The Icarus Syndrome: Why Do Some High Flyers Soar While Others Fall?

Eric Parsons

July 2013

This paper follows a cohort of initially high performing Missouri students from grade-3 through grade-9 and examines which school factors influence their academic success. Three key findings emerge. First, in terms of performance on standardized tests, schools that are effective in promoting academic growth among low performing students are also generally effective with high performing students. Second, high performing students who attend disadvantaged schools are more likely to take Algebra I later relative to their counterparts who attend more advantaged schools. Third, somewhat surprisingly, increasing the number of high performing students in a school negatively affects high performing student outcomes.

The author is in the Economic and Policy Analysis Research Center in the Department of Economics at the University of Missouri – Columbia. The usual disclaimers apply.

1. Introduction

The No Child Left Behind (NCLB) act has had a deep and wide-ranging impact on elementary and secondary education in the United States since its inception in 2001. One of the most important changes brought forth by the law has been the development and growing importance of district and school accountability systems that rely heavily on standardized exam scores to measure student performance. Much of the accountability focus of NCLB has centered on low performing students given the policy objectives of (a) increasing the percentage of students able to perform at the "proficient" level on state standardized exams and (b) "closing the gap" between high performing and low performing students, particularly the portion of the gap that falls along racial and socio-economic lines.

Relatively little attention has been paid in recent policy reforms to students scoring at the top of the score distribution, a point emphasized recently in the New York Times (Finn, 2012). The proficiency-based focus of NCLB and similar legislation have the potential, albeit unintentionally, to harm high performing students. For example, students at the top of the performance distribution already meet proficiency standards and, as such, are already tallied in a school's accountability "plus column." Moreover, they are likely to stay there over time with little or no direct assistance. Hence, a district or school facing tightening accountability requirements and budget constraints has the incentive to transfer resources from high performing students to students closer to the proficiency threshold. Prior research has provided some evidence that such redistributions are occurring (Ballou and Springer, 2008; Reback, 2008; Neal and Schanzenbach, 2010).

The incentives associated with closing the achievement gap also have the potential to harm high achieving students (hereafter, "high flyers"). Notably, school policies designed to help high flyers can *widen* the achievement gap. Hence, even if schools and districts are properly focused on closing the gap from the bottom-up rather than the top-down, they still have little incentive to continually move their goalpost further down the field.

Given that the quality of high performing students has been found to impact economic growth independent of the overall level of human capital (Hanushek and Woessmann, 2012), the increasing policy focus away from high achievers may have adverse implications for the future economic health of the United States. This is only true, however, if (a) schools are actually shifting resources away from high performing students and towards their lower performing peers and (b) this decline in resources is having a deleterious effect on the performance of these top achievers. Point (b) is especially salient given the fact that the parents of high performing students may work to offset any lost resources at the school-level.[1]

The contribution of the present study is to identify the factors that influence the academic success of initially high flying students in public K-12 schools. Initial high flyers are identified based on academic performance in grades 3 and 4 and tracked through grade-9. Three key findings emerge. First, across schools, there is no evidence to suggest that an increase in a school's share of low performing students (i.e., below median) corresponds to worse outcomes for high performing students. Put differently, schools that do a good job educating low performing students also appear to be doing a good job educating their high performing peers. Although this result does not rule out the possibility that resources are being reallocated away from high flyers within schools due to proficiency targeting, it does imply that wholesale specialization where high flyers are "left behind" in disadvantaged schools is not the norm.

[1] Parental substitution of resources may be more likely if high performing students are disproportionately coming from families with high socioeconomic status (SES). Arguments of this type are similar to those made by anti-student tracking advocates, i.e. removing tracking has the potential to help struggling students while causing no harm to high performers (Slavin, 1990). Other research, however, suggests that this second assertion may not be true (e.g. Argys et al., 1996).

A second key finding points to a challenge that high flyers face when surrounded by a larger share of low achieving peers. Specifically, schools that serve a high proportion of low performing students appear to structure their course sequences specifically to meet these students' needs, perhaps unsurprisingly. As a result, high flyers attending these schools are more likely to take Algebra I later relative to their counterparts in other schools. To the extent that this delayed course taking affects college readiness, high flyers attending these schools may be adversely affected. Moreover, these findings for Algebra I may be indicative of more general challenges in course alignment for high flyers in disadvantaged schools. Poor and minority high flyers are the most likely to be affected by course misalignment because they are the most likely to attend schools with a high proportion of low performing students.

Finally, a third key finding is that the academic trajectories of initially high performing students are harmed by the presence of other high performers holding other factors constant. This outcome is seemingly in contrast to previous studies indicating that students at the top of the performance distribution obtain higher academic achievement when they are exposed to more high quality peers (Hoxby and Weingarth, 2006; Imberman et al., 2012; Burke and Sass, 2013). However, a closer reading of these studies suggests that the findings presented here are not as divergent as they initially seem. Some insight into the mechanism underlying this result comes from Lavy et al. (2012), who also find that high performing (male) students are adversely affected by the presence of other top performing students.

2. Prior Research

A primary objective of this study is to examine whether recent federal accountability policies such as NCLB appear to have adversely influenced high flyers. While few studies have focused specifically on outcomes for top students, there is a larger and more general body of research examining how recent, proficiency-based accountability systems affect student performance across the test-score distribution.

Ballou and Springer (2008) use Northwest Evaluation Association (NWEA) testing data and within grade, across year variation between high- and low-stakes years to explore the effect of NCLB accountability on the distribution of student achievement. They find that these policies have led to increased achievement overall, but this increase has occurred despite a small decline in performance among top students (Ballou and Springer, 2008).[2] Using data from Texas, Reback (2008) also finds redistributive effects in schools. Specifically, he shows higher than expected growth among students that are most likely to affect schools' performance ratings. Reback also finds that very low achieving students see larger than expected gains when schools face incentives to increase pass rates, even though these students are unlikely to reach the proficiency threshold; in contrast, high achieving students see lower than expected gains in a similar situation (Reback, 2008). Dee and Jacob (2011) analyze National Assessment of Educational Progress (NAEP) data and find improvements across the distribution on grade-4 mathematics scores in response to NCLB policy. In grade-8, however, they find that mathematics score improvements are limited to the bottom of the distribution. The authors also find no effect of NCLB on reading scores in any grade. In one of the few studies that focuses exclusively on high performing students, Duffett et al. (2008) provide a cross-sectional analysis of NAEP data and find that the performance of the highest achieving students has been stagnant in the post-NCLB era. Given this body of research, there is some indication that schools may be shifting resources in a way that is potentially harmful to the highest achieving students, although the evidence is by no means overwhelming.

[2] Interestingly, the authors also find that the schools *least* likely to be sanctioned experience larger trade-offs than those most likely to fall short of NCLB requirements (Ballou and Springer, 2008). Although somewhat counterintuitive, this finding has potential impacts for the highest performing students, as they are more likely to attend schools with low likelihoods of NCLB sanction.

Xiang et al. (2011) look to improve on previous work using longitudinal data from the NWEA Growth Research Database. They track high performing students from 2004-05 to 2009-10. A primary finding from their study is that many high flyers did not maintain high performing status through the end of their data panel.[3] This result was taken as further evidence to suggest that schools are shifting resources away from top students (Xiang et al., 2011).

The Xiang et al. (2011) study provides a nice entry point into the topic of high flyer performance post-NCLB but has from two key limitations that I address in the present paper. First, the Xiang et al. (2011) study uses a definition for high flyers that is based on a single year of test score outcomes. Given that measurement error can be a large component of individual student exam scores, particularly for those in the tails of the distribution (Boyd et al., 2012; Koedel et al., 2012), this single year approach makes the potential misclassification of high flyers an important problem.[4] In the present study, I limit the role of measurement error in determining the initial group of high flyers by using a definition that depends on data from two consecutive years of test scores. Second, the findings from Xiang et al. (2011) are largely descriptive with only limited analysis of the factors that impact the maintenance of high flyer status. In contrast, the models of high flyer performance studied in this paper include extensive student- and school-level controls, allowing for a more in-depth exploration of the potential risk factors that high flyers face.

3. Data Description

The data for this project are from the Missouri Department of Elementary and Secondary Education's statewide longitudinal data system. This system covers all students that attend a public

[3] The authors also use hierarchical linear modeling to address questions regarding high flyer performance in low performing schools (Xiang, 2011). However, this portion of their analysis uses a within-school high flyers definition that is not paralleled in the current work.

[4] This point was raised by Lee (2011) in his review of the Xiang et al. (2011) report. Xiang et al. (2011) also acknowledge this issue and use conditional standard errors of measurement (CSEMs) to calculate the attrition rate that would be expected if the loss of high flyers over time was solely the result of misclassification due to test measurement error. They note that the actual attrition rate observed in their sample is much higher than the expected rate (Xiang et al., 2011).

elementary or secondary school in the state of Missouri and, by virtue of a unique student identifier, allows for student records to be linked over time and across schools within the state from 2006 onward. In addition to student enrollment data, the system also contains assessment data for all students who have taken from the Missouri Assessment Project (MAP) exam, as well as course assignments for all students.[5]

4. Empirical Strategy

4.1 Identifying High Flyers

The analysis focuses on a single cohort of high flying students – those who were in grade-3 in 2006, the first year in which students in Missouri can be tracked individually. This cohort was chosen because (a) students have complete records of standardized scores from grade-3 through grade-8 and (b) they can be followed into grade-9 in the final year of the data panel, which allows for the analysis to be extended to evaluate course-taking behavior for Algebra I.[6] Algebra I course taking behavior is of interest given that Algebra I is widely viewed as a cornerstone course in many students' K-12 careers (Helfand, 2006; GreatSchools, 2012).

