The international platform of Ludwig-Maximilians University's Center for Economic Studies and the Ifo Institute

Working Papers

www.cesifo.org/wp

Top of the Class: The Importance of Ordinal Rank

Richard Murphy
Felix Weinhardt

CESifo Working Paper No. 4815
Category 5: Economics of Education
May 2014

An electronic version of the paper may be downloaded
- *from the SSRN website: www.SSRN.com*
- *from the RePEc website: www.RePEc.org*
- *from the CESifo website: www.CESifo-group.org/wp*

CESifo
Center for Economic Studies & Ifo Institute

CESifo *Working Paper No. 4815*

Top of the Class: The Importance of Ordinal Rank

Abstract

This paper examines the long-run impact of ordinal rank during primary school on productivity using comprehensive English administrative data. Identification is obtained from variation in test score distributions across cohorts and subjects, such that the same score relative to the class mean can have different ranks. Conditional on cardinal measures of achievement, being ranked highly during primary school has large effects on secondary school achievement, with the impact of rank being more important for boys than girls. Using additional survey data we find that the development of confidence is the most likely mechanisms for this effect on task-specific productivity.

JEL-Code: I210, J240, M540.

Keywords: rank, non-cognitive skills, peer effects, productivity.

Richard Murphy
University College London
London / United Kingdom
richard.murphy@ucl.ac.uk

Felix Weinhardt
Humboldt-University Berlin
Berlin / Germany
felix.weinhardt@hu-berlin.de

This version: 10/05/2014
We thank Johannes Abeler, Esteban Aucejo, Thomas Breda, Andrew Clark, Susan Dynarski, Ben Faber, Eric Hanushek, Brian Jacob, Pat Kline, Steve Machin, Magne Mogstad, Imran Rasul, Olmo Silva, Kenneth Wolpin, Gill Wyness, and participants of the CEP Labour Market Workshop, 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. Weinhardt acknowledges ESRC funding (ES/J003867/1). All remaining errors are our own.

1 Introduction

It is human nature to make comparisons against one's peers. Individuals make comparisons in terms of characteristics, traits and abilities in tasks (Festinger, 1954). However, individuals also often use cognitive shortcuts to simplify decision-making (Tversky and Kahneman, 1974). One such shortcut would be to use simple ordinal rank instead of detailed cardinal information. Rather than working out where one stands in relation to the group mean, one might say 'I am taller than Gill but shorter than Sarah'. In this simplified way of conceptualising the world, when we are making decisions one would be placing weight on ordinal rank as well as relative or absolute information. Indeed, it has recently been shown that ordinal rank, in addition to relative position, is used when individuals make comparisons with others (Brown et al., 2008; Kuziemko et al., 2011; Card et al., 2012). If people are ranking themselves amongst their peers, then ordinal in additional to cardinal information has the potential to affect investment decisions. These could in in turn determine later productivity, through various mechanisms such as learning about ability or the development of confidence.

This paper examines, in the context of education, the additional impact of ordinal rank on subsequent productivity. Students in England take externally marked national exams at the end of primary school at age 11. We use this to calculate their rank amongst their primary school peers in three subjects. These students then start attending secondary schools with a new set of peers and are tested again in the same subjects three years later. We use this setting to estimate the effect of age-11 rank on age-14 test scores, in a new peer environment conditional on prior (age-11) relative test scores, in our main specification. To do so, we use administrative data on the entire English public school population as they move from a primary into secondary education.[1]

The rank parameter is identified from the variation in test score distributions across and within primary schools cohorts, so that the same score relative to a school mean can have different ranks. Our estimates show that being highly ranked amongst your peers in a subject has large and robust effects on later performance in that same subject. Moreover, the impact of rank is significant across the entire rank distribution. These estimates use the school-by-subject-by-cohort variation in rank for a given test score and therefore allow for gains from individuals being ranked highly in one subject to impact on results on other subjects. We also provide more demanding within-student specifications, which absorb the average growth rate of a student between age 11 and 14, therefore removing subject-spillovers and so reflect student specialisation. In these specifications, the variation used for estimation is the within-student-across-subject differences in rank conditional on test scores and average prior peer quality. We argue that conditioning on these age-

[1] Public schools account for 93% of the total student population in England. Comparable data for the remaining 7% attending private schools are not available.

11 test scores, primary-subject-cohort effects and individual student effects, the rank of a student in a subject within primary school is effectively random.

Notably, primary school peers determine the rank measure but we estimate its effects on outcomes after the transition to secondary school. This makes our approach resilient to reflection problems (Manski, 1993) as the average student has 87% new peers in secondary school. We are therefore not relating individual and group outcomes from within the same peer group, as cautioned against by Angrist (2013). Moreover, our estimation sidesteps the standard issue of including a lagged dependent variable and individual effects simultaneously (Nickell, 1981), as the individual effects are recovered from test scores across subjects of a student at age 14, rather than from average test scores over time.

The effects of rank that we present are large in the context of the education literature, with a one standard deviation increase in rank improving age-14 test scores by 0.08 standard deviations. This is of comparable magnitude to being taught by a teacher one standard deviation above average (Aaronson, et al., 2007; Hanushek et al., 2005). As expected, the estimates relying on within-student variation in primary rank, conditional on ability, are smaller. Here, a one standard deviation increase in rank improves subsequent test scores in that subject by 0.055 within student standard deviations. This would mean being ranked at the 75th percentile of your primary school peers in a subject as opposed to the 25th percentile, improves age 14 test scores by 0.2 standard deviations in that same subject.

The paper goes on to examine the nature of these effects and finds that they exist throughout the rank distribution, implying that students accurately place themselves within their class, despite not being explicitly informed of their rank. This is likely to occur due to the repeated interactions among peers throughout the six years of primary school as well as seating arrangements that that reflect rank positions in many English primary schools[2]. Moreover, for nearly all rank positions boys are more affected, both positively and negatively, than girls. Boys at the top of the class in a subject gain four times more than comparable girls. Low-income students also gain more from being top of the class but are less negatively affected by being ranked below the median.

Having presented this range of findings, the paper examines and tests threats to identification such as other forms of peer effects, measurement error and sorting to schools by parents. Using simulations we demonstrate that our findings are robust against non-linear peer effects large measurement errors and are not just a statistical artefact. Using additional survey data, we further show that parental occupations predict subject-specific primary attainment but not rank. For

[2] In English Primary schools it is common for students to be seated at tables of four and for them to be set by pupil ability. Students can be sat at the ‘top table’ or the ‘bottom/naughty table’. This 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.’

example, children of accountants do better in Maths than in English. In contrast, parental occupational background has no relation to ordinal rank conditional on attainment. We interpret this as direct evidence for our main identification assumption that primary rank, conditional on class means and own attainment, is effectively random.

Finally the paper discusses a number of mechanisms that could account for these results: learning about own ability (Ertac, 2006); competitiveness; external (parental) investment by task; and environment favouring certain ranks; but provides evidence that the mechanism that best accommodates all the findings is through the development of non-cognitive skills such as confidence. Combining our administrative data with survey data containing direct measures of subject-specific confidence, we show that those who ranked higher in primary school have larger measures of later confidence, conditional on relative test scores and student effects. Mirroring our findings on attainment, we find that boys' confidence is more affected by their school rank than girls' confidence.

To build intuition for the effect of confidence, consider a child being the best in their neighbourhood at basketball. She will consider herself to be good at basketball, gaining confidence in her basketball abilities and resulting in her enjoying basketball more. This would then lead her to invest more time in playing basketball and so further develop her skills. Similarly, in the labour market, individuals rate their productivity in a task relative to their colleagues, and this in turn could determine in which field they specialize. More relevant to this paper, one might consider one's own school career. Upon starting school, we may not know which subjects we are good at. But, through ranking ourselves relative to our peers, we develop a sense that we are a 'math person' or a 'language person'. A 'math person' would be more confident in solving mathematical problems and enjoy math more, and therefore may invest more time into math homework, all of which could be reflected in their future math test scores.

We believe this paper has two main contributions. First, to the best of the authors' knowledge this is the first large-scale study to document the effects of ordinal rank in a task on later productivity. Critically, this study documents an additional effect of ordinal rank, after controlling for prior achievement and the relative distances between peers, i.e. cardinal measures of performance. Therefore, we believe rank could be considered a new factor in the education production function.[3] Besides implications on partial equilibrium considerations of parents regarding the choice of the best school for their children, this finding has more general implications relating to informational transparency and productivity. For instance managers or

[3] There is a broad range of literature on the determinants of academic achievements, including natural ability (Watkins et al., 2007), family background (Hoxby, 2001), school inputs (Hanushek, 2006; Page et al., 2010), peer effects (Carrell et al., 2009; Lavy et al., 2012), and non-cognitive skills (Heckman et al., 2006); however, rank position has not yet been researched.

teachers could improve productivity by emphasising an individual's local rank position if that individual has a high rank. Alternatively, if an individual is in a high performing peer group and therefore may have a low local rank but ranks high globally, a manager should make the global rank more salient.

Secondly, we believe that the result that ordinal rank matters for later outcomes has the potential to add to the explanation of findings in the following education topics where placing individuals amongst high-performing peers has had mixed results: school integration (Angrist and Lang, 2004; Kling *et al.* 2007) selective schools (Cullen, Jacob and Levitt, 2006; Clark 2010); and affirmative action (Arcidiacono *et al.* 2012; Robles and Krishna, 2012). Moreover the finding that rank may exacerbate early differences in achievement due to individual investment decisions based on relative performance contributes to the literatures on ethnicity (Fryer and Levitt, 2006; Hanushek and Rivkin 2006; 2009), gender (Burgess et al, 2004; Machin & McNally, 2005) and relative age in cohort (i.e. Black *et al.*, 2011).

The remainder of the paper is laid out as follows. Section 2 reviews the literature on social comparisons. Section 3 sets out the empirical strategy and how the rank parameter is separately identified from relative achievement. This is followed by a brief description of the UK educational system, the administrative data, as well as the definition of rank used. Section 5 sets out the main results, nonlinearities and the heterogeneity by gender and parental income. Section 6 discusses and tests threats to identification such as peer effects, measurement error and endogenous sorting. Section 7 discusses potential mechanisms and provides additional survey evidence. Section 8 outlines other topics in education, which corroborate these findings. Finally, we conclude and discuss policy implications.

2 Related Literature

The importance of ordinal rank rather than relative position for individuals was first forwarded by Parducci (1965) with range frequency theory. This has the theoretical prediction that comparisons are based upon ordinal position of items within a comparison set. This prediction has been illustrated empirically recently by Brown *et al.* (2008) and Card *et al.* (2012), who show an individual's rank in addition to relative position in an income distribution is an important determinant of satisfaction. However, the economic literature on rank effects on productivity is sparse.[4]

[4] The discussion of social comparisons is often framed in the form of peer effects (Falk and Ichino, 2006;i.e. Mas and Moretti, 2009, Carrell et al., 2009; Lavy et al., 2012) or the introduction of relative achievement feedback mechanisms (Eriksson et al. 2009; Azmat and Iriberri, 2010). These studies tend to find positive effects of peer quality on contemporaneous productivity, and that relative performance feedback increases productivity when there are piece rate incentives.

A related study on rank and informational transparency finds that providing employees with their productivity rank within the firm increased output throughout the productivity distribution (Blanes i Vidal and Nossol 2011). This is explained by workers becoming concerned about their rank position, as the impact occurred after the feedback policy was announced but before the information was released.[5] Genakos and Pagliero (2012) find that in a tournament setting, where payoffs are based on relative performance and with continuous rank feedback, performance decreases as individuals are ranked higher.[6] In both of these papers, individuals are concerned about their relative positions amongst their immediate peers. The education setting of this study varies in two critical ways. Firstly, students are graded on their absolute performance according to national scales, rather than relative to their peers. Secondly, we are estimating the effect of rank amongst previous peers on contemporaneous test scores, and not the effects of rank within the same peer group. Moreover, whilst both of these papers use rank measurements, neither additionally controls for relative distances, and are therefore not separating rank effects from any cardinal relative effects.

The importance of ordinal rank in addition to relative position has been empirically illustrated by Brown et al. (2008) and Card et al. (2012). Most related to our paper, Clark et al. (2010) compares directly the importance of ordinal rather than relative position on discretionary work effort. They find that an employee's income rank was a stronger determinant of stated work effort compared with the average reference group income and so conclude that comparisons are ordinal rather than cardinal. This is similar to our paper as we also in effect estimate effects of rank and relative position, but different because we observe rank effects in a real effort setting rather than in stated amounts.

