NBER WORKING PAPER SERIES

TOP OF THE CLASS:
THE IMPORTANCE OF ORDINAL RANK

Richard Murphy
Felix Weinhardt

Working Paper 24958
http://www.nber.org/papers/w24958

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2018

We thank Esteban Aucejo, Thomas Breda, David Card, Andrew Clark, Jeff Denning, Susan Dynarski, Ben Faber, Mike Geruso Eric Hanushek, Brian Jacob, Pat Kline, Steve Machin, Magne Mogstad, Imran Rasul, Jesse Rothstein, Olmo Silva, Kenneth Wolpin, Gill Wyness, and participants of the CEP Labour Market Workshop, UC Berkeley Labour Seminar, the Sussex University, Queen Mary University and Royal Holloway-University departmental seminars, the CMPO seminar group, the RES Annual Conference panel, IWAEE, the Trondheim Educational Governance Conference, the SOLE conference, CEP Annual Conference, the UCL PhD Seminar, the BeNA Berlin Seminar, IFS seminar and the CEE Education Group for valuable feedback and comments. Earlier working paper versions of this article are Murphy and Weinhardt (2013) and Murphy and Weinhardt (2014). Weinhardt gratefully acknowledges ESRC seed funding (ES/J003867/1) as well as support by German Science foundation through CRC TRR 190. All remaining errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Top of the Class: The Importance of Ordinal Rank
Richard Murphy and Felix Weinhardt
NBER Working Paper No. 24958
August 2018
JEL No. I21,J24

ABSTRACT

This paper establishes a new fact about educational production: ordinal academic rank during primary school has long-run impacts that are independent from underlying ability. Using data on the universe of English school students, we exploit naturally occurring differences in achievement distributions across primary school classes to estimate the impact of class rank conditional on relative achievement. We find large effects on test scores, confidence and subject choice during secondary school, where students have a new set of peers and teachers who are unaware of the students' prior ranking. The effects are especially large for boys, contributing to an observed gender gap in end-of-high school STEM subject choices. Using a basic model of student effort allocation across subjects, we derive and test a hypothesis to distinguish between learning and non-cognitive skills mechanisms and find support for the latter.

Richard Murphy
Department of Economics
University of Texas at Austin
2225 Speedway, C3100
Austin, TX 78712
and NBER
richard.murphy@austin.utexas.edu

Felix Weinhardt
DIW Berlin
Mohrenstraße 58
10117 Berlin
Germany
fweinhardt@diw.de

1 Introduction

Educational achievement is among the most important determinants of welfare, both individually and nationally, and as a result there exists a vast literature examining educational choices and production. Yet to date, lasting effects of ordinal academic ranks (conditional on achievement) have not been considered. Why might rank matter? It is human nature to make social comparisons in terms of characteristics, traits and abilities (Festinger, 1954). When doing so, individuals often use cognitive shortcuts (Tversky and Kahneman, 1974). One such heuristic is to use simple ordinal rank information instead of more detailed cardinal information. Indeed, recent papers have shown that individuals use ordinal rank position, in addition to relative position, to make comparisons with others and that these positions affect happiness and job satisfaction (Brown et al., 2008; Card et al., 2012). The intuition for such effects is the following: David is smarter than Thomas, who is in turn smarter than Jack. These comparisons focus not on the magnitude of the differences but the ranking of individuals. These comparisons can affect individuals' beliefs about themselves and their abilities. Following this intuition, the way we think of ourselves would partly determined by our immediate environment, and this could affect later outcomes by influencing the actions and investment decisions of ourselves or others.

This paper applies this idea to education and presents the first empirical evidence that a student's academic rank during primary school has impacts throughout secondary school. In our main specifications, we regress outcomes during secondary school on externally marked end-of-primary test scores and on the corresponding ranks within their classroom. We find that primary rank has important independent effects on later test scores, subject-choice and subject-specific confidence, in a new setting with peers and teachers that are unaware of a student's ranking in primary school. Therefore we show that ordinal information in addition to cardinal information has the potential to affect investment decisions. Our analysis proposes novel approaches for isolating rank effects by exploiting idiosyncratic variation in the test score distributions across primary schools combined with the nature of ordinal and cardinal information.

First, idiosyncratic variation in the distribution of primary school peer quality arises naturally because primary school classes are small and students vary in ability. Figure 1 provides a stylized illustration of this. The figure shows two classes of eleven students, with each mark representing a student's test score, increasing from left to right, which can be used to rank students. The classes are very similar having the same mean, minimum, and maximum student test scores. However, two students with the same absolute and relative-to-the-mean test score, can still have different ranks. For example, a student with a test score of Y in Class A would have a lower rank (R=5) than the same test score in Class B (R=2). Similarly, a test score of X would be ranked differently in Classes A and B. Notably, this variation will occur across any classrooms, both within and across cohorts, schools and subjects. For example, in one cohort a student with a test score of 80 would be the second of their class, while in the next cohort the same score would place them fifth. Similarly, Class A and Class B of Figure 1 could represent the same set of students in two different subjects or two classes in the same cohort and subject but in different schools.

Second, in addition to the across class variation in test score distributions we can rely on within class differences between ordinal and cardinal measures of achievement to estimate rank effects. To do this, note that the ordinal ranking of each individual will be coarser than the cardinal test scores in small groups, meaning that a change in test scores would not necessarily cause a change in rank. This differential coarseness allows isolating the rank effect under the assumption that there exists an otherwise smooth relationship between primary and secondary school performance. One can consider this to be analogous to a regression discontinuity approach with primary school test scores as the running variable and rank as treatment status, where rank jumps up by one unit whenever the test score exceeds that of the next student in that group.

These two approaches rely on different assumptions regarding the relation between primary and secondary school test scores. The first approach assumes that test scores can be compared across classrooms. However, the same baseline score of X might represent different underlying abilities in different situations (Figure 1). A student with this score in a class with better resources (e.g. teachers), would have a lower ability than an equivalent student who attended a class with worse resources. This is problematic because such environmental factors would impact cardinal measures of achievement but not ordinal measures. Therefore, the rank parameter may pick up information about unobserved ability. To account for any such mean shifting factors, we always include fixed effects at the primary school-subject-cohort (SSC) level. These fixed effects remove the between SSC-group differences in long run attainment growth due to any group-level factor that enters additively and affects all students in the same way. As the median primary school has only 27 students per cohort and the maxium primary class size is 30, we think of these SSC fixed effects as subject-specific class fixed effects. Therefore, when we refer to 'classes,' we are referring to school-subject-cohorts (SSC),note that students will have the same peers and teachers in these classes within a school-cohort. This takes us back to the original thought experiment, Figure 1, where two students (or the same student in different subjects) with the same test score can have different ranks. As a result, we can control for effects of primary achievement non-parametrically by including a dummy variable for each test score.

The second approach further relaxes the assumption that the same test score represents the same underlying ability in different classes. This comes at the cost of assuming a smooth function for the effect of primary on secondary test scores. Now, in addition to the mean-shifting effects (captured by the SSC effects), we allow the transformation of primary to secondary test scores to vary by cohort, subject or school. This is implemented by interacting smooth polynomials of prior achievement with cohort- subject- or school fixed effects.

In our preferred specifications, we allow for a third-order polynomial in primary test scores, SSC fixed effects and a linear rank effect, the estimates of which are robust to more flexible specifications. In our most demanding specifications, we exploit the fact we measure achievement of each student in three subjects. Specifically, we use the variation in ranks within a student across the three subjects by additionally including student fixed effects. This absorbs the average growth rate of the individual across subjects from primary school to secondary school, absorbing any unobserved individual level characteristics or shocks.

To address the common identification challenges of sorting and reflection and unobserved shocks, we exploit unique features of the English educational system also used by Lavy et al.

(2012) and Gibbons and Telhaj (2016). In England, students moving from primary to secondary school have on average 87 percent new peers. As a result, the primary school test scores can be regarded as predetermined measure of achievement and therefore as unaffected by reflection and common unobserved shocks. In addition, schools are not allowed to choose students based on ability or rank. In our robustness section, we test the validity of both of these claims.

In our main analysis, we use administrative data on five cohorts of the entire English state school sector covering over 2.25 million students from the end of primary school to the end of secondary school. In England, all students are tested in English, Math, and Science at the end of primary school (age 11) and twice during secondary school (ages 14 and 16). We use these scores to construct a dataset tracking student performance over time, with an observation for each student in each subject. These national assessments are marked externally from the school and are intended to be absolute measures of individual achievement; hence, scores are not set to a curve at class, school, district, or country level. We use the age-11 baseline score to rank every student among their primary school cohort peers in the three subjects. Our outcome variables are national assessments in secondary school at age 14, age 16 and final year subject choices at age 18. The census data includes gender, Free School Meals Eligibility (FSME) and student ethnicity, for which we control.

Our first main finding is that conditional on highly flexible measures of cardinal performance in primary school, a student's ordinal rank in primary school is predictive of their academic achievement during secondary school. Students achieve higher test scores in a subject throughout secondary school if they had a high rank in that subject during primary school, conditional on a flexible baseline performance measure and SSC effects. The effects of rank that we present are sizable in the context of the education literature, with a one standard deviation increase in rank improving age-14 and age-16 test scores by about 0.08 standard deviations. Estimates that account for individual unobservables, including average ability and any rank effects constant across subjects, are smaller. Here, a one standard deviation increase in rank improves subsequent test scores in that subject by 0.055 within student standard deviations. These effects vary by pupil characteristics, with boys being more affected by their rank than girls throughout the rank distribution, and with students who are FSME not being negatively impacted by being below the median rank, but gaining relatively more from being ranked highly.

Our second finding is the that primary school rank impacts the choice of subjects taken at the end of secondary school. England has an unconventional system where students typically choose to study only three subjects in the last two years of secondary school. These subject choices have long lasting repercussions, as not choosing any STEM subjects at this point removes the option of studying them at university. Here, we find that conditional on achievement, being at the top of the class in a subject during primary school rather than at the median increases the probability of an individual choosing that subject by almost 20 percent. Moreover, being highly ranked in Maths in primary school means that students will be less likely to choose English. We argue that this highlights an undiscovered channel that contributes to the well-documented gender gaps in the STEM subjects and the consequential labor market outcomes (Guiso et al., 2008; Joensen and Nielsen, 2009; Bertrand et al., 2010).

We perform a number of robustness checks to address remaining concerns that would challenge the interpretation of the rank parameter. First, we test the assumption that parents are not sorting to primary schools on the basis of rank. Second, we show our results are robust to functional form assumptions. Third, we check for balancing of class rank on observable characteristics, conditional on test scores. Finally, we address the broader issue of systematic and non-systematic measurement error in the baseline test scores, which are used to generate the rank measures.

We go on to consider several explanations for what could be causing these estimated rank effects. By combining the administrative data with survey data of 12 thousand students, we test the relevance of competitiveness, parental investment (through time or money), school environment favoring certain ranks (i.e., tracking), and confidence. Using our main specification, our third empirical finding is that primary school rank in a subject has an impact on self-reported confidence in that subject during secondary school. In parallel to what we find with regards to academic achievement, we also find that boys' confidence is more affected by their school rank than girls'.

This higher confidence could be indicative of two mechanisms. First, confidence could be reflective of students learning about their own strengths and weaknesses in subjects, similar to Azmat and Iriberri (2010) or Ertac (2006) where students use their test scores and cardinal relative positions to update beliefs. Alternatively, consistent with the Big Fish Little Pond effect, which has been found in many countries and institutional settings (Marsh et al., 2008), confidence due to rank improves non-cognitive skills and lowers the cost of effort in that subject. The idea is that, when surrounded by people who perform a task worse than oneself, one develops confidence in that area. So, if a student is confident in her maths skills, she will be more resilient and have more grit in solving maths exercises compared to another student of the same ability but different confidence. Using a stylized model of students trying to maximise test scores for a given total effort and ability level across subjects, we derive a test to distinguish the Big Fish Little Pond from the leaning mechanism and find evidence in favor of the former. We find these effects exist in the education setting, but there is no reason to believe that the principal of ordinal rank effects do not occur in other setting such as firms, families and academic departments.

It is important to point out that this paper is complementary to–but distinct from–a number of existing literatures. First, any rank effects are a form of peer effect. The classic peer effects papers consider the mean characteristics of others in the group (Sacerdote, 2001; Whitmore, 2005; Kremer and Levy, 2008; Carrell et al., 2009; Lavy et al., 2012), but other relationships have also been considered. For example, (Lavy et al., 2012) use the same data and find no effects of contemporaneous linear-in-mean peer effects in secondary school but find that a one standard deviation increase in the proportion of "bad peers" in a subject lowers student test scores in that subject by 3.3 percent of a standard deviation. The common theme in all of these papers is that individuals benefit from being surrounded by higher performing individuals. In contrast, we find that having had a one standard deviation higher rank, and therefore worse-performing peers, in a subject in primary school increases secondary school test scores by 0.05 standard deviations.[1] The other core difference from the many of the peer papers is that they focus on estimating contemporaneous effects of

[1] Hoxby and Weingarth (2005) introduce the invidious comparison peer effect, where being surrounded by better peers also has negative effects.

peers. This paper instead estimates the impacts of previous peers on individuals' outcomes when surrounded by new peers, conditional on outcomes from the previous peer environment. In doing so, it averts issues relating to reflection and establishes that these effects are long-lasting in students. This is similar to Carrell et al. (2016), who use cohort-variation during elementary school to estimates the causal effect of disruptive peers on long-run outcomes.

This study is also related to the literature on status concerns and relative feedback. Tincani (2015) and Bursztyn and Jensen (2015) find evidence that students have status concerns and will invest more effort if gains in ranks are easier to achieve and Kuziemko et al. (2014) find evidence for last-place aversion in laboratory experiments.These results are similar to findings from non-education settings where individuals may have rank concerns such as in sports tournaments (Genakos and Pagliero, 2012), and in firms with relative performance accountability systems (i Vidal and Nossol, 2011). We differ from this literature because we estimate the effects of rank in a new environment, where status concerns or information about prior ranks do not matter.[2] With regard to the feedback literature, Bandiera et al. (2015) find that the provision of feedback improves subsequent test scores for college students. Specifically relating to relative feedback measures, Azmat and Iriberri (2010) and Azmat et al. (2015) find that their introduction during high school increases productivity in the short run. In contrast, this paper does not examine the contemporaneous reaction to a new piece of information but rather examines student reactions to previous ranking. Finally, the most closely related literature is that on rank itself. These papers account for relative achievement measures and estimate the additional impact of ordinal rank on contemporaneous measures of well-being (Brown et al., 2008) and job satisfaction (Card et al., 2012). We contribute to this literature by establishing long-run effects on objective outcomes.[3]

Besides policy implications (which we discuss in the conclusion), our findings help to reconcile a number of topics in education. Persistent rank effects could partly speak towards why some achievement gaps increase over the education cycle (Fryer and Levitt, 2006; Hanushek and Rivkin, 2006, 2009). Potentially, rank effects amplify small early differences in attainment. A similar argument could be made for the persistence of relative age effects, which show that older children continue doing better than their younger counterparts (Black et al., 2011). Similarly, research on selective schools and school integration has shown mixed results from students attending selective or predominantly non-minority schools (Angrist and Lang, 2004; Clark, 2010; Cullen et al., 2006; Kling et al., 2007; Abdulkadiroglu et al., 2014). Many of these papers use a regression discontinuity design to compare the outcomes of the students that just passed the entrance exam to those that just failed. The common puzzle is that many of these marginal students do not benefit from attending these selective schools.[4] The findings of this paper suggest that potential benefits of prestigious schools may be attenuated through the development of a drop in confidence among

[2] We show in section 2.1 that contemporaneous rank effects at primary school are controlled for in our setting, and if these were transitory would lead to a downward bias in the long-run estimate. Cicala et al. (2017) show that status concerns in peer groups can affect students contemporaneous behavioural and academic outcomes.

[3] Since the publication of the first working paper version of this manuscript, a new literature has emerged that estimates contemporaneous rank effects using the empirical approach put forward by this paper with survey data, giving us full credit e.g., Elsner and Isphording (2017, 2018)

[4] Similar effects are found in the Higher Education literature with respect to affirmative action policies (Arcidiacono et al., 2012; Robles and Krishna, 2012).

these marginal/bussed students, who are also necessarily the low-ranked students in their new school.

Section 2 discusses the empirical strategy and identification. Section 3 describes the English educational system, the data and the definition of rank. Section 4 presents the main results. Sections 5 and 6 show robustness and heterogeneity. Section 7 lays out potential mechanisms and provides additional survey evidence. In section 8, we conclude by discussing other topics in education that corroborate with these findings and possible policy implications.

2 A Rank-Augmented Education Production Function

This sections builds upon a basic education production function set out how to estimate the long run impact of prior rank on outcomes. The focus here is on the empirical estimation and the assumptions required for the identification of the reduced form rank effect, without any direct interpretations of the mechanism that we discuss in section 7.

2.1 Specification

To begin, we consider a basic contemporaneous education production function using the framework as set out in Todd and Wolpin (2003). For student i in primary school j, studying subject s in cohort c and in time period $t = [0, 1]$:

$$Y_{ijsc}^{t} = \boldsymbol{x_i'\beta} + v_{ijsc}^{t}$$

$$v_{ijsc}^{t} = \mu_{jsc} + \tau_i + \varepsilon_{ijsc}^{t}$$

where Y denotes academic achievement determined by x_i a vector of observable non-time varying characteristics of the student and v_{ijsc}^{t} representing the unobservable factors. Here β represents the permanent impact of these non-time-varying observable characteristics on academic achievement. Applied to our setting, students attend primary school in $t = 0$ and then secondary school in $t = 1$. The error term v_{ijsc}^{t} has three components. μ_{jsc} represents the permanent unobserved effects of being taught subject j in primary school s in cohort c. This could reflect that the effect of a teacher being particularly good at teaching maths in one year but not English, or that a student's peers were good in English but not in science; τ_i represents permanent unobserved student characteristics, which includes any stable parental inputs or natural ability of the child; ε_{ijsc}^{t} is the idiosyncratic time-specific error, which includes secondary school effects. Under this restrictive specification only ε_{ijsc}^{t} could cause relative change in student performance between primary and secondary periods, as all other factors are permanent and have the same impact over time.

This is a restrictive assumption, as the impacts of observable and unobservable characteristics are likely to change as the student ages. For example, neighborhood effects may grow in importance as the child grows older. Therefore, we relax the model by allowing for time-varying effects of these characteristics:

$$Y^t_{ijsc} = \boldsymbol{x_i'\beta} + \boldsymbol{x_i'\beta^t} + \beta_{Rank}R_{ijsc} + \beta^t_{Rank}R_{ijsc} + v^t_{ijsc}$$

$$v^t_{ijsc} = \mu_{jsc} + \mu^t_{jsc} + \tau_i + \tau^t_i + \varepsilon^t_{ijsc}$$

where β^t allows for the effect of student characteristics to vary over time. We have also distinguished the characteristic of interest from other student characteristics, that being the achievement rank of student i in subject s in cohort c and in primary school j, denoted by R_{ijsc}. Like the other characteristics, primary rank can have a permanent impact on outcomes, β_{Rank}, and a period dependent impact, β^t_{Rank}. This specification also allows for the unobservables to have time-varying effects by additionally including τ^t_i and μ^t_{jsc} in the error term.