For this analysis, students are defined as initial high flyers based on their grade-3 and grade-4 mathematics MAP scores. Specifically, students with a score in the top 10% of their grade cohort for one of the two years, and a score not outside the top 20% for the other year, are included as high flyers.[7] For example, a student who ranked in the 88th percentile in grade-3 and the 94th percentile in

[5] All MAP scores used in this analysis are standardized by grade, subject, and year. I look for ceiling effects in each grade-subject cell using the methodology of Koedel and Betts (2010) and find no evidence of their presence in the analyzed data.

[6] Like many states, Missouri's standardized exam regimen begins in grade-3 and ends in grade-8, with end-of-course exams replacing grade-based exams from grade-9 onward.

[7] In an alternative definition, students were flagged as early high flyers if they scored in the top 10% of their grade cohort in one subject on the MAP exam (either mathematics or communication arts) and *no worse than* the top 20% on the other subject in *both* grade-3 and grade-4. Aside from gender composition issues (males are over-represented in the high flyers sample under the primary definition, while females are over-represented under the alternative definition), the results from this alternative definition are very similar to those presented in the main analysis and are available from the author upon request.

grade-4 would be identified as an initial high flyer, while a student that scored in 88th and 81st percentiles, respectively, would not.

This definition is appealing for several reasons. First, the focus on mathematics scores is supported by research indicating that early mathematics performance is the best predictor of future academic success (see Duncan et al., 2007, for a meta-analysis of the research).[8] Moreover, the requirement that students must meet a ranking criterion for two consecutive years reduces the role of measurement error in the designation of high-flyer status (although, of course, it does not eliminate the measurement-error problem entirely).

Table 1 presents the demographic characteristics of the initial high flyers cohort compared to the entire population of Missouri grade-3 students in 2006. As can be seen, even at this relatively early starting point, there are stark differences between high flyers and the general student population. In particular, high flyers are much less likely to be of low socio-economic status (SES) as proxied by free/reduced-price lunch eligibility, or a disadvantaged minority (black or Hispanic).[9] Moving forward, it is important to keep in mind that the results described in this study likely under-estimate the true impact of schooling differences on the performance of high flyers, given the large discrepancies already present by grade-3.[10]

4.2 Outcomes Considered

I begin by examining whether initial high flyers maintain their high flyer status until the end of the standardized testing regime in grade-8. To explore this question, the sample of initial high flyers is divided into "soaring" and "falling" subgroups based on the results on their grade-7 and

---

[8] It should be noted that Duncan et al. (2007) focuses on student skills at school entry rather than grade-3, the initial grade examined in this study.

[9] This "high performance gap", i.e. the under-representation of poor and minority students at the top of the exam score distribution, has been noted by other authors, e.g. see Olszewski-Kubilius and Clarenbach (2012).

[10] Like in most states, standardized testing in Missouri does not begin until grade-3, which means that high flyers cannot be identified prior to this point. However, if earlier testing data become available, extensions of this work using the earlier data will be valuable.

grade-8 mathematics exams. The definition used to determine soaring high flyers is parallel to the definition used above to define initial high flyers. Specifically, initial high flyers must score in the top ten percent on either their grade-7 or grade-8 mathematics exam and not fall outside the top twenty percent in either of those grades to be considered "soaring." Initial high flyers that do not meet these criteria are placed into the "falling" subgroup.[11]

Given the above definitions, the following model is estimated as a probit to explore the factors that determine whether initial high flyers maintain their status through the middle grades:

$$HF_i = Z_i^{06/07}\mathrm{B} + X_i\Gamma + \mathrm{S}_i\Lambda + \varepsilon_i. \tag{1}$$

In (1), the dependent variable, HF_i, is an indicator for whether initial high flyer i is still soaring in grades 7 and 8. $Z_i^{06/07}$ is a vector containing the student's two-year average MAP scores (2006 and 2007), separately in mathematics and communication arts. X_i is a vector of student demographic characteristics including race, gender, free/reduced price lunch (F/RL) eligibility, special education status (as measured by whether the student has an individualized education plan (IEP)), and English as a second language (ESL) status. S_i is a vector of school-level characteristics that includes measures of school quality (described below), average student achievement, the number and share of initial high flyers in high flyer i's grade cohort, total enrollment in high flyer i's grade cohort, and aggregates of the student-level characteristics included in X_i. All school-level control variables are weighted averages of the values for all of the schools attended by the student over the course of the

[11] Given that the high flyer definition used in this paper is based on percentile rankings, high flyer status represents a zero-sum game. Still, to the extent that the composition of the high flyers group changes systematically over time, certain groups within the population are being underserved by the educational system. It should also be noted that other high flyer definitions could be considered. For example, Xiang et al. (2011) use the national norming sample percentile cut-offs, rather than percentile cut-offs based on the estimation sample, when defining high flyers. However, this definition still suffers from many of the criticisms relevant to the definition used in this study. Alternatively, one could define high flyers by setting a specific level of knowledge that any top performing student in a specific grade should have. Unfortunately, for this "knowledge threshold" to be meaningful, the exams used must both be properly vertically scaled across grades and have cut-off values for each grade that represent equivalent, grade-appropriate high achievers' knowledge levels. Research on this issue suggests that many commonly used standardized assessments may not meet the first criterion (Ballou, 2009), and the second criterion is particularly hard to assess.

data panel, where the weights are the number of years enrolled in the given school. More details on these variables are provided below.

It is also of interest to examine longer-term outcomes for high flyers. For example, are soaring high flyers more likely to graduate than their falling counterparts? Do they attend college at different rates, and for those that attend college, are there systematic differences in their choices of college and major?[12] How do these factors affect eventual wage and employment rates? Given that the high flyers studied in this paper will not graduate from high school until 2015, many of these questions must be left for future research. However, an intermediary outcome that can be studied currently is the grade in which initially high performing students take Algebra I. To explore this outcome the following model is estimated via Ordinary Least Squares (OLS):[13]

$$G_i = Z_i^{06/07}\mathrm{B} + X_i\Gamma + \mathrm{S}_i\Lambda + \varepsilon_i \qquad (2)$$

In this equation, G_i is the grade in which student i took Algebra I, and the remaining variables are defined exactly as in equation (1).[14]

4.3 Factors Influencing High Flyer Success

I focus on three school-level factors that are likely to influence outcomes for high flyers – school quality, school achievement, and the presence of additional high flyers in the school. Controls

[12] An important recent study in this literature looks at college application and attendance decisions of high achieving, low-income students (Hoxby and Avery, 2012). The authors find that these students largely do not apply to selective colleges and universities despite the fact that, given generous financial aid packages, they often would pay *less* at these institutions than the less selective colleges to which they do apply. Moreover, these students are accepted and graduate at high rates from these high quality institutions in the instances when they do apply to them.

[13] The model was also estimated as a tobit and ordered probit and qualitatively similar results were obtained.

[14] There is a current policy debate over the optimal grade in which Algebra I should be taken. For example, Clotfelter et al. (2012) find that accelerated Algebra I course taking led to lower average results on Algebra I end-of-course exams and often produced negative results that extended into subsequent math classes such as Algebra II or Geometry. However, correlational evidence from other studies suggests that for students who are academically ready for more rigorous material, accelerated course taking might lead to positive downstream outcomes (Smith, 1996; Ma, 2005a; Ma, 2005b). The fact that nearly 70 percent of the high flyers studied in this paper take Algebra I prior to grade-9 (Table 7) could be interpreted as revealed-preference evidence supporting the latter argument.

for these factors are all included in the S_i vector in equations (1) and (2) above. In this section, I elaborate on how these measures are constructed for the analysis.

The school-quality measures used in equations (1) and (2) are calculated following a three-step procedure, the first step of which is an auxiliary value-added model of the following form (see Harris et al. (2011) for a discussion of value-added models):

$$Z_{ijkt} = Z_{ijk(t-1)}\beta_1 + Z_{ij\tilde{k}(t-1)}\beta_2 + X_{it}\beta_3 + \phi_t + \theta_j + \epsilon_{ijkt} \quad (3)$$

where Z_{ijkt} is the (standardized) exam score from student i at school j in subject k ($\tilde{k}$ represents the off-subject score; e.g., communication arts in the model where the mathematics score is the dependent variable) in time t, X_{it} is a vector of student-level demographic controls for student i in time t, ϕ_t are year effects, θ_j represents a vector of school fixed effects, and ϵ_{ijkt} is the error term.[15] The X-vector contains controls for student F/RL eligibility, race, gender, special education status, ESL status, and an indicator for whether the student was in the school where the exam was taken for the entire school year.[16]

The parameters of interest in equation (3) are the school fixed effects, θ_j. Separate fixed-effect estimates for models of mathematics and communication arts achievement are obtained. These subject-specific value-added measures could be included directly in equations (1) and (2). However, doing so risks creating inference problems resulting from the inclusion of two highly collinear variables in the same regression model. An alternative would be to simply include school quality measures for mathematics, given that the main focus of this paper is on mathematics outcomes, but this results in a loss of predictive information. To overcome these issues, I use a

[15] If a student has a missing off-subject lagged exam score, then the missing value is set to zero (the standardized mean) and a missing score dummy variable is initialized. In addition, this dummy variable is also interacted with the student's same-subject lagged exam score, essentially assigning more predictive weight to this value in the presence of missing data. If the student has a missing same-subject lagged exam score, the student is dropped from the analysis.

