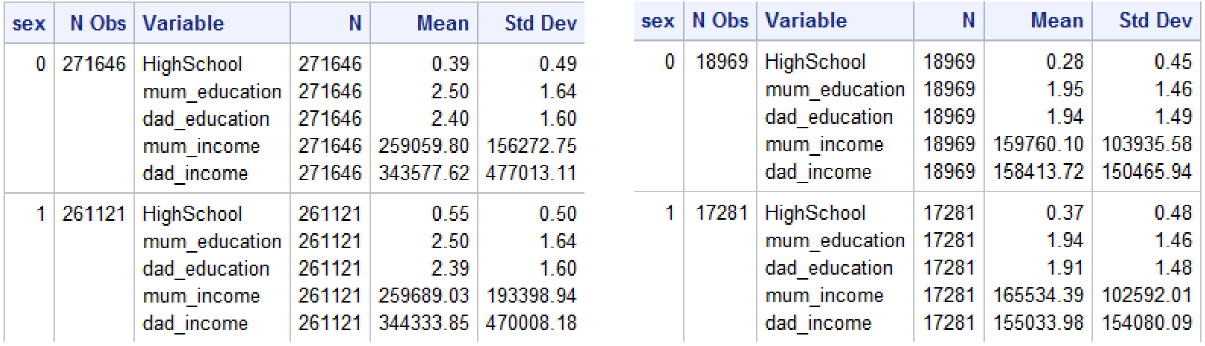
Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	536	191	340	174	58.6	75.5	36.0	38.8	47.7

Appendix F

Kolding and Vejle

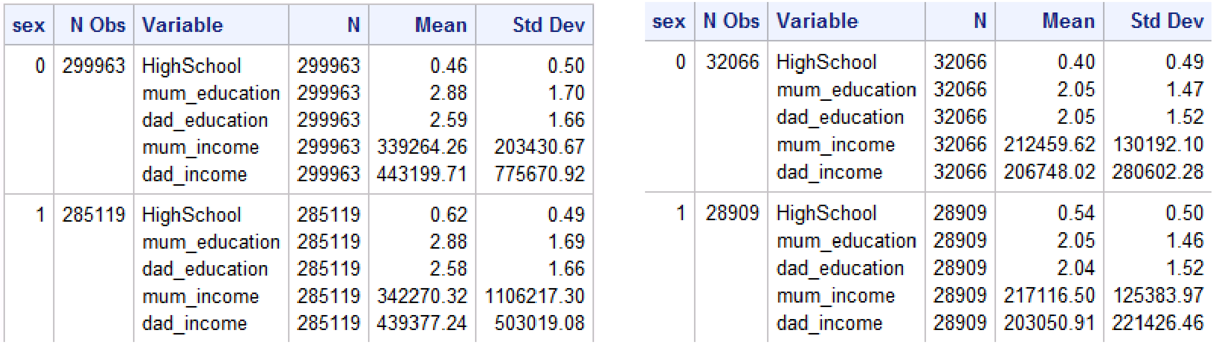
Figure F.1: Left Table: Native origin, Right Table: Immigrant origin. Parental information 2006



sex	N Obs	Variable	N	Mean	Std Dev
0	271646	HighSchool	271646	0.39	0.49
		mum_education	271646	2.50	1.64
		dad_education	271646	2.40	1.60
		mum_income	271646	259059.80	156272.75
		dad_income	271646	343577.62	477013.11
1	261121	HighSchool	261121	0.55	0.50
		mum_education	261121	2.50	1.64
		dad_education	261121	2.39	1.60
		mum_income	261121	259689.03	193398.94
		dad_income	261121	344333.85	470008.18

sex	N Obs	Variable	N	Mean	Std Dev
0	18969	HighSchool	18969	0.28	0.45
		mum_education	18969	1.95	1.46
		dad_education	18969	1.94	1.49
		mum_income	18969	159760.10	103935.58
		dad_income	18969	158413.72	150465.94
1	17281	HighSchool	17281	0.37	0.48
		mum_education	17281	1.94	1.46
		dad_education	17281	1.91	1.48
		mum_income	17281	165534.39	102592.01
		dad_income	17281	155033.98	154080.09

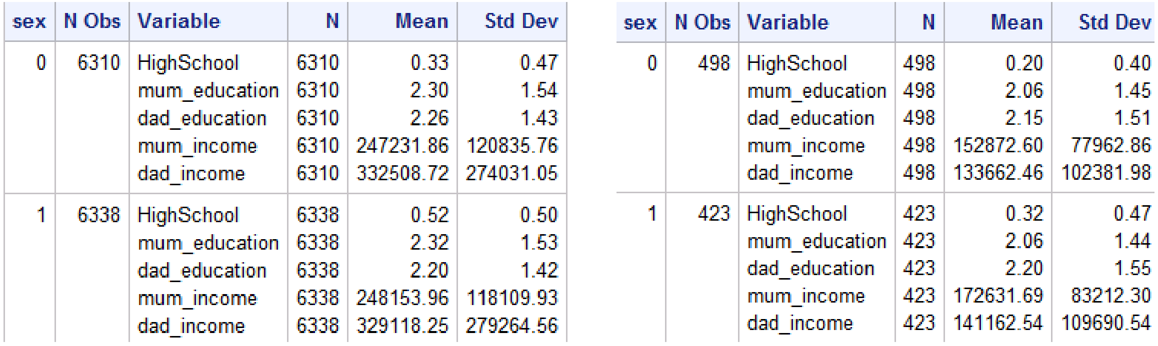
Figure F.2: Left Table: Native origin, Right Table: Immigrant origin. Parental information 2016



sex	N Obs	Variable	N	Mean	Std Dev
0	299963	HighSchool	299963	0.46	0.50
		mum_education	299963	2.88	1.70
		dad_education	299963	2.59	1.66
		mum_income	299963	339264.26	203430.67
		dad_income	299963	443199.71	775670.92
1	285119	HighSchool	285119	0.62	0.49
		mum_education	285119	2.88	1.69
		dad_education	285119	2.58	1.66
		mum_income	285119	342270.32	1106217.30
		dad_income	285119	439377.24	503019.08

sex	N Obs	Variable	N	Mean	Std Dev
0	32066	HighSchool	32066	0.40	0.49
		mum_education	32066	2.05	1.47
		dad_education	32066	2.05	1.52
		mum_income	32066	212459.62	130192.10
		dad_income	32066	206748.02	280602.28
1	28909	HighSchool	28909	0.54	0.50
		mum_education	28909	2.05	1.46
		dad_education	28909	2.04	1.52
		mum_income	28909	217116.50	125383.97
		dad_income	28909	203050.91	221426.46

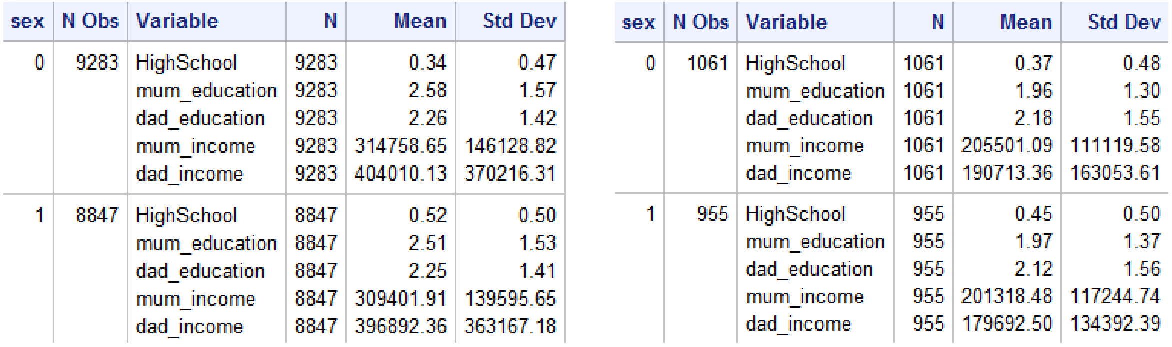
Figure F.3: Left Table: Native origin, Right Table: Immigrant origin. Parental information in Kolding and Vejle 2006



sex	N Obs	Variable	N	Mean	Std Dev
0	6310	HighSchool	6310	0.33	0.47
		mum_education	6310	2.30	1.54
		dad_education	6310	2.26	1.43
		mum_income	6310	247231.86	120835.76
		dad_income	6310	332508.72	274031.05
1	6338	HighSchool	6338	0.52	0.50
		mum_education	6338	2.32	1.53
		dad_education	6338	2.20	1.42
		mum_income	6338	248153.96	118109.93
		dad_income	6338	329118.25	279264.56

sex	N Obs	Variable	N	Mean	Std Dev
0	498	HighSchool	498	0.20	0.40
		mum_education	498	2.06	1.45
		dad_education	498	2.15	1.51
		mum_income	498	152872.60	77962.86
		dad_income	498	133662.46	102381.98
1	423	HighSchool	423	0.32	0.47
		mum_education	423	2.06	1.44
		dad_education	423	2.20	1.55
		mum_income	423	172631.69	83212.30
		dad_income	423	141162.54	109690.54

