

IZA DP No. 7410

## **Gender Differences in the Effects of Behavioral Problems on School Outcomes**

Jannie H. G. Kristoffersen  
Nina Smith

May 2013

# Gender Differences in the Effects of Behavioral Problems on School Outcomes

**Jannie H. G. Kristoffersen**

*Aarhus University*

**Nina Smith**

*Aarhus University  
and IZA*

Discussion Paper No. 7410  
May 2013

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### Gender Differences in the Effects of Behavioral Problems on School Outcomes<sup>\*</sup>

Behavioral problems are important determinants of school outcomes and later success in the labor market. We analyze whether behavioral problems affect girls and boys differently with respect to school outcomes. The study is based on teacher and parent evaluations of the Strength and Difficulties Questionnaire (SDQ) of about 6,000 children born in 1990-92 in a large region in Denmark. The sample is merged with register information on parents and students observed until the age of 19. We find significant and large negative coefficients of the externalizing behavioral indicators. The effects tend to be larger when based on parents' SDQ scores compared to teachers' SDQ scores. According to our estimations, the school outcomes for girls with abnormal externalizing behavior are not significantly different from those of boys with the same behavioral problems. A decomposition of the estimates indicates that most of the gender differences in Reading and Math cannot be related to gender differences in behavioral problems. The large overall gender gap in Reading seems mainly to be the result of gender differences between children without behavioral problems living in 'normal families', i.e. families which are not categorized as low-resource families.

JEL Classification: J16, I29, I19

Keywords: gender differences, education, behavior

Corresponding author:

Jannie H. G. Kristoffersen  
Department of Economics and Business  
Aarhus University  
Building 2621  
Fuglesangs Allé 4  
DK-8210 Aarhus V  
Denmark  
E-mail: [jgroene@econ.au.dk](mailto:jgroene@econ.au.dk)

---

<sup>\*</sup> The authors would like to thank Carsten Obel for making the data on the Strengths and Difficulties Questionnaire available for the Aarhus Birth Cohort. We also thank Helena Skyt Nielsen, Marianne Simonsen, Martin Salm, Anna Sjögren, participants at ESPE 2012, IWAE 2012, and seminar participants at Aarhus University for constructive comments.

# 1 Introduction

During the latest decades there has been a growing focus on why girls seem to perform better in primary schools than boys in most dimensions, only with the exception of math grades where boys are as good as or even outperform girls. Since student achievement in primary school is an important determinant of enrollment into high school and a determinant of success in the educational system, this may be one of the explanations of the emerging gender gap in the educational level which is found in many countries, see for instance Goldin et al. (2006) and Fortin et al. (2011). This is also the case for Denmark where girls are much more successful in the educational system with respect to completing high school and obtaining a qualifying education, i.e. a degree from college or university or a vocational education.

The number of Danish children placed in special needs classes is steeply increasing. A significant proportion of the children in special needs classes have behavioral problems and the majority of them are boys, see Mehlbye (2008). From an economic policy perspective, this development is important and costly. The school outcomes and long-term perspectives for employment of children in special needs classes are less positive than for other children, and 20-33 percent of all expenditures allocated to compulsory schools in Denmark are now spent on student with special needs, see Ministry of Finance (2010).

According to PISA 2009, see e.g. OECD (2010), Danish students have a problematically high score regarding some aspects of behavioral problems. For instance a relatively large proportion of the teachers in the Danish PISA 2009 study claim that in most lessons 'there is noise and disorder' in the classroom.<sup>1</sup> These classroom problems may clearly have negative effects on the learning outcomes of students. The goal of this paper is to investigate the gender differences in the relation between behavioral problems and school outcomes, and how these effects depend on the family background.

In this study, we analyze gender differences in the relation between behavioral problems and student outcomes based on a large sample of Danish children born in 1990-1992. Student outcomes are measured as whether or not they took the 9th grade exit exam (i.e. took at least one of the course exams) and as the performance in terms of grades in Math and Reading, and enrollment into high school (including vocational programs). We focus on whether the effects from behavioral skills follow the same pattern for girls and boys, and which factors such as family background influence this pattern. Our general hypothesis is that the relation between

---

<sup>1</sup>See PISA 2009 Database (ST36Q02).

school outcome and behavioral skills is more sensitive to family and school environment for boys than for girls. For instance we want to test whether boys with disruptive behavior are more negatively affected than girls coming from, what we define as, a low-resource family. Further, we analyze whether teachers' and parents' views on behavioral problems differ systematically between boys and girls and whether this potential difference in perception matters for the gender-specific link between behavior and school outcome.

Our results show that about 11 percent of the boys and 6 percent of the girls in our sample have abnormal or borderline externalizing behavioral problems. For internalizing problems, girls and boys have about the same average scores, and 14 percent of the children are categorized as having abnormal or borderline internalizing problems. These measures are based on the Strength and Difficulties Questionnaire (SDQ) scores assessed by the parents. Teachers tend to categorize more children as having abnormal problems.

We find significant and large negative coefficients of the externalizing behavioral indicators. The effects tend to be larger when based on parents' SDQ scores compared to teachers' SDQ scores. According to our estimations, the school outcomes for girls with abnormal externalizing behavior are not significantly different from those of boys with the same behavioral problems. For borderline externalizing problems and for internalizing problems, the estimated main coefficients are numerically smaller but still significantly negative in most cases, and for these behavioral categories, girls tend to have less negative coefficients.

Children from low-resource families have lower school outcomes. The gender gap in Reading is observed to be largest in the sub-group of 'normal families', i.e. families which are not categorized as low-resource families and smallest in the sub-group of families with a young mother. A decomposition of the estimations indicate that most of the gender differences in Reading and Math cannot be related to gender differences in behavioral problems, neither differences in endowments nor gender differences in estimated coefficients. The overall gender gap in Reading seems mainly to be the result of gender differences between children without behavioral problems living in 'normal families', i.e. families which are not categorized as low-resource families. For Math, the higher overall grades for boys compared to girls seem mainly to be the result of very low Math grades for girls without behavioral problems from low-income families.

## 2 Background

An increasing number of studies have focussed on the importance of behavioral problems or behavioral skills of school children. Even after controlling for school grades, behavioral problems are found to matter for educational outcomes in a number of studies, see for instance Heckman and Rubinstein (2001), Jacob (2002), Heckman et al. (2006), Bertrand and Pan (2011), and the survey in Heckman (2008). Behavioral skills may be measured in a number of dimensions, such as externalizing behavior, hyper activity, self-control, approaches to learning, interpersonal skills and internalizing problems such as loneliness and low self-esteem. Behavioral problems may stem from both nature and nurture, but the most important determinant of behavioral skills seems to be gender: Girls tend to have much less behavioral problems at school age than their male peers, see Bertrand and Pan (2011). A large strand of research stresses the biological ('nature') reasons for behavioral problems, i.e. the development of female brains is different from male brains and this may have consequences at early ages for the observed gender gap in behavior. But 'nurture' also seems to matter. Gender differences in for instance child rearing inputs may affect the way behavioral skills are produced, see the survey in Heckman (2008). Since the focus here is on the latter aspect, we restrict our discussion to 'nurture' explanations of the gender gap, recognizing that biological factors may also play an important role.

The study by Bertrand and Pan (2011) is based on the Early Childhood Longitudinal Study (ECLS) which is a US sample of around 20,000 children who were followed from kindergarten until 8th grade. The gender gap in non-cognitive skills exists already in kindergarten but the gender gap evolves and increases steadily during childhood up until 5th grade. However, when looking at separate groups, the growing gender gap shows to be the result of boys living in single mother families, boys from the lowest social economic status families, and boys born by mothers who had their first child before the age of 24. For other boys and for girls from all social groups the disruptive behavior is stable during childhood ages. The same pattern is found in Carneiro and Heckman (2003). Their study strongly stresses the importance of the family resources which are devoted to boys. Single mothers tend to devote less time resources to their sons compared to their daughters, and boys living in single-parent families (or teenage-mother families or the lowest social economic status families) tend to receive less parental resources compared to other boys and girls. The results in Bertrand and Pan indicate that teacher effects, peer effects and school environment are less important: 'Overall, our findings strongly suggest that boys' deficit in non-cognitive skills is not purely biological but instead subject to very

strong environmental influences, particularly from the home' (Bertrand and Pan (2011, p. 7)).

In a study on UK children, Ermisch (2008) also finds that the gender gap in non-cognitive skills exists early in life, already at the age of 3. His study is based on the Millenium Cohort Study and includes data from the 'Strengths and Difficulties Questionnaire' (SDQ) which is also applied in our study on Danish students but for older children. Ermisch (2008) finds large effects of parental background and effects of parents' activities with their children (e.g. reading to the child on a regular basis). When controlling for these variables, Ermisch (2008) finds that an indicator for 'girl' turns insignificant in his estimation of determinants of behavioral problems.

Jacob (2002) poses an empirical link between non-cognitive skills and educational attainment in a US setting. Based on data from National Education Longitudinal Study (NELS), a representative sample of 8th graders who were followed from 1988 to 1994, he finds an overall gender gap in college enrollment of 5 percentage points (7 points if the sample is restricted to bottom 3 quartiles of socioeconomic groups). About 40 percent of this observed gender gap in college attendance is related to differences in non-cognitive skills between girls and boys. The non-cognitive skills considered in Jacob (2002) are middle-school grades and the number of hours spent on homework per week in 8th grade. In a later study by Heckman et al. (2008), it is found that early treatment of children in the Perry program in Michigan in the 1960s only had temporary effects on the IQ of the treated children but still there was a long-run effect on labor market outcomes of the treatment which was mainly due to non-cognitive effects of the program treatment. Thus, behavioral skills seem to be extremely important for later outcomes in life, and especially for educational outcomes.

### **3 Data: Selection, SDQ Scores, and School Outcomes**

The study is based on a panel survey of 10,907 children born at Aarhus University Hospital, Denmark, between January 1990 and March 1992, denoted the Aarhus Birth Cohort (ABC) data. Aarhus University Hospital is one of the largest maternity wards in Denmark, and it covers the entire region. This implies that the children born here are also likely to enter the same school classes as long as their families also live in Aarhus County. The ABC data are survey data and contain extensive information on child health, behavioral variables (SDQ variables) given by parents as well as school teachers. For a sub-sample of children we have information on the gender of the teacher as well. The ABC data are merged to administrative data hosted

by Statistics Denmark and they contain information on parents' socioeconomic status, income, education and marital status during the period. Since the children were born in the early 1990s, we are able to observe their school outcome up until the age of 15-18. School outcomes are also based on administrative data from Statistics Denmark who collects grades in Math and Reading for all Danish pupils exiting from 9th grade in public and private schools in Denmark. Further, the administrative registers contain information on enrollment into vocational programs and high school after compulsory school.

### 3.1 Description of the Sample

In the late prenatal stages, the mothers who were expected to give birth at Aarhus University Hospital during the period January 1990 - March 1992 were asked to participate in a survey concerning the health of their expected child after child birth. Among these soon-to-be mothers, 98 percent accepted to participate. An illustration of the timing of the set-up can be seen in Figure 1. The first survey was collected when the child was 3.5 years old. This survey included a number of questions relating to the health of the child and parents' time allocation.<sup>2</sup> In 2001, a second survey was collected among the parents. The children were in 2001 in the range of 9 to 12 years old. The 2001 survey covers some health measures for the child along with a range of parental and family measures and early childhood measures. The following year, 2002, all parents, regardless of earlier participation, received a follow-up survey, in which they answered the Strength and Difficulties Questionnaire (SDQ). This survey was relatively short and easy to answer, giving a response rate of 61.7%. At the same time, the children's teachers received the SDQ survey, and for 46% of the sample we have SDQ data reported by both parents and teachers. Thus, each child's behavior is evaluated both by the parents (in 2001 and 2002) and subsequently by the teacher (in 2002).