Given this structure we now state explicitly the conditional independence assumption that needs to be satisfied for estimating an unbiased impact of primary school rank β^1_{Rank} on secondary outcomes Y^1_{ijsc}.

$$Y^1_{ijsc} \perp R_{ijsc} | x_i, \mu_{jsc}, \mu^t_{jsc}, \tau_i, \tau^t_i \in R$$

To achieve this, we require measures of all factors that may be correlated with rank and final outcomes. Conditioning on baseline test scores, Y^0_{ijsc}, absorbs all non-time-varying effects that affect primary period outcomes to the same extent as secondary period outcomes: $\boldsymbol{x_i'\beta}$, $\beta_{Rank}R_{ijsc}$, μ_{jsc}, τ_i. Moreover, Y^0_{ijsc} captures any factors from the first period that impact academic attainment in the primary period only: $\boldsymbol{x_i'\beta^0}$, $\beta^0_{Rank}R_{ijsc}$, μ^0_{jsc}, τ^0_i. Critically, the lagged test scores absorb any type of contemporaneous peer effect during primary school, including that of rank. In this sense, we are not estimating a traditional peer effect with a focus on the contemporaneous impact of peers on academic achievement. Instead, we are estimating the effects of peers in a previous environment on outcomes in a subsequent time period. Therefore, we can express second period outcomes as a function of prior test scores, primary rank, student characteristics and unobservable effects. Using lagged test scores means the estimated parameters are those from the secondary period.

$$\begin{aligned} Y^1_{ijsc} = {} & f(Y^0_{ijsc}(\boldsymbol{x_i'\beta},\ \boldsymbol{x_i'\beta^0},\ \tau_i,\ \beta^0_{Rank}R_{ijsc},\ \beta_{Rank}R_{ijsc},\ \mu_{jsc},\ \tau^0_i,\ \mu^0_{jsc})) \\ & + \beta^1_{Rank}R_{ijsc} + \boldsymbol{x_i'\beta^1} + \mu^1_{jsc} + \tau^1_i + \varepsilon^1_{ijksc} \end{aligned} \tag{1}$$

This leads to our first estimation equation:

$$Y^1_{ijsc} = f(Y^0_{ijsc}) + \beta^1_{Rank}R_{ijsc} + \boldsymbol{x_i'\beta^1} + \boldsymbol{SSC'_{jsc}}\boldsymbol{\gamma^1} + \varepsilon^1_{ijsc} \tag{2}$$

$$\varepsilon^1_{ijsc} = \tau^1_i + v^1_{ijsc}$$

Here, we allow the functional form of the lagged dependent variable to be flexible and examine how changes to the functional form of $f(Y^0_{ijsc})$ change our results. The inclusion of primary School-Subject-Cohort (SSC) dummies, $\boldsymbol{SSC_{jsc}}$, accounts for the lasting impacts of being taught a specific subject in a particular primary school and cohort on secondary period outcomes.[5]

In addition, including SSC fixed effects ensures that we estimate ordinal rank effects rather than cardinal effects. The cardinal effect is the effect on secondary outcomes of being a certain distance from the primary class mean in terms of achievement. This is accounted for by the combination of baseline achievement and SSC effects, leaving $\beta^1_{Rank}$to pick up the impact of ordinal rank only.

The SSC effects remove μ^1_{jsc} from the error term. As a result, the remaining unobservable factor ε^1_{ijsc} is comprised of two components: unobserved individual-specific shocks that occur between $t = 0$ and $t = 1$, τ^1_i, and an idiosyncratic error term υ_{ijsc}. Since we have repeated observations over three subjects for all students, we stack the data over subjects in our main analysis, so that there are three observations per student. To allow for unobserved correlations, in all of our estimations, we cluster the error term at the level of the secondary school.[6] Having multiple observations over time for each student also allows us to to recover τ^1_i, the average growth of individual i in secondary period. However it is worth spending some time interpreting what the rank coefficient represents without its inclusion. Being ranked highly in primary school may have positive spillover effect in other subjects. Allowing for individual growth rates during secondary school period absorbs such spillover effects. Therefore, leaving τ^t_i in the residual means that the rank parameter is the effect of rank on the subject in question and the correlation in rankings from the two other subjects.

Our second estimation specification includes an individual dummy θ_i for each student, which removes τ^1_i from the error term and captures the average student growth rate across subjects. Note that, despite using panel data, this estimates the individual effect across subjects and not over time. When allowing for student effects, we effectively compare rankings within a student across subjects conditional on prior subject specific attainment. This accounts for unobserved individual effects that are constant across subjects such as competitiveness or general confidence.

$$Y^1_{ijsc} = f(Y^0_{ijsc}) + \beta^1_{Rank} R_{ijsc} + \boldsymbol{x'_i \beta^1} + \boldsymbol{SSC'_{jsc}} \gamma^1 + \boldsymbol{\theta'_i \tau^1_i} + \upsilon^1_{ijsc} \tag{3}$$

In this specification the rank parameter represents only the increase in test scores due to subject-specific rank, as general gains across all subjects are absorbed by the student effect. This can be interpreted as the extent of specialisation in subject s due to primary school rank. For this reason, and because of the removal of other co-varying factors, we expect the coefficient of the rank effect in specification 3 to be smaller than those found in 2.

Finally, to also investigate potential non-linearity in the effect of ordinal rank on later outcomes (i.e., effects driven by students who are top or bottom of the class), we replace the linear ranking

[5] In section 7.2 we discuss results that additionally account for secondary-subject-class-level effects as a potential channel.

[6] The treatment occurs at the primary SSC level and therefore a strong argument can be made for this being the correct level at which to cluster the standard errors. However, we chose to cluster the standard errors at the secondary school level for two reasons. The first is that all of the outcomes occur during the secondary school phase, where students from different primary schools will be mixing and will be attending the same secondary school for all subjects. Therefore, we thought it appropriate to partially account for this in the error term. Secondly, clustering at the secondary school level rather than the primary SCC is considerably more conservative, generating standard errors that are 50 percent larger. Standard errors for other levels of clustering including primary, primary SSC, secondary SSC, and two-way clustering at the student and school-subject-cohort, for both primary and secondary schools, are available upon request.

parameter with indicator variables according to vingtiles in rank, $\sum_{\lambda=1}^{20} I\{\lambda_{ijsc}\}$, plus additional indicator variables for those at the top and bottom of each school-subject-cohort (the rank measure is defined in section 3). This can be applied to all the specifications presented. In the case of specification 3, this results in the following estimation equation:[7]

$$Y_{ijsc}^{1} = \beta_{R=0}^{1} Bottom_{ijsc} + \sum_{\lambda=1}^{20} \left(\beta_{\lambda}^{1} I\{\lambda_{ijsc}\} \right) + \beta_{R=1}^{1} Top_{ijsc} + f(Y_{ijsc}^{0}) + \boldsymbol{x_i'\beta^1} + \boldsymbol{SSC_{jsc}'}\boldsymbol{\gamma^1} + \varepsilon_{ijsc}^{1} \tag{4}$$

In summary, if students react to ordinal information in addition to cardinal information, then we expect the rank parameter β_{Rank} to have a significant effect. Given this structure, the conditional independence assumption that needs to be satisfied for estimating an unbiased rank parameter is the following: conditional on prior test scores (accounting for all non-time varying effects and inputs to age 11), student characteristics, and primary SSC level effects, we assume there would be no expected differences in the students' secondary school outcomes except those driven by rank. The remaining concern is that unobserved shocks at $t = 0$ that correlate with rank at the individual-subject level and do not affect Y^0 then do affect Y^1. Examples for such transitory shocks are primary school teachers teaching to specific parts of the distribution whose impact is only revealed in secondary school, non-linear transitory peer, or any other transitory effects potentially generated through measurement error in the age-11 tests. We address these issues in section 5.

2.2 Identification

Specification 2 conditions on baseline achievement, achievement rank within a primary subject school and cohort (SSC) and primary SSC fixed effects. The key question to address is, how can variation in rank conditional on test scores within a group exist? The answer requires functional form assumptions, that rank measures are discrete, and that performance measures are continuous. Naturally occurring variation in the spacing of achievement in primary school classes gives rise to variation in rank within and across classes. Within class, marginal increases in baseline performance will generate discrete increases in rankings within a group, and this non-linearity can be exploited to identify effects of rank. Across classes, we can compare students with identical baseline achievement but different local ranks. To be specific, in order to identify the ordinal rank effect from the cardinal performance measure, we require one of two functional form assumptions. First, that the function between baseline performance and the outcome is smooth. The second is that any non-smooth function is similar across groups, after allowing for mean differences (SSC-effects). Critically, with our data we can relax one of these assumptions in turn.

For the purposes of exposition, let us first consider the simplest specification with only one subject and one class. We also assume a smooth and true linear relationship between the baseline score and the outcome of interest, as well as between rank and the outcome. Let us, for now, also assume that all variables are measured without noise and that group assignment is random.

[7] Estimates are robust to using deciles in rank rather than vingtiles and are available upon request.

$$Y_i^1 = \alpha + \beta^1 Y_i^0 + \beta_{Rank}^1 R_i + \epsilon_i \tag{5}$$

As before, the outcome of individual i in period 1, Y_i^1, is determined by their baseline performance, Y_i^0, and their rank in the group in period $t = 0$ based on the baseline performance, R_i. In this situation, the effect of rank is identified due to a functional form assumption and the nature of the setting. To be precise, specification 5 implies that that a unit increase in Y_i^0 will increase the outcome by β. Assuming this group is small, defined by having fewer individuals than possible performance levels in Y_i^0, a unit increase Y_i^0 will not always lead to increases in rank because there is not always another member in that group with a baseline score that is just one unit higher. As long as members of the group are not equally spaced in terms of baseline performance and as long as there are at least three of them, there will be a non-linear relationship between rank and the performance in that group. These distributional differences occur naturally and allow us to estimate the impact of ordinal rank independent of cardinal achievement measures. One could think of this form of identification as being analogous to a regression discontinuity approach where any changes in the outcome that coincide with the discontinuous change in rank would be attributed to the rank parameter. Indeed, if there existed no impact of rank with $\beta_{Rank}^1 = 0$, there would be a smooth relationship between the baseline and the outcome. Note that because of the smoothness assumption, if there are multiple groups g, it is possible to estimate the effect of the prior test score separately for each group, β_g^1. Applied to our setting, this allows the effects of prior test scores to vary by class, subject and cohort.

One concern with this approach to estimating rank effects is the necessary reliance on the functional form assumption regarding prior performance. Imposing the wrong functional form could cause the rank parameter to pick up the impact of cardinal measures of academic achievement on later outcomes, even if the true effect of rank is zero. To alleviate such concerns, this assumption can be readily relaxed by controlling for Y_i^0 more flexibly by using higher order polynomials. Ideally, one would like to allow the relationship between past and current test scores to be fully flexible or non-parametric, such that every test score can have a distinct impact. This can be achieved by including a dummy variable for each value Y_i^0. However, in a situation with only one class, a non-parametric measure of Y_i^0 and R_i are non-separable. This is because two individuals form the same group g, and with the same score Y_{ig}^0, will always have the exact same rank R_{ig}^0. However, if information on multiple groups is available, it is possible to estimate the following specification:

$$Y_{ig}^1 = f\left(Y_{ig}^0\right) + \beta_{Rank}^1 R_{ig} + \boldsymbol{D_g^{1\prime}} \boldsymbol{\gamma_g} + \epsilon_{ig} \tag{6}$$

There are two key differences compared to specification 5: the baseline score is now allowed to affect Y_{ig}^1 in a flexible way, where in our case a dummy variable is entered for every possible percentile $f\left(Y_{ig}^0\right) = \sum_{p=1}^{100}\left(\beta_p I\{\mathrm{p} = \mathrm{Y}_{\mathrm{ig}}^0\}\right)$. This is done at the cost of having to assume the same functional form across groups so that the functional form $f\left(Y_{ig}^0\right)$ cannot vary between groups. Note that D_g^1 allows for means shifts in the outcomes at the group level even if the flexible function is the same.[8] Specification 6 takes us back to our original thought experiment with two individuals in different groups with different ranks but the identical baseline scores (Figure 1). Comparing such individuals across groups thus allows the estimation of the rank effects even conditional on non-parametric baseline scores. Of course, this is only possible if there are enough groups and enough individuals with identical baseline scores but different ranks in the data. In the following section we provide evidence that we have sufficient support for this strategy when comparing ranks and test scores across SSCs.

In summary, we use the very nature of ordinal rank and cardinal test score measures to generate discontinuities within classrooms, such that marginal increases in test scores only occasionally generate changes in rank. Our specification exploits these discontinuities and attributes corresponding changes in outcomes to the rank measure. By including group fixed effects, the baseline performance measure becomes a measure of relative cardinal achievement, leaving the rank parameter to reflect the impact of ordinal rank. This strategy is related to and builds on a long series of work in empirical education economics that exploits naturally ocurring variation conditional on fixed effects going back to Hoxby (2000). We use the quasi-random variation in the test score distributions within and across classrooms, which allows us to relax functional form assumptions of prior to future performance, either requiring smoothness but allowing it to vary by group, or assuming constant effects over groups but allowing for a non-smooth relationships.

3 Institutional Setting, Data, and Descriptive Statistics

This section explains the administrative data and institutional setting in England that we use to estimate the rank effect using the specifications of section 2.1.

3.1 The English School System

The compulsory education system of England is made up of four Key Stages. At the end of each stage, students are evaluated in national exams. Key Stage 2 is taught during primary school between the ages of seven and eleven. English primary schools are typically small with the median cohort size of 27 students. The mean primary school class size also is 27 students (Falck et al., 2011), so referring to cohort-level primary school rank in a subject is almost equivalent to the class rank in that subject. At the end of the final year of primary school, when the students are age 11, they take tests in English, maths and science. These tests are externally marked on absolute attainment on a national scale of zero to 100 and form our measure of baseline achievement.

Students then transfer to secondary school, where they start working towards Key Stage 3. During this transition, the average primary school sends students to six different secondary schools, and secondary schools typically receive students from 16 different primary schools. Hence, upon arrival at secondary school, the average student has 87 percent new peers. This large re-mixing

[8] Our main specification 2 assumes the same effect across groups, but in Appendix Table A.4 we show results where we allow the coefficients to vary by school, subject and cohort. While there are some differences, none would change the qualitative conclusions we reach in our paper. We regard this as important evidence that the restriction of constant baseline effects across schools are not creating the rank effect.

of peers is useful, as it allows us to estimate the impact of rank from a previous peer group on subsequent outcomes. Importantly, since 1998, it is unlawful for schools to select students on the basis of ability; therefore, admission into secondary schools does not depend on end-of-primary test scores or student ranking.[9] This means that the age-11 exams are low-stakes with respect to secondary school choice. Key Stage 3 takes place over three years, at the end of which all students take again take national examinations in English, maths, and science. Again, these age-14 tests are externally marked out of 100.[10]

At the end of Key Stage 3, students can choose to take a number of subjects (GCSEs) for the Key Stage 4 assessment, which occurs two years later at the age of 16 and marks the end of compulsory education in England. The final grades consist of nine levels (A*, A, B, C, D, E, F, G, U), to which we have assigned points according to the Department for Education's guidelines (Falck et al., 2011). However, students have some discretion in choosing the number, subject and level of GCSEs they study. Thus, GCSE grade scores are inferior measures of student achievement compared to age-14 examinations, which are on a finer scale and where all students are examined in the same compulsory subjects. As students are tested in the same three compulsory subjects at ages 11 and 14, we focus on age-14 test scores as the main outcome measure, but we also present results for the high-stakes age 16 examinations.

After Key Stage 4, some students choose to stay in school to study A-Levels, which are a precursor for university level education. This constitutes a high level of specialisation, as students typically only enroll in three A-Level subjects out of a set of 40. For example, a student could choose to study biology, economics, and geography, but not English or maths. Importantly, students' choices of subjects limit their choice-sets of majors at university and so will have longer run effects on careers and earnings (Kirkeboen et al., 2016). For example, chemistry as an A-Level is required to apply for medicine degrees and math is a prerequisite for studying engineering.[11] To study the long run impact of primary school ranking on students, we examine the impact of rank on the likelihood of choosing to stay on at school and of studying that subject/STEM-subjects at A-Level.

3.2 Student Administrative Data

The Department for Education collects data on all students and all schools in state education in England in the National Pupil Database (NPD).[12] This contains the school's attended and demographic information (gender, Free School Meals Eligible (FSME) and ethnicity). The NPD also tracks student attainment data throughout their Key Stage progression.

We extract a dataset that follows the population of five cohorts students, starting at age of 10/11

[9] The Schools Standards and Framework Act 1998 made it unlawful for any school to adopt selection by ability as a means of allocating places. A subset of 164 schools (five percent) were permitted to continue to use selection by ability. These Grammar schools administer their own admission tests independent of KS2 examinations and are also not based on student ranking within school.

[10] There is no skipping or repeating of grades in the English education system.

[11] For the full overview of subjects that can be chosen, see: http://www.cife.org.uk/choosing-the-right-a-level-subjects.html

[12] The state sector constitutes 93% of the student population in England.

in the final year of primary school when students take their Key Stage 2 examinations through to age 17/18 when they complete their A-Levels. The age-11 exams were taken in the academic years 2000-01 to 2004-05; hence, it follows that the age-14 examinations took place in 2003-04 to 2007-08 and that the data from completed A-Levels comes from the years 2007-08 to 2010-12.[13]

First, we imposed a set of restrictions on the data to obtain a balanced panel of students. We use only students who can be tracked with valid age-11 and age-14 exam information and background characteristics. This comprises 83 percent of the population. Secondly, we exclude students who appear to be double counted (1,060) and whose school identifiers do not match within a year across datasets. This excludes approximately 0.6 percent of the remaining sample (12,900). Finally, we remove all students who attended a primary school whose cohort size was smaller than 10, as these small schools are likely to be atypical in a number of dimensions. This represents 2.8 percent of students.[14] This leaves us with approximately 454,000 students per cohort, with a final sample of just under 2.3 million student observations, or 6.8 million student subject observations (each student-subject pair is distinct observation).

The Key Stage test scores at each age group are percentalised by subject and cohort, so that each individual has nine test scores between zero and 100 (ages 11, 14, and 16). This ensures that students of the same national relative achievement have the same national percentile, as a given test score could represent a different ability in different years or subjects. This does not impinge on our estimation strategy, which relies only on variation in test score distributions at the SSC level.

Table 1 shows descriptive statistics for the estimation sample. Given that the test scores are represented in percentiles, all three subject test scores at age 11, 14, and 16 have a mean of around 50, with a standard deviation of about 28.[15] Almost sixty percent of students decide to stay and continue their education until the A-Levels, which are the formal gateway requirement for university admission. Of the many subjects to choose from, about 14 percent choose to sit an A-Level exam in English, while in maths and science the proportions are about 9 percent and 11 percent, respectively.

Information relating to the background characteristics of the students is shown in the lower panel of Table 1. Half of the student population is male, over four-fifths are white British, and about 15 percent are FSME students, a standard measure of low parental income. The within student standard deviation across the three subjects, English, maths, and science, is 12.68 national percentile points at age 11, with similar variation in the age-14 tests. This is important, as it shows that there is variation within student which is used in student fixed effects regressions.

3.3 Measuring Ordinal Rank

As explained in section 3.1, all students take the end-of-primary national exam at age-11. These are finely and externally graded between zero and 100. We use these scores to rank students in each

[13] The analysis was limited to five cohorts as from year 2008-09 the external age-14 examinations were replaced with teacher assessments.

[14] Estimations using the whole sample are very similar, only varying at the second decimal point. Contact authors for further results.

[15] Age-16 average percentile scores have lower averages due to the coarser grading scheme.

subject within their primary school cohort. Note that these ranks are computed only for students in our estimation sample.

