A big fish in a small pond: Ability rank and human capital investment*

Benjamin Elsner[†] Ingo E. Isphording[‡]

April 26, 2016

Abstract

We study the impact of a student's ordinal rank in a high school cohort on educational attainment several years later. To identify a causal effect, we compare multiple cohorts within the same school, exploiting idiosyncratic variation in cohort composition. We find that a student's ordinal rank significantly affects educational choices later in life. If two students with the same ability have a different rank in their respective cohort, the higher-ranked student is significantly more likely to finish high school and attend college. These results suggest that low-ranked students under-invest in their human capital even if they have a high ability compared to most students of the same age. Exploring potential channels, we find that students with a higher rank have higher expectations about their future career, as well as a higher perceived intelligence.

JEL Codes: I21, I23, J24

Keywords: Human capital, ordinal rank, peer effects, educational attainment

*We would like to thank the editor Philip Oreopoulos, three anonymous referees, as well as Peter Arcidiacono, Jan Bietenbeck, Arnaud Chevalier, Deborah Cobb-Clark, Rajeev Dehejia, Jan Feld, Chris Jepsen, Peter Kuhn, Herb Marsh, Julia Anna Matz, Milena Nikolova, Aderonke Osikominu, Daniele Paserman, Dan Rees, Ying Shi, Derek Stemple, Andreas Steinmayr, Steve Stillman, Ulf Zölitz, and audiences at IZA, RWI, U Mainz, IAB, Modul University, WU Vienna, Concordia University, U Mannheim, U Regensburg, U Southern Denmark, the EALE/SOLE World Congress in Montreal, EEA Mannheim, VfS Münster, SOFI, and Maastricht U for helpful comments.

†Corresponding author. Institute for the Study of Labor (IZA). Address: Schaumburg-Lippe-Str. 5-9, 53113 Bonn, Germany. elsner@iza.org, www.benjaminelsner.com.

‡Institute for the Study of Labor (IZA). Address: Schaumburg-Lippe-Str. 5-9, 53113 Bonn, Germany. isphording@iza.org

1 Introduction

The characteristics of potential classmates are among the decisive factors for parents when choosing a school for their child. It is commonly believed that children learn and achieve more when surrounded by high-ability classmates. In this paper, we explore a channel that runs counter to the positive impact of high-ability peers, namely a student's ordinal rank in her peer group. Students with a low absolute ability yet a high relative ability in their school cohort — big fish in a small pond — may erroneously conclude that they have a high absolute ability and thus invest more in their human capital. Psychologists have labeled this phenomenon the big-fish-in-a-little-pond effect (Marsh, 1987).

In this paper, we test whether being a *big fish* in a high school cohort affects the critical transition from high school to college. Consider two students, Jack and Jim, who have the same absolute ability but a different rank in their respective high school cohort: Jack is among the students with the lowest ability in his cohort, while Jim is among the brightest students; namely, Jim is a big fish in a small pond. In this paper, we analyze whether Jim is more likely than Jack to finish high school, attend college and complete a four-year college degree.

To identify a causal effect, we exploit idiosyncratic changes in the cohort composition within schools over time. We argue that, conditional on attending a given school, the cohort composition is as good as random for a particular student, because school entry is determined by a student's birth date and is beyond the influence of parents or students. The analysis is based on data from the National Longitudinal Study of Adolescent to Adult Health (AddHealth), a representative survey that tracks students in the US from middle and high school to their mid-30s, and contains rich information on cognitive skills and educational outcomes. Key to our identification strategy is the fact that AddHealth covers multiple cohorts within the same high school, allowing us to exploit the within-school variation in cohort composition. Moreover, the survey includes an age-specific standardized ability test, which makes cognitive ability comparable within and across schools and cohorts. Based on these test scores, we compute a student's ordinal rank within a school cohort.

Our central finding is that a student's ability rank in a high school cohort has a strong impact on educational outcomes later in life. In particular, it affects the likelihood of attending college. An increase in the ordinal rank by one within-school standard deviation increases the likelihood of going to college by 1.2 percentage points. Given that cognitive ability, parental education, as well as school and cohort characteristics are held constant, this is a large effect. We also estimate the impact of ordinal rank on high school completion and the likelihood of completing a four-year degree, finding positive yet imprecisely estimated effects. Within a cohort, these effect are the same in the bottom and top half of the ability distribution. Moreover, the effects do not differ between schools with high and low average ability, as well as between schools with high and low variance of ability.

Based on different empirical specifications, we rule out confounding factors suggested by the literature on ability peer effects. Previous research has produced a wealth of evidence suggesting that the mean peer ability affects educational outcomes (Sacerdote, 2011), while more recent research emphasizes the variance of ability as an important channel of peer effects (Tincani, 2015). We demonstrate that the impact of ordinal rank on educational attainment is a peer effect in its own right by showing that the effect is robust to controls for the within-cohort mean and variance of ability. Moreover, in a series of robustness checks, we carefully address several threats to identification. One concern is the strategic delay of school entry, whereby parents might send their child to school one year later to avoid a low-ability cohort. A robustness check based on a sample of children in the 'right' age cohort alleviates this concern. A further confounding factor is unobserved differences across cohorts within a school; for example, in the quality of teachers or the share of disruptive students. We address this concern by estimating a model with school-by-cohort fixed effects, which absorbs any mean differences between school cohorts, and identifies the effect purely based on differences across school cohorts with respect to variance and higher moments of the ability distribution. Finally, the stratified random sampling of the survey introduces a measurement error in the ordinal rank. Based on a Monte Carlo experiment, we show that our estimates would be around 20% higher in the absence of measurement error.

In theory, this result can be explained by at least five mechanisms. First, the rank may provide students with incomplete information regarding their own absolute ability. Students may choose not to attend college if a low perceived ability translates into low expected returns to college. Second, the ordinal rank can affect intrinsic factors. Students with a higher rank may be more motivated and self-confident and hence put more effort into their studies, which subsequently translates into a higher chance of attending college. Third, a student's environment may be responsive to their rank. Teachers, family and friends may offer more support to high-ranked students, leading to better grades and a higher chance of attending college. Fourth, tracking may be an important mechanism, if students with a higher rank are more likely to attend higher-level courses, this may lead to a higher educational attainment. Finally, the result could be explained by selective college admission policies. Colleges often observe a student's grade point average (GPA) rank within a cohort, which is correlated with our rank measure. If admissions officers give priority to students with a higher rank regardless of the school quality, or if colleges automatically admit the top 10% of a school cohort, then this can explain the effect.

While we are unable to fully disentangle these mechanisms, we exploit the rich survey information in AddHealth to provide suggestive evidence that some channels are more important than others. We find strong evidence for the expected returns channel. Applying the same empirical strategy as before, we find that a higher rank has an equally large effect on various measures of career expectations at the age of 16 as on the actual outcomes 14 years later. Moreover, we find that students with a higher rank have a higher perceived intelligence, put more effort into their studies, and are less likely to suffer from mental distress. In terms of support from a student's environment, we find no evidence that the ordinal rank is related to support from teachers, parents and friends. Finally, while we have no information on the type of college to which students are admitted, we can exclude the notion that the effect is purely driven by selective college admissions. When we run our baseline model and additionally control for GPA, the effect of the ability rank on educational attainment remains large and statistically significant, indicating that GPA-based college admissions explain only a fraction of the effect.

This paper contributes to three strands of the literature. First, it extends the growing literature on ordinal rank and education outcomes. A large body of literature in psychology focuses on a student's academic self-concept, showing that students with a higher ordinal rank have a higher perceived ability in various school subjects (Marsh, 1987; Marsh *et al.*, 2007). Several recent studies in economics have stressed the impact of the ordinal rank in education on effort provision (Goulas & Megalokonomou, 2015), test scores (Murphy & Weinhardt, 2014) and disruptive behaviors (Cicala *et al.*, 2015; Elsner & Isphording, 2015).[1] The paper most closely related to ours is by Murphy & Weinhardt (2014), who present the first rigorous causal estimate of ordinal rank on educational performance. Based on administrative school data from England, they find a strong positive impact of ordinal rank in primary school test scores on standardized test scores in secondary school. Our paper is similar in terms of research design, and reinforces their conclusions. Nonetheless, our paper extends their findings along three important dimensions: first, our results show that the ordinal rank in school has a long-run impact that goes beyond compulsory schooling; second, we show that the ordinal rank in school affects actual career choices rather than test scores; and third, we show that career expectations seem to be an important channel through which the ordinal rank affects educational choices.

More broadly, this paper speaks to the literature on peer effects in education. Thus far, there is no consensus concerning whether and to what extent peers matter for student performance. While earlier studies have found that higher peer quality unambiguously has a positive impact on test scores and affects later education choices, more recent studies show that peer effects are non-linear and can even be negative for some students.[2] The ordinal rank effect found in this paper provides one explanation for these ambiguous effects. The positive effect of having better peers, or being in a homogeneous cohort, can be offset by having a lower ordinal rank.

This paper also contributes to the literature on imperfect information and educational choices. The evidence shows that students have imperfect knowledge of their own ability (Stinebrickner & Stinebrickner, 2012, 2014; Zafar, 2011; Bobba & Frisancho, 2014) and are uncertain about their returns to education (Jensen, 2010; Attanasio & Kaufmann, 2015; Wiswall & Zafar, 2015). Our results suggest that the ordinal rank is one of the reasons why students have incorrect beliefs about their absolute ability. In their education decisions, students seemingly place substantial weight on their relative ability, which in turn leads to suboptimal education choices.

[1] Related to findings about the impact of ordinal rank on various outcomes is the work of Azmat & Iriberri (2010) and Azmat *et al.* (2015), showing that a student's relative position in a class, i.e. being above or below the mean, affects effort provision. Jonsson & Mood (2008) find that relative position matters for track choice in secondary school.

[2] The evidence for positive peer effects on student performance ranges from primary schools (Hanushek *et al.*, 2003; Ammermueller & Pischke, 2009) to high schools (Calvó-Armengol *et al.*, 2009; Imberman *et al.*, 2012) and college (Sacerdote, 2001; Zimmerman, 2003; Carrell *et al.*, 2009; De Giorgi & Pellizzari, 2014; Booij *et al.*, 2015). Bifulco *et al.* (2011) and Patacchini *et al.* (2012) show that having better peers also increases the likelihood of attending college. Studies that find a non-linear effect or zero effect are Lavy *et al.* (2012), Koppensteiner (2012), Carrell *et al.* (2013), Burke & Sass (2013), Pop-Eleches & Urquiola (2013), Abdulkadiroglu *et al.* (2014), Feld & Zölitz (2017), Tincani (2015) and Tatsi (2015).

2 Data and Descriptive Statistics

2.1 The AddHealth data

Our primary data source is the restricted-use version of AddHealth, a representative longitudinal dataset of US middle and high schools. Four features of AddHealth are key to our study. First, it covers multiple cohorts within the same school, which is critical for identification, because it allows us to compare students in adjacent cohorts within the same school and exclude selection into schools as a main confounding factor. Second, within every school cohort, we observe a representative sample of students from which we can construct the ability ranking. Third, the longitudinal setup allows us to link the ordinal rank in high school to outcomes 14 years later and observe the critical transition from high school to college, as well as the completion of college. Finally, the survey includes a standardized test that provides us with an objective measure of cognitive ability. Unlike with most other datasets, we can directly measure cognitive ability without having to resort to grades or other self-reported measures as proxies.

To date, four waves of AddHealth are available. The first wave was administered in 1994/1995, when students were between 13 and 18 years old. Follow-ups were run in 1996, in 2000/2001 when most students had left high school, and in 2008/2009, when most had entered the labor market. In the first wave, a representative sample of schools was drawn among all public and private high schools in the US. Within each school, students from grades 7-12 were sampled. Overall, we observe up to six cohorts within a school. All cohorts were interviewed at the same time, such that we only observe each cohort in one grade, i.e. we observe the 1994 entry cohort in grade 7, the 1993 entry cohort in grade 8, the 1992 entry cohort in grade 9, etc. Therefore, we use the terms *cohort* and *grade* interchangably.[3]

The first wave comprised two questionnaires: a basic 'In-school' questionnaire, which was administered to all students in the surveyed schools; and a more comprehensive 'in-home' questionnaire, which was administered to a randomly drawn sub-sample of students within each school. For the in-home sample, 17 boys and 17 girls were randomly drawn from each grade within each school. Additional students were drawn to over-sample groups with certain characteristics: twins, students with disabilities, Blacks from well-educated families, as well as students of Chinese, Cuban and Puerto Rican origin. On average, we observe 22% of a school cohort. This stratification was the general sampling rule applied to most schools. There are some exceptions; for example, schools with grades featuring fewer than 34 students, single-sex schools, and, most importantly, 16 *saturated schools*, from which all students who were present on the day of the survey were included in the in-home sample.[4]

The core of our dataset is the in-home sample of wave I, which we complement with information on educational attainment from wave IV. We drop from the sample all schools with

[3] In schools that integrate high and middle schools and offer grades 7 to 12, all grades were sampled. In high schools that only offer grades 9-12, grades 7 and 8 were sampled from a middle school (so-called *feeder school*) randomly drawn from all surrounding middle schools that send students to the given high school. For further information on the study design and the sampling, see Harris (2009) and Harris *et al.* (2009).

[4] See Appendix D.1 for a more detailed description of the sample design.

20 observations or less and all grades with five students or less. Moreover, due to attrition, we drop all students for whom we do not observe the educational attainment (finished high school, attended any type of college, completed college) or other observable characteristics in wave IV. Our final sample comprises 13,645 students in 130 schools and 432 school-cohort combinations.

2.2 Outcome variables: Educational attainment

We consider three outcome variables that measure different degrees of educational attainment: *completed high school*, *attended college*, *completed a four-year college degree*. These measures are taken from wave IV of AddHealth, where respondents were asked about their highest educational attainment. The categories *attended college* and *completed a four-year college degree* are nested; *completed a four-year college degree* only includes students who completed at least a Bachelor's degree, while *attended college* is broader and also includes students who attended college but finished with less than a Bachelor's, or did not finish at all. Table 1 summarizes the outcome variables for various groups. Among all students, 93% completed high school, while 67% attended college. Around half of those who attended college finished at least with a Bachelor's degree.[5]

The educational attainment considerably differs across sub-groups. In all three measures, women have a higher educational attainment than men. The data also reveal a high correlation between the educational attainment of the parents and their children, whereby children of college-educated parents are four times as likely to complete a college degree and ten times less likely to drop out of high school than children whose parents are high school dropouts. There is less variation in the educational attainment across ethnic groups. Hispanics and Blacks have lower educational attainment than Whites, although the raw differences are less than 13 percentage points. An exception is students of Asian descent, whose educational attainment is considerably higher than that of all other groups.