3 Empirical strategy

3.1 The measurement of rank net of ability and cardinal factors

In order to identify the effect of primary school rank on later outcomes we require variation in rank for a given ability. Moreover, to separately identify the effect of ordinal rank from relative position requires variation in rank for a given distance from mean peer achievement. This comes about through the variation in test score distributions across school cohorts, within and across primary schools, so that students with the same test scores and same distance to peer means can

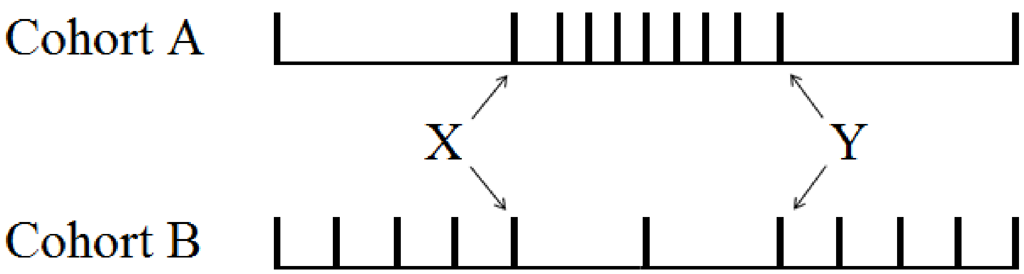
[5] Kosfeld and Neckerman (2011) examine the use of rankings as a non-monetary incentive and find increases in productivity. Specific to education, Jalava, Joensen and Pellas (2013) find that rank based grading increases test performance.

[6] Brown (2011) shows in a tournament setting that when an individual of known outstanding ability (high prior high rank is known) is placed into a group those ranked immediately below them, have a large fall in productivity compared to low ranked participants.

have different ranks. Furthermore, as test score distributions vary across subjects within a school and cohort, a single student with the same score in all three subjects, as well as the same relative distances to peer mean achievements in these subjects, could have different subject ranks. To see this, consider the case illustrated by Figure 1, which shows unimodal and bimodal distributions of hypothetical English test scores for two cohorts of students in a primary school. The school has a very similar intake of students year-on-year, with the same number of students, and moreover the same mean, minimum and maximum student test scores despite having different distributions. As both cohorts have the same mean test scores, students who achieved the same absolute test scores across cohorts (Y), would also have the same relative score compared to the mean of their peers. However, the cohorts have different test score distributions, in the first students are more clusterd around the mean score and in the second test scores are more dispersed and has a bimodal distribution. Due to these different distributions, a student who scored Y in the unimodal cohort is ranked second, whilst the one in the bimodal cohort is ranked fifth.[7]

Note that the rank effect that is identified conditional on the student fixed effect differs in interpretation from school-subject-cohort effects that we just illustrated because it uses the variation in test score distributions *across subjects* within a cohort. The variation used here is analogous to Figure 1 but comparing differences in distributions across subjects rather than within subjects and over time.

Figure 1: Rank dependent on distribution given absolute and relative score



Notes: This figure illustrates that students with the same test score relative to the group mean can have different ranks depending on the distribution of test scores. Two cohorts of eleven students are represented, with each mark representing a student's test score. Test scores are increasing from left to right. Each cohort has the same minimum, maximum and mean test scores. Cohort A has a unimodal distribution and Cohort B has a bimodal test score distribution. A student with a test score of X in Cohort A would have a lower rank than the same test score in Cohort B. Similarly a test score of Y would be ranked differently in Cohorts A and B. Given the definition of rank given in Section 5.2, the rank measurements for score X are R_{xA}= 0.1 and R_{xB}= 0.4 and for Y are R_{yA}= 0.9, R_{yB}= 0.6. This is based on the illustration from Brown et al. (2008).

[7] This is similar to Brown et al. (2008) who rely on the variation in the earnings distributions of a subset of workers across firms to separately identify the effect of relative earnings and ranks in earnings on employee satisfaction.

3.2 A rank-augmented knowledge production function

This section uses the standard education production function approach to derive a rank-augmented value added specification that can be used to identify the effect of primary school rank, measured as outlined in section 3.1, on subsequent outcomes.

To begin, we consider a basic contemporaneous education production function, using the framework as set out in Todd and Wolpin (2003), for student *i* studying subject *s* in primary school *j*, cohort *c* and in time period $t = [1,2]$:

$$Y_{ijsct} = X_i{}'\beta + v_{ijsct} \tag{1}$$

$$v_{ijsct} = \mu_{jsc} + \tau_i + \varepsilon_{ijsct}$$

where *Y* denotes national academic percentile rank in subject *s* at time *t* and X is a vector of observable non-time varying characteristics of the student. Here β represents the permanent impact of these non-time varying observable characteristics on academic achievement. In this specification there are two time periods, in period one students attend primary school and in the second period students attend secondary schools. The error term v_{ijsct} 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 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, this would include any stable parental inputs or natural ability of the child; ε_{ijsct} is the idiosyncratic time specific error which includes secondary school effects. Under this restrictive specification only ε_{ijsct} could cause the national academic rank of a student to change between primary and secondary school, as all other factors are permanent and have the same impact over time.

This is a restrictive assumption, as the impact of observable and unobservable characteristics are likely to change as the student ages. One could imagine that neighbourhood effects may grow in importance as the child grows older, and that the effects of primary school are more important when the child is young and attending that school. Therefore we extend the model by allow for time-varying effects of these characteristics:

$$Y_{ijsct} = \beta_{Rank\,t} R_{ijsc} + X_i{}'\beta + X_i{}'\beta_t + v_{ijsc} \tag{2}$$

$$v_{ijsc} = \mu_{jsc} + \mu_{jsct} + \tau_i + \tau_{it} + \varepsilon_{ijsc}$$

where β_t allows for the effect of student characteristics to vary over time. We have also introduced the parameter of interest $\beta_{Rank\,t}$, which is the effect of having rank R_{ijsc}, in subject s in cohort c and in primary school j on student achievement in that subject in the subsequent period *t*. We are interested in longer-run effects of rank positions students had during early education stages. We therefore assume that there is no effect of rank in the first period *t=1* as there is no prior rank $\beta_{Rank\,1} = 0$. We will hence be estimating $\beta_{Rank\,2}$, the effect of primary school rank on period 2 outcomes. To simplify the notation the time subscript will be dropped, as only one rank parameter is estimated, β_{Rank}.

This specification also allows for the unobservables to have time varying effects. Again τ_i represents unobserved individual effects that capture all time constant effects of a student over time and μ_{jsc} represents the permanent effects of being taught in a specific school-subject-cohort. Now additionally we have τ_{it} and μ_{jsct} allowing for these error components to vary over time so that students can have individual-specific growth rates as they grow older, or that primary school teachers can affect the efficiency of their students to learn a certain subject in the future.

Given this structure we now state explicitly the conditional impendence assumption that needs to be satisfied for estimating an unbiased rank parameter. Conditional on student characteristics, time varying and permanent primary school-subject-cohort level and individual effects, we assume there would be no expected differences in students' outcomes except those driven by rank.

$$Y_{Ri} \perp R_i |\, X_i, \mu_{jsc}, \mu_{jsct}, \tau_i, \tau_{it} \text{ for all R} \tag{3}$$

To achieve this we require measures of all these factors that may be correlated with rank and final outcomes. Conditioning on prior test scores will absorb all non-time varying effects as they will effect period-1 test scores to the same extent as period-2 test scores. Any input, observable or unobservable, that would affect academic attainment is captured in these test scores.[8] Therefore we can express period two outcomes, age 14 test scores, as a function of rank, prior test scores, student characteristics and unobservable effects.

$$\begin{aligned} Y_{ijksc2} = {} & \beta_{Rank} R_{ijsc} + f\left(Y_{ijsc1}\left(X_i'\beta, X_i'\beta_1, \tau_i, \mu_{jsc}, \tau_{i1}, \mu_{jsc1}\right)\right) + \\ & + X_i'\beta_2 + \mu_{jsc2} + \tau_{i2} + \epsilon_{ijksc} \end{aligned} \tag{4}$$

Using lagged test scores means that the remaining factors are those that affect the learning in period 2, between ages 11 and 14 $(X_i'\beta_t, \mu_{jsc2}, \tau_{i2})$. In our regressions, we will allow the functional form of this lagged dependent variable to take two forms, either a 3rd degree polynomial or a fully flexible measure, which allows for a different effect at each national test score percentile. As we can observe certain characteristics and primary school attended, β_2 and μ_{jsc2} can easily be estimated. The interpretation of μ_{jsc2} is that some primary schools are more effective at teaching for a later test than others, in a way that does not show up in the end-of-primary age-11 test scores.

[8] Examples of these effects include students' innate ability, parental investment, teacher effects, peer effects and primary school infrastructure

The discussion of recovering τ_{i2}, the second period academic growth of individual *i* is below, but 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. Any estimation, which allows for individual growth rates during secondary school (second period), would absorb any spillover effects. Therefore, leaving τ_{i2} in the residual means that the rank parameter is the effect of rank of the subject in question and the correlation in rank from the other two subjects, as we have test scores for English, mathematics and science.

In the second period the student will be attending secondary school k which may affect later test scores by subject, π_{ksc}, which is a component of the error term ϵ_{ijksc}, where $\epsilon_{ijksc} = \pi_{ksc} + \varepsilon_{ijksc}$. As stated above conditional on time-varying student effects, prior subject test scores and the other stated factors, we do not expect that these components will be correlated with primary rank. This is primarily because general secondary school effects are absorbed by the time varying student effect but we will return to the issue of secondary school choice in subsection 3.3.1.

The first two specifications that we estimate that will recover the effect of rank due to overall changes in effort which allow for spill-overs between subjects, are the following:

$$Y_{ijksc2} = \beta_{Rank} R_{ijsc} + f\left(Y_{ijsc1}\right) + X_i{}' \beta_2 + \mu_{jsc2} + \epsilon_{ijksc} \tag{5}$$

$$\text{Where } \epsilon_{ijksc} = \tau_{i2} + \pi_{ksc} + \varepsilon_{ijksc}$$

$$Y_{ijksc2} = \beta_{Rank} R_{ijsc} + f\left(Y_{ijsc1}\right) + X_i{}' \beta_2 + \mu_{jsc2} + \pi_{ksc} + \psi_{ijksc} \tag{6}$$

$$\text{Where } \psi_{ijksc} = \tau_{i2} + \varepsilon_{ijksc}$$

Note that we can further augment these regressions by using the average student growth rate across subjects to recover individual growth effects, τ_{i2}. Note that despite using panel data, this is estimating the individual effect across subjects and not over time. Lavy (*et al.* 2012) also use a student-fixed effects strategy to estimate ability peer effects. Applied to this setting, when allowing for student effects, we effectively compare relative rankings within an individual, controlling for national subject-specific ability. The variation used to identify rank is correlation between the differential growth rates by subject within each student and prior subject ranking. Therefore any individual characteristic that is not realised in age-11 test scores but contributes towards age-14 test scores is accounted for, including secondary school attended, as long as the effects are not subject specific.

$$Y_{ijksc2} = \alpha_t + \beta_{Rank} R_{ijsc} + f\left(Y_{ijsc1}\right) + \tau_{i2} + \mu_{jsc2} + \epsilon_{ijksc} \tag{7}$$

$$\text{Where } \epsilon_{ijksc} = \pi_{ksc} + \varepsilon_{ijksc}$$

In these specifications the rank parameter only represents the increase in test scores due to subject specific rank, as any general gains across all subjects would be absorbed by the student effect. This can be interpreted as the extent of specialisation in subject s due to primary school rank. It is for this reason, and the removal of other covarying factors, why we would expect the coefficient of the rank effect in specification (7) to be smaller than those found in (5) or (6).

Finally, to also investigate potential non-linearites in the effect of ordinal rank on later outcomes, i.e. are effects driven by students being top or bottom of the class, we replace the linear ranking parameter with indicator variables according to quantiles in rank plus additional indicator variables for those at the top and bottom of each school-subject-cohort (the rank measure is defined in section 4.2). We allow for non-linear effects according to vingtiles in rank, which can be applied to all the specifications presented. [9]

$$Y_{ijksc2} = \beta_{R=0}Bottom_{ijsk} + \sum_{n=1}^{20} I_n R_{ijsc}\,\beta_{n,Rank} + \beta_{R=1}Top_{ijsc} + f(Y_{ijsc1}) + \tau_i + \mu_{jsc} + \epsilon_{ijksc} \quad (8)$$

In summary, if students react to ordinal information as well as cardinal informaiton, then we would expect the rank in adition to relative achivement to have a signficant effect on later achievement when estimating these equations. This is what is picked up by the β_{Rank}-coefficient. The following sections discuss potential threats to identification, the setting, and how rank is measured before we turn to the estimates.

