Using Data from Schools and Child Welfare Agencies to Predict Near-Term Academic Risks

See https://go.usa.gov/xwGSq for the full report.

Appendix A. Methods

The study team collected and linked five academic years of student-level administrative data from Pittsburgh Public Schools (PPS), Propel Schools, and the Allegheny County Department of Human Services (DHS). The sample included the full population of students enrolled in each local education agency in 2015/16 or 2016/17, and each entity provided any data available on those students for 2012/13–2016/17. The descriptive analyses used data from the two most recent years—considered the "outcome years" for which academic problems are predicted—and the predictive analyses included data from the full five-year period.

Data acquisition

PPS and Propel Schools generated lists of unique state identification numbers (PASecureID) associated with all students enrolled in each local education agency during 2015/16 or 2016/17, which defined the sample. PPS and Propel Schools provided the lists to the Allegheny County DHS, which compiled files with historical data on each student in the sample. PPS, Propel Schools, and the Allegheny County DHS then provided the study team with the data associated with each student for 2012/13–2016/17, identified by PASecureID. DHS data included the entire five-year period 2012/13–2016/17 for each student, regardless of dates of enrollment in PPS or Propel Schools. The data did not include student names, birthdays, addresses, or social security numbers, but the study team took steps to protect the data given that they included PASecureIDs.

Data elements

PPS and Propel Schools provided data for each student on the academic problems examined in the study—absences, suspensions, course performance, and state test performance—as well as demographic characteristics and indicators of eligibility for school services. The agencies provided some data elements on an annual timescale and some on more granular levels—semesters, quarters, or event dates—in separate files. Table A1 shows the types of data elements provided and their timescales.

Table A1. Data elements provided by Pittsburgh Public Schools and Propel Schools, 2012/13–2016/17

	Timescale of data	
Type of data element	Pittsburgh Public Schools	Propel Schools
Demographics (race/ethnicity, gender)	Annual	Annual
Economic disadvantage[a]	Annual	Annual
Grade level	Semester	Annual
English learner status	Event dates	Annual
Special education status[b]	Event dates	Annual
Gifted status	Annual	Annual
Type of disability	Annual	Annual
Type of absence	Event dates	Annual
Behavior incidents and reasons for suspension	Event dates	Event dates
Course grades and cumulative grade point average	Quarter	Quarter
State test score[c]	Annual	Annual
School enrollment and withdrawal events and reasons	Event dates	Event dates

a. Based on eligibility for the national school lunch program.
b. Whether student has an Individualized Education Program.
c. The Pennsylvania System of School Assessment for elementary and middle school students and Keystone exams for high school students.
Source: Authors' compilation.

The Allegheny County DHS provided data on its services, justice system involvement, and public benefits receipt for students in the sample over the five-year study period.

Data preparation

The study team assessed the completeness and quality of the data and then used students' unique state identification numbers (PASecureID) to link school and Allegheny County DHS data. The linked data were used to prepare outcome and predictor data files for the descriptive and predictive analyses.

Outcome data. The study team examined five types of academic outcomes separately for each local education agency. The timescale over which each outcome was defined and analyzed in the descriptive and predictive analyses differed for some outcomes (see box table 2 in table 2 in the main report). For example, *student-quarter* indicates that the analytic file has one observation for each quarter for each student (minus missing data), whereas *student-course* indicates that the file has one outcome for every course taken by each student. Chronic absenteeism (in PPS) and suspensions are calculated by aggregating to a particular time period from the raw (event) data, based on the description in box table 2 in table 2 in the main report. The outcome period was the two years during academic terms (2015/16 and 2016/17) in which a student was enrolled at least half the time. Each outcome is binary, taking a value of 1 if the outcome occurred for the student in the given time period, and 0 otherwise.

Predictor data. For the descriptive analyses the study team defined the predictor period (the time period over which predictors are measured) for each observation to be the period of time immediately preceding (and not overlapping with) the outcome period. The study team created data files of predictor variables aggregated to appropriate time periods for each outcome. For chronic absenteeism, suspensions, course failure, and low grade point average (GPA)—which capture performance over a fixed academic period—the predictor period is the academic period of the same length as the outcome period that immediately precedes it. In most cases, this is the preceding academic term (quarter in PPS and trimester in Propel Schools); the exception is chronic absenteeism in Propel Schools, for which the predictor period is the preceding academic year because data were not available on a shorter time scale. This approach makes it possible to examine the relationships between predictors and outcomes in adjacent periods of equal length.

The outcomes for state tests are different because the test is taken on a single date. PPS and Propel Schools provided data on state test performance for each test taken by each student in each year, including the dates or months when the tests were administered. The study team used the months of test administration to identify the earliest possible date that each test could have been taken. Then, the study team used the test dates or earliest possible dates to define a two-month window preceding each test and created data files of predictor variables for these two-month periods (or most recently completed academic term for term-specific predictors). A two-month time window aligns closely with the nine-week academic period that is used for other outcomes in PPS.

For examining absences and suspensions as predictors, the analysis excluded observations for students who were not enrolled for at least 50 percent of the school days in the predictor period (in addition to the restriction for the outcome period). This is because these predictors are counts of events or percentages of possible days on which events occurred, and they are highly related to the number of days enrolled.

In many cases the level of observation of the predictor is not directly compatible with that of the outcome. For example, suspensions are defined at the term (quarter or trimester) level, but Allegheny County DHS predictors are measured monthly or as date-specific events. In these cases, the study team aggregated the predictors to the appropriate level for each outcome using sensible rules. Monthly flags and events (such as for receiving DHS services) were recalculated at the term level, indicating whether the service or event occurred during any month that overlapped that particular term. The aggregation approach was defined for each predictor.

Analytic sample

After the data preparation stage, the study team created separate analytic sample files for each of 12 analyses (six outcomes for each local education agency; see table A2). These samples consisted of all observed (nonmissing) outcomes during the two-year outcome period (the 2015/16 and 2016/17 school years) that occurred during academic terms in which the student was enrolled for at least 50 percent of possible days. These samples were used for both the bivariate regressions that addressed research question 1 and the predictive models that addressed research question 2. The final sample sizes for each analysis are in table A2, for both the number of unique students and the total number of observations. The number of observations varied for each outcome, based on the time scale of the outcome (for example, annual or quarterly) in addition to data availability.

Table A2. Sample size in Pittsburgh Public Schools and Propel Schools, 2015/16 and 2016/17 (number of observations, except as indicated)

	Pittsburgh Public Schools				Propel Schools			
Type of observation	Elementary school	Middle school	High school	Total (including missing grades)	Elementary school	Middle school	High school	Total (including missing grades)
Descriptive analysis sample (2015/16 and 2016/17)								
Absences	87,957	39,398	49,595	180,529	3,394	1,325	662	6,763
Suspensions	87,957	39,398	49,595	180,529	11,715	5,099	2,748	20,346
Course failures	223,147	155,453	212,139	591,658	40,436	20,266	10,374	72,107
Low grade point average	na	na	47,812	50,967	na	na	2,711	2,972
Score below basic level on state tests[a]	23,230	21,202	16,949	61,381	4,738	4,062	629	9,429
Number of unique students	16,307	8,796	9,348	28,719	3,091	1,527	741	4,614
Predictive model training data (2015/16)								
Absences	44,631	19,971	24,907	92,687	1,684	676	339	3,446
Suspensions	44,631	19,971	24,907	92,687	5,662	2,521	1,380	10,314
Course failures	113,298	78,863	108,410	305,231	19,742	10,045	5,253	36,753
Low grade point average	na	na	24,034	24,465	na	na	1,359	1,410
Score below basic level on state tests[a]	11,685	10,852	8,707	31,244	2,276	1,973	305	4,554
Number of unique students	13,445	6,763	7,064	24,392	2,308	1,082	534	3,562
Predictive model test data (2016/17)								
Absences	43,326	19,427	24,688	90,752	1,710	649	323	3,619
Suspensions	43,326	19,427	24,688	90,752	6,053	2,578	1,368	10,891
Course failures	109,849	76,590	103,729	295,415	20,694	10,221	5,121	38,256
Low grade point average	na	na	23,778	24,324	na	na	1,352	1,428
Score below basic level on state tests[a]	11,545	10,350	8,242	30,137	2,462	2,089	324	4,875
Number of unique students	13,086	6,581	7,081	23,988	2,509	1,106	543	3,784

na is not applicable.

Note: See table 2 in the main report for definitions of outcomes. The sample includes all observations during academic terms in which the student was enrolled for at least 50 percent of possible days in Pittsburgh Public Schools or Propel Schools during either the 2015/16 or 2016/17 school year. The actual sample that contributed to each bivariate regression may be less than the numbers in the table, due to missingness of individual predictors. Grade ranges are K–5 for elementary school, 6–9 for middle school, and 9–12 for high school. Students with missing grade-level information were included in the total sample size columns but not in the columns by grade span; thus, the total number of observations does not equal the sum of students in elementary, middle, and high school.

a. The Pennsylvania System of School Assessment for elementary and middle school students and Keystone exams for high school students.

Source: Authors' calculations using data from Pittsburgh Public Schools, Propel Schools, and the Allegheny County Department of Human Services for school years 2014/15–2016/17.

For the bivariate regression models (research question 1) the sample consisted of all observed outcomes over the two-year period. Since analyzing the bivariate associations required both outcome and predictor data, the actual sample underlying each estimate was smaller than shown in table A2 for analyses for which predictor data were missing. For the predictive modeling the sample was separated into a training set (2015/16) and a test set

(2016/17), as described in the section on predictive modeling methods. Missing predictor data did not affect the sample for these models because the study team implemented a missing data approach known as "missingness incorporated in attributes" (Twala, Jones, & Hand, 2008). This approach creates a binary "missing flag" for each predictor that contains missing values, thereby treating students who are missing that variable as their own category. This avoids excluding these students from the predictive analysis while also allowing for the potential that missingness is informative.