Finally, we consider schools with different average ability and heterogeneity. Unsurprisingly, students from schools with an above-average mean ability have a higher educational attainment. We also check whether more heterogeneous schools are more or less conducive to educational success. If schools are homogeneous with respect to ability, for example, due to tracking or neighborhood segregation, one would expect homogeneous schools to have different outcomes from heterogeneous schools. However, the raw data do not support this conjecture.

2.3 Ranking students

Our regressor of interest is a student's ordinal rank in the ability distribution of a high school cohort. We are interested in estimating a causal effect of rank on educational outcomes, which we identify by comparing students in the same school with the same level of absolute ability who differ in their ordinal rank because they face a different ability distribution in their cohort.

[5] These numbers confirm the representativeness of the survey, as they are very close to the means in the American Community Survey (ACS): 91% have completed high school, 64% attended any type of college, and 31% completed a four-year degree. These calculations are based on the 2007-2011 Public Use File of individuals born between 1976 and 1982 (US natives and immigrants who arrived before 1995).

Table 1: Educational attainment by group

Group	*completed High school* mean	(SD)	*attended College* mean	(SD)	*completed 4-year degree* mean	(SD)	*N*
All	0.93	(0.26)	0.67	(0.47)	0.33	(0.47)	13645
Male	0.91	(0.28)	0.63	(0.48)	0.29	(0.45)	6330
Female	0.94	(0.23)	0.71	(0.45)	0.37	(0.48)	7315
Parental background:							
Less than high-school	0.81	(0.39)	0.45	(0.50)	0.13	(0.33)	1957
High school	0.91	(0.29)	0.55	(0.50)	0.19	(0.39)	3399
Some college	0.94	(0.23)	0.68	(0.47)	0.28	(0.45)	3423
College	0.98	(0.14)	0.85	(0.36)	0.54	(0.50)	4866
Race/Ethnicity:							
White	0.94	(0.25)	0.69	(0.46)	0.36	(0.48)	7733
Asian	0.98	(0.15)	0.78	(0.42)	0.49	(0.50)	882
Hispanic	0.90	(0.30)	0.61	(0.49)	0.23	(0.42)	1961
Black	0.91	(0.28)	0.65	(0.48)	0.28	(0.45)	3069
Average school ability:							
High average ability (above median)	0.96	(0.20)	0.74	(0.44)	0.42	(0.49)	6730
Low average ability (below median)	0.90	(0.30)	0.61	(0.49)	0.24	(0.43)	6915
School heterogeneity (within-school SD in ability)							
High heterogeneity(above median)	0.92	(0.27)	0.67	(0.47)	0.32	(0.47)	7410
Low heterogeneity (below median)	0.93	(0.25)	0.68	(0.47)	0.34	(0.47)	6235

Notes: This table displays the share of students who completed high school, attended college and finished a four-year college degree. The outcome variables are measured in wave IV of AddHealth. Standard deviations are reported in parentheses. Parental background refers to the highest level of education among both parents. Average school ability is the average ability of the entire school, and above/below median refers to the school distribution, i.e. students in the "above median" group attend schools with an above-median ability level. The school heterogeneity is measured by the within-school standard deviation of ability.

This identification strategy requires a standardized ability test that makes students comparable across cohorts within the same school.

Measuring cognitive ability To obtain a standardized measure of cognitive ability, we use the scores of a standardized Peabody Picture Vocabulary Test (PPVT-R; Dunn & Dunn (2007)), of which an abridged version was included in the survey. The Peabody test measures verbal intelligence and scholastic aptitude, and has been shown to strongly correlate with widely-used ability and intelligence tests, such as the Wechsler Adult Intelligence Scale (Anderson & Flax, 1968) or the Armed Forces Qualification Test (AFQT). Moreover, based on data from the NLSY79, it has been shown that Peabody scores are stable over time, i.e. a high score early in life strongly predicts a high score later in life (Baker *et al.*, 1993).

Within AddHealth, the test was administered individually to every student in the in-home sample. The test works as follows: participants are shown a panel with four pictures and are given a stimulus word, which they have to match to the picture that fits best. The test proceeds through a maximum of 87 rounds with increasing difficulty. The test is age-specific, with scores standardized to mean 100 and standard deviation 15 within an age cohort.[6] The scores are computed automatically, without being made available to the interviewer or the respondent. Appendix B provides a more detailed description of the testing procedure.

The ability rank We compute the ordinal rank based on a student's rank position in the ability distribution of his/her school cohort. To make the rank comparable across school cohorts of different size, we convert the absolute ordinal rank (i.e. 1,2,3,...) into a percentile rank, which assigns value 0 to the lowest, and value 1 to the highest-ranked student in a school cohort, and assigns all ranks in between according to the formula[7]

$$\text{percentile rank} = \frac{\text{absolute rank} - 1}{\text{nr of students in school cohort} - 1}. \tag{1}$$

Figure 1 shows that a given level of absolute ability is related to a substantial degree of variation in the within-cohort rank. Each box plot displays the distribution of the local within-cohort rank for a particular decile in the global ability distribution. The variation in the local rank is strongest in the middle of the global distribution, and smaller at the upper and lower end. Students at the upper end of the global ability distribution are most likely at the top of the ability distribution of any school cohort, while students with a low global rank are most likely at the low end of any school cohort. By contrast, students in the fifth or sixth decile of the global distribution can end up virtually anywhere in the local ability distribution, depending on the ability of their peers.

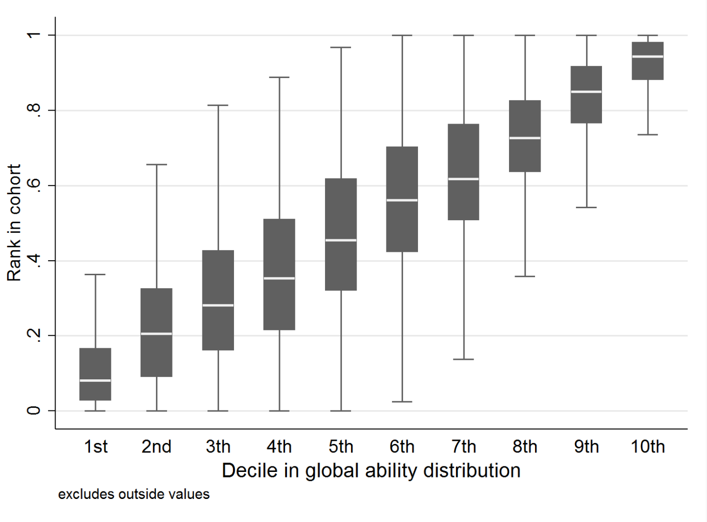
While the variation illustrated in Figure 1 reflects differences in rank within and across schools, our identification relies on variation *within* schools. To further quantify the variation in the rank for a given level of ability, we compute the within-school variance in the rank, conditional on absolute ability and cohort dummies, finding a standard deviation of 0.107.[8] Given that the rank is bound between 0 and 1, this suggests that there is substantial variation in the ordinal rank within a school across cohorts. The same ability can lead to very different ordinal ranks depending on the cohort composition. At the mean school cohort size of 184 students, for a given ability, the rank varies on average by 18 absolute rank positions.

6 Given the standardization within an age cohort, the Peabody score represents a student's rank in the entire age cohort in the population. The scores are comparable across age groups, as long as the ranks in the population are stable. This has been shown to be the case by Baker *et al.* (1993).

7 We first compute the absolute ranks. If a school cohort has 100 students, the student with the highest ability is assigned rank 100, and the student with the lowest ability is assigned rank 1. Students with the same test score are assigned an equal rank. We use a standardized rank measure because we only observe a random sample from each school cohort, such that the absolute rank in the sample and population would not be comparable. The absolute rank would be more important if one was interested in particular rank positions, for example the first or the last place (Gill *et al.*, 2015; Kuziemko *et al.*, 2014, e.g.), whereas we are primarily interested in the average effect of ordinal rank.

8 We obtain these standard deviations by regressing the within-cohort rank on absolute ability a_{isc}, cohort dummies, and school dummies, $r_{isc} = \alpha + \beta a_{isc} + \delta_s + \delta_c$ and then taking the standard deviation of the residuals. In some specifications, we use school-by-cohort fixed effects instead of separate school and cohort dummies. In this case, the residual standard deviation of rank is 0.0998.

Figure 1: Global vs. local rank



Note: This figure illustrates the relationship between the rank in the global ability distribution and the local rank in a school cohort. For each decile of the global ability distribution, the box-whisker plots display the median (white line), 25th and 75th percentile (lower and upper bound of grey box), as well as the minimum and maximum of the local rank distribution.

Discussion of rank variable Ranking students according to their cognitive ability offers the advantage that cognitive ability is formed early in life and remains stable after the age of ten (Jensen, 1998; Cunha *et al.*, 2006). At age 16, which is the average age of adolescents when taking the test, cognitive ability can be seen as predetermined. Under this maintained assumption, the Peabody test scores should not be immediately influenced by school inputs, peer influences, parental investments, a student's effort in school or her ability to learn. This alleviates several important concerns about unobserved factors that may simultaneously affect the ordinal rank and educational outcomes. One concern is that both could be determined by unobserved intentions to attend college. A student with a high motivation to attend college may work harder during high school, which in turn would manifest itself in a higher ordinal rank and a higher likelihood of attending college. A second concern is that the ordinal rank and later outcomes are both driven by tracking within schools. If students take higher-level courses, they may achieve higher test scores and have a higher likelihood of attending college. With cognitive ability being predetermined, the ability rank is unlikely to be influenced by intentions to attend college nor by tracking during high school. If anything, tracking is a channel through which a higher ability rank translates into better outcomes: students with a high rank in their cohort have a higher likelihood of taking higher-level courses, which increases their chances when applying for colleges.[9] A further concern is that the student's environment, for example parental pressure

[9] Another form of tracking occurs across schools, when children with higher test scores are sent to higher-track schools. This will not be a concern, because we only exploit variation *within* schools. For a more detailed description of ability tracking and its underlying mechanisms, see Betts (2011) and Fu & Mehta (2015).

or peer pressure, could affect the test scores. It should be noted, though, that the Peabody test is that it provides no particular incentive for students to attain a high rank, for parents to exert competitive pressure on their children, or for students to put pressure on their peers, because neither students nor their parents learn the test scores. In addition, students take the test individually, making social interactions during the test impossible. It is therefore unlikely that our ordinal rank measure reflects parents' competitive pressure or students' competitive nature rather than their relative ability.[10]

To the extent that students do not learn their test scores, the question remains whether students know how their own ability compares to the ability of others in the same cohort. While we cannot directly infer students' knowledge of their exact ranks from the survey, we have two pieces of evidence showing that students have knowledge about their relative ability in their cohort. First, students with the same absolute ability but a higher ordinal rank have a higher perceived intelligence. When asked if they believe that they are more intelligent than the average, students of a higher rank are significantly more likely to agree than students of lower rank, after controlling for own ability, personal characteristics and only comparing students within schools across cohorts. A second piece of evidence is that students of higher rank have higher expectations about their educational career; for example, they are more likely to expect to finish a college degree later in life. This would hardly be the case if students were unaware of their relative ability.

Cognitive ability is certainly not the only measure based on which we could rank students. An alternative measure would be the GPA, whereby both measure different things: a ranking based on GPA measures relative achievement, while a ranking based on the Peabody test measures relative cognitive ability. While the GPA is more salient than an ability test score, it has several severe limitations that make it less suitable as a base for the ordinal rank. One limitation of GPA is that achievement is determined by many — often unobserved — factors. Cognitive ability is one of these factors,[11] while other factors include effort, motivation, parental investments and school inputs. A failure to control for these unobserved factors would cast doubt on the causal interpretation of the effect of relative achievement on educational outcomes. A further limitation is that grades often depend on the grading scheme applied by the school. Exams are often neither standardized nor graded independently, but are rather *graded on a curve*, which means that teachers apply the same grade distribution to every exam regardless of students' absolute achievement or the average cohort ability. Therefore, we view GPA as a channel through which a student's cognitive ability affects educational choices and outcomes later in life. Smarter children obtain a higher GPA, which makes them more likely to be admitted to college. A third limitation specific to our identification strategy is that GPAs are hardly comparable across grade levels within a school. An 'A' in math in grade 7 does not reflect the

[10] A concern with the Peabody test being a low-stakes test is that students with higher non-cognitive skills may take the test more seriously and thus achieve a higher rank. We address this concern in a robustness check in Section E.

[11] In fact, as shown by Jensen (1998, p. 289), cognitive ability and intelligence are weak predictors of educational achievement.

same achievement as an 'A' in math in grade 12, nor does it reflect the same level of absolute ability. By contrast, the Peabody test has none of these limitations. The test is standardized, independently graded, comparable across grade levels within a school and reflects a trait that is predetermined at the time students take the test.

Sampling variation in the ability rank Given that in the average school cohort we only observe a random sample of 22% of all students, the ordinal rank is measured with error, because the rank in the sample differs from the true rank in the population. On average, the error should be zero, because we attribute to some students a rank that is too high and to others a rank that is too low. However, what matters for measurement error is the sampling variation, i.e. the average deviation of the rank in the sample from the true rank in the population. To quantify the sampling variation in the ordinal rank, we exploit the fact that several large schools were sampled entirely (*saturated schools*) and apply a bootstrap procedure that replicates the stratified random sampling and allows us to compare the true rank in the population of a school cohort to the rank in the random sample. Based on 10,000 random samples from two large saturated schools,[12] we find a standard deviation of $sd = 0.017$, which means that in a school cohort of 100 students, we misclassify every student's rank on average by 1.7 absolute rank positions, or at the average cohort size of 184 students, we misclassify every student by 3.1 absolute rank positions.[13] When compared to the within-school standard deviation of the ordinal rank, $sd = 0.107$, the sampling variation is too small to explain the entire variation in the ordinal rank. Nonetheless, in a linear regression, the resulting measurement error leads to biased estimates. Along with the main results, we will later discuss the direction of this bias and quantify its magnitude in a Monte Carlo experiment.

2.4 Summary statistics

Panel A in Table 2 displays the summary statistics of the main variables of interest. The two columns on the right display the means for students in the bottom and top half of the within-cohort ability distribution. Upon first glance, women and Blacks are over-represented among students in the bottom half of a school cohort, while there is no large difference with respect to average age, the share of Hispanics, students of Asian descent or students with a migration background. A strong correlation appears between ability and parental education. Children of highly educated parents are more likely to have a higher rank within their grade. This suggests that demographic groups self-select differently into schools, which warrants controlling for individual characteristics in the regression.