Table 1 shows the sample selection. The sample is restricted to observations where parental evaluation of the child is non-missing. Thus, there may be a problem of systematic sample attrition since low-resource parents may have lower response rates.

---

<sup>2</sup>Data from the first survey are only available for the younger cohort because the idea of giving a survey when the child was around 3.5 years old came halfway through the period, implying that half of the children were already older.



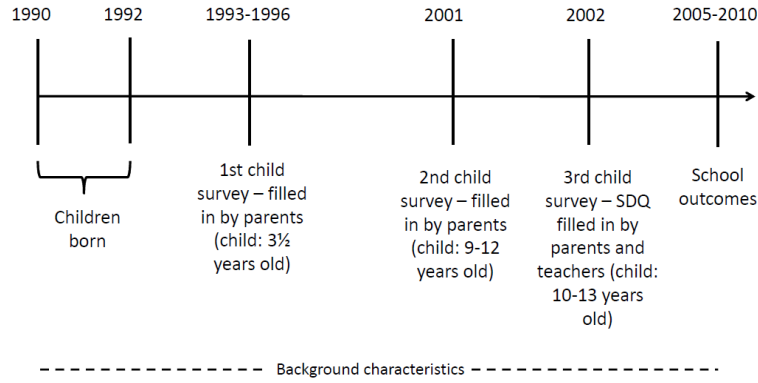


Figure 1: Timing of the set-up

Table 1: Sample selection

	No. obs.	%
Full sample	10,907	100.00
Parental SDQ	6,729	61.69
Teacher SDQ	5,053	46.33
Gender of the teacher	1,662	15.24

### 3.2 The SDQ Scores

The Strengths and Difficulties Questionnaire (SDQ) was first developed by Goodman (1997), and it includes 25 questions<sup>3</sup>, which are answered by both parents and teachers concerning the child’s mental health. Following Goodman et al. (2010), we use 20 of the 25 questions to construct two groups of behavioral problems; externalizing problems (SDQ1) and internalizing problems (SDQ2), see Table 14 in Appendix A for a detailed list of the questions. According to Goodman et al. (2010), this categorization is appropriate for an analysis of a low-risk sample, i.e. ordinary school classes and not special needs classes. What externalizing and internalizing behavior captures, can be seen in Table 14 in Appendix A. Externalizing behavior captures ‘acting out’ behavior. For example, one of the questions in the questionnaire concerns whether the child often has temper tantrums or hot tempers. Internalizing behavior captures internal problems. For example, one of the questions concerns whether the child would rather be solitary and whether the child tends to play alone. So these two behavioral measures capture different aspects of being behaviorally challenged.

<sup>3</sup>SDQ is a well-documented questionnaire investigating children’s behavioral skills, see Goodman (1997), Goodman et al. (2010) and <http://www.sdqinfo.org/a0.html>

Each of the questions can be marked 'not true', 'somewhat true' or 'certainly true', and they are given scores from 0 to 2.

The total score of the 10 (10) measures of SDQ1 (SDQ2) is then found by summing the questions in each category, giving a score ranging from 0 to 20. The scores are given such that the higher the score, the more problematic behavior.

In our main analysis, we use the parental assessment of the child's behavior in order to analyze to which extent there are gender differences in the relation between behavior and school outcomes. We do this as we believe that the parents are more informed on any problems with the child's behavior in different environments than what we would expect from the teachers. Even though the teachers are aware of what happens in school, they might not be aware of what happens in the child's home or with friends, etc. Thus, they may have less focus on internalizing problems than the parents. Teachers may also be more likely than parents to (unintentionally) let current academic performance in the class affect the student's SDQ assessment. However, parents' assessments may suffer from a problem of a 'common standard' for the students since they typically only observe their own child. A recent study by Datta Gupta et al. (2012) shows that the SDQ assessment is sensitive to whether it is the father or the mother who answers the survey questions. In order to test the robustness of our results and also to test if parents and teachers evaluate the children differently, we make the same analysis and estimations using teacher assessments in Section 5, where this pattern is analyzed in more detail.

The distribution of SDQ1 and SDQ2 measures based on parents' evaluation is shown in Figure 2.

As found in other studies, Figure 2 shows that boys tend to score higher values on the externalizing problems (SDQ1) compared to girls, while the distributions are much more alike for internalizing problems (SDQ2).

When estimating the effects of behavioral skills, we focus on problematic behavior. Therefore, the SDQ1 and SDQ2 variables are split into three categories according to Goodman et al. (2010): 'Normal behavior', 'borderline behavior', and 'abnormal behavior'. As shown in Figure 2, externalizing behavior is categorized as normal if the SDQ score is below 8, borderline behavior is between 8 and 9, and abnormal externalizing behavior is a SDQ score above 9. Internalizing behavior is likewise categorized as normal, when the SDQ score is below 6, borderline behavior when the score is between 6 and 7, and abnormal behavior when the SDQ score is above 7. In our empirical analysis, normal behavior is the reference category.

In the psychology literature, it is argued that the thresholds should vary between boys and

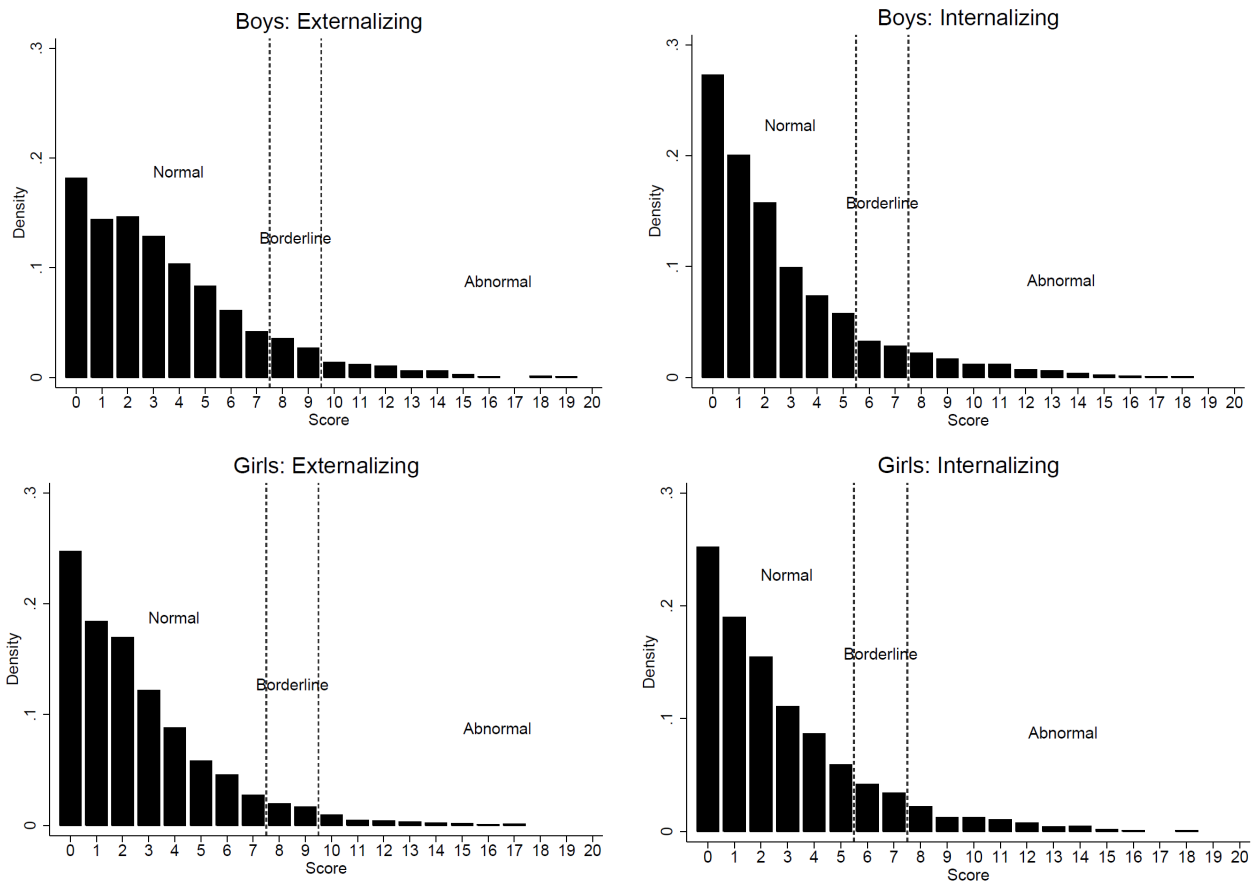


Figure 2: SDQ scores for boys and girls

girls, such that the limits are lowered for girls. However, to our knowledge, no previous study neither use different thresholds nor provide evidence on teachers and parents having different standards in evaluating behaviors of boys and girls. We use the same threshold values for boys and girls, recognizing that this might be a conservative estimate for girls. Since we include a girl indicator in the estimated regression models, we expect this variable to capture a potential difference between boys and girls with respect to the thresholds.

The distributions of the categorizations of behavioral problems are shown in Table 2 for boys and girls. About 11% of the boys have abnormal or borderline externalizing behavioral problems while the figure is 6% for girls. For internalizing behavioral problems the figure is 14% for boys and slightly higher for girls (15%).

Table 2: Frequency of SDQ categories by gender

	Boys		Girls	
	N	%	N	%
<b>Externalizing (SDQ1)</b>				
Normal	3032	88.8	3113	94.0
Borderline	211	6.2	118	3.6
Abnormal	173	5.1	82	2.5
Total	3416	100	3313	100
<b>Internalizing (SDQ2)</b>				
Normal	2931	86.0	2818	85.1
Borderline	205	6.0	249	7.5
Abnormal	274	8.0	244	7.4
Total	3410	100	3311	100

### 3.3 School Outcome Variables

Compulsory schooling in Denmark typically starts at the age of seven<sup>4</sup>, and ends with an exit exam after 9th grade. In 9th grade all students receive annual marks in Reading and Math. At the end of the school year the students take the exit exam, where they receive marks in Reading and Math. After compulsory schooling individuals can choose either to leave the educational system, enroll in high school, or enroll in vocational training programs.

The outcome variables stem from administrative registers on school grades and enrollment into high school (including vocational education). We use four outcome variables: Whether the student took the 9th grade exit exam, the 9th grade exit exam results in Math and Reading, and enrollment in high school or vocational training before the age of 19. This means that we observe the outcome variables several years (5-8 years) after the SDQ questionnaire was filled by the parents. 9th grade is typically passed at the age of 15 (depending on age at school start and class repetition).

In 2005-2006, when the exit exam results are observed for the students in our sample, the Danish grading scale had 10 values, ranging from 0 to 13.<sup>5</sup> As part of the grading, teachers in compulsory school also give term marks. The term marks are expected to show the outcomes of the students during the semester contrary to the exam which is based on a 'point estimate'. However, the exam result is based on the teacher's as well as an external reviewer's evaluation. We expect the exam results to be more 'neutral' than the term marks with respect to behavioral problems which might bias the term marks. Therefore, we prefer to use the exam results.<sup>6</sup> We

<sup>4</sup>In 2009 this was changed to a school starting age of six.

<sup>5</sup>The grading scale was: 00, 03, 5, 6, 7, 8, 9, 10, 11 and 13. The lowest passing grade was 6, and the average grade was 8.