Students for whom grade-level information was missing were not included in any of the results that are separated by grade span, but they are included in other results. Thus, the total number of observations for any given combination of outcome and local education agency does not equal the sum of the observations from elementary, middle, and high school. Additionally, the total number of unique students does not equal the sum of the number of unique students in each grade span because some students are associated with multiple grade levels, even within a single academic year.

Composition of sample

Table A3 contains data on the characteristics of students in the descriptive analysis sample. Tables A4 and A5 show the frequency and duration of Allegheny County DHS involvement for this sample.

Table A3. Demographic characteristics and school service eligibility of sample, Pittsburgh Public Schools and Propel Schools, 2015/16 and 2016/17 (percent of students)

Student characteristic or school service eligibility	Pittsburgh Public Schools (n = 28,719)	Propel Schools (n = 4,614)
Student characteristic		
Gender		
Male	51	49
Female	49	51
Race/ethnicity		
American Indian and Native Hawaiian/Pacific Islander	<1	<1
Black	53	69
Hispanic	3	2
Multiracial	8	8
White	33	21
School service eligibility		
Economic disadvantage (eligible for national school lunch program)	64	82
In special education (has an Individualized Education Program)	21	19
Eligible for English as a second language services	3	<1

Note: Table includes all students in the descriptive analysis sample.
Source: Authors' analysis of administrative data from Pittsburgh Public Schools and Propel Schools for school years 2015/16 and 2016/17.

Table A4. Frequency and duration of student involvement with the Allegheny County Department of Human Services, Pittsburgh Public Schools sample, 2015/16 and 2016/17

Type of involvement	Percent of sample	Mean duration of involvement[a] (standard deviation)	Median duration of involvement[a] (interquartile range)
Duration measured in days			
Behavioral health services			
Outpatient behavioral health services	16	27 (44.0)	12 (4–32)
Counseling services	5	45 (56.3)	28 (8–61)
Inpatient behavioral health services	1	38 (60.4)	14 (8–29)
Duration measured in months			
Child welfare services			
Child welfare nonplacement services	4	8.5 (5.9)	8 (4–11)
Child welfare placement services	2	7.4 (5.5)	6 (3–11)
Housing and family support services			
Any homeless service	3	1.2 (0.4)	1 (1–1)
Any homeless service started	2	1.0 (0.1)	1 (1–1)
Emergency shelter assistance	1	1.0 (0.2)	1 (1–1)
Rental assistance and prevention	2	1.1 (0.4)	1 (1–1)
Head Start	2	2.9 (0.6)	3 (3–3)
Energy assistance[b]	<1	5.2 (5.9)	3 (2–6)
Justice system involvement			
Active case in family court	4	1.3 (0.7)	1 (1–1)
Active case in the juvenile justice system	3	1.2 (0.4)	1 (1–1)
Time spent in county jail	<1	2.6 (2.7)	1 (1–3)
Adult probation	<1	20.3 (17.7)	13 (9–25)
Public benefits			
HealthChoice[c]	63	15.8 (6.2)	20 (11–21)
Supplemental Nutrition Assistance Program	34	10.1 (6.9)	9 (4–16)
Temporary Assistance for Needy Families	15	10.7 (7.1)	10 (5–17)
Section 8 housing choice voucher program	15	16.0 (7.2)	21 (11–22)
Low-income public housing	6	16.0 (7.1)	21 (11–22)

Note: Table includes all students in the descriptive analysis sample.
a. Calculated based on students who have any involvement during the two-year period (excluding zeroes for students with no involvement).
b. Low Income Home Energy Assistance Program.
c. Pennsylvania's managed care program for individuals eligible for Medicaid.
Source: Authors' analysis of administrative data from the Allegheny County Department of Human Services for school years 2015/16 and 2016/17.

Table A5. Frequency and duration of student involvement with the Allegheny County Department of Human Services, Propel Schools sample, 2015/16 and 2016/17

Type of involvement	Percent of sample	Mean duration of involvement[a] (standard deviation)	Median duration of involvement[a] (interquartile range)
Duration measured in days			
Behavioral health services			
Outpatient behavioral health services	13	15 (23.4)	7 (2–16)
Counseling services	5	44 (52.1)	25 (8–63)
Inpatient behavioral health services	1	28 (46.0)	13 (8–26)
Duration measured in months			
Child welfare services			
Child welfare nonplacement services	2	7.0 (4.7)	6 (4–10)
Child welfare placement services	1	6.4 (4.5)	6 (3–10)
Housing and family support services			
Any homeless service	4	1.2 (0.4)	1 (1–1)
Any homeless service started	3	1.0 (0.1)	1 (1–1)
Emergency shelter assistance	<1	1.1 (0.3)	1 (1–1)
Rental assistance and prevention	2	1.2 (0.4)	1 (1–1)
Head Start	4	2.5 (0.6)	3 (2–3)
Energy assistance[b]	<1	na	3 (3–3)
Justice system involvement			
Active case in family court	2	1.1 (0.4)	1 (1–1)
Active case in the juvenile justice system	1	1.1 (0.3)	1 (1–1)
Time spent in county jail	<1	1.3 (0.6)	1 (1–2)
Adult probation	<1	na	1 (1–1)
Public benefits			
HealthChoice[c]s	67	14.7 (6.1)	12 (11–21)
Supplemental Nutrition Assistance Program	34	8.4 (6.1)	6 (4–11)
Temporary Assistance for Needy Families	12	9.5 (6.3)	8 (5–11)
Section 8 housing choice voucher program	16	10.2 (8.0)	11 (1–18)
Low-income public housing	6	11.1 (7.9)	11 (2–21)

Note: Table includes all students in the descriptive analysis sample.
a. Calculated based on students who have any involvement during the two-year period (excluding zeroes for students with no involvement).
b. Low Income Home Energy Assistance Program.
c. Pennsylvania's managed care program for individuals eligible for Medicaid.
Source: Authors' analysis of administrative data from the Allegheny County Department of Human Services for school years 2015/16 and 2016/17.

Prevalence of academic problems in the sample

The prevalence of academic problems varied across outcome and local education agency. For approximately one-third of student-periods in PPS and 40 percent in Propel Schools, students had a GPA below 2.2 (a threshold identified by the stakeholders to identify at-risk students) during the two-year outcome period (2015/16 and 2016/17; table A6). State test performance varied by subject, but in both PPS and Propel Schools, students scored below the basic level on more than 40 percent of Pennsylvania System of School Assessment (PSSA) math tests and Keystone biology exams (figure A1). Other outcomes examined in this study—including chronic absenteeism, suspensions, and core course failure—were less frequent. Across both PPS and Propel Schools, high school students experienced academic problems more frequently than did elementary school students.

Table A6. Frequency of outcomes in 2015/16 and 2016/17, Pittsburgh Public Schools and Propel Schools samples (percent of student-periods)

	Pittsburgh Public Schools (n = 28,719)		Propel Schools (n = 4,614)	
Outcome	K–8	9–12	K–8	9–12
Chronic absenteeism	11	24	5	15
One or more out-of-school suspension	5	8	1	4
Term grade point average below 2.2	na	33	na	40
Score below basic level on state tests[a]				
PSSA English language arts	17	na	14	na
PSSA math	46	na	45	na
PSSA science	31	na	25	na
Keystone algebra I	na	42	na	16
Keystone biology	na	52	na	42
Keystone literature	na	27	na	15
Courses failed[b]				
English language arts	3	12	7	14
Math	4	12	8	11
Science	3	11	4	14
Social studies	2	9	4	12

na is not applicable; PSSA is Pennsylvania System of School Assessment.

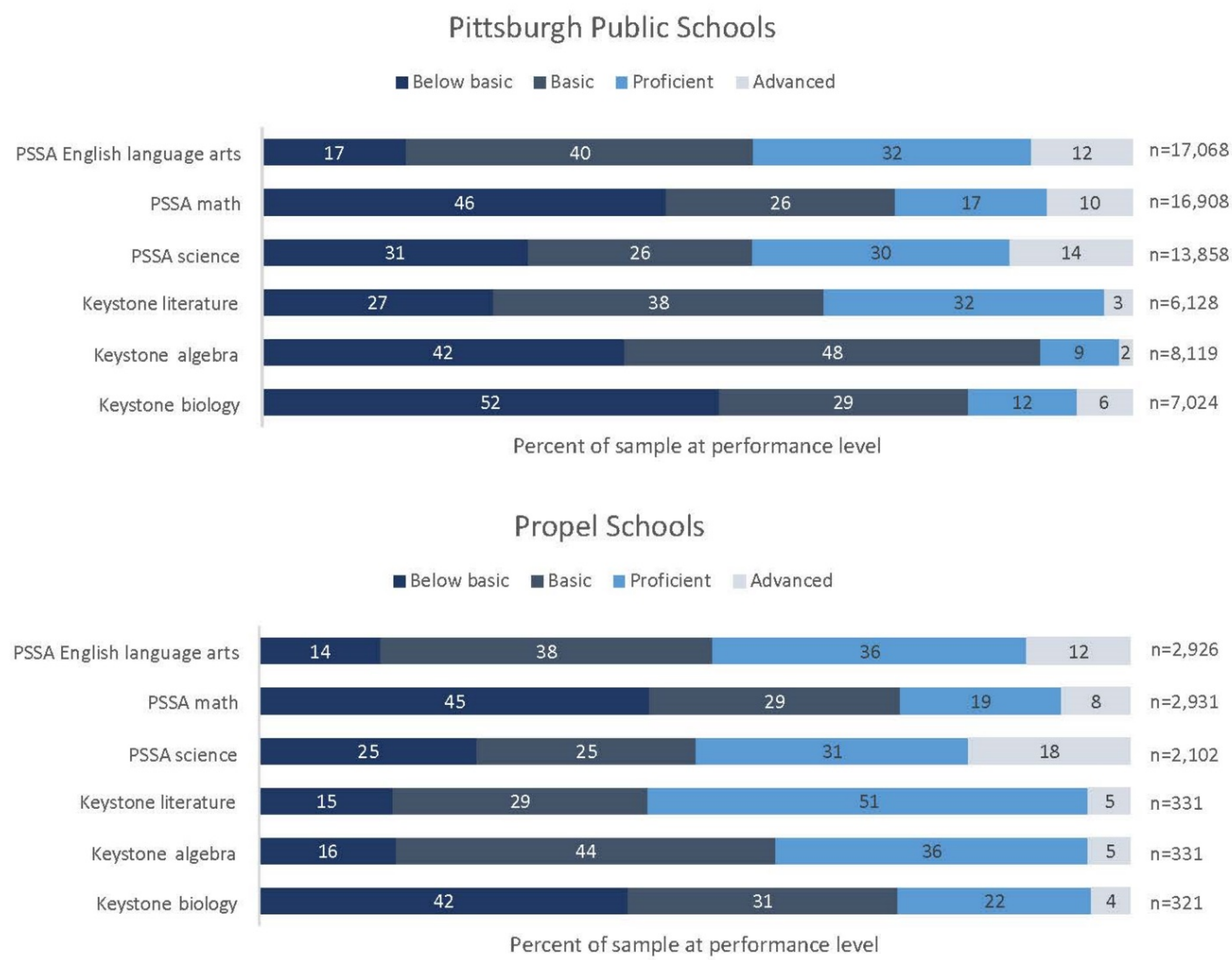
Note: Table includes all students in the descriptive analysis sample. The percentages of chronic absenteeism, suspensions, and low grade point average indicate the proportion of student-periods with each outcome during 2015/16 or 2016/17.