Panel B summarizes the average school and grade characteristics. Schools considerably differ in terms of average ability and heterogeneity. Students in the lowest ability school scored on average 79 on the standardized test, which is three between-school standard deviations below

[12] We only use two schools here, because the remaining 14 of the 16 saturated schools are among the smallest schools in the sample.
[13] In Appendix D.1, we provide more detailed information on the bootstrap procedure.

the mean, whereas the highest ability school scored 116, or 2.5 between-school standard deviations above the mean. To measure heterogeneity in ability, we take the within-school standard deviation of the ability distribution, which varies between 9.2 and 20.5, and is on average twice as large as the between-school standard deviation, which is 6.5.

Depending on the school, the grade size strongly varies: in the population, it ranges from 5 in the smallest grade to 645 in the largest. More relevant for our study is the actual within-grade sample size. The average grade has 40 students in the sample, which is more than the 34 students drawn at random due to over-sampling of minorities and the inclusion of saturated schools.

Table 2: Summary Statistics of the main variables

		All		bottom 50%	top 50%
Variable	*N*	*Mean*	*SD*	*Mean*	*Mean*
A. Individual characteristics					
Cognitive ability	13645	101.14	14.24	91.18	110.84
Age	13645	16.13	1.68	16.25	16.01
Female	13645	0.54	0.50	0.57	0.50
Ever repeated a grade	13645	0.20	0.40	0.28	0.13
Migration background (1st & 2nd gen.)	13645	0.15	0.36	0.16	0.14
Asian descent	13645	0.06	0.25	0.06	0.07
Black	13645	0.22	0.42	0.27	0.19
Hispanic ancestry	13645	0.14	0.35	0.16	0.13
Highest parental education					
Less than high-school	13645	0.14	0.35	0.19	0.10
High-school	13645	0.25	0.43	0.29	0.21
Some college	13645	0.25	0.43	0.24	0.26
College	13645	0.36	0.48	0.29	0.42
B. School and grade characteristics					
School characteristics	*N*	*Mean*	*SD*	*Min*	*Max*
Small (< 401 students)	130	0.22	0.42		
Medium (401-1000 students)	130	0.47	0.50		
Large (> 1000 students)	130	0.31	0.46		
Mean ability	130	100.31	6.46	79.19	115.80
SD ability	130	12.89	2.29	9.24	20.48
Grade characteristics					
Grade size (population)	432	184.50	131.67	5	645
Nr students in sample	460	39.69	44.05	6	545

Notes: Panel A displays the means and standard deviations of the main variables for the entire sample, as well as the means for the students above and below the median ability of their school cohort. Besides the share of students of Asian descent, all differences are statistically significant at the 1%-level. Panel B displays the average school and school-grade characteristics. The school characteristics were reported by the school administrator in a separate survey. When computing the averages, in the case of school characteristics, every school has equal weight, whereas in the case of school-grade characteristics, every school-grade combination has equal weight.

3 Identification and Estimation Strategy

3.1 Identifying variation

Before turning to the empirical strategy, we first describe the identifying variation. To estimate a causal effect, we exploit idiosyncratic variation in cohort composition within schools over time. The idea is that students with the same level of absolute ability have a different rank if they are in cohorts with different ability distributions. The variation in cohort composition can be due to differences in mean ability — some cohorts are on average brighter than others — or differences in the variance or higher moments of the ability distribution. Figure 2 illustrates the identifying variation with respect to the mean and variance of ability, based on a stylized example of two entry cohorts in the same school. In Panel A, the 1994 entry cohort has a lower average ability than the 1995 entry cohort, such that a student with cognitive ability *abil* would have the second rank. If she entered the school in 1995, when the entry cohort was stronger, she would only have the third rank, despite having the same cognitive ability. In Panel B, both cohorts have the same mean ability, although the 1994 cohort has a higher variance than the 1995 cohort. In this case, a student with ability *abil* would have the second rank in 1994 but the first rank in 1995.

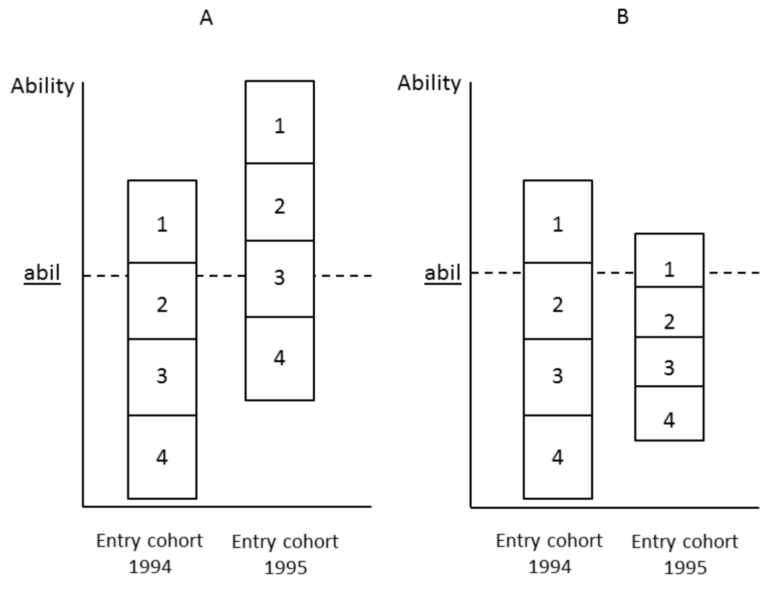


Figure 2: Identifying variation: difference in means (A), difference in variance (B)

Notes: Both panels illustrate how differences in the cohort composition lead to variation in the rank for a given level of ability *abil*. The numbers 1-4 refer to quartiles in the ability distribution. In Panel A, both cohorts differ in their mean ability, while in Panel B, both differ in their variance of ability but have the same mean.

By only using variation within schools, we can rule out the notion that the variation in cohort composition is driven by the systematic self-selection of students into schools.[14] Within a school, the fluctuations in the ability distribution across cohorts can be considered idiosyncratic when they are determined by the timing of births in combination with a cut-off date between school years. As explained by Hoxby (2000), if more children are born before the age cut-off in some years than in others, this leads to fluctuations in the cohort sizes within the same school. Along the same lines, the characteristics of parents may fluctuate from year to year. In some years, the share of children born to highly educated parents is higher than in others, the share of Black or Hispanic children is higher than in others, or in some years more children with a higher innate ability are born than in others. Due to the law of large numbers, such fluctuations may not be pronounced in the entire US but are more pronounced within a school catchment area, where the law of large numbers does not necessarily hold. Our identification relies on this idiosyncratic variation in the population.

3.2 Empirical model and identification

The empirical model explains a student's educational attainment as a function of her ordinal rank in high school,[15]

$$y_{isc} = \alpha + \gamma r_{isc} + g(a_{isc}) + \boldsymbol{X_{isc}\beta} + \delta_c + \lambda_{sc} + \varepsilon_{isc}. \tag{2}$$

The dependent variable y_{isc} is a binary indicator for the educational attainment of student i in high school s in cohort/grade c. The regressor of interest is $r_{isc} \in [0, 1]$, the student's ordinal rank within a school cohort. We control for absolute ability, a_{isc}, with a fourth-order polynomial of the Peabody score. $\boldsymbol{X_{isc}}$ is a vector of individual characteristics, which includes age in months, gender, dummies for race or ethnicity (Asian, Black, Hispanic), a dummy for migration background (1 if a person is a first- or a second-generation migrant), dummies for the highest level of education between both parents (less than high school, high school, some college, college degree), dummies for the the highest occupational status between both parents (not working, blue collar, white collar low-skilled, white collar high-skilled) and a dummy that equals one if a student has ever repeated a grade up until wave I of the survey. The term δ_c represents a full set of cohort dummies controlling for average differences between grade levels 7-12, given that every cohort was sampled in a different grade level. The term λ_{sc} represents a set of school effects that may or may not vary by cohort. Unobserved determinants of educational attainment are summarized in the error term ε_{isc}. To allow for arbitrary within-school correlation in the error term, we cluster the standard errors at the school level.

Our coefficient of interest is γ, which measures the relationship between the ordinal rank and educational attainment. For γ to have a causal interpretation, we need to assume strict

[14] In fact, there is evidence that parents strategically choose schools with their childrens' rank in mind. Cullen *et al.* (2013) show that after automatic admission to the flagship state universities in Texas was granted to the top 10% of a school, parents deliberately sent their children to lower-ability schools to give them a higher chance of being in the top 10%.

[15] The model description in part follows Black *et al.* (2013).

exogeneity conditional on λ_{sc}; namely, $E(\varepsilon_{isc}|r_{isc}, g(a_{isc}), X_{isc}, \delta_c, \lambda_{sc}) = 0$. The plausibility of this assumption depends on the parametrization of the school effects λ_{sc}. We propose two variations of the above model: one with separate school and cohort fixed effects and one with school-by-cohort fixed effects.

Model with separate school and cohort fixed effects One possibility is to assume that the school effects do not vary by cohort, $\lambda_{sc} = \lambda_s$, and estimate a two-way fixed effect model with separate school and cohort fixed effects.[16] This model compares students in the same school, with the same observable characteristics and the same absolute ability, who have a different ordinal rank because they are in different cohorts. The coefficient γ would then be identified through differences in all moments of the ability distribution across cohorts within a school. The school fixed effects rule out selection into schools as a confounding factor, namely the notion that some schools attract on average more intelligent children than others. However, a causal interpretation of γ requires the assumption of strict exogeneity conditional on λ_s to hold. Consider an error term $\varepsilon_{isc} = \rho_{sc} + \eta_{isc}$ that is composed of a systematic school-cohort-specific component ρ_{sc} and a mean-zero random component η_{isc}. Strict exogeneity would require ρ_{sc} to be uncorrelated with all other regressors.

Without controls for school-cohort-specific factors, this assumption is most probably violated. In particular, ρ_{sc} includes the mean and variance of ability of a school-cohort, which have been shown to be important drivers of ability peer effects. The mean cohort ability is mechanically related with the individual within-cohort rank, because a student with a given absolute ability *mechanically* has a lower rank in a cohort with higher average ability, as illustrated in Figure 2a. Controlling for the mean cohort ability breaks this correlation and ensures that γ solely picks up the ordinal rank, rather than the cardinal difference between the test score of a particular student and the cohort mean. Moreover, recent literature suggests that the variance of ability generates peer effects. As shown by Tincani (2015), if students care about their rank, they exert a higher effort if the variance of ability is low, because the same additional effort leads to a greater improvement in one's own rank when the variance is low. By controlling for the mean and variance of ability, we would disentangle the impact of the ordinal rank from the direct impact of mean and variance.[17]

Even if the mean and variance of ability are controlled for, ρ_{sc} could reflect differences in cohort composition within the same school that would not be captured by separate school and cohort fixed effects. One example is dynamic selection into schools, whereby schools may become better or worse over time, such that older cohorts are systematically different from younger cohorts. Cohorts may also differ with respect to gender and ethnic composition, parental backgrounds, immigrant share or the share of disruptive students. Similarly, cohorts may differ in

[16] A model with separate school and cohort fixed effects is a two-way fixed effect model, because the cohort effects are not nested within schools.

[17] As highlighted by one referee, the rank effect could still pick up the direct impact of higher moments of the ability distribution, such as skewness and kurtosis, although it is more difficult to find an intuitive explanation for why this direct effect would occur.

the school inputs that they receive. These differences may be idiosyncratic — one cohort happens to have better teachers than another — or may respond to differences in composition — the cohort with more disruptive students receives more support staff. In addition, it has been shown that IQ can be affected by cohort composition and school inputs (Brinch & Galloway, 2012), which casts further doubt on the validity of the exogeneity assumption. These factors potentially affect the rank as well as educational attainment, and would prevent us from obtaining an unbiased estimate for γ. We are able to control for some observable differences across cohorts, for example the gender composition or the immigrant share, but many other — potentially important — confounders are unobservable to us.

A model with school-by-cohort fixed effects A further possibility is to treat ρ_{sc} as a nuisance parameter and estimate a model with school-specific cohort fixed effects λ_{sc}, which absorb all mean differences between school cohorts, thus alleviating any of the endogeneity concerns discussed in the previous paragraph.[18] In this model, the cohort dummies δ_c would be omitted, because they cannot be separately identified from λ_{sc}. This model is equivalent to a within-transformation of all variables at the school-cohort level, that is, from each variable the school-cohort mean is subtracted. We therefore no longer compare students within the same school across different cohorts, but rather compare students across all school cohorts after having removed all observable and unobservable mean differences between school cohorts. Because the ordinal rank varies at the individual level, the rank effect γ can still be identified from differences in the shape of the ability distribution. As long as the school cohorts differ in the variance, skewness, kurtosis, and potentially even higher moments of the ability distribution, a student with a given absolute ability has a different ordinal ability rank in the respective school cohort. Panel B in Figure 2 illustrates the variation in the rank for two school cohorts a different variance but the same mean. In terms of estimation, this model is much more demanding because 432 fixed effects need to be estimated, compared to 130 school and 6 cohort fixed effects in a model with separate school and cohort fixed effects.

Validation of the identifying assumption The main advantage of the model with school-by-cohort fixed effects is that all confounders at the school-cohort level are ruled out. However, this comes at the cost of treating any difference in the shape of the ability distribution as a nuisance parameter, and prevents us from exploring the determinants of the identifying variation at the school-cohort level. When school-cohort effects are differenced out, we are unable to determine whether differences in the ordinal rank are driven by differences in the mean or the variance of ability.

Whether the less restrictive two-way fixed effect model permits a causal interpretation of γ depends on the support for the identifying assumption in the data, namely whether the school-by-cohort effects in the data-generating process are uncorrelated with all other regressors. To

[18] In addition, the school-by-cohort fixed effects absorb unobservable group shocks, a source of bias that is pervasive in linear-in-means peer effects estimates, as pointed out by Angrist (2014, p.5)

test whether this assumption is supported by the data, we perform a Hausman test, in which we test the null hypothesis of no correlation against the alternative hypothesis of there being a correlation. The test compares the estimates of a model with school-by-cohort random effects to those of a model with school-by-cohort fixed effects. If both produce the same results, the Hausman test fails to reject the null hypothesis of no correlation, in which case the estimates of a two-way fixed effects model could be interpreted as causal. We perform two sets of Hausman tests: one that includes all the regressors in Equation 2 and one in which we additionally control for the cohort mean and variance of ability. The results are displayed in Table 6 in Appendix C.[19] The test results depend on the outcome variable. For college attendance, we reject the null of no correlation in both cases, while for college completion, we cannot reject the null at the 5% significance level. For high school completion, we fail to reject the null in a model without controls for mean and variance, but reject it once these variables are included.