3.3 Threats to identification

3.3.1 Secondary school selection

A concern may remain that students could select secondary schools based on their rank in a particular subject during primary school in addition to their age-11 test scores. If, for example, students who were top of their class in maths aspire to attend to a secondary school that specialises in maths, our estimates could be confounded by secondary school quality. This might seem unlikely because we know that ability sorting for secondary schools in England is largely based on average rather than subject-specific abilities (Lavy *et al.* 2012).

Fortunately, our data allows us to address this concern directly by additionally controlling for secondary school attended. The resulting specification additionally allows for period-2 achievement to vary by secondary school k in subject s of cohort c, π_{ksc2}.[10] Intuitively, this is comparing students who are exposed to the same secondary school influences, thus identifying effects net of any potential subject-rank driven sorting into secondary education. However, secondary school attended can be argued to be an outcome, and therefore should not be conditioned upon. Specifications that include these effects are not our preferred model and should

[9] Estimates are robust to using deciles in rank rather than vigntiles and can be obtained upon request.
[10] We use the Stata command `reg2hdfe` for these estimations (Guimaraes and Portugal, 2010).

only be used as an indication of the extent that secondary school selection has effects on the estimates. As we will see, this modification does not affect our results.

3.3.2 Unobserved individual factors and parental background

Even with this flexible specification, one may still not be convinced that we are identifying the effect of rank on subsequent educational attainment. The rank of a student in primary school may be correlated with other unobserved individual factors that affect students' outcomes. An example of this could be unobserved individual or parental aspirations that correlate with primary rank and later value added, i.e. a competitive child or 'pushy parents'. Furthermore, using across-school variation might be problematic if schools transformed a student's ability into test scores non-monotonically.[11] We believe that the student fixed effects approach outlines by specification (7) addresses most of these concerns as all unobserved factors that affect test scores in all subjects in a similar way are now controlled for.

Notice that while these general factors such as 'pushy parents' that could induce correlations between primary rank and later outcomes are now controlled for, the remaining assumption required for identification is that such unobserved factors are not subject-specific. We return to this issue in Section 6.4 where we show that parental occupations predict primary-subject test scores yet are orthogonal to primary-subject rankings.

3.3.3 Ability peer effects and measurement error

Notice that all of our estimation specifications include primary-subject-cohort effects, which is also necessary to account for potential measurement error in the age-11 test scores arising through unobserved classroom-level shocks. In particular, if there are unobserved primary-school factors, these will create noise in the test score but not in the rank, as the ranking itself is mean-independent. As a result, the ranking variable could start to pick up ability-related information that could not be fully controlled for using only the national percentile test score. Including primary-school effects clears this kind of measurement error from the primary rank variable. We will return to issues in Section 6, where we also examine in detail how these rank effects interplay with ability peer effects, as well as potential threats placed by various kinds of measurement errors. We will conclude that whilst the most obvious candidates, i.e. classroom-level shocks and ability peer effects are controlled for directly in our setting, that higher order issues of measurement error and transitory non-linear peer effects do not invalidate our approach.

[11] If some schools are better at teaching low (high) ability students, then the ranking technology for ability may be different across schools.

4 Institutional setting, data and descriptive statistics

4.1 The English School System

The compulsory education system of England is made up of four Key Stages (KS); at the end of each stage students are evaluated in national exams. Key Stage Two (KS2) is taught during primary school between the ages of 7 and 11. The median size of a primary school cohort and the average primary school class size is 27 students (DFE, 2011). Therefore, when referring to primary school rank, one could consider this as class rank.[12] At the end of the final year of primary school when the students are aged 11 (Year 6), they take KS2 tests in English, math and science. These tests are externally graded on a national scale of between 0-100. This makes it possible to make comparisons in student achievement over time and across schools.

Rather than receiving these raw scores, students are instead given 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, which results in them being a coarse measure, with 85% of students achieve Level 4 or 5. These are non-high stakes exams for students and are mainly used by the government as a measure of school effectiveness.[13] This means that students do not know their underlying exact test score, which we can use to calculate their local ranks. Rather, students infer their rank position in class through repeated interaction, teacher feedback, and often through seating arrangements that reflect ability.

Students then transfer to secondary school, where they start working towards the third Key Stage (KS3). During this transition the average primary school sends students to six different secondary schools and the larger secondary schools receive students from 16 different primary schools. Importantly, admission into secondary schools is generally non-selective and does not depend on end-of-primary KS2 test scores. A subset of schools can select on ability (grammar schools) but these schools administer their own admission tests. The KS2 is thus a low-stakes test with respect to secondary school choice. In the new school, a typical student has 87% new peers upon arrival. This large re-mixing of peers is beneficial, as it allows us to estimate the impact of rank form a previous peer group on subsequent outcomes. Key Stage 3 takes place over three years and at the end of Year 9, all students take KS3 examinations in English, math and science at age fourteen. Again KS3 is not a high-stakes test and is externally marked.

[12] The maximum class size at Key Stage 1 is 30 students. A parallel set of results has been estimated using only cohorts of 30 and below, assuming these are single class cohorts. The results are qualitatively the same and are available from the authors upon request.

[13] The students also appear not to gain academically just from achieving a higher level. Regression discontinuity techniques show no gain for those students who just achieved a higher level. This setting is ideal for a regression discontinuity techniques as the score needed to reach a level changes by year and by subject, which would make it particularly hard to game.

Two years later, students take the national Key Stage 4 test at age 16 (KS4), which marks the end of compulsory education in England. The KS4 is graded from one to eight and students have some discretion in choosing the subjects they study and at what level. Since KS3 is graded on a fine scale [0-100], and students are tested in the same compulsory subjects only, we prefer this as the outcome measure for the purpose of our study. However, our results also hold using KS4 test scores[14].

4.2 Data Construction

4.2.1 Administrative data

The Department for Education (DfE) collects data on all students and all schools in state education in England. The Pupil Level Annual School Census (PLASC) collects student information such as gender, ethnicity, language skills, Special Educational Needs (SEN), or being Free School Meals Eligible (FSME). The number of students and student characteristics are used to determine school funding. The National Pupil Database (NPD) contains student attainment data throughout their Key Stage progression in each of the three compulsory subjects. Each student is given a unique identifier so that they can be linked to schools and followed over time, allowing the government to produce value added measures and publish school league tables. As the functions of both of these datasets are at the school level, no class level data is collected.

We have combined these data to create a dataset following the entire population of five cohorts of English school children. This begins at the age of 10/11 (Year 6) in the final year of Primary School when students take their Key Stage 2 examinations through to age 13/14 (Year 9) when they take Key Stage 3 tests. KS2 examinations were taken in the academic years 2000/2001 to 2005/2006 and so it follows that the KS3 examinations took place in 2003/2004 to 2007/8. From 2009 students were no longer externally assessed, instead teacher assessment was used to evaluate students at the end of Key Stage 3, hence this is the end point of our analysis.

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 KS2 and KS3 exam information and background characteristics, 83% of the population. Secondly, we exclude students who appear to be double counted (1,060) and whose school identifiers do not match (12,900), approximately 0.6% of the remaining sample. 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

[14] Results can be obtained from the authors upon request. They are not presented here due to issues relating to the comparability of these test scores across students as they can be entered into different exams, along with the coarseness of the measures and students choosing to study additional optional subjects.

dimensions; this represents 2.8% of students[15]. 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.

Table 1: Descriptive statistics of the main sample

	Mean	S.D.	Min	Max
Panel A: Student Characteristics				
Age 11 test scores				
KS2 English	50.29	28.03	1	100
KS2 Math	50.52	28.19	1	100
KS2 Science	50.01	28.03	1	100
Within Student KS2 S.D.	12.68	7.70	0	57.16
Age 14 test scores				
KS3 English	51.23	28.18	1	100
KS3 Math	52.89	27.55	1	100
KS3 Science	52.91	27.53	1	100
Within Student KS3 S.D.	11.32	7.19	0	56.59
Panel B: Rank Characteristics				
Rank English	0.488	0.296	0	1
Rank Math	0.491	0.296	0	1
Rank Science	0.485	0.295	0	1
Within Student Rank S.D.	0.138	0.087	0	0.58
Panel C: Background Characteristics				
SEN	0.175	0.380	0	1
FSME	0.146	0.353	0	1
Male	0.499	0.500	0	1
Ethnicity				
White British	0.837	0.370	0	1
Other White	0.019	0.135	0	1
Asian	0.058	0.234	0	1
Black	0.030	0.171	0	1
Chinese	0.003	0.053	0	1
Mixed	0.017	0.128	0	1
Other	0.011	0.104	0	1
Unknown	0.026	0.158	0	1

Notes: 6,815,997 observations over 5 cohorts. Cohort 1 takes Key Stage 2 (KS2) examinations in 2001 and Key Stage 3 (KS3) examinations in 2004. Cohort 5 takes KS2 in 2005 and KS3 in 2008. Test scores are percentalized tests scores by cohort-subject. All test scores come from national exams which are externally marked. The analysis stops in 2008 as after this point Key Stage 3 exams became internally assessed.

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

As described in Section 4, the Key Stage test scores for both levels are percentalized by subject and cohort, so that each individual has six test scores between 0 and 100 (three KS2 and three KS3). This ensures that students of the same nationally relative ability have the same national percentile rank, as a given test score could represent a different ability in different years or subjects. Importantly, this allows for test score comparisons to be made across subjects and across time, this does not impinge on our estimation strategy, which relies only on heterogeneous test score distributions across schools to generate variation in local rank. [16]

We rank students in each subject according to their age 11 national test scores within their primary school by cohort. To have a comparable local rank measurement across schools of different cohort size we transform the rank position of individual *i* with the following normalization:

$$R_{ijsc} = \frac{n_{ijsc}-1}{N_{jsc}-1}, \qquad R_{ijsc} = [0,1] \tag{9}$$

where N_{jsc} is the cohort size of school *j* in cohort *c* of subject *s*. An individual's *i* ordinal rank position within this set is n_{ijsc},, which is increasing in test score. R_{ijsc} is the standardised rank of the student. [17] For example, a student who had the second best score from a cohort of twenty-one students (n_{ijsc}=20, N_{jsc}=21) will have R_{ijsc}=0.95. This rank measure will be approximately uniformly distributed, and bounded between 0 and 1, with the lowest rank student in each school cohort having R=0. In the case of draws of national percentile rank, each of the students is given the lower local rank.

Rank is dependent on students own test scores, but is determined by the scores of others in their school. Again consider the students who scored X and Y in cohorts with different test score distributions from Figure 1. The students who scored Y, being the same distance above the mean in both school cohorts would have a rank of R_{yA}= 0.9 in Cohort A (unimodal distribution) and R_{yB}= 0.6 in Cohort B (bimodal distribution). Similarly students who scored X would have a rank of R_{xA}= 0.1 in Cohort A and R_{xB}= 0.4 in Cohort B. It is the different distribution of peer test scores that allows for the separate identification of the rank effect from a relative ability effect. As there is information for three subjects for each student, a student can have a different rank for

[16] Estimations using standardised rather than percentalized test scores provide similar estimates to the first decimal place in linear specification. For non-linear specifications the effect of rank appears more cubic in nature. However, these estimations suffer from non-comparability as a given test score could represent a different ability in different years or subjects. Year/subject effects would not account for all these differences as there are likely to be distributional differences. Allowing for either functional form of test scores to be interacted by year and subject was extremely computationally intensive, given our already demanding specification. Basic results are available from the authors upon request.

[17] This is rank within school subject cohort, it cannot be done by class as no class level information is available. However, all estimations have been replicated on schools which have cohort sizes of under 30 (maximum class size) and have equivalent results. Obtainable upon request.

each subject within her primary school. This feature of the data allows us to include student fixed effects in some of our regressions.

Note that since the students do not receive their detailed test scores, they will not be able to derive this same rank score themselves, nor will they be given an official ranking. Instead, our measure of local rank is a proxy for the students' experiences over the past six years of interacting with their peers in the classroom. The existence of any effect is driven through student beliefs about their rank position within their class.

4.2.2 Survey data

Additional information about a subsample 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, out of school activities and attitudes.

