Class Rank and Long-Run Outcomes

Jeffrey T. Denning[a], Richard Murphy[b] and Felix Weinhardt[c]

June 2020

Abstract

This paper considers an unavoidable feature of the school environment, class rank. What are the long run effects of a student's ordinal rank in elementary school? Using administrative data from all public school students in Texas, we show that students with a higher third grade academic rank, conditional on achievement and classroom fixed effects, have higher subsequent test scores, are more likely to take AP classes, to graduate from high school, enroll in college, graduate from college, and ultimately have higher earnings 19 years later. The paper concludes by exploring the tradeoff between higher quality schools and higher rank in the presence of these rank-based peer effects.

JEL classification: I20, I23, I28
Keywords: rank, education, subject choice, peer effects

Acknowledgements: We thank Sandra Black and Brigham Frandsen as well as participants of the Society of Labor Economists Annual Meetings, the IZA Junior/Senior Symposium, the CESifo Area Conference on the Economics of Education, University of Utah Department of Finance, and the STATA Texas Empirical Microeconomics Conference for feedback and comments. Weinhardt gratefully acknowledges financial support by the German Science Foundation through CRC TRR 190 (Project number 280092119). All errors are our own. Disclaimer: The research presented here utilizes confidential data from Texas Education Research Center (TERC) at The University of Texas at Austin. The views expressed are those of the authors and should not be attributed to TERC or any of the funders or supporting organizations mentioned herein. Any errors are attributable to the authors alone. The conclusions of this research do not reflect the opinion or official position of the Texas Education Agency, Texas Higher Education Coordinating Board, the Texas Workforce Commission, or the State of Texas. Correspondence: richard.murphy@austin.utexas.edu

[a] Brigham Young University, NBER, IZA, CESifo, jeffdenning@byu.edu
[b] University of Texas at Austin, NBER, IZA, CESifo, CEP richard.murphy@austin.utexas.edu
[c] Humboldt-University Berlin, DIW Berlin, CESifo, IZA, CEP, felix.weinhardt@hu-berlin.de

1. Introduction

There is a large literature examining peer effects in education. This literature typically focuses on either the benefits of high performing peers (Sacerdote, 2001; Whitmore, 2005; Kermer and Levy, 2008; Carrell et al., 2009; Black et al., 2013; Booij et al., 2017), or the negative effects of having disruptive peers (Hoxby and Weingarth, 2006; Lavy et al., 2012; Carrell and Hoekstra, 2010; Carrell et al., 2018). However, there is another mechanism, where having lower-performing peers could *improve* student outcomes—namely a student's rank.[1] We will explore a student's ordinal rank in their classroom and the persistence of this effect on their outcomes into adulthood. Rank is an appealing attribute to study because it occurs naturally in any group of people.

We consider a student's academic rank in third grade (8 to 9 years old), conditional on their achievement, on short and long run outcomes. We use the universe of public-school students in Texas from 1994-2006 and combine this with an identification strategy that leverages idiosyncratic variation in rank. We find that a student's rank in third grade impacts grade retention, test scores, AP course taking, high school graduation, college enrollment, and earnings up to 19 years later.

Academic achievement and rank are highly correlated. So, to isolate the effect of a student's rank, we develop the method used in Murphy and Weinhardt (2014/*forthcoming*) and leverage the idiosyncratic variation in the distribution of test scores across schools, subjects, and cohorts. We define a student's achievement by their test score expressed as a percentile of the state population. We then compute a student's rank within their school, subject, and cohort and express it as a percentile.

Consider the following hypothetical: two students in successive cohorts of the same size and mean attainment, at the same school, who have the same math achievement (as measured by their place in the state-wide distribution). Because school cohorts are small relative to the state cohort, there will be variation in the test score distribution such that one student may be the fifth best student in the class while the other may be the eighth. This is the

[1] This can occur through various channels. These channels can be categorized as internal (learning about ability, development of non-cognitive skills) or external (parental and school investments).

idiosyncratic variation we leverage to identify the effect of rank, which we argue originates from sampling variation in the distribution of human capital within subjects and cohorts of each school. This variation exists because these groups sample relatively small numbers of students only so that distributional differences emerge by chance.[2]

Our idealized experiment would be to compare students with the same level of human capital who are randomly assigned to different classrooms. Sampling variation would lead to naturally occurring variation in rank for students with the same human capital. The question of this paper is whether a student's outcome is dependent upon their rank in the group. In this idealized setting this would be comparing students with the same pre-determined human capital who have the same classroom setting in expectation. We discuss this idealized experiment and its relation to our empirical strategy in detail in section 2.

We perform a battery of further robustness checks to establish the underlying assumptions are valid including balancing of rank on individual-level covariates, alternate functional forms, and definitions of rank. We also present estimates for small schools where there is likely to be one classroom per grade. We address any further concerns about measurement error in student test scores resulting in rank being a proxy for student human capital: we show that individual non-systematic additive measurement error in student test scores, which would be non-rank preserving, would cause a small downward bias to our estimates. Meaning that to the extent that test scores are a noisy measure of human capital and student rank, our estimates should be interpreted as lower bound.

We test for heterogeneous rank effects by gender, parental income, and race. We find the impact of rank on male and female students to be very similar, regardless of outcome. In contrast, we find that disadvantaged students (non-white or Free and Reduced Price Lunch) are significantly more affected by rank than their advantaged counterparts. This is seen throughout a student's life, affecting eighth grade test scores, high school graduation, college enrollment, and earnings.

[2] Hoxby (2000) is an early example of a study using sampling variation in the context of peer effect estimation. Importantly, the variation in the treatment of interest in this paper (rank) occurs within a class which means we can condition on class effects.

A natural question to ask is, why would third grade rank conditional on achievement, impact these outcomes? One possibility is that humans think in terms of heuristics (Tversky and Kahneman, 1974) and therefore rely on ordinal rank position rather than the more detailed cardinal position within a group. Alternatively, ordinal rank may be easier to observe than cardinal position. Regardless of the precise behavioral origins of the effect, its impact can be seen in the findings that an individual's ordinal position within a group predicts well-being (Luttmer, 2005; Brown et al., 2008) and job satisfaction (Card et al., 2012), conditional on cardinal measures of relative standing. Hence, rank may also impact investment decisions and subsequent productivity. In the education literature, this is known as the Big-Fish-Little-Pond effect where individuals gain in confidence, when they are highly ranked in their local peer group (for a review see Marsh et al., 2008).[3]

In this paper however, we are agnostic about what is driving the effects. We define the rank effect to include any reactions to the rank a student acquires from entering a group by any individual. For example, if teachers invest more effort in the worst (or best) students, irrespective of their absolute performance, then this would be included into the rank effect. Students may also adjust their effort as a result a of rank and hence change their academic achievement which in turn changes their rank (Tincani 2017). However, this change in rank operates through a change in achievement, whereas we are primarily focused on the effects of rank conditional on achievement. In summary, anything that is a reaction to a student's rank is a potential mechanism including student effort, parental investment, and teacher investment. Predetermined confounding factors would need to covary with rank conditional on achievement in order to generate a bias. We establish the lasting impact of elementary school rank on long run outcomes. Note that, while we are agnostic about the mechanisms that give rise to the rank effect, we do take great care in establishing that the rank effect is not a product of confounding factors.

Recent studies have documented that a student's relative rank affects short run outcomes independent of achievement. Murphy and Weinhardt (*forthcoming*) document the effect of primary school rank, conditional on achievement, on high school test scores and confidence. Two papers by Elsner and Isphording apply the same idea to the United States

[3] This has been referred to as the invidious comparison peer effect by Hoxby and Weingarth (2006).

using data from the National Longitudinal Study of Adolescent to Adult Health (AddHealth) to study effects of contemporaneous high school rank on high school completion, college going (2017a), and health outcomes (2017b).[4] Zax and Rees (2002) consider the effect of peers and other student characteristics, including IQ and rank within high school on earnings at ages 35 and 53, using a sample of approximately 3,000 male students from Wisconsin. [5]

We make two contributions to this literature. The first is conceptual—we discuss identification and threats to identification in ways that have not been discussed in this literature. We also propose tests for and solutions to identification challenges that are common to many papers in this literature. The second contribution is empirical. We extend the literature by using administrative data on three million individuals to look at the long-term effects of a rank at a young age (in third grade) on adult outcomes.

Explicitly, we consider a student's rank at ages younger than have previously been considered (age 8-9) on outcomes up to 19 years later.[6] Our large dataset allows us to examine nonlinear effects of rank and we find that rank has a nonlinear relationship with outcomes in several instances. We also consider the effects of rank in elementary school. Studying the effects of rank in high school is interesting but conceptually different. Elementary school students are arranged in classrooms where students share the same teacher, peers, and curriculum for the majority of the day; whereas, high school students do not share the same teacher, peers, or curriculum. Hence, rank may affect things like course taking (and peers), which complicates the interpretation of rank in high school.

---

[4] The AddHealth home-survey contains a sample of only 34 students for each school cohort, which are used to compute a measure of rank (Elsner and Isphording 2017a, 2017b). More recently, there is a set of working papers which estimates the impact of rank within college on contemporaneous outcomes in various countries (Elsner et al. 2018; Payne and Smith, 2018; Ribas et al. 2018, Ribas et al. 2020, Delaney Devereux 2019).

[5] A distinct literature has studied the introduction of relative achievement feedback measures in education settings. Azmat and Iriberri (2010) find that providing information on relative performance feedback during high school increases productivity of all students when they are rewarded for absolute test scores. In contrast, Azmat et al. (2019) find relative feedback in college causes significant short run decreases in student performance, but no long run effects.

[6] Murphy and Weinhardt (*forthcoming*) consider the effect of rank on test scores 3 and 5 years later. Elsner and Isphording (2017a) consider rank in high school on college enrollment and graduation—approximately 2 to 8 years later. Elsner and Ishpording (2017b) consider the effect of rank on risky behavior as reported 18 months later. Murphy and Weinhardt consider rank at age 10-11. Elsner and Isphording consider rank at age 14-18.

Our findings contribute to a growing literature that documents how childhood conditions affect adult outcomes. These conditions range from a child's health (Oreopoulos et al., 2008), where a child lives (Chetty et al., 2016), the quality of a student's teacher (Chetty et al., 2014), size of a student's classroom (Chetty et al., 2011), the age of a student when they start school (Black et al., 2011), and the presence of disruptive peers (Carrell et al., 2018; Bietenbeck, 2020) among others. We add to this list that a child's rank in their third grade classroom, independent of their achievement, has meaningful effects on education and earnings in adulthood.

We find that rank effects are larger for historically disadvantaged groups such as non-white students or students eligible for Free or Reduced Price Lunch (FRPL). One implication of this is that, unlike linear in means peer effects, where moving students between groups would have no net impact, rearranging students with rank in mind could improve overall outcomes. However, this sort of exercise merits caution because the changes in classroom distribution will have general equilibrium effects not accounted for in this paper (Carrell et al., 2013).

The more practical implication of this finding is that programs that move disadvantaged students into 'high quality' schooling sometimes comes at a cost that they will be the lowest ranked student. The extensive literature on selective schools and school integration has shown mixed results from students attending selective or predominantly non-minority schools (Angrist and Lang, 2004; Clark, 2010; Cullen et al., 2006; Kling et al., 2007; Abdulkadiroğlu et al. 2014; Dobbie and Fryer, 2014; Bergman, 2018). Our findings would speak to why the potential benefits of prestigious schools may be attenuated among these marginal/bussed students. We explore the magnitude of the tradeoffs between rank and school quality in Section 7.

To examine any such potential rank school-quality tradeoffs, we address the parental question of which school parents should select for their children, given the existence of the rank effects. We find that, if parents were to choose schools solely on the basis of mean peer achievement, rank effects would reduce 39 percent of the potential gains for median performing students in the state. In contrast, when choosing based on value added, rank effects only reduce the gains from choosing a better school by 12 percent.

The rest of the paper is structured as follows. Section 2 briefly sets out the empirical design. Sections 3 describes the data. Section 4 determines the appropriate specification. Section 5 presents the results. Section 6 performs a series of robustness tests, and examines heterogeneity. Section 7 discusses implications for school choice. Lastly, Section 8 concludes.

2. Empirical Design

In this section we will discuss how to identify the effect of class rank. We will first discuss the ideal experiment. Next, we will discuss how reality departs from this ideal experiment and how that motivates our choice of specification.

2.1. Idealized experiment

Our idealized experiment would be to compare students with the same pre-determined level of human capital who were randomly assigned to different classrooms. Sampling variation would lead to naturally occurring variation in rank for students with the same human capital. We are interested in whether a student's outcome is dependent upon their rank in the group. In the ideal setting with randomization to classrooms, our identical students would be assigned to classrooms that are balanced with respect to other characteristics such as teacher quality, peer quality, test score distribution, etc. We then would compare outcomes for students who with the same predetermined human capital but differed in rank due to sampling variation to identify the effect of rank apart from human capital. [7]

With the ideal experiment we would estimate the following equation among students with the same human capital. Let Y_{ijsc} be the outcome of student i who attended elementary school j in subject s from cohort c. Let R be a student's rank.

$$Y_{ijsc} = f(R_{ijsc}) + \epsilon_{ijsc} \tag{1}$$

[7] This sampling variation will cause changes in other relative measures of student achievement. However, they are unlikely to be systematically related to rank. If other relative measures are correlated with rank, they can be thought of as a feature of rank because they are inextricably linked to rank, and so we consider to be part of the rank effect. For example, it would be difficult for a student of very low ability to not have the bottom rank without the presence of other extremely low ability students. Stated differently, if rank is always associated with some other feature of the distribution, we call this the effect of rank. Rank can be thought of as a summary measure of ordinal position.

To identify $f()$ we would need a conditional independence assumption that $E[\epsilon_{ijsc}] = 0$ which would be true in the ideal experiment due to randomization. Equation 1 would identify the effect of rank for students of fixed level of human capital.[8]

We could extend this to students with different levels of human capital by estimating the following where T is a measure of human capital.[9]

$$Y_{ijsc} = f(R_{ijsc}) + g(T_{ijsc}) + \epsilon_{ijsc} \tag{2}$$

Even in this ideal experiment we would consider all reactions to the rank of a student as part of the rank effect. This includes reactions by students, teachers, parents, or other students. Understanding the mechanisms for the rank effect is interesting but not the primary focus of this paper.

We replicate such an ideal experiment with an illustration. Figure 1 shows the Math test score distributions of seven hypothetical elementary school classes. The classes share a similar test score distribution, each has a student scoring 22 and 38 and have a mean test score of 30. In four of the classes there is a student scoring exactly 35, however due to the idiosyncratic variation in the test score distributions, each of these students have a different third grade math rank, ranging between 0.7 and 0.9.[10]

Despite these similarities, reality departs from the idealized experiment in two ways. We do not have direct measures of pre-determined human capital, and we do not have random allocation to schools and classes. These two differences could yield omitted variable bias if we simply compared students with the same achievement measure but with different ranks. We now set out the specifics of how these departures could generate a spurious

[8] This also requires something akin to an "exclusion restriction" where assignment to classroom only affects future outcomes via the effect of rank. In this ideal experiment this is likely to hold since classrooms should be balanced on other factors such as teacher quality, school resources, etc. due to randomization.

[9] Equation 2 adds an assumption that the effect of rank is the same regardless of human capital. This is not required but simplifies the analysis. Moreover, assuming the same effect of rank across subjects which is relaxed in the estimates shown in Figures 5 and 7.

[10] To compare ranks across classrooms of different sizes, we normalize rank to range from zero to one (Section 3.2. for details).

correlation between rank and future outcomes to illustrate how our specification addresses them.

2.2. Measurement error in Human Capital

First, we do not have direct measures of human capital, instead we use student achievement in externally graded statewide examinations.[11] A concern is that despite having the same absolute achievement measures (e.g. the same scores on the same test), these test scores are a noisy measure of true human capital. Measurement error could be at the individual level, or at the class level. Individual-level measurement error would attenuate any rank effects as is common in many settings.[12] However, class-level measurement error in test scores could generate a spurious relationship between rank and achievement. An example of a class level measurement error would be a disturbance on test day. This class shock would affect all students' test scores creating measurement error in our measure of human capital, but would also be rank preserving.

Class level measurement error in test scores, which is rank preserving, would make rank a proxy for underlying ability, because the same absolute achievement measures in different schools will reflect different levels of human capital. In this case, larger measurement error would cause test scores to become less reliable, but would leave rank unaffected, which would lead to rank acting as a proxy for underlying student human capital and generate a spurious "effect" of rank. Since researchers always have imperfect measures of human capital, one must include class fixed effect to account for any mean-shifting effects of the school

[11] We use third grade test scores as it is the earliest achievement measure available. It is possible to estimate the impact of later grade ranks on outcomes. However, as we find that future test scores are increasing in previous rank, it means that later rank itself is an outcome and rank is self-perpetuating. Hence, it is difficult to determine when a child's ranking has the largest impact. Ideally, we would rank students based on a before-school measure of achievement. Rank effects likely operate not only though the score on this particular exam, but more general performance in the class. We are interested in the effect of rank in academic achievement and use third grade test scores as a convenient summary measure.

[12] Despite the measurement error in treatment being non-standard, as it will be correlated non-linearly with achievement, in Section 6.2 (and Appendix Figure 6) we show that adding measurement error to achievement and then recalculating ranks on this measure will lead to a downward bias.

environment such as teachers, resources, peers, etc.[13] We modify equation 2 to account for potential classroom level shocks as follows

$$Y_{ijsc} = f(R_{ijsc}) + g(T_{ijsc}) + \theta_{jsc} + \epsilon_{ijsc} \tag{3}$$

where θ_{jsc} is a School-Subject-Cohort (SSC) fixed effect which we refer to as a class fixed effect.[14]

The concern about measurement error of human capital can be extended to all rank-preserving school-level impacts that affect test scores. Classrooms contain a bundle of (potentially) rank-preserving treatments which impact achievement including class size, teachers, the pre-determined human capital of peers, etc. This will impact the observed measure of achievement of each student in the class, which can be thought as a form of measurement error in the observed measure of human capital. All of these factors that additively impact class level measures of achievement are accounted for with classroom fixed effects.[15]

2.3. Non-random sorting

The second deviation from the idealized experiment is that there is substantial sorting of students to classrooms in the United States. We categorize sorting into two types: active sorting and passive sorting. We define active sorting when parents (or students) choose their classroom on the basis of exact rank. Passive sorting occurs when students are sorted to schools for reasons unrelated to their rank but that may generate a spurious correlation

[13] This accounts for any homogenous peer effect, such as 'shining-light' ('bad-apple ') effects, in which all students in the class gain (lose) from having high (low) achieving student in the class (Hoxby and Weingart, 2006).

[14] We do not have data on class records. We refer to school-by-cohort-by-subject as a classrooms throughout the paper. In small schools with one class per cohort these are the same. We show in Robustness Section 6.2 that our results apply to small and large schools.