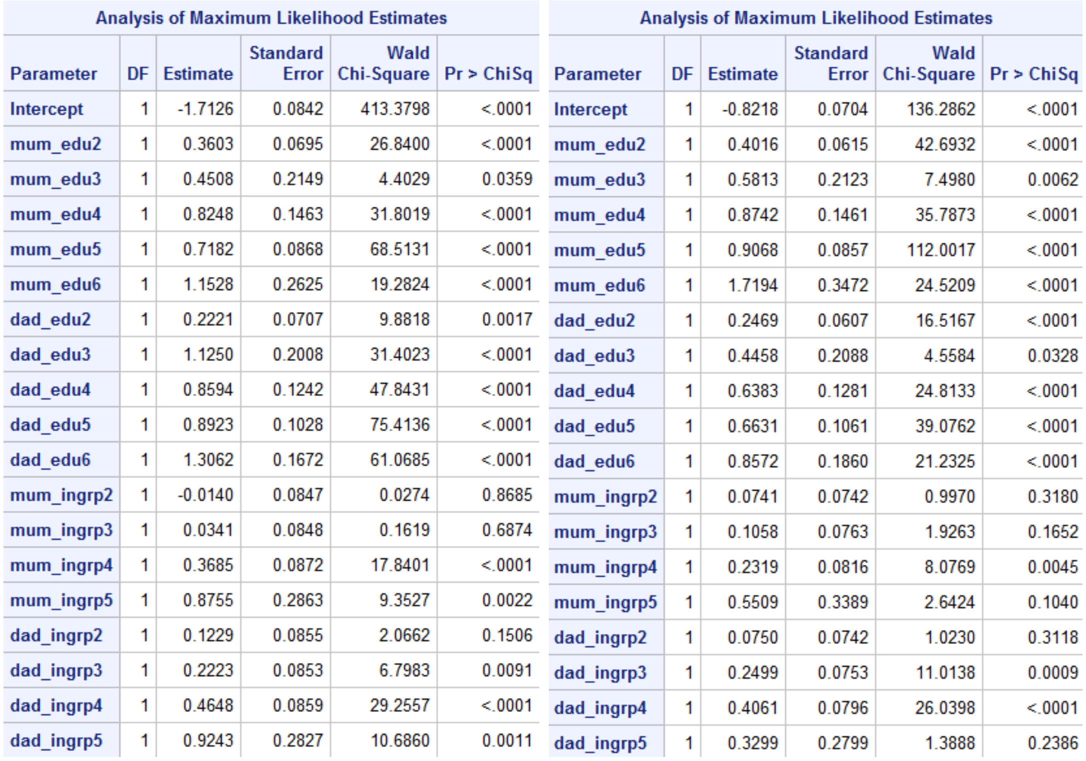
Figure F.4: Left Table: Native origin, Right Table: Immigrant origin. Parental information in Kolding and Vejle 2016



sex	N Obs	Variable	N	Mean	Std Dev
0	9283	HighSchool	9283	0.34	0.47
		mum_education	9283	2.58	1.57
		dad_education	9283	2.26	1.42
		mum_income	9283	314758.65	146128.82
		dad_income	9283	404010.13	370216.31
1	8847	HighSchool	8847	0.52	0.50
		mum_education	8847	2.51	1.53
		dad_education	8847	2.25	1.41
		mum_income	8847	309401.91	139595.65
		dad_income	8847	396892.36	363167.18

sex	N Obs	Variable	N	Mean	Std Dev
0	1061	HighSchool	1061	0.37	0.48
		mum_education	1061	1.96	1.30
		dad_education	1061	2.18	1.55
		mum_income	1061	205501.09	111119.58
		dad_income	1061	190713.36	163053.61
1	955	HighSchool	955	0.45	0.50
		mum_education	955	1.97	1.37
		dad_education	955	2.12	1.56
		mum_income	955	201318.48	117244.74
		dad_income	955	179692.50	134392.39

Figure F.5: Left Table: MLE Boys, Right MLE Table: Girls, native origin data form 2006 Kolding and Vejle



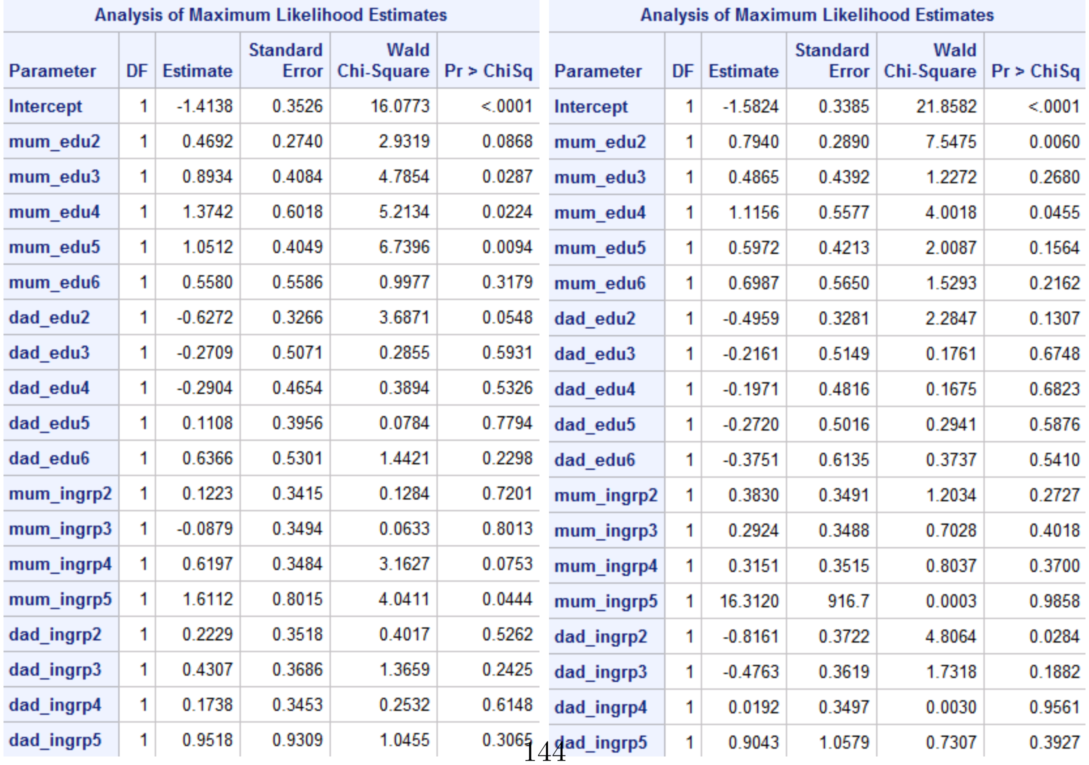
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.7126	0.0842	413.3798	<.0001
mum_edu2	1	0.3603	0.0695	26.8400	<.0001
mum_edu3	1	0.4508	0.2149	4.4029	0.0359
mum_edu4	1	0.8248	0.1463	31.8019	<.0001
mum_edu5	1	0.7182	0.0868	68.5131	<.0001
mum_edu6	1	1.1528	0.2625	19.2824	<.0001
dad_edu2	1	0.2221	0.0707	9.8818	0.0017
dad_edu3	1	1.1250	0.2008	31.4023	<.0001
dad_edu4	1	0.8594	0.1242	47.8431	<.0001
dad_edu5	1	0.8923	0.1028	75.4136	<.0001
dad_edu6	1	1.3062	0.1672	61.0685	<.0001
mum_ingrp2	1	-0.0140	0.0847	0.0274	0.8685
mum_ingrp3	1	0.0341	0.0848	0.1619	0.6874
mum_ingrp4	1	0.3685	0.0872	17.8401	<.0001
mum_ingrp5	1	0.8755	0.2863	9.3527	0.0022
dad_ingrp2	1	0.1229	0.0855	2.0662	0.1506
dad_ingrp3	1	0.2223	0.0853	6.7983	0.0091
dad_ingrp4	1	0.4648	0.0859	29.2557	<.0001
dad_ingrp5	1	0.9243	0.2827	10.6860	0.0011