As previously noted, primary schools are small. However, since there exists some differences in school cohort sizes, in order to have a comparable local rank measure across schools, we cannot use the ordinal rank directly. Instead, we transform the rank position into a local percentile rank in the following way:

$$R_{ijsc} = \frac{n_{ijsc} - 1}{N_{jsc} - 1}, R_{ijsc} \in \{0, 1\} \tag{7}$$

where N_{jsc} is the cohort size of school j in cohort c of subject s. An individual i's ordinal rank position within this set is n_{ijsc}, which increases in test score to a maximum of N_{jsc}. R_{ijsc} is the cohort-size adjusted ordinal rank of students that we use in the estimations. For example, a student who is the best in a cohort of 21 students ($n_{ijsc} = 21$, $N_{jsc} = 21$) has $R_{ijsc} = 1$ and so does a student who is the best in a cohort of 30. Note that this rank measure will be uniformly distributed and bounded between zero and 1, with the lowest ranked student in each school cohort having $R = 0$.[16] Similar to the non-transformed ordinal rank position, this transformed ordinal rank score does not carry cardinal information (i.e. information about relative ability distances). For the ease of exposition, for the reminder of this paper, we will refer to R_{ijsc} as the ordinal rank, rather than as the local percentile rank or as the cohort-size adjusted ordinal rank. Panel A of Table 1 shows descriptive statistics of the rank variable.

Given this measurement of rank, it is relevant to consider how students will know about their academic rank too. In fact, while we as researchers have full access to the test score data, rather than receiving these finely graded scores, students are instead given only one of five broad attainment levels. The lowest performing students are awarded level 1. The top performing students are awarded level 5. These levels are broad and coarse measures of achievement, with 85% of students achieving levels 4 or 5.[17] Therefore neither the students nor teachers are informed of this ranking based on these age-11 test scores. Rather, we take this rank measure as a proxy for perceived ranking based on interactions with peers over the previous six years of primary school, along with repeated teacher feedback. We assume test performance to be highly correlated with everyday classroom performance, and representative of previous performance on any informal class examinations.[18]

While we cannot know if students' academic rank based on these age-11 test scores are good proxy for student perceptions, we have three facts that support this claim. First, there is a long-standing physiological literature that has established that individuals have accurate perceptions of their rank within a group but not of their absolute ability(e.g. Anderson et al., 2006). Second,

[16] In the case of ties in test scores both students are given the lower rank.

[17] The students also appear not to gain academically just from achieving a higher level. Using a regression discontinuity design across these achievement levels, where the underlying national score is the running variable, shows no gains for those students who just achieved a higher level.

[18] In English Primary schools it is common for students to be seated at tables of four and for tables to be set by pupil ability. Students can be sat at the 'top table' or the 'bottom/naughty table'. This would make class ranking more salient and could assist students in establishing where they rank amongst all class members through a form of batch algorithm (e.g. 'I'm on top table, but I'm the worst, therefore I'm fourth best').

we find using merged survey data that, conditional on test scores, students with higher ranks in a subject have higher confidence in that subject (section 7.4.1). And third, if individuals (students, teachers or parents) had no perception of the rankings, then we would not expect to find an impact of our rank measurement at all. To this extent, the rank coefficient β_{Rank} from section 2 would be attenuated and we are estimating a reduced form of perceived rank using actual rank. In section 5.3 we also simulate increasingly large measurement errors in the age-11 test scores, which we use to calculate rank, to document what would occur if these tests were less representative of students abilities and social interactions. We show that increased measurement error in baseline achievement slightly attenuates the rank estimate.

3.4 Survey Data: The Longitudinal Study of Young People in England

Additional information about a sub-sample of students is obtained through a representative survey of 16,122 students from the first cohort. The Longitudinal Survey of Young People in England (LSYPE) is managed by the Department for Education and follows a cohort of young people, collecting detailed information on their parental background, academic achievements and attitudes.

We merge survey responses with our administrative data using a unique student identifier. However, the LSYPE also surveys students attending private schools that are not included in the national datasets. In addition, students that are not accurately tracked over time have been removed. In total 3,731 survey responses could not be matched. Finally, 823 state school students did not fully answer these questions and could not be used for the confidence analysis. Our final dataset of confidence and achievement measures contains 11,558 student observations. Even though the survey does not contain the attitude measures of every student in a school cohort, by matching the main data, we will know where each LSYPE-student was ranked during primary school.[19] This means we are able to determine the effect of rank on confidence, conditional on test scores and SSC fixed effects.

In the LSYPE, at age-14, students are asked how good they consider themselves in the subjects English, maths, and science. We code five possible responses in the following way: 2 "Very Good"; 1 "Fairly Good"; 0 "Don't Know"; -1 "Not Very Good"; -2 "Not Good At All". We use this simple scale as a measure of subject specific confidence. While this is more basic than surveys that focus on confidence, it does capture the essence of the concept with a mean of 0.92 and standard deviation of 0.95 (Table 2 panel A). The LSYPE respondents are very similar to students in the the main sample, with the mean age 11 scores always being within one national percentile point. The LSYPE sample is also has a significantly higher proportion of FSME (18 versus 14.6 percent) and minority (33.7 versus 16.3 percent) students. Although this is to be expected as the LSYPE over sampled students from disadvantaged groups.[20]

The LSYPE also contains a lot of detailed information relating to the parent(s) of the student, which we present in panel B of Table 2. We use information on parent characteristics to test for

[19] This is the first research to merge LSYPE responses to the NPD for primary school information.

[20] Appendix Table A.1 presents the raw differences and their accociated standard errors.

sorting to primary schools on the basis of rank conditional on performance in section 5.1. These characteristics are represented a set of indicator variables, parental qualifications as defined by if any parent has a post secondary qualification (32 percent), and gross household income above £33,000 (21.9 percent). These characteristics are constant within a student. Therefore, to test if there is sorting to primary schools by subject, we have classified parental occupation of each parent to each subject. Then, an indicator variable is created for each student-subject pair to capture if they have a parent who works in that field. For example, a student who has parents working as a librarian and a Science Technician would have parental occupation coded as English and Science.[21]

Finally, information regarding parental time and financial investments in schooling is used to explore possible mechanisms in section 7.2. It is possible that parents may adjust their investments into their child according to student rank during primary school. Therefore we have codified four forms of parent self reported time investment: (1) the number of parents attending most recent parent evening; (2) whether any parent arranged a meeting with the teacher; (3) how often a parent talks to the teacher; and (4) how personally involved does the parent feel in young persons school life. The frequency of meetings with teacher is coded: 1 "Never", 2 "Less than once a term", 3 "At least once a term", 4 "Every 2-3 weeks", 5 "At least once a week" and parental involvement is coded: 1 "Not Involved At All", 2 "Not Very Involved" 3 "Fairly Involved", 4 "Very Involved". In our sample on average, 1.2 parents attended the last parents evening, 23.5 percent had organised a meeting with the teacher, they have meetings less than once a term (2.12) and felt fairly involved in the child's school life (2.97).

.

4 Estimation

Before turning to the estimation results, we illustrate the variation in rank we use for a given test score relative to the mean demonstrated in Figure 1. Figure 2 replicates the stylised example from Figure 1 using six primary school classes in English from our data. Each class has a student scoring the same minimum, maximum, and have a mean test score of 55 (as indicated by the dashed grey line) in the age-11 English exam. Each class also has a student at the 92nd percentile. Given the different test score distributions, each student scoring 92 has a different rank. This rank is increasing from school one to six with ranks R of 0.83, 0.84, 0.89, 0.90, 0.93 and 0.94 respectively, despite all students having the same absolute and relative to the class mean test scores. Figure 3 extends this example of the distributional variation by using the data from all primary schools and subjects in our sample. Here, we plot age-11 test scores, de-meaned by primary SSC, against the age-11 ranks in each subject. The vertical thickness of the points indicates the support throughout the rank distribution. For the median student in a class, we have wide support for in-sample inference from $R = 0.2$ to $R = 0.8$. This means that we have sufficient naturally occurring variation

[21] We use the "Parental Standard Occupational Classification 2000" to group occupations into Science, Math, English and Other in the following way. Science (3.6%): 2.1 Science and technology, 2.2 Health Professionals, 2.3.2 Scientific researchers, 3.1 Science and Engineering Technicians. Math (3.2%): 2.4.2 Business And Statistical Professionals, 3.5.3 Business And Finance Associate Professionals. English (1.4%): 2.4.5.1 Librarians, 3.4.1 Artistic and Literary Occupations, 3.4.3 Media Associate Professionals. Other: Remaining responses.

in our data to include even non-parametric controls for the primary school baseline achievement measure.

4.1 Effect of Rank on Age-14 Test Scores

To begin the discussion of the results, we present estimates of the impact of primary school rank on age-14 test scores. The estimates are reported in the first three columns of Table 3. Column 1 only includes controls for prior test scores and primary SSC effects, column 2 adds student characteristics (ethnicity, gender, ever FSME), and column 3 adds individual fixed effects. All specifications allow for up to a cubic relationship with age-11 test scores.

Column 1 shows that the effect of being ranked top compared to bottom *ceteris paribus* is associated with a gain of 7.946 national percentile ranks (0.29 standard deviations). When accounting for pupil characteristics, there is an insignificant change to 7.894, implying minimal rank-based sorting of students on observables, which we go on to test formally in the section 5. This is a large effect in comparison with other student characteristics typically included in growth specifications. For example, females' growth rate is 1.398 national percentile points higher than males', and FSME students on average have 3.107 national percentile points lower growth rate than non-FSME students. However, given the distribution of test scores across schools, very few students would be bottom ranked at one school and top at another. A more useful metric is to describe the effect size in terms of standard deviations. A one standard deviation increase in rank is associated with increases in later test scores by 0.084 standard deviations or 2.35 national percentile points. Finally, another way to gauge the relative importance of rank compared to traditionally important factors is to examine changes in the mean squared error. In a specification with only prior test scores and SSC effects, including a gender term reduces the mean square error by 0.25, for ethnicity it reduces by 0.28, and the introduction of the rank parameter reduces the mean squared error by 0.31.

Column 3 shows the estimates that also include student fixed effects (specification 3). Recalling from section 2, conditioning on student effects allows for individual growth rates, which absorb all student level characteristics constant across subjects. As a result, male, FSME and minority are dropped from this specification. Since students attend the same primary and secondary school for all subjects, any general school quality or school sorting is also accounted for.

As expected, the within student estimate is considerably smaller as the student effect also absorbs spillover effects gained through high ranks in other subjects, and so is identifying the relative gains in that subject. The effect from moving to the bottom to top of class *ceteris paribus* increases the national percentile rank by 4.562 percentiles, as we see in column 3. Scaling this using the within student standard deviation of the national percentile rank (11.32), this is equivalent to an effect size of 0.40 standard deviations. In terms of effect size, given that a standard deviation of the rank within student is 0.138, for any one standard deviation increase in rank test scores increase by about 0.056 standard deviations.[22] The difference between columns 2 (7.894) and 3 (4.562) can be interpreted as an upper bound of the gains from spillovers between subjects. We examine

[22] For students with similar ranks across subjects the choice of specialisation could be less clear. Indeed, in a sample of the bottom quartile of students in terms of rank differences, the estimated rank effect is 25% smaller than those from the top quartile. Detailed results available on request.

spillovers directly when discussing subject-specific results in section 6.1.

4.2 Effect of Rank on Key Stage 4 Outcomes (Age 16)

Columns 4 to 6 of Table 3 show an equivalent set of results for the same students two years later, taking the national exams at the end of compulsory education. The three core subjects (English, math, and science) are again tested at age 16. The impact of primary school rank on test performance has only marginally dropped in all specifications between ages 14 and 16. Comparing columns 2 and 5, being at the top of class compared to the bottom during primary school increases age 16 test scores by 6.389 percentile points compared with 7.894 at age-14. At age 16, a one standard deviation increase in primary rank improves later test scores by 1.89 national percentiles. Similarly, the impact on test scores using the within student variation has decreased, but remains significant.

4.3 Effect of Rank on A-Level Choices (Age 18)

After the examinations at age 16, students can chose to stay in school and study for A-Levels which are the key qualifications required to study any associated subjects at university. To this end, we estimate the impact of primary school rank in a specific subject on the likelihood of choosing to study that same subject for A-Levels.[23] These results are presented in columns 7 to 9 of Table 3, with a binary outcome variable being whether or not the student completed an A-Level related to that subject. In this linear probability model, conditional on prior test scores, student characteristics and SSC effects, students at the top of the class in a subject compared to being ranked at the bottom are 2.5 percentage points more likely to choose that subject as an A-Level. On average, roughly one in ten students complete these subjects at A-Level. Assuming an linear relationship, a student who was at the 75th rather than 25th rank position at primary in a subject would, therefore, be 11.9 percent more likely to complete a course related to that subject for an A-Level seven years later on.

For A-Level completion, introducing the student fixed effects increases the estimated effect size to 3.5 percentage points in column 9. This may reflect the capacity constraints for this outcome as students are limited to taking three subjects only. Being more likely to take one subject, despite increasing general confidence, could result in being less likely to take another subject, resulting in negative subject-spillover effects. This, as well as the linearity assumption, are investigated in section 6.2. Before proceeding in this direction, we examine the robustness of our main estimate.

5 Robustness

This section examines the robustness of our main results in four dimensions. First, we test for balancing on student level observable characteristics, and go further by testing if parents who will generate high growth in a particular subject systematically sort their children to primary schools, such that their child will have a particularly high/low rank in that subject. Second, we examine

[23] Students that did not take on any of the core subjects or have left school are included in the estimations.

if systematic or non-systematic measurement error in the age-11 achievement test scores would result in spurious rank effects. Third, we check functional form assumptions. Finally, we address miscellaneous concerns such as school sizes, classroom variance and the proportion of new peers at secondary school.

5.1 Rank-Based Primary School Sorting

A core tenet of this paper is that a students rank in a subject is effectively random conditional on achievement. This would not be the case if parents were selecting primary schools based on the rank that their child would have. To do this parents would need to know the ability of their child and of all their potential peers by subject, which is unlikely to be the case when parents are making this choice when their child is only four years old.[24] Moreover, typically parents want to get their child into the best school possible in terms of average grades (Rothstein, 2006; Gibbons et al., 2013), which would work against any positive sorting by rank. We provide two pieces of evidence to test for sorting: by parental characteristics and by student characteristics.

We are most concerned with parental investments that would vary across subjects, because such investments would not be fully accounted for with the student fixed effects specification. One such parental characteristic that could impact investments by subject is the occupation of the parent. Children of scientists may both have a higher initial achievement and a higher growth in science throughout their academic career, due to parental investment or inherited ability. The same could be said about children of journalists for English and children of accountants in maths. This does not bias our results, as long as conditional on age-11 test scores parental occupation is orthogonal to primary school rank. However, if these parents sort their children to schools such that they will be the top of class and also generate higher than average growth, then this would be problematic.

We test for this by using the LSYPE sample, which has information on parental occupation, which we have categorised into subjects (for details see section 3.4). Panel A of Table 4 establishes that this is an informative measure of parental influence by subject, by regressing age-11 test scores on parental occupation and school subject effects. Students have higher test scores in a subject if their parents work in a related field. This is taken one step further in column 2, which shows that even after accounting for additional student fixed effects this measure of parental occupation is a significant predictor of student subject achievement. In the first row of panel B we test for the balancing of parental occupation for violation of the orthogonality condition, by determining if primary school rank predicts predetermined parental occupation, conditional on achievement. We find that there is no correlation between rank and parental occupations, for specifications which do or do not account for student effects. This implies that parents are not selecting primary schools on the basis of rank for their child.

The next two rows of Table 4 test to see if student rank is predictive of other predetermined parental characteristics. These are parental education, as defined by either parent having a post

[24] Parents could infer the likely distributions of peer ability if there is auto-correlation in student achievement within a primary school. This means that if parents know the ability of their children by subject, as well as the achievement distributions of primary schools, they could potentially select a school on this basis.

secondary school qualification (32 percent), and if annual gross household income exceeds £33,000 (21 percent). Neither of these characteristics vary by subject and therefore the balancing tests cannot include student fixed effects. We find that neither parental characteristic is correlated with rank conditional on test scores.

The remaining rows of panel B of Table 4 perform balancing tests of primary school rank on observable student characteristics. Again like parental education and household income , these characteristics do not vary across subjects. Conditional on test scores and SSC effects, rank is a significant predictor of observable student characteristics, although the coefficients are small in size and have little economic meaning. For example, conditional on relative attainment a student at the top of class compared to being at the bottom of class is 0.8 percent more likely to be female (50 percent) and 0.8 percent more likely to be a minority student (16 percent). In addition to these effects being small there is no consistent pattern in terms of traditionally high or low attaining students, with non-FSME, minority and female students being more likely to be higher ranked than other students conditional on test scores. To assess the cumulative effect of these small imbalances the final row tests if predicted age-14 test scores based on student demographic characteristics, age-11 test scores and SSC effects are correlated with primary rank. We find that primary rank does have a small positive relationship with predicted test scores, albeit being about 1/70th of the magnitude of our main coefficient (7.946 and 0.113), implying a one standard deviation increase in rank is associated with a 0.001 standard deviation increase in predicted test scores. This is also reflected by the fact that our main estimates in Table 3 change insignificantly when we include student characteristics as controls.[25]

It appears that parents are not choosing schools on the basis of rank, but there are small imbalances of predetermined student characteristics. These imbalances could instead be caused by different types of students having different rank concerns during primary school as in Tincani (2015) for example. This is because we measure age-11 test scores, and therefore rank, at the end of primary school so rank concerns could impact student effort. We return to resulting measurement issues in section 5.3 and when discussing mechanism of competitiveness in section 7. Regardless of the precise sources of these imbalances, they do not significantly affect our results as they are estimated to be precisely small. As noted above, student demographics are absorbed by specifications that include student fixed effects, therefore these specifications are immune to imbalances related to factors that are constant across subjects.

5.2 Specification Checks

The main specification has a cubic polynomial in prior achievement, but one may be concerned that this functional form is not sufficiently flexible. Appendix Table A.3 shows the main specification with a linear control for baseline achievement in the first column and then each subsequent column progressively includes a higher order polynomial up until we have a sixth-order polyno-

[25] Using the methods proposed by Oster (2017) and conservative assumptions (namely it is possible to achieve an R^2 equal to one and unobservables are one-to-one proportional in their effect to observables, we cannot generate coefficient bounds that include zero for our main effect. This includes scenarios where unobservables explain over 125 times more of our remaining unexplained variation compared to student demographics.

mial in prior test scores in column 6. We find that once there is a cubic relationship, the introduction of additional polynomials makes no significant difference to the parameter estimate of interest.

As described in section 2.2 we can further relax the functional form assumptions by replacing the set of cubic controls for prior achievement with a non-parametric specification using a separate indicator for each age-11 test percentile. Here we would be effectively comparing students with the same test scores to students with different ranks in other classes, albeit still having mean shifts with the inclusion of SSC effects. These results are shown in row 2 of Table 5, where we can see that this does not have a large impact on the rank parameter, changing it from 7.894 (0.147) to 7.543 (0.146). The second column provides the equivalent estimates with additional student effects, these results are also similar to the benchmark specification, changing it from 4.562 (0.107) to 4.402 (0.107). This is comparing students with the same test score across subjects, but having a different rank due to the different test score distributions of peers across subjects.

Specification 2 also imposes that the test score parameters are constant across schools, subjects and cohorts. In Appendix Table A.4, we relax this by allowing for the impact of the baseline test scores to be different by school, subject or cohort, by interacting the polynomials with the different sets of fixed effects. The first two columns use linear and cubic controls for baseline test scores, we find that allowing the slope of prior test scores to vary by these groups does not significantly impact our estimates of the rank effect. In column 3 we use the non-parametric set of controls for the baseline and so allow the impact of any test score to be different by school, subject or cohort. This is effectively relaxing the assumption that a test score value of X in the external exam actually represents the same underlying abilities across different schools, where group-level differences are already captured by the SSC effects. Again, all specifications provide similar results.