[15] The inclusion of classroom fixed effects reduces the amount of unexplained variation there is in test scores, and makes student achievement a relative to the class mean measure. To illustrate that variation in rank conditional on relative achievement remains Appendix Figure 1 plots math achievement de-meaned by school and cohort against class rank. There is a large amount of variation in rank for a given test scores throughout the achievement distribution. This can be considered to be an extension to Figure 1, Panel B to all classrooms which have different distributions, de-meaned by mean achievement.

between rank and other factors. If students are actively or passively sorted to schools then there is the risk of omitted variables producing a spurious rank effect.

Active sorting is only possible if parents can predict their child's rank at a school. However, in practice this would be difficult. Two students with the same achievement, at the same school, in different cohorts will often have meaningfully different ranks. Figure 2 shows the distribution of class rank that a student with median statewide achievement would receive in each of the 13 cohorts, by each school-subject group. Each school-subject is one column on the horizontal axis which we have scaled from 1 to 100 for ease of interpretation.[16] The school-subject groups are sorted such that median class rank of the state median student is increasing, e.g. each school-subject have up to thirteen observations according to the class rank of a student being at the state median, the median of these are used to sort school-subjects..

We plot the class rank of the state median student at given percentiles (10/25/50/75/90th) within the school-subject over the 13 cohorts. There is considerable variation within a school subject group for students at the statewide median level of achievement. For example, median state students enrolled in a school at the 50th Elementary School-Subject group, could have ranks as low as 0.4 (10th percentile) or as high as 0.6 (90th percentile) depending upon which cohort they enrolled in the school. This variation in rank would make it very difficult to actively sort on the basis of rank difficult, even if parents had such preferences, full information about the school environment, and exact measures of student human capital before choosing a school. Hence, we think that active sorting is unlikely to be a problem in our setting. [17]

Even though it is difficult to for a parent to actively sort their child into a school to guarantee that a student's precise rank, the gradient in Figure 2 shows that in some schools a student at the statewide median will consistently have a high class rank (right hand side) and in others would have a low rank. If students with certain characteristics systematically attend schools with a certain test score distribution, then this is what we refer to as passive sorting.

[16] We collapse schools-subject values by 100 percentiles, but the conclusions are unchanged when using finer data.

[17] If motivated parents sorted on the basis of the mean attainment of peers, then this would be negatively correlated with student rank and downward bias the effect.

Passive sorting could generate a spurious rank effect. Panel A of Figure 3 shows two types of schools, both with achievement approximately normally distributed with the same mean, but Type A has high variance in student achievement and Type B has low variance. There are also two types of students, and their type impacts their future outcomes independent of rank e.g. disadvantaged and non-disadvantaged. The red line compare represents two students whose achievement is the same distance below their classroom mean in both distributions. Students at this point in high variance class will have systematically higher rankings than students in the low variance class. If students are randomly assigned to classes, then this rank will be uncorrelated with student characteristics. However, if non-disadvantaged students more often sort into Type A and disadvantaged students more often sort into Type B, then rank becomes correlated with student characteristics (in this case disadvantage) conditional on test scores and class fixed effects. This would generate a spurious rank effects due to the passive rank-sorting into certain types of schools. In this hypothetical example, student characteristics are correlated with rank and future outcomes. This highlights that distributions with different variance can generate a spurious relationship between rank and student characteristics.

This is a simple example where only the variance of the class differed. However, the principal of passive sorting can be applied to schools varying in any higher moment of the test score distribution. All that needs to occur is for students of a certain type and achievement to systematically attend schools with distinct distributions from other types of students. If this is the case, then rank can become correlated with student characteristics conditional on achievement. Figure 3 Panel B provides another illustration of this, where two students with the same achievement relative to the classroom mean will have different ranks due to the test score distribution. As in the Panel A students in school type B will have a higher rank for a given distance to the class mean. This is a problem if students of a certain type systematically attend a school with this type of distribution.

Note that we have discussed passive sorting with regards to student characteristics; however, any factor that is systematically related to the test score distribution of the school may cause a spurious relationship between rank and the outcomes. For example, there is substantial segregation by income in Texas schools and so a low-income student and a high-

income student with the same human capital are likely to have different school environments and distributions.

As illustrated above, the inclusion of class fixed effects would not account for this passive sorting based on higher moments of the distribution. The intuition for this is that by definition every class has the same amount of high and low ranked students, and so these fixed effects will be uncorrelated with the average rank in a class. Therefore, conditioning on something uncorrelated with the treatment, rank, will not resolve this issue. One way to address this is to control for predetermined characteristics such as gender, race, and income to alleviate some of these concerns as shown below.

$$Y_{ijsc} = f(R_{ijsc}) + g(T_{ijsc}) + \theta_{jsc} + X_i\beta + \epsilon_{ijsc} \tag{4}$$

However, it is unlikely that researchers can control for all relevant characteristics because many are likely to be unobserved.

Fortunately, there is a solution to passive sorting—researchers should make comparisons among classrooms of the same type, i.e. with similar test score distributions. If distributions are similar, then there is less scope for features of the distribution to be systematically correlated with rank, and leaving variation in rank to be more likely due to idiosyncratic variation.[18] To accomplish this, we modify our estimating equation by allowing $g(T_{ijsc})$ to vary by characteristics of the test score distribution, indicated by $g_d(T_{ijsc})$.

$$Y_{ijsc} = f(R_{ijsc}) + g_d(T_{ijsc}) + \theta_{jsc} + X_i\beta + \epsilon_{ijsc} \tag{5}$$

In doing so, this addresses passive sorting through making comparisons among similar distributions. If comparisons are among very similar distributions, conditional rank is unlikely to be correlated with features of the distribution. This is close to our ideal experiment where students only experience a different rank due to sampling variation in the test score distribution, rather than systematic differences, because in the ideal experiment, students are randomly assigned to classes. One way to achieve comparisons among similar distributions this is to classify distributions of student achievement into groups based on characteristics of the distribution (for example mean and variance) and interact $g()$ with indicators for these

[18] Note, if schools within a type all had exactly identical distributions, there would be no variation in rank for a given test score, we still require sampling variation to exist within types.

groups of distributions.[19] Because student achievement is measured in percentiles and is uniformly distributed, school mean achievement will be correlated to the skewness of the school distribution. Low mean achievement classes will typically have positive skewness, and high achievement classes typically have negative skewness. The skewness of a distribution can be important for passive sorting; therefore, allowing for the impact of achievement to vary by mean achievement could remain important even with SSC effects due to the correlation of the mean and skewness.

Interacting achievement with features of the class distribution is our preferred specification. The identifying assumption is a conditional independence assumption which is that $E\left[\epsilon_{ijsc} | g_d\left(T_{ijsc}\right), \theta_{jsc}, X_i\beta\right] = 0$.

2.4. Specification Error

The conditional independence assumption can be restated slightly. Consider the following equation where η_{jsc} is specification error.

$$Y_{ijsc} = f\left(R_{ijsc}\right) + g_d(T_{ijsc}) + \theta_{jsc} + X_i\beta + \eta_{ijsc} + \epsilon_{ijsc} \tag{6}$$

If there is any specification error, we must assume that it is unrelated to rank, $\eta_{jsc} \perp R_{jsc}$.[20] One example of a violation of this assumption is if different schools map current achievement to future outcomes differently. As an example, in a high-achieving classroom high test scores would lead to more four-year college enrollment and less two-year college enrollment. In low-achieving classrooms, high test scores might lead to more two-year college enrollment. If human capital has a different relationship to future outcomes in one school versus another but it is modelled as having the same relationship, this would introduce error into the model. If this error is correlated with rank then our estimate of the rank effect would be biased.

The solution to specification error is to allow the mapping of current achievement to future outcomes to be very flexible. Allowing the way current achievement is mapped to future outcomes to be flexible removes specification error from the relationship between rank

[19] Another way would be to estimate equation 5 separately by groups of distributions and aggregate the resulting coefficients.

[20] This assumption is necessary even when students of different abilities are randomly allocated to classrooms.

and future outcomes. Consequently, papers using this strategy should show robustness exercises where results are shown for different flexible specifications of achievement. Stability of estimates for different flexibly specifications of achievement is consistent with $\eta_{jsc} \perp R_{jsc}$.

Our preferred specification will address the problems in the estimation of the effect of rank. First, we include classroom fixed effects to address class level measurement error. Second, we demonstrate that active sorting on the basis of rank would be impractical, and if existent would likely downward bias the rank effect. Third, we show that passive sorting can be addressed by comparing similar distributions. Fourth, we show that misspecification can be addressed by flexibly controlling for the mapping of current achievement to future outcomes. While we describe each of these issues separately, in reality they are likely to be interrelated. For example, students could be passively sorted to schools which have different mapping functions to future outcomes. Notably addressing passive sorting, will not necessarily solve misspecification. We return to determining the exact specification of equation (6) in Section 4.

3. Data

The data we use in this study is the de-identified data from the Texas Education Research Center (ERC), which contains information from many state level institutions.[21] Data concerning students' experience during their school years cover the period 1994–2012, although the primary estimating sample will focus on 1995–2008. These data contain demographic and academic performance information for all students in public K–12 schools in Texas provided by the Texas Education Agency (TEA). These records are linked to individual-level enrollment and graduation from all public institutions of higher education in the state of Texas using data provided by the Texas Higher Education Coordinating Board (THECB).[22] Ultimately, these records are linked to students' labor force outcomes in years 2009-17 using data from the Texas Workforce Commission (TWC). This contains information

[21] For more information on the ERC see https://research.utexas.edu/erc/

[22] We do not observe out of state enrollment or enrollment at private institutions in Texas. Hence, if higher rank causes students to leave the state we will underestimate the effect of rank on college attendance.

on quarterly earnings, employment, and industry of employment for all workers covered by Unemployment Insurance (UI).[23]

3.1. Constructing the Sample

The sample used for this analysis consists of students who took their third grade state examinations for the first time between 1995 and 2008.[24] We focus on students taking their third grade exam for the first time to alleviate concerns regarding the endogenous relationship between class rank and previous retention. We focus on students taking their exams in English, rather than Spanish. During this period, the third grade students took annual reading and math assessments, although the testing regime changed.[25] Consequently, we percentilize student achievement by subject and cohort. This ensures that the test score distribution for each subject is constant and uniform for each cohort.[26] For each student, we generate a rank within their elementary school cohort for math and reading based on their test scores including those who had been retained.

We link students to subsequent outcomes including performance in reading and math in eighth grade. Because rank may affect grade retention, we estimate the impact on eighth grade test scores only for students who took the test on time. We also consider classes taken in high school, including Advanced Placement (AP) courses, and graduation from high school. We then consider whether students enroll in a public college or university in Texas (separately by two-year and four-year schools), if they declare a Science Technology

---

[23] Unemployment insurance records include employers who pay at least $1,500 in gross earnings to employees or have at least one employee during twenty different weeks in a calendar year regardless of the earnings paid. Federal employees are not covered. We do not observe out of state employment. Hence, if higher rank causes students to leave the state we will underestimate the effect of rank on earnings.

[24] Students are defined as taking their third grade exam for the first time if the student was observed not being in the third grade in the previous year.

[25] Until 2002, the Texas Assessment of Academic Skills (TAAS) was used. Starting in 2003, the Texas Assessment of Knowledge and Skills (TAKS) was used. The primary differences had to do with which grades offered which subject tests. This does not affect this study substantively as all students took exams in math and reading for 3rd and 8th grade.

[26] Murphy & Weinhardt (*forthcoming*) show that when estimating rank effects, conditional baseline achievement should be uniformly distributed to prevent the possibility of a certain type of measurement error (that increases multiplicatively further from the mean), generating a spurious rank effect.

Engineering or Mathematics (STEM) major, and whether they graduate from college. Lastly, we examine earnings and the probability of having positive UI earnings.

For binary outcomes, such as AP course taking, high school graduation, and college enrollment, we define the variable as 1 for the event occurring in a school covered by our data and 0 otherwise.[27] For earnings, we consider both average earnings including zeroes as well as excluding zeroes.

To maximize the sample, we consider as many cohorts as possible for each outcome. This means that we have more cohorts for outcomes closer to third grade and fewer cohorts for later outcomes.[28] For K-12 and initial college attended outcomes, we have 13 cohorts of students who took their third grade tests between 1994 and 2006, giving 6,117,690 student subject observations. For graduating college in four years we have 10 cohorts (1994-2003), totaling 4,573,672 student-subject observations. For graduating in 6 years and post college outcomes, for individuals aged 23-27 we have 8 cohorts (1994-2001), or 3,597,340 student-subject observations, and 6 cohorts, or 2,647,240 students, for graduating with a BA within eight years. This explains the discrepancy in sample size across different outcomes.

Table 1 presents summary statistics. The sample is 47 percent White, 35 percent Hispanic, and 15 percent Black. 70 percent of students in the sample eventually graduate from a Texas public high school. 46 percent of students attend a public university or college in the year after "on time" high school graduation.[29] Within three years of on time high school graduation, 23 percent attend a public four-year institution in Texas, and 31 percent attend a Texas community college. When students are 23-27 years old, 65 percent have non-zero earnings, where the average non-zero earnings is $24,818 in 2016 dollars.

3.2. Rank Measurement

We rank each student among their peers within their grade at their school according to their state percentile in standardized tests in each tested subject. Simply, a student with the highest test score in their grade will have the highest rank. However, a simple absolute rank measure

[27] As an example, for college enrollment a student who does not attend any college, or attends a college out of Texas will be coded as a zero.
[28] Results are similar when considering a consistent set of cohorts, see Appendix Figure 4.
[29] On time graduation is defined as graduation if a student did not repeat or skip any grades after third grade.

would be problematic, because it is not comparable across schools of different sizes. Therefore, like state test scores we will percentilize the rank score individual i with the following transformation:

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

where N_{jsc} is the cohort size of school j in cohort c of subject s. An individual i's ordinal rank position within this group is n_{ijsc}, which is increasing in test score. Here, R_{ijsc} is the standardized rank of the student that we will use for our analysis and is the percentile rank in their class (SSC cells). For example, a student who had the second best score in math 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 ties in test scores, each of the students with the same score is given the mean rank of all the students with that test score in that school-subject-cohort.[30] We will calculate this rank measure for each student within a test administration.

Note that this is our measure of the academic rank of a student within their class. Students will not necessarily be told their class rank in these exams by their teachers, nor do we believe that students care particularly about their ranking in these low-stakes examinations. Rather, we interpret our test score rank measure as a proxy for their day-to-day academic ranking in their class. Students learn about their rank through repeated interactions throughout elementary school with their class peers (e.g., by observing who answers the most questions or gets the best grades in assignments).[31] Similar arguments can be made for teachers or parents learning about the rank of students.

[30] Our main analysis limits the sample to students who took their third grade test on time. However, their actual classroom consists of students who are on time and students who are not. In Appendix Table 1 we show that our results are very similar when we calculate rank using all students and using different methods to break ties.

[31] We use rank in third grade as it is the earliest measure of ranking available. It is possible to estimate the impact of later grade ranks on outcomes. However, as we find that future test scores are increasing in previous rank, it means that later rank itself is an outcome and rank is self-perpetuating. Hence, it is difficult to determine when a child's ranking has the largest impact. Ideally, we would rank students based on a before school measure of achievement.

4. Model specification

As described in Section 2, it is imperative that the correct functional form is used to prevent specification error or passive sorting from causing a spurious rank effect. Therefore, a key choice is the proper way to model $g()$. There are two features that this function should meet; 1) make comparisons among similar classrooms to avoid passive sorting; and 2) model the relationship flexibly between current achievement and future outcomes flexibly to avoid specification error. Put another way, what specification satisfies the conditional independence assumption and yields rank that is conditionally as good as random?

A natural way to test for passive sorting similar to a balance test in a randomized controlled trial by replacing Y_{ijsc} in our specification with a predetermined characteristic. If the conditional independence assumption is satisfied, predetermined characteristics should not be correlated with rank. If observable, predetermined characteristics are not correlated with rank, then it plausible that unobserved predetermined characteristics are also not correlated with rank.

We first start with a specification akin to equation (4), omitting student characteristics, with a flexible function of achievement, $g(T_{ijsc})$ which does not vary by school. Specifically, we include nineteen indicators for all-but-one (tenth) ventiles of student achievement. This allows for non-linear relationships between prior test scores and future outcomes.[32] To allow for the comparison across classes with similar distributions we then interact student achievement with indicators for the schools location within the distribution of schools according to mean and variance in achievement, akin to equation (5).

Table 2 shows the correlations between student rank and the following student characteristics; Male, Free and Reduced Price Lunch, English as a Second Language, and Non-White. Each panel shows the correlations for a different specification; Panel A – constant impacts of ventiles of achievement; which then we allow to vary by; quartiles of mean school achievement (Panel B); deciles of school mean achievement (Panel C), quartiles of school variance in achievement (Panel D), deciles of school variance in achievement (Panel E), and

[32] In Figure 11 we show our results are robust to alternate functional forms. We chose ventiles as our main specification as we did not want to allow the rank function to have more flexibility than that of baseline achievement.

the interactions of the quartiles of school mean and variance in achievement (Panel F).[33] Panels A-E show that rank is correlated with predetermined student characteristics. This suggests that these specifications have issues with passive sorting. However, Panel F shows that interacting 16 indicators for school variance and mean quartiles eliminates any imbalance we have. Therefore, our preferred specification is

$$Y_{ijsc} = \sum_{r=1,r\neq 10}^{20} \mathbb{I}_n(R_{ijsc} = r)\,\rho_r + \sum_{D=1}^{16}\sum_{t=1,t\neq 10}^{20} \mathbb{I}_n\big(T_{ijsc} = t\big)\,\mathbb{I}_d(d_s = D)\mu_{nd} + \theta_{\text{jsc}} + X_i\beta + \epsilon_{ijsc} \qquad (7)$$

Where $\mathbb{I}_n(R_{ijsc} = r)$ denotes an indicator function, which takes a value of one when the ventile of class rank R_{ijsc} matches r. This allows for potential non-linearities in the effect of rank on later outcomes by estimating parameters (ρ_r) for each ventile in rank omitting the tenth ventile. We have defined $g_d(.)$ in a similar matter, by allowing for separate impact for each of the ventiles of baseline achievement T_{ijsc}, which can vary at the school distribution level, d_s. Schools are characterized by their quartile in the mean achievement distribution, and quartile of the variance achievement distribution, providing 16 school distribution groups, D.

Appendix Figure 2 shows the remaining variation in class rank for a given test score within each of these distributions. This figure plots the class rank of students against their achievement de-meaned by school within a subject, for each of these 16 types. The relationship between rank and achievement resembles that presented in Appendix Figure 1, but we now see that as the mean and the variance of classes change, the expected rank for a given test score changes. For example, distributions in the top row (with the smallest class variance) have a shallower gradient than those in the fourth row. Regardless, there still exists sufficient variation in rank conditional on achievement even within these distribution types.