[16] Standard errors were clustered at the student-level to control for repeated student observations over time and were calculated to be robust in the presence of heteroskedasticity.

method developed by Lefgren and Sims (2012) that allows for value-added measures from both subjects to be weighted appropriately and included in equations (1) and (2) as a single composite measure. Applying this method, the second step in the three-step procedure is to estimate the following regression model for each school j:

$$\gamma_{1011,j}^{math} = \gamma_{0709,j}^{math}\delta_1 + \gamma_{0709,j}^{com}\delta_2 + \eta_j \tag{4}$$

where $\gamma_{1011,j}^{math}$ is school j's value-added measure for mathematics estimated using pooled 2010 and 2011 data, $\gamma_{0709,j}^{math}$ is school j's value-added measure for mathematics estimated using pooled data from 2007 through 2009, and $\gamma_{0709,j}^{com}$ is school j's value-added measure for communication arts estimated using pooled data from 2007 through 2009.[17] Hence, the model in (4) uses student growth from both subjects in the early part of the data panel to predict student growth in mathematics over the later portion of the panel. In the third step of the process, coefficients from the estimation of equation (4) are applied to the school effects $\hat{\theta}_j$ estimated in equation (3) to create the composite value-added measures included in equations (1) and (2). Specifically, the school quality values are calculated as follows:

$$\hat{\theta}_j^{composite} = \hat{\theta}_j^{math}\hat{\delta}_1 + \hat{\theta}_j^{com}\hat{\delta}_2 \tag{5}$$

where $\hat{\delta}_1$ and $\hat{\delta}_2$ are taken from the estimation of equation (4).[18]

It should be noted that in equation (4), the school's mathematics value-added score is chosen as the outcome variable in the weighting equation. This specification was selected because Lefgren and Sims (2012) find that mathematics value-added is a much stronger predictor (relative to communication arts) of future value-added in both mathematics and in a simple average of mathematics and communications arts. However, alternative composite value-added measures were

[17] The school effect estimates used for these regressions were estimated using *all students* in *all schools* in Missouri over the course of the panel.

[18] For equation (4), $\hat{\delta}_1 = 0.577$ and $\hat{\delta}_2 = -0.038$.

calculated and included in separate models to examine the sensitivity of the main findings. The results from these alternative specifications are largely consistent with the results presented in Tables 5 and 11.[19]

The above procedure has the added benefit of producing shrinkage estimates of the school quality measures (Lefgren and Sims, 2012). Shrunken estimates are preferred when using value-added measures as independent variables in regression analyses – recall that the school quality measures will be inserted as right-hand side variables in equations (1) and (2) – because shrinkage techniques help to correct for measurement error in the effect estimate, reducing attenuation bias (Chetty et al., 2011; Jacob and Lefgren, 2008; for a more extensive treatment of this issue, see Appendix C of Jacob and Lefgren, 2005).[20]

I use two separate value-added measures in the analysis that follows – one estimated via equation (3) for all students attending the school who were *not* in the high flyers' grade cohort, and another estimated using all students in the high flyers' grade cohort who were in the bottom 50 percent of the 2006 grade-3 mathematics score distribution. The inclusion of both of these value-added measures represents an effort to disentangle overall school quality effects from the effects of resource shifting from one side of the achievement distribution to the other.

The non-cohort value-added measure (henceforth referred to as the "overall VAM") is included to capture overall school quality, as it contains information on every student who attended the school over the course of the panel *except* those in the high flyers' grade cohort. The high flyers' own grade cohort is excluded from this measure because of standard concerns about endogeneity

---

[19] One alternative weighting scheme involves using the simple average of the mathematics and communication arts value-added estimates as the outcome variable in equation (4). The resulting weights were 0.301 for mathematics and 0.253 for communication arts. In addition, a third composite value-added measure was calculated using the appropriate weights taken from Lefgren and Sims (2012), Table 2. These values were 0.765 for mathematics and 0.030 for communication arts.

[20] The subject-specific school quality measures presented in the descriptive tables were shrunk using a procedure similar to Koedel et al. (2012).

(i.e., I do not want the school-quality measure I use in equations (1) and (2) to be directly influenced by achievement growth for high flyers), as well as to remove any shocks specific to the high flyers' cohort from the overall school quality measure. The exclusion of the high flyers' grade cohort is conceptually similar to a jack-knife procedure.

In contrast, the value-added estimate based on the sample of below median students in the high flyers' grade cohort (the "low performers VAM") is designed to measure how effectively the school is promoting academic growth among the low performing classmates of the high flyers. As mentioned in the introduction, a reasonable hypothesis is that schools facing accountability pressure might shift resources from high performing students to their lower performing counterparts. If this is the case, one might expect that the schools eliciting high growth from low performing students will see slower achievement growth from top students, holding overall school quality constant.

Next, I turn to measuring school-level student achievement. Total achievement is measured as the average MAP score for *all students* who took the MAP exam in the school in the given year (not just those in the high flyers' grade cohort). Similarly to the school quality measures, a single composite measure of school achievement is used throughout. This composite is calculated using a straightforward extension of the procedure shown in equations (4) and (5). Here, equation (4) is replaced by:

$$M_j^{11} = M_j^{0610}\omega_1 + C_j^{0610}\omega_2 + \mu_j \tag{6}$$

where M_j^{11} is school j's average mathematics MAP score in 2011, M_j^{0610} is school j's average mathematics MAP score from 2006-2010, and C_j^{0610} is the comparable average for communication arts. Paralleling the school-quality calculations, the estimated coefficients from (6) are then applied to the school-level average mathematics and communication arts values to create the composite value.[21]

Finally, I construct measures of peer quality. As noted in the introduction, several studies in the peer effects literature have suggested that high performing students are particularly sensitive to the quality of their peers (Hoxby and Weingarth, 2006; Imberman et al., 2012; Burke and Sass, 2013; Lavy et al., 2012). Although information on the average student in each high flyer's school is captured by the school achievement variable, this is likely an insufficient control for peer effects given their non-linear nature (Hoxby and Weingarth, 2006; Imberman et al., 2012). Hence, both the number of high flyers attending the school as well as their share of overall grade-level enrollment are included as school-level controls in vector S_i.

In addition to these three key variables of interest, student-level controls for race, F/RL eligibility, special education status, and ESL status are also controlled for in equations (1) and (2).[22] For the time-varying characteristics included in this vector (F/RL eligibility, special education status, and ESL status), the measure used in the models is the *total* number of times the condition was met over the course of the panel and, as such, can vary from zero to six. Hence, the marginal effects from these controls can be interpreted as the impact of meeting the relevant criterion for one additional school year. As an example, the marginal effect of F/RL eligibility estimated in equation (1) represents the change in the probability of maintaining high flyer status through the end of the panel when the number of years of F/RL eligibility increases by one. Moreover, all of the student-level controls listed above are aggregated at the school-level and included in S_i.

---

[21] For equation (6), $\widehat{\omega}_1 = 0.832$ and $\widehat{\omega}_2 = 0.118$. Mathematics is chosen as the outcome variable in equation (6) for the reasons described above. However, results from models with an alternative weighting scheme using the simple average of mathematics and communication arts scores as the outcome variable were also estimated, and the results were robust across specifications. The weights from this alternative were 0.492 for mathematics and 0.467 for communication arts.
[22] Special education status indicates that the student has an individualized education plan (IEP), which can cover a wide variety of disabilities including ADHD, dyslexia, and behavioral issues.

5. Results

5.1 Maintaining High Flyer Status

Table 2 presents student-level demographic characteristics of the initial high flyer cohort broken down into the soaring and falling subgroups. Note that nearly two-fifths (2179 out of 5641) of the initial high flyers have lost their high flyer status by the end of the panel. While it is important to keep in mind that the decline in exam scores necessary for this to occur need not to have been dramatic (e.g., a student who scores at the 85^{th} percentile in grades 7 and 8 would be identified as losing high-flyer status by the above definition), Figure 1 shows substantial performance declines are not uncommon. In fact, from grade-5 onward roughly 20 percent of the initial high-flyer cohort has a mathematics test score in the 50^{th} to 80^{th} percentiles, and a smaller group of initial high flyers drop all the way into the bottom half of the score distribution.

Tables 3 and 4 present the characteristics of the schools that high flyers attended, once again broken down by students who do and do not maintain high-flyer status. Table 3 presents school characteristics for high flyers in grade-3 (2006) and grade-8 (2011). To mitigate the influence of district changers, the sample used to construct the table is restricted to only those high flyers who attended schools within the same district for both grades. In this way, the table allows for a fairly straightforward comparison of how the characteristics of schools attended by soaring and falling high flyers change as the result of structural school changes; e.g. moving from small, neighborhood elementary schools to larger, more diverse middle and junior–high schools, as opposed to changes resulting from mobility across districts.

