Erasmus School of Economics
Master Thesis Policy Economics

Better than all the Rest: the Effect of Ordinal Rank on Educational Performance during Primary School

Yrla van de Ven
Student number: 386860

Supervisor: Prof. dr. Dinand Webbink
Second assessor: Dr. Matthijs Oosterveen

Date final version: 6 May 2019

Abstract

In this paper we study the effect of a student's ordinal rank in grade 2 of primary school, at the age of 5, on educational performance in grade 8. We exploit the idiosyncratic variation in the ability distributions that exists both within and across schools and cohorts to estimate a causal effect. Using longitudinal data from primary schools in the Netherlands we show that, after controlling for ability, ordinal rank has a positive effect on both test scores and school advice at a very young age. A one standard deviation increase in rank for the subject Dutch increases test scores with 0.083 standard deviations and school advice with 0.123 standard deviations. The results are stronger for mathematics: a one standard deviation increase in rank increases test scores with 0.226 standard deviations and school advice with 0.256 standard deviations. We find support for three possible mechanisms behind these effects: students with a higher ordinal rank have more confidence, have a better work attitude and receive more extra material from their teacher.

1. Introduction

Teachers and parents often believe that placing a student in a classroom with better performing peers has a beneficial effect on the student. The idea is that students learn more if they interact with smarter students or students who come from a better environment. In this paper we study a different type of peer effect, the effect of having a high ordinal rank within a classroom. This peer effect moves in the opposite direction of traditional peer effects in education: the ordinal rank of a student improves when we remove highly able students from the classroom.

We investigate whether performance during primary school improves by being at the top of the class, rather than having a lower ordinal rank. Using longitudinal data from primary schools in the Netherlands we show that, after controlling for ability, ordinal rank may indeed positively affect test performance early in life. We find that a one standard deviation increase in ordinal rank for the subject Dutch at the age of 5 significantly increases test scores for that same subject at the age of 12 by 0.083 standard deviations. The effect of ordinal rank in mathematics is even larger: a one standard deviation increase in rank for mathematics increases test scores for that subject with 0.226 standard deviations. The ordinal rank of a student in primary school may have long-lasting consequences, since we also find a positive effect of ordinal rank on school advice. High schools admit students based on the school advice given by the teacher in grade 8 of primary school. Only after finishing the highest level of high school, Dutch students can apply for university. A one standard deviation increase in rank for the subject Dutch significantly improves school advice by 0.123 standard deviations. Again, the effects are larger for the subject mathematics: a one standard deviation increase in rank for mathematics significantly improves school advice by 0.256 standard deviations.

These findings are an important addition to the existing literature on peer effects. It shows that the presumption that having better peers leads to better performance is not always true. The positive effect of having better peers might in some cases be outweighed by negative rank effects and vice versa. This finding could also be interesting for policy makers who aim to improve the position of weak students. Placing weak students in a good classroom, where they will have a low ordinal rank, might not have the desired effect on their performance. In addition, we add to the literature on ordinal rank effects by studying the effect of rank at a very young age. The results suggest that students are already aware of their relative position at the age of 5. Furthermore, we provide additional evidence for the mechanisms that could explain the ordinal rank effects found in this paper and others. We find that having a higher ordinal rank improves the confidence level of students, motivates them to exert more effort and increases teacher investment.

From our paper one cannot conclude, however, that we should deliberately place students in classrooms with bad performing peers. Positive peer effects might, in some cases, be stronger than positive rank effects. In addition, one should take other inputs into account, such as having a good teacher. Finally, one should also note that improving the ordinal rank of one student, by changing classroom compositions, might lower the rank of other students.

This thesis is organized as follows. In section 2, we provide an overview of the existing literature on peer effects in general and rank effects in specific. In section 3, we describe the longitudinal data used for this study and provide descriptive statistics. In section 4, we describe the identification strategy and the econometric model that we will use for our analysis. We describe our main results in section 5, followed by several robustness checks, and we investigate possible explanations for the effects in section 6. Finally, we conclude this paper in section 7 by discussing the implications of our results.

2. Literature review

Many recent studies stress the importance of peer effects in educational performance, in addition to traditional inputs such as class size, teacher quality and parental involvement. The general notion is that being in a classroom with high performing peers positively influences a student's own performance, while having bad peers has a negative impact. This idea is supported by a large collection of empirical papers showing that, on average, students benefit from being surrounded by higher able peers (Hoxby, 2000; Sacerdote, 2001; Zimmerman, 2003; Hanushek et al., 2003; Whitmore, 2005; Carrell, 2009; Ammermueller and Pischke, 2009; Bifulco et al., 2011). These papers cover students from primary and secondary school, college and university. Other papers have shown that peer effects are often not linear, and that (larger) effects can be found when including this non-linearity (Burke & Sass, 2013; Gibbons & Telhaj, 2012; Whitmore, 2005). The proportion of students at the tails of the ability distribution, the highest and lowest achievers, may matter more than the average peer quality, as shown by Lavy et al. (2012a) and Lavy et al. (2012b).

A new branch within the peer effects literature challenges the general belief that having good peers improves educational performance by looking at the effect of a student's ordinal rank within a classroom. The ordinal rank of students is determined by the distribution of ability within a classroom. Adding more high achievers to a classroom reduces the ordinal rank of an average student, while removing higher able students and adding lower able students improves the average student's ordinal rank.

To understand why ordinal rank in a classroom could impact performance, first consider a perfect world. In a perfect world where students, parents and teachers have perfect knowledge about their own ability, that of their offspring or their students and act rational, ordinal rank should not have any effect on performance. In the real world, however, people have imperfect knowledge and face uncertainty. To deal with this, people tend to compare themselves to others and use this comparison to evaluate their opinions and ability (Festinger, 1954). When doing so, they often rely on cognitive shortcuts (Tversky & Kahneman, 1974).

One of these shortcuts is to use information about ordinal ranks instead of cardinal information. Given that students might use cognitive shortcuts to deal with uncertainty, they might evaluate their own ability based on their ordinal rank within a classroom, instead of their absolute performance and the magnitude of the difference between them and their classmates. If that is true, having a relatively high ordinal rank then leads to a relatively high perceived ability. Students might even overestimate their ability based on their position in a classroom. A high perceived ability may result in a higher level of confidence, which may in turn positively impact test performance (Mavis, 2001). In that way, ordinal rank may indirectly influence performance. In the psychological literature, this has been described as the Big Fish Little Pond effect (Marsh, 1987).

Alternatively, rank effects could be explained by the fact that teachers and parents also have imperfect knowledge about ability and use the same cognitive shortcut as students. Instead of estimating the ability of their offspring or students based on absolute performance, they update their beliefs about the ability of that student based on ordinal rank. Consequently, parents and teachers may invest more time and money in students with a high ordinal rank, because they expect these students to deliver the highest returns. Or, if the rank effect is negative, parents and teachers might invest more in children with a lower ordinal rank in order to provide these students with the opportunity to catch up.

Several empirical studies find that ordinal rank has a positive effect on educational performance and other short- and long-run outcomes. Murphy & Weinhardt (2018) find that having had a one standard deviation higher rank in a primary school subject increases secondary school test scores for that subject by 0.05 standard deviations. They base their research on administrative data covering 2.25 million students in England from primary school to the end of secondary school. Elsner and Isphording (2017), using survey data from the U.S., find that a student's ordinal rank has a significant and positive effect on educational outcomes later in life. A higher rank significantly increases the chance that a student finishes high school, attends college and completes a 4-year college degree. Denning et al. (2018), using administrative data from public school students in Texas, show that students with a higher rank in third grade, conditional on ability, have higher test scores in following years, are more likely to graduate high school, enroll in college and have higher earnings 19 years later.

In addition to improving educational performance, ordinal rank may influence other outcomes later in life. Elsner and Isphording (2018), using longitudinal data from U.S. high schools, show that ordinal rank impacts the likelihood of engaging in risky behavior. They find that having a higher ordinal rank has a negative effect on the likelihood of smoking, drinking, having unprotected sex and engaging in physical fights.

To estimate a causal effect of ordinal rank, all these papers use a novel approach in which they exploit the idiosyncratic variation in test score distributions across schools and cohorts. They all rely on the assumption that after controlling for prior test scores, school, subject and cohort fixed effects and student characteristics, there would be no expected differences in later (academic) outcomes except those driven by rank.