In summary, using this specification (equation 7) we do not observe imbalance in observable characteristics, and so assume the remaining variation in rank to be orthogonal to

[33] Using school mean has two advantages. First, it allows a different mapping from achievement to future outcomes in different settings which deals with the issues of misspecification. Second, due to the bounded nature of test scores, mean is generally correlated the skewness of a distribution.

unobservable factors that determine our outcomes. We show that our main findings are robust to these choices in Section 6.1 below. We now present our main findings using equation 7.[34]

5. Results

We will primarily present results of equation 7 by plotting the estimate coefficients, ρ_n, along with the 95 percent confidence intervals. All estimates will be relative to the tenth ventile that includes students ranked from 45-50 in their class. Because there are many estimates for each outcome, we present the results visually.[35]

5.1. K-12 Outcomes

We first consider the probability of repeating third grade. Panel A of Figure 4 shows that lower ranked students are more likely to repeat third grade even after conditioning on achievement. The impact of rank on retention decreases as the rank of the student increases. For those in the lowest ventile of the rank distribution, they are six percentage points more likely to be retained, which is double that of the next highest ventile. Moving a student from being ranked last to being ranked in the lower 25th percentile reduces the probability for retention by roughly 5 percentage points. Given the mean retention rate of 1.6 percent, this represents a sizable shift how students are treated by their schools, independent of their human capital. Once students reach the 40th class percentile rank has no significant impact on retention. This suggests that at least some of the rank effect is likely coming from the way the school treats low ranked students.

We next examine the effect of third grade rank on achievement in eighth grade where achievement in eighth grade is measured in state percentiles. Figure 4 Panel B shows an approximately linear effect of rank in third grade on academic performance. Moving from the 25th percentile to the 50th percentile in rank (or from the 50th to the 75th) improves performance by approximately 2.5 percentiles. This is similar to the estimates in Murphy and Weinhardt (2014) that considers outcomes at comparable ages in England, finding the same change in

[34] Note that we observe two test scores for students' math and reading. For most outcomes we stack observations so that each student has two observations. For outcomes that are subject specific we estimate specifications that have both math and reading rank and achievement entered separately to investigate if rank in math and reading have different effects.

[35] The corresponding estimates and standard errors are available in table form upon request.

rank at the end of primary school (age 10/11) improves performance national test scores at age 13/14 by 1.9 percentiles. This is estimated on students who took the test in 8th grade "on time." However, as we have seen rank causes some students to be retained, it will also determine if students are in this estimating sample.[36] Hence, when interpreting the estimates of the effect on 8th grade test scores, one should bear in mind that the sample is selected on the basis of the treatment.[37]

The results on eighth grade test scores are not novel, but they do corroborate that similar rank effects occur in different educational systems, establishing the external validity of each estimate. Moreover, they provide a mechanism for the later outcomes we observe. In particular, student achievement in eighth grade is correlated with many outcomes including high school achievement, class taking, college enrollment and success, and labor market outcomes.

Our first high-school outcome we consider is whether a student takes Advanced Placement courses. In the first two panels of Figure 5 (A & B), we use our standard specification where the two observations for math and reading for each student are stacked, such that we are estimating the mean rank coefficients across subjects. We see that elementary school achievement rank positively impacts the probability of taking AP Calculus and AP English. In both cases, these effects are driven by being highly ranked among classmates during elementary school. The pattern and magnitude of the estimates are similar for taking both AP subjects, with students around the 75th percentile being around 2 percentage points more likely to take AP compared with the median student. Note, that the baseline rate for taking AP Calculus and AP English for our sample is 8.4 percent and 19.0 percent respectively. The exception that there is a discontinuous jump in the probability of taking AP Calculus if the student is in the top ventile of their elementary school. Students in the top ventile are 6 percentage points more likely to enroll in AP calculus compared to the median ranked student.

[36] Appendix Figure 3 that rank affects the probability of taking the eighth grade tests on time.
[37] To bound the estimates, one could consider that not taking the test on time as a the lowest possible score, then our estimates would be a lower bound of the effect of third grade rank on eight grade test scores.

The second two panels of Figure 5 (C & D) consider the effect of elementary school rank in each subject. To do so we control for achievement and rank in third grade math and reading separately and simultaneously. Panel C shows the impact on taking AP Calculus. A higher rank in math causes more students to take AP Calculus. Most of this effect occurs for students above the median in rank; whereas, below the median there are only small difference in the probability of taking AP Calculus. In contrast, a student's rank in reading has very little effect on taking AP Calculus for students with rank above or below the median. The final panel (Panel D) shows the impact of rank on taking AP English courses. As before, rank in Math has a stronger effect than rank in reading. However, here having a high rank in reading does positively impact the probability of taking AP English Courses. This is evidence that some rank effects are subject specific and some rank effects have spillover effects onto other subjects.

The final set of K-12 outcomes we consider is whether a student graduated from high school. The time frame we consider is within three years of "on time" high school graduation which is defined by nine years after their third grade to avoid issues of grade retention. In our sample, 70 percent graduate from high school by this definition. The impact of third grade rank on high school graduation can be seen in Figure 6 Panel A. A higher rank makes students more likely to graduate from high school. The effect is approximately linear, with impact of rank gradually declining as rank increases. Moving from the 25th to 75th percentile in rank increases the probability of graduating from a public Texas high school by 4 percentage points.

In summary, a student's rank in third grade independently affects grade retention, testing performance 5 years later, class selection, and ultimately graduation. As we examine longer term outcomes, these changes throughout schooling will be some of the channels that affect things such as college education and earnings.

5.2. College Outcomes

After high school students face the choice of entering post-secondary education. Figure 6 Panel B presents enrollment in any public college in Texas. Rank has a largely linear effect on the probability of enrolling in college in Texas. Moving from the 25th to 75th percentile in rank leads to an approximately 4 percentage point increase in college going. However, there is a

discontinuous fall in the likelihood of attending college if students were in the lowest ventile, compared to the penultimate ventile of 1.5 percentage points.

To understand this pattern better, we consider enrollment in two-year and four-year schools separately in Figure 6 Panels C and D. Figure 6 Panel C considers enrollment in two-year institutions. Having a low rank reduces the probability of going to community college; whereas high rank is not strongly correlated with attending community college. Again, there is a discontinuous fall in the likelihood of attending college if students were in the lowest ventile. Those in the bottom ventile are 4 percentage points less likely to enroll in community college compared the baseline enrollment rate of 31 percent. Figure 6 Panel D shows the impact on enrollment in a four-year institution. Enrollment is increasing in rank, and this is driven by primarily by students ranked above the median with the effect of rank increasing in rank. The contrast of the effects of enrollment by school type is likely driven by who is at the margin of attending a community college versus a four-year institution.

Once at college, students can declare a major. Given the significant returns to STEM majors we now estimate the ultimate impact of third grade rank on the probability of a student declaring a STEM as their first major. Figure 7 Panel A shows that there is a comparatively small positive relationship between students' rank in elementary school and their likelihood of choosing a STEM major which is driven largely by top ranked students.[38] Like AP choice in high school, major choice is likely to be impacted by students' rankings in particular subjects, and this previous estimate is the average effect of both reading and math ranks. To explore this, Panel B of Figure 7 presents the impacts of third grade math and reading rank on declaring a STEM major. Here we find that the relationship is entirely driven by math rank, with top ranked math students over 2 percent more likely to choose STEM, with the mean STEM uptake rate of 4 percent. In contrast, neither top or bottom reading rank impacts the probability of declaring a STEM major.

Continuing with post-secondary outcomes, the panels in Figure 8 considers graduation with a bachelor's degree within various time frames—4 years (Panel A), 6 years (Panel B), and

[38] We code students who declare a STEM major as 1 and students who do not (including students who do not enroll in college) as 0. Hence, estimates will conflate the effect of rank college enrollment and declaring a STEM major.

8 years (Panel C) after "on-time" graduation from high school We find that students with higher rank in elementary school are more likely to graduate with a bachelor's degree. Consistent with the effects on enrollment, the effect of rank is largely driven by rank above the median. The impact of being top of class in terms of percentage points is smallest for graduating within 4 years at 2.4 percentage points relative to the median ranked student, compared to 3.4 and 3.3 percent for six and eight years respectively.

5.3. Labor Market Outcomes

We consider the effects of rank on a range of labor market outcomes. First, we examine employment outcomes for students ages 23-27 (or 15-18 years after third grade). We consider average annual earnings between the ages for eight cohorts of students who took their third grade examinations between 1994-2001 (employment years 2009-2016).[39] The first panel in Figure 9 considers the probability of having positive earnings from age 23-27. A value of 1 would mean having positive recorded in each of these years, a value of 0 would mean having no recorded earnings between ages 23 and 27. The pattern in Panel A suggests there is little effect of rank on the probability of having positive earnings except for the lowest ranked students who are 1 percentage points less likely to record positive earnings than the median ranked student.[40]

The remaining panels of Figure 9 refer to the effect of rank on earnings. Panel B shows that increasing rank increases average annual earnings between the ages of 23-27. Low ranked students have meaningful earnings penalties, earning $1,850 per year less than the median ranked student. High ranked students see increases in earnings as well with the highest ranked students earning $1,200 more per year than the median student. The lower two panels show the effect on non-zero earnings and log earnings (Panels C and D respectively). Conditional on having positive earnings, we find the negative impact of having a low rank is exacerbated. Students at the bottom now earn $1,250 less per year, compared to students in the middle of the rank distribution. Taking the log of average annual earnings shows that rank

[39] Note, not all individuals have four years over which to average employment. Those from the most recent cohort only have one year of labor outcomes, for instance. We average over all years from age 23 to 27 for which we have earnings. Appendix Figure 4 Panel D, shows the impact on log earnings for the 8 cohorts of student for which we have all five years of earnings (who took third grade 1994-2001).
[40] This shows rank does not cause differential attrition from our earnings sample.

affects the log of average earnings throughout the distribution of rank. Taken together, a student's rank in third grade affects labor market outcomes. Moving from the 25th percentile to the 75th percentile in rank causes log earnings to increase by approximately 7 log points.

6. Robustness

In Section 2 we discussed the concerns relating to correctly specifying the functional so that omitted factors are not driving the results. In section 4 we used balancing regressions to empirically inform the choice of functional form for our regressions. In this section we show that our results are robust to several alternative specifications, as well as to different samples, rank definitions and measurement error in baseline ability.

6.1. Functional form choices

First, we model the relationship between achievement and outcomes using various functional forms. Figure 10 presents results for four main outcomes, eighth grade test scores, graduation from high school, the probability of attending any college, and log earnings. The point estimates are displayed for various functional forms of student achievement, interacted with sixteen indicators for school mean and variance quartiles. In addition to our main specification (which controls for achievement using ventiles in student achievement interacted with indicators for classroom distribution type), we model student achievement using various polynomials from first order to a sixth order interacted with indicators for classroom distribution type. The results are substantively similar once achievement is controlled for with a quadratic—the magnitude does vary somewhat across specifications, but the direction of the relationship is consistent.[41] Hence, our results that rank has a lasting impact on student outcomes are not dependent on the exact functional form chosen to model achievement.

In addition to choices of non linearity in $g_d()$, we also allow of the impact of prior achievement to vary by the school achievement distribution. We have shown in Section 4 and

[41] The fact that only allowing for a linear impact of percentile rank leads to different estimates may be reflective of the pre-determined human capital distribution being normally distributed, which we have transformed into a uniform distribution through percentalization.

Table 2 that the interaction of achievement with 16 indicators for school mean and variance, results in observable covariate balance. As a result, we use this as our main specification. However, we show that our results are not dependent on this specific functional form $g_d()$ for most outocmes.

Appendix Table 2 repeats the six interactions of achievement with ways of classifying school distributions and shows how these impact the estimates on outcomes. Looking at either end of our outcomes chronologically, we can see that the positive impact of rank on repeating 3rd grade, 8th grade test scores, high school graduation and log wages remains regardless of exact specification. The coefficients tend to increase in magnitude as we allow for more flexible interactions with achievement. This is consistent with the imbalance in rank found in Table 2, where more disadvantaged students appeared to have higher rank conditional on achievement, and this imbalance decreased once we allowed for more flexible functional form. Despite these slight differences between specifications the estimates are largely consistent.

Two outcomes are different when comparing our preferred specification to the other specifications, any college and graduation with a bachelor's degree. In both cases, the coefficient on rank becomes negative if we do not allow the impact of achievement to vary by test score distribution. We investigate the sources of this bias in Appendix Figure 5. For the uninteracted specification where there is imbalance, estimates of the effect of rank are very similar when controlling for predetermined characteristics relative to not controlling for them. This suggests that imbalance in predetermined covariates is not generating this bias in the estimates: This highlights the importance of functional form, because the bias in these estimates likely arises due to misspecification of the $g_d()$ function, where there are different mappings of achievement to outcomes across schools. While the imbalance in student characteristics is not driving the bias we observe, imbalance may be correlated to school unobservables which impact the mapping into future outcomes. Therefore, while misspecification and balance are independent concerns, they may often be linked.

6.2. Miscellaneous robustness checks: school size, breaking ties, measurement error

One data limitation is that we do not observe which classroom students are taught in for third-grade students. Hence, our main results use fixed effects for school-subject-cohort (SSC), which we have been referring to as a class. Moreover, given that we are using the

sampling variation in the test score distribution one may be concerned that the results are driven by schools with smaller cohorts as they would experience more sampling variation. Therefore, as a robustness check, we estimate the effects separately by the number of students in the third-grade school-subject-cohort. Figure 11 presents the point estimates for under 30, under 50, 50-75 students, 75-100 students, and 100+ students in an SSC. While the effects tend to be larger for the larger schools, the pattern of results remain regardless of school size. Moreover, the students in schools with fewer than 30 students per school-subject-cohort are likely to all be in the same class during third grade, and the results hold.

In our main specification, we handle ties in rank by assigning students the mean of the rank. We consider other methods including breaking ties including assigning the lowest rank, randomly breaking ties, and a rank only among students who are "on-time" in third grade. Our results are qualitatively similar regardless of our method of dealing with ties. The results tend to be slightly smaller when we break ties randomly, which we attribute to the introduction of noise into our measure of rank (See Appendix Table 1).

We use third-grade student achievement on externally graded state examinations as a measure of human capital of each student. The rank of each student within their class is derived from their test score and all the test scores of their class peers. This means that any individual non-systematic measurement error in student test scores has impacts on the two key explanatory variables rank and achievement. [42] Moreover this measurement error will be correlated. For example, if a student 'got lucky' in an examination, we would record them as having a higher level of human capital and potentially a higher rank than they actually have. Note that measurement error in a test scores does not necessarily mean there will be measurement error in rank, because it may not cause the rank of the student to change if they are sufficiently far (in achievement) from other students. The concern therefore in this situation is that again our rank measurement is simply acting as a less noisy measure of human capital. In addition to this situation, any measurement error in peers test scores will generate additional measurement error in the rank measurement. The combination of these factors makes this a non-standard measurement error problem.

[42] We address any systematic measurement error at the class level, due to class level shocks, with the inclusion of class fixed effects. This is critical as such shocks would be rank preserving.

Murphy and Weinhardt (*forthcoming*) address the same problem and show that additive random noise would non-linearly downward bias the effect of class rank depending on the extent of the measurement error. To reaffirm this result, we follow their approach and add normally distributed random noise to our baseline measure, re-compute the student ranks based on this (more) noisy measure and estimate our main results. Appendix Figure 6 shows the resulting estimates for our main outcomes, where we have added normally distributed noise with zero mean and a standard deviation equivalent to 10%, 20% and 30% of the standard deviation in the 3rd-grade achievement. We find comparable reductions in effect sizes with increased measurement error. The intuition for such a downward bias is that the 'lucky student' will have higher baseline achievement and rank than they should have, we will therefore observe these students (who are artificially highly ranked) having a lower growth in test scores than expected.

6.3. Heterogeneity

In this section we explore the heterogeneity of the rank effects the four main outcomes; (1) eighth grade test scores; (2) graduating high school; (3) enrolling in college; and (4) log average earnings ages 23-27. We explore this for three pre-defined variables; race (white/non-white), gender (male/female), and free and reduce price lunch eligibility in third grade (eligible/non-eligible). We present the estimates for each of these categories in Figures 12, 13, and 14, respectively.

First, is the impact of third grade rank different for white and non-white students? Figure 12 shows that the effects of rank on eighth grade test scores are similar by race. In contrast, the effects of rank on high school graduation are more pronounced for non-white students, for both low and high ranks. There is a more distinct difference when considering the impact on college attendance by race. Having a low rank in elementary school has a considerably larger negative effect on college attendance for non-white students compared to white students. For the lowest ranked non-white student they are 7.5 percentage points less likely to attend college than the median ranked student, in contrast the lowest ranked white students are only 1 percentage point less likely to attend. Once above the median rank, white and non-white students react in in a similar manner to their rank. We find that the effects of rank for

low and high ranks, are larger for non-white students, similar in pattern to high school graduation.

In contrast to the large differences by race, there is no evidence for heterogeneity with respect to gender for test scores, high school graduation, and college attendance. (Figure 13). This is different than Murphy and Weinhardt (*forthcoming*), who show that boys' high school test scores have a more positive reaction to high elementary rank, and a less negative reaction to having a low rank. However, we find that males experiencing a low rank does have a less negative impact on earnings as compared to females.

Finally, the heterogeneity of estimates with regards to FRPL students mirror those of white/non-white students, with the more disadvantaged group being more effected by rank than their counterparts (Figure 14). This is true for test scores, graduation, college enrollment and earnings. This is similar to with Murphy and Weinhardt (*forthcoming*) who find that Free School Lunch students gain more form being highly ranked in England. One explanation is that these sets of students have low academic confidence or a different information set about the achievement and weight their school experience more heavily than non-disadvantaged students. This has important consequences for optimal classroom composition, which we discuss below.

7. Implications on School Choice

We estimate the effects of rank net of SSC fixed effects. Traditional peer effects suggest that better peers should help performance. However, we show that having more better peers also has a negative effect by lowering rank. In this section we quantify the effects of rank as compared to the benefits of having better peers and school environments, and how these effects relate to other findings in the literature.

Consider the following thought experiment. A parent may move their child to a "better" school. This would come with a decrease in their child's rank and a likely increase in the quality of their child's peers. What would be the net gain in test scores from such a move?

To operationalize this, we categorize elementary schools into "good", "bad", and "average" with two different measures of quality, mean achievement and value-added. First, mean third grade achievement is simply third grade achievement on indicators for a student's

third grade school (Columns 1-4 Table 3). Each fixed effect captures many things including, student's own human capital, peer effects, resource differences, and parental investments. Each of these will reflect the average difference in third grade test scores between primary schools and should not be considered causal, although they may reflect what parents consider when choosing an elementary school, as they are relatively easy to observe. For our second quality measure, we calculate elementary schools' third to eighth grade unconditional value added by controlling for students' third grade achievement (Columns 5-8 Table 3). The standard deviation of these school value-added measures is .055. So, if a student moves to a one standard deviation better school, she receives a bump of .055 in eighth grade test scores.