a. Proportion of state tests taken on which students scored below the basic level.

b. Proportion of courses taken for which students received a failing grade.

Source: Authors' analysis of data from Pittsburgh Public Schools and Propel Schools for school years 2015/16 and 2016/17.

Figure A1. Student scores on the Pennsylvania System of School Assessment and Keystone exam, 2015/16 and 2016/17



PSSA is Pennsylvania System of School Assessment.
Note: Figure includes all students in the descriptive analysis sample. The percentages indicate the proportion of total tests taken that resulted in the indicated performance level.
Source: Authors' analysis of administrative data from Pittsburgh Public Schools and Propel Schools for school years 2015/16 and 2016/17.

Descriptive analysis methods

Using the outcome and predictor data files described above, the study team fit separate simple linear regression models for each combination of outcome, predictor, local education agency, and grade span (elementary, middle, and high school). Each model included only one predictor at a time. Because the outcomes are all binary variables, these regression models are known as linear probability models (Wooldridge, 2013).

Most of the predictors in these models are also binary variables. In those cases, the regression coefficient associated with the predictor is reported. The regression coefficient can be interpreted as the absolute difference in the proportion of students who have that outcome when the predictor value is one and the proportion of students who have that outcome when the predictor value is zero. This is an absolute difference, equivalent to subtracting the two proportions. For predictors that are continuous variables, results are presented as the change in the proportion of students who have the outcome for every two standard deviation increase in the value of the predictor. This method makes the results more comparable to those for the binary predictors (Gelman, 2008).

Because of the large number of bivariate analyses, the results are shown in graphical format as heat maps to indicate the direction and strength of each predictor–outcome relationship. For the large number of Allegheny County DHS predictors analyzed, heat maps show the selection of DHS variables displayed in table A5 (see figure 2 in the main report for the PPS sample and figure B7 in appendix B for the Propel Schools sample). Child welfare predictors are presented separately because child welfare services are of high interest to stakeholders. For the other DHS categories for the PPS sample, the figures present the predictors in each category with the greatest probability difference across grades and outcomes and the most frequently observed predictor.

The predictors that meet the "greatest probability difference" and "most frequent" criteria in each category include:

- Behavioral health services: outpatient behavioral health (greatest probability difference and most frequent).
- Housing and family support services: emergency shelter service for seven days or more (greatest probability difference); any homeless service received (most frequent).
- Other housing supports (public benefits): low-income public housing (greatest probability difference and most frequent).
- Juvenile justice: case adjudicated delinquent, assigned day treatment (greatest probability difference); active case in juvenile justice system (most frequent).
- Jail and adult probation: time spent in county jail (greatest probability difference and most frequent).
- Family court: active family court case (greatest probability difference and most frequent).

Additional figures in appendix B include all Allegheny County DHS predictors examined in the descriptive analysis.

Predictive modeling methods

The study team built a predictive model to calculate a risk score for each student for each outcome, with the goal of achieving the best predictions possible. The study team decided against linear or logistic regression models, which rely on strong parametric assumptions on the functional form (such as linearity and additivity of the effects of predictors) in favor of three more flexible machine learning techniques, which take full advantage of the rich data sources available. Machine learning models use data-driven algorithms designed to extract the most relevant information from a dataset, with a focus on maximizing the predictive performance of the model. They are particularly useful when there is no strong theory to guide the way predictors interact, which is common when data come from multiple, loosely related sources. Machine learning approaches are also advantageous when events occur over time and when complex, long-term dependencies exist between predictors and outcomes. All of these features characterize the study data.

Machine learning algorithms. The study team explored the use of three machine learning algorithms: random forests, elastic net logistic regression, and recurrent neural networks.

Random forest (RF). An RF is an ensemble predictive model that is made up of many decision trees. Like decision trees (commonly known as classification and regression trees, or CART models), random forests can identify nonlinear relationships and interactions between predictors. Because they can fit many decision trees (each constructed slightly differently because of randomness), they tend to be more robust than standard CART models. The study team used a grid search and 10-fold cross validation to optimize the tuning parameters.

The input predictors for the RF were taken from the set of aggregated predictors that were used for the descriptive analysis. In some cases, the study team used different forms of these predictors than the one used in the descriptive analysis. For example, there were a number of continuous variables that took the value of 0 for most students, indicating that the students never used a particular service or experienced the event. For the descriptive analysis the study team dichotomized these predictors into zero and greater than zero because these approaches assume linear trends between the predictors and outcomes that are unlikely to hold in their raw form. For the RF model, dichotomization is unnecessary because the RF algorithm automatically identifies relevant thresholds. Therefore, these variables were included as continuous variables in the RF.

Elastic net (EN) logistic regression. An elastic net is a generalized linear model that contains a penalty term on its parameters, making it more appropriate for situations where there are many predictors. This penalty term helps prevent overfitting, typically leading to better predictive performance than an unregularized model.

Because all the outcomes were binary, the generalized linear model used is a logistic regression model. The logistic regression (and thus the EN) assumes a linear relationship between predictors and the log odds of each outcome. Therefore, the study team used the same dichotomized versions of some of these predictors that were used in the descriptive analysis, unlike the RF model, which used their raw (continuous) versions.

Recurrent neural network (RNN). The final machine learning approach used in the study is a recurrent network, designed to optimally leverage the longitudinal nature of the input data. RNN models include a combination of linear and nonlinear functions working together in specified ways; they have gained popularity in recent years for their ability to process sequences of data. Their primary advantage over other predictive models is the ability to learn dependencies between variables and events across long time spans. A student's past predictor history may be important in predicting future outcomes, but it is not clear, a priori, how each past observation can best contribute to the prediction of the outcome at any given time in the future. Traditional approaches to predictive modeling with longitudinal data, including the RF and EN models, require the analyst to summarize the history of each predictor using a small number of time-varying covariates. The number and form of these covariates are typically fixed. In fact, most commonly only the most recent value of a predictor (such as the number of behavioral health services in the most recent time period) is allowed to affect the predicted outcome. RNN models relax these assumptions and instead learn the most appropriate form of these dependencies in a flexible, data-driven manner.

Model selection. After weighing the tradeoffs between each modeling approach, the study team selected the RF models as the primary model presented in this report. This decision is based on several factors, including theory, practicality, and empirical performance.

From a theoretical perspective, the RF and RNN models have distinct advantages over the EN model in that they both allow for nonlinearities and interactions in the relationship between predictors and the outcome. This should allow the model to capture more complicated relationships that the EN model misses. This feature is especially important in the case of continuous predictors and in making predictions for individuals who have outlying predictor values. In these cases the EN model may produce unreasonable predictions because it is extrapolating a linear relationship to a predictor value that does not occur often (if at all) in the training data. That prediction could be based almost entirely on the value of the one predictor. The RF and RNN models, on the other hand, are not subject to this linear extrapolation and thus would be more robust to outlying predictor values. Theory would also lead to the selection of the RNN model over the RF model because the RNN can automatically detect complex, long-term dependencies between predictors and outcomes and does not require aggregating data and potentially throwing away valuable information.

However, the RNN model suffers from a significant practical limitation in that it is much more difficult to train and tune than the RF and EN models. There are more internal model parameters in the RNN model, which require more training data to estimate. In addition, the model contains more hyperparameters to tune, which combined with longer training times, makes model tuning a much more computationally expensive process. Both the RF and EN models, on the other hand, train relatively quickly and easily and have robust, automated methods for tuning that perform well with minimal oversight by a data analyst.

Finally, in comparisons of the relative performance of the three approaches, the two simpler algorithms (RF and EN) outperformed the RNN in every scenario (see figure B14 in appendix B).

The study team concluded that the theoretical benefits of the RNN arising from its flexibility and ability to capture longitudinal dependencies are outweighed by the practical limitations of the difficulty training and tuning the model. This may be because the long-term dependencies are relatively weak, and the more immediate effects (which are captured by all three modeling approaches) are sufficient to express the relationship between predictors and outcomes. Considering theoretical advantages to handle nonlinearities and interactions, ease of implementation, and strong predictive performance, the team selected the RF as the primary model.