This paper adds to the existing literature on peer effects and specifically rank effects by: 1) Estimating the short term effects of ordinal rank on educational performance very early in life, starting at the age of 5. There is, as far as we know, no paper in which ordinal rank effects at such a young age are studied. The results suggest that children are aware of their ordinal rank at a very young age and that the ability distribution within a classroom during the first years of primary school impacts their performance in later years. 2.) Estimating the effect of ordinal rank on school advice by the teacher, which is used by secondary schools to admit students. Choosing the right secondary school has long-lasting effects, since different levels of secondary school education give access to different levels of tertiary education. 3.)

Distinguishing between the effect of ordinal rank on low stakes and high stakes test scores. High stake tests are tests that have important consequences for the student who makes the test. The example that we use in our paper is the CITO end-of-primary test, which has consequences for the level of high school a student can attend, while the other tests in the PRIMA study are less important (low stake). We find that the effect of ordinal rank on test performance is similar for low and high stake tests.

3. Data and descriptive statistics

3.1 Data

This thesis uses data from a longitudinal study commissioned by the *Programmaraad voor het onderwijsonderzoek* (PROO). The study, titled *Primair onderwijs en speciaal onderwijs cohortonderzoeken* (PRIMA), consists of 6 waves of survey data and test results of students in Dutch primary schools. Data was collected during the school years 1994/1995 (PRIMA I), 1996/1997 (PRIMA II), 1998/1999 (PRIMA III), 2000/2001 (PRIMA IV), 2002/2003 (PRIMA V) and 2004/2005 (PRIMA VI). Each wave, data was collected from students in grade 2, 4, 6 and 8. See table 1 for an overview of the Dutch primary school system with the grades and corresponding ages. The dataset covers about 600-700 schools and 50.000-60.000 students per wave. The goal of the researchers who designed the study was to follow the same students and schools for as many years as possible, but this proved to be hard to fulfill.

Table 1: Primary school system in the Netherlands

Age	Grade	Test
Age 4-5	Grade 1	
Age 5-6	Grade 2	First PRIMA test
Age 6-7	Grade 3	
Age 7-8	Grade 4	Second PRIMA test
Age 8-9	Grade 5	
Age 9-10	Grade 6	Third PRIMA test
Age 10-11	Grade 7	
Age 11-12	Grade 8	Fourth PRIMA test & CITO end-of-primary test

Unfortunately, schools that participated in one wave were not always willing to participate in the following waves. In addition, some students switched schools or had to repeat a class during the study and could not be tracked in later years. Therefore, the total number of students that we can follow during their primary school career, from grade 2 until grade 8, is much lower than the total number of participating students.

Figure A.1 in the appendix shows the inflow and outflow of the students that participated in the study. Note that we included all students in the flow diagram, while in the end we can only use the students who started participating in grade 2 and still participated in the survey in grade 8. This means that students who were in grade 4, 6 or 8 in 1994 and students who were in grade 2, 4 and 6 in 2004 are naturally left out of the sample. Each year, the outflow of students who ceased to participate in the following year(s) is quite large. The large majority of these students were dropped from the sample because their entire classroom ceased to participate in the program. This includes students who were not followed in later years because they were already in grade 8. A much smaller group of students dropped out while (some of) their classmates *did* participate in the next wave. This may concern students who switched schools or had to repeat a class and could not be tracked in later years. This outflow of students could be problematic in case individual students that drop out differ from the students in the final sample. For example if students with a low rank are more likely to drop out. In the robustness section we will show in what way this selective attrition influences our results. The main conclusions remain unchanged if we control for selective attrition.

After dropping students who did not participate during their entire primary school career, we made a few other changes to the dataset. We dropped students within cohorts of less than 10 students within a school, since this situation is most likely not representative for the Dutch school environment. This concerns 285 students. Finally, some students could not be tracked because their student number changed and was not updated by the researchers, for example because schools merged. These students were also dropped. This concerns 1711 observations within cohort 1, 1039 in cohort 2 and 292 in cohort 3.

This leaves us with 2,849 students in cohort 1, 2,822 in cohort 2 and 3,645 in cohort 3, who were all followed for 6 years. Since not all the students in our sample participated in all of the tests, for example due to illness, our regression analysis is based on a sample of about 6000 to 8000 students.

3.2 Variables

The variables that we use for our analysis are the following: ordinal rank in percentages, PRIMA test scores in grade 2, 4, 6 and 8, CITO test scores in grade 8, school advice in grade 8, educational level of the parents, socio-ethnic background, school and cohort. In this section, we will explain each of the variables.

3.2.1 Ordinal rank

In the PRIMA study, the ordinal rank of students was not included so we calculated this based on PRIMA test scores for Dutch and mathematics. We calculated rank within a classroom before dropping any observations from our sample. The absolute rank of students in classrooms of different sizes is not comparable. It is much easier to have a high absolute rank in a small classroom than it is in a large classroom. Instead, we compute the percentile rank of students and use this in our study. This allows us to compare students from classrooms with different sizes. Whenever the words 'ordinal rank' are used in this paper, we refer to the ordinal rank expressed in percentages.

$$percentile\ rank = \frac{absolute\ rank\ -\ 1}{nr\ of\ students\ in\ classroom - 1}$$

3.2.2 PRIMA test scores

In grade 2, 4, 6 and 8, students who participated in the PRIMA study made a test, which allows us to track the performance of the students. These tests are externally marked and are not set to a curve at class, school or district level, so that they are comparable between schools. To increase comparability between cohorts, the researchers who collected the data transformed the raw test scores in each grade to ability scores that are roughly on the same scale for each grade. PRIMA test scores in grade 8 are used as an outcome variable in some of our models, while we include PRIMA test scores in grade 2, 4 or 6 as control variables.

3.2.3 CITO end-of-primary test scores

In grade 8, the majority of the students in our sample took the national CITO end-of-primary test, which is also implemented in schools who did not participate in the PRIMA cohort study. It is not mandatory for schools to let their students take the CITO test, and therefore not all of the schools who participated in PRIMA have implemented the CITO test in grade 8. The first cohort made the CITO test in 2000, the second cohort in 2002 and the third cohort in 2004. Table 2 provides a summary of the CITO scores in these years.

Table 2: Summary of the CITO scores for the three cohorts included in the study

Year	Mean CITO score Dutch	Standard deviation	Number of observations
All cohorts	60.588	19.040	6,614
2000	40.535	9.986	2,025
2002	71.866	14.413	2,027
2004	67.516	15.043	2,562
Year	Mean CITO score mathematics	Standard deviation	Number of observations
All cohorts	41.984	11.363	6,591
2000	42.429	11.388	2,026
2002	43.227	11.267	2,066
2004	40.596	11.278	2,499
Year	Mean CITO score all subjects	Standard deviation	Number of observations
All cohorts	533.801	11.676	6,713
2000	534.292	10.028	2,065
2002	533.884	14.463	2,093
2004	533.337	10.259	2,555

3.2.4 School advice

At the end of primary school, the teacher in grade 8 gives each student a school advice. Based on this school advice, secondary schools can select students. Figure 1 gives an overview of the Dutch secondary and tertiary educational system. The school advice corresponds with one of the 8 levels of education or a combination of two levels. The variable can take the value 1 up to 15, where 1 is the lowest level and 15 the highest. Table 3 shows the different values of school advice and their label. Getting a school advice that does not correspond with your true ability can have long term consequences for your education and career, since the level of secondary education you finish determines whether you can apply for university or college.