5.3 Test-Scores as a Measure of Ability

Throughout we have assumed that we can use the age-11 achievement scores as a baseline measure of student ability. In principle, the Key Stage 2 scores should be a particularly good measures of ability, as they are finely graded, and their only purpose is to gauge the achievement of the student on an absolute metric. This means that students are not marked on a curve at the school or national level; hence, test scores are not a function of rank. However, there may be factors that cause these test scores to be a poor measure of ability. As we simultaneously use these test scores to determine rank and as a measure of prior achievement, these factors are of concern to our paper. We consider two cases where test scores do not reflect the underlying student ability due to systematic measurement errors (peer effects, teacher effects) and the general case of non-systematic measurement error (noise).

5.3.1 *Peer Effects*

This paper differs to the existing peer effect literature because we estimate negative effects of better peers, as well as effects of previous peers in a different peer setting. But how do contemporaneous peer effects interact with our estimation? To the extent that peer effects are sizable, they could in principle have meaningful impacts on our results because they would simultaneously impact on a student's rank and age-11 test scores. However, this is not an issue if primary peers have a constant permanent impact on student attainment (e.g., through the accumulation of more human capital, as conditioning on the end of primary test scores fully accounts for this, regardless of the nature of the contemporaneous peer effect).

Instead, consider the situation where an individual has poorly performing peers that affects achievement negatively during primary school and that peer effects fade out over time. Due to this type of transitory peer effect in primary school, students would achieve a lower score than otherwise and would have a higher rank. In secondary school, where in expectation they will have average peers, they will achieve test scores appropriate to their ability. This means that a student with bad primary peers would have a high primary rank and also high gains in test scores. In our value-added specification, this would generate a positive rank effect, even if rank had no impact on test scores.

To alleviate this concern, we include SSC effects in our main specifications, effectively removing any primary class-level long-run impacts on achievement growth and thus taking into account transitory linear-in-means peer effects. This is not exactly a pure linear-in-means peer effect, as the student's own outcome contributes to the fixed effect, but differences between including and excluding oneself for the estimation of linear-in-means effects are negligible.

To assess the remaining possibility that the SSC effects are not sufficient to account for transitory peer effects we implement a data generating process with linear and non-linear peer effects with magnitudes set to be 20 times larger than those found by Lavy et al. (2012) using the same English census data and school setting. Taking mean estimates from 1,000 simulations, we show that when controlling for SSC effects our estimates are indeed unbiased with transitory linear-in-mean peer effects. Moreover, even non-linear transitory peer effects only result in negligible downward bias. The simulations and further discussion can be found in Appendix A.1 and Appendix Table A.2.

5.3.2 *Teacher Effects*

Alternatively, instead of peers influencing the transformation of ability to test scores in a transitory way, it is possible that teachers teach in such a way as to generate false positive rank results. For this to occur, it would require teachers to have different transformation functions of student ability into test scores (e.g., some teachers create more spread in test scores than others). This would be potentially problematic because such transformations could affect the measured test scores but not the rank, therefore leaving rank to pick up information related to ability. Let us consider a simple case, where test scores in school j and subject s, y_{ijs}, are determined by mean teacher effects μ_{js}, and linear teacher transformations b_{js} of student ability a_{is}, $y_{ijs} = \mu_{js} + b_{js}a_{is} + e_{ijs}$. Here, variation in the teachers' production functions will generate differences in the observed test scores for a given ability. If teachers varied only in level effects μ_{js} then these differences would be captured by the SSC effects. However, if teachers also vary in their transformative approach, b_{js}, that would not be captured with these fixed effects, as different students within a class will be affected differently. If there is variation in b_{js}, then the same test score will not represent the same ability, a_{is}, in different schools, and critically rank will preserve some information on ability, which test scores will not. To test if this kind of spreading of test scores is driving the results, we run a set of placebo tests and simulations.

In the first, we use the assumption that this transformative teacher effect (b_{js}) is time invariant across cohorts, e.g. a set of students with a given initial ability distribution would have the same spread in test scores no matter which cohort they were enrolled in at that school, as long as they are taught by the same teacher. This is appropriate in England as teachers generally teach certain year groups rather than following a single cohort through primary school. If it is the teacher specific transformations that are generating the results, then remixing students across cohorts (but to the same teacher) will have no impact on the rank parameter, as rank will contain the same residual information on ability in each cohort. To test for this, we randomly reassign students to cohorts within their own school, such that students will appear in the same schools they attended. Row 3 of Table 5 shows the mean coefficients and standard errors from 1,000 randomization of students within schools across cohorts. Note that the rank parameters are significantly smaller and approximately one fifth of the previous size. As we use five cohorts of data, students will, on average, randomly have a fifth of their true peers in each cohort, and so we would expect this proportion of the rank effect to remain.

Next, if primary teachers varied in their transformations (b_{js}) of students ability into test scores, this would be reflected in the variance of test scores within each classroom that we observe at the end of primary school. For example, teachers with a high value of (b_{js}) would generate a high spread in test scores and therefore a higher variance of test scores within class. To gauge the importance of unobserved teacher transformation affecting the spread, we examine the following question: how much of the rank effects are driven by the variance in achievement within a primary school class? To test if this is driving the effects we include an interaction of the standard deviation of class test scores with each individual's rank. Note that the effect of the standard deviation itself, but not its interaction with rank, is captured by the SSC effect. Row 4 of Table 5 presents estimates of the rank effects that allow for interactions with classroom variance, evaluated at the mean of class standard deviations in test scores. As may be expected, the class fixed effects estimate falls to 5.718 (0.156), as some of this variation is generating the rank effects. The inclusion of the class standard deviation interactions, has a smaller impact on the within student estimates that fall to 4.337 (0.118) from 4.562 (0.107). The rank effects at the 25th and 75th percentiles of the classroom standard deviation of test scores are 6.29 and 5.34 for the SSC fixed effects specification and 4.27 and 4.39 when additionally accounting for individual fixed effects. The key take-away from these two exercises is that the rank effects are not primarily generated through differences in variances across classrooms as they remain relatively constant along this dimension. It follows that unobserved factors during primary school potentially affecting classroom variances in test scores, but leaving ranks unaffected, are not driving our main results. Next, we consider the case of measurement error affecting both, test scores and rank.

5.3.3 *Non-systematic Measurement Error*

When students take a test, the scores will not be a perfect representation of how well they perform academically on a day to day basis, resulting in a noisy measure of ability. This type of non-systematic measurement error is potentially problematic for our paper, as it could generate a mechanical relationship between student rank and gains in test scores. This is because they are both subject to the same measurement error but to different extents. Consider the simplest of situations where there is only one group and two explanatory variables, individuals *i's* ability X_i^*, and class ability ranking R_i^*. Assume X_i^* cannot be measured directly, so we use a test score X_i as a baseline which is a noisy measure of true ability and has measurement error $X_i = l_i(X_i{}^*, e_i)$. We also use this observed test score X_i in combination with the test scores of all others in that group$X_{-i}$to generate the rank of an individual, $R_i = k(X_i, X_{-i})$. We know this rank measure is going to be measured with error e_{i2}, $R_i = h_i(R_i^*, e_{i2})$. The problem is that this error, e_{i2}, is a function on their ability $X_i^*$and their own measurement errore_i, and also depend on the ability and measurement errors of all other individuals in their group (X_{-i}^* and e_{-i}.) $R_i = k(l_i(X_i^*, e_i), l_{-i}(X_{-i}^*, e_{-i})) = f(X_i^*, e_i, X_{-i}^*, e_{-i})$. Therefore, any particular realisation of e_i not only causes noise in measuring X_i^* but also in R_i^*. This means we have correlated and non-linear non-additive measurement error, where $COV(e_i, e_{i2}) \neq 0$. This specific type of non-classical measurement error is not a standard situation, and it is unclear how this would impact the estimated rank parameter.

The standard solution to non-classical measurement error is to obtain repeated measurements or instrumental variables (Schennach, 2016). In our case we do not have another measure of exactly the same test. An alternative would be to use the test scores from another subject as an instrument for X_i. However, as we go on to show in section 6.1, other-subject ranks do not necessarily meet the exclusion restriction as they have spillover effects across subjects.[26]

Therefore, we perform a data-driven bounding exercise to gauge the importance of any measurement error. This involves adding additional measurement error from a known distributiuon to the test score measure of each student, then re-calculating student ranks and then re-estimating the specification at ever increasing variances of measurement error. In doing so we are informed of the direction and the approximate magnitude of the measurement-error induced bias by adding noise to our existing measure. To do this, we again run a Monte Carlo simulation on the entire dataset, now adding artificial measurement error to each student score, which is drawn from a normal distribution with mean zero. The variance of this distribution increases from one percent of the standard deviation in test scores up to thirty percent, in terms of test scores this is an increase from 0.28 up to 11.2. For each measurement error distribution we simulate the data 1000 times and estimate the rank parameter.

Figure 4 shows the simulated estimates of the mean and the 2.5th and the 97.5th percentiles from the sampling distribution of beta for each measurement error level. We see that as measurement error increases, the downward bias in the rank estimate also increases, albeit non-linearly. Small additional measurement error has little impact on the results. The amount of downward bias from increasing the additional error from 1 percent to 20 percent of a standard deviation amounts to the same level of bias as increasing the error from 20 to 25 percent. The intuition for this result is the following: When measurement errors are relatively large, they will impact both test scores and rank measures. Individuals with a mistakenly high (low) test score also have a falsely high

[26] See section 6.1 for a discussion of subject spillovers and IV estimation.

(low) rank. Then, as we are estimating the the growth in test scores, these students would have lower (higher) observed growth which would downward bias the rank parameter. At low levels of measurement error, rankings would not change and it would be possible for the rank measure to pick up some information about ability lost in the test score measure. Appendix Figure A.1 repeats this process with alternative types of measurement errors. Panel A presents mean estimates from a heterogeneous measurement error process, where the impact of the normally distributed error is increasing as the test score is farther away from the national mean. This is reflective of the examinations potentially being less precise at extreme values. This slightly exacerbates the downward bias and the non-linearity of the bias. Next, panel B draws the additional measurement error from the uniform distribution, which results in smaller downward bias. Given that these national tests have been designed to measure ability in a subject and the small downward -but never upward- biases at small levels of different types of measurement error, we conclude that our main estimates are attenuated at most very little.

5.4 Further Robustness Checks: class size, re-mixing of students, and prior peers

To address potential remaining concerns, we present three further robustness checks. First, we have been describing our estimates as the impact of academic standing within ones primary school class. This is because we rank students within their school-subject-cohort and the median cohort size of primary schools is 27. Given that the legal maximum class size for 11-year old students is 30 we argue that cohorts are effectively equivalent to classes. However some schools have larger cohorts and so the ranking would represent their position within their school cohort rather than classroom. Row 5 of Table 5 presents estimates that restrict the sample to only those students who attended a primary school of less than 31 students per cohort. Here, the estimate falls slightly to 6.469 and the standard errors increase due to a smaller sample but the rank effect remains.

Second, row 6 of Table 5 shows an equivalent pair of estimates, which randomly allocates students to primary schools within a cohort. Here, students are very unlikely to be assigned to a class with their actual peers. As expected, this generates precise zero effects, reflecting the ranking of students in different schools has no impact. This implies that there are no mechanical relationships between rank and later achievement driving these results.

The third and final robustness check that we perform relates to the re-mixing of peers when students leave primary school and enter into secondary school. Unlike many other countries, primary schools in England tend to send students to many secondary schools, rather than many being a feeder school to one particular secondary school. This means that the average student has 87% new peers when starting secondary school. We argue that this largely addresses the reflection problem because we are using the peer characteristics from a prior setting and conditioning on test scores from that setting. To fully address reflection concerns we estimate the impact of rank only on the sub-sample of students who attend a secondary school with no prior peers from their primary school. Row 7 of Table 5 shows these results. Both the SSC and student effects estimates are larger in size than the benchmark, implying that any reflection issues would be downward biasing our estimates. However, students such as these may be non-typical and so we limit our interpretation of the parameter to its continued significance.

6 Heterogeneity of the Impact of Rank

Now that we have established the robustness of the estimates, we turn to the variability of the rank effects by considering impacts by subject, spillover effects across subjects, non-linearities and the impact by student demographics.

6.1 Impact by Subject

So far, we have assumed that there is a constant impact across subjects. Panel A of Table 6 shows the estimates of the rank parameter separately for English, Science and Maths (e.g., the impact of a student's rank in English at age 11 on their test scores at age 14). We can see that the impact is similar across subjects, ranging from 7.400 for English up to 8.820 for Maths. That the effects are comparable is important for the student fixed effects specifications which require they be equal.

The main specification also does not allow for any direct spillovers from being highly ranked in another subject. The lower half of Table 6 allows for these effects. Each column in panel B represents a single estimation, where we condition on the rank and prior test scores in each subject separately and present the corresponding rank estimates. The impact of rank on the same subject outcomes are on the main diagonal, and the impact of other subjects are off the diagonal. For each subject, once we allow for the impact of rank from another subject, the main coefficient reduces in size. An explanation is that rank is positively correlated across subjects and that there are positive spillovers across subjects. We also see that the nature of the spillover effects depends on the subject pairing. Science and Maths ranks have small, if any, impacts on English test scores (0.597 and 0.788 respectively), however the impact of Maths rank on Science and Science rank on Maths is considerably larger (3.612 and 3.233 respectively).[27]

6.2 Non-linearities and Heterogeneity

6.2.1 Age-14 Outcomes

The specifications thus far assumed that the effect of rank is linear. Here we allow the effect of rank to change throughout the rank distribution by replacing the rank parameter with a series of 20 indicator variables in the vingtiles in rank, plus top and bottom of class dummies (Specification 4). The equivalent estimates from specification 2 and 3, without and with student fixed effects, are presented in the first panel of Figure 5. For the impact on age-14 test scores, the effect of rank appears to be linear throughout the rank distribution, with small flicks in the tails. This indicates that the effect of rank exists throughout. Students ranked just above the median perform better three years later than those at the median. The within student estimates are smaller in magnitude throughout and have less of a gain for being top of the class.

[27] The only subject-pairing where we do not estimate a statistically significant spillover effect is for effects of ranks in English on outcomes in maths. This suggests that the exclusion restriction is valid for one subject-pairing. If we instrument maths rank and test scores with English rank and test scores, we obtain an IV estimate of the instrumented maths-rank on maths outcomes of 8.07. This is not different to the OLS estimate at conventional levels of statistical significance, which has the value of 7.753.

Continuing to use this non-linear specification with student fixed effects, we now turn to how the effects of rank vary by gender and free school meal eligibility. We estimate these effects by interacting the rank variables, which vary within students, with the dichotomous characteristic of interest.[28] The second panel in Figure 5 shows how rank relates to the gains in later test scores by gender. Males are more affected by rank throughout 95 percent of the rank distribution. Males gain four times more from being at the top of the class, but also lose out marginally more from being in the bottom half. This is within student variation in later test scores, and therefore the coefficient could be interpreted as a specialising term, implying that prior rank is associated with a stronger specialising effect on males than females. Relating these finding to the idea of student self concept, this could be reflecting that males place more importance on their relative ranking in determining their self concept than females. Continuing with this interpretation, the finding that males gain significantly more from being higher ranked than females could be reflecting that males are perceiving themselves to be higher ranked than they actually are. In this reduced form specification, we cannot separate the impact of being highly ranked from the perception of rank. We are assuming that individuals accurately place themselves within a class, and that this ability does not differ by student characteristics.

The final panel in Figure 5 shows the impact by FSME status. This has a very different pattern to the previous panel. Both types of students have a positive relationship with rank and later test scores, however FSME students are less negatively affected by a low rank and more positively affected by a high rank compared to Non-FSME students. The relationship is approximately flat for FSME students in the bottom half of the rank distribution, but they have a very steep gradient in the top half of the distribution with those ranked top in class who gain almost twice as much as Non-FSME students. One interpretation of this is that FSME students already have low confidence in their abilities and thus do not suffer from being lowly ranked. Or, they may already perceive themselves to be lowly ranked. Moreover, the shallower gradient for Non-FSME students could lead to an interpretation that they are less affected by class rank, as these students may have their academic confidence being more affected by factors outside of school. We will return to the interpretation of these effects at the end of the mechanisms section.

6.2.2 A-Level Subject Choices

Figure 6 presents the nonlinear impacts of primary school rank on A-Level specialisation. The three panels in this figure sequentially present the impact of rank in each subject on the likelihood of choosing an English, maths or science A-Level. In each panel we see a positive relationship for the rank of the subject in question on taking that A-Level at the end of secondary school. Unlike the impact on test scores these impacts are highly non-linear with the majority of the impacts occurring in the top decile. There are differences across subjects. For maths, students are just as likely to choose the subject if they were at the bottom of their primary class or at the the middle, it is only students at the very top, for whom rank has an effect. In contrast, for both English and science students are continuously more likely to choose those subjects as primary rank increases.

[28] The figure for the remaining characteristic of minority status is available up request. However, it shows little heterogeneity along this characteristic.

Figure 6 also plots spillover effects across subject. Being lowly ranked in English reduces the chances of students choosing any A-level subject. With regards to choosing English we also see that students are in fact less likely to take English A-level exams if they are highly ranked in either science of maths. This type of negative spillover across subjects does not occur with test scores. An explanation for this is that students are limited to taking only three A-Levels in total, therefore having a high rank in one subject might crowd out the likelihood of taking another subject.

7 Mechanisms

A number of different mechanisms relating to the rank of the student may produce these results. These include competitiveness, environmental favors to certain ranks, parental investment, students learning about their ability, and development of non-cognitive skills. In the following section, we discuss how each of these channels might explain the findings presented in the previous sections.

7.1 Mechanism 1: Competitiveness

Recent research indicates that students may have rank concerns during primary school and that they adjust their efforts accordingly (see Tincani (2015), Hopkins and Kornienko (2004) and Kuziemko et al. (2014) for studies of effects of rank concerns). In our setting, if students work harder during primary school because of these concerns, it will be reflected in higher end-of-primary school achievement scores, which we control for. However, if students want to maximise rank at minimal cost of effort, this potentially could produce the slight imbalances found in section 5.1 but not negative effects of low ranks. One could consider this as a form of measurement error in ability, but one that is caused by the rank of the student.

To illustrate this point, consider students attending primary schools where they face little competition for being at the top of the class. These students would have to spend less effort and thereby obtain a lower test score while still remaining at the top of the class. Then, when facing a more competitive secondary school environment, these previously "lazy" students may exert more effort and will appear to have high growth in test scores. This would generate a positive rank effect as those at the top of class have high gains in test scores. However, if rank concerns during primary school are driving the results, we would see these effects only near the top of the rank distribution. All students in the remainder of the distribution would be applying effort during primary school to gain a higher rank. Consequently, if this competitiveness were driving the results we would not expect to see the effects throughout the rank distribution found in Figure 5.

Alternatively, if there were unobserved heterogeneity in competitiveness, this could cause primary school subject rank to be positively correlated with later test scores. However, in the specification that includes student fixed effects, any general competitiveness of an individual will be accounted for. This competitiveness would need to vary by subject. As previously mentioned, factors that vary by student across subjects conditional on prior test scores could confound, or in this case, explain the results.

7.2 Mechanism 2: Parental investment

Parents may react to the academic rank of their child by altering their investment decisions. Parents can assist the child at home with homework or with other extra-curricular activities, or choose a school specialising in a certain subject. If the parents know that their child is ranked highly in one subject, they may have a tendency to encourage the child to do more activities and be more specialised in that subject. Note that, as we are controlling for student effects, this will need to be subject-specific encouragement, rather than general encouragement pertaining to school work. This mechanism assumes that parents react to achieved primary school rank rather than to prior preferences. For this to be the case, it would require parents to want to specialise their child at the age of 11, rather than to improve their child's weakest subject. If parental investment is focused on the weaker subject (Kinsler et al., 2014), this will reverse the rank effect for these students. To test for this, in this section we use information from the LSYPE survey about parental investments during secondary school in terms of time and money. We also test if, instead of reacting to rank through increased investments, parents respond by sending their child to a school that specialises in that subject area.