Model validation. To assess how the model will perform on a future dataset, the team trained the model only on data through 2015/16. After training, the model was used to predict risk scores for all outcomes in 2016/17, allowing for testing model performance on new data.

To assess model performance in 2016/17, the predicted probabilities returned by the model can be compared with the actual outcomes for students in the sample. There are several ways to compares these. One is to dichotomize the predicted probabilities at a selected threshold. For example, .5 would be an intuitive threshold, meaning that students with a predicted probability of 50 percent or more of having an outcome would be labeled "predicted positive," and those with a predicted probability of less than 50 percent would be labeled "predicted negative." With this thresholding in place, predicted outcomes can be compared with actual (observed) outcomes using a two-by-two contingency table (table A7).

Table A7. Contingency table for comparing predicted outcomes with observed outcomes

	Predicted negative	Predicted positive
Observed negative	True negative	False positive
Observed positive	False negative	True positive

Source: Authors.

The contingency table informs several common summary statistics of model performance:

- Sensitivity (true positive rate, also known as recall) is the proportion of observed positives that are correctly predicted to be positive: true positives / (true positives + false negatives).
- False positive rate is the proportion of observed negatives that are incorrectly labeled as positive: false positives / (false positives + true negatives).
- Precision (also known as positive predictive value) is the proportion of predicted positives that are actually true: true positives / (true positives + false positives).

These performance statistics are contingent on the threshold selected to identify at-risk students (positives) and not at-risk students (negatives). As the threshold rises, both the sensitivity and false positive rate fall (or stay the same) because raising the threshold means that fewer students are classified as positive.

An alternative approach to choosing a threshold is to examine a sensitivity and false positive rate for every possible threshold; a plot of these values against one another is known as a receiver operating characteristic (ROC) curve. In this report ROC curves illustrate the overall predictive performance of a model. A common, one-number summary measure of an ROC curve is the area under the curve (AUC), which ranges from 0 to 1. It can be interpreted as the probability that a randomly selected student with an academic problem is considered at higher risk (by the model) than a randomly selected student without an academic problem. An AUC of 1 indicates perfect prediction (all students with academic problems are predicted to have risk scores higher than all students without observed academic problems), whereas an AUC of .5 indicates that the model is predicting as well as a coin flip.

In cases where sensitivity, false positive rate, or precision was presented, the study team chose the threshold closest to the point of perfect prediction (100 percent sensitivity, 0 percent false positive rate) on the ROC curve.

The study team calculated these performance statistics for each combination of local education agency and outcome. The results are presented overall as well as separated into subgroups, such as by grade span, race/ethnicity, and gender.

To assess the extent to which out-of-school predictors affect model performance, the study team calculated performance metrics for the RF models excluding predictors from Allegheny County DHS data. The team compared these metrics with the performance of models using all available in-school and out-of-school predictors.

Variable Importance. For the final (RF) models, the study team calculated the variable importance for each variable used to train the model, using the Gini impurity index (Nembrini, Koenig, & Wright, 2018). This index ranks each predictor according to its contribution in driving the predictions. Predictors that rank among the 10 most important variables for each model are reported (see table 2 in the main report).

Selecting optimal risk sore cutoff points. To run the predictive models, users have to translate risk scores that range from 0 to 1 into lower and higher risk categories. While many criteria can be applied to select the optimal cutoff for "high risk," the study used a common criterion known as the Youden statistic (Youden, 1950). Maximizing the Youden statistic is equivalent to maximizing the difference between the true positive rate (sensitivity) and the false positive rate.

References

Ballou, D., Sanders, W., & Wright, P. (2004). Controlling for student background in value-added assessments of teachers. *Journal of Educational and Behavioral Statistics*, *29*(1), 37–65. https://eric.ed.gov/?id=EJ727499.

Gelman, A. (2008). Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine, 27*(15), 2865–2873.

Johnson, M. T., Lipscomb, S., and Gill, B. (2015). Sensitivity of Teacher Value-Added Estimates to Student and Peer Control Variables. *Journal of Research on Educational Effectiveness, 8*(1), 60–83. https://eric.ed.gov/?id=EJ1049638.

Nembrini, S., Koenig, I. R. & Wright, M. N. (2018). The revival of the Gini Importance? *Bioinformatics 34*(21), 3711–3718.

Twala, B., Jones, M. C., & Hand, D. (2008). Good methods for Coping with Missing Data in Decision Trees. *Pattern Recognition Letters, 29*(7), 950–956.

Wooldridge, J. M. (2013). A binary dependent variable: The linear probability model. *Introductory econometrics: A modern approach*. Mason, OH: South-Western.

Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer, 3*(1), 32–35.

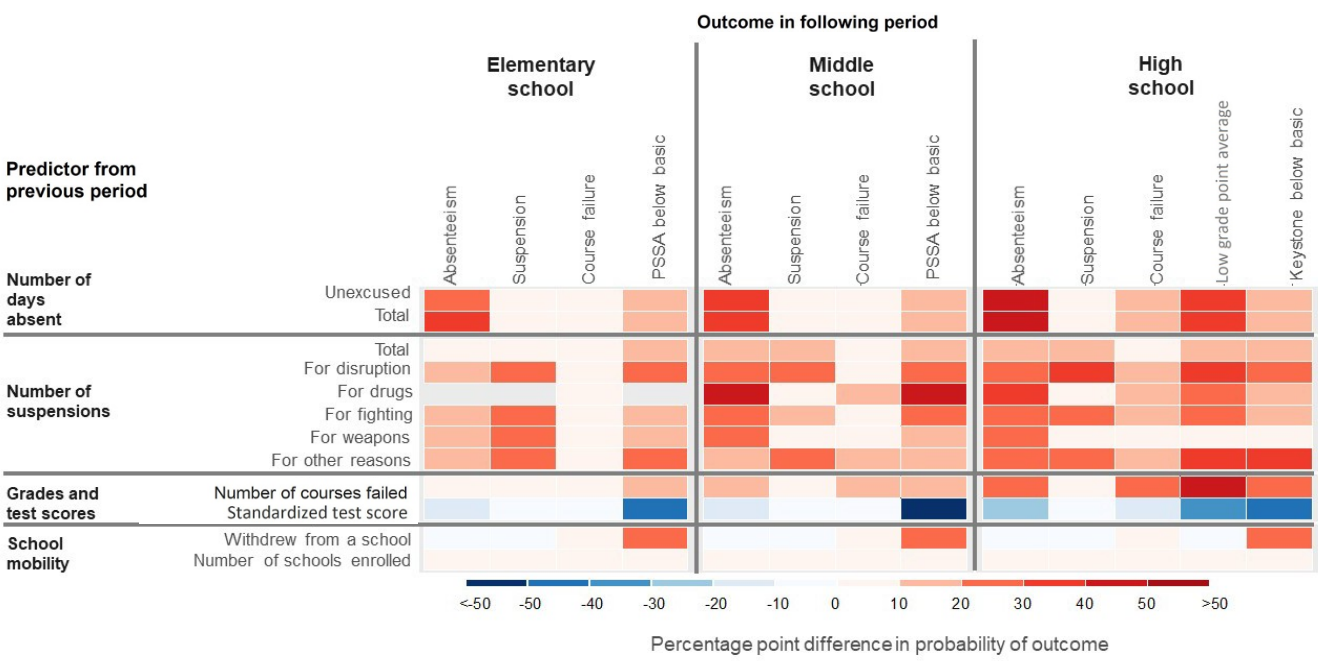
Appendix B. Supporting analyses

Supporting analyses of unadjusted relationships between predictors and outcomes

The heat maps in the main report displaying the relationships between predictors and outcomes show only a handful of the analyses conducted for this study, and only for PPS, the larger local education agency in the study (figures 1 and 2). The predictors included in the main report are those of highest interest to stakeholders, those associated with the greatest probability differences for outcomes, and those most frequently observed in the sample. Figures B1–B12 present findings on bivariate relationships for additional predictors for PPS (figures B1–B5) and all findings for Propel Schools (figures B8–B12). Each heat map presents a different set of predictors: in-school predictors (figures B1 and B8), predictors related to child welfare services (figures B2 and B9), predictors related to involvement in the juvenile justice system (figures B3 and B10), predictors related to behavioral health and housing services (figures B4 and B11), and demographic predictors (figures B5 and B12).

The first two heat maps for Propel Schools display differences in probability of academic problems for students with prior academic problems (figure B6) and difference in probability of academic problems for students with selected types of human services involvement (figure B7) and parallel the figures for PPS in the main report (figures 1 and 2).