Figure 1: Overview of the education system in the Netherlands with different levels of secondary and tertiary education

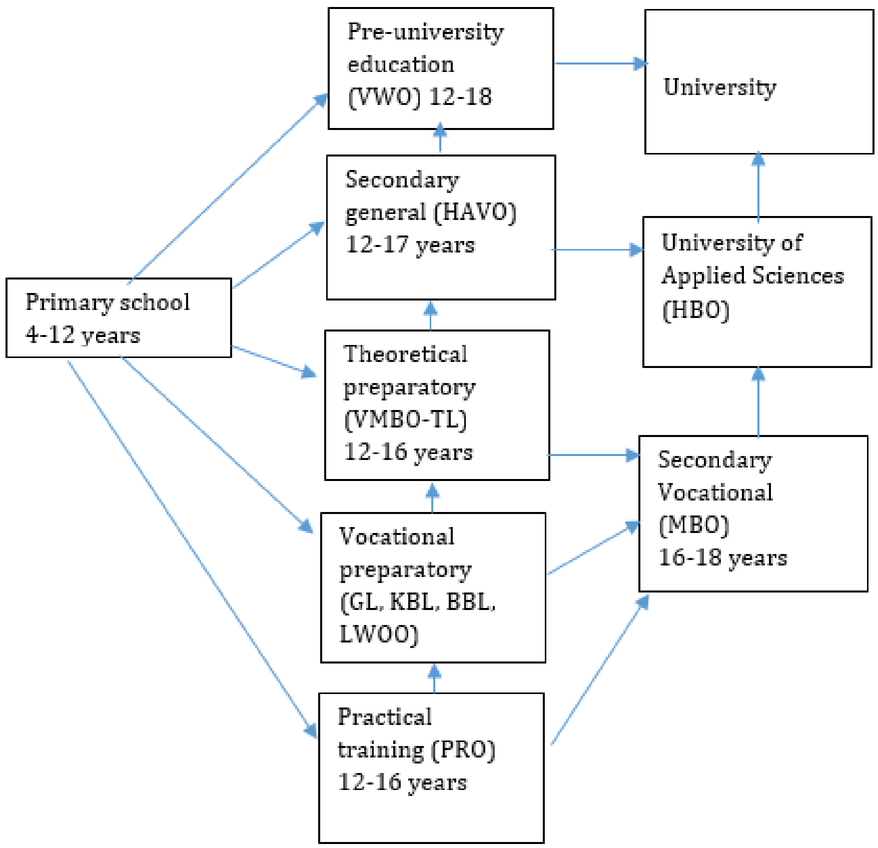


Table 3: The different values of the variable *school advice* and their label

Value	School advice
1	vmbo-pro
2	vmbo-pro/lwo
3	vmbo-lwo
4	vmbo-lwo/bbl
5	vmbo-bbl
6	vmbo-bbl/kbl
7	vmbo-kbl
8	vmbo-kbl/gl
9	vmbo-gl
10	vmbo-gl/tl
11	vmbo-tl
12	vmbo-tl/havo
13	havo
14	havo/vwo
15	vwo

3.2.5 Educational attainment of the father and mother

In our analysis we control for the educational level obtained by the father and mother of the student. The variable educational attainment father (educational attainment mother) can take 5 values: max. Primary education (1), max. Vocational preparatory (2), max. Secondary vocational (3), University or University of Applied Sciences (4) and unknown (5).

3.2.6 Socio-ethnic background

The variable socio-ethnic background allows us to control for a combination of the ethnic and economic background of the student. The variable can take 6 values: highest educational attainment of the parents is vocational preparatory and student is of Turkish/Moroccan descent (1), highest educational attainment of the parents is vocational preparatory and student is of non-Dutch non-Turkish/Moroccan descent (2), highest educational attainment of the parents is vocational preparatory and both parents are Dutch (3), highest educational attainment of the parents is secondary vocational (4), highest educational attainment of the parents is university or university of applied sciences (5) and unknown (6). The values 4, 5 and 6 are not based on the ethnic background of the students but only on the economic background.

3.2.7 School and cohort

We include both school and cohort fixed effects. The final sample consists of 3 cohorts and about 380 schools. The first cohort of students participated in the surveys PRIMA I, II, III and IV. The second cohort in PRIMA II, III, IV and V and the third cohort in PRIMA III, IV, V and VI.

3.2.8 Age

Finally, we control for the age of the student. The variable age was constructed by deducting the year of birth from the year in which the survey was conducted. For example, for the school year 1994-1995, we used the formula: $age = 1994 - year\ of\ birth$. In some cases, the year of birth was not reported or was unrealistically high or low. In these cases, we replaced the variable age by the value of age in the next wave minus 2. By including the variable age, we control for differences in the starting age of students as they enter primary school. Students who had to repeat a class during primary school, after grade 2, were not included in the study.

3.3 Descriptive statistics

Table A.2 in the appendix shows a summary of the most important descriptive statistics. Our sample is quite balanced when it comes to the amount of male and female students. The mean rank of students in our sample is a bit higher (0.525) than should be expected (0.500). Table A.3 shows the different outcomes for different demographic groups. In line with expectations, students with parents that are both not Dutch score lower on all tests. Students with one Dutch and one non-Dutch parent score higher for some tests and lower for others, but the sample size of this group is quite small. On average, girls score slightly lower for the subject mathematics, while boys score slightly lower for Dutch. As a result, the mean ordinal rank of girls is lower than that of boys for mathematics and higher for Dutch. There is a correlation between both test scores and ordinal rank and socio-ethnic background. Students with a socio-ethnic background that is higher than the average value of 3.5 score higher on all tests and have a higher ordinal rank than students with a value below 3.5. The same holds for school advice: the mean school advice is higher for students with a social-ethnical background above average. We will control for socio-ethnic background and gender in our analysis, as well as age and the education level of the father and mother of the student.

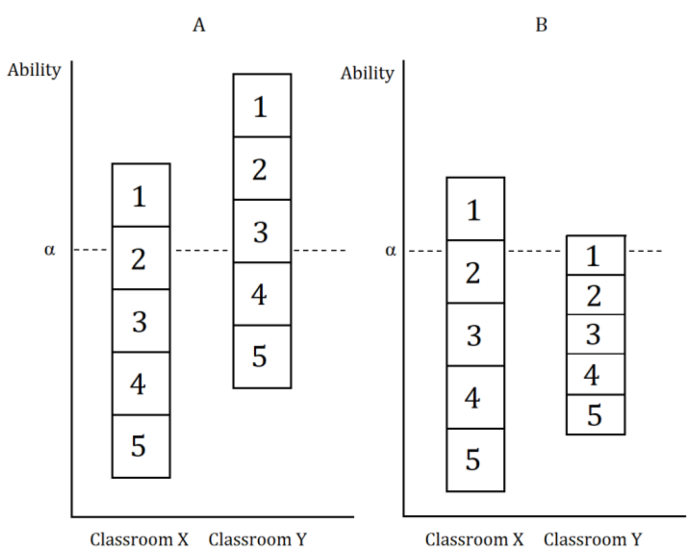
4. Methodology

In this section, we will first explain the identification strategy that allows us to estimate the causal effect of ordinal rank on educational performance. Next, potential threats to identification will be described. Finally, the econometric model will be explained.

4.1 Identification strategy and threats to identification

To measure the impact of ordinal rank on test scores and school advice one cannot use a simple OLS regression analysis, because ordinal rank and test scores are highly correlated. Getting high grades usually results in having a high ordinal rank. Fortunately, we can exploit the variation in grade distributions between classrooms to estimate the effect of ordinal rank. Since classrooms are small and every classroom has a unique grade distribution, students with the same ability may have a different ordinal rank in different classrooms. To illustrate this, consider two students with the same ability α, but different classrooms, X and Y. The classrooms can differ because the mean test score is different or because the shape of the ability distribution is different. As a result, two students with the same ability can have a different ordinal rank, as figure 2 illustrates.

Figure 2: Identifying variation due to differences in means (A) and shape of the ability distribution (B)



This figure is inspired by figure 1 in Elsner & Isphording (2015). Each number represents a quintile within the ability distribution of the classroom.

Since we cannot measure ability directly, we need to rely on a proxy for ability in the form of externally marked test scores. However, test scores are an imperfect measure of ability. Some classrooms might have access to better resources, such as better teachers, which improves the test scores of their students but does not change their ordinal rank. A student from a classroom with good resources might have obtained a lower test score if he or she was placed in another classroom. In this case, students from different classrooms are not perfectly comparable, which makes it impossible to isolate the effect of ordinal rank on outcomes. In other words, ordinal rank might contain hidden information about ability, which could bias our results. To control for differences in resources both within and between schools, we include school and cohort fixed effects. This should absorb the bias in the effect of ordinal rank due to differences in resources between classrooms that affect all students in a classroom in the same way. Including these fixed effects in our model means that we will be looking at whether students who have a higher rank than average, given their school, cohort, prior test score and observable student characteristics, obtain higher test scores.