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.8218	0.0704	136.2862	<.0001
mum_edu2	1	0.4016	0.0615	42.6932	<.0001
mum_edu3	1	0.5813	0.2123	7.4980	0.0062
mum_edu4	1	0.8742	0.1461	35.7873	<.0001
mum_edu5	1	0.9068	0.0857	112.0017	<.0001
mum_edu6	1	1.7194	0.3472	24.5209	<.0001
dad_edu2	1	0.2469	0.0607	16.5167	<.0001
dad_edu3	1	0.4458	0.2088	4.5584	0.0328
dad_edu4	1	0.6383	0.1281	24.8133	<.0001
dad_edu5	1	0.6631	0.1061	39.0762	<.0001
dad_edu6	1	0.8572	0.1860	21.2325	<.0001
mum_ingrp2	1	0.0741	0.0742	0.9970	0.3180
mum_ingrp3	1	0.1058	0.0763	1.9263	0.1652
mum_ingrp4	1	0.2319	0.0816	8.0769	0.0045
mum_ingrp5	1	0.5509	0.3389	2.6424	0.1040
dad_ingrp2	1	0.0750	0.0742	1.0230	0.3118
dad_ingrp3	1	0.2499	0.0753	11.0138	0.0009
dad_ingrp4	1	0.4061	0.0796	26.0398	<.0001
dad_ingrp5	1	0.3299	0.2799	1.3888	0.2386

Figure F.6: Left Table: MLE Boys, Right MLE Table: Girls, Immigrant origin data form 2006 Kolding and Vejle



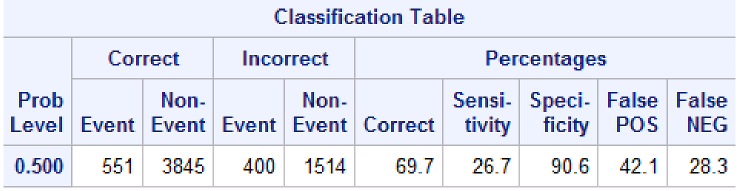
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.4138	0.3526	16.0773	<.0001
mum_edu2	1	0.4692	0.2740	2.9319	0.0868
mum_edu3	1	0.8934	0.4084	4.7854	0.0287
mum_edu4	1	1.3742	0.6018	5.2134	0.0224
mum_edu5	1	1.0512	0.4049	6.7396	0.0094
mum_edu6	1	0.5580	0.5586	0.9977	0.3179
dad_edu2	1	-0.6272	0.3266	3.6871	0.0548
dad_edu3	1	-0.2709	0.5071	0.2855	0.5931
dad_edu4	1	-0.2904	0.4654	0.3894	0.5326
dad_edu5	1	0.1108	0.3956	0.0784	0.7794
dad_edu6	1	0.6366	0.5301	1.4421	0.2298
mum_ingrp2	1	0.1223	0.3415	0.1284	0.7201
mum_ingrp3	1	-0.0879	0.3494	0.0633	0.8013
mum_ingrp4	1	0.6197	0.3484	3.1627	0.0753
mum_ingrp5	1	1.6112	0.8015	4.0411	0.0444
dad_ingrp2	1	0.2229	0.3518	0.4017	0.5262
dad_ingrp3	1	0.4307	0.3686	1.3659	0.2425
dad_ingrp4	1	0.1738	0.3453	0.2532	0.6148
dad_ingrp5	1	0.9518	0.9309	1.0455	0.3065

Analysis of Maximum Likelihood Estimates

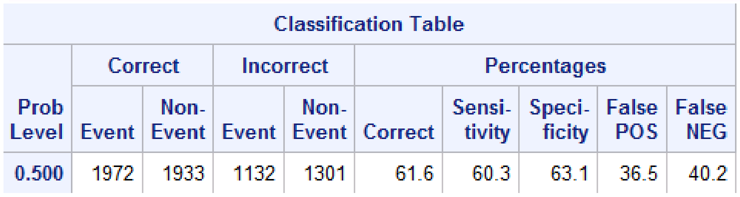
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.5824	0.3385	21.8582	<.0001
mum_edu2	1	0.7940	0.2890	7.5475	0.0060
mum_edu3	1	0.4865	0.4392	1.2272	0.2680
mum_edu4	1	1.1156	0.5577	4.0018	0.0455
mum_edu5	1	0.5972	0.4213	2.0087	0.1564
mum_edu6	1	0.6987	0.5650	1.5293	0.2162
dad_edu2	1	-0.4959	0.3281	2.2847	0.1307
dad_edu3	1	-0.2161	0.5149	0.1761	0.6748
dad_edu4	1	-0.1971	0.4816	0.1675	0.6823
dad_edu5	1	-0.2720	0.5016	0.2941	0.5876
dad_edu6	1	-0.3751	0.6135	0.3737	0.5410
mum_ingrp2	1	0.3830	0.3491	1.2034	0.2727
mum_ingrp3	1	0.2924	0.3488	0.7028	0.4018
mum_ingrp4	1	0.3151	0.3515	0.8037	0.3700
mum_ingrp5	1	16.3120	916.7	0.0003	0.9858
dad_ingrp2	1	-0.8161	0.3722	4.8064	0.0284
dad_ingrp3	1	-0.4763	0.3619	1.7318	0.1882
dad_ingrp4	1	0.0192	0.3497	0.0030	0.9561
dad_ingrp5	1	0.9043	1.0579	0.7307	0.3927

Figure F.7: Top Table: Boys, Bottom Table: Girls, Classification tables native origin data form 2006 Kolding and Vejle



Classification Table

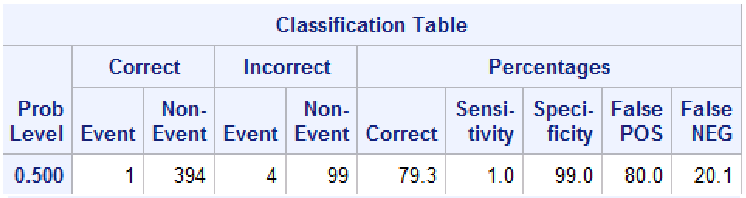
	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	551	3845	400	1514	69.7	26.7	90.6	42.1	28.3



Classification Table

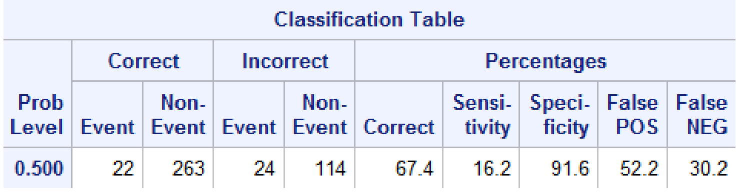
	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	1972	1933	1132	1301	61.6	60.3	63.1	36.5	40.2

Figure F.8: Top Table: Boys, Bottom Table: Girls, Classification tables immigrant origin data form 2006 Kolding and Vejle



Classification Table

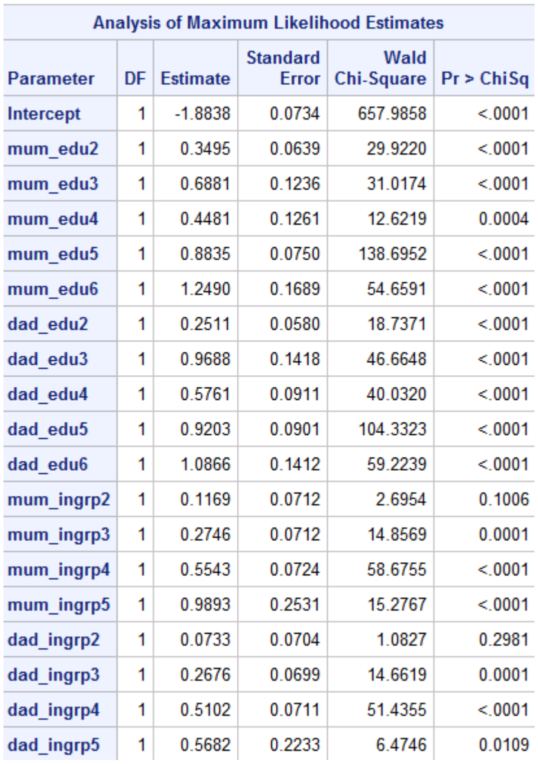
	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	1	394	4	99	79.3	1.0	99.0	80.0	20.1