Figure B1. Heat map showing differences in probability of academic problems for students with prior in-school events in adjacent time periods during the 2015/16 and 2016/17 school years, Pittsburgh Public Schools sample

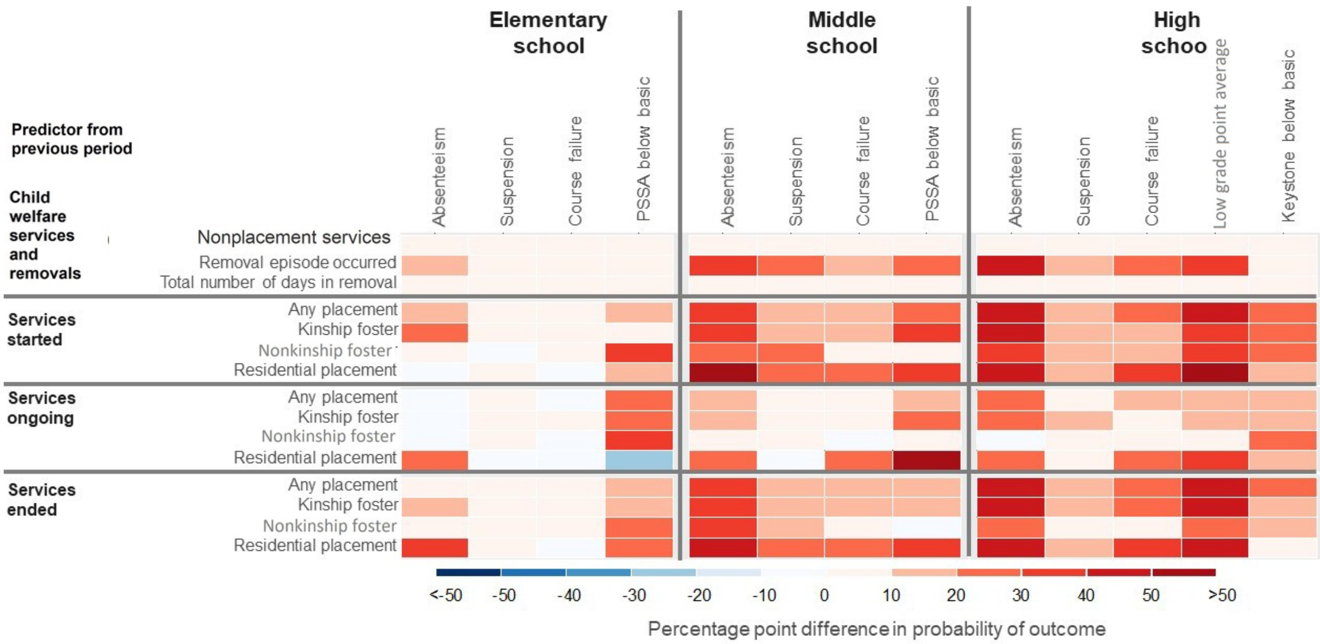


PSSA is Pennsylvania System of School Assessment.

Note: See table 2 of the main report for definitions of outcomes. For binary predictors, saturation of color indicates the difference in probability of experiencing the outcome for students with and without the given predictor. For continuous predictors (such as the number of days absent), the color indicates the difference in probability of the outcome for two students who differ by two standard deviations. Red indicates a positive relationship, blue indicates a negative relationship, and neutral colors indicate that larger values of the predictor are not, on average, associated with higher or lower likelihood of outcomes. See table B5 for the values used to generate the heat map.

Source: Authors' analysis of data from Pittsburgh Public Schools for school years 2015/16 and 2016/17.

Figure B2. Heat map showing differences in probability of academic problems for students with child welfare events in adjacent time periods during the 2015/16 and 2016/17 school years, Pittsburgh Public Schools sample

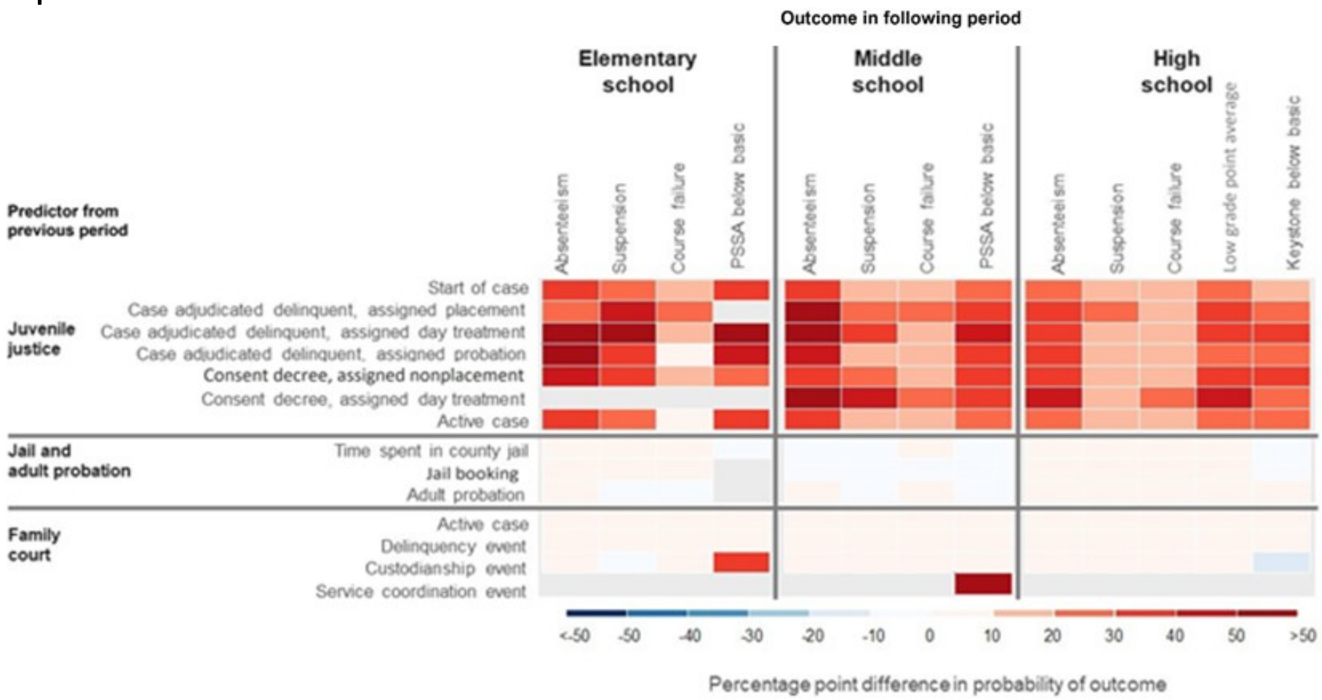


PSSA is Pennsylvania System of School Assessment.

Note: See table 2 of the main report for definitions of outcomes. For binary predictors, saturation of color indicates the difference in probability of experiencing the outcome for students with and without the given predictor. For continuous predictors (such as total number of days in removal), the color indicates the difference in probability of the outcome for two students who differ by two standard deviations. Red indicates a positive relationship, blue indicates a negative relationship, and neutral colors indicate that larger values of the predictor are not, on average, associated with higher or lower likelihood of outcomes. See table B6 for the values used to generate the heat map.

Source: Authors' analysis of data from Pittsburgh Public Schools and the Allegheny County Department of Human Services for school years 2015/16 and 2016/17.

Figure B3. Heat map showing differences in probability of academic problems for students with justice system involvement in adjacent time periods during the 2015/16 and 2016/17 school years, Pittsburgh Public Schools sample

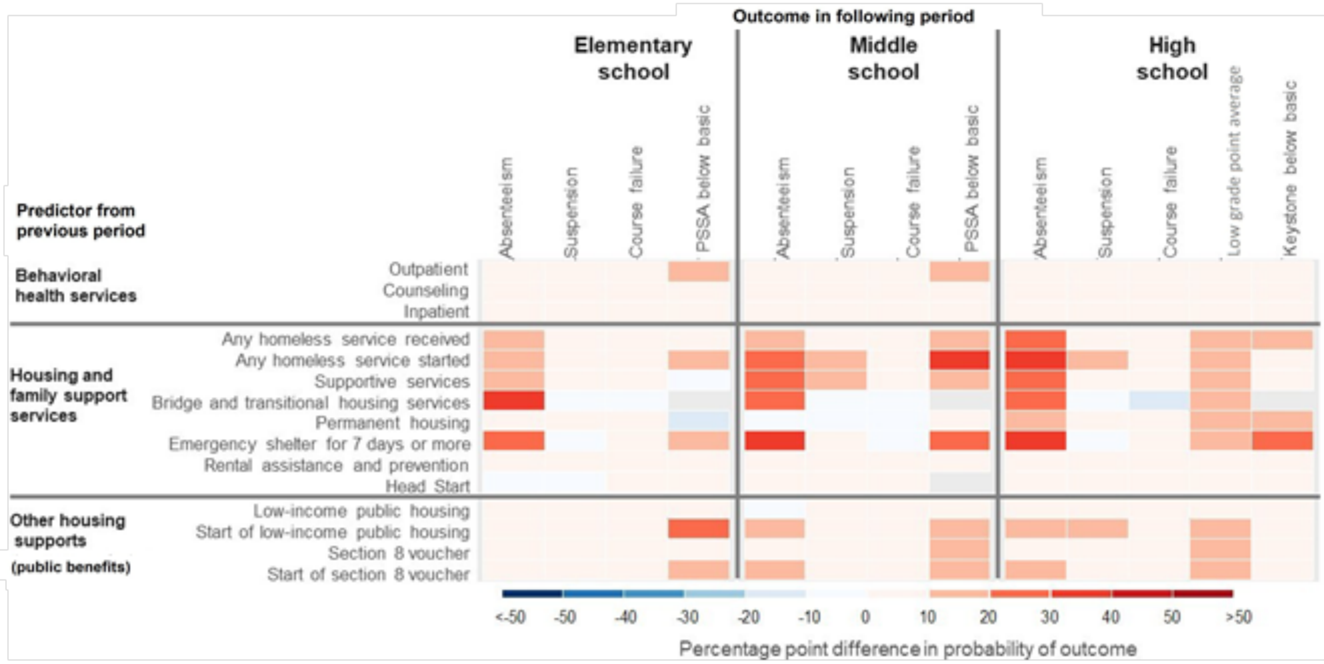


PSSA is Pennsylvania System of School Assessment.

Note: See table 2 of the main report for definitions of outcomes. Adjudicated delinquent is analogous to a “guilty” verdict for an adult and describes the sentence; consent decree is the settlement between the court and the juvenile that typically describes any required community service, day treatment, or nonplacement services. Saturation of color indicates the difference in probability of experiencing the outcome for students with and without the given predictor. Red indicates a positive relationship, blue indicates a negative relationship, and neutral colors indicate that larger values of the predictor are not, on average, associated with higher or lower likelihood of outcomes. See table B7 for the values used to generate the heat map.