Looking first at the school-level achievement metrics, two findings stand out. First, soaring high flyers initially attend higher achieving schools relative to falling high flyers, but the gap is not large. In 2006, the school-level average MAP score in mathematics for schools attended by soaring high flyers is 0.251, while the corresponding value for falling high flyers is 0.234. In communication arts, the comparable numbers are 0.217 and 0.178.[23] Second, both soaring and falling high flyers move from higher achieving schools to lower achieving schools, on average, between grade-3 and grade-8. However, the drop is particularly stark for falling high flyers. The school-level average mathematics MAP score for the 2011 school attended by falling high flyers is 0.147 points lower than in 2006. Communications arts scores also fall by a large amount, 0.083. In contrast, the average scores for soaring high flyers fall by only 0.042 and 0.034, respectively.

To summarize, both soaring and falling high flyers *begin* their school careers in high achieving schools. Although their schools perform worse in grade-8, likely in large part because of the merging of elementary schools, both groups move to middle schools with above average achievement.[24] However, the middle schools attended by falling high flyers are much closer to the statewide average than the middle schools attended by their soaring counterparts.

For the school quality measures, the 2006 results mirror those for school achievement.[25] Specifically, although both groups attend schools with above average quality, soaring high flyers start out in schools that are producing more test score growth than the schools attended by their falling counterparts. As was the case with school achievement, both subgroups see a drop in school quality as they move from grade-3 to grade-8. However, by the end of the panel, falling high flyers are no longer attending above average schools. In fact, the schools attended by falling high flyers in grade-8 are producing below average growth in both subjects. This is particularly true in mathematics, where

[23] The difference between the mathematics averages is marginally significant, while the difference between the communication arts averages is significant at the 0.01 level. Perhaps not surprisingly given the way high flyers are defined in this paper, the mathematics achievement levels are higher than communication arts in both instances.
[24] Examining district-level achievement provides additional insight into this matter. Specifically, soaring high flyers attend schools with average achievement that is higher than the district average in both 2006 and 2011. In contrast, falling high flyers start in schools with achievement that is above the district average but end the panel in schools with achievement that is at or below the district average.
[25] It is also important to note that the values reported in Table 3 are taken from the overall VAM that excludes all students from the high flyers' grade cohort.

the average value-added estimate for schools attended by falling high flyers in 2011 is -0.011, a drop of 0.058 from the grade-3 level.[26]

Table 4 repeats the comparative analysis from Table 3 on the subsample of disadvantaged minority high flyers. Minority high flyers are on average attending schools with lower achievement levels than their non-minority counterparts. In all but one case (grade-3 mathematics), falling minority high flyers are attending schools with below average achievement, while even soaring minority high flyers are attending schools with average achievement that is more than 0.1 standard deviations lower than schools attended by soaring high flyers overall. Interestingly, average achievement in the schools attended by soaring minority high flyers in grade-8 is similar to the school achievement for falling high flyers in the larger sample. For the most part, the school quality measures presented in Table 4 are also lower than in Table 3, particularly in mathematics.[27] Hence, minority high flyers are generally attending schools with lower achievement than non-minority high flyers, and their schools are also doing slightly worse with respect to test score growth.[28]

Aside from the obvious differences between schools attended by minority high flyers and those attended by high flyers on average, many of the other patterns present in Table 3 can also be observed Table 4. Throughout the panel, falling minority high flyers attend lower achieving, lower quality schools than soaring minority high flyers, and although they begin the panel at higher quality schools, by grade-8 falling minority high flyers are attending schools with below average test score growth.

[26] It is worth noting that the 2011 gap in communication arts value-added between schools attended by soaring and falling high flyers is actually smaller than in 2006.

[27] The one exception being that soaring minority high flyers start in schools of the same quality as soaring high flyers in the larger sample.

[28] Interestingly, both soaring and falling minority high flyers attend schools where the average achievement levels are higher than the corresponding district averages. For falling high flyers, the district-level averages are more than 0.1 standard deviations below the statewide average in all instances.

Table 5 presents the average marginal effects of selected variables from equation (1). The table also shows results from an OLS model where the continuous 2011 MAP mathematics score is the dependent variable. Six different specifications of the probit model are presented, from sparse models that only include student-level or school-level variables to the full specification presented in equation (1). First consider the average marginal effect of the low performers VAM on maintaining high flyer status. In column 4, which includes all of the control variables *except* the overall VAM, the average marginal effect of the low performers VAM is large and positive. This is not surprising, considering that it is likely serving as a proxy for overall school quality in the absence of the more comprehensive measure. However, when overall school quality is explicitly controlled for via the inclusion of the overall VAM (column 6), the coefficient on the low performers VAM, although decreasing in absolute size, remains positive, statistically significant, and educationally meaningful. In fact, based on the estimates presented, moving to a school that does 0.25 standard deviations better with low performing students increases the probability that a high flyer maintains his or her status by nearly nine percentage points.[29] From a baseline of 61 percent of high flyers maintaining their status, this corresponds to an over 14 percent increase. In other words, schools that are doing well with their low performing students also seem to be doing well with their high performers.[30] In fact, this statement is further supported by Table 6, which presents correlations between school value-added measures estimated on various subsamples of the population. Specifically, the mathematics school effect estimated from the sample consisting solely of the cohort of high flyers has a correlation of

[29] The mean of the low performers VAM is 0 with a standard deviation of 0.122. Hence, an increase of 0.25 is roughly equivalent to moving from a school one standard deviation below the mean to a school one standard deviation above. For comparative purposes, the mean and standard deviation of the overall VAM are 0 and 0.073, respectively.

[30] Recall from above that the high flyers themselves are not included in either of the value-added measures used in the model. Hence, both the overall VAM and the low performers VAM are estimated independently from the sample used for the probit estimation.

nearly 0.4 when compared to the mathematics school effect estimate derived from the low performers sample.[31]

A second interesting result concerns the effect of the presence of additional high flyers on high flyer performance. In Table 5, as both the number and the share of high flyers increase within a school, high flyers are *less likely* to maintain their high flying status through the end of the panel. Moreover, the effects are sizeable, highly significant and consistent across all model specifications.[32] Hence, given these results, it appears that high flyers may actually do *worse* in the presence of other high flyers, holding overall school quality and school achievement constant.

At first glance, this result appears to be inconsistent with the peer effects literature (Hoxby and Weingarth, 2006; Imberman et al., 2012; Burke and Sass, 2013). However, a closer reading of these papers helps to account for the apparent difference in findings. For example, Hoxby and Weingarth (2006) examine student desegregation reassignments in Wake County School District (North Carolina) and find evidence supporting the boutique model in which all students benefit from having peers of roughly the same academic level.[33] However, when looking only at particularly high achievers – specifically students in the top decile – they find that an increase in the number of additional top decile students in the classroom actually *lowers performance* (see Figure 1 in Hoxby and Weingarth, 2006), a finding that is in line with the results found here.

---

[31] Note that the correlations are generally smaller for communication arts, although they remain positive and statistically significant at the 0.001 level. It is also interesting to note that the high flyer VAM measures are more highly correlated with the same-cohort low performers VAM than with the overall VAM that include high flyers from other grade cohorts. This suggests that cohort specific shocks may well play an important role in measuring value-added.

[32] Given the average values presented in Table 3, an increase of ten high flyers in a school would directly lower the probability of maintaining high flyer status for all high flyers by 3%. Moreover, a nominal increase of this size would boost the share of high flyers in the school by 3-10%, reducing the probability by an additional 2.4-8%. Hence, the total effect would be the between 5.4% and 11%.

[33] However, it should be noted that Hoxby and Weingarth (2006) also find evidence supporting the focus model, in which a student benefits from being in a relatively homogenous environment, even if the student is not a member of the dominant group. Under this model, a high flyer in a classroom dominated by low performers could actually do *worse* following the arrival of additional high performing students if the resulting classroom composition changed by enough to strongly bifurcate the performance distribution.

In a more recent paper, Imberman et al. (2012) analyze the effects of Hurricane Katrina evacuees on native student performance in the receiving communities and find that the arrival of high achieving students improves the performance of all students while the reverse is true for the arrival of low achieving students (i.e. monotonic peer effects). This effect is particularly strong among high performing native students (Imberman et al., 2012). However, the highest performing group in the Imberman et al. (2012) study is drawn from the top quartile of the performance distribution and, hence, encompasses a larger proportion of the student population than the high flyers in this study.[34] This broader definition likely contributes to the seeming contradiction between the Imberman et al. (2012) results and the findings presented here. For example, Hoxby and Weingarth (2006) show that students in the ninth decile see positive peer effects when placed with other ninth decile students; the negative high flyer to high flyer peer effect is only seen at the very top of the performance distribution.

Of the papers in the peer effects literature, the one the most closely parallels this study's focus on high flyers is Lavy et al. (2012), who use national testing data from English secondary students to analyze peer effects, paying particular attention to how the presence of students at the very top and bottom of performance distribution affect their peers. These authors find that the presence of additional top-performing students (in the top 5 percent) has a *negative* effect on other high performing male students (Lavy et al., 2012).