Classification Table

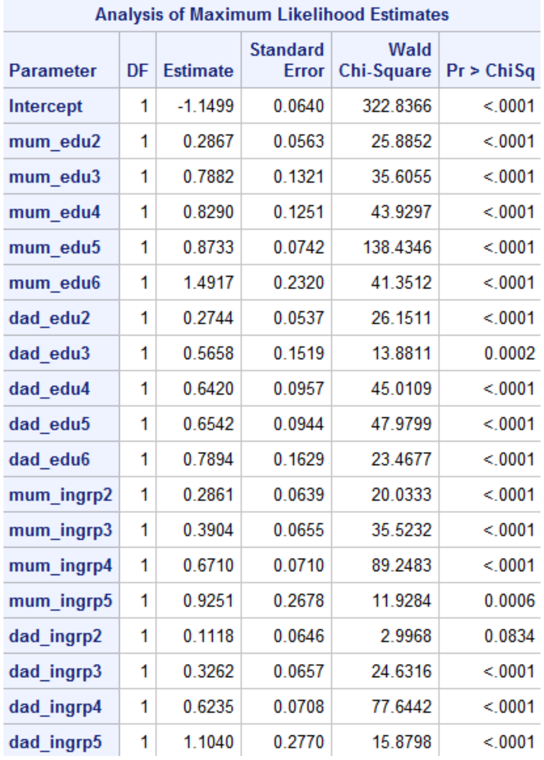
	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	22	263	24	114	67.4	16.2	91.6	52.2	30.2

Figure F.9: Left Table: MLE Boys, Right MLE Table: Girls, native origin data form 2016 Kolding and Vejle



Analysis of Maximum Likelihood Estimates

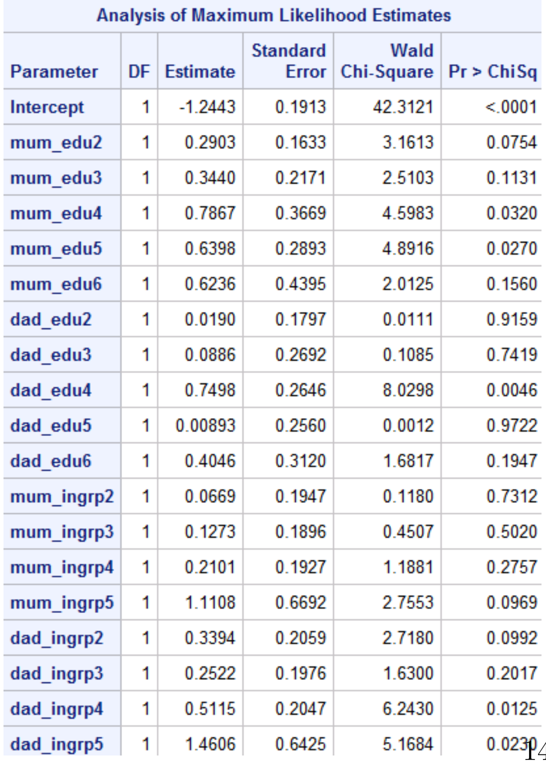
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.8838	0.0734	657.9858	<.0001
mum_edu2	1	0.3495	0.0639	29.9220	<.0001
mum_edu3	1	0.6881	0.1236	31.0174	<.0001
mum_edu4	1	0.4481	0.1261	12.6219	0.0004
mum_edu5	1	0.8835	0.0750	138.6952	<.0001
mum_edu6	1	1.2490	0.1689	54.6591	<.0001
dad_edu2	1	0.2511	0.0580	18.7371	<.0001
dad_edu3	1	0.9688	0.1418	46.6648	<.0001
dad_edu4	1	0.5761	0.0911	40.0320	<.0001
dad_edu5	1	0.9203	0.0901	104.3323	<.0001
dad_edu6	1	1.0866	0.1412	59.2239	<.0001
mum_ingrp2	1	0.1169	0.0712	2.6954	0.1006
mum_ingrp3	1	0.2746	0.0712	14.8569	0.0001
mum_ingrp4	1	0.5543	0.0724	58.6755	<.0001
mum_ingrp5	1	0.9893	0.2531	15.2767	<.0001
dad_ingrp2	1	0.0733	0.0704	1.0827	0.2981
dad_ingrp3	1	0.2676	0.0699	14.6619	0.0001
dad_ingrp4	1	0.5102	0.0711	51.4355	<.0001
dad_ingrp5	1	0.5682	0.2233	6.4746	0.0109



Analysis of Maximum Likelihood Estimates

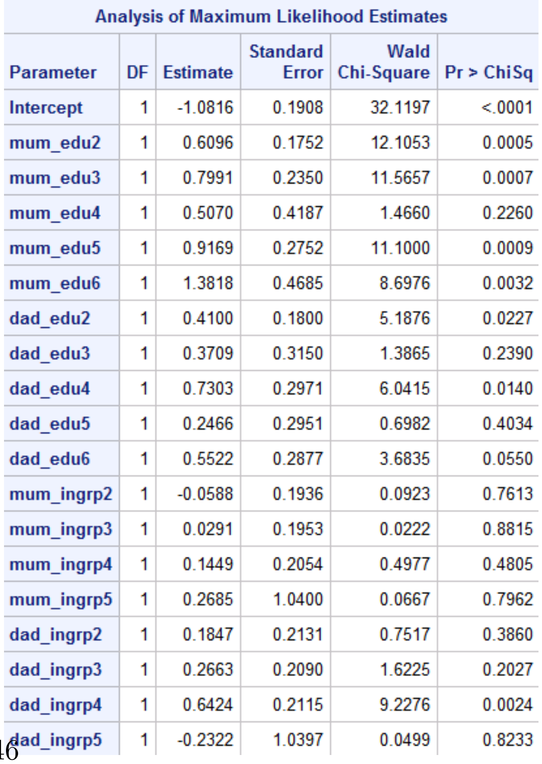
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.1499	0.0640	322.8366	<.0001
mum_edu2	1	0.2867	0.0563	25.8852	<.0001
mum_edu3	1	0.7882	0.1321	35.6055	<.0001
mum_edu4	1	0.8290	0.1251	43.9297	<.0001
mum_edu5	1	0.8733	0.0742	138.4346	<.0001
mum_edu6	1	1.4917	0.2320	41.3512	<.0001
dad_edu2	1	0.2744	0.0537	26.1511	<.0001
dad_edu3	1	0.5658	0.1519	13.8811	0.0002
dad_edu4	1	0.6420	0.0957	45.0109	<.0001
dad_edu5	1	0.6542	0.0944	47.9799	<.0001
dad_edu6	1	0.7894	0.1629	23.4677	<.0001
mum_ingrp2	1	0.2861	0.0639	20.0333	<.0001
mum_ingrp3	1	0.3904	0.0655	35.5232	<.0001
mum_ingrp4	1	0.6710	0.0710	89.2483	<.0001
mum_ingrp5	1	0.9251	0.2678	11.9284	0.0006
dad_ingrp2	1	0.1118	0.0646	2.9968	0.0834
dad_ingrp3	1	0.3262	0.0657	24.6316	<.0001
dad_ingrp4	1	0.6235	0.0708	77.6442	<.0001
dad_ingrp5	1	1.1040	0.2770	15.8798	<.0001

Figure F.10: Left Table: MLE Boys, Right MLE Table: Girls, immigrant origin data form 2016 Kolding and Vejle



