THE SCHOOL DISTRICT OF PHILADELPHIA

Research Brief:
Board Goals & Guardrails

An Analysis of How Enrollment Demographics, School Climate, and School Staffing are Related to Academic Performance

Michael Frisone, M.S., Statistician; Theodore Wills, Ph.D., Director, Research, Policy and Practice

Summary of Key Findings

- A school's enrollment demographics account for over 50% of the variation in school-level Math and ELA achievement.
- Climate and staffing are significant indicators for predicting achievement in Math and ELA in both elementary and high schools.

Introduction

In December 2020, the Philadelphia Board of Education (BOE) established "Goals & Guardrails"[1] that outline what School District of Philadelphia (SDP) students must know and be able to accomplish, and describe the conditions needed in each school to empower all students to succeed in and beyond the classroom. Each of the Goals and Guardrails identifies a key component of a high-quality educational experience and provides a framework for monitoring that component, including targets to be achieved by August 2026. The Goals focus on academic outcomes, and the first four Goals are each tied directly to a Pennsylvania State Assessment (see Table 1 below).

The District has implemented a structured planning and monitoring process for achieving the Goals & Guardrails, called the Progress Monitoring Cycle.

[1] For more information, please see https://www.philasd.org/schoolboard/goals-and-guardrails/, https://www.philasd.org/goalsandguardrails/, and https://www.philasd.org/era/goals-and-guardrails/

At the beginning of the Progress Monitoring Cycle for each Goal, District personnel conduct a thorough review of relevant, school-level data. These data reviews form the foundation of a structured process through which key barriers and facilitators of success are discussed, root causes of inequity are identified, and system-level actions and commitments are adopted.

Table 1. Philadelphia Board of Education Goals 1-5

Goal 1	65% of 3rd-8th grade students will be proficient on the state ELA assessment by August 2025.
Goal 2	62% of 3rd grade students will be proficient on the state ELA assessment by August 2025.
Goal 3	52% of 3rd-8th grade students will be proficient on the state Math assessment by August 2025.
Goal 4	By August 2025, 52% of high school students at the end of 11th grade will be proficient on all three state high school assessments (Algebra, Literature, and Biology).
Goal 5	80% of Career and Technical Education (CTE) students will pass an industry standards-based competency assessment by the end of 12th grade.

When selecting which data to use for the Progress Monitoring Cycle, the District used prior research that identified relationships between certain data points and school-level outcomes. The selected data points reflect a broad range of specific, measurable characteristics of schools, but they can also be seen as three distinct "sets" of data points reflecting *enrollment demographics* (who are the students in our schools?), *school staffing* (who are the staff working in our schools?), and *school climate* (how safe and welcoming are our schools?).

Box 1: Predictive but Not Deterministic

This brief describes a type of analysis (regression) which is easily and frequently misunderstood. A regression analysis answers the question "to what extent does the value of data point A predict the value of data point Z?" If a predictive relationship is found, however, this does not mean that A must be a *cause* of Z.

In some cases, it is easy to keep this distinction in mind. For example, if it were found that national ice cream sales in a given week predicted national bathing suit sales, it would be easy to discount the story that "therefore, ice cream sales cause bathing suit sales—and if I want to boost bathing suit sales, I should do a better job of selling ice cream!" This can be discounted because it does not appear to be plausible, AND because it is easy to imagine an alternative explanation that *is* plausible (there is a kind of weather [hot] that inspires people to buy both ice cream and bathing suits).

In our work, enrollment demographics often follow this same pattern. We frequently find that different enrollment profiles predict a variety of outcomes, but this is rarely confused for a direct causal story. For example, a common finding is that schools with high enrollments of students classified as economically disadvantaged tend to also be schools that have lower average scores on high-stakes assessments. We understand that being classified as economically disadvantaged is a symptom of life circumstances (just as ice cream sales are a symptom of hot weather). Further, we understand that those same life circumstances (hot weather) are also likely to present barriers to academic attainment (bathing suit sales). Demographic data are often very strong *predictors* of educational outcomes, which makes them both critical and necessary components of our analyses—but this does not imply that they are *deterministic*.

It is much harder to maintain disciplined interpretation when the direct story makes intuitive sense. An example of this might be a finding that students with higher rates of attendance also have higher standardized test scores. In this example, the direct story reads "high attendance causes high test scores—if I want to boost test scores, I should do a better job of promoting attendance." This statement is plausible—after all, a student who attends more school has more opportunity to master the material that eventually appears on the standardized test. In fact, this might even turn out to be the best explanation. However, the *analysis* does not give this explanation any more actual weight than the story about ice cream and bathing suits. It requires discipline to continue past the first plausible explanation to other alternate explanations. In this case, a "hot weather" style account might be "students who are disengaged are BOTH less likely to achieve high attendance AND less likely to perform well on standardized tests." In this framing, such students might be induced to attend more often, but if those inducements do not affect general engagement, then this would not be expected to have an effect on test scores.