Another possible threat to our identification strategy is selection into schools and cohorts. If parents select the school of their children based on the ordinal rank they expect for their child, ordinal rank may contain hidden information about a student. This seems unlikely since parents would need information about the ability of their child at the age of 3 and the ability of all students in the potential classrooms. In addition, empirical evidence shows that parents favor schools with high average grades, which is not beneficial for the rank of their child (Gibbons et al., 2013; Rothstein, 2006). This makes selection into schools based on the likelihood of getting a high ordinal rank even less likely.

Selection into cohorts is also highly unlikely since parents would need to influence the date of birth of their child and would need knowledge about the grade distribution within different (future) cohorts. Since our dataset contains data from multiple cohorts of students within the same school, we can still control for selection into schools and cohorts. In our robustness section, we show that including an interaction of school and cohort fixed effects does not change our results.

4.2 Econometric model

To estimate the effect of ordinal rank on educational performance, we will use an econometric model similar to the models used in Murphy & Weinhardt (2018a) and Elsner & Isphording (2017). We will use the following regression setup:

$$Educational\ performance_{ijsc} = \beta_0 + \beta_1\ ordinal\ rank_{ijsc} + \beta_3\ g\left(prior\ test\ score_{ijsc}\right) + \beta_4 X_{ijsc} + School\ FE_s + Cohort\ FE_c + \varepsilon_{ijsc}$$

We will use three different outcome variables that measure educational performance at the age of 12: standardized low-stakes test scores (PRIMA), standardized high-stakes test scores (CITO) and school advice of student i in subject j, school s and cohort c. We perform separate analyses for the subjects Dutch and mathematics, two core courses in Dutch primary schools. The variable of interest is β_1**,** which measures the effect of an increase in ordinal rank in grade 2 (expressed in percentages) for one of the subjects on our outcome variable.

Since we are interested in the effect of ordinal rank exclusively, we condition on prior test scores by including function g. In our main estimates, g is a linear function of the prior test score for Dutch or mathematics. In our robustness section, we will test whether including a polynomial changes our results. Including a polynomial allows for the possibility that the relationship between test scores in year t and educational performance in the following years is non-linear. We will show that our estimate is indeed robust to changing the functional form of g to quadratic, cubic or quartic.

X is a vector of individual control variables. Because the outcome variables (test scores and school advice) are different for different demographic groups, we should control for demographics. For example, children with a migration background tend to have lower test scores, as shown in table A.3. This might bias the results since children with a migration background also tend to have a lower rank. In addition, since students with a better social-ethnical background tend to perform better both in terms of test scores and ordinal rank, we should control for this as well. By including dummies for gender, socio-ethnic background, the education of the father and mother and age in months we limit the bias in the estimate of β_1.

By including fixed effects at the school and cohort level we remove the mean differences between schools and cohorts in educational performance, average cognitive ability and demographic composition. Again, including these fixed effects means that we will be testing whether students with a higher rank than average, given their school, cohort, prior test score and observable student characteristics, perform better in later years.

The last term, ε_{ijsc}, is an error term that captures all unobservable factors that affect educational performance. We cluster the standard errors at the school level since we also calculated rank at the school level and the outcomes of students within a school are likely to be correlated.

5. Results

5.1 Effect of rank in grade 2 on age-12 test scores

We will begin this section by presenting estimates of the impact of ordinal rank in grade 2 on PRIMA test scores in grade 8 of primary school. Table A.4 shows the coefficient of our variable of interest and all control variables using four different models. In column 1, we show the estimates of a linear regression model, without any control variables or fixed effects. In column 2 we added control variables for age, gender, educational level of both parents and socio-ethnic background. In column 3 we show the estimates of a fixed effects model in which we include fixed effects for the school and cohort, but no other control variables. In column 4 we add control variables to this fixed effect model. As indicated by the increased adjusted R^2 in column 4, adding control variables increases the explanatory power of the model. Gender, age, educational level of the father and mother and socio-ethnic status all appear to be important control variables. In all models except the first one, rank is significant at a 1% level. In our most sophisticated fixed effects model, a jump in rank from 0 to 1 increases a student's PRIMA test score for Dutch by 10.216 points and for mathematics by 7.515 points.

Since it is very unlikely that a student will ever move from the very bottom of the class (0) to the top (1), we will get more useful estimates when using standardized rank and test scores. This will also make it easier to compare the separate effects for Dutch and mathematics, since the mean PRIMA test score for Dutch is 1055 and for mathematics 295 (for the complete sample). Therefore, table 4 on the next page shows estimates using standardized rank and test scores, with a mean of 0 and standard deviation of 1. We based table 4 on our extensive fixed effects model. In all estimates, standardized rank is significant at a 1% level. A one standard deviation increase in rank for the subject Dutch increases test scores for that same subject with 0.083 standard deviations. The effect of ordinal rank in mathematics is even larger: a one standard deviation increase in rank for mathematics increases test scores with 0.226 standard deviations.

Table 4: The effect of rank in grade 2 on standardized test scores and school advice in grade 8, using a fixed effects model

	Standardized Age 12 PRIMA scores Dutch	Standardized Age 12 PRIMA scores Math	Standardized Age 12 CITO scores Dutch	Standardized Age 12 CITO scores Math	Standardized school advice (using rank and test score Dutch)	Standardized school advice (using rank and test score Math)
Standardized Rank grade 2	0.083*** *0.019*	0.226*** *0.025*	0.071*** *0.016*	0.246*** *0.019*	0.123*** *0.022*	0.256*** *0.018*
Standardized Test score grade 2	0.408*** *0.042*	1.257*** *0.284*	0.251*** *0.032*	0.970*** *0.168*	0.335*** *0.147*	0.888*** *0.136*
Gender	*-0.063**** *0.023*	-0.313*** *0.023*	0.069*** *0.017*	-0.355*** *0.027*	-0.082*** *0.025*	-0.034 *0.025*
Age	-0.155*** *0.023*	-0.250*** *0.023*	-0.119*** *0.020*	-0.263*** *0.027*	-0.246*** *0.024*	-0.273*** *0.023*
Education father	0.032** *0.013*	0.019 *0.015*	0.034*** *0.010*	0 .034* *0.018*	0 .042** *0.015*	0.043*** *0.014*
Education mother	0.074*** *0.0176*	0.073*** *0.145*	0.050*** *0.014*	0.072*** *0.018*	0.079*** *0.017*	0.077*** *0.017*
Ses	0.122*** *0.016*	0.055*** *0.016*	0.086*** *0.013*	0.047** *0.019*	0.124*** *0.017*	0.127*** *0.017*
School & cohort fixed effects	yes	yes	yes	yes	yes	yes
#n	7,080	6,850	5,506	5,514	6,383	6,419
R^2	0.301	0.336	0.658	0.322	0.311	0.343
Adj. R^2	0.268	0.304	0.641	0.288	0.277	0.311

Standard errors (in Italics) are clustered at the classroom level.
**** p<0.01, ** p<0.05, *p<0.1*

To interpret these results, it is useful to take a closer look at the standard deviations in table A.2. The standard deviation of rank for the subject Dutch in grade 2 is 0.276. A one standard deviation increase in rank therefore roughly corresponds with moving from the bottom of the classroom to first quartile, from the first quartile to the median, from the median to the third quartile or from the third quartile to the top of the class. A one standard deviation increase in PRIMA test scores for the subject Dutch, in turn, corresponds with an increase of 34.970 points (the mean score is 1117.465). In other words, if a student's moves up one quartile in the grade distribution in grade 2 this results in approximately 2.9 extra points[1] on the test in grade 8.

The standard deviation of rank for mathematics in grade 2 is 0.275, while the standard deviation of the test score in grade 8 is 9.349 (with a mean of 117.26). In other words, if a student's rank in grade 2 improves by about one quartile this results in approximately 2.1 extra points on the test in grade 8. Very similar results are found when using CITO test scores as the dependent variable, as column 3 and 4 show. This suggests that the effect of rank is similar for tests where the stakes are low (PRIMA) and tests where the stakes are high (CITO end-of-primary test in grade 8).