To gauge the benefit of attending an elementary school with better attainment, we also record the mean value added of "good", "bad", and "average" schools in terms of attainment. It appears that schools with higher mean achievement also have higher third-to-eighth grade value added, although these gains are only half the size compared to if parents were selecting schools on the basis of value added: 0.026 versus 0.055 (final row of Table 3).

To ascertain the net benefits of attending these schools net of rank effects, we need to consider how a student's rank would change. The values in the columns of Table 3 are the mean rank of students in each state achievement ventile in each school type. For example, consider the median student at the tenth achievement ventile. If they attended an average elementary school in terms of third grade achievement, their expected class rank would be 0.479. Whereas if they attended a "bad" school, they would have an expected rank of 0.613 and 0.346 if they had attended a "good" school. We can see that there is a clear trade off in terms of rank and the quality of school when measured in absolute achievement. As may be expected, this rank-quality tradeoff is higher when elementary school quality is measured in mean achievement compared to value added. When parents move their child from a "bad" value added school to a "good" value added school, the loss in rank is only -0.155. If they instead use mean achievement the change in rank is larger at -0.267 (Table 3 row 10).

We can see that students from higher-up in the state achievement distribution also have higher ranks in their classes. There is not much difference in terms of expected rank at an average school, independent of whether school quality is measured in terms of value added or absolute attainment throughout the achievement distribution. However, this thought

experiment clearly shows that the decrease in rank from moving from a bad to good school is always smaller when considering schools in terms of their value added. Moreover, the loss of rank from attending a 'better' school is largest for the students near the middle of the distribution.

How would these changes in ranks and school environments impact a student's overall attainment? For this, we require one last piece of information, ρ from Equation 1, the relationship between rank and eighth grade test score, which is 0.09. We can now calculate the impact on test scores.

Let us consider the case of parents of a median student (tenth ventile) considering good or bad schools in terms of mean third grade achievement. Sending the child to the better school, would lead to an increase of eighth grade test scores by 6.1 percentiles (using the associated value added scores from the bottom of Table 3 (0.026-(-0.035)). However, sending the child to the better school would reduce the students expected rank by 0.267. This would reduce the student's eighth grade test score by 2.4 percentiles (0.09*0.267). Therefore, the rank effect has reduced the gains from attending a school two standard deviations better by 39 percent, resulting in a net gain would be 3.7 percentiles.

Alternatively, if parents were better informed and selected elementary schools on the basis of value added then there would be a smaller trade off in class rank (0.15) and larger increases in future test scores (0.11=0.055-(-0.055)). In this case, a median student attending a good school rather than a bad school would gain 11 percentile points in the eighth grade test score distribution. In contrast, the student would lose 1.4 percentiles due to their lower rank (-0.014=-0.15*0.09). Hence, if parents were to choose on the basis of value added, there would be a net gain of 9.6 percentiles.

In the case parents choosing on the basis of value added, the effect of school quality is roughly eight times the size of the rank effect. Note, while this rank effect is relatively small, this school quality measure encapsulates *all* observable and unobservable factors that contribute to student value added.

Ideally, we could describe the contribution of rank relative to all other peer effects; however, estimating all other types of peer effects is implausible in our setting. Moreover,

while choosing the best school for their child in value-added is what most parents try to achieve, value-added is difficult to observe. Basing school choice on the basis of mean achievement, we see that the importance of class rank is substantial, reducing any perceived benefits by up to 39 percent. The main message for parents weighing the tradeoff of rank is that choosing schools based on value-added is the best strategy.

Finally, what does the presence of rank effects mean for optimal classroom composition given the heterogeneity of the effects? We find that disadvantaged groups such as non-white students or students eligible for FRPL are more affected by rank. Therefore, unlike linear in means peer effects, where moving students between groups would have no net impact, rearranging students with rank in mind could improve overall outcomes. This would involve creating groups of students such that disadvantaged students predominantly have higher ranks than non-disadvantaged students. Given the heterogeneity of the results the disadvantaged students would gain more from the higher rank, than the non-disadvantaged students would lose from having a lower rank. However, this sort of exercise warrants caution because the changes in classroom distribution may be out of sample for the estimates in this paper (Carrell, et al., 2013).

This finding illustrates a tradeoff for programs that move disadvantaged third grade students into situations where they will be the lowest ranked student. The extensive literature on selective schools and school integration has shown mixed results from students attending selective or predominantly non-minority schools. There are generally two quasi experimental approaches used for the impact of getting into a "high-quality" school, lottery admissions and regression discontinuities around admission thresholds. With respect to our paper the key difference between these approaches is that only in regression discontinuities does attending a "higher quality" school necessarily come at the at the cost of a lower rank. Being admitted to a high value-added school via lottery would not necessarily cause a decrease in rank, but going to a selective school would be. In light of this we now discuss our findings with relation to the existing literature on school quality.

The use of regression discontinuities in estimating the impact of attending an academic selective school ensuring students are similar across the boundary. The preponderance of evidence is that there are no benefits in terms of academic achievement, despite large

improvements in peer quality (Angrist & Lang, 2004; Clark, 2010; Kling et al., 2007; Dobbie & Fryer, 2014 , Abdulkadiroğlu 2014)[43]. In crossing the threshold, in addition to other school factors, at least two types of peer effects would at play. Students have better peers on average, but will have a lower rank. One can think of this as attending a 'better school' as defined by mean achievement in our policy example, but more extreme. In this case the marginal would mechanically go from highest ranked in one school to the lowest ranked in another.

The effectiveness of schools has also been estimated through lottery admissions to oversubscribed charter schools and voucher programs. In contrast to studies that use academic cutoffs, lottery admissions to charter schools do not necessarily pose a rank/school quality trade off, in fact in some cases students may gain in rank and quality. We can think of this as a more extreme version of attending a better school, as defined by value added.

Entry to an oversubscribed charter school often improves student performance (Angrist et al 2012, Angrist et al 2013, Abdulkadiroglu, 2011). Critically, attendance at these high-quality schools may not come at the expense of rank if they are recruiting low attaining students. This can be seen in the heterogeneity of charter school effects for low achievement students, they are positive in lower achieving districts (Angrist et al 2013), and are decreasing in the achievement of peers (Abdulkadiroglu 2011, Cullen et al. 2006). In summary, our results on class rank can help reconcile the different measured effects of high-quality schools across studies. Essentially, rank may "undo" the benefits of attending a high-quality school for students who admitted students who will be the lowest ranked students in their class.

8. Conclusions

We make two contributions in this paper. First, we discuss identification of the effects of class rank. In particular, we discuss active sorting, passive sorting, and misspecification of the achievement mapping as potential sources of bias in this literature. We provide guidance on how to address these issues and reinforce the utility of balance and specification checks. Notably, the issues of misspecification and measurement error also apply to the literature that

[43] There are improvements in non-academic outcomes or college attendance

estimates effects of rank using experiments that randomly allocate students of different abilities into classrooms.

Second, we demonstrate that a students' rank among their peers at a young age has long lasting impacts. This affects a student's performance in school including tests, courses taken, progress through toward graduation. Ultimately, it also affects student graduation from high school. Relative position affects the decision to enroll in post-secondary education. Most strikingly, it affects a student's real earnings in their mid-twenties. We find that a student enrolling in a class where they are at the 75th percentile rather than 25th in third grade increases their real wages between ages 23 and 27 by $1500 per annum, or approximately 7 percent. [44] For comparison, Carrell et al. (2018) look at the long run impact of peers at the same ages, and find that being in a class of 25 with a student who was exposed to domestic violence reduces an individual's earnings by 3 percent.

Our findings add to a growing list of papers that demonstrate conditions for young children have long lasting consequences. In contrast to other papers that focus on policy differences that students face, we document the effect of an unavoidable phenomenon in groups: relative rank. Some of the effect of rank may be coming via teachers and administrator interactions with students. We document that students are more likely to be retained in third grade which is a decision made not by the student but by teachers, administrators, and families.

Moreover, we find that disadvantaged groups gain more from being highly ranked and lose more from being lowly ranked among their peers. Therefore, unlike linear in means peer effects, where moving students between groups would have no net impact, grouping students with rank and these heterogeneous effects in mind could improve overall outcomes. Schools may also influence student performance by manipulating the salience of rank.

Finally, we examine if and to what extent parents should consider rank effects when choosing the best school for their children. Critically, we document a trade off from attending

[44] Using the present discounted value of earnings of $522,000 as in Chetty et al. (2014), which follow Krueger (1999) in discounting earnings gains at a 3 percent real annual rate, we calculate that these rank differences would increases life time earnings by $36,540 in net present value. This figure is based on the point estimates from Chetty et al. (2014) Figure 10 Panel 5. The 5th ventile has a coefficient of -0.032 and the 15th ventile has an estimate of 0.035.

a school with high achieving peers: this mechanically lowers the rank of your own child. We examine this trade off in detail based on the observed student and school allocations in Texas. We find that rank offsets about 40 percent of the benefits of school value added for the median performing student, if parents choose schools based on mean peer achievement. Instead, if parents choose schools based on value-added, the offsetting effects of rank from attending a better school are much smaller.

Future research on rank should focus on the interaction between rank and policies that exaggerate or mediate the effects of rank. Future research should also consider the effect of rank in groups outside of school settings.

References

Andrews, R. J., Li, J., & Lovenheim, M. F. (2016). Quantile treatment effects of college quality on earnings. Journal of Human Resources, 51(1), 200-238.

Abdulkadiroğlu, A., Angrist, J., & Pathak, P. (2014). The elite illusion: Achievement effects at Boston and New York exam schools. Econometrica, 82(1), 137-196.

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

Azmat, G., Bagues, M., Cabrales, A. & Iriberri, N. (2019). "What You Know Can't Hurt You, For Long: A Field Experiment on Relative Performance Feedback" Management Science, Vol. 65, No. 8, p. 3714-3736.

Azmat, G. & 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), p. 435–452.

Bergman P. (2018). "The Risks and Benefits of School Integration for Participating Students: Evidence from a Randomized Desegregation Program" IZA DP No. 11602.

Bietenbeck, J. (2020). "The Long-Term Impacts of Low-Achieving Childhood Peers: Evidence from Project STAR" Journal of the European Economic Association, Vol. 18, No. 1 (February 2020), p. 392-426.

Black, S., Devereux, P. & Salvanes K. (2013). "Under Pressure? The Effect of Peers on Outcomes of Young Adults" Journal of Labor Economics, Vol. 31, No. 1 (January 2013), p. 119-153.

Black, S., Devereux, P., & Salvanes K. (2011). "Too young to leave the nest? The effects of school starting age" Review of Economics and Statistics, 92(2), p. 455—467.

Booij, A., Leuven, E. & Oosterbeek, H. (2017). "Ability peer effects in university: Evidence from a randomized experiment" Review of Economic Studies, 84(2), p. 547-587.

Brown, G., Gardiner, J., Oswald, A. & Qian, J. (2008). "Does wage rank affect employees wellbeing?" Industrial Relations: A Journal of Economy and Society, 47(3), p. 355–389.

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

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

Carlson, N.R., Enzie, M. E., Buskist, W. & C. D. Heth, (2005). "Psychology: the science of behavior" (3rd Canadian ed) Pearson Canada.

Carrell, Scott and Fullerton, Richard L. & West, James, (2009). "Does Your Cohort Matter? Measuring Peer Effects in College Achievement" Journal of Labor Economics 27(3), p. 439-464.

Carrell, S. E., & Hoekstra, M. L. (2010). "Externalities in the classroom: How children exposed to domestic violence affect everyone's kids." American Economic Journal: Applied Economics, 2(1), p. 211-28.

Carrell, S. E., Hoekstra, M., & Kuka, E. (2018). "The long-run effects of disruptive peers" American Economic Review, 108(11), p.3377-3415

Carrell, S. E., Sacerdote, B. I., & West, J. E. (2013). "From natural variation to optimal policy? The importance of endogenous peer group formation" Econometrica, 81(3), p. 855-882.

Chetty, R., Friedman, J., Hilger, N., Saez, E., Schanzenbach, D.W. & Yagan, D (2011). "How does your kindergarten classroom affect your earnings? Evidence from Project STAR" The Quarterly Journal of Economics, 126(4), p. 1593-1660.

Chetty, R., Friedman, J., & Rockoff, J. (2014). "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood" American Economic Review, 104(9), p. 2633-79.

Chetty, R., Hendren, N., & Katz, L. (2016). "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment" American Economic Review, 106(4), p. 855-902.

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

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

Cullen, J. B., Long, M. C., & Reback, R. (2013). "Jockeying for position: Strategic high school choice under Texas' top ten percent plan" Journal of Public Economics, 97, p. 32-48.

Delaney, J., & Devereux, P. J. (2019). The Effect of High School Rank in English and Math on College Major Choice.

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

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

Elsner, B., & Isphording, I. E. (2017b). "Rank, Sex, Drugs and Crime" Journal of Human Resources, 0716-8080R.

Elsner, B., Isphording, I.E. & Zölitz, U. (2018). "Achievement Rank Affects Performance and Major Choices in College" SSRN working paper.

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

Greenwood, Michael J. (1997). "Internal migration in developed countries." Handbook of population and family economics 1, p. 647-720.

Hastings, J., Kane, T. J., & Staiger, D. O. (2009). "Heterogeneous preferences and the efficacy of public school choice" NBER Working Paper 2145, p. 1-46.

Hoxby, C. (2000). *Peer effects in the classroom: Learning from gender and race variation* (No. w7867). National Bureau of Economic Research.

Hoxby, C. and Weingarth G. (2006). "Taking race out of the equation: School reassignment and the structure of peer effects" Working paper, Harvard University.

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

Kendler, S. K,, Turkheimer, E., Ohlsson, H., Sunquist, J. and K. Sunquist (2015). "Family environment and the malleability of cognitive ability: A Swedish national home-reared and adopted-away cosibling control study" Proceedings of the National Academy of Sciences 112(15), p. 4612-4617.

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

Levy, D. (2008). "Peer effects and alcohol use among college student" The Journal of Economic Perspectives 22(3), p. 189-206.

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

Luttmer, E. F. P. (2005). "Neighbors as negatives: Relative earnings and well-being" The Quarterly Journal of Economics, 120(3), p. 963–1002.

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

Murphy R. & Weinhardt F. (2014). "Top of the Class: The Importance of Ordinal Rank" CESIFO Working Paper NO. 4815.

Murphy R. & Weinhardt F. (forthcoming). "Top of the Class: The Importance of Ordinal Rank" The Review of Economic Studies, https://doi.org/10.1093/restud/rdaa020.

Oreopoulos, P., Stabile, M., Walld, R., & Roos, L. L. (2008). "Short-, medium-, and long-term consequences of poor infant health an analysis using siblings and twins" Journal of Human Resources, 43(1), p. 88-138.

Payne, A. & Smith, J. (2018). "Big Fish, Small Pond: The Effect of Rank at Entry on Post-Secondary Outcomes" Mimeo.

Ribas, R. P., Sampaio, B., & Trevisan, G. (2020). "Short-and Long-term Effects of Class Assignment: Evidence from a Flagship University in Brazil" *Labour Economics*, 101835.

Ribas, R.P., Breno, S., and Giuseppe, T. (2018). "Can Better Peers Signal Less Success? The Disruptive Effect of Perceived Rank on Career Investment" SSRN working paper, http://dx.doi.org/10.2139/ssrn.3135824.

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

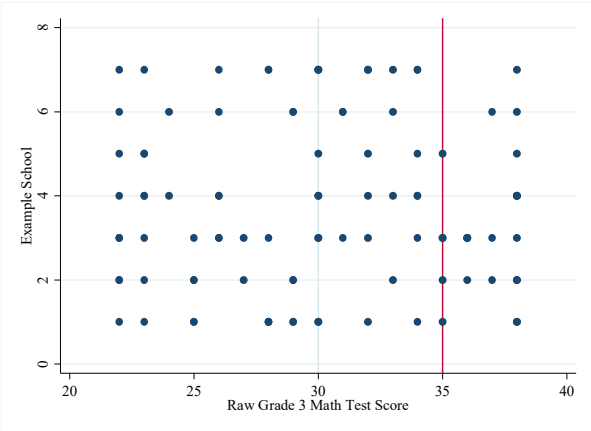
Tincani, M. M. (2017). "Heterogeneous peer effects and rank concerns: Theory and evidence." CESifo Working Paper Series 6331.

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

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

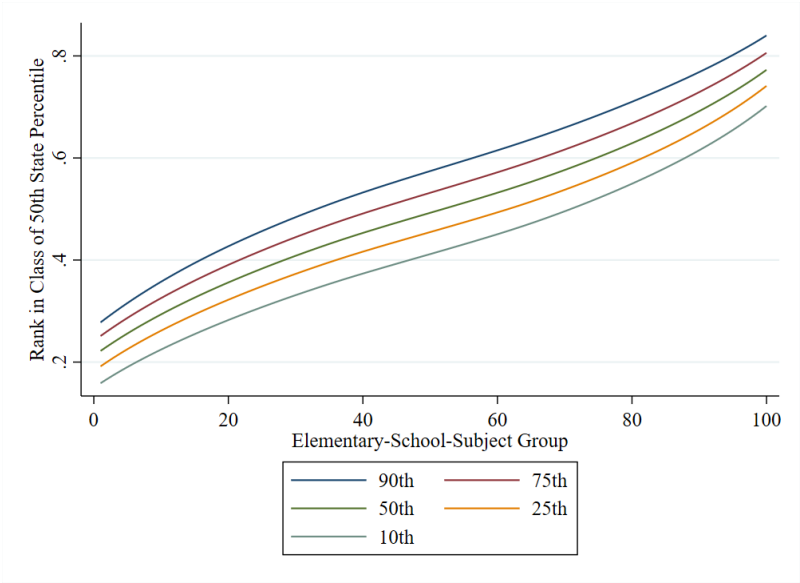
Zax, J. S., and Rees, D. I. (2002). IQ, academic performance, environment, and earnings. *Review of Economics and Statistics*, *84*(4), 600-616.

Figure 1: Class Test Score Distributions



Note: These figures are based on hypothetical data, based on schools in Texas. This example shows classrooms with the same min, max, and mean scores, presenting seven with the same mean, max, with students who have the same achievement having different rank.