In summary, the analyses in this report help us identify data points that are interrelated. This is a valuable tool that helps to identify the signals (bathing suits, ice cream) that warrant closer examination to find root causes (hot weather).

Research Questions

The analyses in this brief focus on three specific sets of data points, reflecting three general components of schools. Those three components (or constructs) are *enrollment demographics* (who are the students in our schools?), *school staffing* (who are the staff working in our schools?), and *school climate* (how safe and welcoming are our schools?). The guiding research questions are:

1. **In schools serving grades K-8:** Which sets of predictor variables [enrollment demographics, school climate, and school staffing], best predict the outcome variables [student ELA and Math performance], and, therefore, provide the best "signals" for identifying groups of schools with common underlying root causes?
2. **In schools serving grades 9-12:** Which predictor variables [percent of students with economic disadvantage status, percent of students with 95% attendance, and teacher years of experience], best predict the outcome variables [student ELA and Math performance], and, therefore, provide the best "signals" for identifying groups of schools with common underlying root causes?

Methods

To answer the research questions, we employed ordinary least square (OLS) regression analyses.[2] Regressions are the process of using one or more variables (known as predictors) to estimate the most likely value of another variable of interest (known as the outcome). Regression analyses are based on correlations. Predictor variables that correlate more strongly with the outcome variable will provide better estimations of expected outcomes. In statistical terms, better predictors will account for more of the variance (or variation) found in the outcome variables.

The analyses summarized in this brief use a type of regression approach called hierarchical setwise regression. Setwise regression is the procedure of systematically adding sets of variables into a model in order to determine how much of the overall variance in the outcome variable is accounted for by each *set* of input variables.[3] Each set of variables (known as variates) are linked by a common conceptual construct. For example, student suspensions, serious incident occurrences, and student

[2] The boundaries of our outcome variables, which represent the percent of students meeting a test standard, cannot be lower than 0% or larger than 100%. OLS regression does not account for these boundaries in the data, allowing for predicted values greater than 100% and lower than 0%. To assess whether this technical limitation resulted in substantive distortion of our analyses we also constructed beta regressions, an alternative technique which does apply boundaries to predicted values (results of beta regressions may be found in appendix B). In all cases, findings were consistent between OLS and beta regressions, confirming that in these specific analyses the OLS regressions were not compromised by the bounding limitation. Because OLS regression has other advantages over beta regression, particularly the potential for effect size analysis, the OLS regressions were retained as the main analytic technique.

[3] Unlike stepwise regression, setwise regression is motivated by testing theory (not optimizing total variance explained).

attendance are all included in our *school climate* construct.[4] The setwise regression technique allows us to evaluate the unique predictive contribution of each conceptual construct (regardless of the number of individual variables contained in that construct).

In total, 166 schools serving grades K-8, and 52 schools serving grades 9-12 in the School District of Philadelphia were included in the analyses.[5] OLS regression typically requires a sample size of ten data units (in this case schools) for each predictor variable in the model. For the K-8 analyses, the sample size of 166 allowed as many as 16 predictor variables, which allowed us to proceed with all of our available predictors. However, a sample of only 52 high schools restricted us to a maximum of five predictor variables. For this reason, the decision was made to include only one variable from each of our three conceptual constructs, as two per construct, for a total of six, would exceed the maximum allotment.

All predictor and outcome variables used for these analyses are from the winter of the 2020-2021 school year, except for third through eighth grade students' Math performance, which was collected in the spring of the 2020-2021 school year (Tables 1 and 2).[6] To account for ELA and Math performance, we used data from the two academic screeners used by the District during this timeframe, called aimswebPlus and Star.[7] These two screeners have differences, but both assess whether a student is at or above grade level for both ELA and Math, separately. The outcome variable was the percent of students scoring at or above grade level in either screener; in some models for ELA, and in others for Math.