We merge survey responses with our administrative data using a unique student identifier. This results in a dataset where we can track students from a primary school, determine their academic ranks and then observe their later measurements of confidence and attainment, allowing us to test if rank effects confidence conditional on attainment. This is the first research to merge LSYPE responses to the NPD for primary school information.

At age 14 the students are asked how good they consider themselves to be in English, maths and science, with 5 possible responses that we code 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 self-concept. Whilst it is much more basic than surveys that focus on self-concept, it does capture the essence of the concept.

The matching between the NPD and LSYPE was perfect. However, the LSYPE also surveys students attending private schools that are not included in the national datasets; moreover, as students not accurately tracked over time have been removed, a further 3,731 survey responses could not match. Finally, 1,017 state school students did not fully complete these questions and so could not be used for the self-concept analysis. Our final dataset contains 11,898 student observations with self-concept measures. Even though the survey will not contain the attitude measures of every student in a school cohort, by matching the main data we will know where that student was ranked. This means we will be able to determine the effect of rank on self-concept, conditional on age 11 test scores and school-subject-cohort fixed effects.

4.3 Descriptive statistics

4.3.1 Main sample

The data has the complete coverage of the state student population from age 10 to 14. We follow each student from their primary school through to secondary school, linking their rank in their school to their later outcomes. so that they can be linked to schools and followed over time, allowing the government to produce value added measures and publish school league tables. As the functions of both of these datasets are at the school level, no class level data is collected.

We have combined these data to create a dataset following the entire population of five cohorts of English school children. This begins at the age of 10/11 (Year 6) in the final year of Primary School when students take their Key Stage 2 examinations through to age 13/14 (Year 9) when they take Key Stage 3 tests. KS2 examinations were taken in the academic years 2000/2001 to 2005/2006 and so it follows that the KS3 examinations took place in 2003/2004 to 2007/8. From 2009 students were no longer externally assessed, instead teacher assessment was used to evaluate students at the end of Key Stage 3, hence this is the end point of our analysis.

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 KS2 and KS3 exam information and background characteristics, 83% of the population. Secondly, we exclude students who appear to be double counted (1,060) and whose school identifiers do not match (12,900), approximately 0.6% of the remaining sample. 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% of students. 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.

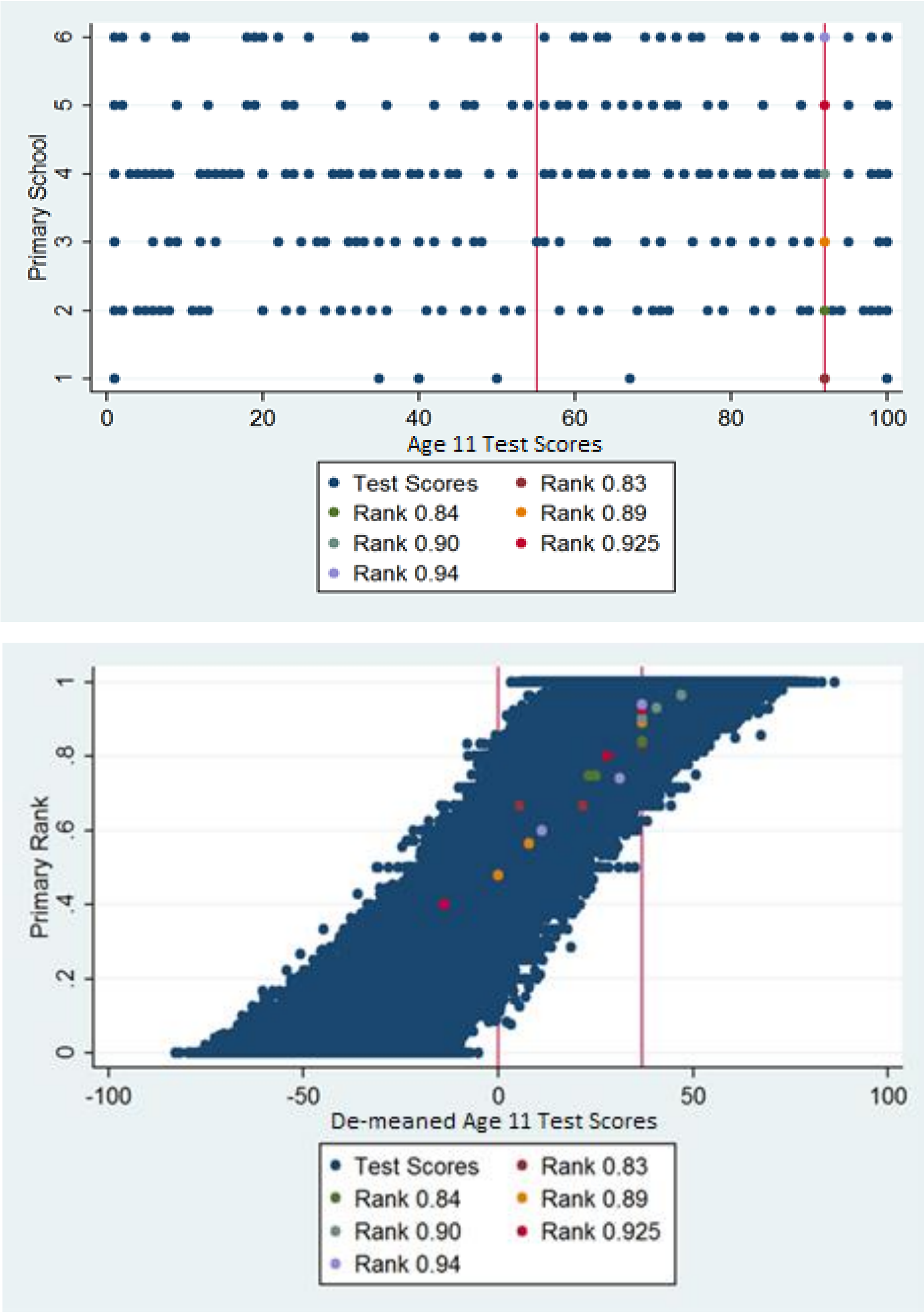
Table 1 shows summary statistics for all students that are used the analysis. Given that the test scores are represented in percentiles, all three subjects test scores at age 11 and 14 have a mean of 50 with a standard deviation of 28.

The within-student standard deviation across the three subjects English, math 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 effects regressions.

Information relating to the background characteristics of the students is shown in the lower panel of Table 1 half the student population is male, over four-fifths are white British and about 15 per cent are Free School Meal Eligible (FSME) a standard measure of low parental income.

We use this variation of test scores across schools to identify the effect of rank separately from relative ability. This was previously illustrated for a theoretical case in Figure 1, which shows the rank of an individual is dependent on the distribution of test scores even when maximum, minimum and mean test scores are the same in both schools. In the top panel of Figure 2 we replicate this with actual student test score data from six primary schools. Each point represents a student's age 11 English test score. Each row represents a school which has a student ranked in the 1st and 100th national percentiles, has a mean percentile of 54 and a student in the 93rd percentile in English. This is a very specific case, but in each the student at the 93rd percentile has a different rank. For the estimations, we use all subjects and the distributions of test scores across all primary schools whilst accounting for mean school-subject-cohort test scores. Therefore the lower panel of Figure 2 plots the rank of every student in each subject by de-meaned test score. The vertical thickness of the distribution of points indicates the support at throughout the rank distribution. For the mean students there is nearly full rank support.

Figure 2: Rank distributions in schools and across subjects



Notes: In the upper panel each point represents a student's Key Stage 2 test score. The six schools that are represented have the same mean (54), minimum (0) and maximum (100) tests scores in English, and also have a student with a test score of 93. Each student with the test score of 93 has a different rank. The lower panel shows all students in our data. The Y-axis is the primary rank of students and the X-axis shows the de-meaned test scores by primary school-subject-cohort. The colored points represent the three different test scores and ranks of students from Figure 5 with a test score of 93 in English. Note that the number of students per school as well as individual test scores have been randomly altered enough to ensure anonymity of individuals and schools. They are for illustrative purposes only and in no way affects the interpretation of these figures.

That there are differences in test score distributions across schools will be the result of many factors. One example is that a school in a rural area where there is little school choice may have a wider test score distribution than a school in a city where there is more parental sorting. However, conditional on school-subject-cohort and student effects, we are confident that these factors would no bias our results.

4.3.2 Longitudinal Study of Young People in England

We use the LSYPE sample to estimate rank effects on a direct measure of self-concept. The LSYPE respondents are representative of the main sample, although mean age 11 test scores are slightly lower and the proportion of Free School Meal Eligible is higher than the national at 18.6% and 14.6% respectively (Online Appendix Table 3).

The LSYPE students are asked to rate themselves in each of the subjects from 'Not good at all' to 'Very Good,' which is summarized in Online Appendix Table 5. Our measure of self-concept is coarse, with only five categories to choose from and around 60% choosing "fairly good". We can see that students do think about their own ability, with less than 0.2% not having an opinion. As would be expected, those who considered themselves to be poor performers did tend to have lower average national KS2 percentile rank and lower rank within their school. However, there is also large variance in these ranks within these self-evaluated categories. For every subject, each self-assessment category with an opinion has at least one individual in the top 9% nationally, including those who considered themselves 'Not Good'. Similarly, each category has an individual in the lowest performing percentile nationally, even those who consider themselves very good.[18]

5 Main Results

5.1 Effect of Rank: comparing across school cohorts

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 Panel A of Table 2, with the specifications becoming increasingly flexible moving across columns to the right. The first row shows estimates of the rank parameter using fully flexible set of controls for age 11 test scores, allowing each percentile score to have a different effect on later test scores. Due to computational constraints we

[18] In Online Appendix Table 5 we also show the performance of the top and the bottom 10% of students within each self-assessment category that are less affected by outliers. We continue to see very large variance within categories. Consider Science in Panel C: of those who consider themselves 'Very Good' the bottom 10% performers in this category are ranked at the 17 percentile point nationally, whereas the top 10% of performers in the category that rated themselves 'Not very good at all' ranked at 64th percentile nationally.

are unable to run all specifications using this functional form and therefore the second row instead uses a third order polynomial of age 11 test scores. It appears that this is sufficient approximation to account for the effect of age 11 test scores. All estimates control for a set of student characteristics and have standard errors clustered at the secondary school level[19].

The first column is a basic specification, which only controls for age 11 test scores, student characteristics, along with cohort and subject fixed effects. This shows a comparatively large estimate: a student at the top of their cohort has an 11.6 larger national percentile rank gain in test scores compared to a student ranked at the bottom, *ceteris paribus*. However, this regression does not condition on school-subject-cohort effects and therefore the parameter cannot be interpreted as pure rank effect as it will also capture the effects of relative ability. Furthermore, it uses variation in average quality of students across schools, which might correlate to family background characteristics, later school quality, or other unobserved variables.

Indeed, this is what we find in column (2) which is significantly smaller and additionally allows for any primary school-subject-cohort effects (Specification 5). Using this specification, the effect of being ranked top compared to bottom *ceteris paribus* is associated with a gain 7.96 national percentile ranks (0.29 standard deviations) conditional on a cubic of age 11 test scores. This can be interpreted as the additional ordinal rank effect. Given the distribution of test scores across schools, very few students would be bottom ranked at one school and top at another school. 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 0.085 standard deviations or 2.36 national percentile points. Note that any determinant that has a permanent effect on student outcomes would be absorbed by prior test scores, this is the growth in national percentiles between the ages of 11 and 14 due to primary school rank.[20] In comparison with other student characteristics, females' growth rate is 1.01 national percentile points higher than males and free school meal eligible students on average lose 2.96 national percentile points (Online Appendix Table 6).[21]

We see that when additionally allowing for secondary school-subject-cohort effects (Specification 6) there is only a marginal impact on the estimates and are not significantly different from those in column 2. This is evidence that that there is negligible sorting into secondary schools by subject rank, conditional on student test scores.

[19] Student characteristics are ethnicity, gender, ever Free School Meal Eligible (FSME) and Special Educational Needs (SEN)

[20] Using teacher assessment data on student ability to rank students near start of primary school, at age 7. We find that students who are consistently top of their primary school do additionally better in age 14 test. Results available upon request, not presented as main result due to coarseness and reliability of age 7 test scores.

[21] Including the rank parameter in this specification reduces the Mean Square Error by 0.31. This is more than the reduction from allowing for a gender growth term (0.25) or an ethnicity growth term (0.28).