These results suggest that we have to be cautious when interpreting the results from a two-way model as causal, and rather rely on a model with school-by-cohort fixed effects. Seemingly, there are some systematic differences across school cohorts that are correlated with the other regressors. These are absorbed in the model with school-by-cohort fixed effects. Nevertheless, in the empirical analysis to follow, we will present results from both approaches. As our most conservative specification, the model with school-by-cohort fixed effects will serve as the benchmark model, although for comparison we will also present the results from a two-way fixed effects model with various sets of controls.

Further threats to identification Besides the threats to identification mentioned above, there are several more. First, parents may strategically delay their child's school entry, which would violate the identifying assumption that being in one cohort or another is as good as random. Second, the estimates could be biased by selective attrition, because students of low ability are potentially more likely to drop out of the sample. Third, the initial rank in high school may affect the ability test scores measured in higher grade levels. Fourth, there may be individual-level confounders that affect the rank through their impact on the Peabody test, while at the same time affecting educational attainment, for example students' risky behaviors, or personality traits. We will address these threats in a series of robustness checks. A further source of bias is measurement error, which we will address in a Monte Carlo experiment.

[19] Because standard errors are clustered at the school level, the assumptions of a standard Hausman test are not valid. Therefore, we apply the cluster-robust alternative to the standard Hausman test, as discussed in Wooldridge (2002, p.291).

4 Results

4.1 Baseline results: ordinal rank and educational attainment

We estimate various specifications of the linear probability model in Equation (2), with the results shown in Table 3.[20] Each coefficient is the result of a separate regressions of each of the three outcome variables — dummies for having completed high school, attended college and completed a four-year college degree — on the ordinal rank of a student in her high school cohort.

Table 3: OLS regression results: the importance of rank position

	(1)	(2)	(3)	(4)	(5)
Dependent variable					
Completed high school	0.054*	0.034	0.054*	0.059*	0.048
	(0.028)	(0.027)	(0.032)	(0.032)	(0.033)
Attended College	0.139***	0.103**	0.126***	0.122**	0.112**
	(0.045)	(0.041)	(0.047)	(0.048)	(0.049)
Completed 4-year degree	0.121**	0.101**	0.081	0.082	0.082
	(0.052)	(0.048)	(0.057)	(0.057)	(0.059)
Controls:					
Individual ability (quartic)	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	No
Cohort FE	Yes	Yes	Yes	Yes	No
Individual controls	No	Yes	Yes	Yes	Yes
Average cohort ability	No	No	Yes	Yes	No
Ability variance	No	No	No	Yes	No
School × Cohort FE	No	No	No	No	Yes
Goodness of fit:					
R^2 Completed high school	0.09	0.15	0.15	0.15	0.17
R^2 Attended College	0.15	0.23	0.23	0.23	0.25
R^2 Completed College	0.18	0.26	0.26	0.26	0.28

Note: This table displays the results of separate OLS regressions of the dependent variables *completed high school*, *attended college*, and *completed college* on the percentile rank. Each coefficient is the result of a separate regression. From left to right, more controls and fixed effects are being introduced. Standard errors, clustered at the school level, are displayed in parentheses, with significance levels * p<0.10, ** p<0.05, *** p<0.01.

The coefficients are best interpreted in relation to the within-school standard deviation in the ordinal rank, which equals 0.1, or one decile of the support of the ordinal rank, $r_{isc} \in [0, 1]$. In Column (1), we begin with a parsimonious model with separate school and cohort fixed effects, as well as without individual controls. A one decile increase in the within-cohort rank is associated with a 0.5 percentage point higher likelihood of completing high school, a 1.4 percentage point higher likelihood of attending college, as well as a 1.2 percentage point higher likelihood of completing a four-year college.

[20] We choose a linear probability model over logit and probit due to the large number of fixed effects, which would be difficult to estimate in a non-linear model.

In Column (2), when we add individual controls, all three coefficients are smaller and the coefficient for completing high school becomes statistically insignificant. The change in the size of the coefficients, as well as the increase in the R^2, underline the importance of including individual controls. The coefficients are smaller because some demographic groups are over- or under-represented at high ranks, as shown in the summary statistics in Table 2. For instance, students whose parents have a college degree are more concentrated in the top half of the within-cohort ability distribution, while Black students are more likely to be in the bottom half of a cohort. The individual controls account for these systematic differences in the ordinal rank.

In Column (3), we additionally control for the mean cohort ability, which has been identified as an important determinant of educational performance in the peer effects literature. Including the mean ability in the regression disentangles the impact of ordinal rank from linear peer effects. Moreover, it breaks the mechanical negative correlation between the ordinal rank and the mean cohort ability. The coefficient for high school completion is now larger than in Column (2), and statistically significant at the 10%-level. A one decile increase in the ordinal rank increases the probability of finishing high school by half a percentage point. The coefficient of college attendance is large and statistically significant. A one decile increase in the ordinal rank increases the likelihood of attending college by 1.26 percentage points. The effect on college completion is smaller, and imprecisely estimated.

In Column (4), we also control for the within-cohort variance in ability. Recent literature has highlighted the importance of variance for the strength of peer effects in high schools; for example, because students can more easily attain a higher rank in classrooms with a low variance (Tincani, 2015). The estimates from this specification are similar to those in Column (3). Including the variance adds no explanatory power to the model, suggesting that the direct effect of the within-cohort variance on educational attainment is small. Moreover, the point estimates remain almost identical, suggesting that omitting the variance from the regression would not bias the estimate of γ.

Finally, in Column (5), we present the estimates from our most conservative model with school-by-cohort fixed effects. The fixed effects absorb all school-cohort-specific variables, including the mean and variance of ability. Thus, we compare students with the same ability across all school cohorts, holding all school-cohort-specific influences constant. The coefficient of college attendance remains large and statistically significant. A one decile increase in the within-cohort rank increases the likelihood of attending college by 1.1 percentage points. The effects on high school completion and completion of a four-year degree are positive but imprecisely estimated.

The central finding that emerges from these results is a strong positive impact of the ordinal rank in high school on the likelihood of attending college. A one decile increase in the within-cohort rank increases the likelihood of going to college by more than one percentage point. This is a large effect, given that we compare students with the same cognitive ability, the same observable characteristics, who attend the same school but happen to be in cohorts with different ability distributions. For the other outcome variables, high school completion and completion of a four-year college degree, we also find large positive effects, although these are statistically insignificant in the more conservative specifications.

The fact that we find a large and statistically significant impact on the likelihood to attend yet a smaller and imprecisely estimated impact on the likelihood to complete college points to an influence of ordinal rank on educational choices rather than educational achievement. The decision to attend college depends, among other factors, on a person's expected returns to attending college, which may be influenced by one's relative ability in high school. Once a person enters college, she is surrounded by new peers, and presumably, for finishing a four-year degree, other factors like the relative ability among the new peers become more important.

These results suggest that the impact of ordinal rank on educational outcomes reflects a peer effect in its own right. Specifications (3) and (4) account for the most important peer effects highlighted in the traditional literature on linear peer effects (Sacerdote, 2011, e.g.), as well as the recent literature on non-linear peer effects (Tincani, 2015). The effect of ordinal rank on college attendance is robust to controls for mean and variance of ability. Moreover, Columns (3)-(5) make us more comfortable about the causal interpretation of the estimates from a two-way fixed effects model, once the mean and variance are controlled for. Although the Hausman test casts doubt on the causal interpretation of the results in Columns (2)-(4), it emerges that once we control for the mean cohort ability, the results are similar to those from a model with school-by-cohort fixed effects in Column (5).

The results reinforce the findings of Murphy & Weinhardt (2014) based on English administrative data, showing that the ordinal rank in achievement at the end of primary school has a significant positive effect on standardized test scores at the end of secondary school. Given the detail of their data in terms of number of students, schools and student-subject observations, they are able to exploit both the variation in rank for multiple cohorts within a school-subject combination and the variation in rank within student across subjects. They find that a one standard deviation increase in the within-school-cohort rank increases test scores three years later by between 5% and 10% of a standard deviation, which is a similar magnitude as the impact of having a teacher whose value-added is one standard deviation above the average. Our results confirm the importance of ordinal rank as an input in educational production. Moreover, our paper adds three important pieces of evidence on the impact of ordinal rank. First, we show that the ordinal rank in ability — as opposed to rank in achievement — is an additional dimension along which ordinal rank in school affects educational outcomes. Second, we show that a student's ordinal rank in school has long-term consequences that extend beyond compulsory education. Third, and related, our paper departs from the notion that test scores are the main outcome and shows that the ordinal rank in school affects actual education choices that have long-lasting consequences for individual careers.

4.2 Heterogeneous effects

In the previous section, we have shown that a student's ordinal rank has a robust positive effect on educational attainment — on average. In Figure 3, we explore whether this effect differs by cohort size, peer quality and peer heterogeneity, as well as whether it differs at the bottom

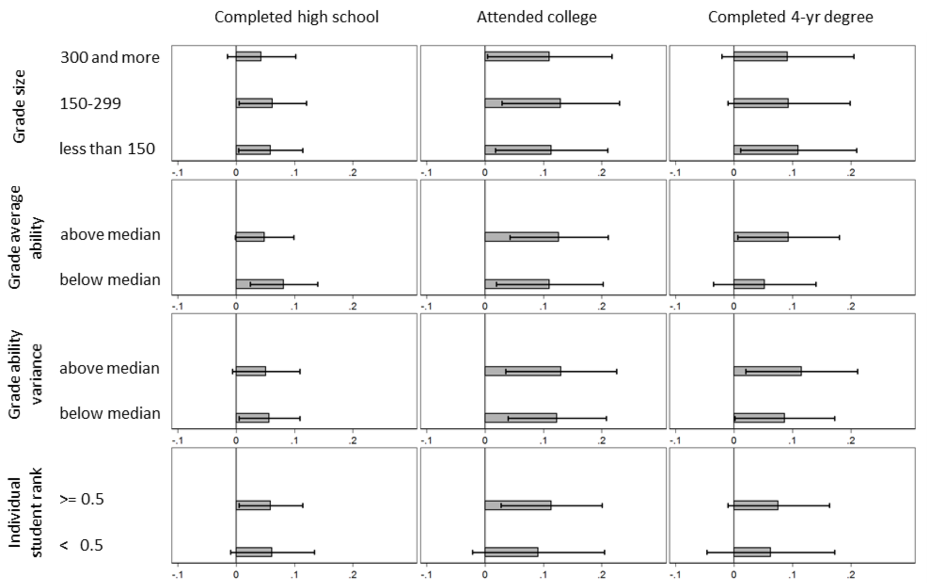


Figure 3: OLS results: heterogeneous effects

Note: The first three rows display the marginal effects of rank on educational attainment, obtained from the estimation of Equation (2) with school-by-cohort fixed effects, and the rank interacted with dummies for the respective groups. The fourth row displays the marginal effects based on an interaction of rank and a dummy for being in the top half of a school cohort, obtained from the estimation of Equation 2 with separate school and cohort fixed effects, and controls for mean and variance of ability.

and the top of the within-cohort ability distribution. Each graph displays the point estimates and 95%-confidence intervals for the sub-sample listed on the left. The results are based on the estimation of Equation (2) with school-by-cohort fixed effects and an interaction of the ordinal rank with dummies for the respective groups.

The first row shows the effect for different school cohort sizes. The rank effect could differ by cohort size because students in small cohorts may be more aware of their rank and thus respond more to differences in the rank. We split the sample in three equally sized bins. For all three outcome variables, the point estimates are similar across cohort sizes. The effect of rank is statistically significant for all outcome variables in small cohorts, and insignificant in large cohorts.

In the second row, we explore whether the ordinal rank has a stronger impact in school cohorts with a higher average ability. This could be the case if there was a greater sense of competition in schools with a high average ability. To test for this difference, we split the sample into school cohorts with an average ability above and below the sample median. The point estimates indicate that the impact of rank on high school completion is slightly larger in low-ability cohorts, while for the completion of a four-year college degree, the impact is larger in high-ability cohorts. For college attendance, the point estimates are almost equal. In terms of statistical significance, none of these coefficients are different from each other.

In the third row, we analyze the difference in the effect between cohorts with high and low variance in ability. Our prior, shaped by Carrell *et al.* (2013) and Tincani (2015), is that the rank effect should be more pronounced in cohorts with a low variance, because it is easier for students to attain a higher rank. The point estimates do not support this prior, given that the estimates are similar in cohorts with a high and low variance in ability. This is in line with the effects shown in Columns (3) and (4) of Table 3, where controlling for the variance of ability has hardly any influence on the point estimates.

Finally, in the fourth row, we analyze whether the rank effect differs between the bottom and top half of the within-cohort ability distribution. A priori, we would expect the rank to matter more at the top than at the bottom, because the difference between the 10th and the 20th rank in a cohort of 100 students is presumably more important for education decisions than the difference between the 70th and 80th rank. In the linear specification in Equation (2), both differences would receive equal weight. In the fourth row of Figure 3, we interact the ordinal rank with a dummy for being in the top half of a cohort, and estimate Equation (2) with separate school and cohort fixed effects and controls for mean and variance of ability. The point estimates of the marginal effects are similar in the top and bottom of the ability distribution. These results are in line with Murphy & Weinhardt (2014), who show that the impact of ordinal rank within schools is approximately linear along the test score distribution.

In general, the heterogeneous effects have to be interpreted with some caution, because we lack the statistical power to carry out a more fine-grained analysis owing to the sample size. Nonetheless, the similar point estimates indicate that the effects of ordinal rank do not differ greatly across school size, cohorts with different ability levels and variance, as well as along the ability distribution within a cohort.

4.3 Measurement error in the ordinal rank

The stratified sampling of AddHealth introduces measurement error in the ordinal rank. Specifically, measurement error may arise from three sources: the over-sampling of minorities, the stratification by gender and the random sampling of students within a school cohort. In the following, we quantify the bias resulting from these three types of measurement error.

Random sampling within school cohorts The most severe source of measurement error is random sampling within school cohorts. AddHealth only samples on average 22% of all students from a school cohort, which, as shown in Section 2.3, introduces measurement error in the ordinal rank. Given that the sampling within a school cohort is random, it introduces a classical measurement error, leading to a downward bias in the estimated coefficients.

To verify the direction of the bias and quantify its magnitude, we carry out a series of Monte Carlo experiments based on artificial school data with similar properties as AddHealth. We first construct the population school data, assuming a data-generating process in which the ordinal rank has a positive impact on the outcome. From each school cohort in this dataset, we draw a random sample of 40 students and run the model of Equation (2) with school-by-cohort fixed effects. We repeat this procedure 1,000 times and compare the average estimated coefficient from the random sample to the coefficient in the data-generating process. To account for the fact that AddHealth samples a constant number of students from school cohorts of varying size, we run the same experiment for school cohort sizes between 40 and 400 students.