The findings presented here, in conjunction with the Hoxby and Weingarth (2006) and Lavy et al. (2012) results, warrant some discussion of possible mechanisms. In contrast to the competition for scarce resources hypothesis discussed in the introduction – i.e., that low performing and high performing students are competing with one another for scarce school resources – it may be the

[34] Burke and Sass (2013) also use a wider definition for their top performing group, specifically the top quintile of all students.

case that the competition for resources is taking place *within* the high flyers group. For example, teacher attention might be in limited supply, and as a result, the addition of more high flyers into a school setting could result in less attention for each student. In addition, schools with many high flyers may be less concerned – or less able to notice – if one of them slips moderately in the academic standings, especially if they are still firmly above average. Such a mechanism is in line with what Lavy et al. (2012) refer to as a "crowding out effect". Finally, peer effects could explain this negative result if they are of the "invidious comparison" variety. This might particularly be the case in male-dominated high flyers' environments, as males seem particularly sensitive to these types of negative comparisons (Lavy et al., 2012).

5.2 Grade in which Algebra I is Taken

Moving onto the second outcome measure, Table 7 presents the distribution of Algebra I course taking for the high flyers sample.[35] Before turning to the results, however, a few points regarding the data are worth noting. Given that student course records are only available in Missouri from 2009 onward, the precise grade in which Algebra I was taken for students who took the course prior to grade-7 cannot be determined. Hence, students are placed in a "grade-7 and under" category in these instances (this category accounts for 3.6 percent of the analytic sample). On the other end of the spectrum students who had not taken Algebra I by grade-9, the last year of the data panel, are

[35] The results presented in this section explore the factors that determine the specific grade in which high flyers take Algebra I. A related question explores the factors that determine the grade in which high flyers take Algebra I *relative to the district mode*. This alternative analysis examines the characteristics of schools that have the resources to accelerate high flyers' coursework given their own district curricular policies. Although the primary results from this parallel research question are similar to those presented in the main body of this paper, there are some interesting differences. First, the effect of school achievement is roughly half the size of that reported in Table 11. In addition, the overall VAM effect is now large and negative, indicating that students attending high quality schools are more likely to take Algebra I early relative to the district mode even though they are not more likely to take Algebra I early in a global sense.

placed into a single "grade-10 or higher" category, which accounts for an additional 5.3 percent of high flyers.[36]

Table 7 shows that nearly two-thirds of high flyers took Algebra I in grade-8, while another quarter took it in grade-9.[37] Given that the statewide modal Algebra I grade is ninth grade (approximately half of students in Missouri take Algebra I in grade-9), these results indicate that high flying students are taking the course earlier than the typical student in the state.

Table 8 presents student-level characteristics of high flyers broken down by whether they took Algebra I before, during, or after grade-9. From this table, it is apparent that soaring high flyers are more likely to take Algebra I early. Specifically, nearly 69 percent of initially high performing students who took Algebra I early are still high performing according to standardized exam performance in grade-8. In contrast, less than fifty percent of the high flyers cohort that took Algebra I at or after grade-9 are still high flyers in grade-8. Of course, this comparison is strictly descriptive, as maintaining high flyer status is endogenous to Algebra I course taking. That said, there are also important differences along demographic lines. For example, high flyers who take Algebra I early are significantly less likely to be F/RL eligible than those who take it during grade-9, while the reverse is true for those who take Algebra I late.

There are a number of interesting patterns in the data in Tables 9 and 10. One consistent finding is that the average school attended by high flyers who take Algebra I late is serving a student

[36] Additionally, 163 students (2.9 percent of the larger high flyers sample) attend districts that do not offer a traditional Algebra I course. These students were dropped from this portion of the analysis. As a final data note, some students appear in the course record files with no Algebra I course records but course records for a higher math class. In these cases, the students were assigned an Algebra I grade for the same grade in which the higher math course was taken. These students account for 7.6% of all cases. However, models that exclude these students produce results consistent with those presented below.

[37] Interestingly, although all of the students analyzed are high performing mathematics students, less than four percent took Algebra I before grade-8. In fact, a slightly larger percentage (5.3 percent) took Algebra I in grade-10 or later than took it in grade-7 or earlier. Clotfelter et al. (2012) find negative impacts among high performing students who were accelerated into Algebra I in grade-7. Hence, the fact that very few high flyers in Missouri are taking Algebra I prior to grade-8 can be seen as a positive outcome.

population with lower overall achievement than the average school attended by high flyers who take Algebra I early or in grade-9. However, the differences in school quality across types are typically smaller than the differences in achievement levels and are often insignificant, particularly for mathematics, suggesting that schools attended by high flyers *are* doing comparably well with the student populations that they are serving. Hence, it may be the case that schools are structuring their course sequences to serve the typical student, so that disadvantaged schools are pushing students to take Algebra I later, with high flyers in these schools inadvertently getting caught up in this policy.

Table 11 presents the results from equation (2). The model is estimated via OLS; however, tobit and ordered probit specifications were also estimated and returned qualitatively similar results. Like in Table 5, Table 11 presents the results from a variety of model specifications, some of which omit certain sets of variables from the full specification shown in equation (2).

The most interesting results from Table 11 are the effects of school achievement levels and the low performers VAM. Looking at the full model, there is a large, positive, and highly significant coefficient on the low performers VAM and a large, negative, and highly significant coefficient on the school achievement variable. In other words, high flyers who attend high achieving schools are more likely to take Algebra I early, while high flyers who attend schools that do particularly well with low performing students (those in the bottom half on the 2006 grade-3 mathematics distribution) are more likely to take Algebra I late. For example, if school achievement increases by 0.8 standard deviations, high flyers are likely to take Algebra I more than one-half of a school year earlier. In contrast, if a high flyer moves to a school that does 0.25 standard deviations better with low performing students, the grade in which Algebra I is taken increases by nearly three-tenths of a school year.[38] It is also worth noting that the overall VAM is not significant in the full model, whereas it is marginally significant and positive in the model where the low performers VAM is omitted. In this case (column 5 of Table 11), the overall VAM is likely picking up part of the effect of low performing student growth on the grade in which Algebra I is taken.

The results from Tables 9, 10, and 11 present a fairly clear picture of the factors that affect the grade in which initial high flyers take Algebra I. Specifically, high flyers who take algebra late appear to be attending schools that serve a more disadvantaged and lower achieving student population. As a result, these schools seem to be delaying Algebra I instruction. This strategy appears to be in line with improving the academic growth of low performing students. However, as an unintended side effect, it is also resulting in high flyers delaying their algebra course taking.

6. Discussion and Conclusion

I follow a cohort of high performing students from the time of their first statewide assessment in grade-3 into their early high school years. I examine the factors that affect whether these students continue to attain high scores on standardized achievement tests and at what point in their schooling careers they take Algebra I. Three important results emerge from this analysis.

First, I find no evidence of large-scale resource shifts away from high performing students to low performing students. Schools that do well with low performing students are also generally supporting academic growth among high flyers. Given this result, concerns that federal policy focusing on proficiency rates and closing the achievement gap has harmed high performing students may be overstated.

However, when attention shifts from exam performance to course sequencing, these issues become more salient. Specifically, high flyers who attend schools that serve low achieving students

[38] As noted previously (footnote 29), the standard deviation of the low performers VAM is 0.122. For the average student achievement variable, the school-level standard deviation is roughly 0.4. Hence, the numbers presented above represent similarly-sized moves within the respective distributions of the two variables.

are likely to take Algebra I later than comparable high flyers attending higher achieving schools, although schools serving both types of students appear to be eliciting similar levels of mathematics growth from their respective student bodies, at least in the eighth grade. One possible explanation is that low achieving schools are purposefully slowing the mathematics course sequence for their students. This policy seems to be effective in promoting growth for the average student in these schools. However, as an unintended consequence, it also appears to slow student progress in mathematics for high achieving students who are ready to proceed at a more advanced pace.

Finally, this paper finds that high flyers may perform worse when they are exposed to more high flying peers. This result has received little attention in the peer effects literature, and runs counter to the belief that top students benefit from being around other top students. Resource allocation issues may help to explain this result. In an environment where high flyers compete for scarce resources, like teacher attention, the presence of more competitors can have negative consequences that may counter any positive spillover effects.

These findings have important policy implications for those concerned with the progress and development of high performing students. First, one of the best ways to help high flying students is simply to improve schools overall. Specifically, soaring high flyers are attending schools that are significantly outperforming the schools attended by falling high flyers. Moreover, schools that do well in promoting student growth appear to be doing well with *all* of their students, from the bottom of the distribution to the top. Hence, simply improving overall school quality will do a great deal to help high achievers.

Second, as high flyers move into the middle and upper grades, policies that allow them to transfer to or take specific courses at schools that serve a more academically prepared student population should be considered. Such a change would give these high flyers, many of whom live in disadvantaged environments, opportunities to accelerate their coursework that they might not otherwise have if they remain in their home schools. This could provide substantial benefits in terms of college readiness, as well as potentially opening wider fields of study for these students once they proceed into postsecondary education.