---

[4] In general usage, the term "school climate" refers to the extent to which a school's culture and environment are welcoming, safe, and positive. In this analysis, we expect to see this reflected in measures of student behavior. In other contexts, this same concept is sometimes measured via self-report (e.g., in SDP's annual District Wide Survey: https://www.philasd.org/research/wp-content/uploads/sites/90/2021/10/DWS-Technical-Report-2020-21-October-2021.pdf

[5] The analyses included a total of 213 unique schools (not 166+52=218), because some schools serve one or more grades in both the K-12 and 9-12 bands. Additionally, three schools were excluded due to insufficient data, and an additional school (Philadelphia Virtual Academy) was removed from analyses because it has a unique operational model that is strongly atypical of District schools. Further, exploratory analyses which included Philadelphia Virtual Academy confirmed that it distorted results, and produced regression models that did not adhere to the normality of residual assumption.

[6] This difference was simply due to the timing and sequence in which specific Goals were the subject of the review process.

[7] For more information about the the aimswebPlus and Star screeners see: https://www.philasd.org/research/2021/09/03/assessing-student-performance-before-and-during-virtual-learning-a-cohort-comparison-of-student-performance/

Table 1. Summary of outcome variables

Goal	Indicator	Description	School Year Collected
1 and 2	Star/ aimsWeb	Percent of Students in School Meeting Grade-Specific ELA Test Standard	2020-21
3	Star/ aimsWeb	Percent of Students in School Meeting Grade-Specific Math Test Standard	2020-21
4	Star	Percent of Students in School Meeting Grade Specific Standards for Both ELA and Math Tests	2020-21

Table 2. Summary of predictor variables

Variate	Indicator	Description	School Year Collected
Demographics	% Special Education	Percent of students in school with (non-gifted) Individual Education Plans (IEPs)	2020-21
	% Economically Disadvantaged	Percent of students in school with Economic Disadvantaged status[8]	2020-21
	% English Learner	Percent of students in school with English Learner status	2020-21
Climate	No Out of School Suspensions (OOS)	School either had no out-of-school suspension, or at least one (binary)	2020-21
	Serious Incidents	Number of Serious Incidents at school, per 100 students	2020-21
	% 95% Attendance	Percent of students attending 95% or more of enrolled days	2020-21
Staffing	Teacher Retention	Percent of teachers returning from previous year	2020-21
	Teacher Years of Experience (YOE)	Average years of experience for teachers in school	2020-21
	Principal Tenure	Number of years that current principal has led school	2019-20

Note: *% Economically Disadvantaged, % 95% Attendance*, and *Teacher YOE* variables were selected to be used for High School analyses.

[8] The School District of Philadelphia classifies students as economically disadvantaged if their household receives certain types of government assistance (such as SNAP, TANF, and Medicaid).

To account for enrollment demographics in the K-8 analyses, three variables were included: percent of students in special education programs, percent of students classified as economically disadvantaged, and percent of English Learners (ELs). Only one of these variables was selected for the high school analyses: percent of students classified as economically disadvantaged.[9] To represent school climate in K-8 schools, three variables were included: an indicator for whether a school had issued zero out-of-school suspensions, the rate of serious incidents in each school, and the percent of students with 95% or higher attendance. The school climate variable used for the high school analyses was the percent of students with 95% or higher attendance. Finally, to represent school staffing in the K-8 analyses, the three variables included were: principal years of tenure at the school (as of the 2019-2020 school year), teachers' average years of experience, and percent of teachers retained from the previous year.[10] For the high school analyses, teachers' average years of experience was used.

In these analyses, our first step was to look at the relationship between school-level enrollment demographics and the outcome variables. As the result of hundreds of years of systemic racism and structural inequities in the United States of America, we know that schools vary in ways that are directly related to the demographics of the enrolled student population, so we always look at the difference in outcomes for different student groups. In this analysis we also seek to identify the school characteristics above and beyond student demographic characteristics that provide information about student performance. After this step, we added the climate and school staffing constructs to explore how these predict Math and ELA achievement above and beyond enrollment demographics.

[9] For high school analyses, the smaller sample size (fewer schools) required the use of fewer predictor variables. The variable that was found to be the strongest predictor among each respective variate was retained.

Findings

Box 2. Statistical Significance and Effect Sizes in Regression Analysis[11]

Statistical significance: This component of the analyses answers the question "there appears to be a relationship between these two variables—how confident are we that this apparent relationship is not just due to randomness in our data?" The analysis provides a p-value, and the smaller the p-value the less likely it is that the results are due to randomness.

Effect Size: A second component of regression analysis (and other statistical procedures), effect sizes are a method of presenting the practical importance of findings. They answer the question: "Assuming this effect is legitimate, to what extent would it be noticeable and impactful in the real world? Or would it be too small to actually matter in real-world contexts?" Generally, effect sizes are characterized as small, medium or large according to the following:

- **Small**: At least 0.2 (an effect size below this value would be of questionable practical importance)
- **Medium**: Greater than 0.5
- **Large**: Greater than 0.8

RQ1: In schools serving grades K-8: Which sets of predictor variables [enrollment demographics, school climate and school staffing], best predict the outcome variables [student ELA and Math performance], and, therefore, provide the best "signals" for identifying groups of schools with common underlying root causes?