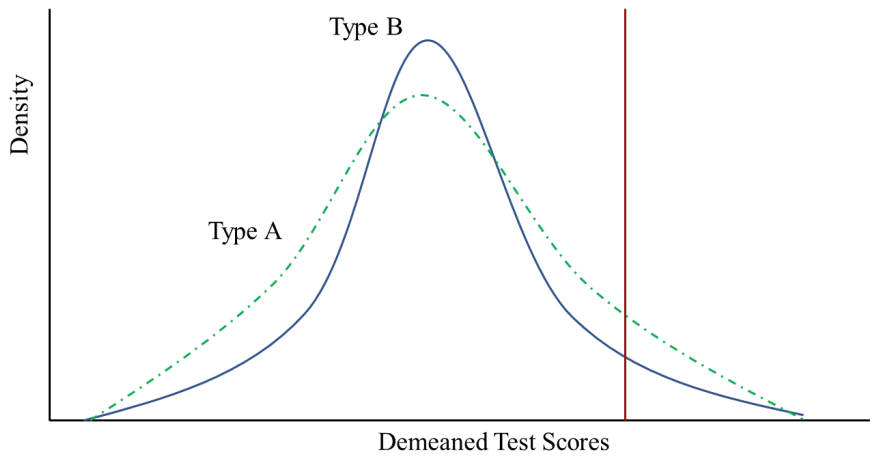
Figure 2: Variation of Local Rank of Median Student Within School-Subject Groups



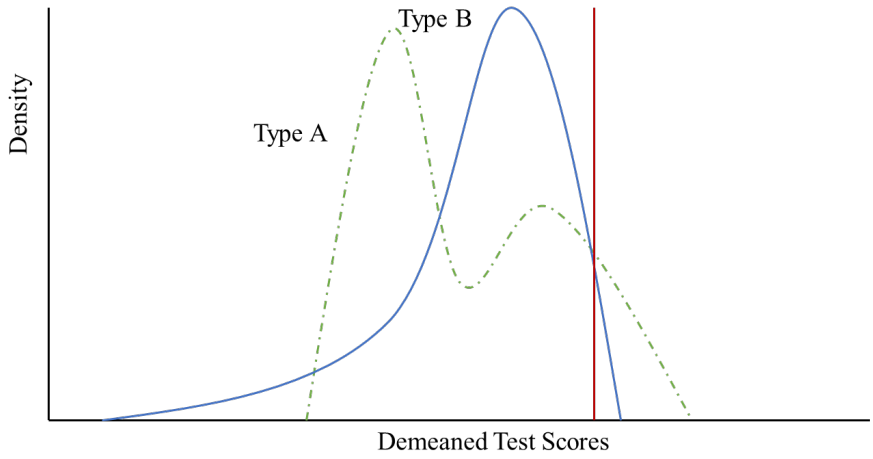
Note: This figure plots the distribution of rank for students at the 50th percentile of the statewide achievement distribution at different schools-subject groups. The horizontal axis plots elementary school-subjects as percentiles so that the school with the lowest rank for the statewide median student has a value of 1 and the school with the highest rank for the statewide median student has a value of 100. The variation in class rank for the state median student comes from the 13 cohorts. For each school-subject percentile, we plot the class rank of 10th percentile of students with state median achievement. The equivalent lines are drawn for the 25th, 50th, 75th, 90th percentiles.

Figure 3: Passive Sorting and Rank

A. High/Low Variance



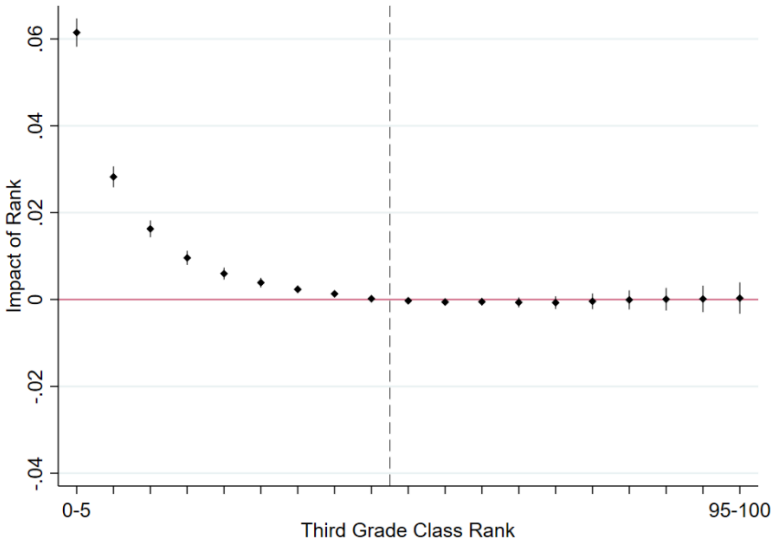
B. Non-Standard Distribution



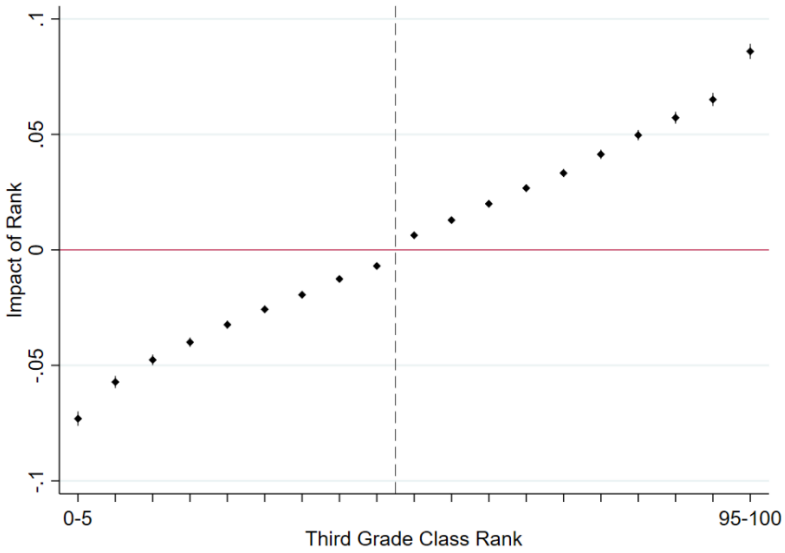
Note: This figure displays several hypothetical test score distributions to illustrate the issue of passive sorting. Each panel shows the test score distribution of two types of schools, Type A and Type B, that have been transformed such that they have the same mean. The red line indicates an arbitrary test score of students a set distance from the type average. In both panels, due to the shape of the test score distributions students with the same relative to the mean test score the student would have a higher rank in in Type B compared to Type A. Passive sorting can occur if students with different characteristics systematically attended schools with different distribution types. Passive sorting would then cause a bias if the characteristics directly impact future outcomes.

Figure 4 – Third Grade Rank on K-12 Outcomes, 1

A. Repeat Third Grade

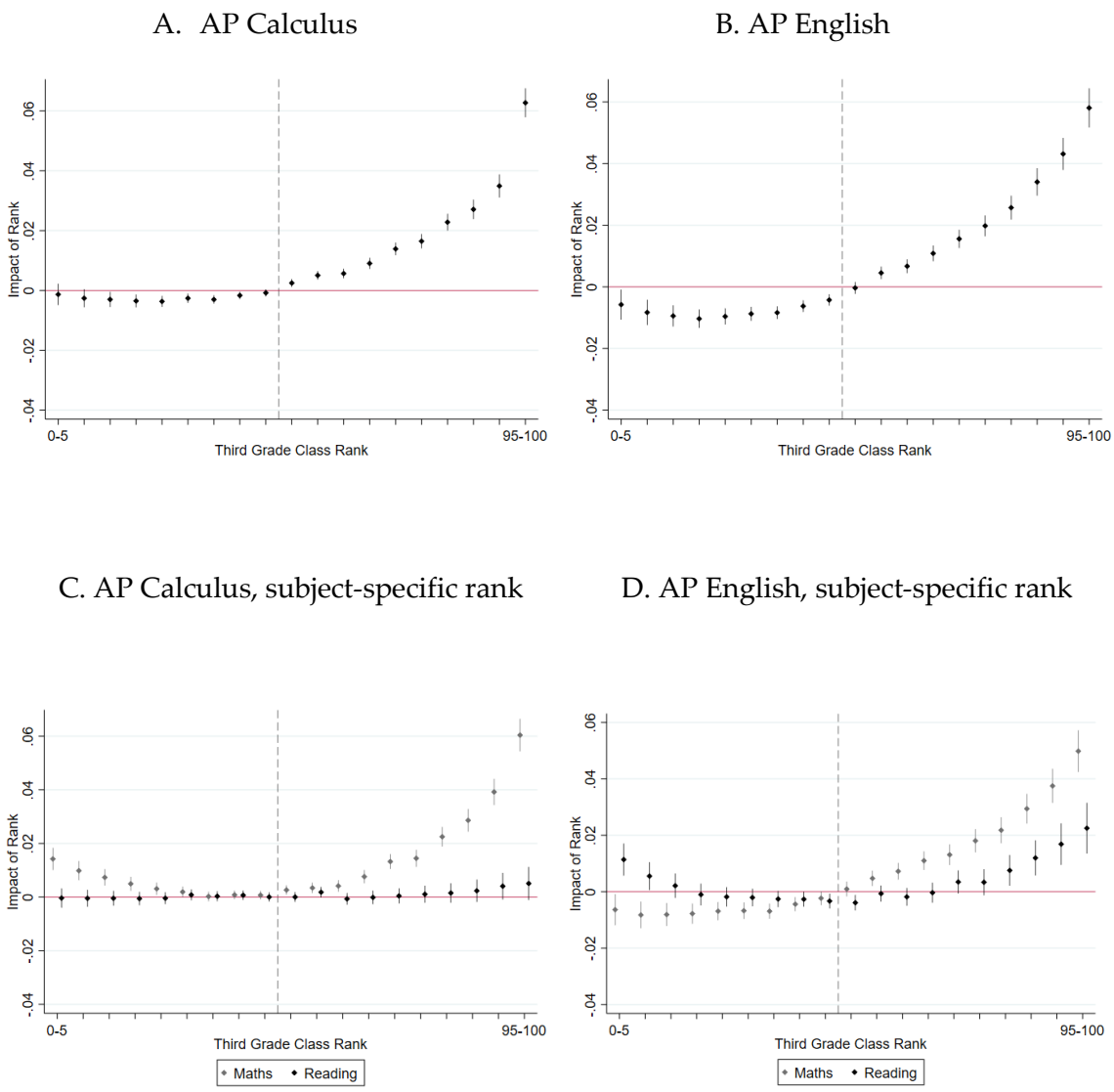


B. Eighth Grade Test Scores



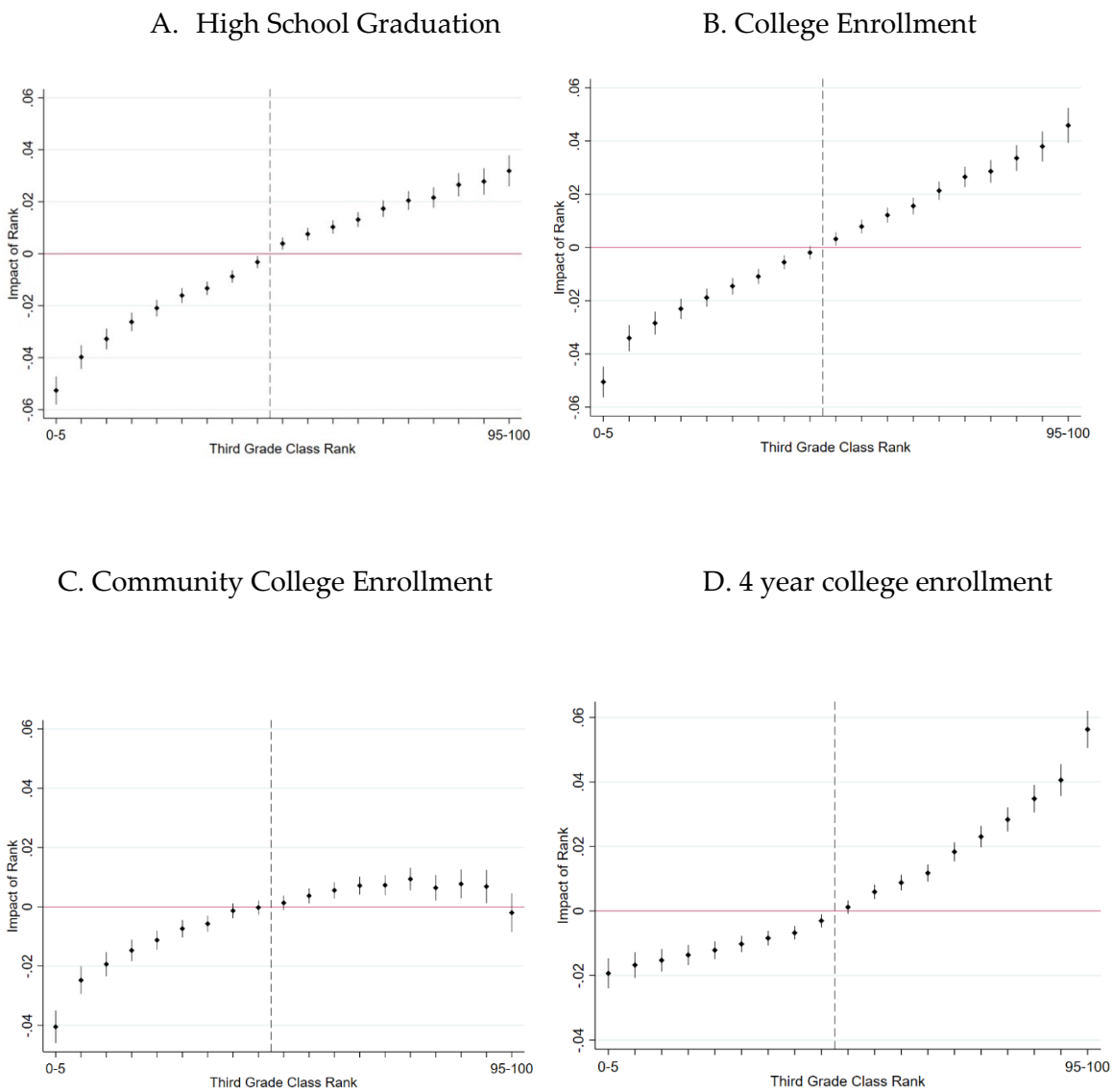
Note: These figures plot the coefficient for ventiles of class rank with 95% confidence intervals calculated using standard errors clustered at the school level. The 45th-50th percentile is the omitted category. Estimates come from Equation 7, which includes controls for race, gender, ESL status, and indicators for ventiles of student achievement interacted with type of elementary school test-score distribution. The mean retention rate is 1.6.

Figure 5 – Third Grade Rank on K-12 Outcomes, 2



Note: These figures plot the coefficient for ventiles of class rank with 95% confidence intervals calculated using standard errors clustered at the school level. The 45th-50th percentile is the omitted category. Estimates come from Equation 7, which includes controls for race, gender, ESL status, and indicators for ventiles of student achievement interacted with type of elementary school test-score distribution. Panels C and D come from a specification which controls for achievement and rank in third grade in both subjects simultaneously.

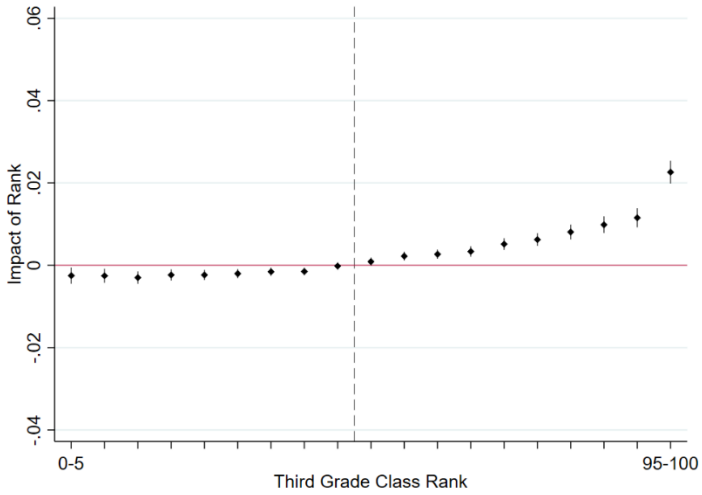
Figure 6 – Third Grade Rank on High School Graduation and College Enrollment



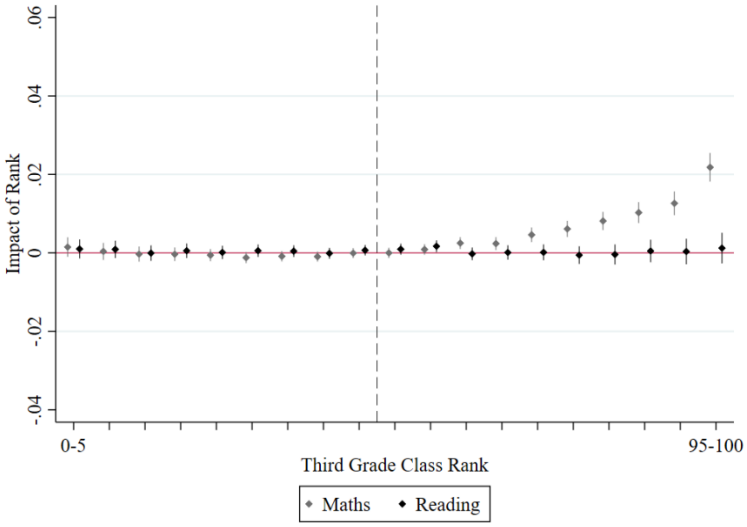
Note: These figures plot the coefficient for ventiles of class rank with 95% confidence intervals calculated using standard errors clustered at the school level. The 45th-50th percentile is the omitted category. Estimates come from Equation 7, which includes controls for race, gender, ESL status, and indicators for ventiles of student achievement interacted with type of elementary school test-score distribution. The mean high school graduation rate of 71 percent and a college enrollment rate of 47 percent. The mean two-year college enrollment rate is 31 percent, and the mean four-year college enrollment rate is 23 percent.