The results confirm that measurement error attenuates the estimates. With an average school cohort size of 184 students, we under-estimate the true effect by around 20%. Moreover, the bias levels off as the cohort sizes become larger. At a cohort size of 400, we would under-estimate the true effect by 25%. In Appendix D.3, we provide a more detailed account of the stratified random sampling, the resulting sampling variation and the Monte Carlo experiment.

Over-sampling of minorities The over-sampling of minorities introduces measurement error if minority students have a lower average ability than white Americans in the same school cohort, in which case white Americans would be assigned a higher ordinal rank than under random sampling. To assess the bias resulting from this type of measurement error, we exploit the sequencing of the sampling in AddHealth and the precise documentation of the sampling in the in-home sample. The in-home sample includes an indicator that flags students that have been sampled randomly at first and minority students that have been added later, such that we observe for each student in the sample the rank with and without over-sampling. The correlation in the percentile ranks in both samples is almost perfect (corr=0.9867), indicating that measurement error from over-sampling is not a concern.

Gender stratification Gender stratification could introduce measurement error, because within each school cohort equal numbers of boys and girls were drawn, regardless of the underlying gender distribution in the population of a school cohort. This is problematic if the underlying gender distribution is skewed. For instance, take a cohort of 80 males and 20 females. If we drew 17 male and 17 female students from this cohort, we would sample 85% of all females but only 21% of all males in a cohort. A priori, it is unclear in what direction this measurement error would bias the estimates. To quantify the bias, we compare the estimates based on the full in-home sample to the estimates obtained from a sub-sample of school cohorts with a fairly even gender balance in the underlying population. We include in the sub-sample all school cohorts in which the share of girls in the population ranges between 40% and 60%.[21] The results, displayed in Row 4) of Table 4, suggest that stratified sampling indeed introduces measurement error and leads to a downward bias in the estimates: the effects are around 20% larger for the sub-sample compared to the sample with all school cohorts.

4.4 Further robustness checks

Thus far, we have interpreted the effect of ordinal rank on educational attainment as causal, given that the school-by-cohort fixed effects rule out all possible confounding factors at the school-cohort level, as well as based on the assumption that being in one school cohort or another is as good as random. In this section, we discuss scenarios under which this assumption may not hold and provide evidence that the results are robust to these threats to identification. Table 4 displays the results for these robustness checks. All specifications include individual controls, a quartic in absolute ability and school-by-cohort fixed effects. As a benchmark, Row 1) reproduces the results from Table 3, Column (6).

Are the results affected by strategic delay of school entry? The central identifying assumption is that, conditional on being in a particular school, being in a given cohort is as good as random. This assumption holds if the year of school entry is determined by a student's birthday. However, as shown by Deming & Dynarski (2008), academic redshirting — delaying school entry to allow children to mature for another year — is widespread in the US. Similarly, students could voluntarily repeat a grade, in which case they would be older than most other students in their cohort. To test whether the results are affected by these age differences, we restrict the sample to students whose birth date is sufficiently close to the cohort average whereby we can plausibly exclude redshirting or grade repetition. In Column 3) of Table 4, we restrict the sample to age bands of 0.4 years around the mean age of an entire age cohort. The results show that strategic delay is not a threat to our identification. Compared to the baseline, the effects on high school completion and college attendance are slightly smaller, while the effect

[21] We calculate these distributions from the in-school sample. Ideally, we would want to only include school cohorts in which the gender balance is exactly 50-50. However, given that no school cohort has exactly this distribution, and to have a sufficiently larger number of observations, we include all school grades with a gender balance close to 50-50 and set the boundaries at 40-60 and 60-40, which accounts for 71.4% of the sample.

Table 4: Robustness checks

	Dependent variable		
	Completed high school	*Attended college*	*Completed 4-year degree*
Regressor: percentile rank within school cohort			
1) Baseline estimates	0.048	0.112**	0.082
	(0.033)	(0.049)	(0.059)
2) Keep grades with female share 40-60%	0.061	0.147**	0.115*
	(0.037)	(0.059)	(0.068)
3) Keep if age = 0.4 ± mean age	0.037	0.106*	0.104
	(0.034)	(0.056)	(0.069)
4) Control for conscientiousness and neuroticism	0.045	0.099**	0.077
	(0.033)	(0.049)	(0.058)
5) Control for marijuana use and drinking	0.043	0.095*	0.059
	(0.034)	(0.050)	(0.059)
Alternative regressors			
6) Relative rank within gender group	0.038*	0.126***	0.067*
	(0.021)	(0.034)	(0.040)
7) Rank among those who completed high-school		0.124***	0.098*
		(0.046)	(0.056)

Note: This table displays estimation results for Equation (2). All regressions include individual controls and school-by-cohort fixed effects. In Rows 1)-5), the regressor of interest is the percentile rank within a school cohort. In Rows 6) and 7), the regressor of interest is the percentile rank within same-gender members of the same cohort (Row 6), and the percentile rank among all students who eventually finished high school (Row 7). Standard errors, clustered at the school level, are displayed in parentheses, with significance levels * p<0.10, ** p<0.05, *** p<0.01.

on college completion is moderately larger. The effect on college attendance remains statistically significant.

Selective attrition A further source of bias is selective attrition, which arises because 25% of all students attrit from the in-home sample between waves I and IV. This attrition may lead to biased estimates if it is correlated with the ordinal rank. We address this concern in a robustness check in Appendix E, in which we estimate Equation (2) with the attrition status as the dependent variable, finding no evidence of a systematic relationship between rank and attrition.

An additional type of attrition may have occurred before and up to wave I: high school dropouts. Because we observe each school cohort in a different grade, high school dropouts should induce a greater attrition in older rather than younger cohorts. To the extent that dropouts predominantly come from the low end of the within-cohort ability distribution, the ability distribution may shift to the right over time. To rule out this type of attrition as a concern, we carry out a robustness check in which we recalculate the rank measure using only students who have indicated in wave IV that they completed high school. The results, displayed in Row 6), are moderately larger than in the baseline, and college completion is now statistically significant at the 10%-level.[22] The difference in the results suggests that dynamic attrition leads to a small under-estimation of the actual results.

Does the initial rank affect ability test scores in later years? Given that we observe each cohort in a different grade, one may be concerned that the initial rank position in high school affects the results of an ability test in later years and hence changes the ordinal rank measured in higher grade levels. However, given that we are interested in the impact of the *ordinal* rank, this concern does not necessarily invalidate our identification strategy. While it is possible that a higher initial rank may lead to a higher measured ability, this does not necessarily change the ordinal ranking within a group. For example, if ability is a monotonically increasing function in the initial rank, then the cardinal difference in test scores between students may increase, but the ordinal ranking remains constant. In Appendix E, we further address this concern in a robustness check, showing that the ordinal rank in wave I has no statistically significant impact on the results of a Peabody test in wave III two years later.

Individual-level omitted confounders Thus far, we have assumed that the Peabody test score is predetermined, thus being immutable by parents, peers, or by students' personality traits and behaviors. If this assumption does not hold, two sources of omitted variable bias can arise: First, a variable may directly influence the ordinal rank, while at the same time affecting educational attainment. A second, more subtle form of omitted variable bias can occur if an unobserved factor simultaneously affects the Peabody score and the outcome. Because the ordinal rank is based on the Peabody score, this could lead to a bias in the estimated rank effect.

Candidates for the first type of bias are parental investments that may differ by a student's rank, or a student's risky behavior, such as smoking marijuana and binge drinking, which may differ by rank while affecting educational attainment.[23] Similarly, rank could be correlated with non-cognitive skills that also affect educational decisions. While these factors could potentially lead to an upward-bias in the estimates of the rank effect, they do not bias the estimates as long as the assignment of students to ranks is quasi-random conditional on the absolute ability score and on being in a given school cohort. Given that we control for absolute ability and school-by-cohort fixed effects in the regression, the results should not be affected by this form of omitted variable bias. In Rows 4) and 5) of Table 4, we estimate the baseline model with school-by-cohort fixed effects, and additionally control for non-cognitive skills and risky behaviors.[24] The

[22] We do not report the result for high school completion here, because the definition of the sub-sample is that students completed high school.

[23] The available evidence on risky behaviors and test scores is mixed. Marie & Zölitz (2015) find that access to marijuana impedes performance at university; Meier *et al.* (2012) show that regular marijuana use can lead to cognitive decline, although their results have been challenged by Rogeberg (2013) who shows that the effect is not causal, and is most likely zero. Elsner & Isphording (2015) show that a low ordinal rank is associated with a higher likelihood of binge drinking, but not with higher marijuana use.

[24] The risky behaviors are coded as follows. Marijuana use: a dummy equaling 1 if the person has smoked marijuana in the last 30 days. Binge drinking: a dummy equaling 1 if the person drank at least once per week five drinks in a row over the last 12 months.

point estimates are marginally lower compared to the baseline results. Yet we do not interpret this difference as evidence of omitted variable bias, but rather see non-cognitive skills and risky behaviors as channels through which a higher rank translates into higher educational attainment, thus mediating the rank effect.

The second type of omitted variable bias could occur if an individual factor affects the Peabody test score as well as the outcome. Candidates for such factors are pressure from parents or peers, or a student's willingness to exert effort in a low-stakes test. These factors indirectly affect a student's ordinal rank because we rank students according to their measured ability, which potentially biases the estimated rank effect if these factors also affect the outcome. However, any factor that affects the measured ability, i.e. the Peabody test score, does not affect the estimate of the rank effect, because we control for measured ability in the regression. Thus, any unobserved factor that affects rank through ability is implicitly controlled for.[25]

Are same-gender peers more important? We are also interested in whether all students in the same school cohort are the relevant comparison group or whether students with the same gender are more important. It may matter more if a girl is the best among all girls rather than being the best among everyone in the grade. In Row 6) of Table 4, we replace the ordinal rank in a school cohort with the ordinal rank within a gender group within a cohort. The results are almost identical as in the baseline specification, indicating that same-gender peers are as important as all students in the same school cohort.

5 Potential channels

The baseline results reveal a significant effect of ordinal rank in high school on human capital investment later in life. The question remains which mechanisms can explain this reduced-form relationship. In this section, we present theoretical arguments for four mechanisms and use the rich survey information provided in AddHealth to analyze which of these mechanisms is supported by the data. We run two sets of regressions. First, we explore the extent to which rank affects a number of intermediate outcomes. We re-estimate the model in Equation (2) using as the dependent variable a dummy that equals one if the student strongly agrees to a given statement, and zero otherwise. The results of this exercise are displayed in Table 5, Column (1). In Columns (2)-(4), we present estimates for the degree to which the estimated total causal effect of rank is mediated through a specific channel. We estimate the degree of mediation by first running an auxiliary regression in which we add the respective mediator as additional explanatory variable in Equation (2). The importance of each mediator is subsequently computed as the product of the coefficient from a regression of rank on the mediator (reported in Column 1) and the coefficient of the mediator on in the auxiliary regression. A higher percentage indicates

[25] We verified this intuition in a series of Monte Carlo simulations, which show that the bias is absent even for large assumed direct effects on the outcome and on the test scores. In additional simulations, we assume that the test scores are measured with error, which would lead to a systematic downward bias in the estimates. The results are available from the authors upon request.

Table 5: Regression results: rank position and intermediate outcomes, wave I

Dependent Variable	Coefficient (1)	SE	% of total effect mediated: Compl. HS (2)	Att. Coll. (3)	Compl. Coll. (4)
Self-concept					
1(I am more intelligent than the average)	0.223***	(0.057)	14.9%***	24.0%***	42.1%***
Expectations					
1(I will have a college degree by the age of 30)	0.110**	(0.048)	7.2%**	10.3%**	14.2%**
Mental distress					
CES-D depression scale (0-60)	-1.684**	(0.802)	6.9%**	7.4%**	12.6%**
Effort					
1(I was absent at school without excuse)	-0.149***	(0.050)	16.7%***	13.0%***	20.7%***
Support from others					
1(I feel that teachers care about me)	0.079	(0.053)	3.7%	4.2%	6.9%
1(I feel that parents care about me)	-0.020	(0.036)	0.1%	-0.6%	-0.9%
1(I feel that friends care about me)	-0.004	(0.020)	-0.3%	-0.2%	-0.3%
Grades					
GPA	0.167***	(0.064)	22.9%**	22.8%***	40.8%***

Note: This table displays the results for separate OLS regressions of the outcomes listed in the first column on the ability rank within a school cohort (Column 1). Columns (2)-(4) display the percentage of the total effect mediated by the specific mediator. With the exception of GPA and depression scale, each outcome is a dummy variable with value 1 if a student strongly agrees to a statement, and zero otherwise. All regressions include school-by-cohort fixed effects, a quartic in absolute ability, and individual control variables. Standard errors, clustered at the school level, are displayed in parentheses, with significance levels * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

a stronger role for a specific channel in mediating the total effect.[26]

Overall, we explore four channels, the first three of which affect the demand for schooling. In a basic human capital model, a student chooses her level of schooling optimally by equating the marginal costs of one more year of schooling to the marginal gains. In such a model, the ordinal rank can affect the marginal costs of investing in education in various ways, which we will explore in the following. The fourth channel works through the supply side. Colleges may

[26] Under fairly strong assumptions, this quantity then can be interpreted as the fraction of the total effect mediated by single mechanisms. Specifically, in addition to our maintained identification assumptions outlined above, we have to assume that 1) equations characterizing mediator and outcome conditional on mediator are correctly specified and 2) conditional on individual (pre-treatment) characteristics $\boldsymbol{X_{isc}}$, mediators are independent of one another and the rank. For a deeper discussion of these assumptions to identify causal pathways, see Imai *et al.* (2011) and Huber (2014). Mediated percentages of the total effect do not add up to one, because assumption (2) is likely to be violated by cross-correlations between different mediators, although they provide a useful approximation of the relative importance of mediators.

place restrictions on the number of slots available and choose students with higher ranks.

Rank provides incomplete information of one's own ability One mechanism could be that the rank provides students with an imperfect signal about their own ability. There is ample evidence that students have imperfect knowledge about their actual ability (Jensen, 2010; Zafar, 2011; Stinebrickner & Stinebrickner, 2012; Bobba & Frisancho, 2014). The ordinal rank can be one reason for this imperfect knowledge. Students may infer their *absolute ability* from their *relative ability* within their cohort, whereby this misinterpretation may affect their education choices later in life. A low rank may lead a student to believe that she has low gains from education, which is why she may choose not to attend college even if she is brighter than most students of her age cohort.[27]

To assess the importance of this channel, we provide evidence that the ordinal rank affects perceived ability and expectations. We first assess whether students with a higher rank have a higher perceived ability, a frequent finding in the psychology literature (Marsh, 1987). As shown in the first panel of Table 5, this is indeed the case. In wave I of AddHealth, students were asked if they think that they are more intelligent than the average. Conditional on absolute ability, students with a 10 percentage points higher rank are 0.22 percentage points more likely to believe that they are more intelligent than the average. The estimates in Columns (2)-(4) indicate that perceived intelligence is indeed a very important mechanism in mediating the total causal effect of rank on all outcomes, but especially for college completion.