In a regression analysis, multiple predictor variables may overlap with each other, such that their combined predictive power is less than the sum of their individual predictive power. In such cases, the analysis can isolate the unique contribution of a predictor variable by evaluating the increase to the explained variance when that variable is added to an existing model. In this analysis, we consider the combined predictive value of the model, but also unique contributions of school climate and school staffing when they are added to a model that already includes enrollment demographics (a summary of results may be found below in Table 3, and detailed regression tables may be found in the appendix).

Overall, enrollment demographics, school climate, and school staffing account for about 75% of the total variance in ELA performance and 79% of Math performance, which translates to effects sizes (ES) of 2.99 and 3.77 respectively (See Box 2). In this case, these are very large effect sizes. They

[11] For more information about statistical significance and effect sizes, see: Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

contribute a great deal of the variance when explaining ELA and Math performance. In other words, if we know the enrollment demographics, school climate, and staffing at a particular school, we can reasonably estimate what student ELA and Math performance will be at that school.

Table 3. Percent variance explained by variate sets for K-8 schools (n = 166)

Outcome Variable	Variate Set(s)	% Variance Explained by Variate Set(s)
ELA	Demographics	50.1***
	Demographics + Climate	67.8***
	Demographics + Staffing	65.0***
	All Three Variate Sets	74.9***
Math	Demographics	53.5***
	Demographics + Climate	77.8***
	Demographics + Staffing	62.7***
	All Three Variate Sets	79.1***

Note: *p<.05; **p<.01; ***p<.001

How to read this table: The column *% Variance Explained by Variate Set* refers to total amount of variance in the outcome variable that is explained by the variable set(s) found in the *Variate(s)* column. If the Variate Set(s) column contains only "Demographics," then this column presents the percent of variance in Math or ELA scores that is explained by the Demographics variate set (% Special Education, % English Learner, and % Economically Disadvantaged) by itself. If the column reads "Demographics + Climate," then this column displays the percent of variance in the outcome variable that is explained by the combined contributions of the variables in both variate sets. "All three variate sets" means the full, combined contributions of the Demographics, Climate, and Staffing sets.

By itself, the demographics set accounts for 50.1% (ES= 2.00) of the variance in student ELA performance and about 53.5% (ES=2.55) of the total variance in Math performance. School staffing accounts for 14.9% (ES=.60) of the variance in ELA performance and 9.3% (ES=.44) of the variance in Math Performance when accounting for student demographics. Further, staffing accounts for 7.1% (ES=.28) of the overall ELA variance and 1.3% (ES=.06) of Math variance when holding both demographic and climate indicators constant.

Similarly, climate indicators account for 17.7% (ES=.71) of the variance in ELA performance and 24.3% (ES=1.16) of the variance in Math performance when accounting for demographic indicators. Further, climate indicators accounted for 9.9% (ES=.36) of the total ELA variance and 16.3% (ES=.78) of Math variance above and beyond just considering demographic and climate indicators. All findings were statistically significant. Effect sizes for school climate ranged from medium to very large and effects sizes related to school staffing were small to medium (Table 3).

Taken together, school climate and staffing indicators account for 24.8% of the variance in ELA performance and 25.6% of the variance in Math performance when accounting for demographic indicators (which explain 50.1% of the variance in ELA performance and 53.5% of the variation in Math performance).

RQ2: In schools serving grades 9-12: Which predictor variables [percent of students with economic disadvantage status, percent of students with 95% attendance, and teacher years of experience], best predict the outcome variables [student ELA and Math performance], and, therefore, provide the best "signals" for identifying groups of schools with common underlying root causes?

When accounting for the single indicator that was retained from each construct for the high school analyses (percent of students who are economically disadvantaged, percent of students with 95% or more attendance, and teacher average years of experience[12]), about 88% of the total variance in ELA performance and 90% of Math performance are explained (Table 4). This translates to very large effects sizes of 7.18 and 8.64 respectively.