Analysis of Maximum Likelihood Estimates

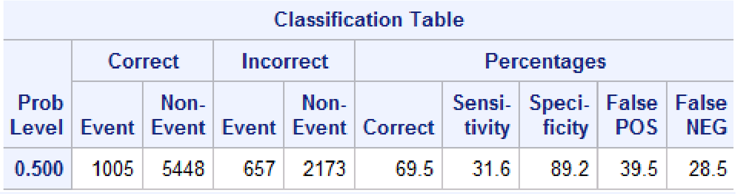
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.2443	0.1913	42.3121	<.0001
mum_edu2	1	0.2903	0.1633	3.1613	0.0754
mum_edu3	1	0.3440	0.2171	2.5103	0.1131
mum_edu4	1	0.7867	0.3669	4.5983	0.0320
mum_edu5	1	0.6398	0.2893	4.8916	0.0270
mum_edu6	1	0.6236	0.4395	2.0125	0.1560
dad_edu2	1	0.0190	0.1797	0.0111	0.9159
dad_edu3	1	0.0886	0.2692	0.1085	0.7419
dad_edu4	1	0.7498	0.2646	8.0298	0.0046
dad_edu5	1	0.00893	0.2560	0.0012	0.9722
dad_edu6	1	0.4046	0.3120	1.6817	0.1947
mum_ingrp2	1	0.0669	0.1947	0.1180	0.7312
mum_ingrp3	1	0.1273	0.1896	0.4507	0.5020
mum_ingrp4	1	0.2101	0.1927	1.1881	0.2757
mum_ingrp5	1	1.1108	0.6692	2.7553	0.0969
dad_ingrp2	1	0.3394	0.2059	2.7180	0.0992
dad_ingrp3	1	0.2522	0.1976	1.6300	0.2017
dad_ingrp4	1	0.5115	0.2047	6.2430	0.0125
dad_ingrp5	1	1.4606	0.6425	5.1684	0.0230



Analysis of Maximum Likelihood Estimates

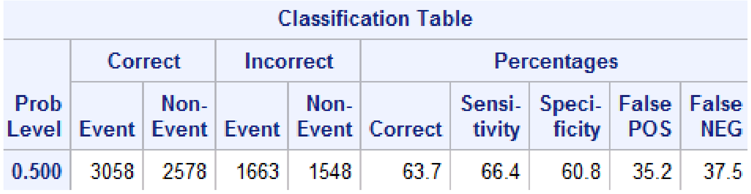
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.0816	0.1908	32.1197	<.0001
mum_edu2	1	0.6096	0.1752	12.1053	0.0005
mum_edu3	1	0.7991	0.2350	11.5657	0.0007
mum_edu4	1	0.5070	0.4187	1.4660	0.2260
mum_edu5	1	0.9169	0.2752	11.1000	0.0009
mum_edu6	1	1.3818	0.4685	8.6976	0.0032
dad_edu2	1	0.4100	0.1800	5.1876	0.0227
dad_edu3	1	0.3709	0.3150	1.3865	0.2390
dad_edu4	1	0.7303	0.2971	6.0415	0.0140
dad_edu5	1	0.2466	0.2951	0.6982	0.4034
dad_edu6	1	0.5522	0.2877	3.6835	0.0550
mum_ingrp2	1	-0.0588	0.1936	0.0923	0.7613
mum_ingrp3	1	0.0291	0.1953	0.0222	0.8815
mum_ingrp4	1	0.1449	0.2054	0.4977	0.4805
mum_ingrp5	1	0.2685	1.0400	0.0667	0.7962
dad_ingrp2	1	0.1847	0.2131	0.7517	0.3860
dad_ingrp3	1	0.2663	0.2090	1.6225	0.2027
dad_ingrp4	1	0.6424	0.2115	9.2276	0.0024
dad_ingrp5	1	-0.2322	1.0397	0.0499	0.8233

Figure F.11: Top Table: Boys, Bottom Table: Girls, Classification tables native origin data form 2016 Kolding and Vejle



Classification Table

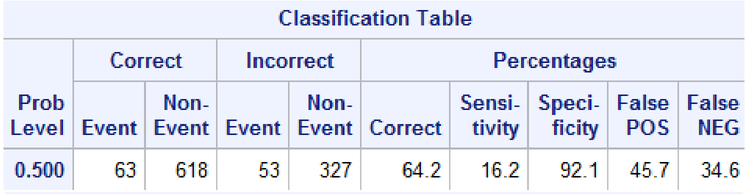
	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	1005	5448	657	2173	69.5	31.6	89.2	39.5	28.5



Classification Table

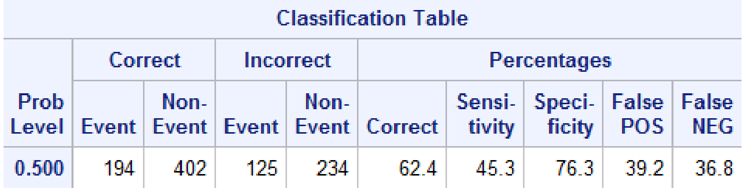
	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	3058	2578	1663	1548	63.7	66.4	60.8	35.2	37.5

Figure F.12: Top Table: Boys, Bottom Table: Girls, Classification tables immigrant origin data form 2006 Kolding and Vejle



Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	63	618	53	327	64.2	16.2	92.1	45.7	34.6



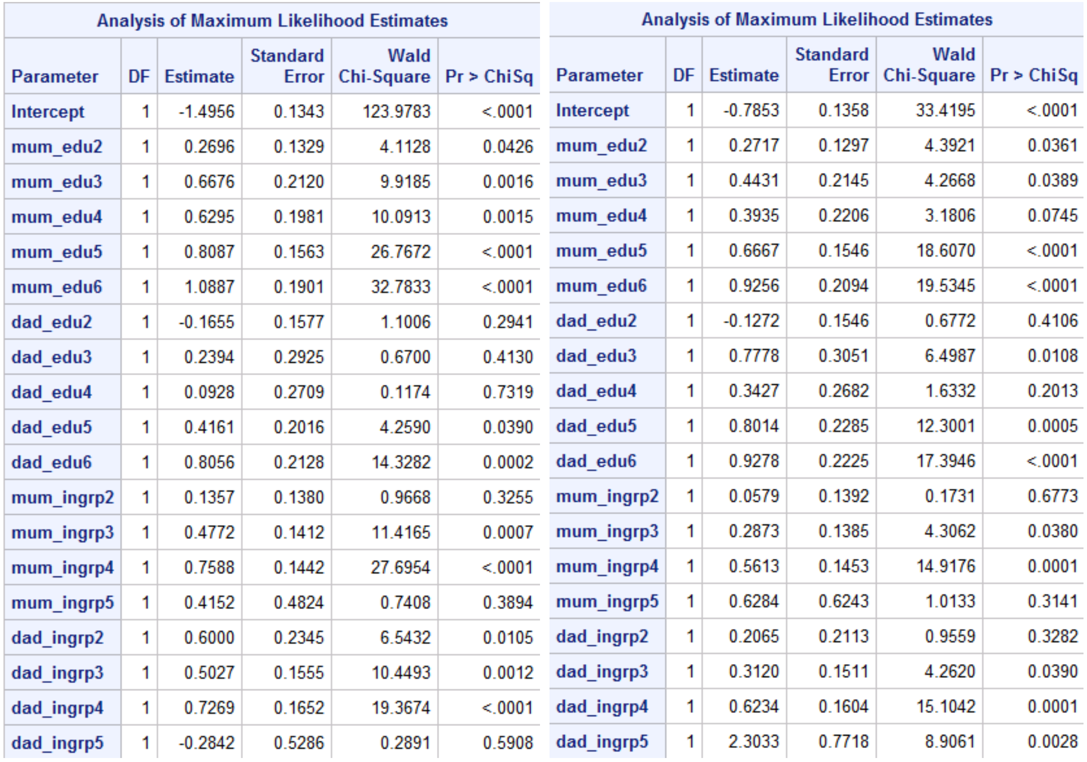
Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	194	402	125	234	62.4	45.3	76.3	39.2	36.8

Appendix G

Western/non-Western origin

Figure G.1: Left Table: MLE Boys, Righ Table: MLE Girls, Western origin data form 2006