Source: Authors’ analysis of data from Pittsburgh Public Schools and the Allegheny County Department of Human Services for school years 2015/16 and 2016/17.

Figure B4. Heat map showing differences in probability of academic problems for students receiving behavioral health and housing services in adjacent time periods during the 2015/16 and 2016/17 school years, Pittsburgh Public Schools sample

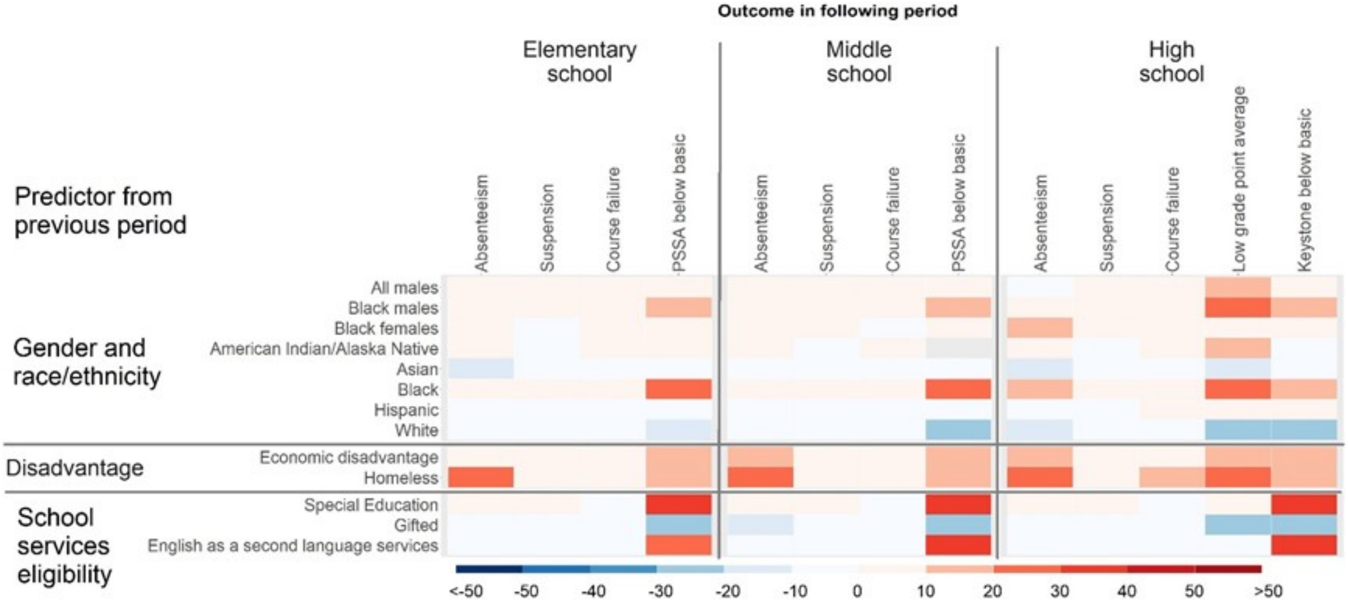


PSSA is Pennsylvania System of School Assessment.

Note: See table 2 of the main report for definitions of outcomes. For binary predictors, saturation of color indicates the difference in probability of experiencing the outcome for students with and without the given predictor. For continuous predictors (such as the number of days absent), the color indicates the difference in probability of the outcome for two students who differ by two standard deviations. Red indicates a positive relationship, blue indicates a negative relationship, and neutral colors indicate that larger values of the predictor are not, on average, associated with higher or lower likelihood of outcomes. See table B6 for the values used to generate the heat map.

Source: Authors' analysis of data from Pittsburgh Public Schools and the Allegheny County Department of Human Services for school years 2015/16 and 2016/17.

Figure B5. Heat map showing differences in probability of academic problems in adjacent time periods during the 2015/16 and 2016/17 school years, by student demographic characteristic, Pittsburgh Public Schools sample

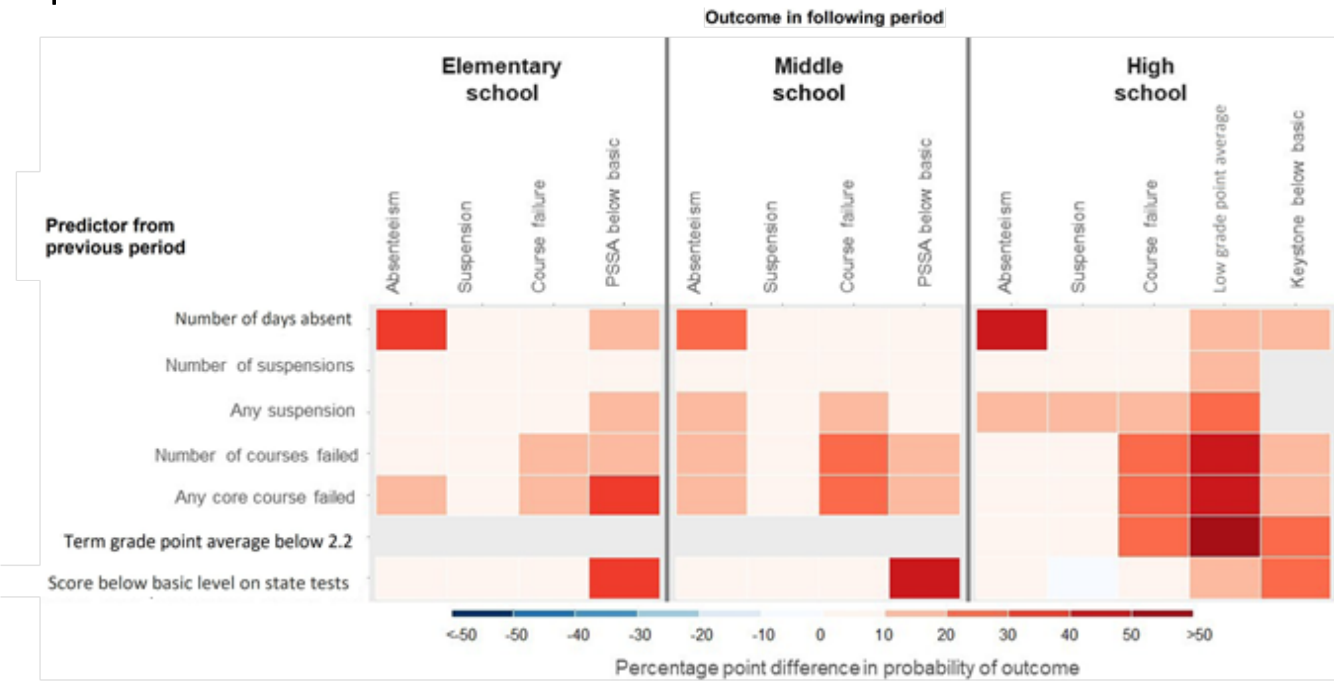


PSSA is Pennsylvania System of School Assessment.

Note: See table 2 of the main report for definitions of outcomes. Saturation of color indicates the difference in probability of experiencing the outcome for students with and without the given predictor. Red indicates a positive relationship, blue indicates a negative relationship, and neutral colors indicate that larger values of the predictor are not, on average, associated with higher or lower likelihood of outcomes. See table B8 for the values used to generate the heat map.

Source: Authors' analysis of data from Pittsburgh Public Schools for school years 2015/16 and 2016/17.

Figure B6. Heat map showing differences in probability of academic problems for students with prior academic problems in adjacent time periods during the 2015/16 and 2016/17 school years, Propel Schools sample

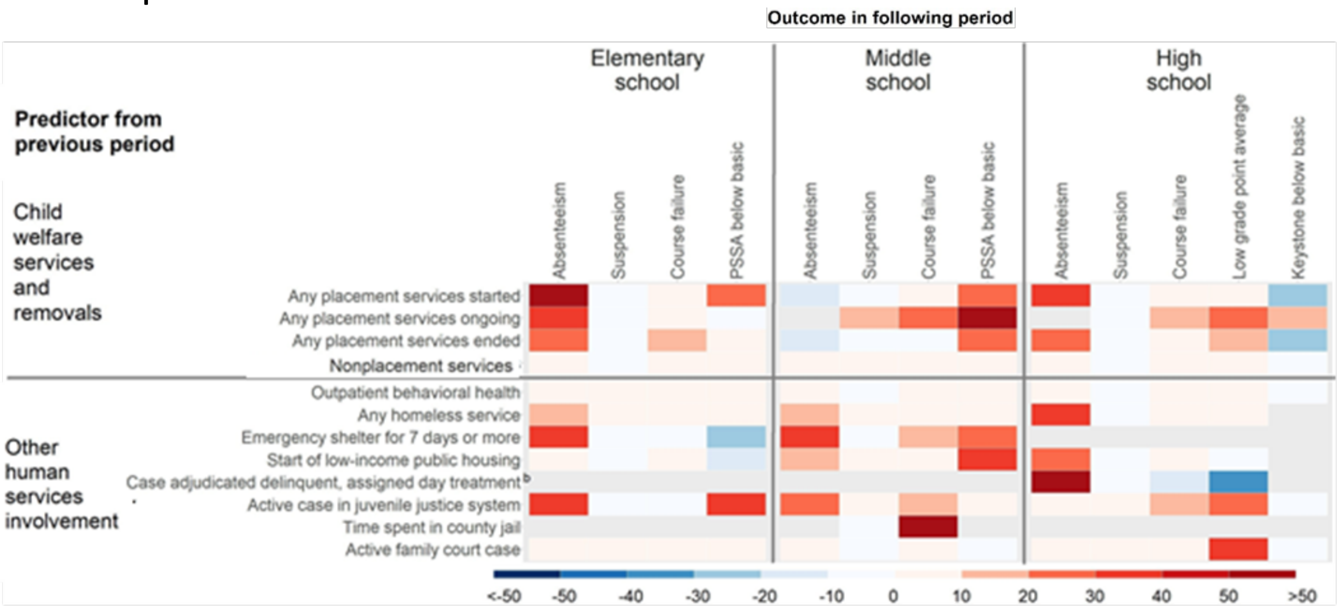


PSSA is Pennsylvania System of School Assessment.