We also evaluated the unique, added, predictive value of school staffing (teacher years of experience) and school climate (95% attendance rate) when layered onto a model that already includes enrollment demographics (percent of students with economic disadvantage status). To begin, the percent of students who are economically disadvantaged (the demographic variable used in the high school analyses), by itself, accounts for 78.6% (ES=6.43) of the total variance in ELA performance and 57.2% (ES=5.52) of the total variance in Math performance. These are very large effect sizes. After accounting for economic disadvantaged status, student attendance (climate variable) accounts for an additional 7.2% (85.8-78.6; ES=.59) of the variance in ELA performance, and 29.4% (86.6-57.2; ES=2.85) of the variance in Math performance. Further, when controlling for both economic disadvantage status and teacher experience (staffing variable), student attendance uniquely accounts for 4.9% (ES=.40) of the variance in ELA performance and 23.1% (ES= 2.23) of the variance in Math performance. Teacher experience accounts for 4.3% (ES=.35) of the ELA variance and 9.3% (ES=.89) of the Math variance when holding economic disadvantage status constant. Finally, when holding both economic disadvantage status and student attendance constant, teacher experience accounts for 2.0% (ES=.16) and 2.8% (ES=.27) of ELA and Math variance respectively. These findings are all statistically significant. The climate related effect sizes range from medium to very large and the staffing related effect sizes are small to medium.

[12] Recall there was a relatively small sample of schools serving grades 9-12. As a result, only one variable for each of the three constructs was retained for analyses of these schools.

Table 4. Percent variance explained by variate sets for high schools (n = 166)

Outcome Variable	Variate(s)	% Variance Explained by Variate(s)
ELA	% Economically Disadvantaged	78.6***
	% Economically Disadvantaged + 95% Attendance	85.8***
	% Economically Disadvantaged + Teacher YOE	82.9***
	All Three Variates	87.8***
Math	% Economically Disadvantaged	57.2***
	% Economically Disadvantaged+ 95% Attendance	86.6***
	% Economically Disadvantaged + Teacher YOE	66.5***
	All Three Variates	89.5***

Note: *p<.05; **p<.01; ***p<.001

How to read this table: The column *% Variance Explained by Variate(s)* refers to total amount of variance in the outcome variable that is explained by the variable(s) found in the *Variate(s)* column. If the Variate(s) column contains only "% Economically Disadvantaged", then this column presents the percent of variance in Math or ELA scores that is explained by % Economically Disadvantaged by itself. If the column reads "% Economically Disadvantaged + 95% Attendance," then this column displays the percent of variance in the outcome variable that is explained by the combined contributions of the two variates. "All three variates" means the full, combined contributions of "% Economically Disadvantaged, 95% Attendance, and Teacher YOE.

Conclusions and Next Steps

These results show that measures of enrollment demographics, school climate, and school staffing are strongly related to student ELA and Math performance. Further, climate and staffing measures are related to student ELA and Math performance even after holding student demographics constant.

Limitations and Considerations

There are other conclusions that might be tempting to draw if we do not pay close attention to the limitations of these models. For example, it might be tempting to conclude that if school climate is important, then targeting student attendance (one of the climate variables) would be a straight-line solution for improving Math performance. However, within each construct the individual variables are better viewed as a complex, overlapping, interweaving, and at times redundant set of indicators which influence student performance. There are likely many, *many* other variables or data points, both those directly related to climate and those highly correlated with climate, that would also account for most of the *same* variance that student attendance is representing in these models. So, for example, though 23% of high school Math performance is predicted by the percent of students with 95% attendance, this does not mean that attendance *determines* or *controls* student performance in

Math (though it is likely that it *contributes*). In this case, our analysis cannot establish whether attendance is the direct cause ("hot weather") of Math performance, or if they are both co-symptoms ("ice cream" and "bathing suits") of a mutual, underlying cause (see Box 1).

Further, while a school-level analysis is useful, it does bring with it challenges related to statistical power. In this case, the unit of analysis is an entire school, and there were 166 schools that serve students in grades K-8, and 52 schools that serve grades 9-12 in our analytic sample. These sample sizes limit the number of predictor variables that can be used to represent the three larger constructs. Future analyses will incorporate student-level data, which will significantly increase statistical power, and allow exploration of a wider range of variables. Further, by doing these analyses at the student level, we gain access to additional nuance in our data. When we aggregate to the school level, we must ignore student-to-student variance. This approach may also clarify why the effect sizes we obtained are extraordinarily high. By excluding student-to-student variance, we may be inflating our effect size estimates because important sources of variability are "hidden" from a school-level analysis.

Appendix A: Regression Model Details

Presented below are regression tables for each model broken down by variable. The first column in each table (Model 1) is restricted to demographic information only. The second column (Model 2) includes both demographic information and climate factors. The third column (Model 3) includes demographic indicators and staffing indicators. The fourth column (Model 4) includes the variables from all three major areas.