Table 2: Age 14 Test Scores on Primary School Rank (Age 11)

	Raw (1)	Primary (2)	Primary-Secondary (3)	Primary-Student (4)
Panel A: The effect of primary rank				
Primary Rank Flexible Age 11 Test Scores	11.551** *0.293*	7.662** *0.145*		
Primary Rank Cubic Age 11 Test Scores	11.001** *0.298*	7.960** *0.145*	7.901** *0.146*	4.562** *0.132*
Panel B: The effect of placebo rank				
Placebo Rank Flexible Age 11 Test Scores	0.0055 *0.100*	0.015 *0.011*		
Placebo Rank Cubic Age 11 Test Scores	0.0045 *0.100*	0.013 *0.116*	0.016 *0.119*	-0.008 *0.137*
Student characteristics	✓	✓	✓	Abs
Age 11 Test Scores	✓	✓	✓	✓
Primary-cohort-subject Effects		✓	✓	✓
Secondary Effects			Abs	Abs
Secondary-cohort-subject Effects			✓	
Student Effects				✓

Notes: Results obtained from twelve separate regressions based on 2,271,999 student observations and 6,815,997 student-subject observations. The dependent variable is by cohort by subject percentalized KS3 test scores. All specifications control for Key Stage 2 results, student characteristics, cohort effects and subject effects. Student characteristics are ethnicity, gender, free school meal (FSME) and special educational needs (SEN). Coefficients in columns (2) and (3) are estimated using Stata command reg2hdfe allowing two high dimensional fixed effects to be absorbed. Standard errors in italics and clustered at 3,800 secondary schools. Abs indicates that the effect is absorbed by another estimated effect. ** 1% sig.

5.2 Effect of Rank: within student analysis

We now turn to estimates that use the within student variation to estimate the rank effect (Specification 7). Conditioning on student effects allows for individual growth rates, which absorb any student level characteristic. Since students attend the same primary and secondary school for all subjects, any general school quality or school sorting is also accounted for. Subject specific primary school quality is absorbed by the primary school-subject-cohort effects. This uses the variation in the relative growth rates across subjects within student according to differing rank in primary school.

Besides removing potential biases, the inclusion of student effects changes the interpretation of the rank parameter. The student effect will also absorb any spillover effects gained through high ranks in other subjects and is only identifying the relative gains in that subject. Accordingly the within student estimate is considerably smaller. The effect from moving to the bottom to top of class *ceteris paribus* increases national percentile rank by 4.56 percentiles, as we see in Panel A, column (4) of Table 2. To make a comparison in terms of standard deviations this effect is scaled by the within student standard deviation of national percentile rank (11.32). Therefore, conditional on student and school-subject-cohort effects, the maximum effect of rank is 0.40 standard deviations. This is a very large effect, but a change from last to best rank *within student* represents an extreme treatment. It is more conceivable for a student to move 0.5 rank points, e.g. being at the 25th percentile in one subject and 75th at another. Our estimates imply that this student would improve their test scores in that subject by 0.20 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]

Again, if there were any general gains through achieving a high rank in one subject, this would be absorbed in the within student estimates, and thus could be interpreted as the between subject substitutions of effort allocation, or a lower bound of the effect of being highly ranked. The difference between the within school estimates (7.96) and the within student estimates (4.56) can be interpreted as an upper bound of the gains from spilovers between subjects.

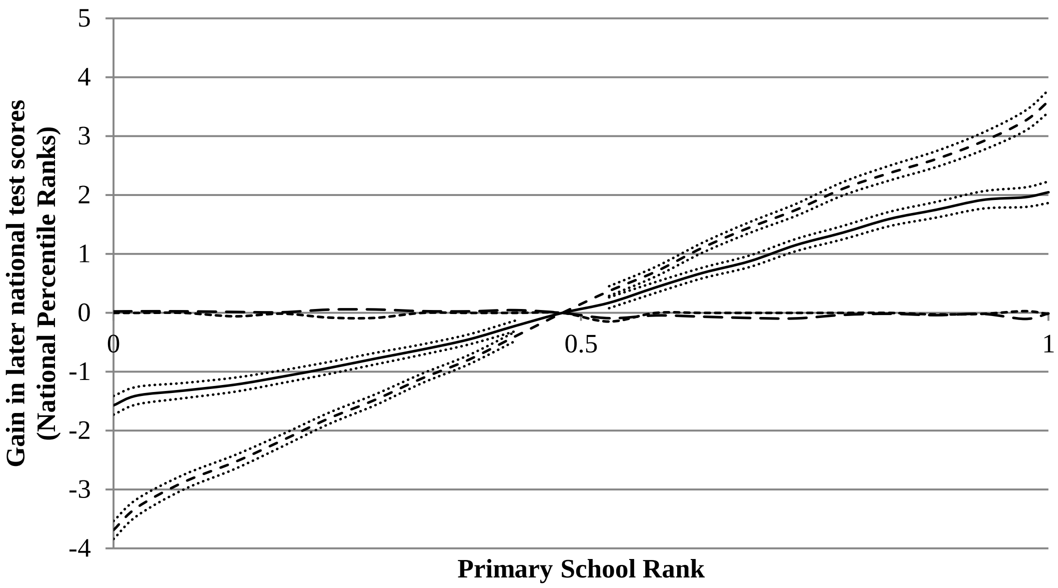
5.3 Non-linear Effects

The specifications thus far assumed the effect of rank is linear, however, it is conceivable that the effects of rank change throughout the rank distribution (Brown, 2011). To address this we allow for non-linear effects of rank by replacing the rank parameter with a series of 20 indicator variables for to the vingtiles in rank, plus top and bottom of class dummies (Specification 8).

The equivalent estimates from specification (5) and (7), i.e. without and with student fixed effects, are presented in Figure 3. The effect of rank appears to be almost linear throughout the rank distribution, with small flicks in the tails. Reassuringly, the placebo ranks (to be discussed in Section 6.3) are also insignificant when allowing for non-linear effects. In comparison, all rank coefficients are significantly different from the reference group of the median-ranked students (10th vingtile). This indicates that the effect of rank exists throughout; even those students ranked just above the median perform better three years later than those at the median. Given that students are not informed of rank, our interpretation of this is that students are good at ranking themselves within the classroom. This ranking developed through the constant exposure to peers over the length of primary school, which continually reinforces the effect on self-concept such that by the end of primary school they have strong beliefs about where they rank. Finally, the fact that the rank effect exists throughout the distribution is in line with the idea that self-concept forms according to rank position.

[22] Students with similar ranks across subjects the choice of specialization would be less clear. Indeed, the impact of rank is 25% smaller form students in bottom quartile than those from the top quartile.

Figure 3: Effect of Primary School rank on Secondary School outcomes



- - - - Rank conditional on Primary-Subject-Cohort Effects

——— Rank conditional School-Subject-Cohort and Pupil Effects

------- Placebo Rank conditional on Primary-Subject-Cohort Effects

— — — Placebo Rank conditional on Primary-Subject Cohort & Pupil Effects

Notes: Non-linear effect with dummies for the vingtiles of rank plus a dummy for being top or bottom of school-subject-cohort. All specifications have subject specific rank and test score across three subjects. Placebo rank generated from actual test scores but randomly allocated peers, using the actual distribution of primary school size. All standard errors clustered at the actual secondary school attended. Specification 1: Student characteristics and primary, subject and cohort effects. Specification 2: Primary-subject-cohort group effects and student effects. Dashed lines represent 95% confidence intervals.

5.4 Heterogeneity by gender and parental income

We now turn to how the effects of rank vary by student characteristics using the student fixed effects specification (7) with non-linear rank effects and interacting the rank variable with the dichotomous characteristic of interest.[23] The student characteristics are Male: Female and, FSME: Non-FSME. The baseline group coefficients and the interaction plus baseline coefficients are plotted to show the effect of rank on test scores for both groups, illustrating how the different groups react to primary school rank.[24]

[23] Interacting student characteristics rather than estimating the effects separately, ensures that students who attend the same school have the same relative. Use of interactions is preferred over separate regressions as the school-subject-cohort effects will be shared across groups and so relative test scores according to that school's mean will be the same for both.

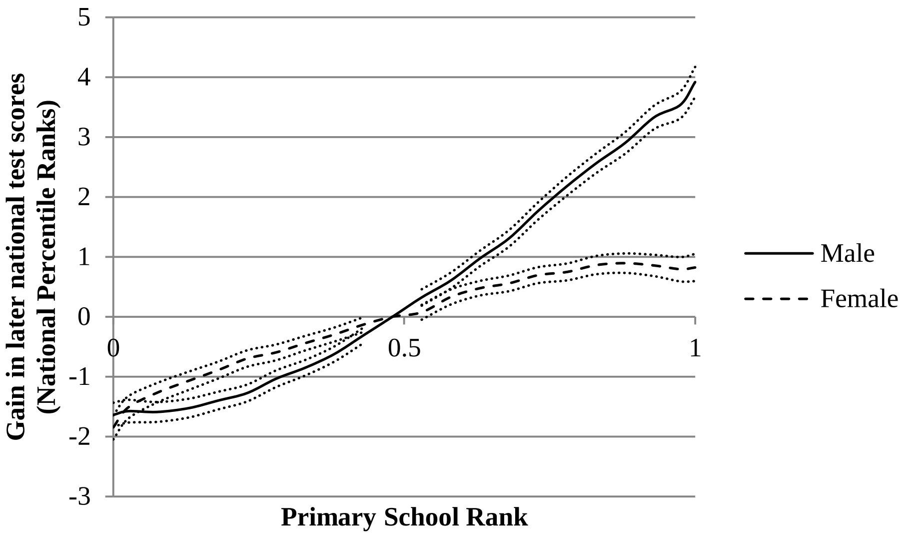
[24] The student characteristics themselves are not included in the estimations, as they are absorbed by the student effects. These characteristics interacted by rank, however, are not absorbed by student effects, because there is variation within the student due to having different ranks in each subject.

The first panel in Figure 4 shows the how rank relates to the gains in later test scores by gender. Males are more affected by rank throughout 95% of the rank distribution, this is shown by the steeper gradient of the rank effect. 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. As this is within a student variation in later test scores, the coefficient could be interpreted as a specialising term, implying that prior rank has a stronger specialising effect on males than females.

The second panel in Figure 4 shows that Free School Meal Eligible (FSME) students are less negatively affected by rank and more positively affected than Non-FSME students. FSME students with a high rank gain more than Non-FSME students, especially those ranked top in class, who gain almost twice as much. FSME students who are below the median have limited negative effects on later test scores. This could be interpreted as these students already having a low self-concept for other reasons and therefore the negative effects of low rank have less of an effect. Moreover the shallower gradient for Non-FSME students could also be interpreted that they are less affected by class rank as these students may have their academic self-concept being more be affected by factors outside of school.

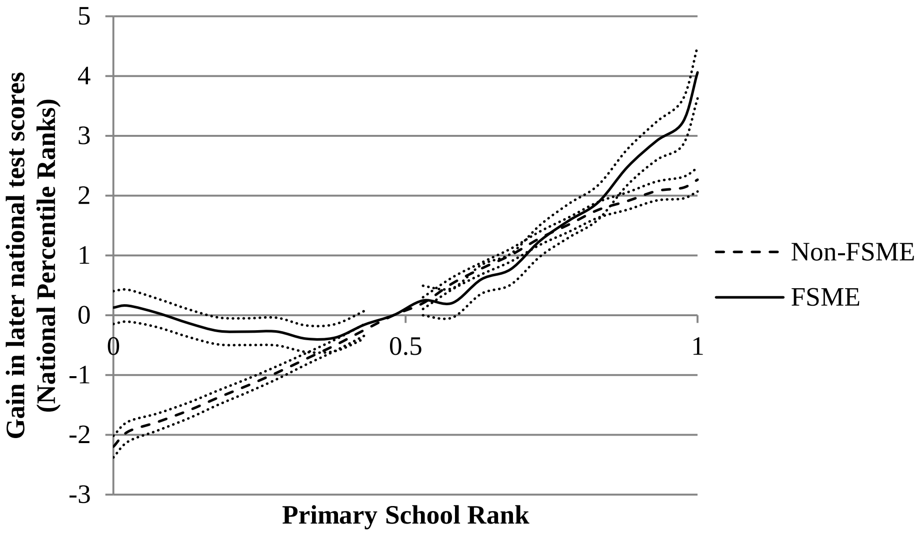
Figure 4: Effect of Primary School rank on Secondary School outcomes by student characteristics

Panel A: Student Gender



Notes: Effects obtained from estimating the effect of rank on Female students and the interaction term with Male students. Non-linear effect with dummies for the vingtiles of rank plus a dummy for being top or bottom of school-subject-cohort. All estimates use subject specific rank and test score across three subjects and condition on Primary-subject-cohort group effects and student effects. Dashed lines represent 95% confidence intervals.