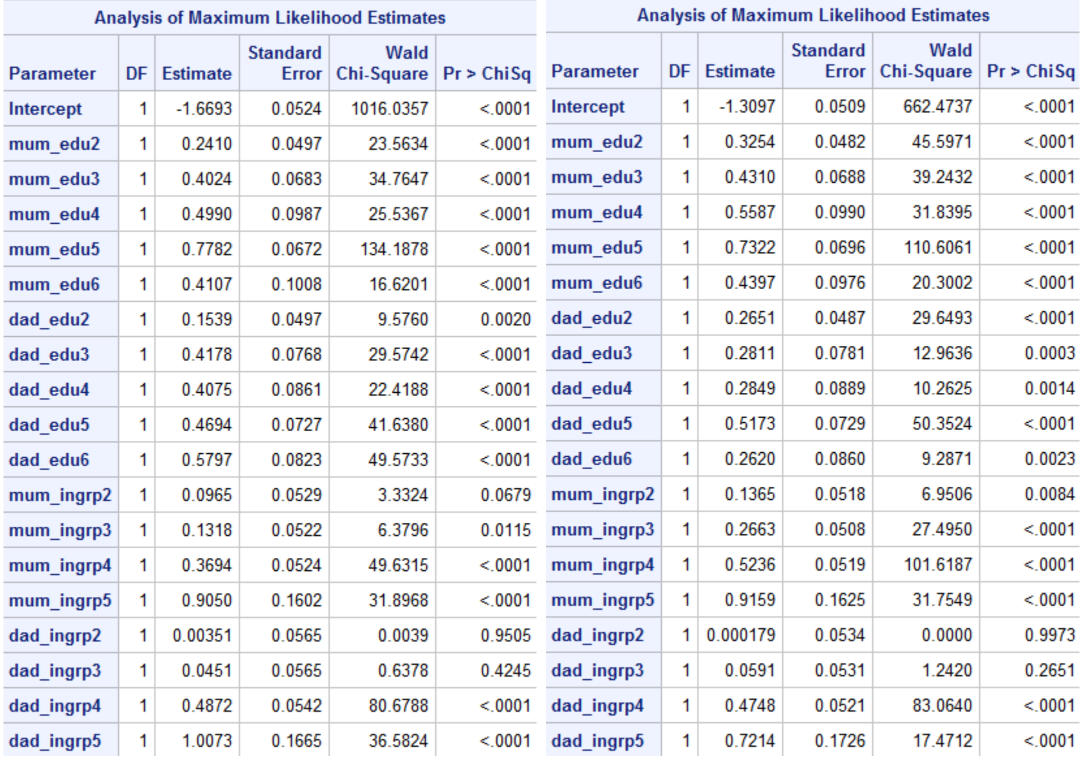
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.4956	0.1343	123.9783	<.0001
mum_edu2	1	0.2696	0.1329	4.1128	0.0426
mum_edu3	1	0.6676	0.2120	9.9185	0.0016
mum_edu4	1	0.6295	0.1981	10.0913	0.0015
mum_edu5	1	0.8087	0.1563	26.7672	<.0001
mum_edu6	1	1.0887	0.1901	32.7833	<.0001
dad_edu2	1	-0.1655	0.1577	1.1006	0.2941
dad_edu3	1	0.2394	0.2925	0.6700	0.4130
dad_edu4	1	0.0928	0.2709	0.1174	0.7319
dad_edu5	1	0.4161	0.2016	4.2590	0.0390
dad_edu6	1	0.8056	0.2128	14.3282	0.0002
mum_ingrp2	1	0.1357	0.1380	0.9668	0.3255
mum_ingrp3	1	0.4772	0.1412	11.4165	0.0007
mum_ingrp4	1	0.7588	0.1442	27.6954	<.0001
mum_ingrp5	1	0.4152	0.4824	0.7408	0.3894
dad_ingrp2	1	0.6000	0.2345	6.5432	0.0105
dad_ingrp3	1	0.5027	0.1555	10.4493	0.0012
dad_ingrp4	1	0.7269	0.1652	19.3674	<.0001
dad_ingrp5	1	-0.2842	0.5286	0.2891	0.5908

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.7853	0.1358	33.4195	<.0001
mum_edu2	1	0.2717	0.1297	4.3921	0.0361
mum_edu3	1	0.4431	0.2145	4.2668	0.0389
mum_edu4	1	0.3935	0.2206	3.1806	0.0745
mum_edu5	1	0.6667	0.1546	18.6070	<.0001
mum_edu6	1	0.9256	0.2094	19.5345	<.0001
dad_edu2	1	-0.1272	0.1546	0.6772	0.4106
dad_edu3	1	0.7778	0.3051	6.4987	0.0108
dad_edu4	1	0.3427	0.2682	1.6332	0.2013
dad_edu5	1	0.8014	0.2285	12.3001	0.0005
dad_edu6	1	0.9278	0.2225	17.3946	<.0001
mum_ingrp2	1	0.0579	0.1392	0.1731	0.6773
mum_ingrp3	1	0.2873	0.1385	4.3062	0.0380
mum_ingrp4	1	0.5613	0.1453	14.9176	0.0001
mum_ingrp5	1	0.6284	0.6243	1.0133	0.3141
dad_ingrp2	1	0.2065	0.2113	0.9559	0.3282
dad_ingrp3	1	0.3120	0.1511	4.2620	0.0390
dad_ingrp4	1	0.6234	0.1604	15.1042	0.0001
dad_ingrp5	1	2.3033	0.7718	8.9061	0.0028

Figure G.2: Left Table: MLE Boys, Righ Table: MLE Girls, non-Western origin data form 2006



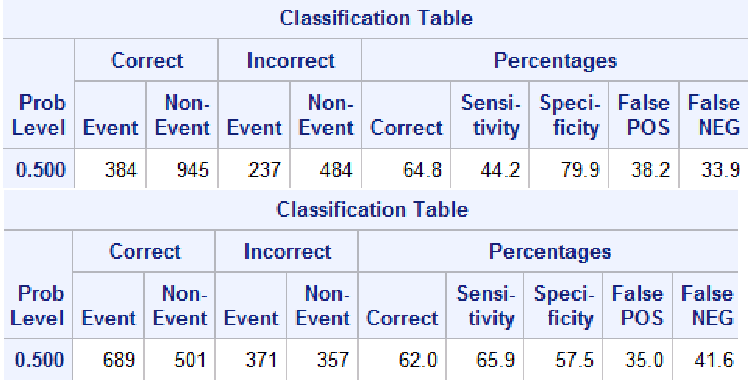
Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.6693	0.0524	1016.0357	<.0001
mum_edu2	1	0.2410	0.0497	23.5634	<.0001
mum_edu3	1	0.4024	0.0683	34.7647	<.0001
mum_edu4	1	0.4990	0.0987	25.5367	<.0001
mum_edu5	1	0.7782	0.0672	134.1878	<.0001
mum_edu6	1	0.4107	0.1008	16.6201	<.0001
dad_edu2	1	0.1539	0.0497	9.5760	0.0020
dad_edu3	1	0.4178	0.0768	29.5742	<.0001
dad_edu4	1	0.4075	0.0861	22.4188	<.0001
dad_edu5	1	0.4694	0.0727	41.6380	<.0001
dad_edu6	1	0.5797	0.0823	49.5733	<.0001
mum_ingrp2	1	0.0965	0.0529	3.3324	0.0679
mum_ingrp3	1	0.1318	0.0522	6.3796	0.0115
mum_ingrp4	1	0.3694	0.0524	49.6315	<.0001
mum_ingrp5	1	0.9050	0.1602	31.8968	<.0001
dad_ingrp2	1	0.00351	0.0565	0.0039	0.9505
dad_ingrp3	1	0.0451	0.0565	0.6378	0.4245
dad_ingrp4	1	0.4872	0.0542	80.6788	<.0001
dad_ingrp5	1	1.0073	0.1665	36.5824	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.3097	0.0509	662.4737	<.0001
mum_edu2	1	0.3254	0.0482	45.5971	<.0001
mum_edu3	1	0.4310	0.0688	39.2432	<.0001
mum_edu4	1	0.5587	0.0990	31.8395	<.0001
mum_edu5	1	0.7322	0.0696	110.6061	<.0001
mum_edu6	1	0.4397	0.0976	20.3002	<.0001
dad_edu2	1	0.2651	0.0487	29.6493	<.0001
dad_edu3	1	0.2811	0.0781	12.9636	0.0003
dad_edu4	1	0.2849	0.0889	10.2625	0.0014
dad_edu5	1	0.5173	0.0729	50.3524	<.0001
dad_edu6	1	0.2620	0.0860	9.2871	0.0023
mum_ingrp2	1	0.1365	0.0518	6.9506	0.0084
mum_ingrp3	1	0.2663	0.0508	27.4950	<.0001
mum_ingrp4	1	0.5236	0.0519	101.6187	<.0001
mum_ingrp5	1	0.9159	0.1625	31.7549	<.0001
dad_ingrp2	1	0.000179	0.0534	0.0000	0.9973
dad_ingrp3	1	0.0591	0.0531	1.2420	0.2651
dad_ingrp4	1	0.4748	0.0521	83.0640	<.0001
dad_ingrp5	1	0.7214	0.1726	17.4712	<.0001