Each cell includes three pieces of information. The first number is the regression weight. This represents the linear pattern that we see between each indicator and the outcome variable of interest. For example, in Table A1 Model 1, the percent of students in a special education program has a regression weight of -0.26. This means that for every 1% increase in a school's percent of students in a special education program, on average, we expect to see about a quarter of a percent *decrease* in the number of students in the school who scored at or above benchmark in ELA (if this value had been positive, it would have predicted an *increase* in the number of students at or above benchmark). Please note, these regression weights are correlations. Correlation does not necessarily mean there is a causal connection.

The second number in each cell (in parentheses) is the standard error for the regression weight; this is a representation of the variability around regression weight. In other words, the standard error informs us of how precise our regression weights are at summarizing our sample. For example, in Table A1 Model 1, the percent of students in a special education program has a standard error of .11. This means that the average dispersion around the regression rate is about a tenth of a percent.

Finally, the asterisks represent statistical significance. Statistical significance is a representation of the likelihood we could have found the results we have, assuming these indicators are actually *not* related. For example, in Table A1, Model 1, assuming the percent of students in a special education program and percent of students at or above benchmark in ELA are independent, there is a less than 5% probability that we could have gotten the results we got (or more extreme results). The greater the number of asterisks, the lower the probability. Generally, we do not consider findings to be reliable if they are not statistically significant.

Table A1: K-8 school ELA performance predicted by Demographics, Climate, and Staffing indicators

	Model 1: Demographics *B*(SE)	**Model 2: Demographics + Climate *B*(SE)**	**Model 3: Demographics + Staffing *B*(SE)**	**Model 4: Demographics + Climate + Staffing *B*(SE)**
Intercept	1.16 (0.07)***	-0.70(20.29)	0.65(0.09)***	0.74(18.23)
% Special Education	-0.26(0.11)*	-0.11(0.09)	-0.27(0.09)**	-0.12(0.08)
% Economically Disadvantaged	-0.82(0.08)***	-0.41(0.08)***	-0.70(0.07)***	-0.41(0.07)***
% English Learner	0.09(0.08)	-0.05(0.07)	-0.01(0.07)	-0.08(0.06)
School has No OOS	-	1.19(20.30)	-	-0.43(18.23)
Serious Incident	-	0.01(0.04)	-	0.001(0.03)
% 95% Attend	-	0.49(0.05)***	-	0.42(0.05)***
Teacher Retention	-	-	0.26(0.07)***	0.08(0.07)
Teacher YOE	-	-	0.02(0.003)***	0.02(0.003)***
Principal Tenure	-	-	-0.001(0.002)	-0.004(0.002)*

$*p<.05;**p<.01;***p<.001$

Table A2: K-8 school Math performance predicted by Demographics, Climate, and Staffing indicators

	Model 1: Demographics *B*(SE)	**Model 2: Demographics + Climate *B*(SE)**	**Model 3: Demographics + Staffing *B*(SE)**	**Model 4: Demographics + Climate + Staffing *B*(SE)**
Intercept	1.28(0.08)***	22.84(19.30)	0.78(0.11)***	21.53(19.08)
% Special Education	-0.31(0.12)**	-0.10(0.08)	-0.34(0.11)**	-0.12(0.08)
% Economically Disadvantaged	-0.96(0.08)***	-0.41(0.07)***	-0.83(0.08)***	-0.41(0.07)***
% English Learner	0.24(0.09)**	-0.04(0.06)	0.13(0.08)	-0.02(0.06)
School has No OOS	-	-22.46(19.31)	-	-21.26(19.08)
Serious Incident	-	0.02(0.03)	-	0.02(0.03)
% 95% Attend	-	0.66(0.05)***	-	0.61(0.06)***
Teacher Retention	-	-	0.37(0.08)***	0.10(0.07)
Teacher YOE	-	-	0.01(0.004)	0.01(0.003)
Principal Tenure	-	-	0.001(0.003)	-0.002(0.002)

$*p<.05;**p<.01;***p<.001$

Table A3: 9-12 school ELA performance predicted by Demographics, Climate, and Staffing indicators

	Model 1: Demographics *B*(SE)	**Model 2: Demographics + Climate *B*(SE)**	**Model 3: Demographics + Staffing *B*(SE)**	**Model 4: Demographics + Climate + Staffing *B*(SE)**
Intercept	1.43(0.09)***	0.90(0.13)***	1.13(0.12)***	0.76(0.13)
% Economically Disadvantaged	-1.31(0.10)***	-0.96(0.11)***	-1.22(0.09)***	-0.95(0.10)***
% 95% Attend	-	0.34(0.07)***	-	0.29(0.07)***
Teacher YOE	-	-	0.02(0.006)**	0.01(0.005)**

p<.05;**p<.01;p<.001**

Table A4: 9-12 school Math performance predicted by Demographics, Climate, and Staffing indicators