Note: See table 2 of the main report for definitions of outcomes. For binary predictors, saturation of color indicates the difference in probability of experiencing the outcome for students with and without the given predictor. For continuous predictors (such as the number of days absent), the color indicates the difference in probability of the outcome for two students who differ by two standard deviations. Red indicates a positive relationship, blue indicates a negative relationship, and neutral colors indicate that larger values of the predictor are not, on average, associated with higher or lower likelihood of outcomes. See table B9 for the values used to generate the heat map.

Source: Authors' analysis of data from Propel Schools for school years 2015/16 and 2016/17.

Figure B7. Heat map showing differences in probability of academic problems for students with selected types of human services involvement in adjacent time periods during the 2015/16 and 2016/17 school years, Propel Schools sample



PSSA is Pennsylvania System of School Assessment.

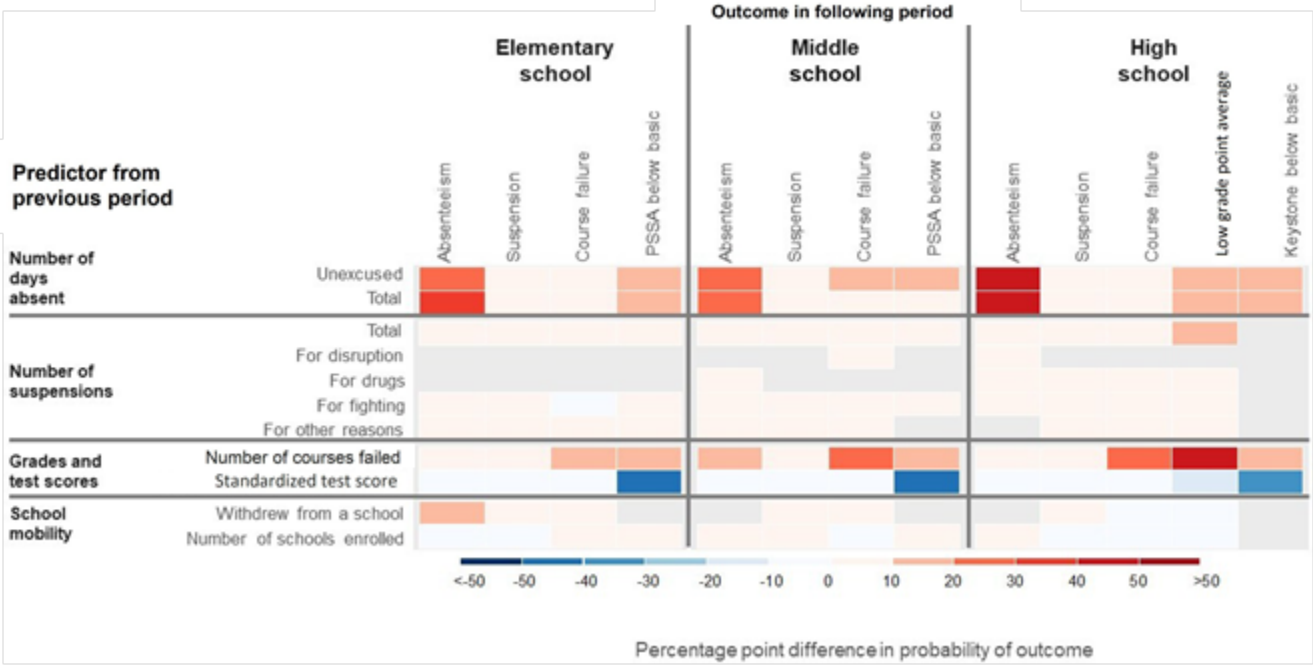
Note: See table 2 of the main report for definitions of outcomes. Saturation of color indicates the difference in probability of experiencing the outcome for students with and without the given predictor. Red indicates a positive relationship, blue indicates a negative relationship, and neutral colors indicate that larger values of the predictor are not, on average, associated with higher or lower likelihood of outcomes. See tables B10 and B11 for the values used to generate the heat map.

a. Services provided in the community or home to children with an active child welfare case. Services include housing supports, counseling and behavioral health treatment, after school programming, youth mentoring, and crisis interventions.

b. A juvenile justice case adjudicated "delinquent" is analogous to a "guilty" verdict in an adult case.

Source: Authors' analysis of data from Propel Schools and the Allegheny County Department of Human Services for school years 2015/16 and 2016/17.

Figure B8. Heat map showing differences in probability of academic problems for students with prior in-school events in adjacent time periods during the 2015/16 and 2016/17 school years, Propel Schools sample

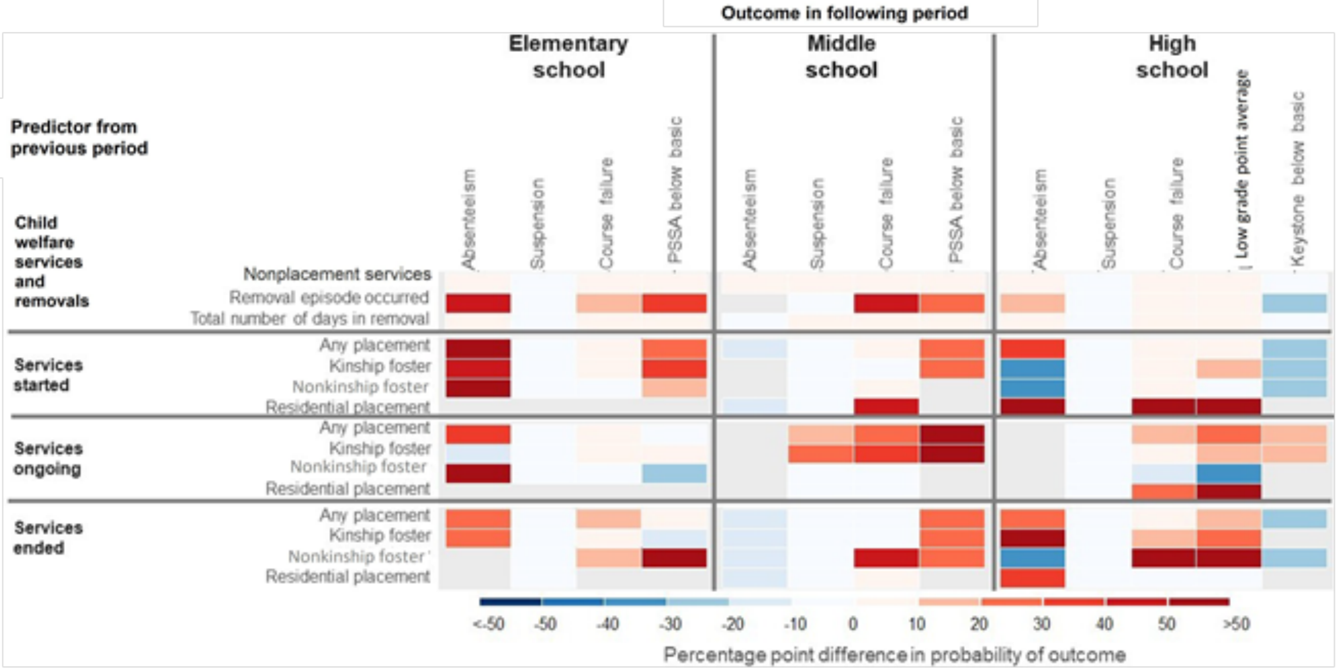


PSSA is Pennsylvania System of School Assessment.

Note: See table 2 of the main report for definitions of outcomes. For binary predictors, saturation of color indicates the difference in probability of experiencing the outcome for students with and without the given predictor. For continuous predictors (such as the number of days absent), the color indicates the difference in probability of the outcome for two students who differ by two standard deviations. Red indicates a positive relationship, blue indicates a negative relationship, and neutral colors indicate that larger values of the predictor are not, on average, associated with higher or lower likelihood of outcomes. See table B9 for the values used to generate the heat map.

Source: Authors' analysis of data from Propel Schools for school years 2015/16 and 2016/17.