However, one potential stumbling block to this policy is that schools serving many high flyers may end up doing poorly by them, as an increased number of high achieving students appears to lower performance. Hence, any policy that increases the concentration of high flyers in certain schools should be approached cautiously. To the extent that the negative peer effects among high flyers are the result of a competition for scarce high-end resources, additional policy levers could be applied to correct for this. For example, schools seeing an influx of high flyers could receive additional funds designed to increase advanced course availability, as well as provide services specifically designed for the high flyer population. Of course, to the extent that these negative effects are the result of invidious-comparison peer effects, a particular issue in male dominated environments (see Lavy et al., 2012), the solution is less simple and may require maintaining a proper balance of high flyers across schools.

References

Aaronson, Daniel, Lisa Barrow and William Sander. 2007. Teachers and Student Achievement in the Chicago Public High Schools. *Journal of Labor Economics* 25(1), 95-135.

Argys, L.M., Rees, D.I., & Brewer, D.J. 1996. Detracking America's schools: equity at zero cost? *Journal of Policy Analysis and Management*, *15*(4), 623–645.

Bailey, D. H., Littlefield, A., & Geary, D. C. 2012. The co-development of skill at and preference for use of retrieval-based processes for solving addition problems: Individual and sex differences from first to sixth grade. *Journal of Experimental Child Psychology, 113*, 78-92.

Ballou, Dale. 2009. Test Scaling and Value-added Research. *Education Finance and Policy* 4(4): 351-383.

Ballou, Dale and Matthew Springer. 2008. *Achievement Trade-Offs and No Child Left Behind*. Nashville, TN: Peabody College of Vanderbilt University.

Boyd, Donald, Hamilton Lankford, Susanna Loeb, and James Wyckoff. 2012. Measuring Test Measurement Error: A General Approach. Cambridge, MA: NBER Working Paper No. 18010.

Burke, Mary and Tim Sass. 2013. Classroom Peer Effects and Student Achievement. *Journal of Labor Economics* 31(1): 51-82.

Chetty, Raj, John N. Friedman and Jonah E. Rockoff. 2011. The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood. National Bureau of Economic Research Working Paper No. 17699.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2012. The Aftermath of Accelerating Algebra: Evidence from a District Policy Initiative. Cambridge, MA: NBER. WP 18161 (June).

Dee, Thomas S. and Brian Jacob. 2011. The Impact of No Child Left Behind on Student Achievement. *Journal of Policy Analysis and Management* 30: 418–446.

Duffett, Ann, Steve Farkas, and Tom Loveless. 2008. *High-Achieving Students in the Era of NCLB*. Washington, D.C.: Thomas B. Fordham Institute. .

Duncan, Greg J., Chantelle J. Dowsett, Amy Claessens, Katherine Magnuson, Aletha C. Huston, Pamela Klebanov, Linda S. Pagani, Leon Feinstein, Mimi Engel, Jeanne Brooks-Gunn, Holly Sexton, Kathryn Duckworth and Crista Japel. 2007. School Readiness and Later Achievement. *Developmental Psychology* 43(6), 1428-1446.

Finn, Chester. 2012. Young, Gifted and Neglected. *New York Times*. September 18, 2012.

GreatSchools. 2010. Why is Algebra So Important? Retrieved from http://www.greatschools.org/students/academic-skills/354-why-algebra.gs

Hanushek, Eric and Ludger Woessmann. 2012. Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of Economic Growth* 17: 267-321.

Harris, Douglas N., Tim R. Sass and Anastasia Semykina. 2011. Value-Added Models and the Measurement of Teacher Productivity. Unpublished Manuscript, Florida State University.

Helfand, Duke, A Formula for Failure in LA Schools, *Los Angeles Times* (1-30-2006).

Hoxby, Caroline and Christopher Avery. 2012. The Missing "One-Offs": The Hidden supply of High-Achieving, Low Income Students. Cambridge, MA: NBER. WP 18586 (December).

Hoxby, Caroline M., and Gretchen Weingarth. 2006. Taking race out of the equation: School reassignment and the structure of peer effects. Unpublished manuscript, Department of Economics, Harvard University.

Imberman, Scott A., Adriana D. Kugler and Bruce I. Sacerdote. 2012. Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees. *American Economic Review* 101(5), 2048-82.

Jacob, Brian A., and Lars Lefgren. 2005. Principals as Agents: Subjective performance measurement in education. Cambridge, MA: NBER WP 11463 (July).

-- 2008. Can Principals Identify Effective Teachers? Evidence on Subjective Performance Evaluation in Education. *Journal of Labor Economics* 26(1): 101-136.

Koedel, Cory and Julian R. Betts. 2010. Value-Added to What? How a Ceiling in the Testing Instrument Influences Value-Added Estimation. *Education Finance and Policy* 5(1): 54-81.

Koedel, Cory, Rebecca Leatherman and Eric Parsons. 2012. Test Measurement Error and Inference from Value-Added Models. *The B.E. Journal of Economic Analysis & Policy* 12(1). (Topics).

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

Lee, Jaekyung. 2011. Review of Do High Flyers Maintain Their Altitude. Boulder, CO: NEPC (October).

Lefgren, Lars and David Sims. 2012. Using Subject Test Scores Efficiently to Predict Teacher Value-Added. *Educational Evaluation and Policy Analysis* 34(1): 109-121.

Ma, X. 2005a. Early Acceleration of Students in Mathematics: Does It Promote Growth and Stability of Growth in Achievement Across Mathematical Areas? *Contemporary Educational Psychology* 30: 439-460.

Ma, X. 2005b. A Longitudinal Assessment of Early Acceleration of Students in Mathematics on Growth in Mathematics Achievement. *Developmental Review* 25: 104-131.

Neal, Derek and Diane Whitmore Schanzenback. 2010. "Left Behind by Design: Proficiency Counts and Test-Based Accountability. *Review of Economics and Statistics* 92: 263-283.

Olszewski-Kubilius, Paula and Jane Clarenbach. 2012. Unlocking Emerging Talent: Supporting High Achievement of Low-Income, High-Ability Students. Washington DC: NAGC.

Reback, Randall. 2008. Teaching to the Rating: School Accountability and the Distribution of Student Achievement. *Journal of Public Economics* 92: 1394-1415.

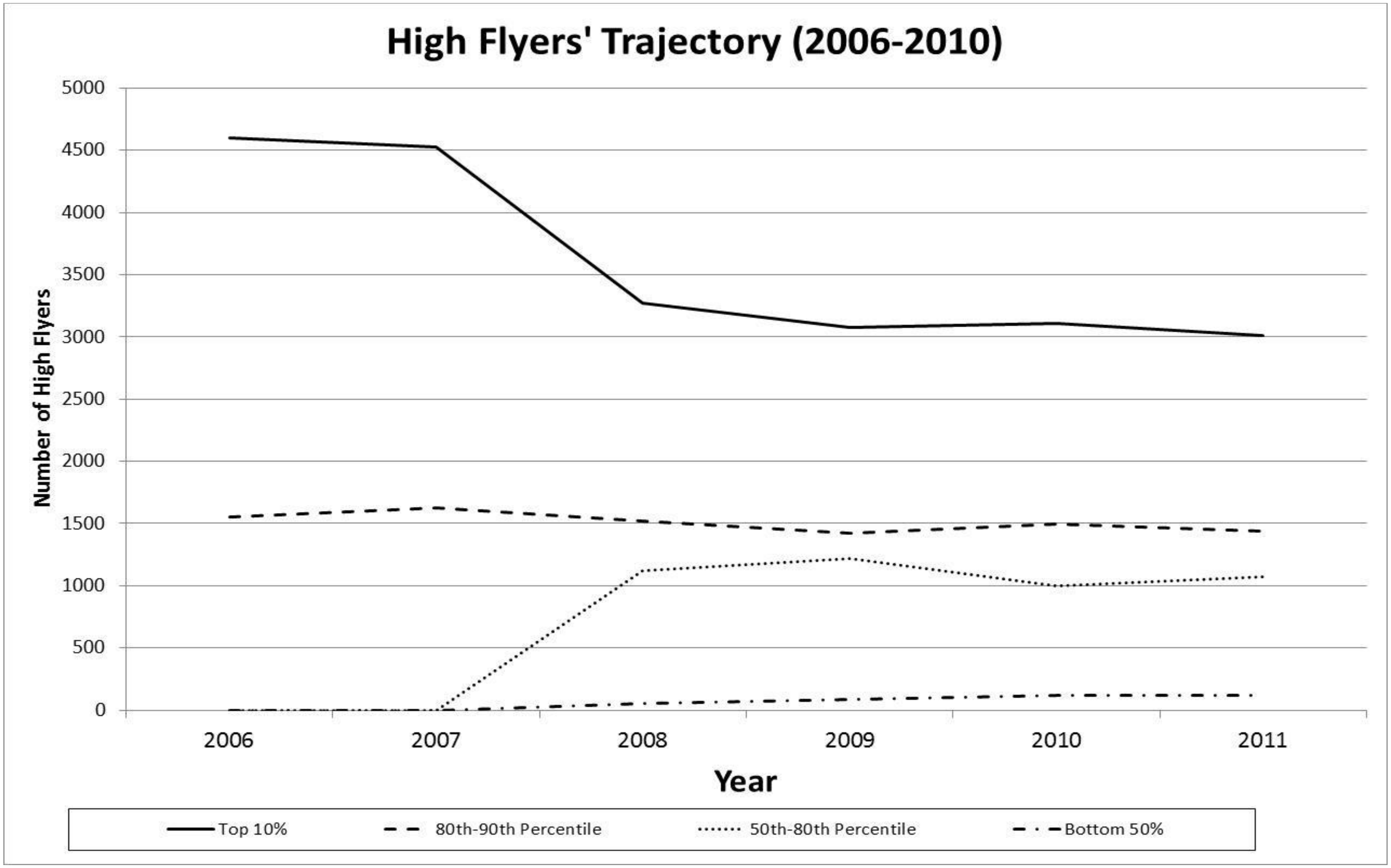
Reback, Randall, Jonah Rockoff and Heather L. Shwartz. 2012. Under Pressure: Job Security, Resource Allocation, and Productivity in Schools Under NCLB. Columbia University Working Paper.