<sup>6</sup>Boys get higher grades at exam compared to the term mark, and opposite for girls, while the enrollment

standardize the scale to an outcome measure with zero mean and a standard deviation of 1 (for the two grading outcomes). The standardization is based on the full sample of students for whom we have register information.

Table 3: School outcomes by gender

Outcomes	Girls			Boys			t-value
	N	Mean	Std.	N	Mean	Std.	
9th grade exit exam	3185	0.9758	0.1536	3220	0.9776	0.1479	-0.48
Reading (exit exam)	3142	0.4378	0.9235	3167	0.0892	0.9577	14.72
Math (exit exam)	3125	0.2572	0.8949	3183	0.3833	0.8752	-5.66
High school enrollment	3313	0.9185	0.2736	3416	0.9013	0.2982	2.46

Table 3 shows the standardized average outcome variables for girls and boys in the sample.<sup>7,8</sup> As found in other studies, girls get higher grades in Reading and boys get higher grades in Math. The gender differences are less significant with respect to taking the 9th grade exit exam. However, girls who are categorized as having a borderline externalizing behavior have a very high proportion (30 percent) who do not take the 9th grade exit exam (significantly higher than boys). In all sub-samples, girls tend to enroll in high school at a marginally (though significantly) higher rate than boys. Since the mean grades are (significantly) positive, this indicates that the sample of students in Table 3 for which we have SDQ information is positively selected in the grade distribution from the full sample of students.

Table 4 shows the average values of the four outcome variables by gender and by category of externalizing and internalizing behavioral problems. The last column gives the t-test for differences between girls and boys. The overall picture is that in all categories of behavioural problems, girls perform better than boys in Reading and opposite in Math. There is no tendency that the gender gap in school outcomes is mainly found among children with behavioral problems. The gender gap is also highly significant for 'normal' children.

rate at high school is significantly higher for girls than for boys. The difference between exit exam grades and term marks may indicate that girls and boys perform differently during their daily school activities and when they are tested at exam. See e.g. Niederle and Vesterlund (2010), who argue that the gender gap reflects that boys are more competitive and perform better in competitive environments than girls.

<sup>7</sup>The Danish compulsory school conducts exams in oral Danish, written Danish and written Math. The students must write an essay at the exam in written Danish and this is combined with a spelling grade. In this analysis, we are interested in the students' ability to read and understand, which is the content of the oral exam. For this reason, we focus on the oral exam in Danish and the written exam in Math.

<sup>8</sup>Note that children in the sample in general perform better compared to the national average. This is a result of selection in who participated in the SDQ survey, where as an example immigrant students are underrepresented.

Table 4: SDQ and school outcomes by gender

	Girls			Boys			t-value
	N	Mean	Std.	N	Mean	Std.	
<b>Externalizing - Abnormal</b>							
9th grade exit exam	59	0.9322	0.2536	128	0.9141	0.2814	0.44
Reading	58	-0.4131	0.9484	120	-0.4617	0.9631	0.32
Math	56	-0.8022	0.9778	120	-0.3775	1.0316	-2.64
HS Enrollment	82	0.6463	0.4810	173	0.6821	0.4670	-0.56
<b>Internalizing - Abnormal</b>							
9th grade exit exam	204	0.9559	0.2059	211	0.9194	0.2728	1.54
Reading	202	0.0129	0.9879	199	-0.1633	1.0448	1.74
Math	195	-0.0761	0.9487	205	0.0145	0.9573	-0.95
HS Enrollment	244	0.8074	0.3952	274	0.7336	0.4429	2.00
<b>Externalizing - Borderline</b>							
9th grade exit exam	102	0.8824	0.3238	165	0.9515	0.2154	-1.91
Reading	96	0.0459	0.8028	160	-0.3327	0.9303	3.44
Math	92	-0.1960	1.0001	161	-0.1929	0.8792	-0.02
HS Enrollment	118	0.7034	0.4587	211	0.7393	0.4400	-0.69
<b>Internalizing - Borderline</b>							
9th grade exit exam	230	0.9739	0.1597	173	0.9422	0.2340	1.53
Reading	224	0.2619	0.9089	165	-0.1605	0.9257	4.48
Math	227	0.0360	0.9880	163	0.1564	0.8692	-1.27
HS Enrollment	250	0.8560	0.3518	205	0.7756	0.4182	2.19
<b>Externalizing - Normal</b>							
9th grade exit exam	3024	0.9798	0.1406	2927	0.9819	0.1334	-0.58
Reading	2988	0.4670	0.9161	2887	0.1354	0.9464	13.64
Math	2977	0.2911	0.8739	2902	0.4467	0.8411	-6.96
HS Enrollment	3113	0.9338	0.2486	3032	0.9251	0.2632	1.33
<b>Internalizing - Normal</b>							
9th grade exit exam	2752	0.9771	0.1496	2834	0.9841	0.1250	-1.90
Reading	2716	0.4844	0.9103	2801	0.1210	0.9486	14.52
Math	2703	0.3007	0.8747	2813	0.4228	0.8617	-5.22
HS Enrollment	2820	0.9333	0.2495	2935	0.9257	0.2623	1.13

## 3.4 Covariates

We use an extensive list of covariates relating to family composition, health, and school information in our conditioning set in order to control for as much unobserved variation in relevant background factors as possible. A complete variable description with a list of included covariates and outcomes can be found in Table 12 in Appendix A. All covariates included are measured before the school start (when the child is 0-6 years old) except for the behavioral measures (SDQ) which are measured at the age of 10-12, i.e. before school outcomes are measured at the age of 15-19. This is, of course important, as we want to make sure, that the effect does not work the other way around.

### 3.4.1 Basic characteristics

First of all, we include indicators for the gender of the child, and whether the child is a native Dane, whether the child has any younger or older siblings, and how many younger or older siblings. The measure of siblings is determined when the child is six years old, which makes sure that no effects can go through having more siblings after starting in school.

Since day care is widely used in Denmark, it is natural to include variables for the type of day care the child is in. We determine the type of day care when the child is four years old. Parents with problematic children may move between municipalities more often than other parents. For that reason, we include an indicator for whether or not the family has moved between municipalities, and a variable for how many times the family has moved from one municipality to another. We measure this just before the child is 7 years old.

We also control for age at completing compulsory schooling. We do this for several reasons. First of all, the children in this sample are born in three consecutive years, which might influence the given grades. Also, as girls tend to start earlier in school than boys, it is important to take the year effects into account.

Further, we include health information from the mother's pregnancy and the birth of the child, which include the length of the pregnancy, whether the child was born prematurely, extremely prematurely, whether there was any complications at birth and the birth weight. We also include a dummy for the birth year.

The register data allow us to control for whether the child has been diagnosed with a mental or behavioral disorder. We construct the variables such that a diagnosis should occur before the child is 7 years old. One could argue that examining behavioral problems would only capture

the effect for those children who are diagnosed with a behavioral disorder, but this is not the case in this study, as we are able to control for that as well.

A child with depression or cardiovascular problems might have different behavior and school outcomes than other children. Therefore, we also include information on whether the child receives cardiovascular and antidepressant medicine before school age.

### **3.4.2 Parental characteristics**

Not only a child's own characteristics may be important for behavioral problems and school outcomes. Therefore, we include a range of parental characteristics as well. Parental employment might be an important factor, and we include employment along different dimensions to capture different aspects of the influence on the child. We include the degree of employment when the child is 4, 5 and 6 years old. This captures to what extent the parents have been employed during the year in the years up until the child enters school. The sector of employment is also included, i.e. private versus public, and whether the parents are full-time or part-time employed. Furthermore, the occupation is included, e.g. a manager versus an unskilled worker.

Besides employment measures, the highest obtained education for the parents might matter in the formation of skills for the child, which is why this is also included in our conditioning set.

As for each child, we include three specific health measures regarding the mental health. That is, we include an indicator for whether the parents have been diagnosed with a mental or behavioral disorder, an indicator for being prescribed antidepressant medicine, and an indicator for cardiovascular medicine. These health measures for the parents are determined before the child is 6 years old. In addition, we have information from the crime registers, enabling us to identify parents with criminal behavior before the child turns 7 years old. All of the parental characteristics are included for both the mother and the father.

Since family background characteristics may be important for the child's development, we include a number of family characteristics. First, the mother's age at birth might matter for the ability and skill formation of the child, which might suggest a quantity-quality trade-off, see e.g. Miller (2009). We further include the age difference between the father and the mother. We include three variables which may proxy family resources: (1) Having a young mother, defined by an indicator being 1 if the mother was 23 years old or younger at the time of birth, 0 else. (2) An indicator for having a non-intact family defined by an indicator taking the value of 1 if the biological parents are cohabiting up until the child is six years old, 0 else. (3) An indicator



for whether it is a low-income family, measured as 50% of the median income.

### 3.4.3 School fixed effects and class information

School and class data are obtained for all children who were born in the region of Aarhus, and who at one point in time entered compulsory school in that region. This means that we do not include children who were born in the region of Aarhus but never went to school in the Aarhus area. Even though the majority of our sample has information on school and class, there are still some missing data. Individuals with missing school and class data are kept in the sample and grouped in a 'missing group'. (An indicator for missing data is included).

In order to capture variation in class characteristics, we include the fraction of boys in class, the fraction of boys in class squared, and a variable telling how many times the child changed classes before 2002, i.e. before behavior is measured. Because the class environment may affect behavior and vice versa, children with problematic behavior might be more prone to enroll in another school and class.

## 3.5 Descriptive Statistics

Table 13 in Appendix A shows the sample means for all included covariates, the behavioral measures and the outcomes. As found in other studies, see for instance Bertrand and Pan (2011), there are no significant differences between boys and girls regarding the parental background characteristics for the majority of covariates. However, for some variables differences for boys and girls are present. The birth weight of boys is larger than for girls, which is a matter of physiological differences. Also, boys are more likely to have a psychiatric diagnosis, which is also found in Gaub and Carlson (1997). Besides that, girls finish 9th grade sooner than boys, which corresponds to boys enrolling later than girls.

## 4 Estimation Results

The estimated model of school outcomes and behavior is based on the following specification with ordinary least squares

$$\begin{aligned}
 Outcome_{is} = & \alpha + Girl_{is}\alpha^g + \sum_{j=1}^4 (SDQ_{j,is}\gamma_j + SDQ_{j,is}Girl_{is}\gamma_j^g) \\
 & + X_{is}\beta + \eta_s + \varepsilon_{is},
 \end{aligned} \tag{1}$$

*Outcome* is the standardized exit exam outcome or enrollment indicator, *Girl* is an indicator for being a girl which is interacted with all SDQ variables in order to allow the behavioral parameters in the model to be gender-specific.  $SDQ_j$  denotes SDQ measure  $j$  (i.e. abnormal and borderline externalizing and internalizing behavior),  $X$  is a vector of other covariates, including child, parental, family background characteristics, and class characteristics, etc., and  $\eta_s$  is school fixed effects. The parameters of interest are  $\gamma_j^g$  which indicates whether there are gender differences in the effects of behavioral skills on school outcomes.

In order to identify causal effects from behavior to school outcomes, we either have to have valid instruments of behavior, experimental data or be able to identify all relevant covariates, and all covariates related to behavior other than the variables of interest have to be predetermined, see Angrist and Pischke (2009). We do have a large battery of background characteristics and child information on the children from their birth up until the age of 18-20. Behavior is measured at the age of 10-12 years, school outcomes are measured at the age of 15-18, and social background control variables when the child is aged 0-6. But since we do not have any valid instruments at hand nor do we have experimental data, we have to rely on the conditional independence assumption which we cannot test. Thus, though we sometimes use the term 'effect' in the following sections, we do not claim to present causal effects.