Each row of Table 7 shows the estimated impact of rank on a different type of parental investment. All estimates in the second column are from our main specification, the first column omits the flexible student achievement covariates. Panel A relates to parental time investments during secondary school. We see that rank in primary school is positively correlated with the number of parents who attended the most recent parents' evening, but, once we condition on primary school achievement, there is no longer a significant relationship. A similar pattern occurs with parents having special meetings with the teacher about the child and the frequency of talking with the child's teacher: both of these are significantly negatively correlated with the students rank in primary school, but after conditioning on test scores there appears to be no relation to rank. The fourth investment measure about self-reported feelings about parental involvement with the child's school life is not correlated with rank in either specification. We construct an index of these parental time investments using principal components analysis, which we standardise, and again find that once conditioning on prior attainment parental investment is uncorrelated with prior rank.

Panel B of Table 7 presents how parental financial investments during secondary school are related to primary rank. The outcome of the first row of panel B is an indicator variable for parents investing in any form of out of school tuition for the student (23.4 percent). When not controlling for prior attainment, this is positively correlated with student rank, which could be reflective of coming from wealthier families. However, once prior attainment is accounted for, in the second column, this significant correlation ceases. The second row of panel B examines weather student rank induces parents to pay for out of school tuition in the associated subject, allowing the investment to vary within the student across subjects. We see that students ranked low in a subject during primary school indeed have more subject-specific tuition in secondary school. But again, conditional on attainment, this correlation is removed.

In the final panel of Table 7, we test for sorting to secondary schools by subject according to primary rank by calculating each secondary school's subject-specific value added measure in terms of age 11 to 14 growth in test score percentiles. The first row uses a raw value added measure and the second uses one that conditions on student demographics, and each have been standardised to mean zero and standard deviation one. Both measures provide similar results: students ranking high in a subject tend to enroll in school with a high value added in that subject, but, like the previous panels, when we account for age 11 achievement, there is no significant relationship [29]

Instead of testing for sorting directly, the final two rows of Table 5 explore how the main rank coefficient changes when we condition on different secondary school characteristics in order to determine the overall importance of the secondary school attended. Because these school characteristics can be interpreted as possible outcomes of primary rank in their own right, these parameters should be interpreted with caution. We proceed by first removing all students who attend secondary schools labeled specialist in English, math, or science from the sample and re-estimating the effects. This consists of eight percent of all secondary schools at the age of entry, or 575,000 observations in our sample. The removal of these schools has a negligible impact on the rank effects (Row 8, Table 5). Second, parents may not be reacting to the labels, but choosing schools with high actual gains in certain subjects. If the rank effect is due to this type of parental sorting, then additional conditioning on secondary SSC effects would reduce the size of the rank parameter. In the final row of Table 5, we see that these additional controls only have a small impact on the rank effect, as would be expected given the results from panel C of Table 7.

This section has shown that while parental investments during secondary school are correlated with a student's rank during primary school only when not conditioning on achievement. We take this as direct evidence against rank-based parental investments causing the positive long-run effects of primary rank during secondary school. More generally, parents are unlikely to be accurately informed of their child's specific class rank in the English context. Teacher feedback to parents will convey some information, such as the student being the best or worst in class, but it is doubtful that they would be able to discern a difference from being near the middle of the cohort rankings. Our results, however, show significantly different effects from the median for all vingtiles with school-subject-cohort effects. Taken together, parental investments are an unlikely mechanism.

7.3 Mechanism 3: The environment favors certain ranks

Another possible explanation for the positive impact of rank is that there are rank-based investments. For example, one can imagine primary school teachers focusing their attention on low ranked students or schools providing extra resources to these students. This does not impact our results if these investments have a permanent impact on student achievement, as they are realised in primary school test scores, on which we condition. This mechanism requires that any improvement in test scores is temporary, either being realised only in primary school or secondary school. Like the competitive mechanisum, one can consider this as a form of rank generated measurement error.

[29] Appendix Table A.5 additionally tests for school level value added, with and without student controls. None of the eight specifications find a significant relationship with rank conditional on prior attainment.

For example, if teachers targeted low ranked students, these students would achieve higher age-11 test scores than otherwise, but retain their low ranking. These students would also have low achievement growth between ages 11 and 14 and therefore this would generate a positive rank correlation at the bottom of the distribution.[30] For a positive rank correlation to occur at the top of the distribution, we would require the opposite situation, where increased investment in top students during primary school was *not* reflected in primary school test scores but only in later achievement. This would be captured by our rank effect, β^1_{Rank}, and would be considered a potential mechanism.

Both of these channels require the targeting of teacher effort, however in our setting teachers are in a performance related pay system, where they are rewarded for generating better than average improvements in performance (Atkinson et al., 2009). Principals and teachers agree on a target achievement levels for each individual student and teachers are then rewarded for exceeding them. As a result, there are no financial incentives for teachers to target low performing or low ranked students, similar to a pay-for-percentile system as discussed in Barlevy and Neal (2012). Moreover, it has been found that teachers fail to target specific groups of students, even when faced with explicit incentive schemes (Chakrabarti, 2014; Reback et al., 2014). Ultimately, as awards are based on student growth, teachers would have incentives to generate contemporaneous, rather than unrealised gains in test scores. Given these inconsistencies and the required assumptions that any beneficial effects are transitory, we have doubts that this is the dominant reason for the effect, but we cannot exclude this mechanism. Of course, if there was no measurement error in the age-11 test scores, and no fade out of teacher investments, this mechanism could be ruled out.

7.4 Mechanism 4: Student Confidence

A simple mechanism for the rank effect is that being highly ranked among peers makes an individual more confident in school generally or in a specific subject. We will first provide evidence that rank, conditional on test scores, increases confidence and follows similar patterns of heterogeneity found with the main effects. To explain the positive effects of primary rank on confidence and attainment, we then propose two models of student effort allocation based on learning and non-cognitive skills, which we test against each other.

7.4.1 Impact on Confidence

We link our administrative data to the LSYPE data which contains questions regarding student confidence in each of the subjects of interest at the age of 14. This allows us to test directly if rank position within primary school has a lasting effect on subject confidence, conditional on attainment. The specification is equivalent to specification 2 with the dependent variable now being subject-specific confidence. Since this survey was run for only one cohort, the SSC effects are replaced by school-by-subject effects. Also, the LSYPE data was not sampled to be representative

[30] Note if primary teachers taught to the median student in such a transitory way, those at both extremes would lose out. So instead of a linear effect, we would find a U-shaped curve with both students at the bottom and the top of the distribution gaining relatively more during secondary school.

of the student population, so we first replicate our main result using students for the LSYPE sample only in the first column of Table 8. In this sample of students we obtain a rank-effect of 8.977, which is slightly, but not statistically significantly, higher than our baseline estimates.

The second column of Table 8 shows that students with a higher primary school rank position are significantly more likely to say that they are good in that subject at the 5-percent level of statistical significance. Moving from the bottom of the class to the top increases confidence by 0.196 points on a five-point scale, which corresponds to 20 percent of a standard deviation of this confidence measure. This suggests that students develop a lasting sense of strengths and weaknesses depending on their local rank position, conditional on relative test scores. This is despite there being very few LSYPE respondents per primary school (4.5 students conditional on at least one student being in the survey), which severely limits the degrees of freedom in each primary school subject group. Columns 3 and 4 present results separately for boys and girls. In line with our results from the administrative data, we find that boys are more strongly affected by primary rank.

Overall, given the effects of rank on the direct measures of student confidence and the heterogeneity of effects found in the main results, it seems likely that confidence matters. This is in line with the psychological literature which finds that academic confidence is thought to be especially malleable at the primary school age (Tidemann, 2000; Rubie-Davies, 2012; Leflot et al., 2010).

7.4.2 Learning or Non-Cognitive Skills?

We conclude the analysis by testing how this increased confidence improves test scores by comparing two channels, learning about ability and non-cognitive skills. The first is that students use their ordinal class ranking in addition to their absolute achievement to learn about their own subject-specific abilities, similar to a model proposed by Ertac (2006). Students then use this information when making effort investment decisions in secondary school across subjects. Are they relatively more productive in math or English? Critically, this mechanism does not change an individual's education production function, only their perception of it. We will argue below that this feature allows us to test the learning model.

The second is that a student's ordinal ranking during primary school has an impact on their academic confidence and hence non-cognitive skills. In the educational literature, this effect is known as the Big Fish Little Pond Effect, and it has been found to occur in many different countries and institutional settings (see Marsh et al. (2008) for a review).[31] This confidence can differ over tasks, so a student can consider herself good in English, but bad in maths (Marsh et al., 1988; Yeung and Lee, 1999). Confidence generates non-cognitive skills in a subject such as grit, resilience, and perseverance Valentine et al. (2004). The importance of such non-cognitive skills for both academic attainment and non-academic attainment is now well established (Borghans et al., 2008; Lindqvist and Vestman, 2011; Heckman and Rubinstein, 2001). We propose to model these increased non-

[31] The psychological-education literature uses the term self-concept, which is formed through our interactions with the environment and peers O'Mara et al. (2006). Individuals can have positive or negative self-concept about different aspects of themselves and students with a high self-concept would also develop positive non-cognitive skills..

cognitive skills as a decrease in the costs of effort for that task.[32]

In order to test the learning and non-cognitive skills models, we propose a simple conceptual framework which can accommodate both of them. We again have two periods, an experience phase and an action phase, representing primary school and secondary school respectively. In the first period, students carry out tasks (subjects) in a small group and compare their performance to others, which determines their confidence when entering the second period.[33] In the second period, we model students as total grade maximising agents for a given total cost of effort and subject ability level . The total grade Y, of student i is the sum of their grades across subjects s. For simplicity of notation, assume that there are only two subjects, $s = \{e, m\}$. The production of grades in each subject is a function of subject specific ability A_{is} and effort E_{is}^{κ}, where we assume decreasing returns to effort, $0 < \kappa < 1$. The productivity of these factors is additionally affected by subject specific school factors μ_s. Accordingly, the total grade of individual i is a separable production function and can be represented as:

$$Y_i = f(A_{ie}, E_{ie}) + f(A_{im}, E_{im}) = \mu_{ie} A_{ie} E_{ie}^{\kappa} + \mu_{im} A_{im} E_{im}^{\kappa}$$

The students maximises total grades subject to a cost function. This cost function is determined by their cost of effort in each subject C_{is} and a general cost of academic effort C_{ig}. This general cost reflects a student's attitude towards education in general and is linear in the sum of effort applied across all subjects, $E_{im} + E_{ie}$. The total cost of effort T that a student can apply is fixed; however, the inclusion of a general cost of academic effort term means that the total effort applied by a student is very flexible.

$$T_i \geq C_{im} E_{im} + C_{ie} E_{ie} + C_{ig}(E_{im} + E_{ie})$$

These factors determine how students choose how to allocate effort across the two subjects. In both models their optimal choice will be affected by their previous experience, either through perceived ability or the cost of effort so that a grade-maximising student chooses optimal effort E^*:

$$E_{is}^{*} = \left[\frac{\lambda(C_{is} + C_{ig})}{(\kappa \mu_s A_{is})} \right]^{(1/(\kappa-1))} \tag{8}$$

where $\lambda > 0$ is the marginal grade per effort. As $0 < \kappa < 1$, any increase in subject ability A_{is} will increase the optimal effort allocated to that subject, and higher costs would decrease effort. This case is represented in Figure 7 panel A, where the relative costs of effort C_s, C_g determine the gradient of the iso-cost line and abilities determine the shape of the iso-quant Q. A student chooses optimally to invest E_e in English effort and E_m in maths effort. We can now use this simple framework to distinguish between the two models.

In the learning-about-ability model, students know their costs of effort but do not know their abilities. Therefore, in the experience phase, students use their test scores and relative ranking to

[32] The alternative is to have non-cognitive skills impact the ability (returns to effort), by subject, which leads to the same predictions and testable hypothesis.

[33] These comparisons can be in terms of cardinal and ordinal performance.

form beliefs about their abilities and the shape of their isoquant. Students experiencing a high rank in English in the first period believe that they have a high ability $A_e^{\sim}$ in that subject and therefore optimise according to the new isoquant $Q^{\sim}$ and devote more effort to English in the following period, $E_e^{\sim}$ believing that it generates high marginal returns to effort (Figure 7 panel B).

In the non-cognitive skills model, students know about their costs of efforts and abilities in each subject. Here, in the experience phase, students develop confidence in the subjects in which they are ranked highly and develop positive non-cognitive skills in that subject, which we model as reducing the cost of effort $C_e > C_e^{'}$. This shifts the students' iso-cost line out along the English effort axis to point $T/(C_e^{'} + C_g)$. Consequently, they are able to reach higher isoquant $Q^{'}$, and optimally invest more effort in English, $E_e^{'} > E_e$(Figure 7, panel C). Rank can also impact students' general cost of effort C_g, which we assume to be a decreasing function of ranks in all subjects. If there are any general gains in confidence due to having a high rank in English this reduces the cost of any academic effort $C_g > C_g^{'}$ and causes a parallel shift out of the iso-cost line, with intercepts $T/(C_m + C_g^{'})$ and $T/(C_e^{'} + C_g^{'})$. This results in the students providing more effort to all subjects $E_e^{''} > E_e^{'} > E_e$. Note that with this channel, it will be impossible for students to misallocate effort across subjects, as they are perfectly informed about their costs and abilities.

Given these two ways of interpreting confidence, we now consider the case where students have a high rank locally in their class but have a low global rank nationally. Under the learning hypothesis, students with large differences between local and national ranks (in absolute terms) would have more distorted information about their true abilities, assuming national test scores are a good measure of ability. These students would then be more likely to misallocate effort across subjects, thereby achieving lower average grades compared to students whose local ranks happen to closely align with national ranks. Turning back to panel B of Figure 7, this is represented by the student believing that she is on iso-quant $Q^{\sim}$with resulting effort choices $E_e^{\sim}$and $E_m^{\sim}$, which are not optimal because her perceptions are incorrect. Such a student would therefore over-invest in English. The local ranking provides a distortion that shifts effort allocation from the optimal allocation ($E_e$and E_m) to an non-optimal allocation $E_e^{\sim}$and $E_m^{\sim}$. This means that due to the misinformation about relative abilities students ultimately end up on a lower iso-quant than they could achieve $Q^{'} < Q$.

This gives rise to a testable hypothesis to distinguish between the confidence and learning channel. In situation where local ranks are very different from national ranks, and thus less informative about actual abilities, the misinformation can result in students obtaining lower total grades. Conversely, if the rank effects are caused by actual changes in the costs associated with the education production function, even if local rank is different from national rank, this would not lead to a misallocation of effort in terms of maximising grades, and so total grades would not decrease. To summarise, if a student experiences a local rank in English higher than their national rank, under both models they allocate more effort into English in the second period, but under the non-cognitive skills case this would not lead to a misallocation of effort and lower average grades.

We do not have direct data on perceptions versus reality of costs. However, we can test for misallocation of effort indirectly by examining how average grades achieved are correlated with the *degree of misinformation*. More precisely, we compute a measure of the degree of misinformation for students in each subject using their local rank $R_{ijsc} \in \{0,1\}$ and national percentile rank $Y^0_{ijsc} \in \{0,100\}$ at age-11. Both are uniformly distributed in the aggregate and, therefore, we simply define the degree of misinformation Mis_{ijsc} as the absolute difference between the two after re-scaling percentile rank:

$$Mis_{ijsc} = |R_{ijsc} - \frac{Y^0_{ijsc}}{100}| \quad where \quad Mis \in \{0,1\} \tag{9}$$

This measure takes the value zero for students where their local rank happens to correspond exactly to the national rank. Here, there are no differences between the predictions of the learning and confidence models. A large value for Mis_{ijsc}, on the other hand, indicates large differences between local and national rank. Here, total grades obtained should be lower if students use this information to form beliefs. Averaging this metric across subjects within student provides a mean indicator of the degree of misinformation for each student. To test directly whether or not a student with a large amount of misinformation does significantly worse, we use a specification similar to specification 2 but with the by-subject variation removed, as we are examining the effect on average test scores. We estimate the following specification:

$$\bar{Y}^1_{ijc} = \beta^1_{Rank}\bar{R}_{ijc} + f(Y^0_{ijc}) + \beta^1_{Mis}\bar{M}is_{ijc} + \boldsymbol{x'_i\beta} + \boldsymbol{j'_{jc}\varphi} + \eta_{ijc} \tag{10}$$

where

$$\eta_{ijc} = \tau_i + \upsilon_{ijc}$$

Here, $\bar{Y}^1_{ijc}$ denotes the average test scores across subjects in period 1, $\bar{R}_{ijc}$is average rank in primary school,$\bar{M}is$ the additional misinformation variable, $\boldsymbol{x_i}$a vector of individual characteristics and j_{jc} primary school-cohort fixed effects. To clearly re-state our hypothesis: if the amount of misinformation causes students to misallocate effort across subjects we expect $\beta_{Mis} < 0$. Alternatively, the null hypothesis that local rank causes changes to the actual production function means $\beta_{Mis} = 0$.

$$H_1: \ Learning \ \ \beta_{Mis} < 0$$
$$H_0: \ Null \ \ \beta_M is = 0$$

We obtain the following estimates using our full sample of 2,271,999 students. For benchmarking purposes, we first estimate a version of specification 10 without the additional misinformation variable in Table 9. The effect of average rank on average test score is estimated at 10.765, and is highly statistically significant. Column 2 adds the coefficient for the effect of misinformation, which is estimated to be small, positive and statistically insignificant at conventional levels while the rank parameter remains almost unchanged. We repeat this in columns 3 and 4 additionally controlling for student characteristics and find no meaningful differences. Given these results, we fail to reject the null hypothesis that the degree of misinformation does not cause students to misallocate effort. Therefore, we conclude that the learning mechanism alone is unlikely to generate our results.

Given this finding, we turn back to re-interpret the main results. The non-cognitive skills model is consistent with the empirical results found in section 4. Moreover, under this model the smaller estimates from the pupil fixed effects specification 3 have general confidence effects absorbed ($I^{''}$ to $I^{'}$ in Figure 7). Therefore, these estimates pick up only the effect of within student reallocation of effort across subjects.

8 Conclusion

This paper established that an individual's ordinal position within a group impacts later objective outcomes, controlling for cardinal achievement. In doing so, we have introduced a new factor in the education production function, showing that rank position within primary school has significant effects on secondary school achievement and the likelihood of completing STEM subjects. There is significant heterogeneity in the effect of rank, with males being influenced considerably more. Moreover, a higher rank seems also linked to important non-cognitive skills, such as confidence.

What are the policy implications of these findings? With specific regard to education, these findings leads to a natural question for a parent deciding where to send their child (in partial equilibrium). Should my child attend a "prestigious school" or a "worse school" where she will have a higher rank? Rank is just one of the many factors in the education production function. Therefore, choosing solely on the basis of rank is unlikely to be a correct decision.

To gauge the relative importance of rank for choosing a primary school, we follow an approach similar to Chetty et al. (2011) to estimate the overall effect of class quality (which includes class size, teacher quality, peer quality, etc.) on long-run outcomes by using the size of our SSC effects from specification 2 as a benchmark. Attending a primary school with a one standard deviation higher quality, net of the effect of rank, is associated with 0.269 standard deviation higher growth rate at the age-14 exam. This means that increasing primary rank by one standard deviation has effects equivalent to increasing general primary school quality by about 0.3 standard deviations.