Slavin, Robert E. 1990. Achievement Effects of Ability Grouping in Secondary Schools: A Best Evidence Synthesis. *Review of Educational Research* 60(3).

Smith, J. 1996. Does an Extra Year Make Any Difference? The Impact of Early Access to Algebra on Long-Term Gains in Mathematics Attainment. *Educational Evaluation and Policy Analysis* 18: 141-53.

Xiang, Yun, Michael Dahlin, John Cronin, Robert Theaker, and Sarah Durant. 2011. Do High Flyers Maintain Their Altitude? Performance Trends of Top Students. Washington DC: Thomas B. Fordham Institute (September).

Figure 1. Math MAP Percentile Trajectory of High Flyers.



Notes: This figure charts the number of high flyers that fall into various percentile groupings on the math MAP examination for each year from 2006 to 2011. For example, note that no high flyers fall below the 80th percentile in 2006 or 2007. This is by definition. From 2008 onwards, roughly 3000 high flyers score in the top 10 percent of the exam each year, approximately 1500 score in the 80th-90th percentiles, slightly over 1000 fall between the 50th and 80th percentiles, and a small additional number fall into the bottom half of the score distribution.

Table 1. High Flyer Cohort Demographics.

	Cohort of All Missouri Third Graders in 2006	*High Flyers*
Total Number of Students	64369	6151 (9.6%)
Percent Female	49.1%	47.3%
Percent Free/Reduced Lunch Eligible	44.4	17.8
Percent Black	18.1	3.8
Percent Hispanic	3.5	1.3
Percent Asian	1.8	3.5
Percent White	76.2	91.0

Notes: Students are flagged as high flyers if they scored in the top 10 percent of their state-wide grade cohort on the mathematics MAP examination in either grade-3 or grade-4 (2006 or 2007) and no worse than the top 20 percent in the other grade. All demographic values are taken from the 2006 student records.

Table 2. Student Demographic Characteristics of Soaring versus Falling High Flyers.

	High-Flyers	
	Soaring (n=3462)	Falling (n=2179)
Percent White	92.2%	91.3%
Percent Black	2.3**	5.5**
Percent Hispanic	1.2	1.5
Percent Asian	4.0**	1.2**
Percent Female	46.7†	49.3†
Percent FR/L Eligible	23.1**	39.6**

Notes: ** indicates that the means are significantly different at the 0.01 level, * at the 0.05 level, and † at the 0.10 level. For purposes of this table, students were categorized as FR/L eligible if they were *ever* FR/L eligible over the course of the entire panel.

Table 3. Average Characteristics of School Attended for Soaring versus Falling High Flyers. 2006 and 2011 schools attended.

	High-Flyers	
	Soaring (n=3178)	Falling (n=1881)
2006 School Attended		
Ave. MAP Math Score	0.251†	0.234†
Ave. MAP Com Arts Score	0.217**	0.178**
Ave. VAM Math Effect	0.059**	0.047**
Ave. VAM Com Arts Effect	0.045**	0.025**
Percent Female	48.9%	48.9%
Percent FR/L Eligible	29.5%**	35.8%**
Percent Minority	12.2%†	13.1%†
Number of High Flyers	13.2*	12.7*
Share of High Flyers	16.0%	16.2%
2011 School Attended		
Ave. MAP Math Score	0.209**	0.087**
Ave. MAP Com Arts Score	0.183**	0.095**
Ave. VAM Math Effect	0.016**	-0.011**
Ave. VAM Com Arts Effect	0.008**	-0.005**
Percent Female	49.0%	49.0%
Percent FR/L Eligible	35.4%**	41.4%**
Percent Minority	15.8%	15.9%
Number of High Flyers	33.1	32.8
Share of High Flyers	13.1%**	12.5%**

Notes: ** indicates that the means are significantly different at the 0.01 level, * at the 0.05 level, and † at the 0.10 level. The initial year VAM estimates are taken from the school attended for the 2007 school year, rather than 2006, as a number of high flyers attended schools in 2006 for which VAM school-level values could not be estimated. In addition, the sample of students in this table is limited to those that attended schools within the *same district* in both 2006 and 2011. Hence, the differences in the average school characteristics between 2006 and 2011 indicate structural school differences by grade configurations *within* districts, i.e. students moving from smaller, neighborhood elementary schools to larger, more diverse middle schools, rather than differences observed when mobile students move from one district to another over time.

Table 4. Average Characteristics of School Attended for Soaring versus Falling High Flyers. 2006 and 2011 schools attended. Minority Students Only.

	High-Flyers	
	Soaring (n=110)	Falling (n=130)
2006 School Attended		
Ave. MAP Math Score	0.119*	0.018*
Ave. MAP Com Arts Score	0.096*	-0.018*
VAM Math Effect	0.059	0.039
VAM Com Arts Effect	0.048†	0.028†
Percent Female	49.8%	49.2%
Percent FR/L Eligible	41.8%*	49.7%*
Percent Minority	32.0%**	45.0%**
Number of High Flyers	10.5	9.8
Share of High Flyers	13.8%	12.7%
2011 School Attended		
Ave. MAP Math Score	0.086**	-0.082**
Ave. MAP Com Arts Score	0.071**	-0.077**
Ave. VAM Math Effect	0.000*	-0.024*
Ave. VAM Com Arts Effect	0.002†	-0.015†
Percent Female	50.0%	50.1%
Percent FR/L Eligible	47.5%*	55.4%*
Percent Minority	34.8%**	46.9%**
Number of High Flyers	29.0	25.1
Share of High Flyers	11.0%	9.9%

Notes: ** indicates that the means are significantly different at the 0.01 level, * at the 0.05 level, and † at the 0.10 level. The initial year VAM estimates are taken from the school attended for the 2007 school year, rather than 2006, as a number of high flyers attended schools in 2006 for which VAM school-level values could not be estimated. In addition, the sample of students in this table is limited to those that attended schools within the *same district* in both 2006 and 2011. Hence, the differences in the average school characteristics between 2006 and 2011 indicate structural school differences by grade configurations *within* districts, i.e. students moving from smaller, neighborhood elementary schools to larger, more diverse middle schools, rather than differences observed when mobile students move from one district to another over time.

Table 5. Effects of Student- and School-Level Characteristics on the Probability of Retaining High Flyer Status and on 8th Grade Mathematics MAP Scores.

	Probability of Maintaining High-Flyer Status (Probit)						2011 MAP Math Score (OLS)
	Model 1	*Model 2*	*Model 3*	*Model 4*	*Model 5*	*Model 6*	*Model 7*
	n=5624	n=5641	n=5624	n=5616	n=5624	n=5616	n=5616
							R-squared = 0.244
Student-Level							
2006/2007 Average Math	0.190**		0.200**	0.201**	0.200**	0.201**	0.361**
MAP Score	(0.145)		(0.015)	(0.015)	(0.015)	(0.015)	(0.020)
2006/2007 Average Com	0.170**		0.169**	0.170**	0.173**	0.173**	0.268**
Arts MAP Score	(0.011)		(0.011)	(0.011)	(0.011)	(0.011)	(0.016)
School-Level							
Composite Average MAP		0.520**	0.440**	0.282**	0.107	0.068	0.318**
Score		(0.061)	(0.058)	(0.066)	(0.070)	(0.072)	(0.106)
Number of High Flyers		-0.005**	-0.005**	-0.004**	-0.004**	-0.003**	-0.008**
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Share of High Flyers		-0.010**	-0.010**	-0.009**	-0.008**	-0.008**	-0.013**
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Cohort Enrollment /100		0.040*	0.039*	0.035*	0.033†	0.032†	0.047†
		(0.018)	(0.017)	(0.017)	(0.017)	(0.017)	(0.025)
Composite VAM – Bottom				0.790**		0.352*	0.468*
50% of Cohort Students				(0.156)		(0.168)	(0.236)
Composite VAM – All Non-					1.928**	1.729**	1.934**
Cohort Students					(0.227)	(0.246)	(0.346)
Constant							1.000**
							(0.186)
Student-Level Demographic Controls Included	X		X	X	X	X	X
School-Level Demographic Controls Included		X	X	X	X	X	X

Notes: ** represents significance at the 0.01 level, * at the 0.05 level, and † at the 0.10 level. 61.4 percent of students in the sample retained their high-flyer status over the course of the panel. Values presented for the probit models represent average marginal effects for each of the independent variables.