Furthermore, we find strong support that rank affects expected returns to education. In wave I, students were asked whether they expect to have a college degree at the age of 30. Students with a higher rank are more likely to state that they will have a college degree by the age of 30. Remarkably, the impact of rank on college expectations at age 16 is equally as large as the impact of rank on actual college outcomes more than 10 years later.[28] Furthermore, the degree of mediation suggests that the expectations channel is quantitatively important, albeit more for college completion than high school completion and college attendance.

Mental distress and effort The effect could also be explained by intrinsic factors. As suggested by the literature on relative comparisons and effort provision, a higher rank may give students a greater motivation, make them more self-confident and ultimately induce them to exert more effort in their studies (Clark *et al.*, 2010; Azmat & Iriberri, 2010; Gill *et al.*, 2015). In turn, effort would increase the marginal gains from schooling and induce students to invest more in education. The third panel in Table 5 displays the impact of ordinal rank on intrinsic factors, exploiting questions from a survey module on mental distress.

Students with a higher rank score significantly lower on the Center for Epidemiological Studies Depression Scale (CES-D), based on 19 items on recent feelings and moods. To proxy

[27] This would mainly work through the marginal costs of schooling. A low perceived ability leads to a higher marginal costs, in which case students *optimally* choose less education.

[28] The impact of ordinal rank on these expectations is 0.1, the same as the impact of rank on attending college, shown in Table 3.

for effort, we use self-reported information on school absences and construct a dummy that equals one if the student has been absent without excuse at least once in the previous school year. We find that students with a higher rank are significantly less likely to be absent without excuse, indicating that they take their studies more seriously and put more effort into it. Both factors significantly mediate the effect of rank on the outcomes, with mental distress playing a weaker role than differences in effort provision.

Behavioral responses from teachers, parents and friends A further potential channel is behavioral responses from a student's environment. As shown by Pop-Eleches & Urquiola (2013), teachers and parents are responsive to a student's relative position within their school. They compare marginal students who just made it into a high-quality school to those who did not and find that parents provide less effort when their child attends a better school. Moreover, if teachers are time-constrained, they may not be able to give equal treatment to all students, rather providing more support to students with a higher rank. More support from the environment lowers the marginal costs of schooling, and — all else being equal — should lead to better educational outcomes.

In the fifth panel of Table 5, we show the effect of the ordinal rank on perceived support from teachers, parents and friends. Students were asked whether they believe that these groups care about them. In neither of these cases do we find a statistically significant effect of the ordinal rank. Moreover, we do not find any support for a significant mediation of the rank effect on the different outcomes.

Selective college admissions Finally, the effect can be explained by college admissions. Colleges may have a fixed number of slots and/or a fixed amount of financial aid available and may give priority to students with a higher within-school rank. In light of the channels analyzed thus far, college admissions, if anything, can only explain part of the effect. We have shown that rank affects a student's perceived ability, ambition, effort and expectations before students even apply for colleges. If supply-side restrictions were the only explanation for the effect, then we should not observe a significant effect of rank on these other outcomes. Moreover, the outcome variable *attended college* includes all colleges in the US; namely, it also includes a vast number of non-selective community colleges for which these restrictions typically do not apply. If restrictions were the only explanation, we would expect to find no effect on attending any college.

A further supply-side channel is affirmative action, through which colleges may give preferential access to students from minorities or schools in disadvantaged areas. While affirmative action has been shown to significantly distort sorting into colleges (Arcidiacono, 2005), it should not explain our results, because we control for many characteristics that define the minorities targeted by affirmative action, such as Blacks or Hispanics. However, following prominent lawsuits in the mid-1990s, many state colleges in the US have abandoned affirmative action. Instead, California, Texas and Florida introduced ten-percent plans, granting students in the top 10 percent of their high school cohort automatic access to flagship state universities.[29] The 10%-plans were introduced three years after the first wave of AddHealth was collected and thus should only affect the youngest cohorts, if at all.[30]

Besides the 10%-plans that specifically apply to students with a given rank, the effect of rank on college outcomes can more generally be explained by selective college admission policies. Students typically apply for college with their 11th-grade results, which often state the percentile of a student in the GPA distribution of her grade. If college admission officers have this information, and if GPA rank is positively correlated with the ability rank, then our result could reflect a pure mechanical effect: colleges only admit students with a higher rank, which is why we observe higher college attendance rates for highly ranked students. Indeed, as displayed in the second row of Table 5, the ordinal rank is positively related with a student's self-reported GPA. The importance of GPA in mediating the total causal effect puts it on par with perceived intelligence as the main causal pathways between ordinal rank and educational achievements. Controlling for GPA the rank coefficients remain positive and the estimate for college attendance remains statistically significant, suggesting that college admissions can — if at all — only partly explain the results. Interestingly, the mediating effect is largest for college attendance, which indicates that GPA does not purely affect achievements through college admissions but also plays an additional role through providing additional information on one's ability.

6 Conclusion

In this paper, we show that a student's ordinal rank in a high school cohort is an important determinant for educational attainment later in life. Comparing students across cohorts within the same school and flexibly controlling for cognitive ability, we find that students with a higher ordinal rank in their cohort are more likely to complete high school, attend college and complete a four-year degree. These results complement the findings of recent research by Murphy & Weinhardt (2014), who show that a student's ordinal rank in school has a significant positive impact on test scores, as well as Tincani (2015), who shows that rank concerns are an important determinant of peer effects. Our paper reinforces their general conclusion and shows that the ordinal rank in high school affects important career choices.

In addition, by exploiting rich survey information, we provide evidence on the mechanisms that help to explain this result. We find that students with a higher rank have higher expectations about their future career outcomes, are more optimistic and self-confident. These results suggest that students partly base their schooling decisions on their relative rather than their absolute ability. The ordinal rank provides students with noisy information about their

[29] Daugherty *et al.* (2014) for Texas and Arcidiacono *et al.* (2014) for California provide evidence that the introduction of these plans changed the composition of students at flagship state colleges. In Texas, attendance and completion rates at these colleges increased, but more so for students from high-ability high schools. In California the college attendance rates of Blacks increased, but larger shares of Blacks went to lower-ranked colleges.

[30] It is not possible to test whether the results differ between states with and without a 10%-plan, because AddHealth contains no state identifiers.

own ability, which in turn distorts the trade-off between the short-run costs and the long-run benefits from schooling.

The rank effect runs counter to most of the literature on peer effects, which finds that being exposed to better peers has a positive effect on educational attainment. In our econometric model, we are able to isolate the rank effect from linear effects that work through the mean and variance of peer ability, showing that the rank effect is a channel of peer effect in its own right. It can help explain seemingly puzzling findings of the more recent literature on non-linear peer effects, which shows that students at the low end of the peer ability distribution are negatively affected by better peers (Carrell *et al.*, 2013). This could be the case because the negative effect from a low rank is greater than the positive effect of being exposed to better peers.

Our results should concern parents and policy-makers alike. Given that a student's relative position in a cohort affects outcomes later in life, parents should take the rank of their child into account when choosing a school for their child. The highest-ranked school may not be beneficial for children at the margin who only just make it into the school. However, the results should be taken with the caveat that rank is only one input in educational production. Our results reflect local effects, which we obtained by comparing students within the same school but in different cohorts. If parents chose schools exclusively based on their child's rank, they would ignore many other school inputs that are equally important and differ between schools, such as teacher quality, average peer ability and school resources.

Policy-makers should be concerned because our results show that students with a high ability but a low rank — small fish in a big pond — under-invest in their human capital. Students seemingly place too much weight on their relative ability and too little on their absolute ability when making college decisions. It is difficult to think of a pareto-improving algorithm that reduces this inefficiency by regrouping students into schools, because the ordinal rank is a relative measure and improving one student's rank means worsening the rank of another.

A more efficient policy could be to give more support to students at lower ranks of the ability distribution to compensate for the negative impact of their rank. Especially for students with a low rank in high-ability schools, providing them with information on their absolute ability could be an inexpensive and effective way to increase their educational attainment. Recent experimental studies by Azmat & Iriberri (2010), Tran & Zeckhauser (2012), Hastings & Weinstein (2008), Bettinger *et al.* (2012), Oreopoulos & Dunn (2013) and Wiswall & Zafar (2015) have shown that students are indeed responsive to these type of interventions.

References

Abdulkadiroglu, Atila, Angrist, Joshua D., & Pathak, Parag A. 2014. The Elite Illusion: Achievement Effects at Boston and New York Exam Schools. *Econometrica*, **82**(1), 137–196.

Ammermueller, Andreas, & Pischke, Jörn-Steffen. 2009. Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study. *Journal of Labor Economics*, **27**(3), 315–348.

Anderson, Darrell E., & Flax, Morten E. 1968. A Comparison of the Peabody Picture Vocabulary Test with the Wechsler Intelligence Scale for Children. *Journal of Educational Research*, **62**(3), 114–116.

Angrist, Joshua. 2014. The Perils of Peer Effects. *Labour Economics*, **30**(C), 98–108.

Arcidiacono, Peter. 2005. Affirmative Action in Higher Education: How do Admission and Financial Aid Rules Affect Future Earnings? *Econometrica*, **73**(5), 1477–1524.

Arcidiacono, Peter, Aucejo, Esteban, Coate, Patrick, & Hotz, V. Joseph. 2014. Affirmative Action and University Fit: Evidence from Proposition 209. *IZA Journal of Labor Economics*, **3:7**.

Attanasio, Orazio P., & Kaufmann, Katja M. 2015. Education Choices and Returns to Schooling: Mothers' and Youths' Subjective Expectations and their Role by Gender. *Journal of Development Economics*, **109C**, 203–216.

Azmat, Ghazala, & Iriberri, Nagore. 2010. The Importance of Relative Performance Feedback Information: Evidence from a Natural Experiment using High School Students. *Journal of Public Economics*, **94**, 435–452.

Azmat, Ghazala, Bagues, Manuel, Cabrales, Antonio, & Iriberri, Nagore. 2015. What You Know Can't Hurt You (for Long): A Field Experiment on Relative Performance Feedback in Higher Education. *Aalto University, mimeo.*

Baker, Paula C., Keck, Canada K., Mott, Frank L., & Quinlan, Stephen V. 1993. *NLSY Child Handbook, Revised Edition: A Guide to the 1986-1990 NLSY Child Data.* The Ohio State University, Center for Human Resource Research.

Bettinger, Eric P., Long, Bridget Terry, Oreopoulos, Philip, & Sanbonmatsu, Lisa. 2012. The Role of Application Assistance and Information in College Decisions: Results from the H&R Block FAFSA Experiment. *The Quarterly Journal of Economics*, **127**(3), 1205–242.

Betts, Julian R. 2011. The Economics of Tracking in Education. *Pages 341–381 of:* Hanushek, Eric A., Machin, Stephen, & Woessmann, Ludger (eds), *Handbook of the Economics of Education*, vol. 3. North Holland.

BIFULCO, ROBERT, FLETCHER, JASON M., & ROSS, STEPHEN L. 2011. The Effect of Classmate Characteristics on Post-Secondary Outcomes: Evidence from the Add Health. *American Economic Journal: Economic Policy*, **3**, 25–53.

BLACK, SANDRA E., DEVEREUX, PAUL J., & SALVANES, KJELL G. 2013. Under Pressure? The Effect of Peers on Outcomes of Young Adults. *Journal of Labor Economics*, **31**(1), 119–153.

BOBBA, MATTEO, & FRISANCHO, VERONICA. 2014. Learning About Oneself: The Effects of Signaling Academic Ability on School Choice. *Inter-American Development Bank, mimeo.*

BOOIJ, ADAM, LEUVEN, EDWIN, & OOSTERBEEK, HESSEL. 2015. Ability Peer Effects in University: Evidence from a Randomized Experiment. *IZA Discussion Paper*, **8769**.

BRINCH, CHRISTIAN N., & GALLOWAY, TARYN ANN. 2012. Schooling in Adolescence Raises IQ Scores. *Proceedings of the National Academy of Sciences*, **109**(2), 425–430.

BURKE, MARY A., & SASS, TIM A. 2013. Classroom Peer Effects and Student Achievement. *Journal of Labor Economics*, **31**(1), 51–82.

CALVÓ-ARMENGOL, ANTONI, PATACCHINI, ELEONORA, & ZENOU, YVES. 2009. Peer Effects and Social Networks in Education. *Review of Economic Studies*, **76**, 1239–1267.

CARRELL, SCOTT E., FULLERTON, RICHARD L., & WEST, JAMES E. 2009. Does Your Cohort Matter? Measuring Peer Effects in College Achievement. *Journal of Labor Economics*, **27**(3), 429–464.

CARRELL, SCOTT E., SACERDOTE, BRUCE I., & WEST, JAMES E. 2013. From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation. *Econometrica*, **81**(3), 855–882.

CICALA, STEVE, FRYER, JR., ROLAND G., & SPENKUCH, JÖRG. 2015. Comparative Advantage in Social Interactions. *Northwestern University, mimeo.*

CLARK, ANDREW, MASCLET, DAVID, & VILLEVAL, MARIE-CLAIRE. 2010. Effort and Comparison Income: Experimental and Survey Evidence. *Industrial and Labor Relations Review*, **63**(3), 407–426.

CULLEN, JULIE BERRY, LONG, MARK C., & REBACK, RANDALL. 2013. Jockeying for Position: Strategic High School Choice under Texas' Top Ten Percent Plan. *Journal of Public Economics*, **97**, 32–48.

CUNHA, FLAVIO, HECKMAN, JAMES J., LOCHNER, LANCE, & MASTEROV, DIMITRIY. 2006. Interpreting the Evidence on Life Cycle Skill Formation. *Chap. 12, pages 697–812 of:* HANUSHEK, ERIC A., & WELCH, FINIS (eds), *Handbook of the Economics of Education*, vol. 1. Elsevier.

Daugherty, Lindsay, Martorell, Paco, & McFarlin Jr, Isaac. 2014. Percent Plans, Automatic Admissions, and College Outcomes. *IZA Journal of Labor Economics*, **3:10**.

De Giorgi, Giacomo, & Pellizzari, Michele. 2014. Understanding Social Interactions: Evidence from the Classroom. *Economic Journal*, **124**(579), 917–953.

Deming, David, & Dynarski, Susan. 2008. The Lengthening of Childhood. *Journal of Economic Perspectives*, **22**(3), 71–92.

Dunn, Lloyd M., & Dunn, Leota M. 2007. *The Peabody Picture Vocabulary Test*. 4th edn. Bloomington, MN: NCS Pearson, Inc.