We now make some comparisons with effect sizes found by the literature. We first compare the rank effect to the impact of teachers for one year. Being taught by a teacher who is one standard deviation better than average for one year improves student test scores by 0.1 to 0.2 standard deviations (Rivkin et al., 2005; Aaronson et al., 2007). Comparatively, we find that a student with one standard deviation higher rank throughout six years of primary school increases the age-14 achievement by 0.08 standard deviations. Note however, that these teacher effects on test scores are contemporaneous and fade out over time (Chetty et al., 2014), whereas the rank effect is long-lasting. There are few papers that look directly at the long run impact of elementary school peers on later outcomes, with the exception of Carrell et al. (2016). They estimate the causal effect of a single disruptive peer in a class of 25 throughout elementary school to reduce test scores in grades 9 and 10 by 0.02 standard deviations. In comparison, we estimate the long-run effect of a one standard deviation increase in rank throughout primary schools to be equivalent to adding four disruptive peers.

Next, how do our lasting peer effects compare in size to estimates of contemporaneous peer effects? Lavy et al. (2012) estimate peer effects on test score growth using the same administrative data and conditional on student effects. They find that each additional new peer at secondary school of the bottom 5 percent of the national age-11 achievement distribution decreases age-14 test scores by 0.008 standard deviations of the within-pupil achievement distribution. Equivalently, a one standard deviation increase in contemporaneous 'bad' peers decreases age-14 value added by 0.033 within student standard deviations. In comparison, a one standard deviation higher rank during primary school increases age-14 value added by 0.055 within student standard deviations.[34]

Our results also show that primary school rankings affect A-Level subject choices and thus have long run implications, as A-Level choices are linked directly to university admissions. Hastings et al. (2013) uses data from Chile to estimate thate choosing STEM subjects at univerity over humanities increases later earnings by 12 percent. Of course, if males and females, on average, had the same primary rankings across all subjects, these findings would not contribute to the gender STEM-gap. However, females outrank males in English subjects during primary school, so females have on average higher ranks in non-STEM subjects.[35] Taking absolute differences in combined math and science versus English ranks, males have a 0.166 higher rank in STEM subjects than females. Based on our estimates, this difference would mean that males are about 0.66 percentage points more likely choose a STEM A-Levels, conditional on achievement. Given the low share of students taking STEM subjects, if we were to equalize the primary rankings in subjects across genders, this would reduce the total STEM gender gap by about 7 percent. One direct way to achieve this would be to separate primary schools by gender, ensuring there will be same amount of females and males being on top of their class in STEM and English.

Building upon the results of section 7.4.1, where your rank impacts confidence and non-cognitive skills, there would be general implications for productivity and informational transparency. To improve productivity, it would be optimal for managers or teachers to highlight an individual's local rank position if that individual has a high local rank. If an individual is in a high-performing peer group and therefore may have a low local rank but high global rank, a manager should make the global rank more salient. For individuals who have low global and local ranks, managers should focus on absolute attainment, or focus on other tasks where the individual has higher ranks.

Besides policy implications, our findings also help to reconcile a number of topics in education.These persistent rank effects could partly speak towards why some achievement gaps increase over the education cycle. Widening education gaps have been documented by race (Fryer and Levitt, 2006; Hanushek and Rivkin, 2006, 2009). With rank effects, small differences in early overall attainment could negatively affect general academic confidence, which would lead to decreased investment in education and exacerbate any initial differences. A similar argument could be made for the persistence of relative age-effects, which show that older children continue better compared to their younger counterparts (Black et al., 2011).

Finally, research on selective schools and school integration has shown mixed results from students attending selective or predominantly non-minority schools (Angrist and Lang, 2004; Clark, 2010; Cullen et al., 2006; Kling et al., 2007; Dobbie and Fryer, 2014). Many of these papers use a regression discontinuity design to compare the outcomes of the students that just passed the

[34] The Lavy et al. (2012) calculations are based on within student estimates from Table 3 column 4 row 4, and the within-student standard deviation in treatment status found in Table 2.

[35] The average primary male's (female's) rank in English is 0.440 (0.535), in math 0.515 (0.468), and in science 0.477 (0.474).

entrance exam to those that just failed. The common puzzle is that many papers find no benefit from attending these selective schools.[36] However, our findings would speak to why the potential benefits of prestigious schools may be attenuated through the development of a fall in confidence among these marginal/bussed students, who are also necessarily the low-ranked students. This is consistent with Cullen et al. (2006, p. 1194), who find that those whose peers improve the most gain the least: "Lottery winners have substantially lower class ranks throughout high school as a result of attending schools with higher achieving peers, and are more likely to drop out".

References

Aaronson, D., Barrow, L., and Sander, W. (2007). Teachers and Student Achievement in the Chicago Public High Schools. *Journal of Labor Economics*, 25:95–135.

Abdulkadiroglu, A., Angrist, J., and Pathak, P. (2014). The elite illusion: Achievement effects at boston and new york exam schools. *Econometrica*, 82(1):137–196.

Anderson, C., Srivastava, S., Beer, J. S., Spataro, S. E., and Chatman, J. A. (2006). Knowing your place: Self-perceptions of status in face-to-face groups. *Journal of Personality and Social Psychology*, 91(6):1094–1110.

Angrist, J. D. and Lang, K. (2004). Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program. *American Economic Review*, 94(5):1613–1634.

Arcidiacono, P., Aucejo, E., and Spenner, K. (2012). What happens after enrollment? an analysis of the time path of racial differences in gpa and major choice. *IZA Journal of Labor Economics*, 1(1):5.

Atkinson, A., Burgess, S., Croxson, B., Gregg, P., Propper, C., Slater, H., and Wilson, D. (2009). Evaluating the impact of performance-related pay for teachers in england. *Labour Economics*, 16:251–261.

Azmat, G., Bagues, M., Cabrales, A., and Iriberri, N. (2015). What You Know Can't Hurt You, For Long: A Field Experiment on Relative Performance Feedback. Mimeo, Queen Mary University.

Azmat, G. and Iriberri, N. (2010). The importance of relative performance feedback information: Evidence from a natural experiment using high school students. *Journal of Public Economics*, 94(7-8):435–452.

Bandiera, O., Larcinese, V., and Rasul, I. (2015). Blissful Ignorance? A Natural Experiment on the Effect of Feedback on Students' Performance. *Labour Economics*, 34:13–25.

Barlevy, G. and Neal, D. (2012). Pay for percentile. *American Economic Review*, 102(5):1805–31.

Bertrand, M., Goldin, C., and Katz, L. F. (2010). Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics*, 2(3):228–55.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2011). Too Young to Leave the Nest? The Effects of School Starting Age. *The Review of Economics and Statistics*, 93(2):455–467.

[36] Similar effects are found in the Higher Education literature with respect to affirmative action policies (Arcidiacono et al., 2012; Robles and Krishna, 2012).

Borghans, L., Duckworth, A. L., Heckman, J., and ter Weel, B. (2008). The economics and psychology of personality traits. *Journal of Human Resources*, 43(4):972–1059.

Brown, G., Gardiner, J., Oswald, A., and Qian, J. (2008). Does wage rank affect employees well-being? *Industrial Relations: A Journal of Economy and Society*, 47(3):355–389.

Bursztyn, L. and Jensen, R. (2015). How Does Peer Pressure Affect Educational Investments? *The Quarterly Journal of Economics*, 130(3):1329–1367.

Card, D., Mas, A., Moretti, E., and Saez, E. (2012). Inequality at work: The effect of peer salaries on job satisfaction. *American Economic Review*, 102(6):2981–3003.

Carrell, S., Fullerton, R. L., and West, J. (2009). Does your cohort matter? measuring peer effects in college achievement. *Journal of Labor Economics*, 27(3):439–464.

Carrell, S. E., Hoekstra, M., and Kuka, E. (2016). The long-run effects of disruptive peers. Technical report, NBER Working Paper No. 22042.

Chakrabarti, R. (2014). Incentives and responses under no child left behind: Credible threats and the role of competition. *Journal of Public Economics*, 110:124–146.

Chetty, R., Friedman, J. N., Hilber, N., Saez, E., Schanzenbach, D. W., and Yagan, D. (2011). How does your kindergarten classroom affect your earnings? evidence from project star. *Quarterly Journal of Economics*, 126(4):1593–1660.

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014). Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American Economic Review*, 104(9):2593–2632.

Cicala, S., Fryer, R., and Spenkuch, J. L. (2017). Self-selection and comparative advantage in social interactions. *Journal of the European Economic Association*, forthcoming.

Clark, D. (2010). Selective schools and academic achievement. *The B.E. Journal of Economic Analysis & Policy*, 10(1):1–40.

Cullen, J. B., Jacob, B. A., and Levitt, S. (2006). The Effect of School Choice on Participants: Evidence from Randomized Lotteries. *Econometrica*, 74(5):1191–1230.

Dobbie, W. and Fryer, R. (2014). The impact of attending a school with high-achieving peers: Evidence from the new york city exam schools. *American Economic Review*, 6(3):58–75.

Elsner, B. and Isphording, I. E. (2017). A big fish in a small pond: Ability rank and human capital investment. *Journal of Labor Economics*, 35(3):787–828.

Elsner, B. and Isphording, I. E. (2018). Rank, sex, grugs and crime. *Jounal of Human Resources*, 53(2):356–381.

Ertac, S. (2006). Social Comparisons and Optimal Information Revelation: Theory and Experiments. Mimeo, University of California Los Angeles.

Falck, O., Gold, R., and Heblich, S. (2011). Class Size and Education in England. Research Report DFE-RR169, Research Report.

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

Fryer, R. G. and Levitt, S. D. (2006). The Black-White Test Score Gap Through Third Grade. *American Law and Economics Review*, 8(2):249–281.

Genakos, C. and Pagliero, M. (2012). Interim Rank, Risk Taking, and Performance in Dynamic Tournaments. *Journal of Political Economy*, 120(4):782 – 813.

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

Gibbons, S. and Telhaj, S. (2016). Peer effects: evidence from secondary school transitions in england. *Oxford Bulleting of Economics and Statistics*, 78(4):548–575.

Guiso, L., Monte, F., Sapienza, P., and Zingales, L. (2008). Culture, gender, and math. *Science*, 320(5880):1164–1165.

Hanushek, E. A. and Rivkin, S. G. (2006). School Quality and the Black-White Achievement Gap. NBER Working Papers 12651, National Bureau of Economic Research, Inc.

Hanushek, E. A. and Rivkin, S. G. (2009). Harming the best: How schools affect the black-white achievement gap. *Journal of Policy Analysis and Management*, 28(3):366–393.

Hastings, J., Neilson, C., and Zimmerman, S. (2013). Are some degrees worth more than others? evidence from college admission cutoffs in chile. Technical report, NBER Working Paper 19241.

Heckman, J. J. and Rubinstein, Y. (2001). The Importance of Noncognitive Skills: Lessons from the GED Testing Program. *American Economic Review*, 91(2):145–149.

Hopkins, E. and Kornienko, T. (2004). Running to keep in the same place:consumer choice as a game of status. *American Economic Review*, 91(2):1085–1107.

Hoxby, C. (2000). Peer effects in the classroom: learning from gender and race variation. Technical report, NBER Working Paper 7867.

Hoxby, C. and Weingarth, G. (2005). Taking race out of the equation: school reassignment and the structure of peer effects. Technical report, Department of Economics, Harvard University.

i Vidal, J. B. and Nossol, M. (2011). Tournaments without prizes: Evidence from personnel records. *Management Science*, 57(10):1721–1736.

Joensen, J. S. and Nielsen, H. S. (2009). Is there a causal effect of high school math on labor market outcomes? *Journal of Human Resources*, 44(1):171–198.

Kinsler, J., Pavan, R., and DiSalvo, R. (2014). Distorted Beliefs and Parental Investment in Children. Mimeo, SOLE-JOLE 2015.

Kirkeboen, L. J., Leuven, E., and Mogstad, M. (2016). Field of study, earnings, and self-selection. *The Quarterly Journal of Economics*, 131(3):1057–1111.

Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental Analysis of Neighborhood Effects. *Econometrica*, 75(1):83–119.

Kremer, M. and Levy, D. (2008). Peer effects and alcohol use among college students. *The Journal of Economic Perspectives*, 22(3):189–206.

Kuziemko, I., Buell, R., Reich, T., and Norton, M. (2014). Last-place aversion: Evidence and redistributive implications. *Quarterly Journal of Economics*, 129(1):105–149.

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

Leflot, G., Onghena, P., and Colpin, H. (2010). Teacher-child interactions: Relations with children's self-concept in second grade. *Infant and Child Development*, 19(4):385–405.

Lindqvist, E. and Vestman, R. (2011). The labor market returns to cognitive and noncognitive ability: Evidence from the swedish enlistment. *American Economic Journal: Applied Economics*, 3(1):101–128.

Marsh, H., Seaton, M., Trautwein, U., Ludtke, O., Hau, O'Mara, A., and Craven, R. (2008). The Big Fish Little Pond Effect Stands Up to Critical Scrutiny: Implications for Theory, Methodology, and Future Research. *Educational Psychology Review*, 20(3):319–350.

Marsh, H. W., Byrne, B. M., and Shavelson, R. J. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of Educational Psychology*, 80(3):366–380.

Murphy, R. and Weinhardt, F. (2013). The Importance of Rank Position. CEP Discussion Paper 1241, Centre for Economic Performance, London School of Economics.

Murphy, R. and Weinhardt, F. (2014). Top of the Class: The Importance of Ordinal Rank. CESifo Working Paper 4815, CESifo Munich.

O'Mara, A., Green, J., and Marsh, H. (2006). Administering self-concept interventions in schools: No training necessary? A meta-analysis. *International Education Journal*, 7:524–533.

Oster, E. (2017). Unobservable selection and coefficient stability, theory and evidence. *Journal of Business and Economic Statistics*, 0(0):1–18.

Reback, R., Rockhoff, J., and Schwartz, H. L. (2014). Under pressure: Job security, resource allocation, and productivity in schools under no child left behind. *American Economic Journal: Economic Policy*, 6(3):207–41.

Rivkin, S. G., Hanushek, E. A., and Kain, J. F. (2005). Teachers, Schools, and Academic Achievement. *Econometrica*, 73(2):417–458.

Robles, V. C. F. and Krishna, K. (2012). Affirmative Action in Higher Education in India: Targeting, Catch Up, and Mismatch. NBER Working Papers 17727, National Bureau of Economic Research.

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

Rubie-Davies, C. (2012). Teacher Expectations and Student Self-Perceptions: Exploring Relationships. *Psychology in the Schools*, 43(5):537–552.

Sacerdote, B. (2001). Peer effects with random assignment: Results for dartmouth roommates. *Quarterly Journal of Economics*, 116(2):681–704.

Schennach, S. M. (2016). Recent advances in the measurement error literature. *Annual Review of Economics*, 8(1):341–377.

Tidemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *Journal of Educational Psychology*, 92(1):144–151.

Tincani, M. M. (2015). Heterogeneous peer effects and rank concerns: Theory and evidence. Mimeo, University College London.

Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157):1124–1131.

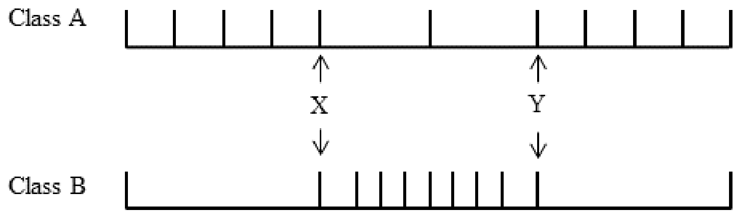
Valentine, J. C., DuBois, D. L., and Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, 39(2):111–133.

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

Yeung, A. S. and Lee, F. L. (1999). Self-concept of high school students in china: Confirmatory factor analysis of longitudinal data. *Educational and Psychological Measurement*, 59(3):431–450.

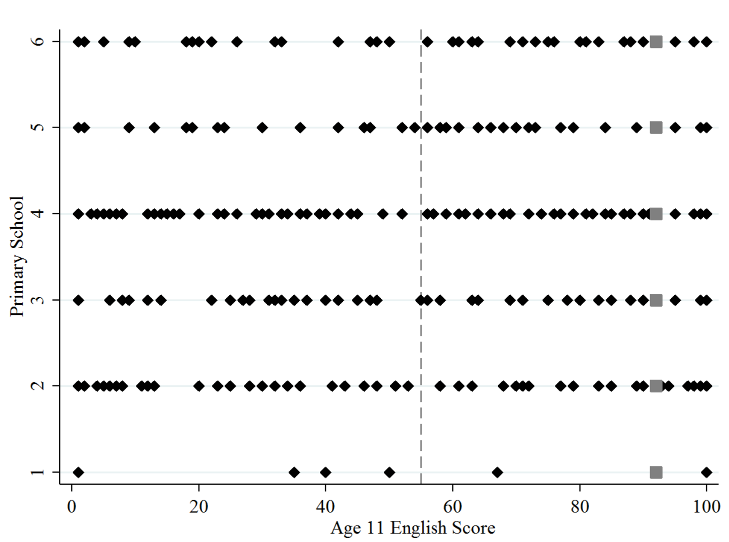
Figures

Figure 1: Rank Dependent on Distribution of Test Scores



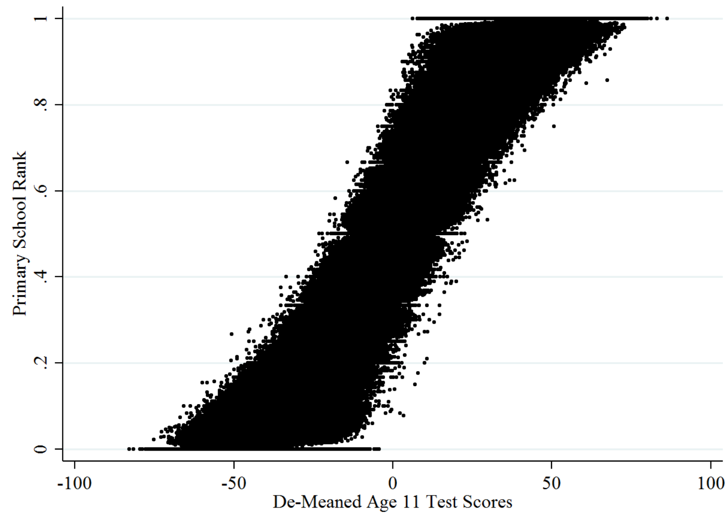
Notes: The figure shows two classes of eleven students, with each mark representing a student's test score, which are increasing from left to right. The classes have the same mean, minimum, and maximum student test scores. However, two students with the same absolute and relative-to-the-mean test scores can still have different ranks. For example, a student with a test score of Y in class A would have a lower rank (ranked fifth out of eleven) than a student with the same test score in class B (ranked second out of eleven). Similarly, a test score of X would be ranked differently in classes A and B.

Figure 2: Test Score Distributions Across Similar Classes



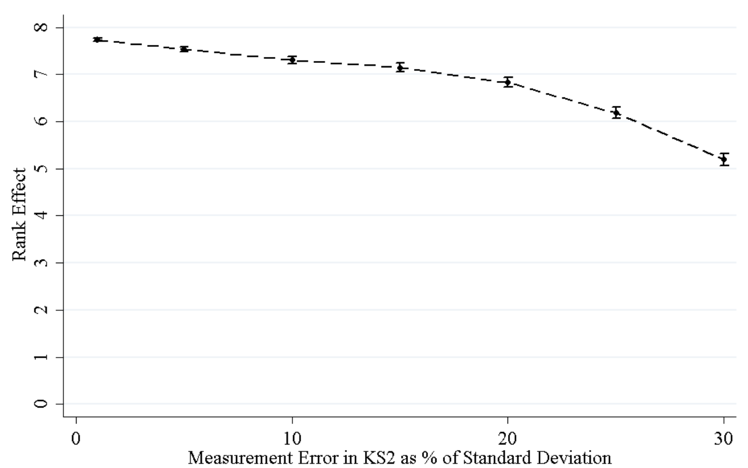
Notes: This graph presents data from six primary school English classes which all have a student scoring the minimum and maximum and have a mean test score of 55 (as indicated by the dashed grey line). Each diamond represents a student score, and gray squares indicate all students who scored 92. Given the different test score distributions, each student scoring 92 has a different rank. This rank is increasing from school 1 through to school 6 with ranks of 0.83, 0.84, 0.89, 0.90, 0.925 and 0.94 respectively, despite all students having the same absolute and relative-to-the-class-mean test score. Note that individual test scores have been randomly altered enough to ensure anonymity of individuals and schools, which is for illustrative purposes only and in no way affects the interpretation of these figures.