Table 6. Correlations between School-Level VAM Estimates

Mathematics

	All Cohort Students	All Non-Cohort Students	High Flyers	Bottom 50% of Cohort Students	Composite VAM – Bottom 50%	Composite VAM – Non-Cohort
All Cohort Students	1	0.592	0.521	0.912	0.884	0.579
All Non-Cohort Students	--	1	0.314	0.546	0.545	0.980
High Flyers	--	--	1	0.390	0.377	0.310
Bottom 50% of Cohort Students	--	--	--	1	0.961	0.539
Composite VAM – Bottom 50%	--	--	--	--	1	0.552
Composite VAM – Non-Cohort	--	--	--	--	--	1

Communication Arts

	All Cohort Students	All Non-Cohort Students	High Flyers – Def. 1	Bottom 50% of Cohort Students	Composite VAM – Bottom 50%	Composite VAM – Non-Cohort
All Cohort Students	1	0.628	0.387	0.922	0.432	0.371
All Non-Cohort Students	--	1	0.208	0.590	0.344	0.572
High Flyers	--	--	1	0.246	0.168	0.098
Bottom 50% of Cohort Students	--	--	--	1	0.457	0.363
Composite VAM – Bottom 50%	--	--	--	--	1	0.552
Composite VAM – Non-Cohort	--	--	--	--	--	1

Note: All correlations have a p-value of < 0.0001

Table 7. Distribution of the Grade in which Algebra I is Taken among High Flyers.

Grade Taken	*Frequency*	*Percentage*
7th or earlier	198	3.6%
8th	3516	64.2
9th	1473	26.9
10th or later	291	5.3

Table 8. Student Characteristics of High Flyers by the Distribution of the Grade in which Algebra I is Taken.

	Base-line Sample Averages	Grade Taken Relative to Grade-9		
		Before	During	After
	n=5478	n=3714	n=1473	n=291
Soaring High-Flyer	61.2%	68.8%**	44.2%	49.5%†
Percent White	92.0	91.7†	93.3	88.7*
Percent Black	3.5	3.3	3.7	4.5
Percent Hispanic	1.3	1.2	1.0	3.8*
Percent Asian	2.9	3.5**	1.4	2.8
Percent Female	47.8	49.1*	45.4	43.0
Percent FR/L Eligible	29.6	26.5**	35.0	42.6*

Notes: ** indicates that the means are significantly different at the 0.01 level, * at the 0.05 level, and † at the 0.10 level. For the "Grade Taken Relative to Grade-9" panel, both the "before" and "after" means are compared to the means from the "during" group. To avoid confusion, significant differences in comparison to the "during" group are only marked in the before and after columns. In addition, percent FR/L eligible value is based on if the student *ever* met that criterion over the course of the panel.

Table 9. Average School Characteristics by the Distribution of the Grade in which Algebra I is Taken. 2006 and 2011 schools attended.

	Grade Taken Relative to Grade-9		
	Before	During	After
	n=3371	n=1282	n=253
2006 School Attended			
Ave. MAP Math Score	0.249	0.241	0.210
Ave. MAP Com Arts Score	0.212**	0.189	0.129**
VAM Math Effect	0.057	0.053	0.035**
VAM Com Arts Effect	0.041**	0.033	0.019**
Percent Female	49.0%	48.8%	48.9%
Percent FR/L Eligible	30.3%**	34.6%	38.2%*
Percent Minority	12.7%**	10.0%	21.4%**
Number of High Flyers	13.4**	12.5	11.5*
Share of High Flyers	15.9%*	16.6%	16.3%
2011 School Attended			
Ave. MAP Math Score	0.178**	0.144	0.097*
Ave. MAP Com Arts Score	0.171**	0.130	0.060**
Ave. VAM Math Effect	0.006	0.003	0.007
Ave. VAM Com Arts Effect	0.006**	0.001	-0.008*
Percent Female	49.0%**	49.3%	48.2%*
Percent FR/L Eligible	36.0%**	40.0%	45.1%**
Percent Minority	15.7%**	12.8%	27.2%**
Number of High Flyers	33.2	32.7	29.5*
Share of High Flyers	12.9%	12.9%	12.3%†

Notes: ** indicates that the means are significantly different at the 0.01 level, * at the 0.05 level, and † at the 0.10 level. Both the "before" and "after" means are compared to the means from the "during" group. To avoid confusion, significant differences in comparison to the "during" group are only marked in the before and after columns. The initial year VAM estimates are taken from the school attended for the 2007 school year, rather than 2006, as a number of high-flyers attended schools in 2006 for which VAM school-level values could not be estimated. Furthermore, the sample of students in this table is limited to those that attend schools within the *same district* in both 2006 and 2011. Hence, the differences in the average school characteristics between 2006 and 2011 indicate structural school differences by grade configurations *within* districts, i.e. students moving from smaller, neighborhood elementary schools to larger, more diverse middle schools, rather than differences observed when mobile students move from one district to another over time.

Table 10. Average School Characteristics by the Distribution of the Grade in which Algebra I is Taken. 2006 and 2011 schools attended. Minority Students Only.

	Grade Taken Relative to Grade-9		
	Before	During	After
	n=144	n=64	n=20
2006 School Attended			
Ave. MAP Math Score	0.039*	0.183	-0.215**
Ave. MAP Com Arts Score	0.016*	0.138	-0.269**
VAM Math Effect	0.050	0.076	-0.023**
VAM Com Arts Effect	0.039	0.049	-0.006*
Percent Female	49.5%	49.5%	49.8%
Percent FR/L Eligible	48.2%**	37.4%	62.3%**
Percent Minority	43.8%**	23.2%	62.2%**
Number of High Flyers	9.5*	12.5	6.3**
Share of High Flyers	12.3%**	16.0%	9.4%**
2011 School Attended			
Ave. MAP Math Score	-0.033	0.068	-0.059
Ave. MAP Com Arts Score	-0.025	0.055	-0.057
Ave. VAM Math Effect	-0.015	-0.008	-0.016
Ave. VAM Com Arts Effect	-0.011*	0.007	-0.001
Percent Female	50.1%*	49.1%	52.4%*
Percent FR/L Eligible	54.5%**	42.0%	69.3%**
Percent Minority	45.8%**	26.7%	62.2%**
Number of High Flyers	23.0**	36.4	17.4**
Share of High Flyers	9.8%*	12.2%	9.4%†

Notes: ** indicates that the means are significantly different at the 0.01 level, * at the 0.05 level, and † at the 0.10 level. Both the "before" and "after" means are compared to the means from the "during" group. To avoid confusion, significant differences in comparison to the "during" group are only marked in the before and after columns. The initial year VAM estimates are taken from the school attended for the 2007 school year, rather than 2006, as a number of high-flyers attended schools in 2006 for which VAM school-level values could not be estimated. Furthermore, the sample of students in this table is limited to those that attend schools within the *same district* in both 2006 and 2011. Hence, the differences in the average school characteristics between 2006 and 2011 indicate structural school differences by grade configurations *within* districts, i.e. students moving from smaller, neighborhood elementary schools to larger, more diverse middle schools, rather than differences observed when mobile students move from one district to another over time.

Table 11. Effects of Student- and School Level Characteristics on the Grade in which Algebra I is Taken.

	Grade in which Algebra I is Taken					
	Model 1	*Model 2*	*Model 3*	*Model 4*	*Model 5*	*Model 6*
	n=5461	n=5478	n=5461	n=5453	n=5461	n=5453
	R-sq=0.041	R-sq=0.037	R-sq=0.068	R-sq=0.076	R-sq=0.068	R-sq=0.076
Student-Level						
2006/2007 Average Math MAP Score	-0.083**		-0.090**	-0.094**	-0.089**	-0.094**
	(0.019)		(0.019)	(0.019)	(0.019)	(0.019)
2006/2007 Average Com Arts MAP Score	-0.091**		-0.088**	-0.085**	-0.087**	-0.085**
	(0.015)		(0.015)	(0.015)	(0.015)	(0.015)
School-Level						
Composite Average MAP Score		-0.463**	-0.426**	-0.707**	-0.521**	-0.698**
		(0.087)	(0.085)	(0.095)	(0.105)	(0.106)
Number of High Flyers		0.007**	0.007**	0.009**	0.007**	0.009**
		(0.002)	(0.002)	(0.001)	(0.002)	(0.001)
Share of High Flyers		0.006*	0.006*	0.007**	0.007**	0.007**
		(0.003)	(0.003)	(0.002)	(0.003)	(0.002)
Cohort Enrollment/100		-0.092**	-0.088**	-0.102**	--0.090**	-0.102**
		(0.023)	(0.022)	(0.022)	(0.022)	(0.022)
Composite VAM – Bottom 50% of Cohort Students				1.154**		1.173**
				(0.210)		(0.223)
Composite VAM – All Non-Cohort Students					0.547†	-0.077
					(0.328)	(0.346)
Constant	8.542**	8.603**	8.755**	8.172**	8.742**	8.174**
	(0.032)	(0.249)	(0.253)	(0.191)	(0.254)	(0.191)
Student-Level Demographic Controls Included	X		X	X	X	X
School-Level Demographic Controls Included		X	X	X	X	X

Notes: ** represents significance at the 0.01 level, * at the 0.05 level, and † at the 0.10 level.