Figure 7 – Third Grade Rank on Major Choice

A. Declaring a STEM Major



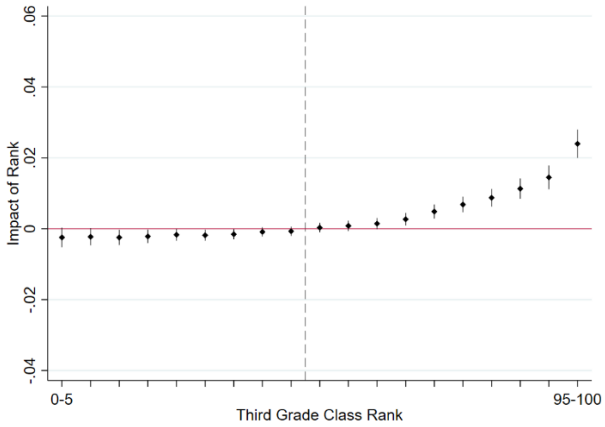
B. Declaring a STEM by subject specific rank



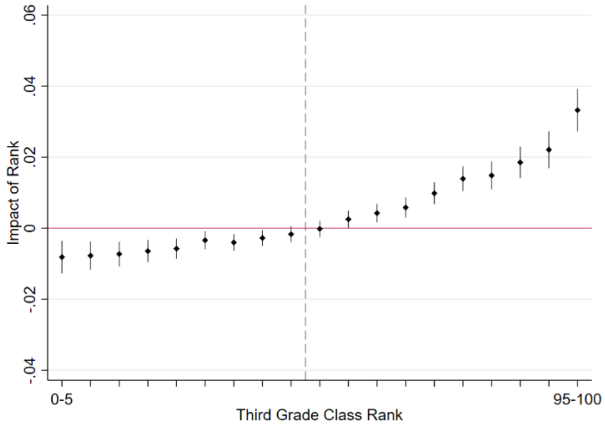
Note: These figures plot the coefficient for ventiles of class rank with 95% confidence intervals calculated using standard errors clustered at the school level. The 45th-50th percentile is the omitted category. Estimates come from Equation 7, which includes controls for race, gender, ESL status, and indicators for ventiles of student achievement interacted with type of elementary school test-score distribution. Panel B from a specification which controls for achievement and rank in third grade in both subjects simultaneously. Mean STEM enrolment is 0.04.

Figure 8 – Third Grade Rank on Bachelor's Degree Receipt

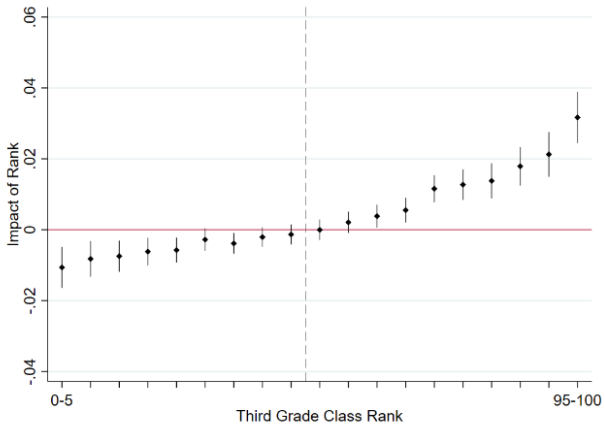
A. Graduate 4 Year College in 4 years



B. Grad 4 Year College in 6 years

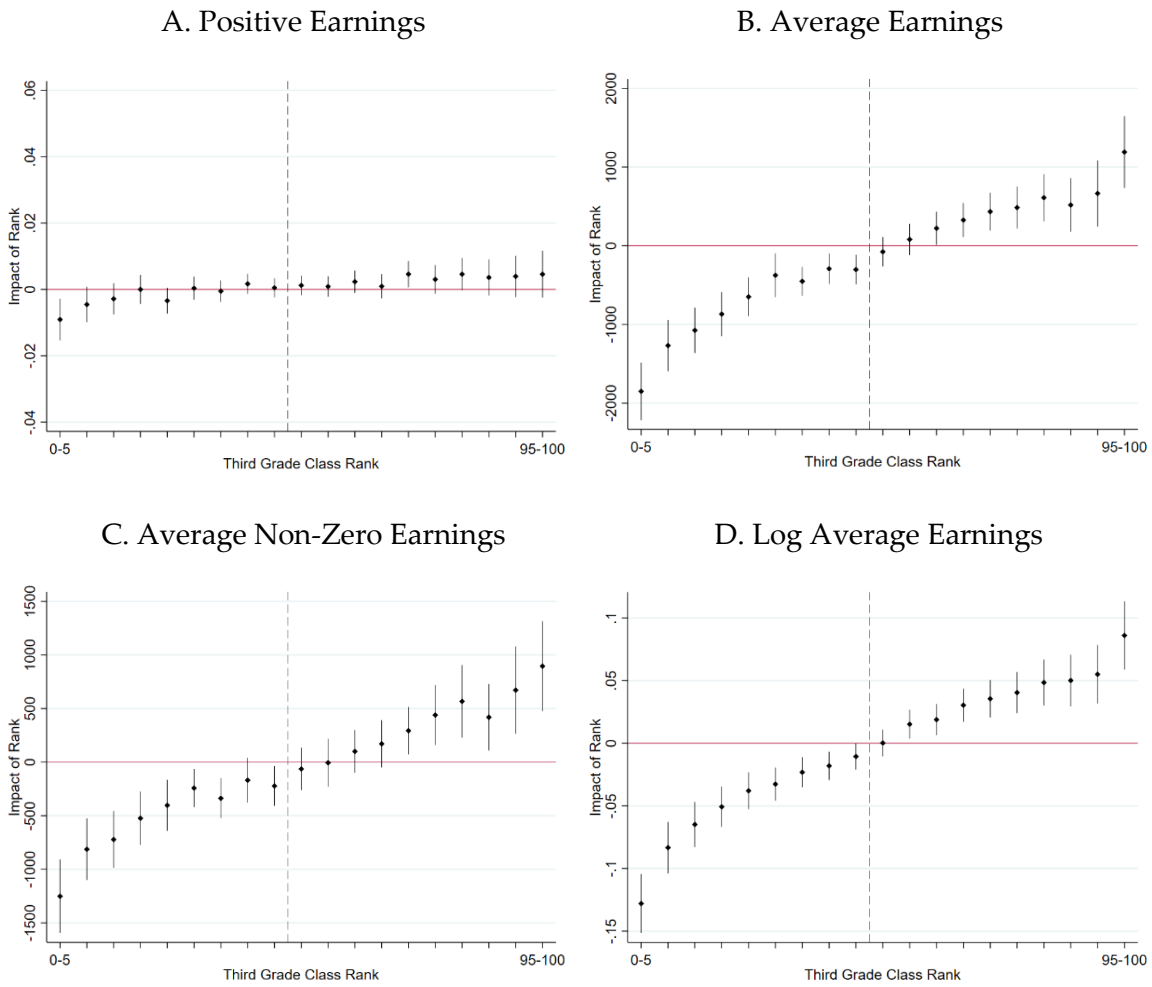


C. Grad 4 Year College in 8 years



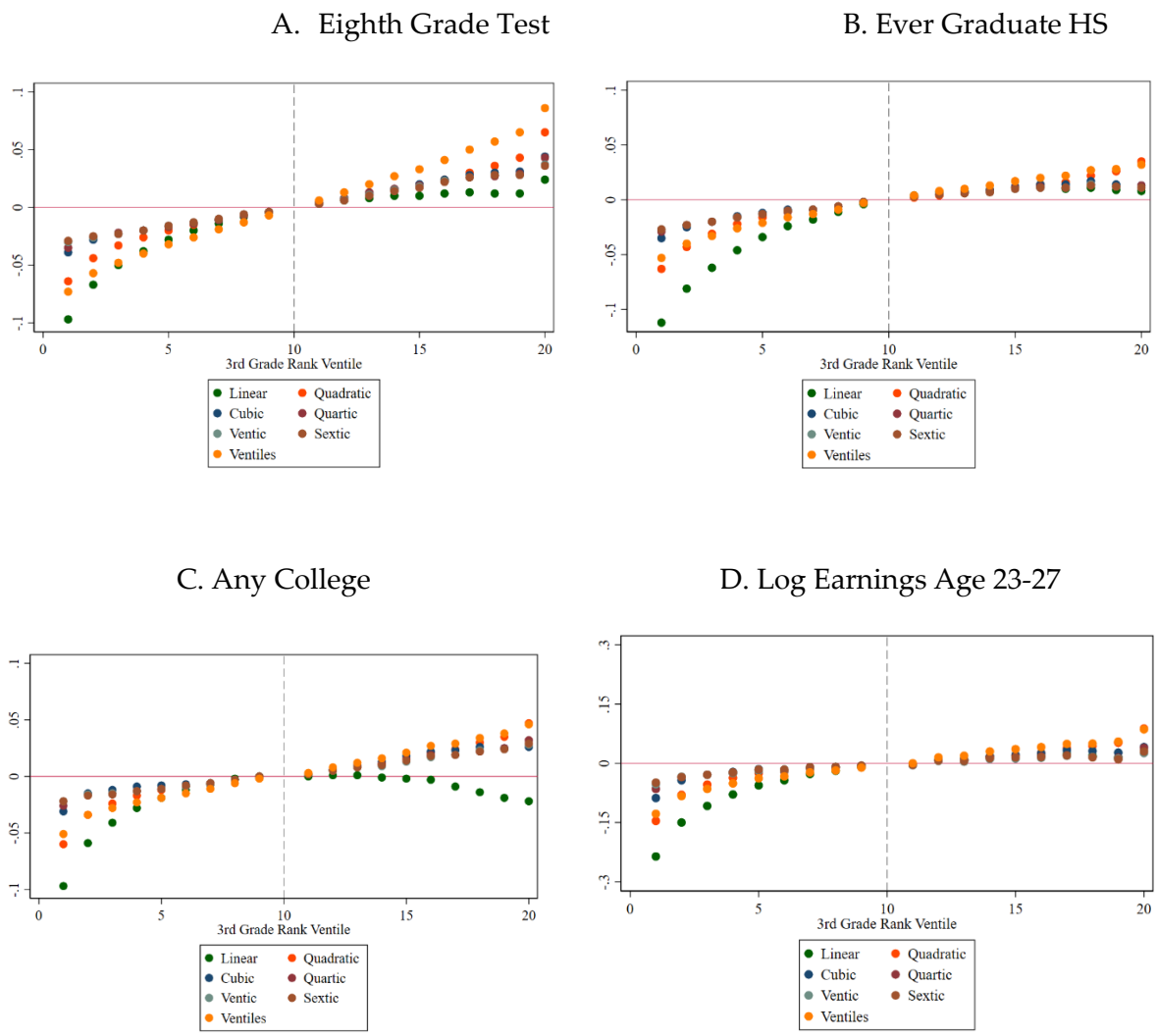
Note: These figures plot the coefficient for ventiles of class rank with 95% confidence intervals calculated using standard errors clustered at the school level. The 45th-50th percentile is the omitted category. Estimates come from Equation 7, which includes controls for race, gender, ESL status, and indicators for ventiles of student achievement interacted with type of elementary school test-score distribution. Mean graduation rate in 4/6/8 years is 4/14/16 percent.

Figure 9 – Third Grade Rank on Labor Market Outcomes (Age 23-27)



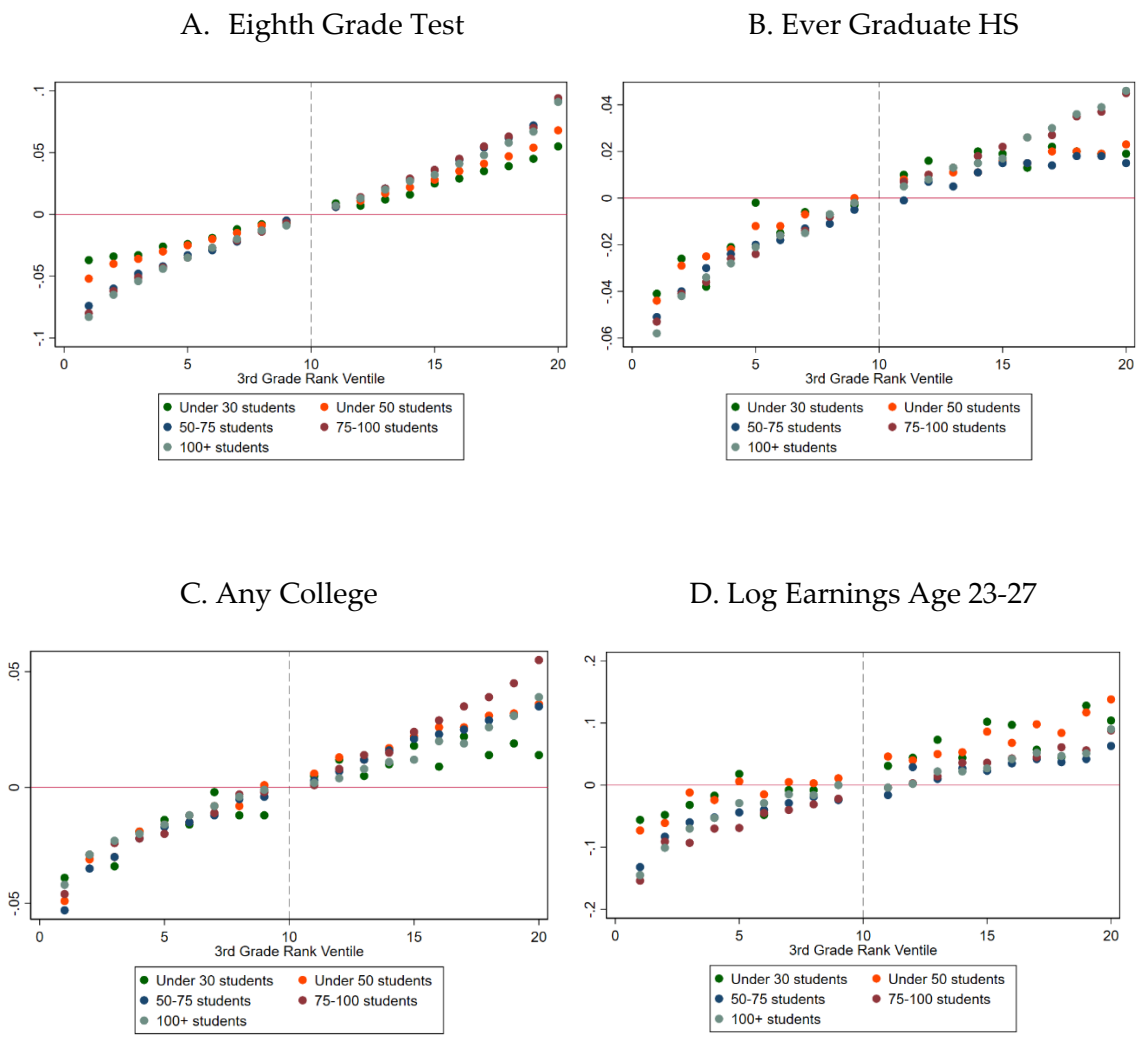
Note: These figures plot the coefficient for ventiles of class rank with 95% confidence intervals calculated using standard errors clustered at the school level. The 45th-50th percentile is the omitted category. Estimates come from Equation 7, which includes controls for race, gender, ESL status, and indicators for ventiles of student achievement interacted with type of elementary school test-score distribution. The mean positive earnings between 23-27 are $24,912. Mean earnings are $17,365.

Figure 10 – Flexible controls for Achievement



Note: These figures plot the coefficient for ventiles of class rank. The 45th-50th percentile is the omitted category. Estimates come from Equation 7, which includes controls for race, gender, ESL status, and indicators for ventiles of student achievement.

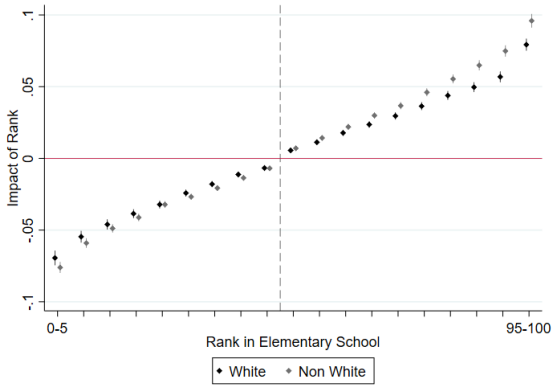
Figure 11 – Results by Class Size



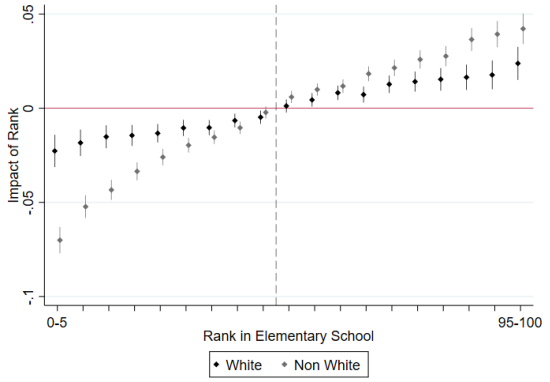
Note: These figures plot the coefficient for ventiles of class ran with 95% confidence intervals calculated using standard errors clustered at the school level k. The 45th-50th percentile is the omitted category. Estimates come from Equation 7, which includes controls for race, gender, ESL status, and indicators for ventiles of student achievement.

Figure 12 – Heterogeneity by Race

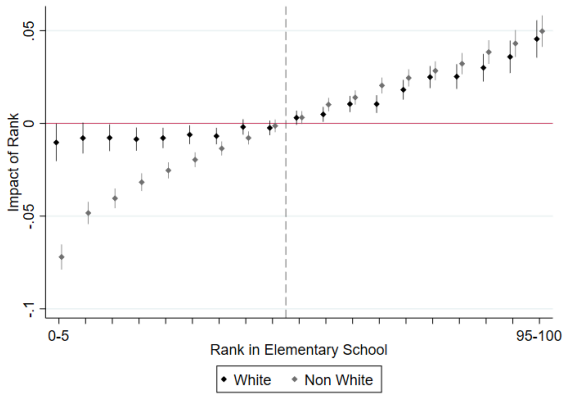
A. Eighth Grade Test



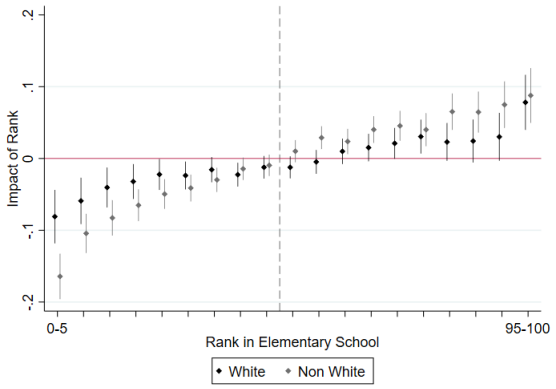
B. Ever Graduate HS



C. Any College



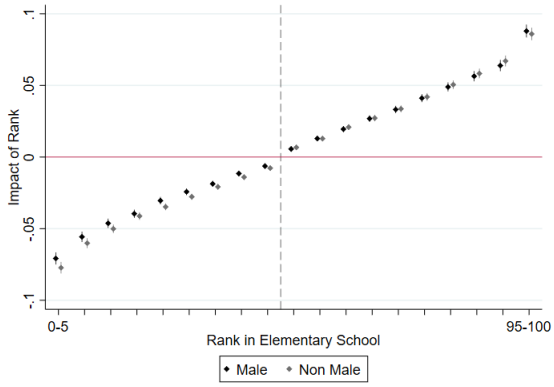
D. Log Earnings Age 23-27



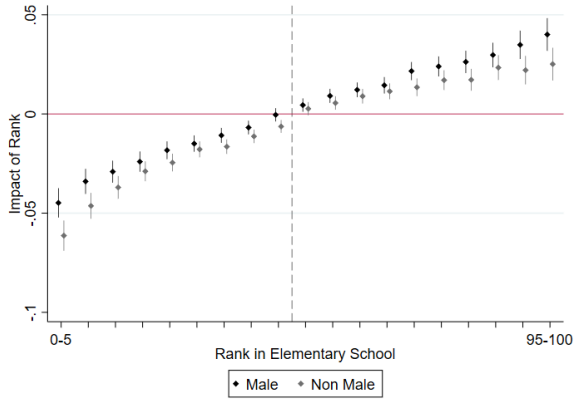
Note: These figures plot the coefficient for ventiles of class rank with 95% confidence intervals calculated using standard errors clustered at the school level. The 45th-50th percentile is the omitted category. Estimates come from Equation 7, which includes controls for race, gender, ESL status, and indicators for ventiles of student achievement.