## 4.1 Main Model

The results from the main model are presented in Table 5 for the four school outcomes. Most of the SDQ externalizing variables have significantly negative coefficients. In column 1, the results for the probability of taking the 9th grade exit exam are shown. Students with abnormal externalizing behavior are estimated to have a 3.6 percentage points lower chance of taking the 9th grade exit exam, their score in Reading is 0.36 lower, in Math 0.57 lower and their high school enrollment chance 4.1 percentage point lower than children without behavioral problems. For children with borderline externalizing behavior the estimated coefficients are of the same sign, but numerically smaller in most cases.

The coefficients of the interaction effect between the girl indicator and the indicators for externalizing behavior are surprising. The estimations indicate that the girls with abnormal externalizing behavior tend to have the same school outcome as boys in this category. But girls with borderline externalizing behavior seem to have a significantly lower chance of taking the 9th grade exit exam compared to boys (8 percentage points lower), while girls with borderline externalizing behavior seem to get significantly higher grades in Math and Reading than boys

with borderline externalizing behavior.

For children with a high SDQ score on internalizing behavior, the story seems to be different. The coefficients of the main effect variables for internalizing behavior in rows 3-4 are numerically small and in most cases insignificant. The girl interaction effects also tend to be insignificant, except for girls with abnormal internalizing behavior where the coefficient is significantly positive for taking the 9th grade exit exam and high school enrollment.

Table 5: Estimates of behavioral effects (Parents' SDQ) on school outcomes

	9th grade exit exam	Reading (exit exam)	Math (exit exam)	HS Enrollment
Abnormal Externalizing Behavior (SDQ1)	-0.0363* (0.0190)	-0.3554*** (0.1090)	-0.5716*** (0.0807)	-0.0406 (0.0263)
Borderline Externalizing Behavior (SDQ1)	-0.0180 (0.0156)	-0.3122*** (0.0666)	-0.4837*** (0.0648)	-0.0574*** (0.0219)
Abnormal Internalizing Behavior (SDQ2)	-0.0452*** (0.0167)	0.0053 (0.0793)	-0.0159 (0.0536)	-0.0265 (0.0204)
Borderline Internalizing Behavior (SDQ2)	-0.0297 (0.0195)	-0.1277* (0.0766)	-0.0632 (0.0770)	-0.0419** (0.0201)
Girl	-0.0080*** (0.0030)	0.3637*** (0.0282)	-0.1275*** (0.0209)	-0.0033 (0.0048)
Abnormal Externalizing Behavior*Girl (SDQ1)	0.0070 (0.0267)	-0.0059 (0.1438)	-0.1305 (0.1609)	-0.0518 (0.0416)
Borderline Externalizing Behavior*Girl (SDQ1)	-0.0841* (0.0494)	0.1710* (0.1013)	0.2556** (0.1061)	-0.0526 (0.0428)
Abnormal Internalizing Behavior*Girl (SDQ2)	0.0430** (0.0205)	-0.2027 (0.1225)	0.0251 (0.0700)	0.0450* (0.0234)
Borderline Internalizing Behavior*Girl (SDQ2)	0.0391 (0.0236)	0.0461 (0.0728)	0.0202 (0.0806)	0.0344 (0.0278)
<i>N</i>	6122	6031	6031	6122
<i>R</i> <sup>2</sup>	0.087	0.201	0.257	0.166

Standard errors in parentheses and clustered at the school level.

Conditioning variables: Basic, parental, School FE, and class variables.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

## 4.2 The Importance of the Conditioning Set

In order to evaluate the significance of a large conditioning set, Table 6 shows the estimates of behavioral problems on Math (exit exam), where additional controls are gradually added.

Column (5) in Table 6 represents the results shown in Table 5.<sup>9</sup>

Table 6: Estimates of behavioral effects (Parents' SDQ) on Math (exit exam)

	(1)	(2)	(3)	(4)	(5)
	No controls	(1) + Basic	(2) + Parental	(3) + School FE	(4) + Class Var.
Abnormal Externalizing Behavior (SDQ1)	-0.7493*** (0.0831)	-0.7009*** (0.0825)	-0.5598*** (0.0811)	-0.5693*** (0.0808)	-0.5716*** (0.0807)
Borderline Externalizing Behavior (SDQ1)	-0.5934*** (0.0823)	-0.5565*** (0.0833)	-0.4754*** (0.0616)	-0.4801*** (0.0651)	-0.4837*** (0.0648)
Abnormal Internalizing Behavior (SDQ2)	-0.2047*** (0.0559)	-0.1499*** (0.0498)	-0.0301 (0.0531)	-0.0168 (0.0539)	-0.0159 (0.0536)
Borderline Internalizing Behavior (SDQ2)	-0.1330 (0.0806)	-0.1026 (0.0663)	-0.0305 (0.0822)	-0.0688 (0.0756)	-0.0632 (0.0770)
Girl	-0.1469*** (0.0219)	-0.1509*** (0.0214)	-0.1378*** (0.0205)	-0.1363*** (0.0207)	-0.1275*** (0.0209)
Abnormal Externalizing Behavior*Girl (SDQ1)	-0.2350 (0.1595)	-0.1889 (0.1607)	-0.1139 (0.1465)	-0.1398 (0.1605)	-0.1305 (0.1609)
Borderline Externalizing Behavior*Girl (SDQ1)	0.1760 (0.1172)	0.1652 (0.1191)	0.2781*** (0.1023)	0.2527** (0.1078)	0.2556** (0.1061)
Abnormal Internalizing Behavior*Girl (SDQ2)	0.0061 (0.0759)	-0.0015 (0.0707)	0.0262 (0.0711)	0.0264 (0.0705)	0.0251 (0.0700)
Borderline Internalizing Behavior*Girl (SDQ2)	-0.0760 (0.1100)	-0.0762 (0.0890)	-0.0305 (0.0849)	0.0255 (0.0795)	0.0202 (0.0806)
<i>N</i>	6308	6258	6031	6031	6031
<i>R</i> <sup>2</sup>	0.054	0.093	0.225	0.256	0.257

Standard errors in parentheses and clustered at the school level.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Basic characteristics are for instance type of day care, health, native, siblings, etc.

For the full list of included covariates, see Section 3.4

Column (1) shows the estimates of behavioral problems on Math (exit exam), where no additional controls are added. Including the basic characteristics of the child (such as day care and health measures) in column (2) does not change the estimates much. Including the parental characteristics in column (3) lowers the estimates in magnitude. It is worth noting that the inclusion of conditioning variables changes the explanatory power of the model considerably. In column (1) where no controls are added the R-squared is 5 percent, while including family and parental characteristics increases this figure to 23 percent, including school fixed effects

<sup>9</sup>We only present the estimates of gradually including covariates for the Math outcome. The pattern is the same for the Reading outcome, except that the level effect is of opposite sign. For the outcome of attending high school, there is no gender difference in levels either. The estimations are available upon request.

almost doubles the R-squared to 26 percent, but adding class variables does not change the R-squared much. During this process of expanding the conditioning set of variables, the estimated coefficients of the SDQ variables are fairly stable. Thus, the estimates in column (5) might be considered conservative and fairly robust.<sup>10</sup>

## 5 Are Teachers' SDQ Different from Parents' SDQ?

In this section, we investigate whether parent and teacher evaluations are different in the assessment of child behavior. Parents and teachers might have different views of what defines bad behavior. Further, they are assessing the child in different settings, which might also generate different answers. Parents may be more prone to exaggerate the positive parts of their child's behavior and underreport problematic behavior, whereas teachers might evaluate the child relative to the other children in the class, giving different results. Further, since the teacher is the responsible agent in the class room, it may change the relationship between SDQ measures and student outcomes which was observed in the previous section because teachers may 'compensate' in different ways when he or she is aware of behavioral problems. Or the opposite: If the teacher has daily conflicts with some students with (externalizing) behavioral problems, this may have effects on both teacher SDQ scores and school outcomes. Parents do not have easy access to a comparison to their child or a standard for normal behavior. In contrast to this, teachers actually do have a comparison group. Further, there may be gender differences in these assessment patterns. According to Gillberg et al. (2012) who analyze children with externalizing behavior (ADHD diagnosis), parents more often tend to identify girls with these symptoms than teachers do. Therefore, we look into how the teacher evaluations of behavior affect the school outcomes for boys and girls in order to evaluate whether teachers and parents tend to agree on behavioral problems, and whether it matters for the results if they do not agree.

Figure 3 shows how the students are evaluated by both parents and teachers. The picture is clear: Teachers are more prone to assess 'extreme' behavior, i.e. an SDQ score of 0 or above 13, whereas the parental evaluations are more evenly distributed.

In the same spirit, Table 7 shows the categorization of students in abnormal, borderline and normal behavior by parent and teacher evaluations. The picture is the same: teachers

---

<sup>10</sup>The corresponding tables with 9th grade exit exam, Reading and high school enrollment as outcomes show the same pattern. These tables are available on request from the authors.

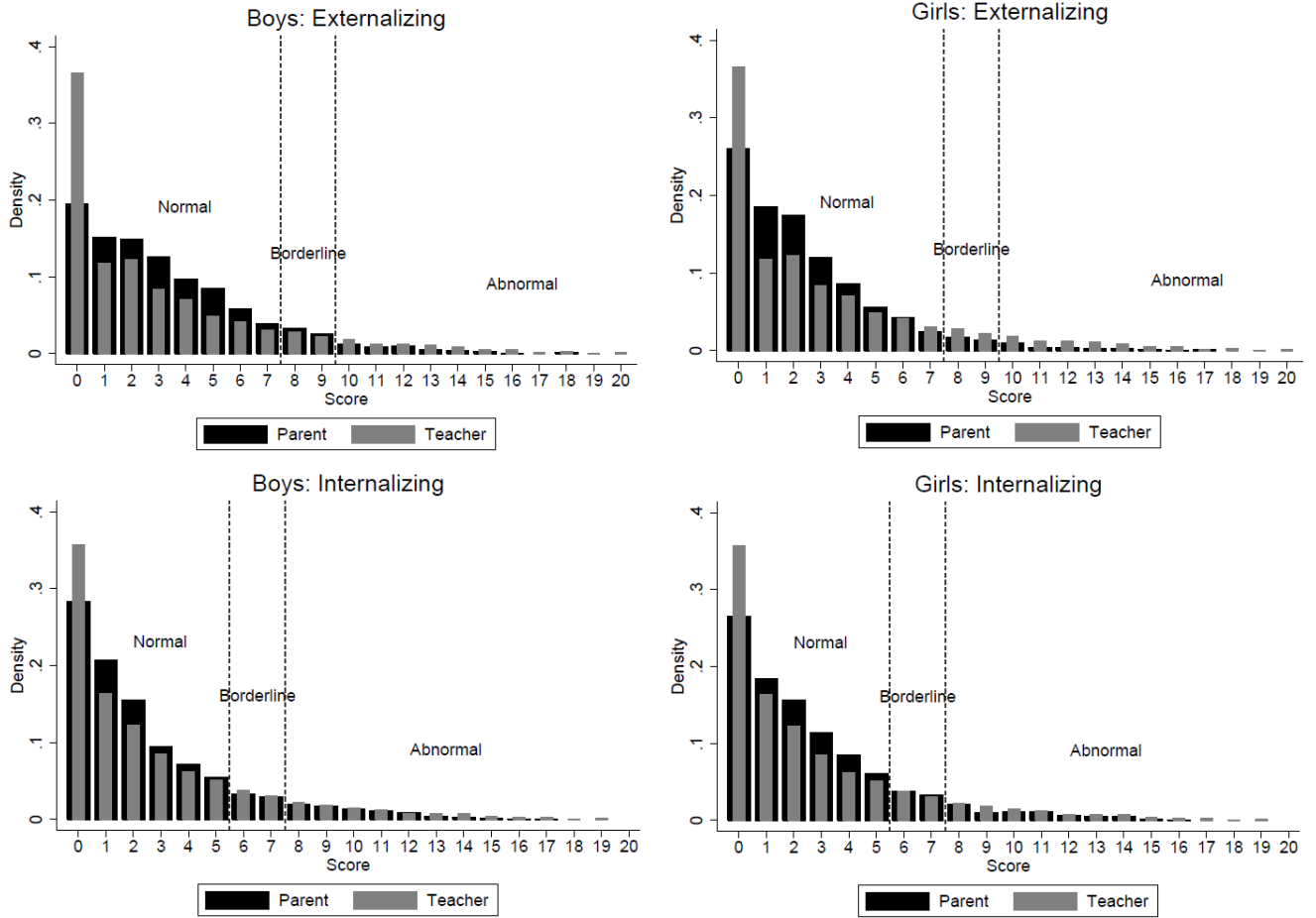


Figure 3: SDQ scores for boys and girls evaluated by parents and teachers.