Figure G.3: Top Table: Boys, Bottom Table: Girls, Classification tables Western origin data form 2006



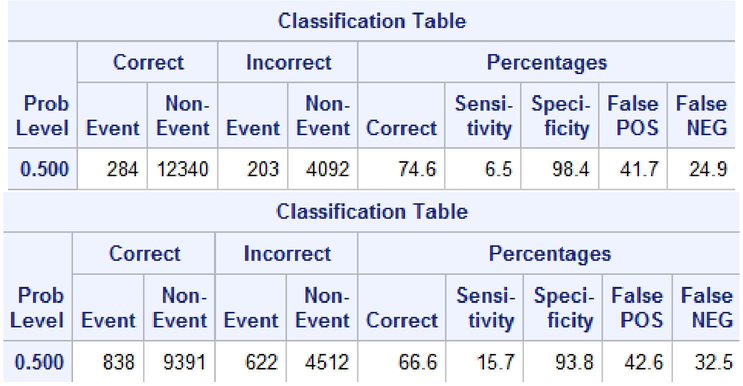
Classification Table

Prob Level	Correct		Incorrect		Percentages				
	Event	Non-Event	Event	Non-Event	Correct	Sensitivity	Specificity	False POS	False NEG
0.500	384	945	237	484	64.8	44.2	79.9	38.2	33.9

Classification Table

Prob Level	Correct		Incorrect		Percentages				
	Event	Non-Event	Event	Non-Event	Correct	Sensitivity	Specificity	False POS	False NEG
0.500	689	501	371	357	62.0	65.9	57.5	35.0	41.6

Figure G.4: Top Table: Boys, Bottom Table: Girls, Classification tables non-Western origin data form 2006



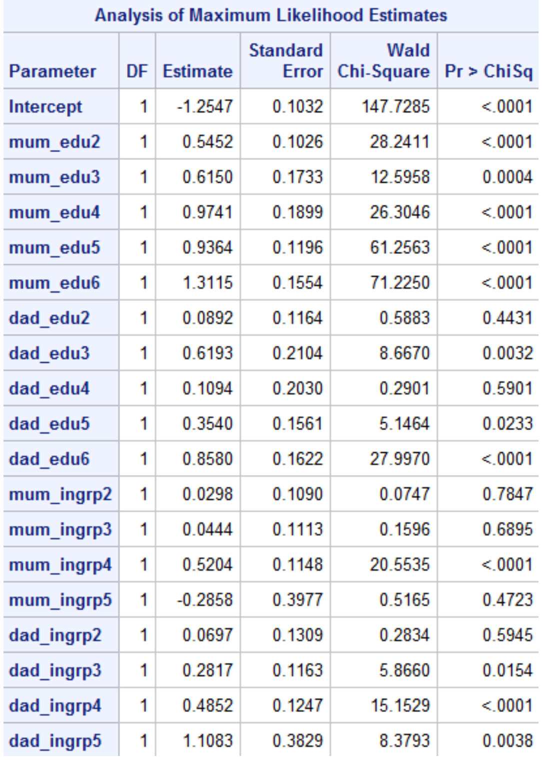
Classification Table

Prob Level	Correct		Incorrect		Percentages				
	Event	Non-Event	Event	Non-Event	Correct	Sensitivity	Specificity	False POS	False NEG
0.500	284	12340	203	4092	74.6	6.5	98.4	41.7	24.9

Classification Table

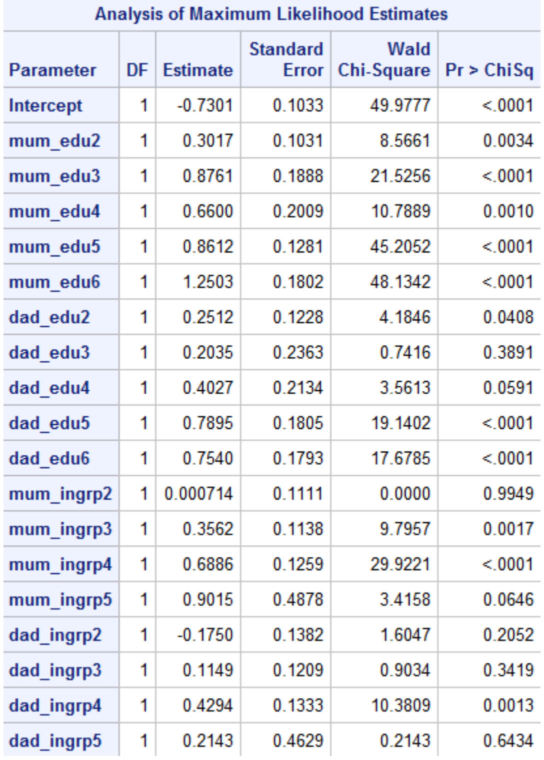
Prob Level	Correct		Incorrect		Percentages				
	Event	Non-Event	Event	Non-Event	Correct	Sensitivity	Specificity	False POS	False NEG
0.500	838	9391	622	4512	66.6	15.7	93.8	42.6	32.5

Figure G.5: Left Table: MLE Boys, Righ Table: MLE Girls, western origin data form 2016



Analysis of Maximum Likelihood Estimates

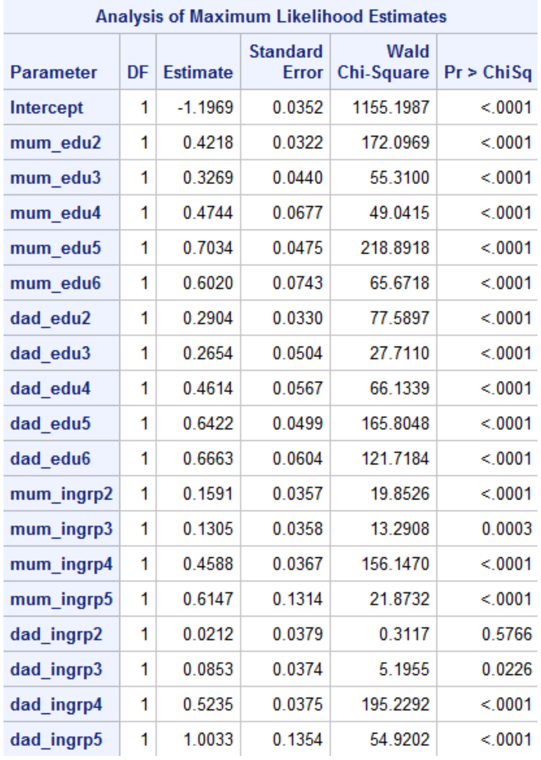
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.2547	0.1032	147.7285	<.0001
mum_edu2	1	0.5452	0.1026	28.2411	<.0001
mum_edu3	1	0.6150	0.1733	12.5958	0.0004
mum_edu4	1	0.9741	0.1899	26.3046	<.0001
mum_edu5	1	0.9364	0.1196	61.2563	<.0001
mum_edu6	1	1.3115	0.1554	71.2250	<.0001
dad_edu2	1	0.0892	0.1164	0.5883	0.4431
dad_edu3	1	0.6193	0.2104	8.6670	0.0032
dad_edu4	1	0.1094	0.2030	0.2901	0.5901
dad_edu5	1	0.3540	0.1561	5.1464	0.0233
dad_edu6	1	0.8580	0.1622	27.9970	<.0001
mum_ingrp2	1	0.0298	0.1090	0.0747	0.7847
mum_ingrp3	1	0.0444	0.1113	0.1596	0.6895
mum_ingrp4	1	0.5204	0.1148	20.5535	<.0001
mum_ingrp5	1	-0.2858	0.3977	0.5165	0.4723
dad_ingrp2	1	0.0697	0.1309	0.2834	0.5945
dad_ingrp3	1	0.2817	0.1163	5.8660	0.0154
dad_ingrp4	1	0.4852	0.1247	15.1529	<.0001
dad_ingrp5	1	1.1083	0.3829	8.3793	0.0038