5.2 Effect of rank in grade 2 on age-12 school advice

In addition to the positive effect on test scores, ordinal rank is associated with a positive effect on school advice, as shown by column 5 and 6 of table 4. A one standard deviation increase in rank for Dutch increases school advice by 0.123 standard deviations. Again, for the interpretation of this result, it is useful to take a look at table A.2. School advice can take the values 1 to 15, with a mean of 10.492 and a standard deviation of 3.568. If a student's rank improves with one quartile in grade 2 this results in a school advice in grade 8 that is about 0.4 points higher. Again, the effects are larger within the subject mathematics: a one standard deviation increase in rank for mathematics increases school advice by 0.256 standard deviations. The effect of ordinal rank on school advice can have long-lasting consequences, since the school advice determines which level of high school a student can attend. The level of high school in turn determines whether a student can apply for university, university of applied sciences or a lower level of tertiary education.

The effect of rank in grade 2 on school advice in grade 8 is larger than the effects found in section 5.1. If rank effects exist because teachers invest more in students with a higher ordinal rank, one would expect the effect of rank on school advice (a subjective measure of performance) to be larger than the effect on PRIMA and CITO test scores (objective measures of performance). The results are in line with this expectation. In section 6 we will take a closer look at the mechanisms through which ordinal rank could influence performance.

5.3 Effect of rank in grade 4 and 6 on age 12 test scores and school advice

Since the dataset provides us with data on grades and rank in grade 2, 4, 6 and 8, we are also able to estimate whether the results become stronger when we base them on rank in grade 4 or 6 instead of grade 2. Table 5 shows the effect of rank in grade 4, conditional on test scores in grade 4, on test scores and school advice in grade 8. These short term effects are indeed stronger than the effect of rank in grade 2 and all estimates are significant at the 1% level. Table 6 shows the effect of rank in grade 6 - conditional on test scores in grade 6 - on test scores and school advice in grade 8. Some of these very short term effects are stronger than the effects of ordinal rank in grade 4 and grade 2, while others are not.

[1] $34.970 \times 0.083 \approx 2.9$

Table 5: The effect of standardized rank in grade 4 on standardized test scores and school advice in grade 8, using a fixed effects model

	Standardized Age 12 PRIMA scores Dutch	Standardized Age 12 PRIMA scores Math	Standardized Age 12 CITO scores Dutch	Standardized Age 12 CITO scores Math	School advice (using rank and test score Dutch)	School advice (using rank and test score Math)
Standardized Rank grade 4	0.145*** *0.023*	0.487*** *0.019*	0.097*** *0.021*	0.477*** *0.019*	0.129*** *0.023*	0.455*** *0.016*
Standardized Test score grade 4	0.373*** *0.046*	0.520*** *0.197*	0.264*** *0.042*	0.644*** *0.216*	0.369*** *0.046*	0.391** *0.164*
Gender	-0.061*** *0.022*	-0.152*** *0.022*	0.080*** *0.017*	-0.191*** *0.024*	-0.066*** *0.025*	0.129*** *0.022*
Age	-0.100*** *0.022*	-0.151*** *0.019*	-0.078*** *0.018*	-0.150*** *0.024*	-0.192*** *0.023*	-0.181*** *0.021*
Education father	0.028** *0.013*	0.018 *0.013*	0.026** *0.011*	0.030* *0.016*	0.038** *0.015*	0.035*** *0.012*
Education mother	0.067*** *0.018*	0.059*** *0.014*	0.046*** *0.013*	0.057*** *0.016*	0.077*** *0.018*	0.071*** *0.014*
Ses	0.106*** *0.016*	0.042*** *0.014*	0.076*** *0.013*	0.034** *0.016*	0.107*** *0.017*	0.111*** *0.014*
School & cohort fixed effects	yes	yes	yes	yes	yes	yes
#n	7,230	6,952	5,597	5,552	6,505	6,483
R^2	0.331	0.461	0.671	0.460	0.332	0.441
Adj. R^2	0.300	0.435	0.655	0.433	0.300	0.414

Standard errors (in Italics) are clustered at the classroom level.
**** p<0.01, ** p<0.05, *p<0.1*

Table 6: The effect of standardized rank in grade 6 on standardized test scores and school advice in grade 8, using a fixed effects model

	Standardized Age 12 PRIMA scores Dutch	Standardized Age 12 PRIMA scores Math	Standardized Age 12 CITO scores Dutch	Standardized Age 12 CITO scores Math	School advice (using rank and test score Dutch)	School advice (using rank and test score Math)
Standardized Rank grade 6	0.209*** *0.022*	0.077* *0.040*	0.180*** *0.018*	0.150*** *0.024*	0.946*** *0.077*	0.657*** *0.083*
Standardized Test score grade 6	0.624*** *0.049*	25.81*** *1.801*	0.413*** *0.035*	22.568*** *1.146*	1.720*** *0.147*	66.010*** *3.530*
Gender	-0.074*** *0.020*	-0.082*** *0.017*	0.058*** *0.015*	-0.108*** *0.019*	-0.279*** *0.077*	0.694*** *0.066*
Age	-0.053*** *0.019*	-0.093*** *0.015*	-0.045*** *0.016*	-0.082*** *0.019*	-0.541*** *0.073*	-0.417*** *0.064*
Education father	0.021** *0.011*	-0.005 *0.011*	0.023*** *0.009*	0.010 *0.013*	0.099** *0.048*	0.059 *0.038*
Education mother	0.047*** *0.015*	0.038*** *0.014*	0.027** *0.011*	0.043*** *0.014*	0.187*** *0.054*	0.198*** *0.048*
Ses	0.077*** *0.013*	0.017 *0.011*	0.050***	0.010 *0.011*	0.296*** *0.051*	0.313*** *0.044*
School & cohort fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
#n	7,242	6,788	5,605	5,415	6,510	6,309
R^2	0.459	0.659	0.743	0.660	0.459	0.595
Adj. R^2	0.748	0.643	0.730	0.642	0.433	0.575

Standard errors (in Italics) are clustered at the classroom level.
**** p<0.01, ** p<0.05, *p<0.1*

Why are (most of) the short term effects stronger? It might be that students become more aware of their relative position in class when they get older, so that rank plays a more important role in grade 4 and 6. This might also be the result of changing teaching methods: in grade 2 students learn through play and communication, mostly soft skills, while in grade 4 and 6 students learn how to read and do math, which can be classified as hard skills. One can imagine that it is easier to compare yourself with others based on hard skills than on soft skills.

5.4 Robustness checks

5.4.1 Functional form of g

As a first robustness check we will test whether our estimate is robust for different functional forms of g(prior test scores). There is, as also described by previous empirical studies in this field, a possibility that ability (measured by prior test score) influences test scores in a non-linear way. Ignoring this possibility might result in a biased estimate of ordinal rank. Therefore, we will test whether including a different polynomial for prior test scores changes our results.

In our main estimates, prior test scores entered linearly. Table A.5 shows the estimates using different polynomials of prior test scores: quadratic, cubic and quartic. The estimates change very little, while the adjusted R-squared stays as good as constant. Therefore, we conclude that our model is robust to different functional forms of prior test scores.

5.4.2 Inclusion of school x cohort fixed effects

One concern when measuring ordinal rank effects is that average cohort characteristics, like average peer ability, might bias the estimates. Remember figure 2 in which we showed that the identifying variation in ordinal rank may come from differences in means or differences in ability distributions. Therefore, as a robustness check, we exploit a method similar to Elsner & Isphording (2017) and include school x cohort fixed effects. By including an interaction of school and cohort fixed effects we take school specific mean differences across cohorts into account. The estimate is now based only on the variation of the ability distribution within schools across cohorts. This corresponds with the right side (B) of figure 2. The effect of ordinal rank for the subject Dutch on PRIMA scores and school advice increases only very slightly, as table A.6 shows, so that these estimates seem to be robust. The other estimates seem to be less robust, although the significance and sign stay the same. Since all estimates become stronger when including cohort x school fixed effects and they stay both positive and significant, there is not much reason to worry.

5.4.3 Selective attrition

As shown in section 3.1 of this paper, the outflow of students that ceased to participate in the next year(s) of the study is large. The majority of the students that dropped out of the study did so because their whole classroom or school ceased to participate in the study. This should not cause any problems for our analysis. However, the smaller group of students who dropped out while their classmates *did* participate in the next wave may be problematic. This may concern students who moved to a different school or had to repeat a class. If the individual students that dropped out differ from the students in the final sample our results may be biased. One may expect, for example, that students who dropped out because they had to repeat a class have a lower ordinal rank than the average student in the sample. Table 7 shows that, indeed, students who stopped to participate in the survey after one year do have a lower ordinal rank than the mean in table A.2.