Table 7: Behavioral assessment of parents and teachers

Teacher evaluation						
Externalizing	Girls			Boys		
<b>Parental evaluation</b>	Abnormal	Borderline	Normal	Abnormal	Borderline	Normal
Abnormal	23	5	33	59	16	34
Borderline	6	3	67	43	21	83
Normal	49	46	2,307	194	157	1,907
Teacher evaluation						
Internalizing	Girls			Boys		
<b>Parental evaluation</b>	Abnormal	Borderline	Normal	Abnormal	Borderline	Normal
Abnormal	75	19	78	82	31	77
Borderline	38	19	120	28	26	99
Normal	132	132	1,925	128	110	1,930

categorize more children as normal and as abnormal, whereas parents report more borderline behavior. This might either reflect that parents have a more nuanced view of their child's behavior, or that teachers are more aware of, what is normal and not normal behavior.

Table 8 shows the estimates of the main model for the four school outcomes using teacher evaluations. The model estimated in Table 8 is fully comparable to the model on the parental SDQ scores in Table 5, except that there are fewer observations included in Table 8.<sup>11</sup>

Table 8: Estimates of behavioral effects (Teachers' SDQ) on school outcomes

	9th grade exit exam	Reading (exit exam)	Math (exit exam)	HS Enrollment
Abnormal Externalizing Behavior (SDQ1)	0.0048 (0.0096)	-0.1914*** (0.0522)	-0.5269*** (0.0459)	-0.0290* (0.0169)
Borderline Externalizing Behavior (SDQ1)	-0.0006 (0.0105)	-0.3678*** (0.0745)	-0.3892*** (0.0709)	-0.0366 (0.0222)
Abnormal Internalizing Behavior (SDQ2)	0.0012 (0.0162)	-0.1316** (0.0537)	-0.0699 (0.0659)	-0.0428** (0.0206)
Borderline Internalizing Behavior (SDQ2)	-0.0204 (0.0157)	-0.0814 (0.0727)	-0.0211 (0.0619)	-0.0332 (0.0221)
Girl	-0.0012 (0.0041)	0.3347*** (0.0308)	-0.1565*** (0.0273)	-0.0052 (0.0062)
Abnormal Externalizing Behavior*Girl (SDQ1)	-0.0179 (0.0348)	-0.1271 (0.1445)	-0.2166 (0.1315)	-0.0098 (0.0486)
Borderline Externalizing Behavior*Girl (SDQ1)	-0.0452 (0.0522)	0.1516 (0.2027)	-0.1433 (0.1092)	-0.0791 (0.0562)
Abnormal Internalizing Behavior*Girl (SDQ2)	-0.0218 (0.0172)	-0.0625 (0.0858)	0.0960 (0.0966)	0.0138 (0.0286)
Borderline Internalizing Behavior*Girl (SDQ2)	0.0322* (0.0184)	-0.0223 (0.1045)	-0.0816 (0.0824)	-0.0007 (0.0348)
<i>N</i>	4625	4565	4562	4625
<i>R</i> <sup>2</sup>	0.086	0.203	0.269	0.165

Standard errors in parentheses.

Conditioning variables: Basic, parental, School FE, and class variables.

Standard errors are clustered at the school level.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Most of the results in the upper part of Table 8 based on teachers' SDQ scores are similar

<sup>11</sup>In alternative estimations, not shown here, we have restricted the estimations based on parents' SDQ scores to the students for whom we also have the teacher's SDQ scores shown in Table 8. The results do not change in any notable way when restricting the sample in this way. Thus, we conclude that the differences in coefficients between teachers' and parents' SDQ scores on student outcomes are not due to fewer observations. The results are available upon request from the authors.

to the results found in Table 5 using the parents' SDQ scores, implying that no matter who evaluates the student, behavioral problems tend to have a negative impact on school outcomes. In the lower part of Table 8 we find only insignificant coefficients, except for one (girls with borderline internalizing behavior in taking the 9th grade exit exam). Thus, our results indicate, that it matters whether the SDQ scores are given by the parents or the teachers with respect to the relation to student outcomes. The gender-specific coefficients tend to disappear when we base our behavioral measures on teachers' evaluation. Based on Table 8 we are not able to explain which mechanisms drive this result. It may be that teachers are more gender-neutral in evaluating the behavior of boys and girls, or it may be that teachers tend to moderate or compensate for the potential negative effects of bad behavior when they are aware of these behavioral problems. The results may also reflect that parental SDQ scores (and also teachers' SDQ scores) are related to their own social background, and this relationship may be different for boys and girls if parents from different social backgrounds have different gender-specific expectations regarding the behavior of their children. We look closer at this question in Section 6.

For 1,662 students we also observe the gender of the teacher. It might be that the gender of the teacher could have an effect on the different school outcomes. In the study by Datta Gupta et al. (2012) it is found that mothers and fathers have significantly different SDQ assessment of their children. Thus, a hypothesis may be that the SDQ scores depend on the gender of the person who makes the assessment.<sup>12</sup> Alternatively, the gender of the teacher may have effects on the behavior of the children. Therefore, we have estimated an alternative model (not shown here, but the results are available on request from the authors) where we include as additional right-hand-side variables an indicator for the teacher being male, and interactions with this variable with SDQ-score, Girl, and SDQ-score\*Girl. In general we do not find many significant effects relating to the gender of the teacher. One obvious reason may be that our sample is too small when restricted to only 1,662 observations to document significant effects. When significant effects appear, our results indicate that male teachers improve the student outcomes for children with externalizing behavioral problems and opposite for children with internalizing problems. We do not find notable gender effects for this relation.

---

<sup>12</sup>Unfortunately, we do not have information of the gender of the parent who filled the parents' survey.



## 6 Social Background, Behavior and Student Outcome

It is a well known fact that student outcome is highly sensitive to social background. Bertrand and Pan (2011) show that especially boys from low-resource families are worse off when they have behavioral problems, i.e. the gender-specific impact of behavioral problems on later school outcomes interacted with social background. In this section we look at this interaction between gender, social background and behavior and the impact on student outcome. Our a priori expectation is - in line with Bertrand and Pan (2011) - that boys from low-resource families are relatively more vulnerable than boys from other families and more vulnerable than girls from low-resource families. Thus, we expect that the observed girl-SDQ interaction coefficient is more negative for children born in low-resource families than in other families.

Table 9 shows the descriptive statistics for means of outcomes by three categories of low-resource families (low-income families, with a young mother, and non-intact families, i.e. the biological parents stopped living together before the child enters school) separately for boys and girls.<sup>13</sup> We merge the categories 'abnormal' and 'borderline' in Table 9 and the regression analyses. We also restrict the analysis to Reading and Math outcomes.

As expected, students from low-resource families have lower outcomes than 'other students'. This pattern is mainly evident for grades in Math where children (girls as well as boys) with young mothers have the lowest mean grades, children in non-intact families have higher grades than children with young mothers, and children in low-income families have higher grades than children from non-intact families and lower grades than 'other families', i.e. families which are not low-resource families. (The differences in mean grades between the four sub-groups are all significant). Looking at the gender gap in Math grades, boys score significantly higher mean grades in all four sub-groups. But the lowest gap is found for the group 'other families'.

For Reading grades the picture is slightly different. In all four sub-groups, girls obtain higher mean grades in Reading than boys. But children from low-income families obtain almost as high grades as 'other children', and boys obtain even higher mean grades. The gender gap in Reading is fairly stable across sub-groups, except for children in families with a young mother where the gap is insignificant, i.e. girls obtain almost as low grades as boys from this family type. This result is the opposite of what was expected, i.e. that boys in low-resource families experience larger difficulties in the school system than girls.

In order to test the relationship between the observed pattern in SDQ scores and Math and

---

<sup>13</sup>The three categories may overlap in the sense that some students may show up in more of the low-resource family categories.

Table 9: Descriptives: Outcomes and SDQ by gender for low resource families

	All	Other	Low income	Non-intact	Young mother
<b>SDQ</b>					
<b>Proportion with externalizing behavior</b>					
Girls	0.0604	0.0418	0.0931	0.1231	0.1399
Boys	0.1123	0.0895	0.1372	0.1888	0.1742
<b>Proportion with internalizing behavior</b>					
Girls	0.1489	0.1218	0.1861	0.2334	0.2937
Boys	0.1405	0.1154	0.1722	0.2126	0.2516
<b>Outcomes</b>					
<b>Reading</b>					
Girls	0.4379	0.4890	0.4177	0.2236	-0.2337
Boys	0.0889	0.1118	0.0993	-0.0379	-0.1587
<b>Diff: Girls-Boys</b>	<b>0.3490</b>	<b>0.3772</b>	<b>0.3184</b>	<b>0.2615</b>	<b>-0.0749</b>
<b>Math</b>					
Girls	0.2578	0.3640	0.1000	-0.0949	-0.3516
Boys	0.3829	0.4424	0.3024	0.1510	-0.0058
<b>Diff: Girls-Boys</b>	<b>-0.1251</b>	<b>-0.0784</b>	<b>-0.2024</b>	<b>-0.2459</b>	<b>-0.3458</b>

Reading outcomes, we replicate the regressions above for each sub-sample of students. The results for the SDQ variables and the girl indicator are shown in Table 10 for the full sample (column 1) and each sub-group (columns 2-5). As found above, for the full sample there is a strong and significantly negative coefficient of the indicator for externalizing behavioral problems on both Reading and Math (the estimated coefficients are -0.51 and -0.68, i.e. more than one half standard deviation of the scores in Reading and Math). For internalizing behavioral problems, the coefficients are numerically smaller, but still negative and significant (-0.18 to -0.23). When splitting the full sample into sub-groups of low-resource families, an interesting pattern emerges: The significantly negative coefficients of the internalizing problem indicators disappear, except for one coefficient (Math, Low income). Further, there is no clear indication that externalizing behavioral problems are more negatively related to low grades in Math and Reading in low-resource families compared to families which are not observed as low-resource families. Also, we do not find, that the coefficients of the girl\*SDQ measures are positive (i.e. that boys with behavioral problems should be more vulnerable than comparable girls in low-resource families). All interaction coefficients are insignificant! The only positively significant interaction is found for the sub-group of 'other families' (Reading, Internalizing behavior). Of course, a part of the explanation is that the sample size is reduced, when we look at these sub-groups. Actually, we observe that many of the point estimates are the same for the low-resource