Elsner, Benjamin, & Isphording, Ingo E. 2015. Rank, Sex, Drugs, and Crime. *IZA Discussion Paper*, **9478**.

Feld, Jan, & Zölitz, Ulf. 2017. On the Nature, Estimation and Channels of Peer Effects. *Journal of Labor Economics*, **forthcoming**.

Fu, Chao, & Mehta, Nirav. 2015. Ability Tracking, School and Parental Effort, and Student Achievement: A Structural Model and Estimation. *University of Wisconsin-Madison, mimeo.*

Gill, David, Kissová, Zdenka, Lee, Jaesun, & Prowse, Victoria L. 2015. First-Place Loving and Last-Place Loathing: How Rank in the Distribution of Performance Affects Effort Provision. *IZA Discussion Paper*, **9286**.

Goulas, Sofoklis, & Megalokonomou, Rigissa. 2015. Knowing Who You Are: The Effect of Feedback Information on Exam Placement. *University of Warwick, mimeo.*

Hanushek, Eric A., Kain, John F., Markman, Jacob M., & Rivkin, Steven G. 2003. Does Peer Ability Affect Student Achievement? *Journal of Applied Econometrics*, **527-544**.

Harris, Kathleen Mullan. 2009. The National Longitudinal Study of Adolescent to Adult Health. *doi: 10.3886/ICPSR27021.v9.*

Harris, K.M., Halpern, C.T., Whitsel, E., Hussey, J., Tabor, J., Entzel, P., & Udry, J.R. 2009. The National Longitudinal Study of Adolescent to Adult Health: Research Design [WWW document]. *URL: http://www.cpc.unc.edu/projects/addhealth/design.*

Hastings, Justine S., & Weinstein, Jeffrey M. 2008. Information, School Choice, and Academic Achievement: Evidence from Two Experiments. *The Quarterly Journal of Economics*, **123**, 1373–1414.

Hoxby, Caroline. 2000. The Effects of Class Size on Student Achievement: New Evidence from Population Variation. *The Quarterly Journal of Economics*, **115**(4), 1239–1285.

Huber, Martin. 2014. Identifying Causal Mechanisms (Primarily) Based on Inverse Probability Weighting. *Journal of Applied Econometrics*, **29**(6), 920–943.

Imai, Kosuke, Keele, Luke, Tingley, Dustin, & Yamamoto, Teppei. 2011. Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies. *American Political Science Review*, **105**(11), 765–789.

Imberman, Scott, Kugler, Adriana D., & Sacerdote, Bruce. 2012. Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees. *American Economic Review*, **102**(5), 2048–2082.

Jensen, Arthur R. 1998. *The g Factor : the Science of Mental Ability*. Praeger Publishers.

Jensen, Robert. 2010. The Perceived Returns to Education and the Demand for Schooling. *The Quarterly Journal of Economics*, **125**(2), 515–548.

Jonsson, Jan O., & Mood, Carina. 2008. Choice by Contrast in Swedish Schools: How Peers' Achievement Affects Educational Choice. *Social Forces*, **87**(2), 741–765.

Koppensteiner, Martin Foureaux. 2012. Class Assignment and Peer Group Effects: Evidence from Brazilian Primary Schools. *University of Leicester Working Paper*, **03**.

Kuziemko, Ilyana, Buell, Ryan W., Reich, Taly, & Norton, Michael I. 2014. Last-place Aversion: Evidence and Redistributive Implications. *The Quarterly Journal of Economics*, **129**(1), 105–149.

Lavy, Victor, Silva, Olmo, & Weinhardt, Felix. 2012. The Good, the Bad and the Average: Evidence on Ability Peer Effects in Schools. *Journal of Labor Economics*, **30**(2), 367–414.

Marie, Olivier, & Zölitz, Ulf. 2015. 'High' Achievers? Cannabis Access and Academic Performance. *IZA Discussion Paper*, **8900**.

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

Marsh, Herbert W., Trautwein, Ulrich, Lüdtke, Oliver, Baumert, Jürgen, & Köller, Olaf. 2007. The Big-Fish-Little-Pond Effect: Persistent Negative Effects of Selective High Schools on Self-Concept After Graduation. *American Educational Research Journal*.

Meier, Madeline H., Caspi, Avshalom, Ambler, Antony, Harrington, HonaLee, Houts, Renate, Keefe, Richard S. E., McDonald, Kay, Ward, Aimee, Poulton, Richie, & Moffitt, Terrie E. 2012. Persistent Cannabis Users Show Neuropsychological Decline from Childhood to Midlife. *Proceedings of the National Academy of Sciences*, **109**(40), E2657–E2664.

Murphy, Richard, & Weinhardt, Felix. 2014. Top of the Class: The Importance of Ordinal Rank. *CESifo Working Paper*, **4815**.

Oreopoulos, Philip, & Dunn, Ryan. 2013. Information and College Access: Evidence from a Randomized Field Experiment. *Scandinavian Journal of Economics*, **115**(1), 3–26.

Patacchini, Eleonora, Rainone, Eduardo, & Zenou, Yves. 2012. Student Networks and Long-Run Educational Outcomes: The Strength of Strong Ties. *CEPR Discussion Paper*, **9149**.

Pop-Eleches, Cristian, & Urquiola, Miguel. 2013. Going to a Better School: Effects and Behavioral Responses. *American Economic Review*, **103**(4), 1289–1324.

Rogeberg, Ole. 2013. Correlations between Cannabis Use and IQ Change in the Dunedin Cohort are Consistent with Confounding from Socioeconomic Status. *Proceedings of the National Academy of Sciences*, **110**(11), 4251–4254.

Sacerdote, Bruce. 2001. Peer Effects with Random Assignment: Results for Dartmouth Roommates. *The Quarterly Journal of Economics*, **116**(2), 681–704.

Sacerdote, Bruce. 2011. Peer Effects in Education: How Might They Work, How Big Are They, and how much Do We Know thus far? *In:* Hanushek, Eric A., Machin, Stephen, & Woessmann, Ludger (eds), *Handbook of the Economics of Education*, vol. 3. North-Holland.

Stinebrickner, Todd, & Stinebrickner, Ralph. 2012. Learning about Academic Ability and the College Drop-out Decision. *Journal of Labor Economics*, **30**(4), 707–748.

Stinebrickner, Todd, & Stinebrickner, Ralph. 2014. A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout. *Review of Economic Studies*, **81**(1), 426–472.

Tatsi, Eirini. 2015. Endogenous Social Interaction: Which Peers Matter? *Goethe University Frankfurt, mimeo.*

Tincani, Michela. 2015. Heterogeneous Peer Effects and Rank Concerns: Theory and Evidence. *University College London, mimeo.*

Tran, Anh, & Zeckhauser, Richard. 2012. Rank as an Inherent Incentive: Evidence from a Field Experiment. *Journal of Public Economics*, **96**, 645–650.

Wiswall, Matthew, & Zafar, Basit. 2015. Determinants of College Major Choices: Identification from an Information Experiment. *Review of Economic Studies*, **82**(2), 791–824.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross- Section and Panel Data.* The MIT Press.

Young, J. Kenneth, & Beaujean, A. Alexander. 2011. Measuring Personality in Wave I of the National Longitudinal Study of Adolescent Health. *Frontiers in Psychology*, **2**(158).

Zafar, Basit. 2011. How do College Students Form Expectations? *Journal of Labor Economics*, **29**(2), 301–348.

Zimmerman, David J. 2003. Peer Effects in Academic Outcomes: Evidence from a Natural Experiment. *Review of Economics & Statistics*, **85**(1), 9–32.

WEB APPENDIX

A Disclaimer

This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due to Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis.

B The Peabody Picture Vocabulary Test

AddHealth includes an abriged version of the Peabody Picture Vocabulary Test Revised (PPVT-R), which was developed in the 1950s to measure verbal intelligence and scholastic aptitude (Dunn & Dunn, 2007). The revised version of the test, introduced in 1981, comprises 175 tasks given to the respondent in a sequence of increasing difficulty. Each task involves a stimulus word and an image plate with four pictures, whereby respondents have to match the word to the picture that fits best. The tasks are the same for all age groups, although older age groups start with a higher task. The age-specific starting point is suggested by the survey manual.

A respondent's score is based on the number of correct matches of words and pictures between an individual-specific lower bound (*basal*) and an upper bound (*ceiling*). Both bounds are determined based on number of subsequent correct and incorrect answers, respectively. To determine the basal, the test begins at the suggested starting point and proceeds until the respondent has given the first incorrect answer after having correctly answered eight subsequent tasks. The basal is then given by the highest task of these eight correctly answered tasks. If a respondent does not give eight consecutive correct answers, the test proceeds backwards until eight consecutive correct answers were given, and failing that, gives a basal of one.

The ceiling is determined by the number of incorrect answers. The test stops once a respondent has incorrectly answered six out of eight tasks. The ceiling is then defined as the highest of this series of eight tasks. If a respondent never incorrectly answers more than five out of eight answers, he/she is given a ceiling of 175.

The final score of the test is computed based on the number of correct answers below the ceiling. All answers below the basal count as correct, regardless of whether the respondent has actually given a correct answer. The version included in AddHealth is based on the same tasks as the PPVT-R, although it only uses every other task: odd-numbered tasks from 1 to 87, and even-numbered tasks from 90 to 175, amounting to a total of 87 tasks. The criteria for the basal and the ceiling required half the number of answers of the long version: four consecutive correct answers to define the basal and three out of four tasks answered incorrectly to define the ceiling.[31]

C Validation of the identifying assumption

In Section 3.2, we present two ways of modeling the school cohort effects λ_{sc}, by either assuming that they are constant across cohorts or estimating a model with school-by-cohort fixed effects. As explained in the same section, both models have their virtues but come at their cost. The two-way fixed effects model, i.e. a model with separate school and cohort fixed effects, exploits variation in the rank that comes from differences in all moments of the ability distribution across cohorts within a school. In this model, γ only has a causal interpretation if there are no systematic differences between cohorts within a school. A model with school-by-cohort fixed effects treats the average differences across school cohorts as a nuisance parameter and does not allow us to explore the source of the variation in rank because all moments of the ability distribution are absorbed in the fixed effects.

To test whether a two-way model is sufficient for a causal interpretation of the results, we perform a cluster-robust Hausman test, based on which we test whether the school-cohort effects are correlated with all other regressors in Equation (2). Under the null hypothesis, there is no correlation, whereas under the alternative hypothesis, the school-cohort effects are correlated with at least some of the regressors. The Hausman test essentially compares a model with school-by-cohort random effects to a model with school-by-cohort fixed effects. Under the null hypothesis of no correlation, both models should yield the same results. If we fail to reject the null hypothesis, we can interpret the results of the two-way fixed effects model as causal.

In Table 6, we present two sets of results. In Column (1), we include the same controls as in Equation (2). In Column (2), the random effects model additionally includes controls for the mean and variance of ability. Given that the rank is mechanically correlated with the mean ability, we would expect to reject the null hypothesis purely on these grounds if the mean ability is not included. By including these variables, we test whether the remaining school-cohort effects are uncorrelated with the regressors. The test results are mixed and depend on the outcome. For college attendance, we reject the null in both specifications. For high school completion, we reject the null once we include the mean and variance of ability, whereas for college completion, we fail to reject the null hypothesis at the 5% level in both cases.

Overall, the results from the Hausman test leave doubts about the causal interpretation of the marginal effect from a two-way fixed effects model, and rather suggest relying on a more conservative model with school-by-cohort fixed effects.

[31] We would like to thank Joyce Tabor, the AddHealth Data Manager, for providing us with detailed information on the test.

Table 6: Cluster-robust Hausman tests

Dependent variable	(1)	(2)
Completed High School	1.43	2.09
	[0.13]	[0.01]
Attended college	4.12	4.02
	[0.00]	[0.00]
Completed High School	1.61	1.55
	[0.07]	[0.08]
Controls		
Individual ability	Yes	Yes
Individual controls	Yes	Yes
Mean cohort ability	No	Yes
Cohort variance of ability	No	Yes

Note: This table presents F statistics and p-values (in parentheses) of a cluster-robust Hausman test as proposed by Wooldridge (2002). The H_0 states that in a regression $y_{isc} = M_{isc} + \rho_{sc} + \eta_{isc}$ the school-by-cohort effects ρ_{sc} are uncorrelated with all other regressors M_{isc}, $Cov(\rho_{sc}, M_{isc}) = 0$, where M_{isc} includes all regressors in Equation 2.

D Measurement error in the rank variable

Given the survey design, we only observe a random sample of every school cohort, based on which we compute the ordinal rank. This inevitably introduces measurement error in the rank variable, because based on the sample we attribute to some students a rank that is higher than their true rank in the population, and to others a rank that is lower. In this section, we first provide additional details on the sample design and quantify the extent to which this design introduces sampling variation in the ordinal rank. Finally, based on Monte Carlo experiments, we show how the measurement error that arises from the sample design affects the estimates.

D.1 Sample design

In wave I of AddHealth, a stratified random sample of students from each school-cohort combination was included in the in-home sample. Within each school cohort, 17 boys and 17 girls were drawn at random, while several students from minorities were included to facilitate the analysis of these groups. In 16 so-called *saturated schools*, all students that were present on the interview day were included in the sample. Most saturated schools are small.

Figure 4 displays some properties of the sample. Each data point represents a within-school average, such that the graphs display the distribution across schools. The top-left panel displays the distribution of the number of students sampled per grade. In most schools, between 10 and 60 students were sampled per cohort, with an average of 40. The outliers represent the saturated schools. The top-right panel displays the distribution of average grade sizes across schools. The mean size is 184 students, with a standard deviation of 131. Both panels at the top illustrate the sample design: from schools of different size, shown in the top-right panel, a more-or-less constant number of students was drawn, shown in the top-left panel. The bottom-left panel shows how the stratified sampling from schools with different cohort sizes translates into sampling ratios, i.e. the share of students from a grade that have been sampled. On average, 34% of a cohort have been sampled, with a standard deviation of 25%.[32] Finally, the bottom-right panel shows how the sample ratio and the average grade size are related. In larger schools, a smaller share of students was sampled. The outliers represent the saturated schools.