Table 7: Rank and test scores of students who dropped out after one wave

	Mean	Standard deviation	Observations
Test scores grade 2			
Dutch	961.641	36.516	19,622
Mathematics	597.394	395.118	19,471
Rank grade 2			
Dutch	0.412	0.299	19,569
Mathematics	0.401	0.299	19,408

As a robustness check, we include all students who participated in the study while they were in grade 2 and did not participate in any of the following years. Since we do not observe the test scores and school advice of these students in grade 8, we predicted these based on the observations from our 'normal' sample. To predict the test score for Dutch in grade 8, for example, we started with regressing the test scores in grade 8 on test scores in grade 2, using the sample from our results section. We used the constant (β_0) and coefficient of interest (β_1) from this regression to calculate predicted test scores for the students who did not participate after grade 2.

$$Predicted\ test\ score\ grade\ 8 = \beta_0 + \beta_1 * test\ score\ grade\ 2$$

In order to calculate the predicted school advice, we regressed school advice in grade 8 on both the test score for Dutch and mathematics in grade 2 and used the constant and coefficient of interest from this regression.

Next, we performed the same regression analyses as in table 4 of our results section while including the predicted test scores from the students who ceased to participate after grade 2. If selective attrition biased our previous results, we would expect the effect of ordinal rank to change when including the students who stopped participating. As table A.7 shows, the magnitude of the effects does decrease slightly, but the sign does not and ordinal rank is still significant in all estimates. Therefore, we conclude that the bias from selective attrition is limited.

6. Potential mechanisms

In this section we will explore three possible mechanisms through which ordinal rank could positively influence educational performance. The first mechanism that we will discuss is an increased level of confidence. Alternatively, rank might impact performance by motivating students to exert more effort, which is the second mechanism that we will discuss. Finally, the third mechanism that we will discuss is increased teacher investment, correlated with ordinal rank. We will test the importance of all three of these mechanisms using survey data from the PRIMA study.

6.1 Confidence

One of the channels through which ordinal rank may influence performance is an increased level of confidence. In the presence of uncertainty about their own ability, students might update their perceived ability based on ordinal rank, in addition to absolute performance. A high perceived ability may result in more confidence, which may in turn positively impact test performance, as also shown by Mavis (2001). Therefore, an increase in ordinal rank might influence performance through an increase in the confidence level of students. In the psychological literature, this has been described as the Big Fish Little Pond effect (Marsh, 1987).

Teachers from the schools that participated in the study filled in a survey in which they graded the confidence level of their students, on a scale of 1 to 5. The mean confidence score of all students in the sample was 3.656 (21,231 observations). Column 1 in table 8 and column 1 in table 9 show that, conditional on test scores, confidence and ordinal rank for Dutch and mathematics are correlated. Of course, confidence might also be correlated with confounding factors that are also correlated with rank, such as socio-ethnic background. Therefore, we will use the econometric model from our methodology section but with confidence as the outcome variable and without lags. Column 2 in table 8 and column 2 in table 9 show that, using this extensive model, rank does indeed have a positive effect on confidence. A one standard deviation increase in rank for Dutch improves confidence with 0.078 points on a 5 point scale. A one standard deviation increase in rank for mathematics improves confidence with 0.142 points on a 5 point scale.

Table 8: The effect of rank on confidence, work attitude and teacher investment for the subject Dutch, using OLS and fixed effects

	Confidence Simple OLS	Confidence Fixed effects	Work attitude Simple OLS	Work attitude Fixed effects	Extra material Simple OLS	Extra material Fixed effects
Rank Dutch (standardized)	0.067*** *0.006*	0.078*** *0.009*	0.184*** *0.006*	0.138*** *0.009*	0.272*** *0.008*	0.156*** *0.015*
Test score Dutch (standardized)	0.053*** *0.006*	0.032* *0.019*	-0.003 *0.006*	0.051*** *0.016*	0.120*** *0.009*	0.329*** *0.027*
Fixed effects	no	yes	no	yes	no	yes
Control variables	no	yes	no	Yes	no	yes
n	20,923	19,418	22,894	21,034	15,968	14,973

Standard errors (in Italics) are clustered at the classroom level.
**** p<0.01, ** p<0.05, *p<0.1*

Table 9: The effect of rank on confidence, work attitude and teacher investment for the subject mathematics, using OLS and fixed effects

	Confidence Simple OLS	Confidence Fixed effects	Work attitude Simple OLS	Work attitude Fixed effects	Extra material Simple OLS	Extra material Fixed effects
Rank mathematics (standardized)	0.145*** *0.005*	0.142*** *0.006*	0.229*** *0.005*	0.240*** *0.006*	0.425*** *0.007*	0.356*** *0.015*
Test score mathematics (standardized)	-0.045*** *0.005*	0.016* *0.010*	0.014*** *0.005*	0.005 *0.010*	1.429*** *0.112*	3.716** *0.456*
Fixed effects	no	yes	no	yes	no	yes
Control variables	no	yes	no	yes	no	yes
n	20,613	19,159	22,569	20,767	15,683	14,740

Standard errors (in Italics) are clustered at the classroom level.
**** p<0.01, ** p<0.05, *p<0.1*

6.2 Effort

The effect of ordinal rank on performance might also be explained by other intrinsic factors that are linked to confidence, such as effort. It might be that students with a higher ordinal rank (conditional on ability) exert more effort and have a better work ethic. For example, because they expect the return to effort to be higher. Clark et al (2010), using experimental evidence combined with survey data, find that individual effort on the job depends on both one's own income and the individual's rank in the income distribution. They find that individuals with a higher rank exert more effort. Something similar might happen in an educational setting: students with a higher rank might exert more effort in class. This is also what Elsner & Isphording (2015) find.

In the survey the teachers filled in they also graded the work attitude of their students, on a scale of 1 to 5. The mean work attitude score of all the students in the sample was 3.486 (23,805 observations). We will use this score as a proxy for the effort exerted by students. Column 3 in table 8 and column 3 in table 9 show that, conditional on test scores, work attitude and ordinal rank for Dutch and mathematics are correlated. Again, work attitude might be correlated with confounding factors that are also correlated with rank. Column 4 in table 8 and column 4 in table 9 show that, using the extensive fixed effects model, rank does indeed have a positive effect on the work attitude of students. This supports the idea that students exert more effort if they have a high ordinal rank, which could positively impact their performance in later years.

6.3 Teacher investment

Another possible explanation that we can test with our data is increased teacher investment. Just as students might base their perceived ability on their ordinal rank, teachers might estimate the ability of their students based on ordinal rank, in addition to absolute performance. Consequently, teachers may invest more time and effort in students with a high ordinal rank, because they expect these students to deliver the highest returns. This might improve performance of students with a high ordinal rank.

Primary schools in the Netherlands often offer extra material for smart students, which might positively affect their grades in the longer term. This is measured by the variable *Extra material* in our dataset. Students that are not offered any extra material get a score of 1, while students with the maximum amount of extra material get a score of 5. The mean value of *Extra material* is 2.832 (21,188 observations). We use *Extra material* as a proxy for teacher investment in smart students. Conditional on prior test scores, ordinal rank and extra material are correlated, as shown by column 5 in table 8 and 9. Again, extra material might be correlated with confounding factors that are also correlated with ordinal rank. Therefore, we will again use our extensive fixed effects model, now with extra material as the outcome variable. A one standard deviation increase in ordinal rank for Dutch increases extra material with 0.156 points on a 5 point scale, as column 6 in table 8 shows. A one standard deviation increase in ordinal rank for mathematics increases extra material with 0.356 points, as column 6 in table 9 shows. These effects are stronger than the ones in section 6.1 and 6.2, which would indicate that teacher investment based on ordinal rank is a slightly more important mechanism than confidence and effort.

Of course, there may be other channels through which ordinal rank influences performance. Confidence, effort and teacher investment in the form of offering extra material are just three of these channels. Unfortunately, our dataset is not extensive enough to investigate other channels, such as increased parental investment.

7. Conclusion

In this paper we studied the effect of ordinal rank in grade 2 of primary school on educational performance in grade 8. The paper is based on data from a longitudinal study that included a large sample of students from primary schools in the Netherlands. To estimate a causal effect, we exploited the idiosyncratic variation in test score distributions across schools and cohorts, a method also used by Murphy & Weinhardt (2018) and Elsner & Isphording (2017).