Table 10: Estimates of behavioral effects (Parents' SDQ) on Reading (exit exam) and Math (exit exam) for subgroups

	All	Other	Low income	Non-intact	Young mom
<b>Reading</b>					
Externalizing Behavior (SDQ1)	-0.5058*** (0.0599)	-0.5083*** (0.1238)	-0.5173*** (0.1482)	-0.4514*** (0.1364)	-0.5900** (0.2322)
Internalizing Behavior (SDQ2)	-0.2326*** (0.0487)	-0.2204*** (0.0729)	-0.2466 (0.1498)	-0.1814 (0.1094)	-0.0505 (0.2725)
Girl	0.3851*** (0.1039)	0.6335*** (0.1478)	-0.0112 (0.2201)	0.2493 (0.1813)	0.0889 (0.3814)
Externalizing Behavior*Girl (SDQ1)	0.0220 (0.1186)	0.1539 (0.1802)	-0.1226 (0.2121)	-0.0225 (0.1481)	0.3167 (0.2694)
Internalizing Behavior*Girl (SDQ2)	0.0306 (0.0830)	0.1462* (0.0826)	-0.2450 (0.2410)	-0.0251 (0.1626)	-0.3101 (0.3828)
<i>N</i>	4773	3157	870	930	286
<i>R</i> <sup>2</sup>	0.061	0.057	0.090	0.059	0.042
<b>Math</b>					
Externalizing Behavior (SDQ1)	-0.6804*** (0.0729)	-0.6511*** (0.0877)	-0.5610*** (0.1417)	-0.7377*** (0.1572)	-0.4739*** (0.1665)
Internalizing Behavior (SDQ2)	-0.1797*** (0.0448)	-0.1447** (0.0657)	-0.2665* (0.1475)	-0.1479 (0.1098)	-0.1360 (0.1454)
Girl	-0.0243 (0.1622)	0.1529 (0.2126)	-0.4498** (0.1932)	0.0099 (0.2007)	-0.0640 (0.2699)
Externalizing Behavior*Girl (SDQ1)	0.0900 (0.1254)	0.2003 (0.1716)	-0.2786 (0.1784)	0.2671 (0.1944)	-0.0875 (0.3033)
Internalizing Behavior*Girl (SDQ2)	0.0333 (0.0813)	0.0733 (0.0977)	0.0080 (0.1927)	0.0065 (0.1535)	0.3052 (0.2628)
<i>N</i>	4769	3157	868	930	287
<i>R</i> <sup>2</sup>	0.046	0.030	0.078	0.070	0.054

Standard errors in parentheses.

Standard errors are clustered at the school level.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

samples.

Table 10 further indicates that the large positive coefficient of the girl indicator in Reading in full sample (+0.39) stems from the observations in the group of 'other families' (+0.63), but we do not find this result in low-resource families where all coefficients of the girl indicator are numerically smaller and insignificant. Thus, our results indicate that the large and positive gap in Reading grades between girls and boys is mainly due to a 'girl-gap' in normal families, i.e. families which are not categorized as low-resource families. For Math, the results are different. The girl indicator is significantly negative in the sub-group of low-income families, but it is insignificant and sometimes even positive in all other groups.<sup>14</sup>

In order to sum up our results, we perform an Oaxaca decomposition<sup>15</sup> of the gender gap in Reading and Math scores based on the estimation in Table 10. The mean outcome gap between girls and boys is decomposed into three components, an 'endowment component' (differences in characteristics), a 'coefficient component', and an interaction component (which implies that the three components add up to the total raw gap in school outcome). We use boys' coefficient and endowments as the reference point (for simplicity we drop the subscripts):<sup>16</sup>

$$\overline{Outcome}_{girls-boys} = \overline{Z}^g \widehat{\lambda}^g - \overline{Z}^b \widehat{\lambda}^b = (\overline{Z}^g - \overline{Z}^b) \widehat{\lambda}^b + \overline{Z}^b (\widehat{\lambda}^g - \widehat{\lambda}^b) + (\overline{Z}^g - \overline{Z}^b) (\widehat{\lambda}^g - \widehat{\lambda}^b)$$

where superscripts  $g$  and  $b$  refer to girls and boys, respectively, and  $Z$  and  $\lambda$  are characteristics (endowments) and parameters in the estimated model. The results relating to the SDQ behavioral variables are shown in Table 11.<sup>17</sup>

Table 11 indicates that the gender gap in Reading and Math cannot be ascribed to gender differences in behavioral problems, either the endowment component or the coefficient component.<sup>18</sup> Though the previous tables show that behavioral problems are clearly related to school outcomes, we do not find many significant components in Table 11, and in most cases where some of the components are significant, their absolute size is marginal compared to the observed gender gap. Especially for the Reading gap, the SDQ components only account for

---

<sup>14</sup>The analysis is also done for high school enrollment (or vocational education) and the results show no indication of a pattern in the gender gap in behavioral problems on further education. The results are available upon request.

<sup>15</sup>Introduced in Oaxaca (1973).

<sup>16</sup>The decomposition is performed by a STATA program, see Jann (2008).

<sup>17</sup>The results from the estimations, the full decomposition, and the two other school outcome variables are available upon request from the authors.

<sup>18</sup>In Appendix A, Table 15 shows the results of a decomposition relating behavioral problems to taking any further education (i.e. either high school or vocational education). The results show no gender gap in behavioral problems on further education.

Table 11: Oaxaca Decomposition for samples of students from low-resource families: Reading and Math

	All	Other	Low income	Non-intact	Young mother
<b>Reading</b>					
Difference: Girls-Boys	0.3536*** (0.0251)	0.3730*** (0.0303)	0.3112*** (0.0684)	0.2919*** (0.0678)	0.0630 (0.0919)
Percentage	100.00	100.00	100.00	100.00	100.00
Endowments					
SDQ Externalizing	0.0148*** (0.0031)	0.0167*** (0.0060)	0.0086 (0.0081)	0.0098 (0.0068)	0.0236 (0.0270)
Percentage	4.19	4.48	2.76	3.36	37.46
SDQ Internalizing	-0.0017 (0.0015)	-0.0002 (0.0013)	-0.0019 (0.0046)	-0.0047 (0.0052)	0.0047 (0.0181)
Percentage	-0.48	-0.05	-0.61	-1.61	7.46
Coefficients					
SDQ Externalizing	0.0114 (0.0094)	0.0216 (0.0134)	-0.0088 (0.0197)	-0.0091 (0.0235)	0.0658 (0.0425)
Percentage	3.22	5.79	-2.83	-3.12	104.44
SDQ Internalizing	0.0032 (0.0119)	0.0099 (0.0101)	-0.0218 (0.0255)	-0.0182 (0.0293)	-0.0544 (0.0468)
Percentage	0.90	2.65	-7.01	-6.24	-86.35
<i>N</i>	4691	3132	846	881	278
<b>Math</b>					
Difference: Girls-Boys	-0.1165*** (0.0279)	-0.0815** (0.0331)	-0.1945*** (0.0521)	-0.2334*** (0.0657)	-0.2242** (0.0972)
Percentage	100.00	100.00	100.00	100.00	100.00
Endowments					
SDQ Externalizing	0.0214*** (0.0038)	0.0227*** (0.0047)	0.0093 (0.0078)	0.0252* (0.0132)	0.0109 (0.0155)
Percentage	-18.37	-27.85	-4.78	-10.80	-4.86
SDQ Internalizing	-0.0002 (0.0012)	0.0004 (0.0008)	-0.0037 (0.0048)	0.0057 (0.0055)	-0.0139 (0.0135)
Percentage	0.17	-0.49	1.90	-2.44	6.20
Coefficients					
SDQ Externalizing	0.0150* (0.0089)	0.0208** (0.0090)	-0.0276 (0.0189)	0.0293 (0.0290)	-0.0189 (0.0464)
Percentage	-12.88	-25.52	14.19	-12.55	8.43
SDQ Internalizing	0.0074 (0.0091)	0.0034 (0.0102)	0.0189 (0.0207)	0.0029 (0.0246)	0.0889** (0.0396)
Percentage	-6.35	-4.17	-9.72	-1.24	-39.65
<i>N</i>	4688	3132	845	882	279

Standard errors in parentheses.

Standard errors are clustered at the school level.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

a few percentages of the observed reading gap (for instance for the subgroup 'other families' 4 percent of the gender gap in Reading are estimated to be due to the endowment component. For young mothers, the estimated percentage is higher, but here the component is insignificant). For Math, differences in behavioral problems are estimated to be relatively larger. The sign of the estimated components are positive for the SDQ components while the observed gaps in Math are negative, i.e. the differences in 'endowment of behavioral problems' are in favor of girls' grades in Math. But since the observed Math gap is negative, the unexplained gap is actually larger than the observed gap, according to the estimations.

In most cases, the SDQ components are insignificant for the low-resource families (and the absolute size of the coefficients are also relatively small, i.e. the insignificance is generally not because of fewer observations in these sub-groups).

Thus, the overall conclusion from this section is that only a minor proportion of the 'girl gap' in Reading can be related to differences in abnormal or borderline behavioral problems between boys and girls and there is no tendency that behavioral problems account for a larger proportion of the gender gap in Reading in low-resource families. The large and positive gap in Reading grades between girls and boys is mainly due to a 'girl-gap' in normal families. For Math, our results indicate that the SDQ endowment components slightly favor girls' grades, mostly in 'normal families'. For low-income families, we find a very negative coefficient of the girl indicator, i.e. girls in low-income families seem to get much lower grades than boys from low-income families, irrespective of behavioral endowments.

## 7 Conclusion

The results in this study document that behavioral problems are significantly related to lower exit exam grades in Reading and Math and a lower probability of taking the 9th grade exit exam and being enrolled at high school. About 11 percent of the boys and 6 percent of the girls in our sample have abnormal or borderline externalizing behavioral problems. For internalizing problems, girls and boys have about the same average scores, and 14 percent of the children are categorized as having abnormal or borderline internalizing problems.

The estimations show significantly and large negative coefficients of the externalizing behavioral indicators. Boys with abnormal externalizing behavior are estimated to have 3.6 percentage points lower chance of taking the 9th grade exit exam, and their score in Reading is estimated to be 0.36 standard deviations lower than children without behavioral problems.

For Math this figure is 0.57 of a standard deviation. According to our estimations, the school outcomes for girls with abnormal externalizing behavior are not significantly different from those of boys with the same behavioral problems. For borderline externalizing problems and for internalizing problems, the estimated main coefficients are numerically smaller but still significantly negative in most cases, and for these behavioral categories girls tend to have less negative coefficients.

We document that measurement of behavioral problems depends on whether it is the parents or the teachers who report the problems and this evidence has an effect on the estimated relationship between gender, behavioral problems and school outcomes. The negative outcome effects of behavioral problems are numerically smaller and the gender differences less significant when teachers' Strength and Difficulties Questionnaire (SDQ) scores are applied in the estimation. This may reflect that teachers who are aware of behavioral problems try to compensate for these problems.

Splitting the sample into subgroups of low social resource families, we find that children from low-resource families have lower school outcomes. The gender gap in Reading is observed to be largest in the subgroup of families who are not categorized as low-resource families and smallest in the subgroup of families with a young mother. Thus, we do not find that boys with behavioral problems from low-resource families seem to be more vulnerable than girls coming from low-resource families.

A decomposition of the estimations indicates that most of the gender differences in Reading and Math cannot be related to gender differences in behavioral problems. The overall gender gap in Reading seems mainly to be the result of gender differences between children without behavioral problems living in 'normal families', i.e. families which are not categorized as low-resource families. For Math, the higher overall grades for boys compared to girls seem mainly to be the result of very low Math grades for girls without behavioral problems from low-income families.