Analysis of Maximum Likelihood Estimates

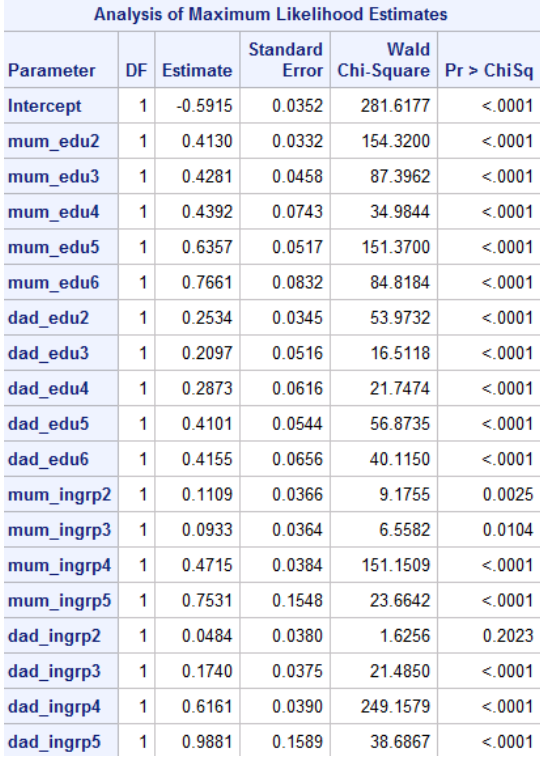
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.7301	0.1033	49.9777	<.0001
mum_edu2	1	0.3017	0.1031	8.5661	0.0034
mum_edu3	1	0.8761	0.1888	21.5256	<.0001
mum_edu4	1	0.6600	0.2009	10.7889	0.0010
mum_edu5	1	0.8612	0.1281	45.2052	<.0001
mum_edu6	1	1.2503	0.1802	48.1342	<.0001
dad_edu2	1	0.2512	0.1228	4.1846	0.0408
dad_edu3	1	0.2035	0.2363	0.7416	0.3891
dad_edu4	1	0.4027	0.2134	3.5613	0.0591
dad_edu5	1	0.7895	0.1805	19.1402	<.0001
dad_edu6	1	0.7540	0.1793	17.6785	<.0001
mum_ingrp2	1	0.000714	0.1111	0.0000	0.9949
mum_ingrp3	1	0.3562	0.1138	9.7957	0.0017
mum_ingrp4	1	0.6886	0.1259	29.9221	<.0001
mum_ingrp5	1	0.9015	0.4878	3.4158	0.0646
dad_ingrp2	1	-0.1750	0.1382	1.6047	0.2052
dad_ingrp3	1	0.1149	0.1209	0.9034	0.3419
dad_ingrp4	1	0.4294	0.1333	10.3809	0.0013
dad_ingrp5	1	0.2143	0.4629	0.2143	0.6434

Figure G.6: Left Table: MLE Boys, Righ Table: MLE Girls, non-Western origin data form 2016



Analysis of Maximum Likelihood Estimates

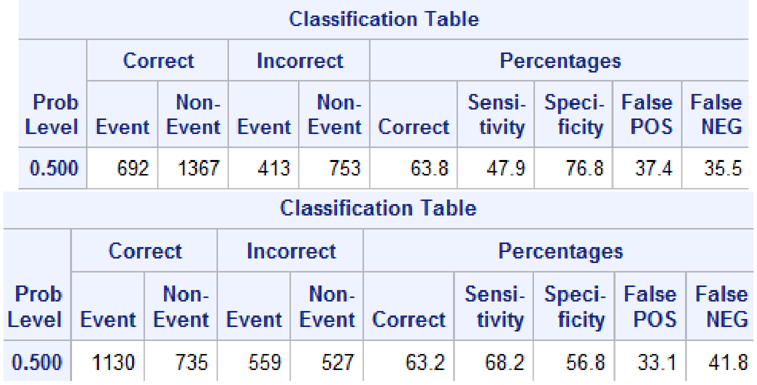
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.1969	0.0352	1155.1987	<.0001
mum_edu2	1	0.4218	0.0322	172.0969	<.0001
mum_edu3	1	0.3269	0.0440	55.3100	<.0001
mum_edu4	1	0.4744	0.0677	49.0415	<.0001
mum_edu5	1	0.7034	0.0475	218.8918	<.0001
mum_edu6	1	0.6020	0.0743	65.6718	<.0001
dad_edu2	1	0.2904	0.0330	77.5897	<.0001
dad_edu3	1	0.2654	0.0504	27.7110	<.0001
dad_edu4	1	0.4614	0.0567	66.1339	<.0001
dad_edu5	1	0.6422	0.0499	165.8048	<.0001
dad_edu6	1	0.6663	0.0604	121.7184	<.0001
mum_ingrp2	1	0.1591	0.0357	19.8526	<.0001
mum_ingrp3	1	0.1305	0.0358	13.2908	0.0003
mum_ingrp4	1	0.4588	0.0367	156.1470	<.0001
mum_ingrp5	1	0.6147	0.1314	21.8732	<.0001
dad_ingrp2	1	0.0212	0.0379	0.3117	0.5766
dad_ingrp3	1	0.0853	0.0374	5.1955	0.0226
dad_ingrp4	1	0.5235	0.0375	195.2292	<.0001
dad_ingrp5	1	1.0033	0.1354	54.9202	<.0001



Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.5915	0.0352	281.6177	<.0001
mum_edu2	1	0.4130	0.0332	154.3200	<.0001
mum_edu3	1	0.4281	0.0458	87.3962	<.0001
mum_edu4	1	0.4392	0.0743	34.9844	<.0001
mum_edu5	1	0.6357	0.0517	151.3700	<.0001
mum_edu6	1	0.7661	0.0832	84.8184	<.0001
dad_edu2	1	0.2534	0.0345	53.9732	<.0001
dad_edu3	1	0.2097	0.0516	16.5118	<.0001
dad_edu4	1	0.2873	0.0616	21.7474	<.0001
dad_edu5	1	0.4101	0.0544	56.8735	<.0001
dad_edu6	1	0.4155	0.0656	40.1150	<.0001
mum_ingrp2	1	0.1109	0.0366	9.1755	0.0025
mum_ingrp3	1	0.0933	0.0364	6.5582	0.0104
mum_ingrp4	1	0.4715	0.0384	151.1509	<.0001
mum_ingrp5	1	0.7531	0.1548	23.6642	<.0001
dad_ingrp2	1	0.0484	0.0380	1.6256	0.2023
dad_ingrp3	1	0.1740	0.0375	21.4850	<.0001
dad_ingrp4	1	0.6161	0.0390	249.1579	<.0001
dad_ingrp5	1	0.9881	0.1589	38.6867	<.0001

Figure G.7: Top Table: Boys, Bottom Table: Girls, Classification tables Western origin data form 2016



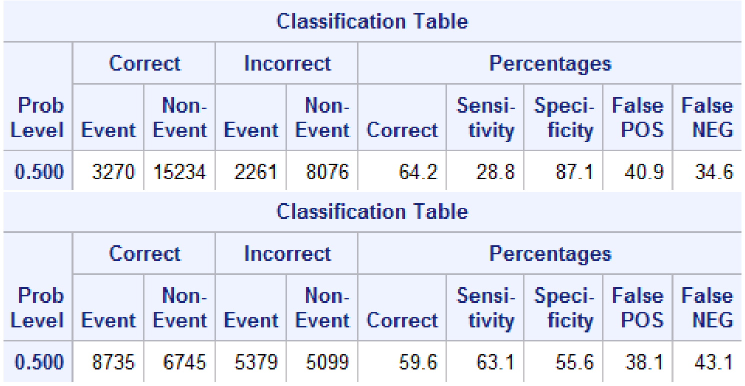
Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	692	1367	413	753	63.8	47.9	76.8	37.4	35.5

Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	1130	735	559	527	63.2	68.2	56.8	33.1	41.8

Figure G.8: Top Table: Boys, Bottom Table: Girls, Classification tables non-Western origin data form 2016



Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	3270	15234	2261	8076	64.2	28.8	87.1	40.9	34.6

Classification Table

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	8735	6745	5379	5099	59.6	63.1	55.6	38.1	43.1