The results indicate that ordinal rank does indeed have a positive effect on educational performance in the form of test scores and school advice for the subjects Dutch and mathematics. In other words, being better than the rest of the students in your classroom matters. A one standard deviation increase in ordinal rank for Dutch increases test scores in grade 8 with 0.083 standard deviations. The results are stronger for the subject mathematics, suggesting that rank effects play a more important role within this subject. A one standard deviation in ordinal rank for mathematics increases test scores with 0.226 standard deviations. Furthermore, the effect of rank on school advice is stronger than the effect on both low and high stake test scores. The school advice students receive at the end of primary school is provided by the teacher. Although teachers should incorporate the CITO score in their advice, it is much more subjective than national test scores. Therefore, the results are a first indication that teachers do update their beliefs about the ability of their students based on rank. This might cause teachers to invest more time and effort in students with a high ordinal rank, a mechanism that we tested in the last section.

In our main analysis we controlled for prior test scores, school and cohort fixed effects and several observable student characteristics. There is a possibility that ability, measured by prior test scores, influences performance in a non-linear way. To check whether this influences our result, we performed a first robustness check in which we included different polynomials of prior test scores. We conclude that our model is robust to different functional forms of prior test scores. Furthermore, the results also prove to be robust when including an interaction of cohort and school fixed effects. This controls for school specific mean differences across cohorts. As a last robustness check we included all students who participated in the study while they were in grade 2 and did not participate in <u>any</u> of the following years. We predicted the test scores and school advice they would have obtained in grade 8 and included these in our analysis. The magnitude of the effects decreases but they are still highly significant and positive, so we conclude that the bias from selective attrition is limited.

In the last section of this paper we studied three possible mechanisms that could explain the effect of ordinal rank on educational performance. Having a higher ordinal rank may increase the confidence level of students, it may stimulate them to exert more effort or it may result in more time and effort from the teacher. All three mechanisms, confidence, effort and increased teacher investment, are supported by the survey data. Increased teacher investment appears to be slightly more important than confidence and effort. Further research into the mechanisms behind the positive effect of ordinal rank could help us to better understand the positive rank effects found in this and other papers.

References

Ammermueller, A., & Pischke, J. (2009). Peer effects in European primary schools: Evidence from the progress in international reading literacy study. *Journal of Labor Economics*, 315-348.

Bifulco, R., Fletcher, J., & Ross, S. (2011). The Effect of Class-mate Characteristics on Post-Secondary Outcomes: Evidence from the Add Health. *American Economic Journal: Economic Policy*, 3, 25-53.

Burke, M., & Sass, T. (2013). Classroom Peer Effects and Student Achievement. *Journal of Labor Economics*, 31(1), 51-82.

Carrell, S., Fullerton, R., & West, J. (2009). Does Your Cohort Matter? Measuring Peer Effects in College Achievement. *Journal of Labor Economics*, 27(3), 439-464.

Clark, A., Masclet, D., & Villeval, M. (2010). Effort and Comparison Income: Experimental and Survey Evidence. *Industrial and Labor Relations Review*, 407-426.

Denning, J., Murphy, R., & Weinhardt, F. (2018). Class Rank and Long-Run Outcomes. *IZA Discussion Papers*.

Elsner, B., & Isphording, I. (2017). A Big Fish in a Small Pond: Ability Rank and Human Capital Investment. *Journal of Labor Economics*, 35(3), 787-828.

Elsner, B., & Isphording, I. (2018). Rank, Sex, Drugs and Crime. *Journal of Human Resources*, 53(2), 356-381.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7(2), 117-140.

Gibbons, S., & Telhaj, S. (2012). Peer Effects: Evidence from Secondary School Transition in England. *IZA Discussion Papers*, No. 6455.

Gibbons, S., Machin, S., & Silva, O. (2013). Valuing school quality using boundary discontinuities. *Journal of Urban Economics*, 75, 15-28.

Hanushek, E., Kain, J., Markman, J., & Rivkin, S. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics*, 18(5), 527-544.

Hoxby, C. (2000). Peer Effects in the Classroom: Learning from Gender and Race Variation. *NBER Working Papers*, No. 7867.

Lavy, V., Paserman, D., & Schlosser, A. (2012a). Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers In the Classroom. *The Economic Journal*, 112(559), 208-237.

Lavy, V., Silva, O., & Weinhardt, F. (2012b). The Good, the Bad, and the Average: Evidence on Ability Peer Effects in Schools. *Journal of Labor Economics*, 30(2), 367-414.

Marsh, H. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Education Psychology*, 79(3), 280.

Mavis, B. (2001). Self-Efficacy and OSCE Performance Among Second Year Medical Students. *Advances in Health Sciences Education*, 6(2), 93-102.

Murphy, R., & Weinhardt, F. (2018). Top of the Class: The Importance of Ordinal Rank. *NBER Working papers*, No. 24958.

Rothstein, J. (2006). Good principals or good peers? Parental valuation of school characteristics, Tiebout equilibrium, and the incentive effects of competition among jurisdictions. *American Economic Review*, 96(4), 1333-1350.

Sacerdote, B. (2001). Peer Effects with Random Assignment: Results for Dartmouth Roommates. *The Quarterly Journal of Economics*, Vol. 116, Issue 2, p. 681-704.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185(4157), 1124-1131.

Whitmore, D. (2005). Resource and peer impacts on girls' academic achievement: Evidence from a randomized experiment. . *American Economic Review*, 199-203.

Zimmerman, D. (2003). Peer effects in academic outcomes: Evidence from a natural experiment. Review of Economics and statistics. *Review of Economics and Statistics*, 9-23.

Appendix

Figure A.1: Flow diagram with in- and outflow of students who participated in the study

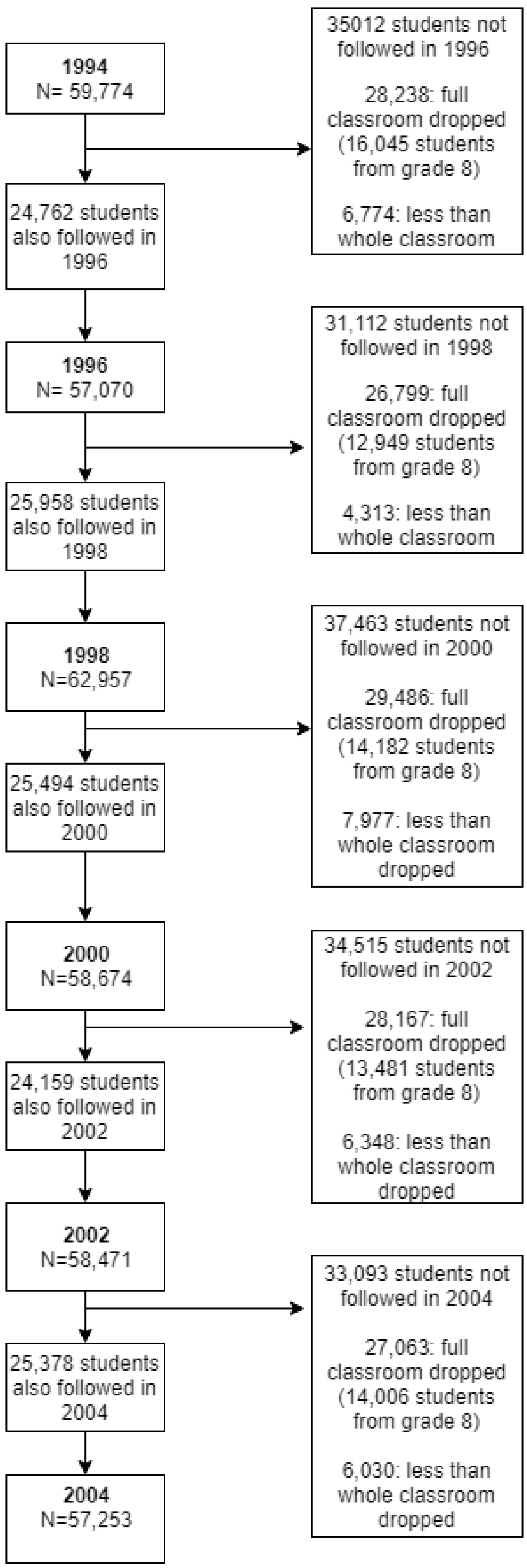


Table A.2: Summary of demographics, rank and outcomes of the students in our final sample