## References

- [1] Angrist, J. and J-S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press.
- [2] Bertrand, M. and J. Pan. 2011. "The Trouble with the Boys: Social influences and the Gender Gap in Disruptive Behavior". Working paper.
- [3] Carneiro, P. and J. Heckman. 2003. Human Capital Policy, in *Inequality in America: What Role for Human Capital Policies*, edited by James Heckman and Alan Krueger, MIT Press.
- [4] Datta Gupta, N., M. Lausten and D. Pozzoli. 2012. "Does Mother Know Best? Parental Discrepancies in Assessing Child Functioning", IZA Working Paper no. 6962.
- [5] Ermisch, J. 2008. "Origins of social immobility and inequality: parenting and early child development", *National Institute Economic Review*, 205, pp. 62-71.
- [6] Fortin, N., P. Oreopoulos and S. Phipps. 2011. "Leaving Boys Behind: Gender Disparities in High Academic Achievement", working paper.
- [7] Gaub, M., and C. L. Carlson. 1997. "Gender differences in ADHD: A meta-analysis and critical review", *Journal of the American Academy of Child and Adolescent*, 36(8): 1036-45.
- [8] Gillberg, C., E. Heiervang, C. Obel, M. Posserud and A. Ullebø. 2012. "Prevalence of the ADHD phenotype in 7- to 9-year-old children: effects of informant, gender and non-participation". *Social Psychiatry and Psychiatric Epidemiology* 47: 763-769.
- [9] Goldin, C., L. F. Katz and I. Kuziemko. 2006. "The Homecoming of American College Women: The Reversal of the College Gender Gap," *Journal of Economic Perspectives*, 20, pp. 133-156.
- [10] Goodman, R. 1997. "The Strengths and Difficulties Questionnaire: A Research Note", *Journal of Child Psychology and Psychiatry* 38, 581-586.
- [11] Goodman A, D. L. Lamping and G. B. Ploubidis. 2010. "When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the Strengths and Difficulties Questionnaire (SDQ): data from British parents, teachers and children." *Journal of Abnormal Child Psychology*, 38, 1179-1191.



- [12] Heckman, J. 2008. "Schools, Skills, and Synapses". *Economic Inquiry*, Vol. 46, No. 3, p. 289-324.
- [13] Heckman, J. J., L. Malofeeva, R. R. Pinto, P. Savelyev and A. Yavitz. 2008. "The Impact of the Perry Preschool Program on Noncognitive Skills of Participants." Unpublished manuscript, University of Chicago, Department of Economics
- [14] Heckman, J. and Y. Rubinstein. 2001. "The Importance of Noncognitive Skills: Lessons from the GED Testing Program," *American Economic Review*, American Economic Association, vol. 91(2), pages 145-149, May.
- [15] Heckman, J., J. Stixrud and S. Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics* 24(3):411–82.
- [16] Jacob, B. 2002. "Where the Boys Aren't: Non-cognitive Skills, Returns to School and the Gender Gap in Higher Education", *Economics of Education Review* 21, 589-598.
- [17] Jann, B. 2008. "The Blinder-Oaxaca decomposition for linear regression models". *The Stata Journal* 8(4): 453-479.
- [18] Mehlbye, J. 2008. Specialundervisningsselevers skolegang og tiden efter. AKF Working paper, 2008(2), AKF.
- [19] Miller, A. R. 2009. Motherhood Delay and the Human Capital of the Next Generation, *American Economic Review*, 99(2), 154{58.
- [20] Ministry of Finance. 2010. Budgetredegørelse 2010, Copenhagen.
- [21] Niederle, M. and L. Vesterlund. 2010. "Explaining the Gender Gap in Math Test Scores: The Role of Competition". *Journal of Economic Perspectives*, Vol 24, No. 2, pp.129-144.
- [22] Oaxaca, R. 1973. "Male-Female Wage Differentials in Urban Labor Markets". *International Economic Review* 14: 693–709.
- [23] OECD. 2010. Education at a glance 2010, OECD, Paris

# A Appendix A

Table 12: Description of variables

Variable name	Description	Source
<b>Basic characteristics</b>		
Boy	Boy (0/1)	Registers
Native	Native Dane (0/1)	Registers
Birth weight (g)	The child's birth weight in grams	Registers
Length of pregnancy	Length of pregnancy in weeks	Registers
Born prematurely	The child is born before the 37th gestational week (0/1)	Registers
Born extremely prematurely	The child is born before the 28th gestational week (0/1)	Registers
Complications at delivery	Complications during delivery of the child based on an APGAR score of 7 or above. (0/1)	Registers
Year of birth: 1990	The child is born in 1990 (0/1)	Registers
Year of birth: 1991	The child is born in 1991 (0/1)	Registers
Year of birth: 1992	The child is born in 1992 (0/1)	Registers
Number of younger siblings	Number of younger siblings before the child is 7 years old, including half-siblings.	Registers
Number of older siblings	Number of older siblings before the child is 7 years old, including half-siblings.	Registers
Younger siblings	The child has younger siblings before the age of 7 (0/1)	Registers
Older siblings	The child has older siblings measured before the age of 7 (0/1)	Registers
Psychological diagnosis	The child is diagnosed with a mental or behavioral disorder before the age of 7. (0/1)	Registers
Cardiovascular medicine	The child is prescribed cardiovascular medicine before turning 7 years old. (0/1)	Registers
Antidepressant medicine	The child is prescribed antidepressant medicine before turning 7 years old. (0/1)	Registers
Finished 9th grade in 2005	Completion of 9th grade in 2005 (0/1)	Registers
Finished 9th grade in 2006	Completion of 9th grade in 2006 (0/1)	Registers
Finished 9th grade in 2007	Completion of 9th grade in 2007 (0/1)	Registers
Finished 9th grade in 2008	Completion of 9th grade in 2008 (0/1)	Registers
Finished 9th grade in 2009	Completion of 9th grade in 2009 (0/1)	Registers
Finished 9th grade in 2010	Completion of 9th grade in 2010 (0/1)	Registers
Moved to another municipality	Moved to another municipality before the age of 7 (0/1)	Registers
Number of moves between municipalities	Number of moves between municipalities before the age of 7.	Registers
Not registered day care	Day care is not registered (0/1)	Registers
Centralized day care	The child is enrolled in centerbased day care at the age of 4 (0/1)	Registers
No outside home day care	The child is taken care of in the home by either parents or grandparents at the age of 4 (0/1)	Registers
Private day care	The child is enrolled in private family day care at the age of 4	Registers

Table 12: Description of variables - continued from previous page

Variable name	Description	Source
<b>Mother's characteristics</b>		
Committed crime	Crime committed before the child is 7 years old(0/1).	Registers
<b>Mother's degree of employment</b>		
Degree of employment, 4 years old	Degree of year employed when the child is 4 years old. = 1000 if unemployed the whole year.	Registers
Degree of employment, 5 years old	Degree of year employed when the child is 4 years old. = 1000 if unemployed the whole year.	Registers
Degree of employment, 6 years old	Degree of year employed when the child is 4 years old. = 1000 if unemployed the whole year.	Registers
<b>Mother's employment sector</b>		
Private sector	Works in the private sector (0/1)	Registers
City authorities	Works in the city authorities (0/1)	Registers
Public sector	Works in the public sector (0/1)	Registers
<b>Mother's employment</b>		
Full-time employed	Is full-time employed (>25 h/week) (0/1)	Registers
Part-time employed	Is part-time employed (<25 h/week) (0/1)	Registers
Other	Is in other form of employment (0/1)	Registers
<b>Mother's occupation</b>		
Employer	Is employer when the child is 6 years old (0/1)	Registers
Manager	Is employed at management level when the child is 6 years old (0/1)	Registers
White-collar worker	Is employed at medium level as a white-collar worker when the child is 6 years old (0/1)	Registers
Unskilled worker	Is employed as an unskilled worker at lower level when the child is 6 years old (0/1)	Registers
Not in the labor force	Not in the labor force	Registers
<b>Mother's highest education</b>		
No education	Compulsory school, corresponding to 9 years of education when the child is 6 years old (0/1)	Registers
High school education	High school, corresponding to 12 years of education when the child is 6 years old (0/1)	Registers
Vocational education	Vocational education, corresponding to 12 years of education when the child is 6 years old (0/1)	Registers
Short cycle education	Short cycle education, corresponding to 14 years of education when the child is 6 years old (0/1)	Registers
Long cycle education	Long cycle education, corresponding to 17 years of education when the child is 6 years old (0/1)	Registers
On-going education	On-going education. If enrolled in education program when the child is 6 years old (0/1)	Registers

Table 12: Description of variables - continued from previous page

Variable name	Description	Source
<b>Mother's health</b>		
Psychological diagnosis	Diagnosed with a mental or behavioral disorder before the child is 6 years old (0/1)	Registers
Cardiovascular medicine	Prescribed cardiovascular medicine before the child is 6 years old (0/1)	Registers
Antidepressant medicine	Prescribed antidepressant medicine before the child is 6 years old (0/1)	Registers
<b>Father's characteristics</b>		
Committed crime	Crime committed before the child is 7 years old(0/1).	Registers
<b>Father's degree of employment</b>		
Degree of employment, 4 years old	Degree of year employed when the child is 4 years old. = 1000 if unemployed the whole year.	Registers
Degree of employment, 5 years old	Degree of year employed when the child is 4 years old. = 1000 if unemployed the whole year.	Registers
Degree of employment, 6 years old	Degree of year employed when the child is 4 years old. = 1000 if unemployed the whole year.	Registers
<b>Father's employment sector</b>		
Private sector	Works in the private sector (0/1)	Registers
City authorities	Works in the city authorities (0/1)	Registers
Public sector	Works in the public sector (0/1)	Registers
<b>Father's employment</b>		
Full-time employed	Is full-time employed (>25 h/week) (0/1)	Registers
Part-time employed	Is part-time employed (<25 h/week) (0/1)	Registers
Other	Is in other form of employment (0/1)	Registers
<b>Father's occupation</b>		
Employer	Is employer when the child is 6 years old (0/1)	Registers
Manager	Is employed at management level when the child is 6 years old (0/1)	Registers
White-collar worker	Is employed at medium level as a white-collar worker when the child is 6 years old (0/1)	Registers
Unskilled worker	Is employed as an unskilled worker at lower level when the child is 6 years old (0/1)	Registers
Not in the labor force	Not in the labor force	Registers

Table 12: Description of variables - continued from previous page

Variable name	Description	Source
<b>Father's highest education</b>		
No education	Compulsory school, corresponding to 9 years of education when the child is 6 years old (0/1)	Registers
High school education	High school, corresponding to 12 years of education when the child is 6 years old (0/1)	Registers
Vocational education	Vocational education, corresponding to 12 years of education when the child is 6 years old (0/1)	Registers
Short cycle education	Short cycle education, corresponding to 14 years of education when the child is 6 years old (0/1)	Registers
Long cycle education	Long cycle education, corresponding to 17 years of education when the child is 6 years old (0/1)	Registers
On-going education	On-going education. If enrolled in education program when the child is 6 years old (0/1)	Registers
<b>Father's health</b>		
Psychological diagnosis	Diagnosed with a mental or behavioral disorder before the child is 6 years old (0/1)	Registers
Cardiovascular medicine	Prescribed cardiovascular medicine before the child is 6 years old (0/1)	Registers
Antidepressant medicine	Prescribed antidepressant medicine before the child is 6 years old (0/1)	Registers
<b>Family background</b>		
Mother's age at birth	The mother's age when she gave birth to the child	Registers
Age difference between father and mother	The age difference between the father and the mother at the time of birth	Registers
Young mother (age < 24)	The mother is younger than 24 years old (0/1)	Registers
Non-intact family	The biological parents are not cohabiting up until the child is 6 years old (0/1)	Registers
Average household income	Average household income measured at time of birth.	Registers
Low-income household (< 50 % of the median)	Low-income household, having an average income below 50% of the median income (0/1)	Registers