Figure 13 – Heterogeneity by Gender

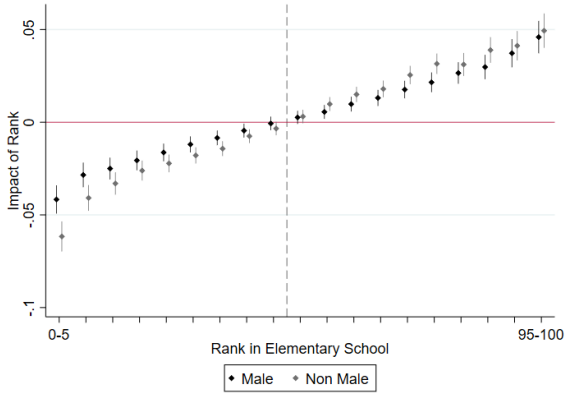
A. Eighth Grade Test



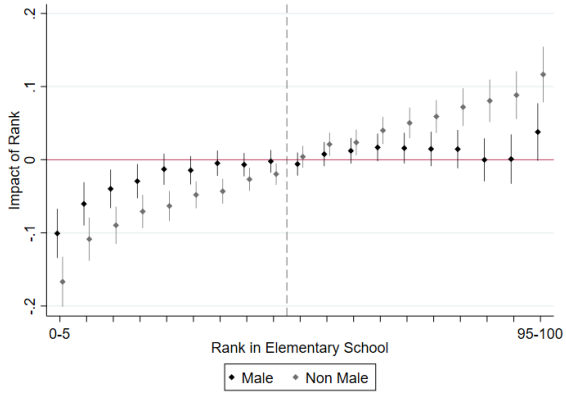
B. Ever Graduate HS



C. Any College

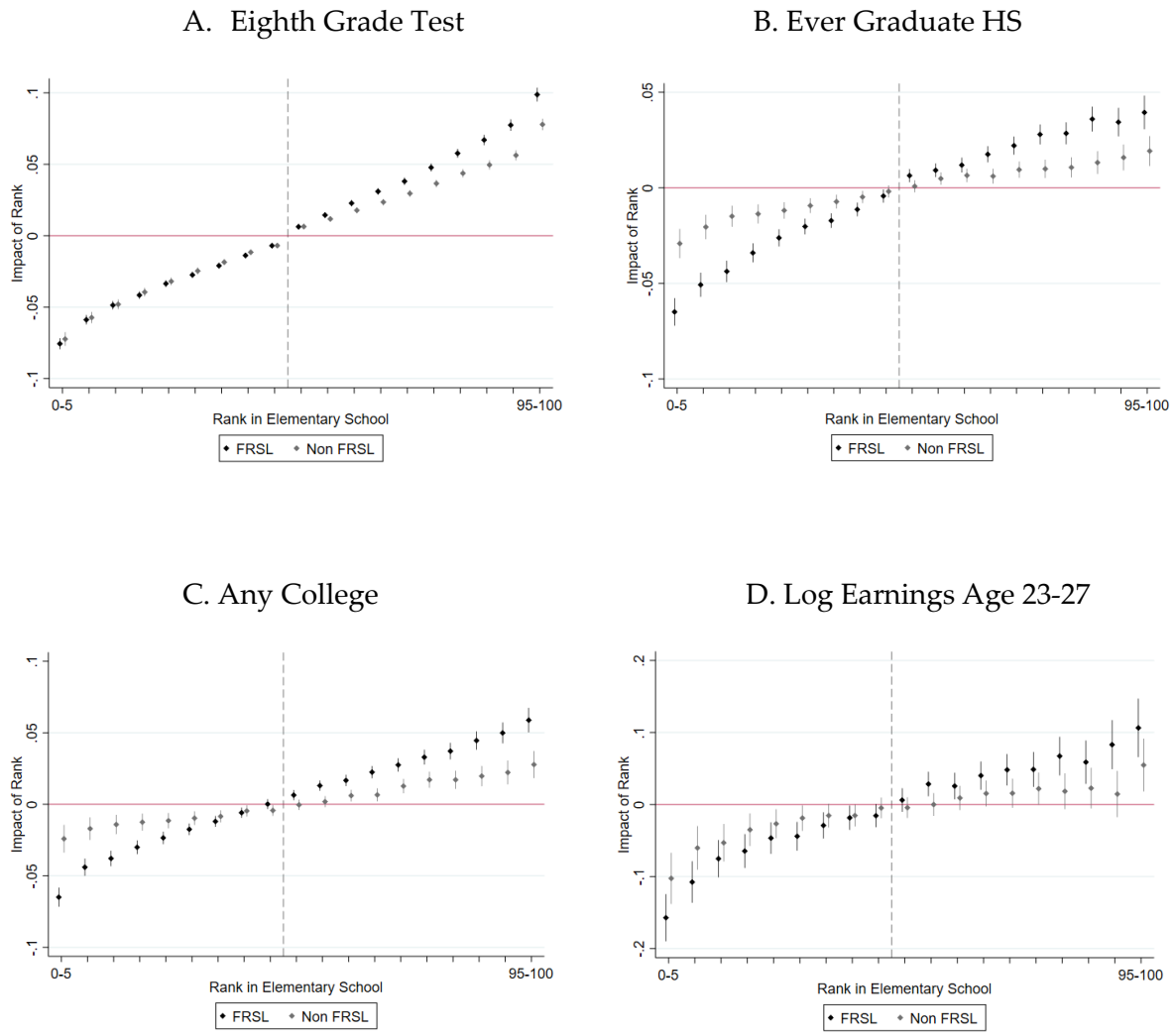


D. Log Earnings Age 23-27



Note: These figures plot the coefficient for ventiles of class rank. The 45th-50th percentile is the omitted category with 95% confidence intervals calculated using standard errors clustered at the school level. Estimates come from Equation 7, which includes controls for race, gender, ESL status, and indicators for ventiles of student achievement.

Figure 14 – Heterogeneity by Free and Reduced-Price Lunch



Note: These figures plot the coefficient for ventiles of class rank. The 45th-50th percentile is the omitted category with 95% confidence intervals calculated using standard errors clustered at the school level. Estimates come from Equation 7, which includes controls for race, gender, ESL status, and indicators for ventiles of student achievement.

Table 1: Summary Statistics

	Mean	Standard Deviation	Observations
Demographics – 13 Cohorts			
Male	0.50	0.50	6,117,690
Economic Disadvantage	0.49	0.50	6,117,690
English as a Second Language	0.12	0.32	6,117,690
White	0.47	0.50	6,117,690
Asian	0.03	0.17	6,117,690
Black	0.15	0.36	6,117,690
Hispanic	0.35	0.48	6,117,690
Size of Third Grade SSC	93.7	46.6	6,117,690
Repeat Third Grade	0.02	0.13	6,117,690
K-12 Outcomes – 13 Cohorts			
State Test Percentile, eighth grade	0.55	0.28	4,919,673
Ever Graduate High School	0.70	0.46	6,117,690
AP Calculus	0.08	0.28	6,117,690
AP English	0.18	0.39	6,117,690
Any College	0.46	0.50	6,117,690
Enroll, 4 yr college	0.23	0.42	6,117,690
Enroll, 2 year college	0.31	0.46	6,117,690
College Outcomes – 10/8/6 Cohorts			
Declare STEM Major	0.041	0.198	6,117,690
BA in 4 years	0.06	0.24	4,573,672
BA in 6 years	0.14	0.34	3,597,340
BA in 8 years	0.16	0.37	2,647,240
Age 23-27 Labor Outcomes – 8 Cohorts			
Non Zero Wages	0.65	0.43	3,597,340
Real Wages	17,300	24,093	3,597,340
Real Non-Zero Wages	24,818	25,372	2,652,284

Note: This table contains summary statistics for the main estimating sample of third graders from 1995-2008. Some outcomes are only available for early cohorts which generates the differences in sample size. *Enroll, 4yr college* means enrollment within 3 years of "on-time" high school graduation and is similarly defined for two-year colleges.

Table 2 – Balance Test

	Low Income	Male	ESL	White	Asian	Black	Hispanic
				A. Un-interacted			
Rank	0.097***	-0.011**	0.030***	-0.085***	-0.001*	0.047***	0.037***
	(0.005)	(0.004)	(0.003)	(0.005)	(0.000)	(0.004)	(0.004)
				B. School Mean Quartiles			
Rank	0.030***	-0.004	0.010*	-0.028***	-0.001	0.024***	0.006
	(0.006)	(0.005)	(0.004)	(0.006)	(0.002)	(0.005)	(0.005)
				C. School Mean Deciles			
Rank	0.025***	-0.001	0.006	-0.024***	-0.004	0.023***	0.006
	(0.006)	(0.005)	(0.004)	(0.006)	(0.002)	(0.005)	(0.005)
				D. School Variance Quartiles			
Rank	0.039***	-0.015***	0.020***	-0.039***	0.000	0.021***	0.018***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.001)	(0.004)	(0.004)
				E. School Variance Deciles			
Rank	0.030***	-0.014***	0.020***	-0.031***	-0.001	0.016***	0.015***
	(0.004)	(0.004)	(0.003)	(0.004)	(0.001)	(0.004)	(0.004)
				F. School Quartiles of Mean and Quartiles of Variance			
Rank	-0.006	-0.006	0.002	0.008	0.003*	-0.003	-0.009*
	(0.005)	(0.005)	(0.004)	(0.005)	(0.002)	(0.004)	(0.004)
N	6,117,651	6,117,651	6,117,651	6,117,651	6,117,651	6,117,651	6,117,651

Notes: This table regresses predetermined characteristics on a linear effect of rank, SSC fixed effects, and achievement controls. Panel A does not interact ventiles of achievement with anything. Panel B interacts ventiles of achievement with indicators for quartiles of school mean achievement. Panel C interacts achievement with deciles of school mean. Panel D interacts achievement with quartiles of school variance. Panel E interacts achievement with deciles of school variance. Panel F is our preferred specification and interacts achievement with 16 indicators for school mean and variance (quartiles of mean X quartiles of variance). Standard errors are clustered at the school level.

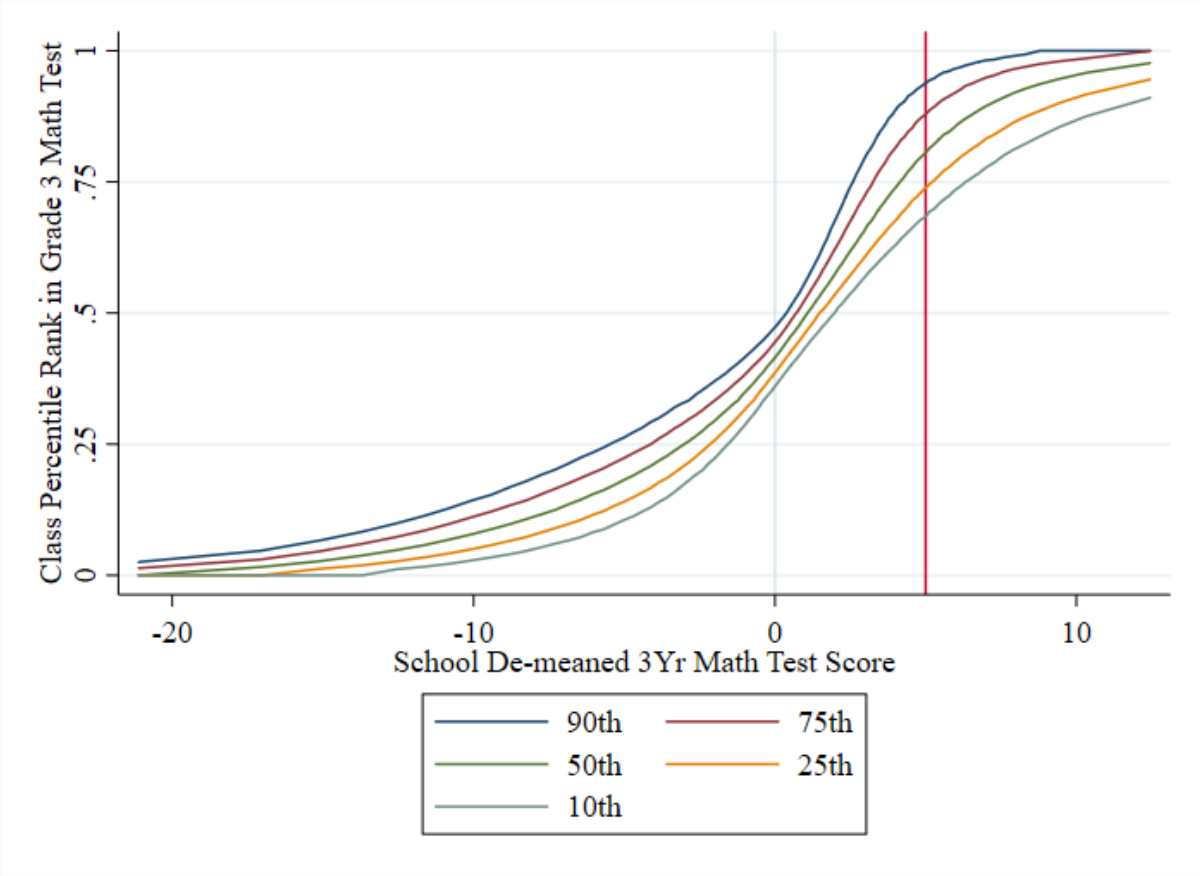
Table 3: The Distribution of Rank by Elementary School Effectiveness Measures

	Third Grade Attainment				Third to Eighth Grade Value Added			
State Ventile	Average School	Bad School	Good School	Rank Change	Average School	Bad School	Good School	Rank Change
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1	0.028	0.054	0.014	*-0.040*	0.049	0.075	0.035	*-0.040*
2	0.074	0.137	0.039	*-0.098*	0.106	0.145	0.075	*-0.070*
3	0.123	0.211	0.068	*-0.143*	0.160	0.203	0.113	*-0.090*
4	0.172	0.279	0.100	*-0.179*	0.213	0.257	0.153	*-0.104*
5	0.222	0.341	0.133	*-0.208*	0.261	0.309	0.188	*-0.121*
6	0.272	0.401	0.170	*-0.231*	0.310	0.354	0.227	*-0.127*
7	0.323	0.458	0.210	*-0.248*	0.358	0.399	0.264	*-0.135*
8	0.374	0.509	0.250	*-0.259*	0.403	0.447	0.304	*-0.143*
9	0.426	0.565	0.294	*-0.271*	0.449	0.493	0.349	*-0.144*
10	0.479	0.613	0.346	*-0.267*	0.497	0.542	0.387	*-0.155*
11	0.527	0.650	0.390	*-0.260*	0.541	0.581	0.435	*-0.146*
12	0.577	0.703	0.452	*-0.251*	0.590	0.624	0.485	*-0.139*
13	0.627	0.737	0.495	*-0.242*	0.630	0.667	0.527	*-0.14*
14	0.685	0.787	0.561	*-0.226*	0.684	0.718	0.587	*-0.131*
15	0.726	0.810	0.613	*-0.197*	0.724	0.749	0.634	*-0.115*
16	0.784	0.858	0.685	*-0.173*	0.778	0.799	0.693	*-0.106*
17	0.828	0.889	0.745	*-0.144*	0.821	0.841	0.752	*-0.089*
18	0.876	0.916	0.808	*-0.108*	0.870	0.881	0.819	*-0.062*
19	0.930	0.995	0.889	*-0.106*	0.924	0.930	0.890	*-0.040*
20	0.962	0.973	0.934	*-0.039*	0.959	0.961	0.940	*-0.021*
Value Added	-0.001	-0.035	0.026	*0.061*	0.000	-0.055	0.055	*0.11*

Notes: This table categorizes elementary schools into good, bad and average in terms of average third grade attainment, and third to eighth grade value added. Good/bad are defined as being one standard deviation above/below the average (with tolerance of 0.0045). Each row the average class rank of students in this type of school for that ventile. The Rank Change column is the average rank change of students in that ventile from moving from a bad to a good school. The final row presents the third to eighth grade school level value added for this type of school. Value added calculated conditional on a cubic of third grade percentile.

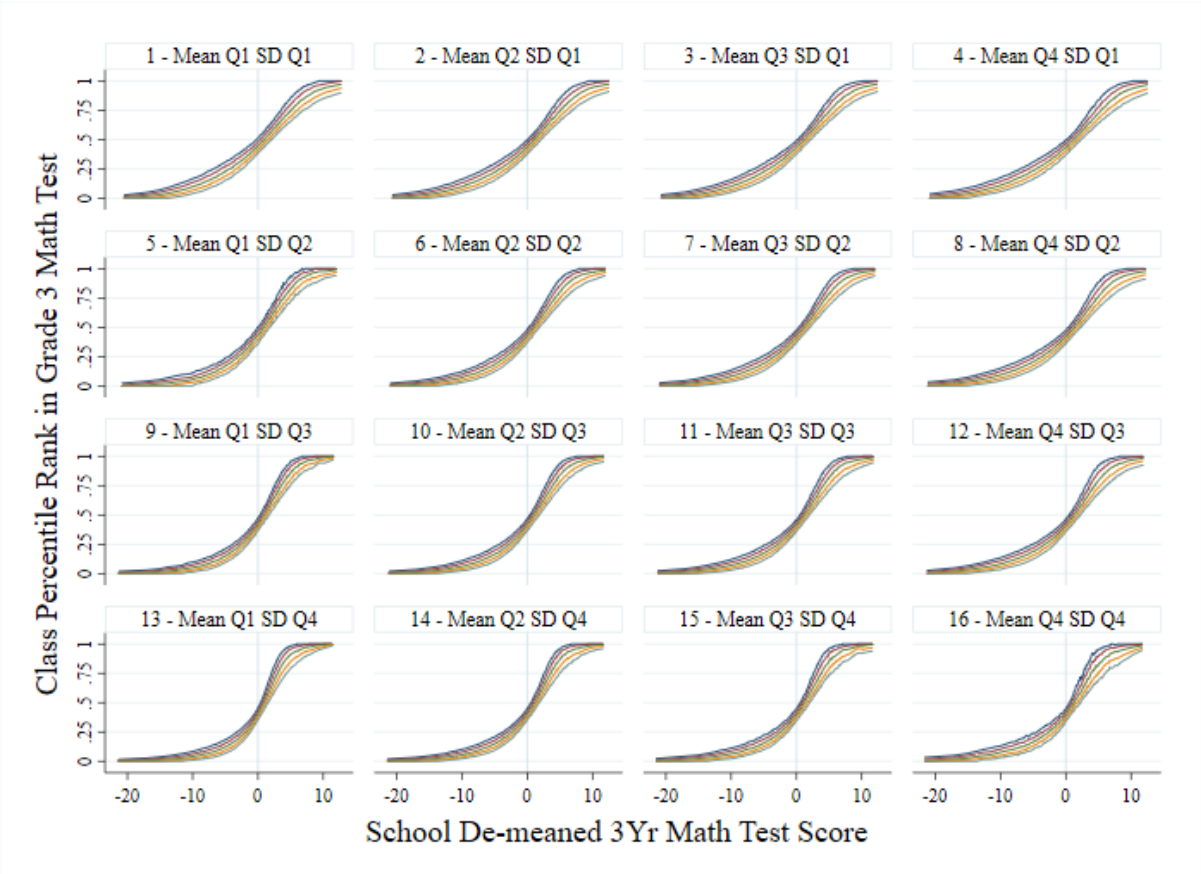
Appendix Figures and Tables

Appendix Figure 1: Common Support of Class Test Score



Notes: This shows the rank for a student for a given test score relative to the mean. We show the differences in rank across SSCs. Each line represents a percentile of students with that test score relative to the mean. (e.g. 10th percentile). The vertical thickness shows the different rank values students with the same relative position within their classes have. The red line represents where the students from Figure 1, who are five points above their class mean would reside.