Panel B: Student Free School Meal Eligibility Status



Notes: FSME stands for Free School Meal Eligible student. Effects obtained from estimating the effect of rank on Non-FSME students and the interaction term with FSME students. Non-linear effect with dummies for the vingtiles of rank plus a dummy for being top or bottom of school-subject-cohort. All estimates use subject specific rank and test score across three subjects and condition on Primary-subject-cohort group effects and student effects. Dashed lines represent 95% confidence intervals.

6 Robustness

Some non-trivial empirical challenges arise when estimating the effect of rank conditional on test score because we do not independently observe both a students' rank and a student's ability. Instead, we rely on externally marked and nationally standardised tests at the end of primary school to derive a student's local rank during primary education and also use this measure to control for a student's subject-specific achievement. This may cause problems relating to the influence of peers, parents and measurement error on test scores.

6.1 Peer Effects

Firstly, given that we are discussing an atypical peer effect, it is important to address the issues associated with such.[25] Any primary school peer effects that have a permanent effect on test scores do not bias the estimates as they are captured in the end-of-primary school test scores. Furthermore, we can account for contemporaneous secondary peer quality with the inclusion of secondary school-subject-cohort effects.[26]

[25] The standard reflection problem is not a first order issue in this situation, as students are surrounded by 87% new peers when they transfer to secondary school, and the rank effect is generated by primary school peers.

[26] This has almost zero effects on our coefficients partly because of the large re-mixing of students during the primary-to-secondary transition, and the sorting to school not being rank dependent.

However, if peer effects have a transitory effect on test scores i.e. only current peers matter, any estimation of the effect of primary rank on age 14 test scores whilst controlling for primary test scores could be biased. This is to the extent that both the conditioning variable and rank will be correlated with primary peer effects. The intuition for this is as follows: in the presence of transitory peer effects, a student with lower quality peers would attain a lower primary school results than otherwise and also have a higher rank than otherwise. Thus, when controlling for primary test scores in the estimations, those who previously had low quality peers would appear to gain more as they now have a new peer group, who on average would be better. Since rank is negatively correlated with peer quality in primary test scores, it would appear that those with high rank make the most gains. Therefore having a measure of ability confounded by transitory peer effects would lead to an upward biased rank coefficient.

This is shown to be the case in the Online Appendix 1, where we create a data generating process in which we specify that subsequent test scores are not effected by rank. Instead test scores are only a function of ability and individual linear or non-linear peer effects. To be cautious we allow for these peer effects to be 20 times larger than those found in Lavy et al. (2012). We simulate these data 1000 times and estimate the rank parameter with different sets of controls. This shows that not controlling for the primary school peer group generates biased results, but that this bias is negligible when allowing for mean school-subject-cohort effects, even with these large non-linear peer effects.

6.2 Measurement Error

In addition to peer effects, individual test scores may be imperfect measures of inputs up until that point in time. Given that both rank and test scores will be affected by the same measurement error, but to different extents due the heterogeneous test score distributions, calculating the size of the bias is intractable. To gauge the extent of measurement error we again simulate the data assuming 20% of the variation in test scores is random noise, 70% student ability and 10% school effects, these proportions reflect that 80% of the variance of test scores is within schools and 20% across schools (Online Appendix 2). This shows that normally distributed individual-specific measurement error would work against finding any effects.

The intuition for this is the following: a particular student having a large positive measurement error would result in both an inflated end-of-primary score and a higher local rank measure. Both of these would work against finding positive effect of rank on later outcomes, as we explicitly control for prior attainment. This student’s later test scores would hence be benchmarked against other students' with the same end of primary result but higher actual ability. Since the student only got a high local rank because of the measurement error, this would downward bias any positive rank effect estimate.

6.3 Is rank just picking up ability?

Our estimates of primary school subject-specific rank are relatively large, given that we are conditioning prior test scores and individual growth. As rank is highly correlated with student ability and test scores, there could be a concern that measurement error in the test scores for ability may be recovered in the rank measurement, if rank is measured with less error than test scores.

Note that this is different from the measurement error concern discussed above. To address the specific measurement error problem of rank having less measurement error than test scores and thus containing residual ability information, we perform placebo tests. This involves generating a placebo-rank measure that uses underlying ability, but would not reflect the social comparison experiences of students. To achieve this we re-assigned randomly all students into primary schools by cohort and re-calculated the ranks that they would have had in these schools with their original age-11 test scores but with peers that they never actually interacted with. These placebo-ranks are highly correlated with age-11 test scores. If they were found to be significant determinants of later achievement, this would indicate that rank is picking up ability not captured in end of primary school outcomes. We re-estimate all the specifications fifty times using new placebo-ranks each time and present the mean results in Panel B of Table 2, and the non-linear effects in Figure 3. We find no effects of these placebo ranks on later test scores. From these simulation results we conclude that our findings are unlikely to be mechanically driven by measurement error in test scores.

6.4 Are student effects enough? Primary school sorting and parental occupation

The causal interpretation that we give to estimates relies on the conditional independence assumption. That a student’s rank needs to be orthogonal to other subject-varying determinants of a student's later achievement. Given the student effects, the variation need not be orthogonal to general determinants of the student's achievement, but would need to vary within a student across subjects. A prime example of this could be the occupational background of the parents. Children of scientists may have a higher learning curve in science throughout their academic career for reasons of parental investment or inherited ability. Similarly children of journalists for English and children of accountants in math. This will not bias our results as long as conditional on age 11 test scores parental occupation is orthogonal to primary school rank. Or more broadly, there would be a problem if conditional on other factors, rank was correlated to subject-varying determinates of future achievement. This might well be the case if parents strongly aspire for their child to rank top in that subject and also have a higher academic growth rate in that subject between the ages 11 and 14.

Typically parents are trying to get their child into the 'best school' possible in terms of average grades. This would work against any positive sorting by rank as higher average achievement would decrease the probability of their child having a high rank. This sorting on general achievement would be accounted for by the student fixed effect. However, if parents wanted to maximise their child's rank in a particular subject, this could bias the results. In order to do this they would need to know the ability of their child and all potential peers by subject. This is unlikely to be the case, particularly for such young children who have yet to enter formal education at age 4. Parents could possibly infer the likely distributions of peer ability if there is autocorrelation of the student achievement within a primary school. This means that if parents did know the ability of their child by subject, and the achievement distributions of primary schools they could potentially select a school on this basis.

We test for this by using the LSYPE sample which has information on parental occupation. All parental occupations are classified into English, math, science, or 'other' and then an indicator variable is created for each student-subject if they have a parent who works in that field [27]. This is taken as an indicator for the parents' subject preference. We then regress age-11 test scores on parental occupation, school-subject effects and student effects (Table 3, Panel A). This establishes that this measure of parental occupation is a significant predictor of student subject achievement even when allowing for individual effects. Then using rank as the dependent variable we test for a violation of the orthogonality condition in Panel B of Table 3. Here we see that whilst parental occupation does predict student achievement by subject, it does not predict rank conditional on test scores. This implies that parents are not selecting schools on the basis of rank for their child. We therefore do not reject that the orthogonality condition does not hold with respect to parental background. This does not rule out other co-varying factors that may bias the results but it provides us with confidence that this likely large factor does not.

[27] Parental Standard Occupational Classification 2000 grouped in Science, Math, English and Other. **Science** (3.5%); 2.1 Science and technology, 2.2 Health Professionals, 2.3.2 Scientific researchers, 3.1 Science and Engineering Technicians. **Math** (3.1%); 2.4.2 Business And Statistical Professionals, 3.5.3 Business And Finance Associate Professionals. **English** (1.5%): 2.4.5.1 Librarians, 3.4.1 Artistic and Literary Occupations, 3.4.3 Media Associate Professionals. **Other**: Remaining responses.

Table 3: Balancing by parental occupation

	Primary (1)	Primary-Student (2)
Panel A: Effects on age-11 tests		
Parental Occupation	7.722**	1.706*
	0.840	*0.783*
Panel B: Effects on Ordinal Rank		
Parental Occupation	-0.004	0.000
	0.005	*0.034*
Primary-subject Effects	✓	✓
Student Effects		✓

Notes: Results obtained from regressions based on 31,050 subject-student observations for which parental occupations could be identified from the LSYPE. Detailed occupational coding available from the authors on request. Panel A has KS2 as dependent variable, in Panel B KS2 with polynomials up to cubic are included as controls. All regressions control for student characteristics and subject effects. Regressions in column (2) estimated using Stata command reg2hdfe. ** 1%,* 5% significant.

7 Mechanisms

A number of different mechanisms could produce similar results; competitiveness; environmental favours certain ranks; external (parental) investment by task; students learn about their ability. In the following, we discuss how each coincides with the results presented so far.

7.1 Hypothesis 1: Competitiveness

If the goal of individuals was to be better than their peers, maximise rank, this could produce some of our results, but not the full pattern.

To see this, consider two students of the same ability who attend the same secondary school but went to primary schools of different peer quality. The student attending the primary school of low quality peers could provide less effort in their end of primary school tests and still be ranked top. This student would then achieve lower end of primary school test scores than the student who faced competition in primary school. At secondary school when they have the same level of competition, and due to their same ability they will have the same expected age 14 test scores. In our estimation, controlling for prior test scores will make it appear that the student who faced lower competition and was ranked higher, had larger growth and thus generate the positive effect of rank.

However, if these mechanisms were driving the results, we would only expect to see these effects near the top of the rank distribution as it only applies to students who far exceed their peers and so get a lower than would be expected age-11 test scores. All those in the remainder of the distribution would be applying effort during primary school to gain a higher rank and so we should not see an effect. However given the result that the rank effect is approximately linear throughout it is unlikely that this type of competition mechanism is causing the effect.

It could still be the case that primary school subject rank is positively correlated with the degree of competitiveness of the student. Then those who are the most competitive increase their effort the most when entering secondary school and so have higher test score growth. Note that in the student effects specification any general competitiveness of an individual would be accounted for, this competiveness would need to vary by subject. As previously mentioned, any factor that varies by student across subjects conditional on prior test scores could confound –on in this case, explain- the results.

7.2 Hypothesis 2: The environment favours certain ranks

Another possible explanation for this finding is that the environment could favour the growth of certain ranks of agents. As an example, one can imagine primary school teachers teaching to the low ability students if faced with a heterogeneous class group[28]. If this were the case, teachers may design their classes with the needs of the lowest ranked students in mind. This means that these students would achieve higher age 11 test scores than they otherwise would have done and students further from the bottom lose out.

What would this mean for the rank effect estimates? Again consider two students of the same ability who attend to the same secondary school but different primary schools, where one was top of year. The top student would get less attention during primary school and therefore get a lower grade than they otherwise would have done. At secondary school they have the same attention due to their same ability and get the same age 14 test scores. In our estimation, controlling for prior test scores will make it appear that the top student had higher growth and thus generate the positive effect of rank. Therefore, teachers teaching to the bottom student could also generate a positive rank effect. This would require primary school teachers only being effective with lowest ranked students and secondary school teachers teaching to each ability level equally. Note if primary teachers taught to the median student, 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.

[28] We have run estimations controlling for the within school-subject-cohort variance to take into account that high variance classes may be more difficult to teach. However, these cannot include school-subject-cohort or student effects, and thus the estimates should not be cleanly interpreted as ordinal rank affects. Therefore these specifications only allowed for general school effects or no school effects. The inclusion of a school-subject-cohort variance into these specifications does not significantly alter the rank parameter. Our findings can be presented upon request.

If this was mainly due to the teacher focusing on those of low rank we would not necessarily expect to see large differences by gender, or free school meal status. We saw that males are more affected by rank than females, which would imply that males are more negatively affected by having the subject content not tailored to them e.g. top males under-achieve more during primary but catch up during secondary school. This is conceivable, however it runs counter to our estimate that males on average have lower growth in test scores between 11 and 14 (Online Appendix Table 6). Moreover, this does not also easily explain why free school meal students up to the middle of the class rankings are not negatively affected by the focus on the bottom, and those at the top of class are. Given these inconsistences, and that it relies on primary school teachers focusing solely on the lowest rank student and secondary school being tailored to student ability to generate similar effects, we doubt that his is the dominant reason for the effect.