Figure 3: Rank Distributions Across School-Subject-Cohorts Subjects



Notes: The Y-axis is the primary rank of students and the X-axis shows the de-meaned test scores by primary school-subject-cohort (SSC). Note that individual test scores have been randomly altered enough to ensure anonymity of individuals and schools, which is for illustrative purposes only and in no way affects the interpretation of these figures.

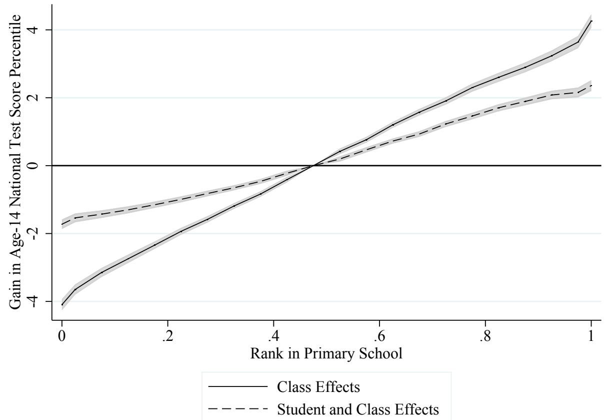
Figure 4: Estimates from Monte Carlo Simulations with Additional Measurement Error in Baseline Test Scores



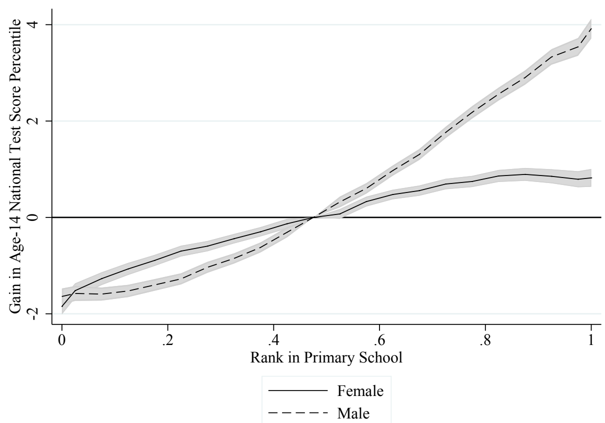
Notes: This figure plots the mean rank estimate from a 1000 simulations of Specification 2 with increasing additional measurement error added to student baseline test scores before computing ranks. The measurement error is drawn from a normal distribution with mean zero and a standard deviation that is proportional to the standard deviation in baseline test scores (28.08). The measurement error of each subject within a student is independently drawn. The error bars represent the 2.5th and the 97.5th percentiles from the sampling distribution of beta for each measurement error level.

Figure 5: Impact of Primary Rank on age-14 Test Scores

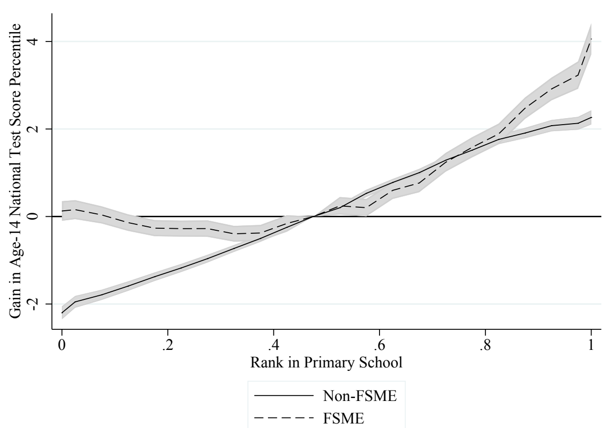
Figure 5.1: Non-Parametric Estimation



5.2: Non-Linear Effects by Gender



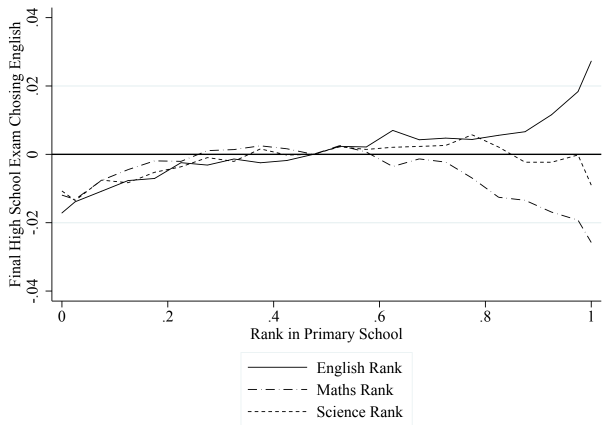
5.3 Non-Linear Effects by Free School Meal Eligibility



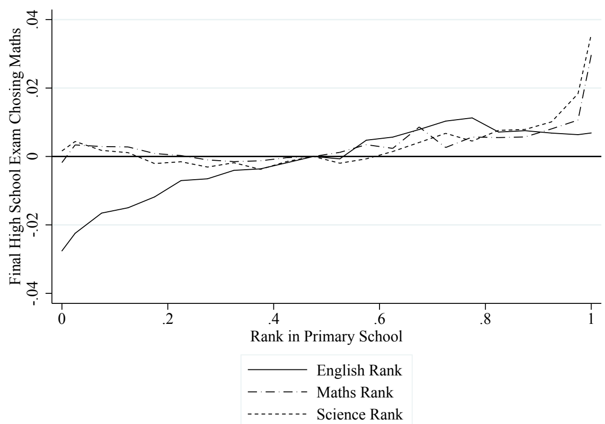
Notes: The panels show the impact of rank in primary school using versions of Specification 4 allowing the effect of rank to vary by ventile and a dummy for being top or bottom of class (SSC). The reference ventile are those from the 45-50th percentiles. FSME stands for Free School Meal Eligible student. Effects obtained from estimating the effect of rank on Non-FSME (Female) students and the interaction term with FSME (Male) students. All estimates have cubic controls for baseline test scores and condition on SSC effects and student effects. Shaded area represents 95% confidence intervals. Standard errors are clustered at the secondary school level.

Figure 6: Impact of Primary Subject-Rank on A-Level Choices

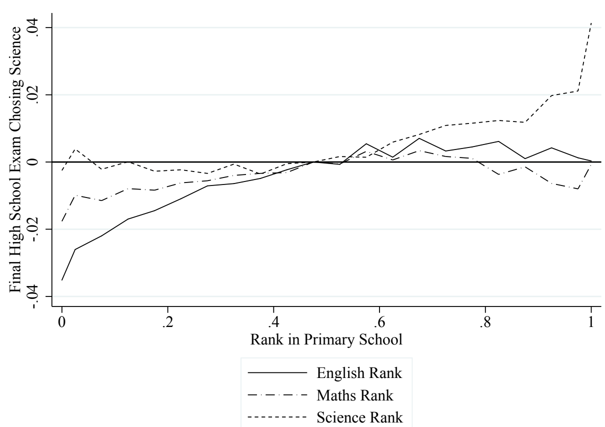
Figure 6.1: English



6.2: Maths

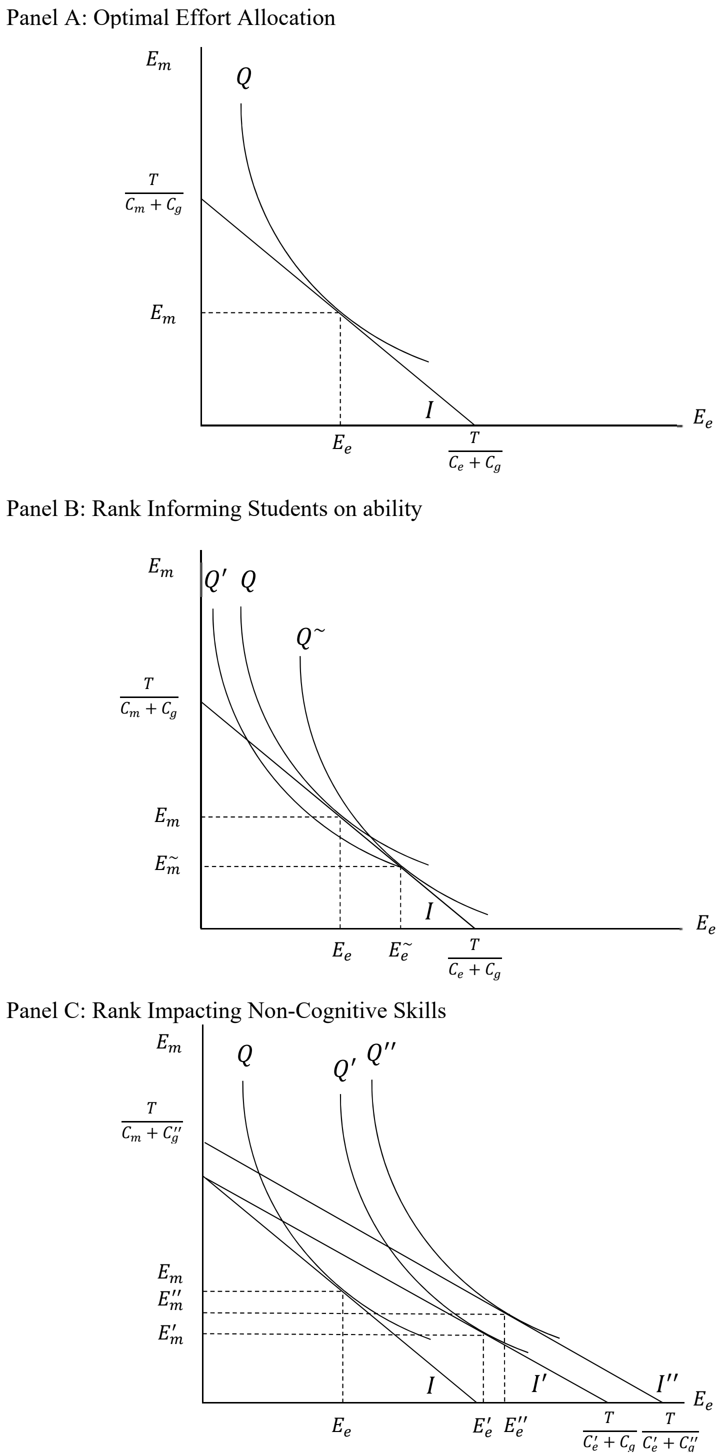


6.3 Science



Notes: The panels show the impact of rank in primary school for each of the three subjects on the likelihood of completing each of the subjects at the A-Level. The estimates come from a version of Specification 4 allowing the effects of rank to vary by ventile. The reference ventile are those from the 45-50th percentiles. Each panel represents the results from three separate regressions for each subject rank (and baseline test scores) in primary school. The take up rate for the A-Level subjects are English 12.3%, Maths 8.4% and Science 10.8%.

Figure 7: Optimal Allocation of Effort Across Subjects, Confidence Model



Notes: These figures show students' optimal effort allocation between maths $E_m$and English E_e. Students have costs of effort for each subject C_m, C_e and a general cost of effort C_g. Students are willing to provide a total cost of effort T, have perceived iso-quants Q and iso-costs I. In panel B, under the learning about abilities model $Q^{\sim}$represents perceived iso-quant, and $E_e^{\sim}$ $E_m^{\sim}$the resulting chosen effort levels in English and maths respectively. In panel C, under the non-cognitive skills model Q' and I' represent the iso-quant and iso-cost lines with lower costs of English effort, where $E_e^{'}$ and $E_m^{'}$ are the resultant chosen effort levels. Finally, Q'' and I'' represent the iso-quant and iso-costs with lower costs of English effort and lower costs of general academic effort, where E_e'' and E_m'' are the resultant chosen effort levels. For more details, see section 7.4.2.

Tables

Table 1: Descriptive Statistics of the Main Estimation Sample

	Mean	S.D.	Min	Max
Panel A: Student Test Scores				
Age 11 National Test Scores Percentile				
English	50.285	28.027	1	100
Maths	50.515	28.189	1	100
Science	50.005	28.026	1	100
Age 11 Rank				
English	0.488	0.296	0	1
Maths	0.491	0.296	0	1
Science	0.485	0.295	0	1
Within Student Rank SD	0.138	0.087	0	0.577
Age 14 National Test Scores Percentile				
English	51.233	28.175	1	100
Maths	52.888	27.545	1	100
Science	52.908	27.525	1	100
Age 16 National Test Scores Percentile				
English	41.783	26.724	1	94
Maths	43.074	27.014	1	96
Science	41.807	26.855	1	94
Age 18 Subjects Completed				
English	0.123	0.328	0	1
Maths	0.084	0.277	0	1
Science	0.108	0.31	0	1
Panel B: Student Background Characteristics				
FSME	0.146	0.353	0	1
Male	0.499	0.5	0	1
Minority	0.163	0.37	0	1
Panel C: Observations				
Students	2,271,999			
Primary Schools	14,500			
Secondary Schools	3,800			

Notes: 6,815,997 student-subject observations over 5 cohorts. Cohort 1 takes age 11 examinations in 2001, age 14 examinations in 2004, age-16 examinations in 2006 and A-Levels at age-18 in 2008. Test scores are percentalised by cohort-subject and come from national exams which are externally marked. Age-16 test scores mark the end of compulsory education. Age-18 information could be merged for a sub-sample of 5,147,193 observations from cohorts 2 to 5. For a detailed description of the data, see section 3.

Table 2: LSYPE Sample: Descriptive Statistics

	Mean	S.D.	Min	Max
Panel A: Student Descriptive Statistics				
How good do you think you are at...				
English	0.928	0.928	-2	2
Maths	0.944	0.917	-2	2
Science	0.904	1.008	-2	2
Age 11 National Test Scores Percentile				
English	50.114	27.725	1	100
Maths	50.783	28.378	1	100
Science	49.453	28.287	1	100
Age 11 Rank				
English	0.496	0.295	0	1
Maths	0.501	0.297	0	1
Science	0.489	0.294	0	1
Within Student Rank SD	0.137	0.089	0	0.575
Student Characteristics				
FSME	0.180	0.384	0	1
Male	0.498	0.500	0	1
Minority	0.337	0.473	0	1
Panel B: Parental Descriptive Statistics				
Any Post-Secondary Qualification	0.323	0.468	0	1
Gross Household Income>£33,000	0.219	0.413	0	1
Occupation				
English	0.014	0.119	0	1
Maths	0.031	0.175	0	1
Science	0.036	0.185	0	1
Parental Time Investment in Schooling				
Number Attending Parents Evening	1.213	0.713	0	2
Special Meeting with Teacher About Child	0.235	0.424	0	1
Frequency of Talking with Child's Teacher	2.124	0.957	1	5
Involved in Child's School Life	2.969	0.782	1	4
Parental Financial Investment in Schooling				
Paying for Any Out of School Tuition	0.234	0.423	0	1
English Tuition	0.057	0.232	0	1
Maths Tuition	0.027	0.163	0	1
Science Tuition	0.019	0.137	0	1
Panel C: Observations				
Students	10,318			
Primary Schools	4,137			
Secondary Schools	780			

Notes: The LSYPE sample consists of 34,674 observations from the cohort 1 who took age-11 exams in 2001 and age-14 exams in 2004. For a detailed description of the data see section 3.

Table 3: Rank Effects on Student Outcomes

	Age 14 Test Scores			Age 16 Test Scores			Complete Subject Age 18		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Primary Rank	7.946***	7.894***	4.562***	6.468***	6.389***	3.129***	0.025***	0.025***	0.035***
	0.147	*0.147*	0.107	0.154	0.152	0.107	0.002	0.002	0.002
Male		-1.398***	..		-2.526***	..		-0.012***	..
		0.045	..		*0.056*	..		*0.000*	..
FSME		-3.107***	..		-4.329***	..		-0.017***	..
		0.030	..		*0.041*	..		*0.000*	..
Minority		1.978***	..		4.329***	..		0.045***	..
		0.054	..		*0.083*	..		*0.001*	..
Cubic in Age 11 Test Scores	✓	✓	✓	✓	✓	✓	✓	✓	✓
Primary SSC Effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Secondary School Effects			Abs.			Abs.			Abs.
Student Effects			✓			✓			✓

Notes: Results obtained from nine separate regressions based on 2,271,999 student observations and 6,815,997 student-subject observations. In columns 1-3 the dependent variable is by cohort by subject percentalised KS3 test scores. In columns 4-6 the dependent variable is by cohort by subject percentalised KS4 test scores. In columns 7-9 the dependent variable is an indicator for completing an A-Level at age 18 in the corresponding subject. SSC effects are fixed effects for each primary school-by-subject-by-cohort combination. Standard errors in italics and clustered at the secondary school level (3,800). "Abs." indicates that the effect is absorbed by another estimated effect. *** 1%,** 5%, * 10% significance

Table 4: Balancing Table

	(1)	(2)
Panel A: Parental occupation		
Age-11 test scores	8.943***	1.534*
	1.207	*0.846*
Panel B: Balancing by rank		
Parental Occupation	0.004	0.001
	0.004	*0.004*
Parental Education	-0.004	-
	0.063	
Household Income>£33k	-0.014	-
	0.052	
Male	-0.008**	-
	0.004	
FSME	-0.008***	-
	0.003	
Minority	0.008***	-
	0.002	
Predicted Age 14 Test Scores	0.113***	-
	0.031	
Cubic in Age 11 Test Scores (excluding Panel A)	✓	✓
Primary SSC Effects	✓	✓
Student Effects		✓

Notes: Rows 1 and 2 are based on 31,050 subject-student observations for which parental occupations could be identified from the LSYPE (for details see section 3.4). Panel A establishes the relevance of the parental occupation variable by estimating effects on age-11 test scores. Panels B shows balancing of rank with the listed student characteristics as dependent variables. Predicted age-14 test scores are generated from a linear projection of student observables and SSC effects. Primary SSC effects are fixed effects for each school-by-subject-by-cohort combination. As parental occupation varies across subjects within a student, regressions in column 2 include additional student fixed effects. Standard errors in italics and clustered at the secondary school level (3,800). *** 1%,** 5%, * 10% significance.