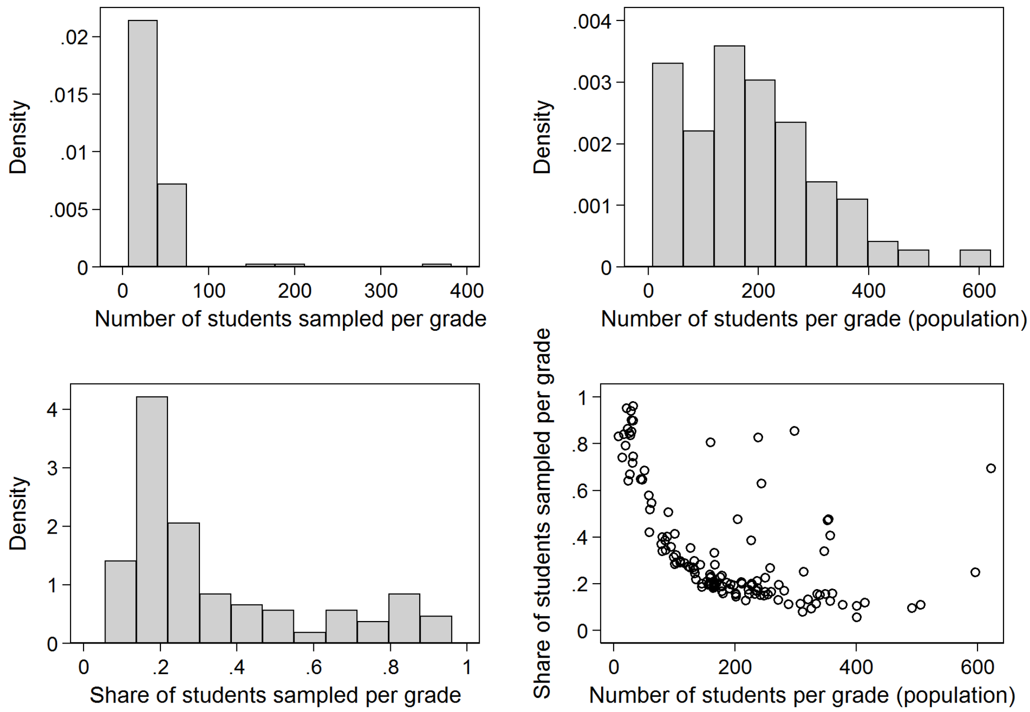


Figure 4: Sample design

Note: This graph displays some features of the sample design. Each data point represents a within-school average. The panels show the distribution of the number of students sampled per school grade (top left), the distribution of the population grade size (top right), the distribution of the share of students sampled per grade (bottom left) and the relationship between the share of students sampled per grade and the grade size in the population (bottom right).

D.2 Quantifying sampling variation in the ordinal rank

Next, we proceed to quantify the degree of sampling variation, using data from the two large 'saturated schools'. In most schools included in AddHealth, all students provided basic information in the in-school sample, while a random sample from each school cohort were included in the in-home sample and administered the detailed questionnaire, including the Peabody test. By contrast, in saturated schools, all students were included in the in-home sample. Based on the saturated schools, we can quantify the amount of sampling error by comparing the true

[32] These figures differ from those provided in the summary statistics in Table 2, because the figures presented here are within-school averages, such that all schools have equal weight, whereas in Table 2, all school-grade combinations have equal weight.

rank in the population to the rank based on the random sample. For this purpose, we set up a bootstrap procedure, in which we draw a large number of random samples from the two large saturated schools and calculate the amount of sampling variation.

AddHealth has 16 saturated schools, 14 of which are small schools with 150 students or less. For our analysis, we only consider the two largest schools. School A has 786 students in grade 9-12, with an average cohort size of around 200. School B has 1,530 students in grades 10-12, and an average cohort size of 510 students. In both schools, the cohort size is above the average of 184, which means that we will most likely over-estimate the degree of sampling variation.

The bootstrap procedure works as follows. We first compute the true within-cohort rank based on the population of students in both schools. Subsequently, in each replication, we draw a random sample of 17 boys and 17 girls from each school-cohort combination, calculate the rank based on this random sample and compute the sampling error as the difference between the true rank and the rank in the random sample. Figure 5 displays the distribution of the sampling error for 10,000 replications. In line with random sampling, the sampling error is centered around zero, which means that for roughly one half of the students we over-estimate the rank, while for the other half we under-estimate it.

More important than the mean is the standard deviation of the sampling error, which measures by how much on average the rank from the random sample deviates from the true rank. We find a standard deviation $sd = 0.017$, based on a rank variable defined between zero and one, $r_{isc} \in [0, 1]$. In a grade of 100 students, this means that we misclassify every student by 1.7 absolute rank positions. Scaled up to the average population cohort size of 184 students, we misclassify each student's rank on average by 3.1 positions.

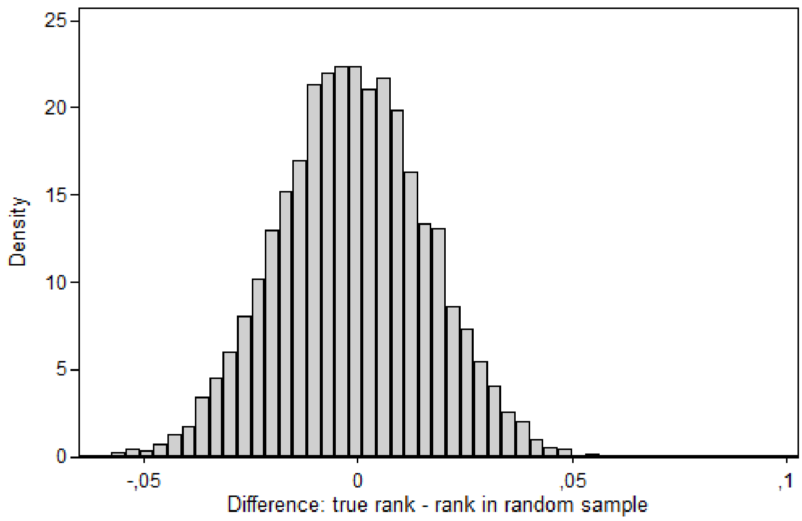


Figure 5: Sampling variation in the rank variable

Note: This graph displays the distribution of the sampling error based on 10,000 random samples drawn from two large saturated schools.

D.3 Monte Carlo simulations

Having quantified the amount of measurement error in the ordinal rank, we now analyze the extent to which this measurement error biases the estimated impact of ordinal rank on educational attainment. To this end, we carry out a series of Monte Carlo experiments, in which we first create artificial school data similar to the base sample of AddHealth, draw a random sample from each school cohort in this dataset, and estimate the model in Equation (2) with school-by-cohort fixed effects based on the random sample. Given that AddHealth samples a constant number of students from school cohorts of varying sizes, we carry out separate experiments and compute the bias as a function of the school-cohort size.

We generate the population school data with similar properties as the schools in AddHealth, based on 500 school-cohort combinations. We assume that in the population the ordinal rank has a positive effect on the outcome, $\gamma_{true} = 0.1$. The assumed data-generating process is

$$y_{isc} = 0.1 r_{isc} + 0.6 a_{isc} + \delta_{sc},$$

where y_{isc} is the outcome of student i in school s in cohort s, for example the likelihood to go to college, r_{isc} is a student's percentile rank in the ability distribution of the school cohort, a_{isc} is the cognitive ability, and δ_{sc} is a school-cohort-specific intercept. We draw the ability from a normal distribution $a_{isc} \sim N(101, 14)$, and construct the percentile rank as described in Section 2.3.[33]

Each experiment is based on 1,000 replications of the same procedure. Each replication $r = 1, \ldots, 1000$ proceeds as follows:

- Generate the population school dataset as described above.
- Draw a random sample of 40 students per school cohort, and compute the rank based on the ability distribution of the sample.
- Run a linear regression of the outcome on rank, absolute ability, and a vector of school-cohort fixed effects to obtain an estimate for the effect of the ordinal rank on the outcome, $\hat{\gamma}_r$.

To quantify the bias from measurement error, we compare the average estimate from the random sample, $\bar{\hat{\gamma}} = 1000^{-1} \sum_r \hat{\gamma}_r$ to the benchmark coefficient from the data-generating process, $\gamma_{true} = 0.1$. We begin with a school-cohort size of 40, in which case the entire school cohort is sampled, and increase the school cohort size in steps of 40, up to a school cohort size of 400, in which case only 10% of all students are sampled.

Figure 6 displays the results of this exercise. The bias arising from measurement error is the difference between the horizontal line at $\gamma_{true} = 0.1$ and the average estimate from the random sample. When we sample the entire cohort of 40, there is no measurement error and

[33] For the value δ_{sc}, we use the school identifiers $s = 1, \ldots, 500$. The value of δ_{sc} is irrelevant for the analysis, because it will be absorbed by school-cohort fixed effects in all regressions.

thus no attenuation bias. Once we increase the school cohort size but keep the number of sampled students constant, the attenuation bias increases, but eventually flattens out. The baseline results are labeled as *homogeneous school cohorts*, because the individual ability of every student was drawn from the same distribution. At a cohort size of 184, the average in the AddHealth sample, the attenuation bias amounts to 30%; at a cohort size of 400, the attenuation bias is 33%.

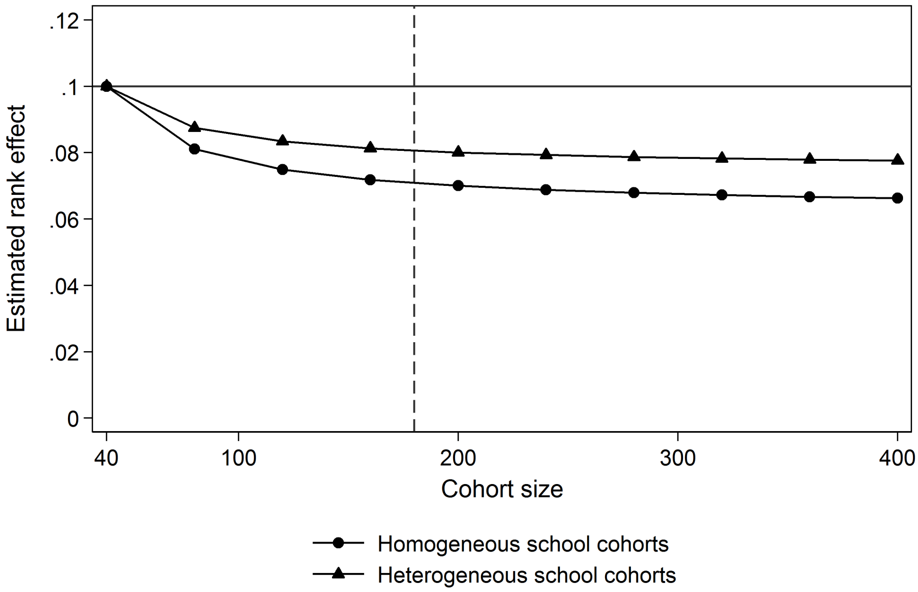


Figure 6: Monte Carlo simulations

Note: This graph displays results for separate Monte Carlo experiments. Each data point is the average coefficient from the baseline regression, based on a sample of 40 students drawn from a cohort size indicated on the horizontal axis. In the experiments with homogeneous school cohorts, the ability of each student was drawn from the overall ability distribution. In the experiments with heterogeneous school cohorts, the school cohorts systematically differ in the mean and variance of ability. The vertical dashed line indicates the average school cohort size of 184.

In a further series of Monte Carlo experiments, we relax the assumption that every student's ability is drawn from the same distribution, and account for heterogeneity across school cohorts in terms of mean and variance of ability. As shown in the summary statistics in Table 2, there is significant heterogeneity in the mean and variance of ability across schools. The school data are now created in multiple steps. We first draw the mean $\bar{a}$ and standard deviation sd_a of ability for every school cohort, using the distribution observed in AddHealth, $\bar{a} \sim N(101, 7)$, and $sd_a \sim N(12, 2.5)$, which gives us for every school cohort sc a new ability distribution $a_{sc} \sim N(\bar{a}_{sc}, sd_{a,sc})$. We then draw the individual ability from this distribution, and otherwise follow the same procedure as before. In this setup, ability has a lower variance within school cohorts and a higher variance between.

With heterogeneous school cohorts, the attenuation bias is smaller. At the average cohort size of 180 students, it amounts to 20%, and flattens out at around 25%. The bias is smaller with heterogeneous school cohorts because the within-cohort variance in ability is smaller, while the between-cohort variance is completely absorbed by the school-by-grade fixed effects. Due to the smaller within-cohort variance, a the ability distribution of random sample more accurately approximates the distribution of a school cohort in the population.[34]

E Further robustness checks

Table 7: Further robustness checks

	Dependent variable		
	Percentile rank	*Attrition dummy*	*Wave 3 PPVT-R*
Neuroticism	-0.002		
	(0.001)		
Conscientiousness	0.000		
	(0.001)		
Percentile rank		0.027	1.073
		(0.042)	(1.107)

Note: This table presents additional robustness checks. In the left column, the rank is regressed on non-cognitive skills, individual controls, absolute ability and school-by-cohort fixed effects. In Column 2) an attrition dummy is regressed on the ordinal rank, individual controls, ability and school-by-grade fixed effects. In Column 3) Peabody test scores from AddHealth Wave 3 are regressend on the ordinal rank, individual controls, ability and school-by-grade fixed effects. Standard errors, clustered at the school level, are displayed in parentheses, with significance levels * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

In this section, we present further robustness checks. The first check addresses the concern that the Peabody score is partly driven by non-cognitive skills, which would be an omitted variable in the regression. The second check provides evidence against selective attrition.

Peabody test scores and non-cognitive skills One concern with using the Peabody test scores as ability measure is that Peabody is a low-stakes test in which students have little incentive to do well. In light of this, more conscientious students may perform better on the test simply because they take the test more seriously. In Column 1), we regress the rank on measures for conscientiousness and neuroticism — the only non-cognitive skills available in wave I — as well as all other controls of Equation (2) and school-by-cohort fixed effects. If non-cognitive skills were important confounders, their coefficient would be large and statistically significant. This is clearly not the case.

Selective attrition In Section 4.4, we explained that the results could be biased due to selective attrition, for example if we are less likely to observe lower-ranked students in wave IV. To assess the importance of selective attrition, we re-estimate the baseline model on the

[34] In an additional Monte Carlo experiment, we refined the data-generating process by accounting for the negative correlation between the mean and standard deviation in ability across school cohorts. However, this refinement does not significantly change the results. The results are available from the authors upon request.

full sample of wave I, and use as outcome variable an attrition dummy that equals one if the person is *not* in the sample in wave IV. As shown in Column 2) of Table 7, selective attrition is unrelated to rank and should not lead to a systematic bias in our estimates.

Does the initial rank affect IQ? In Section 4.4 we discuss a potential issue that could arise because the ordinal rank earlier in a person's life may affect the person's IQ in subsequent years. To alleviate this concern, we exploit that the Peabody test was administered a second time in wave III, two years after wave I. In the regression shown in Column 3) of Table 7, we estimate the model in Equation (2) with school-by-grade fixed effects, and the Peabody score in wave III as the dependent variable. The point estimate is positive, but small and imprecisely estimated, which does not support the view that IQ is influenced by previous rank effects.