	Mean	Standard deviation	Observations
Demographics			
Female	0.505	0.500	35,620
Social-ethnic background (ses)	3.525	1.332	36,093
Share of students with two Dutch parents	0.672	0.469	36.133
Age in grade 2	5.355	0.512	9,073
Education father	2.683	1.134	34,358
Education mother	2.544	1.068	34,344
Test scores grade 2			
Dutch	978.368	34.288	8,569
Mathematics	580.002	418.773	8,583
Rank grade 2			
Dutch	0.525	0.276	8,559
Mathematics	0.529	0.275	8,573
Outcomes grade 8			
PRIMA Dutch	1117.465	34.970	8,569
PRIMA mathematics	117.256	9.349	8,222
CITO Dutch	60.588	19.040	6,614
CITO mathematics	41.984	11.363	6,591
School advice	10.492	3.568	7,715

Table A.3: Rank, test scores and school advice for the different demographic groups within our sample

	CITO score Dutch		CITO score mathematics		PRIMA score Dutch		PRIMA score mathematics		Percentile rank Dutch		Percentile rank mathematics		School advice	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Male	59.407	(19.300)	43,916	(10.707)	1052.488	(62.517)	296.905	(374.578)	0,479	(0.283)	0,527	(0.276)	10.494	(3.561)
Female	61.762	(18.670)	40.065	(11.661)	1056.974	(61.171)	294.609	(373.213)	0.512	(0.284)	0,466	(0.283)	10.479	(3.572)
Both parents are Dutch	62,179	(19.195)	43.046	(11.092)	1061.887	(60.692)	310.393	(384.497)	0.515	(0.282)	0.507	(0.280)	10.928	(3.355)
Father is Dutch, mother is not	59.175	(20.590)	42.423	(12.246)	1058.206	(62.728)	432.127	(429.595)	0.536	(0.282)	0.524	(0.292)	11.168	(3.492)
Mother is Dutch, father is not	63.932	(16.686)	40.419	(10.759)	1057.719	(58.862)	222.803	(317.661)	0.536	(0.280)	0.496	(0.280)	9.952	(3.623)
Both parents are not Dutch	56.124	(18.237)	39.686	(11.700)	1035.906	(61.452)	265.721	(352.092)	0.435	(0.282)	0.466	(0.283)	9.382	(3.855)
Ses <3.5	54.744	(18.280)	38.640	(11.777)	1043.352	(59.387)	301.323	(376.406)	0.451	(0.283)	0.456	(0.281)	8.961	(3.678)
Ses >3.5	65.030	(18.398)	44.557	(10.323)	1062.878	(62.334)	292.777	(374.326)	0.527	(0.281)	0.524	(0.278)	11.562	(3.066)

Standard errors in parentheses

Table A.4: The effect of rank in grade 2 on test scores in grade 8, using OLS and fixed effects

	Age 12 PRIMA scores Dutch				Age 12 PRIMA scores mathematics			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Rank grade 2	-1.927	8.859***	10.972***	10.216***	10.445***	10.309***	7.483***	7.515***
	2.433	*2.111*	*2.347*	*2.339*	*0.455*	*0.480*	*0.867*	*0.820*
Test score grade 2	0.400***	0.238***	0.261***	0.231***	0.001	0.001	0.030***	0.031***
	0.024	*0.021*	0.024	*0.024*	*0.001*	*0.001*	*0.006*	*0.007*
Gender		*-1.896***		-2.200***		-2.801***		-2.927***
		0.791		*0.798*		*0.221*		*0.213*
Age		-5.224***		-5.408***		-2.187***		-2.339***
		0.808		*0.810*		*0.249*		*0.213*
Education father		1.008**		1.125**		0.278**		0.181
		0.420		*0.456*		*0.136*		*0.138*
Education mother		3.294***		2.577***		0.782***		0.679***
		0.544		*0.616*		*0.144*		*0.136*
Ses		5.639***		4.273		0.823***		0.515***
		0.481		*0.549*		*0.174*		*0.146*
School & cohort fixed effects	no	no	yes	yes	no	no	yes	yes
#n	8,015	7,080	8,015	7,080	7,701	6,850	7,701	6,850
R^2	0.147	0.233	0.255	0.301	0.097	0.179	0.285	0.336
Adj. R^2	-	-	0.224	0.268	-	-	0.254	0.304

Standard errors (in Italics) are clustered at the classroom level.

**** p<0.01, ** p<0.05, *p<0.1*

Columns 1 and 2 are based on an OLS regression, columns 3 and 4 on a Fixed Effects model.

Table A.5: The results are robust when using different functional forms of g, a function of the test score in grade 2

	Standardized Age 12 PRIMA scores Dutch	Standardized Age 12 PRIMA scores Dutch	Standardized Age 12 PRIMA scores Dutch	Standardized Age 12 PRIMA scores math	Standardized Age 12 PRIMA scores math	Standardized Age 12 PRIMA scores math
Standardized Rank grade 2	0.076*** *0.022*	0.065*** *0.023*	0.062*** *0.023*	0.149*** *0.029*	0.149*** *0.029*	0.135*** *0.032*
Functional form test scores grade 2	*Quadratic*	*Cubic*	*Quartic*	*Quadratic*	*Cubic*	*Quartic*
Control variables	yes	yes	yes	yes	yes	yes
School & cohort fixed effects	yes	yes	yes	yes	yes	yes
#n	7,077	7,077	7,077	6,849	6,849	6,849
R^2	0.301	0.302	0.302	0.340	0.341	0.342
Adj. R^2	0.268	0.268	0.269	0.309	0.309	0.311

Standard errors (in Italics) are clustered at the classroom level.
**** p<0.01, ** p<0.05, *p<0.1*
Note: to improve readability we only included estimates with PRIMA test scores as outcome variables. In our other models, the estimate of ordinal rank also proves to be robust.

Table A.6: The results are robust when including school x cohort fixed effects

	Standardized Age 12 PRIMA scores Dutch	Standardized Age 12 PRIMA scores math	Standardized Age 12 CITO scores Dutch	Standardized Age 12 CITO scores math	School advice (using rank and test score Dutch)	School advice (using rank and test score Math)
Standardized Rank grade 2	0.090*** *0.019*	0.321*** *0.013*	0.140*** *0.026*	0.312*** *0.014*	0.462*** *0.081*	1.151*** *0.046*
Control variables	yes	yes	yes	yes	yes	yes
School x cohort fixed effects	yes	yes	yes	yes	yes	yes
#n	7,028	6,849	5,502	5,510	6,376	6,412
R^2	0.301	0.326	0.535	0.318	0.311	0.339
Adj. R^2	0.268	0.294	0.512	0.284	0.277	0.307

Standard errors (in Italics) are clustered at the classroom level.
**** p<0.01, ** p<0.05, *p<0.1*

Table A.7: The effect of ordinal rank on educational performance when including students who ceased to participate after grade 2, as a robustness check for selective attrition

	Standardized Age 12 PRIMA scores Dutch	Standardized Age 12 PRIMA scores Math	Standardized Age 12 CITO scores Dutch	Standardized Age 12 CITO scores Math	Standardized school advice (using rank and test score Dutch)	Standardized school advice (using rank and test score Math)
Standardized Rank grade 2	0.057*** *0.011*	0.158*** *0.012*	0.056*** *0.012*	0.139*** *0.011*	0.073*** *0.012*	0.307*** *0.010*
Standardized Test score grade 2	1.399*** *0.039*	0.054 *0.105*	1.418*** *0.044*	0.123 *0.086*	1.394*** *0.042*	1.541*** *0.092*
Control variables	yes	yes	yes	Yes	Yes	yes
School & cohort fixed effects	yes	yes	yes	Yes	Yes	yes
#n	24,677	24,300	23,101	22,961	22,837	22,873
R^2	0.713	0.662	0.780	0.718	0.718	0.462
Adj. R^2	0.702	0.649	0.771	0.706	0.706	0.439

Standard errors (in Italics) are clustered at the classroom level.
**** p<0.01, ** p<0.05, *p<0.1*