Table 5: Alternative Specifications

		Age 14 Test Scores	
		(1)	(2)
(1)	Main Specifications - Benchmark	7.894***	4.562**
		0.147	*0.107*
(2)	Fully Flexible Age 11 Test Scores	7.543***	4.402***
		0.146	*0.107*
(3)	Randomised Cohort Within School	2.132***	0.869**
		0.154	*0.140*
(4)	Accounting for Primary Class Variance	5.718***	4.337***
		0.156	*0.118*
(5)	Small Primary Schools (Single Class) Only	6.469***	3.646***
		0.176	*0.146*
(6)	Randomised School Within Cohort	-0.099	-0.227
		0.130	*0.150*
(7)	No Prior Peers in Secondary School	10.461**	5.011***
		0.449	*0.415*
(8)	Excluding Specialist Secondary Schools	7.875***	4.586***
		0.155	*0.112*
(9)	Accounting for Secondary-Cohort Subject FX	7.942***	4.471***
		0.146	*0.106*
Student Characteristics		✓	Abs
Cubic in Age 11 Test Scores		✓	✓
Primary SSC Effects		✓	✓
Student Effects			✓

Notes: This table is discussed in section 5 (row 1-7) and in section 7.2 (row 8 and 9). Results obtained from 18 separate regressions. Rows 1, 2, 3, 4, 6 and 9 use the main sample of 6,815,997 student-subject observations. Row 5 uses a reduced sample of 2,041,902 student-subject observations who attended primary schools with cohort sizes of less than 31. Row 7 is estimated using a sample of 452,088 observations, which have no primary peers in secondary school. Row 8 uses a reduced sample of 6,235,806 student-subject observations who did not attend a secondary school classified as specialist. The dependent variable is by cohort by subject percentalised KS3 test scores. Student characteristics are ethnicity, gender, and Free School Meal Eligiblity (FSME). SSC effects are fixed effects for each school-by-subject-by-cohort combination. Standard errors in italics and clustered at the secondary school level. *** 1%,** 5%, * 10% significance

Table 6: Subject-specific Educational Production

	Age 14 Test Scores		
	English	Science	Maths
	(1)	(2)	(3)
Panel A:Same Subject Effects			
Rank	7.400***	8.373***	8.820***
	0.202	*0.192*	*0.188*
Panel B: Cross Subject Effects			
English Rank	5.423***	1.417***	-0.140
	0.189	*0.163*	*0.138*
Science Rank	0.597***	5.566***	3.233***
	0.171	*0.167*	*0.137*
Maths Rank	0.788***	3.612***	7.753***
	0.178	*0.174*	*0.173*
Student Characteristics	✓	✓	✓
Cubic in Age 11 Test Scores	✓	✓	✓
Primary SSC Effects	✓	✓	✓

Notes: Each column of panel A estimates specification 2 separately by subject. Each column of panel B additionally allows for cross-subject effects of ranks and test scores. Standard errors in italics and clustered at the secondary school level (3,800 schools). *** 1%,** 5%, * 10% significance.

Table 7: Parental Investments and Secondary Schools

	(1)	(2)
Panel A: Parental Time Investments		
Number of Parents Attending Parents Evening	0.248***	-0.167
	0.033	*0.102*
Special Meeting with Teacher About Child	-0.119***	0.053
	0.018	*0.062*
Frequency of Talking with Child's Teacher	-0.215***	0.170
	0.041	*0.129*
Involved in Child's School Life	0.028	-0.114
	0.034	*0.110*
Index of Parental School Involvement	-0.224***	0.042
	0.086	*0.146*
Panel B: Parental Financial Investments		
Paying for Any Out of School Tuition	0.148***	-0.066
	0.018	*0.062*
Paying for Out of School Tuition by Subject	-0.029***	-0.007
	0.006	*0.026*
Panel C: Sorting to Secondary by Subject Value Added		
Unconditional Value Added	*0.365****	*-0.002*
	0.016	*0.009*
Conditional Value Added	*0.367****	*-0.002*
	0.016	*0.010*
Cubic in Age 11 Test Scores (excluding Panel A)		✓
Primary SSC Effects	✓	✓

Notes: Results obtained from 18 separate regressions. Regressions in panel A and B are based on 11,558 student observations and 34,674 student-subject observations from the LSYPE sample. For descriptives, see Table 2. Regressions from panel C use the main sample. Secondary school subject specific value added is calculated in terms of age 11 to 14 growth in test score percentiles recovered from a secondary school subject fiexed effect. These fixed effects have been standardised to mean zero and standard deviation one. Cohort effects are not included because the LSYPE data is only available for one cohort. Standard errors in parenthesis and clustered at the secondary school level (796/3800). *** 1%,** 5%, * 10% significance.

Table 8: Student Confidence On Rank

	Age 14 Test Scores	Subject Confidence	Male – Subject Confidence	Female – Subject Confidence
	(1)	(2)	(3)	(4)
Primary Rank	8.977***	0.196**	0.285*	-0.009
	1.602	*0.093*	*0.168*	*0.210*
Male	-1.939***	0.124***		
	0.259	*0.014*		
FSME	-2.588***	0.046**	0.032	0.026
	0.383	*0.019*	*0.037*	*0.038*
Minority	2.107***	0.159***	0.113***	0.219***
	0.414	*0.022*	*0.038*	*0.045*
Cubic in Age 11 Test Scores	✓	✓	✓	✓
Primary SSC Effects	✓	✓	✓	✓

Notes: Results obtained from four separate regressions based on 11,558 student observations and 34,674 student-subject observations from the LSYPE sample (17,415 female, 17,259 male). For descriptives, see Table 2. The dependent variable is a coarse measure of confidence by subject. Cohort effects are not included because the LSYPE data is only available for one cohort. Standard errors in parenthesis and clustered at the secondary school level (796). *** 1%,** 5%, * 10% significance.

Table 9: Is the Degree of Misinformation Harmful?

	Age 14 Test Scores			
	(1)	(2)	(3)	(4)
Primary Rank	10.765***	10.768***	10.678***	10.672***
	0.226	*0.225*	*0.225*	*0.224*
Misinformation		0.103		-0.172
		0.266		*0.264*
Student Characteristics			✓	✓
Cubic in Age 11 Test Scores	✓	✓	✓	✓
Primary-cohort Effects	✓	✓	✓	✓

Notes: Results obtained from four separate regressions based on 2,271,999 student observations averaged over subjects where columns 2 and 4 include an additional explanatory variable of the degree of misinformation. The dependent variable is age-14 test scores. Rank is the average rank across English, maths and science in primary school. The misinformation measurement is the average absolute difference between local rank and national percentile rank for each student in the end-of-primary school test score as explained in section 7.4.2. Student characteristics are ethnicity, gender, and Free School Meal Eligibility (FSME). Standard errors in italics and clustered at 3,800 secondary schools. *** 1%,** 5%, * 10% significance.

Appendix (for online publication)

A.1 Peer Effects

There are concerns that, with the existence of peer effects, peer quality jointly determines both a student's rank position and their age-11 test scores. This mechanical relationship could bias our estimation because, in the presence of peer effects, a student with lower quality peers would attain lower age-11 test scores and gain a higher rank than otherwise. Thus, when controlling for prior test scores in the age-14 estimations, when students have a new peer group, those who previously had low quality peers in primary school would appear to gain more. Since rank is negatively correlated with peer quality in primary school, it would appear that those with high rank experience the largest gains. Therefore, having a measure of ability confounded by peer effects would lead to an upward-biased rank coefficient.

We propose a solution through the inclusion of subject-by-cohort-by-primary school controls. These effects will absorb any average peer effects within a classroom. However, they will not absorb any peer effects that are individual-specific. This is because all students will have a different set of peers (because they cannot be a peer to themselves). Therefore, including class level effects will remove only the average class peer effect. The remaining bias will be dependent on the difference between the average peer effect and the individual peer effect and its correlation with rank. We are confident that the remaining effect of peers on the rank parameter will be negligible, given that the difference between average and individual peer effect decreases as class size increases. The bias will be further attenuated because the correlation between the difference and rank will be less than one, and both effects are small.

We test this by running simulations of a data generating process, where test scores are not affected by rank and are only a function of ability and school/peer effects. We then estimate the rank parameter given this data. We allow for the data-generating process to have linear mean-peer effects, as well as non-linear peer effects Lavy et al. (2012). The non-linear peer effect is determined by the total number of peers in the class in the bottom 5 percent of students in the population. We are conservative and assume extremely large peer effects, allowing both types of peer effects to account for 10 percent of the variance of a student's subject-specific outcome. Given that the square root of the explained variance is the correlation coefficient, this assumption implies that a one standard deviation increase in peer quality improves test scores by 0.31 standard deviations. In reality, Lavy et al. (2012) find a one standard deviation increase in peers increases test scores by only 0.015 standard deviations, which is one 20th the size.

The data generating process is as follows:

- We create 2,900 students attending 101 primary schools and 18 secondary schools of varying sizes.
- A range of factors are used to determine achievement. Each of these factors are assigned a weight, such that the sum of the weights equals one. This means that weights can be interpreted as the proportion of the explained variance.
- Students have a general ability α_i and a subject specific ability δ_{is} taken from normal distributions with mean zero and standard deviation one. Taken together, they are given a weight of 0.7, as the within school variance of student achievement in the raw data is 0.85. They are given a weight of 0.6 where rank effects exist.
- All schools are heterogeneous in their impact on student outcomes. These are taken from normal distributions with mean zero and standard deviation one. School effects are given a weight of 0.1, as the across school variance in student achievement in the raw data is 0.15.

- Linear mean peer effects are the mean subject and general ability of peers not including themselves. Non-linear peer effect is the negative of the total number of peers in the bottom 5 percent of students in the population in that subject. Peer effects are given a weight of 0.1, which is much larger than reality.

- We allow for measurement error in test scores to account for 10 percent of the variance.

- We generate individual i's test scores as a function of general ability α_i, subject specific ability δ_{is}, primary peer subject effects ρ_{ijs} or secondary peer subject effects σ_{iks}, primary school effects μ_j or secondary school effects π_k, age-11 and 14 measurement error ε_{ijs} or ε_{ijks} , and primary school Rank ω_{ijs}.

  – age-11 test scores
  $$Y^0_{ijs} = 0.7(\alpha_i + \delta_{is}) + 0.1\mu_j + 0.1\rho_{ijs} + 0.1\varepsilon_{ijs}$$

  – age-14 test scores where rank has no effect (panel A):
  $$Y^1_{ijks} = 0.7(\alpha_i + \delta_{is}) + 0.1\pi_k + 0.1\sigma_{iks} + 0.1\varepsilon_{ijks}$$

  – age-14 test scores where rank has an effect (panel B):
  $$Y^1_{ijks} = 0.6(\alpha_i + \delta_{is}) + 0.1\pi_k + 0.1\sigma_{iks} + 0.1\omega_{ijs} + 0.1\varepsilon_{ijks}$$

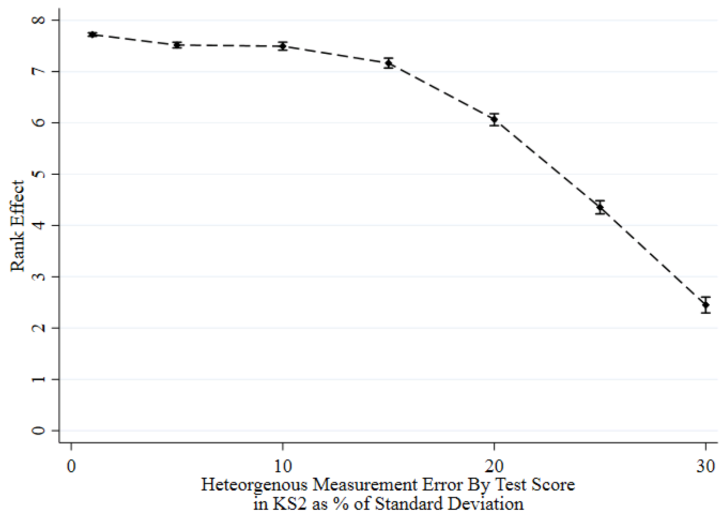
  We simulate the data 1,000 times and each time estimate the rank parameter using the following specifications with and without school-subject effects, with and without school-subject effects.
  $$Y^1_{ijks} = \beta^1_{Rank} R_{ijs} + \beta^1_y Y^0_{ijs} + \varepsilon_{ijks}$$

The results from these estimations can be found in Appendix Table A.2. Panel A assumes that there is no rank effect, and we would expect $\beta_{Rank} = 0$. Panel B has a rank effect in the data generating process of 0.1, so we would expect $\beta_{Rank} = 0.1$. Columns 1 & 2 show estimates with linear-in-means peers effects, and columns 3 & 4 show estimates with non-linear peer effects. With these inflated peer effects, we see upward bias in the rank effect, showing a rank effect where none exists (panel A columns 1 and 3). When including SSC effects, this positive bias is removed (columns 2 & 4). With large linear-in-mean peer effects, there is no remaining bias. With non-linear peer effects twenty times greater than those found in reality, the inclusion of SSC effects introduces a slight negative bias; therefore, our results can be considered upper bounds.

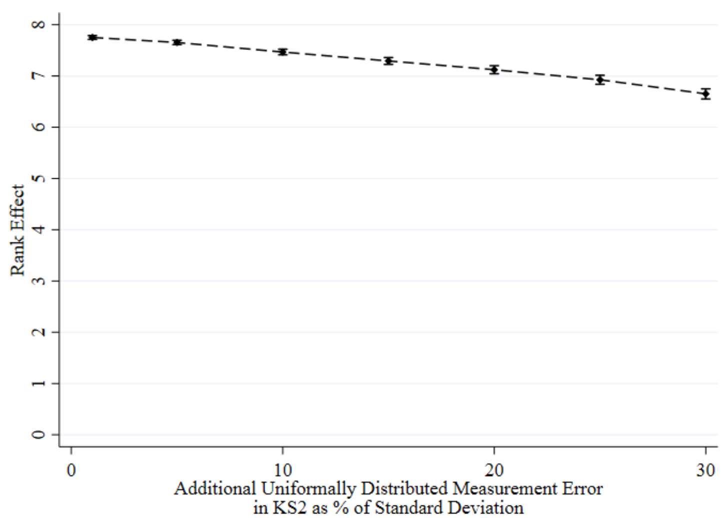
Appendix Figures and Tables

Figure A.1: Rank estimate with different types of additional noise in the baseline score

Panel A: Noise by ability



Panel B: Uniform Noise



Notes: These figures plot the mean rank estimate from a 1,000 simulations of specification 2 with increasing additional measurement error added to student baseline test scores. The measurement error for each student within a student is independently drawn. The error bars represent the 2.5th and the 97.5th percentiles from the sampling distribution of beta for each measurement error level. The extent of measurement error in panel A is increasing linearly in distance from the 50th percentile in the national test score distribution, such that students at the 50th percentile experience no measurement error and that students at the 1st and 100th percentile experience additional measurement error drawn from a normal distribution with a mean zero and standard deviation equal to the proportion of the standard deviation in baseline test scores represented on the X-axis. The measurement error in panel B is drawn from a uniform distribution with mean zero and a standard deviation that is proportional to the standard deviation in baseline test scores.

Table A.1: Differences in Characteristics Between Main and LSYPE Samples

	Main Sample	LSYPE Sample	Difference	Standard Error
Panel A: Student Test Scores				
Age 11 National Test Scores Percentile				
English	50.285	50.114	-0.171	0.277
Maths	50.515	50.783	0.268	0.278
Science	50.005	49.453	-0.552	0.277
Age 11 Rank				
English	0.488	0.496	0.008	0.003
Maths	0.491	0.501	0.010	0.003
Science	0.485	0.489	0.004	0.003
Age 14 National Test Scores Percentile				
English	51.233	51.376	0.143	0.278
Maths	52.888	53.779	0.891	0.272
Science	52.908	53.051	0.143	0.272
Panel B: Student Background Characteristics				
FSME	0.146	0.180	0.034	0.002
Male	0.499	0.498	-0.001	0.003
Minority	0.163	0.337	0.174	0.002
Panel C: Observations				
Students	2,271,999			
Primary Schools	14,500			
Secondary Schools	3,800			

Notes: The table presents means characteristics from the main and the LSYPE samples, and their raw differences. The standard errors are unclustered.

Table A.2: Simulation of Rank Estimation with Peer Effects

	Mean peer effects		Non-linear Peer Effects	
	(1)	(2)	(3)	(4)
Panel A: Rank has no effect β_{rank}=0.0				
Mean $\hat{\beta}_{rank}$	0.046	0.000	0.302	-0.041
Mean SE of $\hat{\beta}_{rank}$	*0.014*	*0.018*	*0.015*	*0.019*
SE of $\hat{\beta}_{rank}$	*0.015*	*0.019*	*0.031*	*0.020*
95% Lower Bound	0.015	-0.037	0.243	-0.079
95% Upper Bound	0.077	0.035	0.364	-0.003
Panel B: Rank has an effect β_{rank}=0.1				
Mean $\hat{\beta}_{rank}$	0.099	0.100	0.304	0.068
Mean SE of $\hat{\beta}_{rank}$	*0.014*	*0.017*	*0.014*	*0.018*
SE of $\hat{\beta}_{rank}$	*0.015*	*0.018*	*0.027*	*0.018*
95% Lower Bound	0.069	0.066	0.252	0.033
95% Upper Bound	0.129	0.133	0.358	0.104
KS2 and Rank	✓	✓	✓	✓
School-Subject-Effects		✓		✓

Notes: 1,000 iterations, 95% confidence bounds are obtained from 2.5th and 97.5th estimate of ordered rank parameters. For details see Appendix section A.1.

Table A.3: Specification Check 1/2: Functional form of baseline test scores

	Age 14 Test Scores					
Degree of Polynomial	Linear (1)	+Quadratic (2)	+Cubic (3)	+Quartic (4)	+Quintic (5)	+Sextic (6)
Primary Rank	10.704***	10.724***	7.894***	7.859***	7.622***	7.619***
	0.178	*0.151*	*0.147*	*0.147*	*0.146*	*0.146*
Minority	1.847***	1.969***	1.978***	1.979***	1.980***	1.980***
	0.054	*0.054*	*0.054*	*0.054*	*0.054*	*0.054*
Male	-1.401***	-1.394***	-1.398***	-1.398***	-1.398***	-1.398***
	0.045	*0.046*	*0.045*	*0.045*	*0.045*	*0.045*
FSME	-3.238***	-3.167***	-3.107***	-3.104***	-3.104***	-3.104***
	0.031	*0.031*	*0.030*	*0.031*	*0.031*	*0.031*
Primary SSC Effects	✓	✓	✓	✓	✓	✓

Notes: The specification in the first row of column 3 corresponds to column 2 of the main results Table 3. The specifications differ in the degree of polynomials that are allowed for the baseline.

Table A.4: Specification Check 2/2: Interacting past performance with school, subject of cohort effects

	Linear (1)	Cubic (2)	Fully Flexible (3)
Age 14 Test Scores	10.802***	7.946***	7.589***
	0.179	*0.147*	*0.147*
Age 14 Test Scores interacted with primary school FX	11.393***	11.482***	12.095***
	0.180	*0.126*	*0.153*
Age 14 Test Scores interacted with subject FX	11.219***	8.299***	7.890***
	0.179	*0.147*	*0.146*
Age 14 Test Scores interacted with cohort FX	10.809***	7.939***	7.575***
	0.179	*0.147*	*0.147*

Notes: The specification in the first row of column 2 corresponds to column 1 of the main results Table 3. The column headings refer to how we control for prior performance at age-11, moving from linear to cubic to fully flexible. In rows 2 to 4, these age-11 test score controls are additionally interacted with different sets of fixed effects, relaxing the assumption that the baseline test score has identical effects across all primary schools, subjects or cohorts.

Table A.5: Sorting to Secondary Schools

	Unconditional Value Added		Conditional Value Added	
	(1)	(2)	(3)	(4)
Panel A: Secondary School				
Primary Rank	-0.016	-0.018*	-0.017	-0.019
	0.011	*0.011*	*0.011*	*0.011*
Panel B: Panel A: Secondary School Subject				
Primary Rank	-0.002	-0.004	-0.002	-0.004
	0.009	*0.009*	*0.010*	*0.009*
Cubic in Age 11 Test Scores	✓	✓	✓	✓
Primary SSC Effects	✓	✓	✓	✓
Student Characteristics		✓		✓

Note: This table shows the results from eight estimations of primary school rank on secondary school value added measures. In columns 1 and 2 of panel A the secondary school value added measures are the school fixed effects in a raw estimation of age-14 test scores on age-11 test scores. In columns 3 and 4 these are again the school fixed effects in an estimation of age-14 test scores on age-11 test scores, but additionally controlling for student demographics. Panel B uses a parallel set of value added measures, but are at the secondary school-subject area, and so use the secondary school by subject fixed effects.