7.3 Hypothesis 3: External (parental) investment by task

It may not be the students that are applying different effort by subject but that parents of the students are. Parents can assist the child at home with homework or other extra-curricular activities. If the parents know that their child is ranked highly in one subject, they might encourage the child to do more activities and be more specialised in this subject. Note that as we are controlling for student effects, this must be subject specific encouragement rather than general encouragement regarding schoolwork, and the additional investment must take place between ages 11 and 14. As we have already shown that conditional on test scores, parental occupation does not predict student rank, this hypothesis assumes parents react to achieved primary school rank rather than prior preferences.

However, we believe there are two further counter arguments for this mechanism. Firstly, whilst some parents may choose to specialise their child, others may want to improve their child's weakest subject. If parental investment focused on the weaker subject, this would reverse the rank effect for these students. To explain the positive rank effect, one would need to assume that the majority of parents wanted their child to specialise, which seems unrealistic for the ages eleven to fourteen. Secondly, parents are unlikely to be highly informed of their child's exact rank in class in the English context. Teacher feedback to parents will convey some information for the parents to act upon, such as the student being the best or worst in class, but may not 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 at all vingtiles with school-subject-cohort effects.[29]

[29] Information on the within student comparative advantage by subject would be easier for a teacher to communicate, and so parents could use this to specialize the student. However, these effects would then appear less significant in the school-subject-cohort effects specifications.

7.4 Hypothesis 4: Students learn about their ability

Another possibility is that students use the information obtained by their local rank to learn about their subject-specific abilities, and as a result allocate effort accordingly. This is similar to the model proposed by Ertac (2006) where individuals do not know their own ability and therefore use their own absolute and relative performance to learn about it. Note that 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 (and fail to reject that) the learning model has no effect. Thus we cannot provide no evidence in favour of the learning model. .

Under the learning hypothesis students additionally use local rank information to make effort investment decisions across subjects, applying more effort according to where there is higher perceived ability. This would produce the same predictions by subject as a mechanism that changed the production function, however it has a different prediction for average grades. Students with larger differences between local and national ranks (in absolute terms) would have a more distorted information about their true abilities. These students would then have a higher misallocation of effort across subjects under the learning model, assuming diminishing returns to learning in each subject and that students want to allocate effort where they are most productive. Those with higher misallocation of effort would thus achieve lower overall grades, compared to students whose local ranks happen to closely align with national ranks. This is because this misallocation would lead to inefficient effort allocation across subjects and thus reduce average grades obtained. Whereas, if the rank effects were caused by actual changes in the education production function (and not just learning and changes in perceptions), even if local rank was different from national rank, this would not lead to a misallocation of effort in terms of maximising grades.

We do not have direct data on perceptions versus reality of costs, however we can test for misallocation of effort by examining how average grade achieved is correlated with misinformation. More precisely, we compute a measure of misinformation for students in each subject using their local rank $R_{ijsc} = [0,1]$ and national percentile rank $Y_{ijsc1} = [1,100]$ at age 11. Both are uniformly distributed and therefore we simply define misinformation Mis_{ijsc1} as the absolute difference between the two after rescaling percentile rank:

$$Mis_{ijsc1} = \left| R_{ijsc} - {Y_{ijsc1}}/{100} \right| , where\ Mis = [0,1] \tag{10}$$

This measure takes the value zero for students where their local rank happens to correspond exactly to the national rank. A large value, on the other hand, indicates large differences between local and national rank. Averaging this metric across subjects within student provides a mean indicator of misinformation for each student. To test directly if a student with a large amount of disinformation does significantly worse, we use a specification similar to (5) but with the by subject variation removed as we are examining the effect on average test scores. We estimate the following specification:

$$\bar{Y}_{ijc2} = \beta_{Rank}\bar{R}_{ijc} + f\left(\bar{Y}_{ijc1}\right) + X_i{}'\beta_2 + \varphi_{jc2} + \beta_{Mis}\overline{M\imath s}_{ijc1} + \epsilon_{ijksc} \quad (11)$$

$$\text{Where } \epsilon_{ijc} = \tau_{i2} + \varepsilon_{ijc}$$

where $\bar{Y}_{ijct}$ is average test scores across subjects in period t, $\bar{R}$ is average rank, φ_{jc} are primary school-cohort effects and $\overline{M\imath s}$ the additional misinformation variable. If the amount of misinformation caused them to misallocate effort over subjects we would expect $\beta_{Mis} < 0$, alternatively the null hypothesis local rank causes changes to the actual production function and $\beta_{Mis} = 0$.

$$H_1\colon Learning\ \beta_{Mis} < 0$$

$$H_0\colon Null\ \beta_{Mis} = 0$$

We obtain the following estimates using our full sample of 2,271,999 students. For benchmarking purposes, we first estimate a version of specification (11) without the additional misinformation variable (Table 4). The effect of average rank on average test score is estimated at 10.7 and highly statistically significant.[30] Column (2) adds the coefficient for the effect of misinformation, which is estimated to be small and statistically insignificant whilst the rank parameter remains almost unchanged. Given this specification we fail to reject the null hypothesis that the amount of misinformation does not cause students to misallocated effort. We therefore conclude that the learning mechanism alone is unlikely to generate our results, though we fully acknowledge the limitations of this test, in particular that we cannot control for primary-cohort-subject or student fixed effects in this specification.

[30] This is about three points lager compared to our previous estimates of Table 2 (column 2). Note, however, that this specification does not allow controlling for Primary-cohort-subject effects. Instead, only Primary-cohort effects can be included.

Table 4: Average Age 14 Test Scores on Average Primary School Rank (Age 11) and Misinformation

	Raw (1)	Primary (2)
Primary Rank	10.710**	10.694**
	0.223	*0.223*
Misinformation	-	-0.361
	-	*0.233*
Student characteristics	✓	✓
Age 11 Test Scores (cubic)	✓	✓
Primary-cohort-subject Effects	✓	✓

Notes: Results obtained from two separate regressions based on 2,271,999 student observations averaged over subjects where column (2) includes an additional explanatory variable on misinformation. The dependent variable is by cohort by subject percentalized average KS3 test scores. The misinformation measurement is the average absolute difference between local rank and national percentile rank for each student. Student characteristics are ethnicity, gender, free school meal (FSME) and special educational needs (SEN). Coefficients are estimated using Stata command reg2hdfe allowing two high dimensional fixed effects to be absorbed. Standard errors in italics and clustered at 3,800 secondary schools. ** 1% sig.

7.5 Main hypothesis: Rank position develops self-concept

An alternative explanation is that when surrounded by people who perform a task worse (better) than oneself, one develops a positive (negative) self-concept in that area. 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.

Applied to our setting, we envisage that students with higher rank would develop positive academic self-concept. Self-concepts can be subject specific as well as for academic work generally, so that a student can consider themselves good a school but still bad at math (Marsh et al., 1988; Yeung et al., 2000). Valentine et al. (2004) found that students with a high self-concept would also develop positive non-cognitive skills such as confidence, resilience, and perseverance. There is also broad agreement in the psychological literature that academic self-concept is most malleable before age 11 (Tiedemann, 2000; Lefot et al., 2010; Rubie-Davis, 2011), which is when we measure rank. The importance of such non-cognitive skills for both academic attainment and non-academic attainment is now well established (Heckman and Rubinstein, 2001; Borghans et al., 2008; Lindqvist and Vestman, 2011).

Therefore, the hypothesised mechanism is that an individual's relative rank in a task amongst peers affects self-concept. This in turn has an impact on non-cognitive skills like resilience, persistence and confidence which affects the costs of effort for that task or task-specific productivity directly. An exemplary basic behavioural model that works through the changing-cost channel is provided in Online Appendix 3.

To provide evidence for this mechanism we link the administrative data to the Longitudinal Survey of Yong People in England (LSYPE). We are able to match approximately twelve thousand students from the survey who answer questions on their self-concepts in each subject. This allows us to test directly if rank position within primary school has an effect on this measure of self-concept, conditional on attainment. The specifications are equivalent to (5) and (7) with the dependent variable now being student confidence. Since this survey was only run for one cohort, the school-subject-cohort effects are replaced by school-subject effects.

Panel A of Table 5 presents these results and demonstrates that conditional on age 11 test scores students with a higher primary school rank position are significantly more likely to say that they are good in that subject (column 1). Controlling for school-subject effects, the impact of moving from the bottom of class to the top is 0.196 points on a five point scale (-2, 2), or about twenty per cent of a standard deviation in our self-concept measure (see column 2).[31] This suggests that students develop a clear sense of their strengths and weaknesses depending on their local rank position, conditional on relative test scores.

While we would prefer to have a measure of self-concept directly at age 11 at the end of primary school, these measures are only available to us just prior to the age 14 tests. Therefore, in Panel B we additionally control for contemporaneous attainment at age 14, which is an outcome. To cautiously interpret these estimates, students with 'the same' age 11 and 14 results have higher self-concept if they have had a higher local rank in that subject in primary school.

Note column (2), the specifications allowing for primary-subject effects cannot reject the null hypothesis that rank has no effect on self-concept. A reason for this is that there are few students per primary school in this survey (4.5 students conditional on at least one student being in the survey); as the survey was conducted at secondary schools. The small number of students per school severely limits the degrees of freedom in each school-subject group, the lack of variation is exacerbated due to the coarseness of the self-concept variable. This is exacerbated further when additionally conditioning on individual student effects column 3. To obtain a clearer view of the effect of rank on contemporaneous self-concept we estimate how rank based on age-14 test scores within a secondary school subject affects subject confidence conditional on secondary-subject effects and individual effects. The advantage of this is that there are on average 20 students for

[31] The standard deviation of the self-concept measure is 0.99.

each school that has students in the survey.[32] These results can be found in Panel C, where we see that conditional on school-subject effects, moving from bottom to top of class improves confidence by 0.43 on the 5 point scale. Allowing for individuals to have different levels of confidence and only using the variation between subjects reduces the parameter to 0.38 but remains significant at 1% (column 3).

Furthermore, we examine the heterogeneity of these effects by estimating the effect of age 14 rank on confidence separately by gender, conditional on student and school-subject effects (lower part of Panel C). We find that the effect on male confidence is five times larger than the effect on females ($\hat{\beta}_{rank\ male}$ =0.61, $\hat{\beta}_{rank\ female}$ =0.12), which mirrors the results we find for the effect of rank on later test scores. Unfortunately due to the smaller sample size of the LSYPE, we are unable to produce the effects non-linearly or by FSME status.

Table 5: Self-Concept on Rank

	(1)	(2)	(3)
Panel A: Self-Concept on Age 11 Test Scores			
Primary Rank	0.563**	0.196*	0.056
	0.038	*0.117*	*0.18*
Panel B: Self-Concept on Age 11 & 14 Test Scores			
Primary Rank	0.436**	0.109	0.014
	0.039	*0.115*	*0.079*
Panel C: Self-Concept on Age 14 Test scores			
Secondary Rank	0.897**	0.427**	0.382**
	0.048	*0.099*	*0.155*
Secondary Rank – Male Students	0.754***	0.530***	0.606***
	0.059	*0.126*	*0.206*
Secondary Rank – Female Students	1.067***	0.317*	0.115
	0.071	*0.166*	*0.233*
School-by-subject effects		✓	✓
Student Effects			✓

Notes: Results obtained from fifteen 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 Online Appendix Table 3. The dependent variable is a course measure of self-concept by subject. All specifications in columns 1 and 2 control for observable student characteristics, these are absorbed by the student effect in column 3. Student characteristics are ethnicity, gender, free school meal (FSME) and special educational needs (SEN).Panels A and B condition on age 11 test scores (cubic) and primary school by subject effects. Panels B and C condition on age 14 test scores (cubic) and secondary school by subject effects. Cohort effects are not included because the LSYPE data is only available for one cohort. Standard errors in parenthesis and clustered at 796 secondary schools ** 1% sig. * 10% sig.

[32] The reason why we do not look at the effect of KS3 rank on later outcomes is due to the tracking by subject in secondary school, which will be related to rank. This is not an issue with primary school rank, because even if there were tracking in primary schools, when moving to secondary school, students with the same test scores (but different primary ranks) would be assigned to the same track.

The magnitudes of the secondary school ranks effects on secondary confidence are large, but we may expect the contemporaneous effect of primary rank on confidence at age 11 to be even larger, as self-concept is thought to be more malleable at this age (Tiedemann, 2000; Lefot et al., 2010; Rubie-Davis 2011). Moreover, we find indicative evidence that later confidence is affected by previous primary school rank.

Overall given the effects of rank on direct measures of self-concept and the heterogeneity of effects found in the main results we are confident in our conclusion that self-concept forms according to rank position and that this affects later investment decisions.