	Model 1: Demographics *B*(SE)	**Model 2: Demographics + Climate *B*(SE)**	**Model 3: Demographics + Staffing *B*(SE)**	**Model 4: Demographics + Climate + Staffing *B*(SE)**
Intercept	1.70(0.16)***	0.37(0.15)*	1.16(0.20)***	0.17(0.15)
% Economically Disadvantaged	-1.37(0.17)***	-0.51(0.12)***	-1.21(0.16)***	-0.49(0.11)***
% 95% Attend	-	0.83(0.08)***	-	0.77(0.07)***
Teacher YOE	-	-	0.04(0.01)***	0.02(0.006)***

p<.05;**p<.01;p<.001**

Appendix B: Beta Regression Model Details

Presented below are beta regression tables for each model broken down by variable. The purpose of this section is to demonstrate that models which assume boundary conditions resulted in similar results to those presented in the main body of this report. The first column in each table (Model 1) is restricted to demographic information only. The second column (Model 2) includes both student demographic information and climate factors. The third column (Model 3) includes demographic indicators and staffing indicators. The fourth column (Model 4) includes the variables from all three major areas.

Each cell includes three pieces of information. The first number is the regression weight. Unlike the tables in Appendix A, these tables present log odds because beta regression is type of logistic regression. By taking the exponent of each of these log odds, we are able to obtain the odds ratio for that covariate. Interpreting these ratios in accessible terms is challenging. Consider a hypothetical school in which 20% of students are meeting the testing standard in Math which is equivalent to an odds ratio of 1:4 (one "yes" per four "no's", or one of every five). Continuing the example, in Table B2, Model 4, the percent of students with 95% attendance has a log odds of .028. This number can be translated into a conventional odds ratio, through the formula $e^{.028}$, yielding a value of 1.03. This, in turn, means that a 1% increase in a school's 95% attendance rate is associated with a 3% increase in the *odds ratio* that expresses the percent of students at that school meeting the Math testing standard. In this case, the updated odds would be 1.03:4, which translates to 1.03 of every 5.03, and the updated percent of students meeting the testing threshold would move from 20.0% to 20.5%.

The second number in each cell (in parentheses) is the standard error for the regression coefficient; this is a representation of the variability around the logit. In other words, the standard error informs us of how precise our coefficients are as summaries of our sample. For example, in Table B2, Model 4, the coefficient associated with the percent of students with 95% attendance has a standard error of .27. Standard errors in these models are on a log odds scale.

Finally, the asterisks represent statistical significance. Statistical significance is a representation of the likelihood of obtaining the results we have, assuming these indicators are actually *not* related. For example, in Table B2, Model 2, assuming the percent of students in a special education program and percent of students at or above benchmark in ELA are independent, there is a less than 5% probability (one asterisk) that we could have gotten the results we got (or more extreme results). The greater the number of asterisks, the lower the probability that the results could have occurred by chance. Generally, we do not consider findings to be reliable if they are not statistically significant.

Table B1: K-8 school ELA performance predicted by Demographics, Climate, and Staffing indicators

	Model 1: Demographics ***B*(SE)**	**Model 2: Demographics + Climate** ***B*(SE)**	**Model 3: Demographics + Staffing** ***B*(SE)**	**Model 4: Demographics + Climate + Staffing** ***B*(SE)**
Intercept	3.13 (0.355)***	-8.86(104.23)	0.42(0.46)	-1.70(92.908)
% Special Education	-0.01(0.005)**	-0.008(0.005)	-01(0.005)**	-0.008(0.004)
% Economically Disadvantaged	-0.04(0.004)***	-0.02(0.004)***	-0.03(0.003)***	-0.02(0.003)***
% English Learner	0.004(0.004)	-0.003(0.003)	-0.001(0.003)	-0.004(0.003)
School has No OOS	-	0.09(1.043)	-	0.01(0.929)
Serious Incident	-	0.001(0.002)	-	0.0002(0.002)
% 95% Attendance	-	0.02(0.003)***	-	0.02(0.003)***
Teacher Retention	-	-	0.01(0.004)***	0.01(0.003)
Teacher YOE	-	-	0.0008(0.00)***	0.001(0.000)***
Principal Tenure	-	-	-0.007(0.01)	-0.02(0.009)*

p<.05;**p<.01;p<.001**

Table B2: K-8 school Math performance predicted by Demographics, Climate, and Staffing indicators