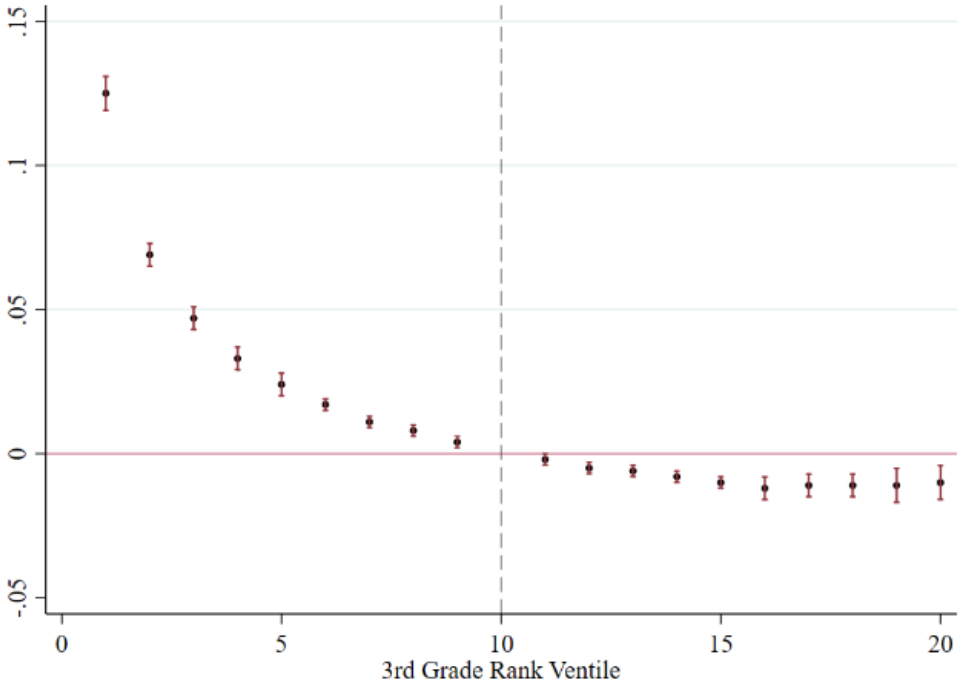
Appendix Figure 2 – Variation of Local Rank of Median Student Rank Within School-Subject Groups



Notes: This shows the rank for a student for a given test score relative to the mean in each school distribution type $g_a(\quad)$. We show the differences in rank across SSCs. Each line represents a percentile of students with that test score relative to the mean. (e.g. 10th percentile). The vertical thickness shows the different rank values students with the same relative position within their classes have.

Appendix Figure 3 – Testing for Missingness

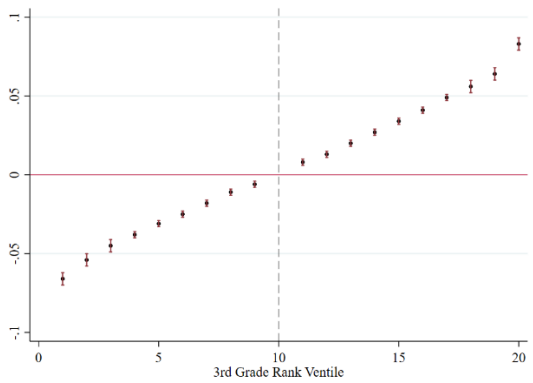
A. Missing 8th Grade Test



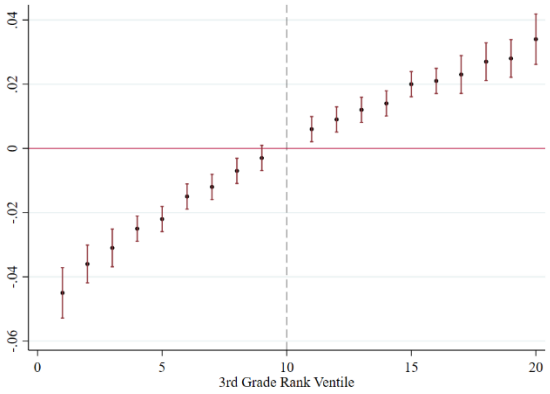
Note: These figures plot the coefficient for ventiles of class rank with 95% confidence intervals calculated using standard errors clustered at the school level. The 45th-50th percentile is the omitted category. Estimates come from Equation 7, which includes controls for race, gender, ESL status, and indicators for ventiles of student achievement. The mean retention rate is 1.6.

Appendix Figure 4 - Balanced Sample for Full Set of Outcomes

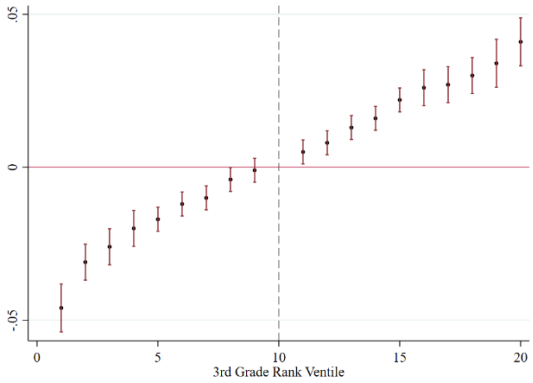
A. Eighth Grade Test



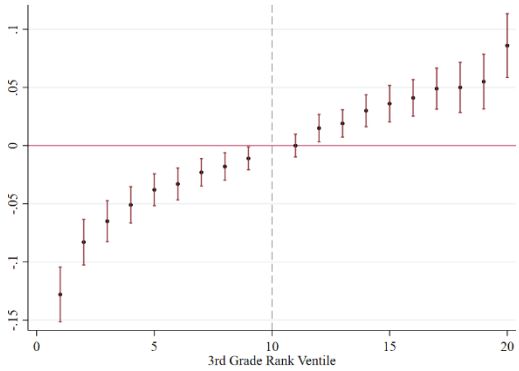
B. Ever Graduate HS



C. Any College

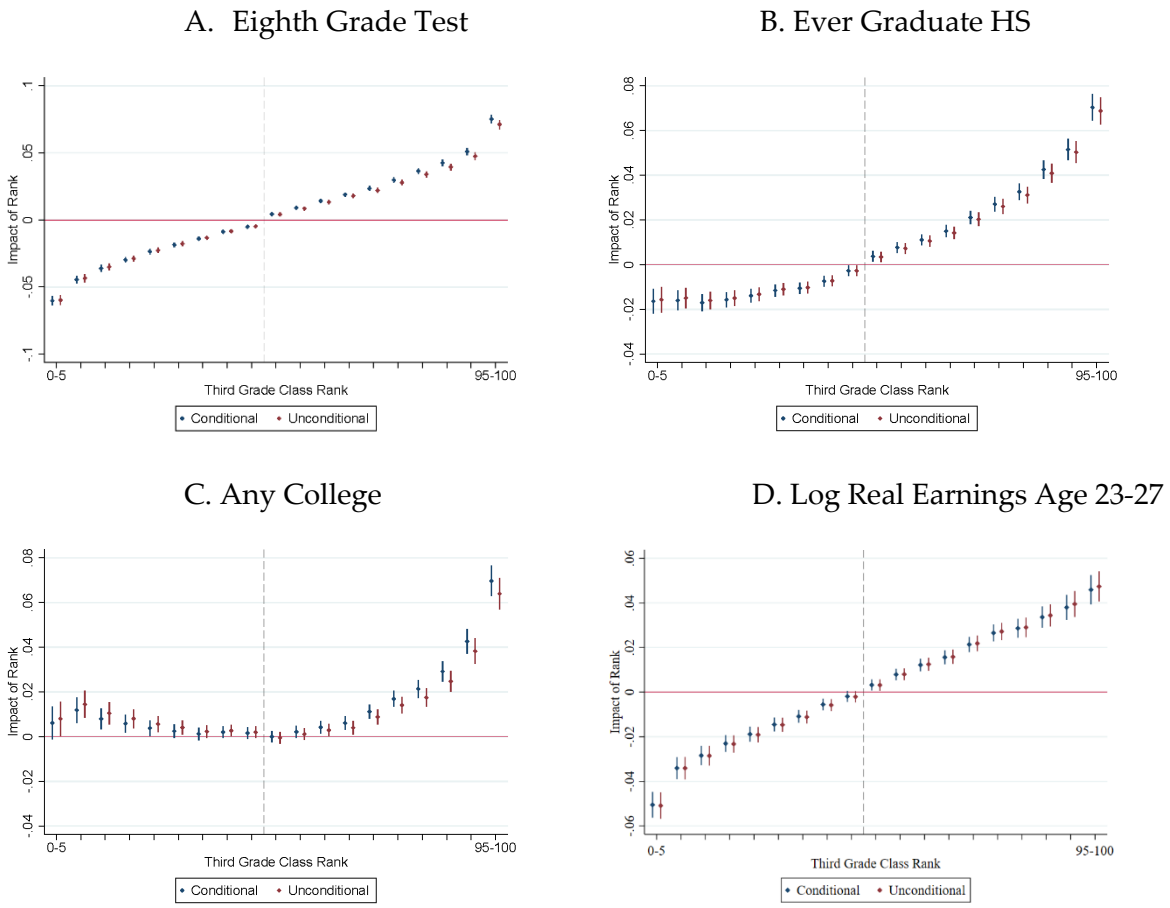


D. Log Real Earnings Age 23-27



Notes: These figures are on the reduced balanced Sample for Full Set of Outcomes 8 cohorts Third Grade 1994-2001. These figures plot the coefficient for ventiles of class rank. The 45th-50th percentile is the omitted category. Estimates come from Equation 7, which includes controls for race, gender, ESL status, and indicators for ventiles of student achievement.

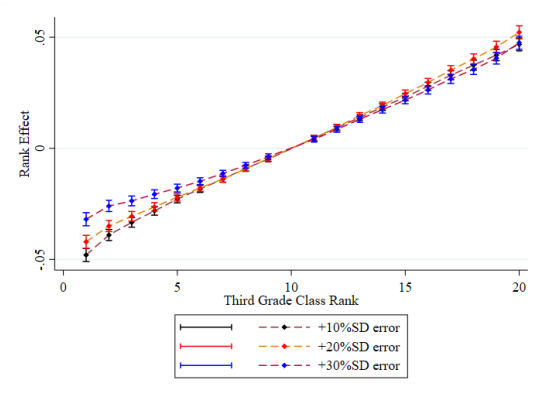
Appendix Figure 5 - Conditional/unconditional estimates.



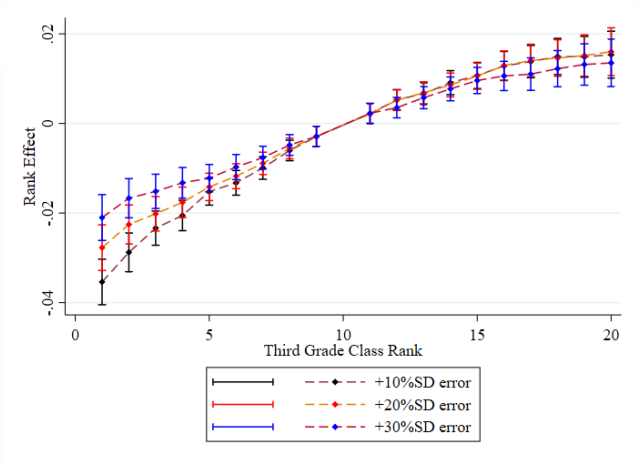
Notes: These figures plot the coefficient for ventiles of class rank with 95% confidence intervals calculated using standard errors clustered at the school level. The 45th-50th percentile is the omitted category. Estimates come from Equation 4, which includes controls for race, gender, ESL status, and indicators for ventiles of student achievement, not interacted by school test score distribution. We show how the estimates change when including controls for predetermined characteristics including race, gender, and ESL status (our preferred specification) and when we do not.

Appendix Figure 6 – Measurement Error Simulations

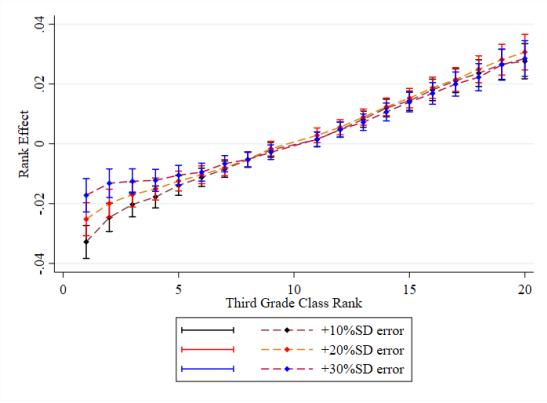
A. Eighth Grade Test



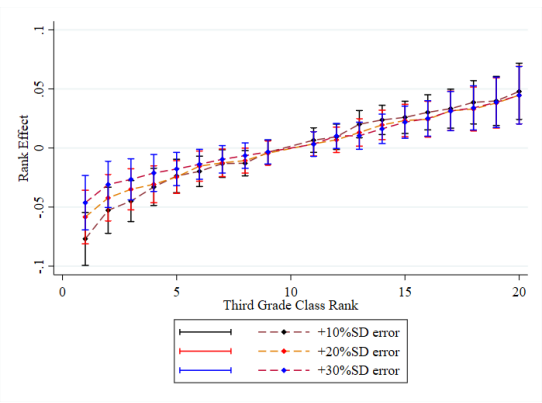
B. Ever Graduate HS



C. Any College



D. Log Real Earnings Age 23-27



Notes: This figure shows estimates for our main specification where we added normally distributed noise with zero mean and a standard deviation equivalent to 10%, 20% and 30% of the standard deviation in the 3rd-grade achievement before calculating ranks. The resulting non-linear and correlated measurement error in 3rd-grade achievement and the rank-measure results in a non-linear downward-bias in the rank estimate. This replicates a result by Murphy and Weinhardt (*forthcoming*), who discuss this non-traditional noise for other types of noise that are not normally distributed, which also lead to downward- bias.

Appendix Table 1 – Alternate Measures of Rank on Main Outcomes

	Method for Calculating Rank			
	Mean Rank	Bottom Rank	Random Rank	On Time Students
	(1)	(2)	(3)	(4)
Outcome				
Test Scores	0.149***	0.155***	0.100***	0.147***
	(0.003)	(0.002)	(0.002)	(0.003)
Grad High School	0.085***	0.078***	0.059***	0.083***
	(0.004)	(0.004)	(0.004)	(0.004)
Any College	0.088***	0.090***	0.060***	0.086***
	(0.005)	(0.004)	(0.004)	(0.005)
Real Earnings	1853.1***	1983.5***	1482.3***	1785.8***
	(251.1)	(244.6)	(213.7)	(247.7)

Notes: This table presents the rank estimates from 16 different regressions, using four different rank measures on four outcomes. Mean Rank – assigns the average rank to all students tied with the same score (this is the measure we use in the paper). Bottom Rank – assigns the bottom rank to all students tied with the same score. Random Rank – assigns a random rank to all students with tied with the same score. On time students – Assigns rank on a mean rank basis only among students who took their third grade exam on time, rather than all students in their class who took the exam. Standard errors are clustered at the school level.

Appendix Table 2 – Outcomes by Function of Distribution

	Repeat 3rd	Grade 8 Test	Grad HS	Any College	Grad BA in 8 yrs	Log Wage
			A. Un-interacted			
Rank	-0.037***	0.089***	0.061***	-0.007	-0.013**	0.121***
	(0.005)	(0.003)	(0.004)	(0.005)	(0.005)	(0.017)
			B. School Mean Quartiles			
Rank	-0.038***	0.141***	0.084***	0.084***	0.019***	0.171***
	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)	(0.020)
			C. School Mean Deciles			
Rank	-0.040***	0.152***	0.088***	0.092***	0.027***	0.185***
	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)	(0.021)
			D. School Variance Quartiles			
Rank	-0.039***	0.106***	0.083***	0.041	-0.002	0.144***
	(0.003)	(0.002)	(0.004)	(0.004)	(0.005)	(0.017)
			E. School Variance Deciles			
Rank	-0.040***	0.113***	0.089***	0.050***	0.005	0.152***
	(0.003)	(0.002)	(0.004)	(0.004)	(0.005)	(0.018)
			F. School Quartiles of Mean and Quartiles of Variance			
Rank	-0.036***	0.147***	0.087***	0.088***	0.029***	0.173***
	(0.003)	(0.003)	(0.004)	(0.005)	(0.005)	(0.020)
N	6,117,651	4,919,628	6,117,651	6,117,651	2,647,234	2,652,264

Notes: This table considers the linear effect of rank for different specification of $g_d()$. Panel A, is controls for student prior achievement with ventiles for third-grade achievement. Panel B interacts achievement with quartile of school mean achievement. Panel C interacts achievement with deciles of school variance. Panel D interacts achievement with quartile of school achievement variance. Panel E interacts achievement with deciles of school achievement variance. Panel C. Panel F is our preferred specification and interacts achievement with 16 indicators for school mean and variance (quartiles of mean X quartiles of variance). Standard errors are clustered at the school level.