8 Corroborating research

The finding that higher peer quality could have negative effects on later outcomes may seem controversial, but there are a number of topics in education that have findings which corroborate this hypothesis.

Research on selective schools and school integration have shown mixed results from students attending selective or predominantly non-minority schools (Cullen, et al., 2006; Angrist and Lang, 2004; Kling *et al.*, 2007). 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 general puzzle is that many papers find no benefit from attending these selective schools. However, our findings would speak to why the potential benefits of prestigious schools may be attenuated through the development of negative self-concepts amongst these marginal/bussed students, who necessarily would also be the low ranked students. This is consistent with Cullen, Jacob and Levitt (2006), 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'. Similar effects are found in the Higher Education literature with respect to affirmative action policies (Arcodiacono *et al.,* 2012; Robles and Krishna, 2012).

The early formation of self-concept and specialisation could also partly explain why some achievement gaps increase over the education cycle. Widening overall education gaps have been documented by race (Fryer and Levitt, 2006; Hanushek and Rivkin 2006; 2009), small differences in early overall attainment could negatively affect general academic self-concept, which would lead to decreased investment in education and exacerbate any initial differences. In the case of gender a gap occurs by subject, where males are overly represented in mathematics and science by the age of 18, despite girls outperforming boys at early ages in these subjects (Burgess et al, 2004; Machin & McNally, 2005). Even with girls performing better in all subjects, if boys do comparatively less badly in mathematics and are more affected by rank for investment decisions, then they would chose to invest more in those subjects. Finally the literature on age-effects in education shows that older children do better compared to their younger peers (i.e. Black *et al.*, 2011). The development of positive self-concepts of the older children at an early age due to initial differences is a potential mechanism for the continuation of these effects as the students grow older.

9 Conclusions

Individuals continuously make social comparisons, which can affect our beliefs and investment decisions. If individuals make these comparisons using ordinal as well as cardinal information, then an individual's rank amongst their peers could impact on their investments and later productivity.

This paper examined how, conditional on relative achievement, rank amongst peers affects subsequent performance. Applied to an education setting we establish a new result, that rank position within primary school has significant effects on secondary school achievement. Moreover, a higher rank also improves students' confidence, an important non-cognitive skill. These rank effects are in addition to any effect caused by a student's relative distance from the class mean.

The approximately linear impact of rank implies that students are very good at determining their rank amongst their peers. Furthermore, there is significant heterogeneity in the effect of rank with males being influenced considerably more. We find male confidence in a subject is five times more affected by their rank amongst their peers compared to females. Accordingly, male students specialise more according to their primary school rank than females. To the extent that boys gain four times more in later test scores from being top of the class compared to comparable female students. Contrastingly, students with low parental income background are not negatively affected by low rank positions during primary education. Together, we take this as evidence that an individual's rank amongst their peers during primary school affects their self-concepts over many dimensions which in turn are likely to impact on the development of task specific non-cognitive skill and subsequent investment decisions.

We cannot fully exclude other mechanisms, such as learning about ability, to generate parts of these results. However, we have shown that differences between local and national ranks have no negative impact on average performance. This speaks in favour of mechanisms that change the actual grade production function either through shifting task-specific productivity or cost, and against learning models where only student perceptions are affected. Given the impact of rank on a direct measure of confidence, we thus believe that rank is most likely to affect later results through non-cognitive skills.

It is worthwhile to think about policy implications of this finding.

With specific regards to education, these findings leads to a natural question for a parent deciding on 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 many factors in the education production function, and therefore choosing solely on the basis of rank is unlikely to be correct. The authors are currently not aware of any study that identifies the effectiveness of schools in terms of standard deviations[33]; therefore, we use estimates of the impact of teachers as an indicative measure for effects of school quality for this benchmarking exercise. A teacher who is one standard deviation better than average improves student test scores by 0.1 to 0.2 standard deviations (Aaronson, et al. 2007; Rivkin et al. 2005). Comparatively we find that a student with one standard deviation higher rank in primary school will score 0.08 standard deviations better at age 14.[34] Forthcoming work will look at the longer run impacts of primary school rank, as well as changes in school ranks from moving schools.

We believe these findings have 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 had a high local rank. If an individual is in a high-performing peer group and therefore may have a low local rank but a 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.

Finally these findings have general implications regarding the formation of non-cognitive skills and productivity. Given the heterogeneous effects of rank it would be possible to organise groups by individuals characteristics and abilities to maximise output. However this would be very cumbersome and administratively intensive. Therefore the key implication is that non-cognitive skills such as confidence, perseverance and resilience have large effects on productivity. Rank can be thought of as just one treatment that impacts on these behaviours, however there are many other interventions that could have positive effects on all individuals within a group and not just those above the median.

[33] Evaluations of school effectiveness using admission lotteries (i.e. Hoxby et al. 2009, Angrist et al. 2010, Dobbie and Fryer 2011, Abdulkadiroglu et al. 2011) are comparing effectiveness between types rather than the whole distribution of effectiveness.

[34] Note that these are still not directly comparable because the effect of the teacher is annual and quickly fades out, whereas the rank treatment lasts the duration of primary school (5 years) and the effect is found three years later.

References

Aaronson, D., L. Barrow, and Sander, W. (2007). “Teachers and student achievement in the Chicago public high schools”, *Journal of Labor Economics*, 25(1), pp. 95–135.

Abdulkadiroglu, A., Angrist, J., Dynarski, S., Kane, T., and Pathak, P. (2011). “Accountability and flexibility in public schools: Evidence from Boston's charters and pilot", The *Quarterly Journal of Economics*, 126(2): pp. 699-748.

Angrist, J. (2013). “The Perils of Peer Effects”, NBER Working Paper No. 19774.

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

Angrist, J., Dynarski, S., Kane, T., Pathak, P., and Walters, C. (2010). “Inputs and impacts in charter schools: KIPP Lynn", *American Economic Review*, 100(2), pp. 239-243.

Arciciacono, P., Aucejo, E. and Spenner, K. (2012). “What Happens After Enrolement? An Analysis of the Time Path of Racial Differences in GPA and Major Choice”, *IZA Journal of Labor Economics,* 1(5).

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), pp. 435-452.

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), pp. 455–467.

Borghans, L., Duckworth A., Heckman J., Weel ter Bas (2008). “The Economics and Psychology of Personality Traits”, *Journal of Human Resources*, 43(4), pp.872-1059.

Brown G., Gardner, J., Oswald, A. and Qian, J. (2008). “Does Wage Rank Affect Employee’s Well-being?”, *Industrial Relations*, 47(3), pp. 355-389.

Brown J., (2011). “Quitters Never Win: The (Adverse) Incentive Effects of Competing with Superstars”, *Journal of Political Economy*, 119(5), pp. 982-1013.

Burgess S., McConnell B., Propper C., Wilson D., “Girls Rock, Boys Roll: An Analysis of the Age 14-16 Gender Gap in English Schools”, *Scottish Journal of Political Economy*, Scottish Economic Society, 51(2), pp. 209-229.

Card, D., Mas, M., Moretti, E., and Saez, E. (2012). “Inequality at Work: The Effect of Peer Salaries on Job Satisfaction”, *American Economic Review*, 102(6), pp.2981-3003.

Carrell, S.E., Fullerton R.L., and West, J.E. (2009). “Does Your Cohort Matter? Estimating Peer Effects in College Achievement”, *Journal of Labour Economics*, 27(3), pp.439-464.

Clark, A., and Oswald, A. (1996). “Satisfaction and Comparison Income", *Journal of Public Economics*, 61(September), pp. 359-381.

Clark, A., Masclet, D. and Villeval M.C., (2010). “Effort and comparison income: experiment and survey evidence”, *Industrial and Labour Relations Review*, 63(3) pp. 407-426.

Clark, D. (2010) “Selective Schools and Academic Achievement” *The B.E. Journal of Economic Analysis and Policy*, 10(1), pp. 1935-1682.

Cullen, J.B., Jacob, B.A. and Levitt S., (2006). “The effect of school choice on participants: evidence from randomised lotteries”, *Econometrica*, 74(5), pp.1191-1230.

DFE (2011). “Class Size and education in England evidence report”, Research Report DFE-RR169.

Dobbie, W. and Fryer, R. (2011). “Are high quality schools enough to close the achievement gap? Evidence from the Harlem Children's Zone", *American Economic Journal: Applied Economics*, 3(3), pp. 158-187.

Eriksson, T., Poulsen, A., Villeval, M., (2009). “Feedback and incentives: experimental evidence”, *Labour Economics*, 16, pp. 679–688.

Falk, A., Ichino, A., (2006) “Clean evidence on peer pressure” *Journal of Labor Economics* 24 (1), pp. 39–57

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

Fryer, G. and Levitt, S. (2006). “The Black-White Test Score Gap Through Third Grade”, American Law and Economics Review, 8(2), pp. 249-281.

Genakos, C. and Paglierio M. (2012) Interim Rank, Risk Taking, and Performance in Dynamic Tournaments. *Journal of Political Economy*, University of Chicago Press, vol. 120(4), pages 782 - 813.

Guimaraes, P. and Portugal, P. (2010) "A Simple Feasible Alternative Procedure to Estimate Models with High-Dimensional Fixed Effects", *Stata Journal*, 10(4), pp. 628-649.

Hanushek, E. (2006). “School Resources”, *Handbook of the Economics of Education*, 2(Chapter 14).

Hanushek, E., Kain, J. and Rivkin, S. (2005). “Teachers, Schools, and Academic Achievement”, *Econometrica*, 73(2), pp. 417–458.

Hanushek, E. and Rivkin, S. (2006). “School Quality and the Black-White Achievement Gap”, NBER Working Paper No. 12651.

Hanushek, E. and Rivkin, S (2009). “Harming the Best: How Schools Affect the Black-White Achievement Gap”, *Journal of Policy Analysis and Management*, 28(3), pp. 336-393.

Heckman J., Stixrud, J. and Urzua, S. (2006). "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior", *Journal of Labor Economics*, 24(3), pp. 411-482.

Hoxby, C. (2001). "If Families Matter Most, Where Do Schools Come In?" in T. Moe, ed. A Primer on American Schools. Stanford: Hoover Institution Press.

Hoxby, C., Murarka, S., and Kang, J. (2009). “How New York City's charter schools affect achievement.", Cambridge, MA: New York City Charter Schools Evaluation Project.

Kling, J., Liebman, J. and Katz, L. (2007). “Experimental Analysis of Neighborhood Effects”, *Econometrica*, 25(1), pp. 83-119.

Kosfeld, M. and Neckermann, S. (2011). “Getting More Work for Nothing? Symbolic Awards and Worker Performance”, *American Economic Journals: Microeconomics*, 3(3), pp. 86-99.

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*, 20(2), pp.367-414.

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), pp. 101-28.

Kuziemko, I, Buell, R., Reich, T., and Norton, M. (2011). “Last-place Aversion: Evidence and Redistributive Implications", NBER Working Paper No. 17234.

Machin, S., and McNally, S. (2005). “Gender and student achievement in English schools”, *Oxford review of economic policy*, 21(3). pp. 357-372.

Manski, C. (1993). “Identification of Endogenous Social Effects: The Reflection Problem”, The Review of Economic Studies, 60(3), pp. 531-542.

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

Mas, A. and Moretti, E., (2009). “Peers at work”, *American Economic Review*, 99(1), pp. 112–145.

Nickell, S. (1981). “Biases in Dynamic Models with Fixed Effects”, *Econometrica*, 49(6), pp. 1417-1426.

Page, M., Carrell, S., and West, J. (2010). “Sex and Science: How Professor Gender Affects the Gender Gap", *Quarterly Journal of Economics*, 125(3), pp.1101-1144.

Parducci, A (1965). "Category Judgment: A Range-frequency Theory", *Psychological Review*, (72), pp. 407-418.

Robles, F. V.C. and Krishna, K. (2012). “Affirmative Action in Higher Education in India: Targeting, Catch Up, and Mismatch”, NBER Working paper No.17727.

Tiedemann, 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), pp 144-151.

Todd, P.E. and Wolpin K.I., (2003). “On The Specification and Estimation of The Production Function for Cognitive Achievement”, *The Economic Journal*, 113(1), pp. 3-33.

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

Valentine, J., DuBois, D. and Cooper, H. (2004). “The relations between self-beliefs and academic achievement: A systematic review”, *Educational Psychologist*, (39), pp.111–133.

Watkins, M., Lei, P. and Canivez, L. (2007). “Psychometric intelligence and achievement: A cross-lagged panel analysis”, *Intelligence*, 35, pp.59-68.

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