Table 12: Description of variables - continued from previous page

Variable name	Description	Source
<b>Behaviour</b>		
Abnormal Externalizing behavior	Showing behavioral problems in terms of abnormal externalizing behavior based on the SDQ score (0/1)	2002 Questionnaire
Borderline Externalizing behavior	Showing behavioral problems in terms of borderline externalizing behavior based on the SDQ score (0/1)	2002 Questionnaire
Abnormal Internalizing behavior	Showing behavioral problems in terms of abnormal internalizing behavior based on the SDQ score (0/1)	2002 Questionnaire
Borderline Internalizing behavior	Showing behavioral problems in terms of borderline internalizing behavior based on the SDQ score (0/1)	2002 Questionnaire
<b>Outcomes</b>		
Taking 9th grade exit exam	Taking the 9th grade exit exam (0/1)	Registers
Reading (exit exam)	Reading exam standardized test score	Registers
Math (exit exam)	Math exam standardized test score	Registers
High school enrollment	Enrolled in high school before turning 19 years old	Registers

Table 13: Descriptive statistics of covariates, behavior and school outcomes

Variable	Girls		Boys		Diff
	Mean	Std.dev.	Mean	Std.dev.	
Native	0.97	0.16	0.98	0.14	
Birthweight	3426.66	573.08	3548.66	599.09	***
Length of pregnancy	39.53	1.82	39.51	1.88	
Born prematurely	0.09	0.29	0.10	0.30	
Born extremely prematurely	0.00	0.03	0.00	0.04	
Complications at delivery	0.01	0.10	0.01	0.11	
Year: 1990	0.40	0.49	0.40	0.49	
Year: 1991	0.39	0.49	0.41	0.49	
Year: 1992	0.20	0.40	0.19	0.39	
No. older siblings	0.85	1.01	0.86	1.00	
No. younger siblings	0.63	0.68	0.64	0.68	
Older siblings	0.55	0.50	0.56	0.50	
Younger siblings	0.53	0.50	0.54	0.50	
Psychological diagnosis	0.00	0.06	0.01	0.10	***
Cardiovascular medicine	0.00	0.06	0.00	0.06	
Antidepressant medicine	0.01	0.11	0.01	0.10	
Finished 9th grade in 2005	0.01	0.11	0.00	0.05	***
Finished 9th grade in 2006	0.36	0.48	0.33	0.47	**
Finished 9th grade in 2007	0.39	0.49	0.40	0.49	
Finished 9th grade in 2008	0.23	0.42	0.25	0.43	
Finished 9th grade in 2009	0.00	0.05	0.01	0.10	***
Finished 9th grade in 2010	0.00	0.02	0.00	0.02	
Moved to another municipality	0.23	0.42	0.23	0.42	
No. of moves between municipalities	0.30	0.60	0.29	0.58	
Not registered day care	0.02	0.16	0.02	0.15	
Centralized day care	0.90	0.30	0.90	0.29	
No outside home day care	0.04	0.19	0.03	0.18	
Private day care	0.04	0.19	0.04	0.19	
Mother's age at birth	29.76	4.58	29.73	4.47	
Age difference between father and mother	2.41	4.23	2.37	4.35	
Young mother (age < 24)	0.07	0.25	0.07	0.25	
Non-intact family	0.21	0.41	0.21	0.41	
Average household income	178741.05	98256.22	179826.73	96787.37	
Low-income household (< 50 % of median)	0.19	0.39	0.18	0.39	

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$



Table 13: Descriptive statistics of covariates, behavior and school outcomes - continued from previous page

Variable	Girls		Boys		Diff
	Mean	Std.dev.	Mean	Std.dev.	
Mother degree of year employed (4 years old)	101.42	219.15	100.45	218.52	
Mother degree of year employed (5 years old)	95.46	209.61	87.42	203.11	
Mother degree of year employed (6 years old)	88.18	201.70	79.91	193.37	*
Mother employed in private sector	0.29	0.46	0.30	0.46	
Mother employed in the city authorities	0.65	0.48	0.63	0.48	*
Mother employed in public sector	0.06	0.24	0.07	0.26	**
Mother full-time employed	0.83	0.37	0.84	0.37	
Mother part-time employed	0.03	0.18	0.04	0.19	
Mother other type of employment	0.14	0.34	0.13	0.33	
Mother employer	0.01	0.12	0.01	0.11	
Mother manager	0.30	0.46	0.29	0.45	
Mother white-collar worker	0.31	0.46	0.32	0.47	
Mother unskilled worker	0.15	0.36	0.15	0.36	
Mother unemployed	0.11	0.31	0.11	0.32	
Mother not in labor force	0.11	0.31	0.10	0.30	
Mother no education	0.19	0.39	0.18	0.39	
Mother long cycle education	0.08	0.28	0.08	0.27	
Mother high school education	0.15	0.36	0.14	0.35	
Mother vocational education	0.28	0.45	0.29	0.45	
Mother short cycle education	0.30	0.46	0.31	0.46	
Mother on-going education	0.10	0.30	0.09	0.28	
Mother psychological diagnosis	0.02	0.14	0.03	0.16	
Mother cardiovascular medicine	0.14	0.35	0.15	0.36	
Mother antidepressant medicine	0.23	0.42	0.23	0.42	
Father degree of year employed (4 years old)	65.84	191.11	63.11	189.70	
Father degree of year employed (5 years old)	55.65	175.23	49.78	166.62	
Father degree of year employed (6 years old)	46.83	156.90	44.86	153.68	
Father employed in private sector	0.56	0.50	0.55	0.50	
Father employed in the city authorities	0.36	0.48	0.35	0.48	
Father employed in public sector	0.09	0.28	0.10	0.29	
Father full-time employed	0.86	0.35	0.87	0.34	
Father part-time employed	0.01	0.11	0.01	0.11	
Father other type of employment	0.13	0.33	0.12	0.33	

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table 13: Descriptive statistics of covariates, behavior and school outcomes - continued from previous page

Variable	Girls		Boys		Diff
	Mean	Std.dev.	Mean	Std.dev.	
Father employer	0.05	0.22	0.05	0.23	
Father manager	0.34	0.47	0.35	0.48	
Father white-collar worker	0.14	0.35	0.14	0.35	
Father unskilled worker	0.17	0.38	0.16	0.37	
Father unemployed	0.07	0.25	0.06	0.24	
Father not in labor force	0.08	0.27	0.08	0.27	
Father no education	0.19	0.39	0.17	0.38	
Father long cycle education	0.16	0.36	0.16	0.37	
Father high school education	0.12	0.33	0.11	0.31	*
Father vocational education	0.34	0.47	0.35	0.48	
Father short education	0.19	0.40	0.20	0.40	
Father on-going education	0.11	0.31	0.10	0.30	
Father psychological diagnosis	0.03	0.16	0.02	0.15	
Father cardiovascular medicine	0.09	0.29	0.08	0.28	
Father antidepressant medicine	0.16	0.36	0.16	0.37	
Fraction of boys in class	0.29	0.29	0.33	0.28	***
Fraction of boys in class squared	0.14	0.14	0.19	0.18	***
Fraction of boys in class missing	0.41	0.41	0.39	0.49	*
Class moves before 2002	0.40	0.40	0.37	0.91	
Abnormal Externalizing Behavior (Parent eval.)	0.02	0.16	0.05	0.22	***
Borderline Externalizing Behavior (Parent eval.)	0.04	0.19	0.06	0.24	***
Abnormal Internalizing Behavior (Parent eval.)	0.07	0.26	0.08	0.27	
Borderline Internalizing Behavior (Parent eval.)	0.08	0.26	0.06	0.24	**
Abnormal Externalizing Behavior (Teacher eval.)	0.02	0.15	0.09	0.28	***
Borderline Externalizing Behavior (Teacher eval.)	0.02	0.13	0.06	0.23	***
Abnormal Internalizing Behavior (Teacher eval.)	0.07	0.26	0.07	0.25	
Borderline Internalizing Behavior (Teacher eval.)	0.05	0.22	0.05	0.22	

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table 14: Strength and Difficulties Questionnaire (SDQ)

Measures	Description
<b>Externalizing problems</b>	<p>Measures acting out behavior.</p> <p>10 questions (or statements) rated as "certainly true", "somewhat true" or "not true".</p> <p>The questions include:</p> <ul style="list-style-type: none"> <li>- Restless, overactive, cannot stay still for long</li> <li>- Often has temper tantrums or hot tempers</li> <li>- Generally obedient, usually does what adults request</li> <li>- Constantly fidgeting or squirming</li> <li>- Often fights with other children or bullies them</li> <li>- Easily distracted, concentration wanders</li> <li>- Often lies or cheats</li> <li>- Thinks things out before acting</li> <li>- Steals from home, school or elsewhere</li> <li>- Sees tasks through to the end, good attention span</li> </ul>
<b>Internalizing problems</b>	<p>Measures the apparent internal problems.</p> <p>10 questions (or statements) rated as "certainly true", "somewhat true" or "not true".</p> <p>The questions include:</p> <ul style="list-style-type: none"> <li>- Often complains of headaches, stomach-aches or sickness</li> <li>- Rather solitary, tends to play alone</li> <li>- Many worries, often seems worried</li> <li>- Has at least one good friend</li> <li>- Often unhappy, down-hearted or tearful</li> <li>- Generally liked by other children</li> <li>- Nervous or clingy in new situations, easily loses confidence</li> <li>- Picked on or bullied by other children</li> <li>- Gets on better with adults than with other children</li> <li>- Many fears, easily scared</li> </ul>

Table 15: Oaxaca Decomposition for samples of students from low-resource families: Further education (High school and vocational education)

	All	Other	Low income	Non-intact	Young mother
<b>Further education</b>					
Difference: Girls-Boys	0.0013 (0.0065)	0.0095 (0.0077)	-0.0311 (0.0239)	-0.0311 (0.0197)	-0.0053 (0.0314)
Percentage	100.00	100.00	100.00	100.00	100.00
<b>Endowments</b>					
SDQ Externalizing	0.0014 (0.0011)	0.0020* (0.0012)	-0.0000 (0.0017)	0.0002 (0.0031)	0.0014 (0.0050)
Percentage	107.69	21.05	0.00	-0.64	-26.42
SDQ Internalizing	-0.0002 (0.0004)	0.0000 (0.0004)	0.0013 (0.0017)	-0.0001 (0.0023)	-0.0030 (0.0058)
Percentage	-15.38	0.00	-4.18	0.32	56.60
<b>Coefficients</b>					
SDQ Externalizing	-0.0050 (0.0032)	0.0033 (0.0027)	-0.0269** (0.0125)	-0.0190* (0.0110)	-0.0149 (0.0176)
Percentage	-384.62	34.74	86.50	61.09	281.13
SDQ Internalizing	0.0045 (0.0028)	0.0037 (0.0035)	-0.0016 (0.0086)	0.0068 (0.0136)	0.0046 (0.0108)
Percentage	346.15	38.95	5.14	-21.86	-86.79
<i>N</i>	4753	3168	858	901	281

Standard errors in parentheses.

Standard errors are clustered at the school level.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$