	Model 1: Demographics ***B*(SE)**	**Model 2: Demographics + Climate** ***B*(SE)**	**Model 3: Demographics + Staffing** ***B*(SE)**	**Model 4: Demographics + Climate + Staffing** ***B*(SE)**
Intercept	4.01(0.415)***	118.69(93.77)	1.48(0.555)**	107.20(92.49)
% Special Education	-0.02(0.006)**	-0.01(0.005)*	-0.02(0.001)***	-0.01(0.005)*
% Economically Disadvantaged	-0.05(0.005)***	-0.02(0.004)***	-0.04(0.004)***	-0.02 (0.004)***
% English Learner	0.01(0.004)*	0.0004(0.004)	0.004(0.004)	-0.0004(0.003)
School has No OOS	-	-1.19(0.938)	-	-1.08(0.925)
Serious Incident	-	0.001(0.002)	-	0.001(0.002)
% 95% Attendance	-	.03(0.002)***	-	.03(0.003)***
Teacher Retention	-	-	0.02(0.004)***	0.006(0.004)
Teacher YOE	-	-	0.0003(0.000)	0.0003(0.000)
Principal Tenure	-	-	0.004(0.01)	-0.001(0.01)

p<.05;**p<.01;p<.001**

Table B3: 9-12 school ELA performance predicted by Demographics, Climate, and Staffing indicators

	Model 1: Demographics ***B*(SE)**	**Model 2: Demographics + Climate** ***B*(SE)**	**Model 3: Demographics + Staffing** ***B*(SE)**	**Model 4: Demographics + Climate + Staffing** ***B*(SE)**
Intercept	5.15(0.582)***	1.91(0.890)*	3.09(0.721)***	0.55(0.835)
% Economically Disadvantaged	-0.07(0.006)***	-0.05(0.007)***	-0.06(0.006)***	-0.05(0.006)***
% 95% Attendance	-	0.02(0.005)***	-	0.02(0.004)***
Teacher YOE	-	-	0.001(0.000)***	0.001(0.000)***

p<.05;**p<.01;p<.001**

Table B4: 9-12 school Math Performance Predicted by Demographics, Climate, and Staffing Indicators

	Model 1: School Enrollment ***B*(SE)**	**Model 2: Demographics + Climate** ***B*(SE)**	**Model 3: Demographics + Staffing** ***B*(SE)**	**Model 4: Demographics + Climate + Staffing** ***B*(SE)**
Intercept	6.02(0.816)***	0.12(0.908)	3.90(0.982)***	-1.04(0.869)
% Economically Disadvantaged	-0.07(0.009)***	-0.03(0.008)***	-0.06(0.008)***	-0.03(0.007)***
% 95% Attendance	-	0.04(0.004)***	-	0.04(0.004)***
Teacher YOE	-	-	0.001(0.000)***	0.001(0.000)***

p<.05;**p<.01;p<.001**

Appendix C: Selecting Input Variables for Schools Serving Grades 9-12.

As noted previously, the relatively small sample of schools serving grades 9-12 constrained the number of input variables that should be retained. For each construct (demographics, climate, and staffing) the individual variables were screened to identify which was the strongest predictor of the outcome variable, and should therefore be retained (Tables C1 and C2).

Table C1: 9-12 school ELA performance predicted by Demographics, Climate, and Staffing indicators

Variate Groups	Variables	*B*	SE
	Intercept	0.988***	0.176
Demographics	% Special Education	-0.290*	0.109
	% Economically Disadvantaged	**-0.965*** **	**0.102**
	% English Learner	-0.150	0.113
Climate[1]	**% 95% Attendance**	**0.097**	**0.094**
Staffing	Teacher Retention	-0.041	0.146
	Teacher YOE	**0.017** **	**0.006**
	Principal Tenure	-0.002	0.004

[1]Pecent of Students with 0 out-of-school suspensions and serious incidents rates were not included for 9-12 grade students.

p<.05;**p<.01;p<.001**

Table C2: 9-12 school Math performance predicted by Demographics, Climate, and Staffing indicators

Variate Groups	Variables	*B*	SE
	Intercept	0.315	0.193
Demographics	% Special Education	-0.391**	0.119
	% Economically Disadvantaged	**-0.516*** **	**0.112**
	% English Learner	-0.051	0.124
Climate[1]	**% 95% Attendance**	**0.513*** **	**0.103**
Staffing	Teacher Retention	-0.186	0.160
	Teacher YOE	**0.019** **	**0.006**
	Principal Tenure	-0.004	0.005

[1]Pecent of Students with 0 out-of-school suspensions and serious incidents rates were not included for 9-12 grade students.

p<.05;**p<.01;p<.001**