Figure B9. Heat map showing differences in probability of academic problems for students with child welfare events in adjacent time periods during the 2015/16 and 2016/17 school years, Propel Schools sample

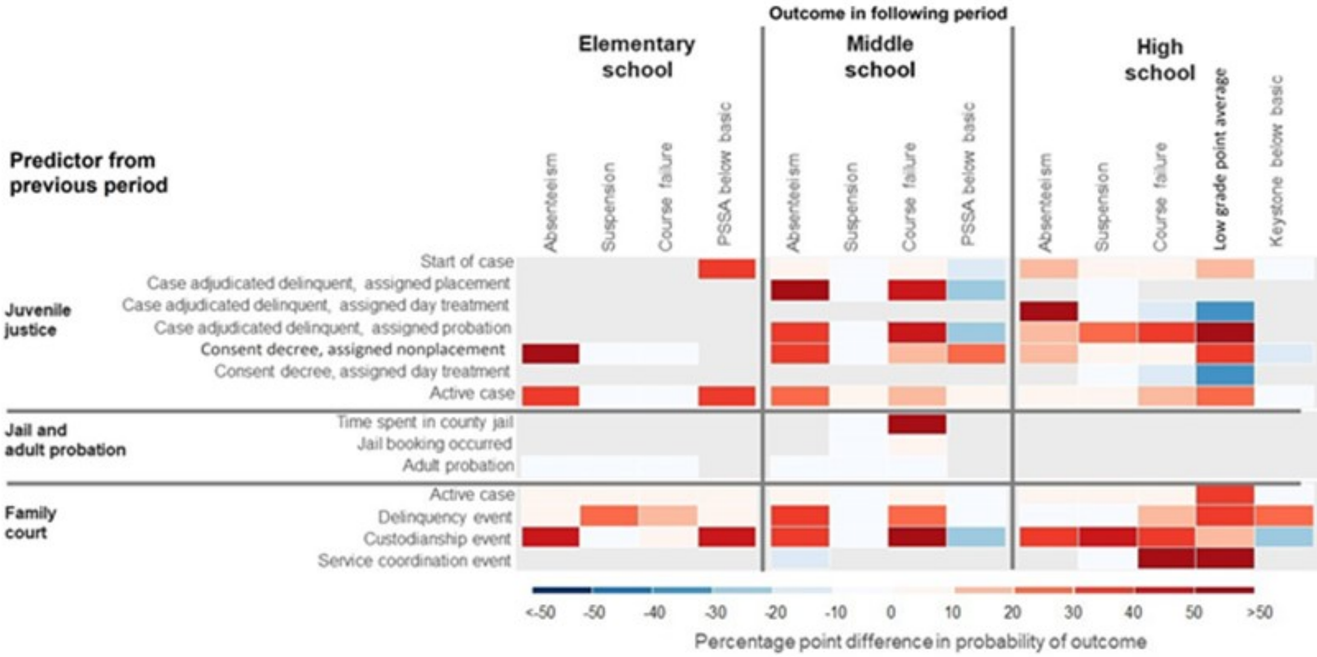


PSSA is Pennsylvania System of School Assessment.

Note: See table 2 of the main report for definitions of outcomes. For binary predictors, saturation of color indicates the difference in probability of experiencing the outcome for students with and without the given predictor. For continuous predictors (such as total number of days in removal), the color indicates the difference in probability of the outcome for two students who differ by two standard deviations. Red indicates a positive relationship, blue indicates a negative relationship, and neutral colors indicate that larger values of the predictor are not, on average, associated with higher or lower likelihood of outcomes. See table B10 for the values used to generate the heat map.

Source: Authors' analysis of data from Propel Schools and the Allegheny County Department of Human Services for school years 2015/16 and 2016/17.

Figure B10. Heat map showing differences in probability of academic problems for students with justice system involvement in adjacent time periods during the 2015/16 and 2016/17 school years, Propel Schools sample

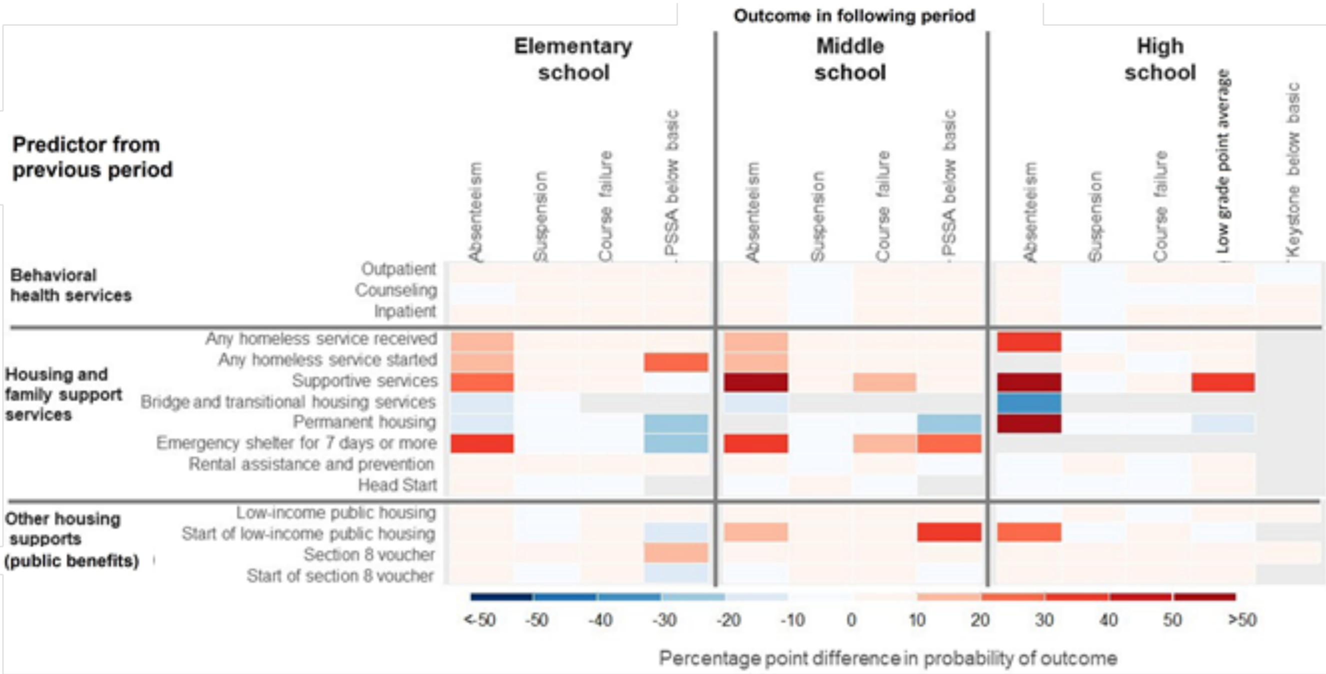


PSSA is Pennsylvania System of School Assessment.

Note: See table 2 of the main report for definitions of outcomes. Adjudicated delinquent is analogous to a “guilty” verdict for an adult and describes the sentence; consent decree is the settlement between the court and the juvenile that typically describes any required community service, day treatment or nonplacement services. Saturation of color indicates the difference in probability of experiencing the outcome for students with and without the given predictor. Red indicates a positive relationship, blue indicates a negative relationship, and neutral colors indicate that larger values of the predictor are not, on average, associated with higher or lower likelihood of outcomes. See table B11 for the values used to generate the heat map.

Source: Authors’ analysis of data from Propel Schools and the Allegheny County Department of Human Services for school years 2015/16 and 2016/17.

Figure B11. Heat map showing differences in probability of academic problems for students receiving behavioral health and housing services in adjacent time periods during the 2015/16 and 2016/17 school years, Propel Schools sample

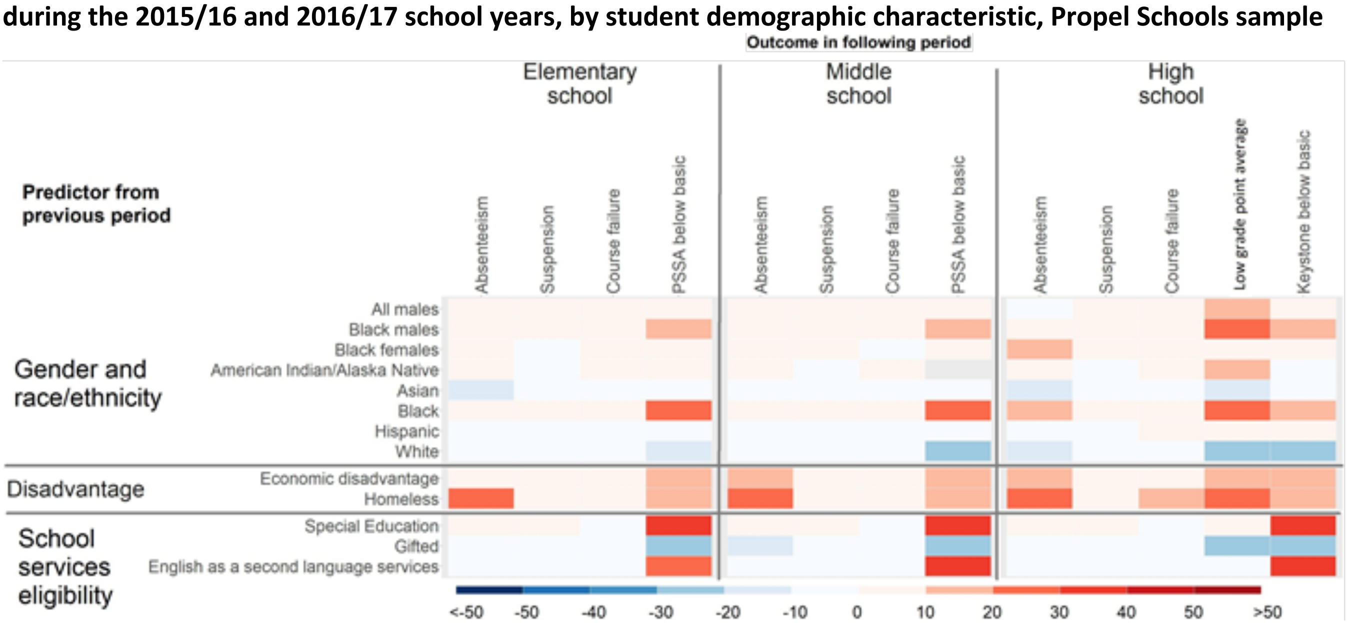


PSSA is Pennsylvania System of School Assessment.

Note: See table 2 of the main report for definitions of outcomes. Saturation of color indicates the difference in probability of experiencing the outcome for students with and without the given predictor. Red indicates a positive relationship, blue indicates a negative relationship, and neutral colors indicate that larger values of the predictor are not, on average, associated with higher or lower likelihood of outcomes. See table B10 for the values used to generate the heat map.

Source: Authors' analysis of data from Propel Schools and the Allegheny County Department of Human Services for school years 2015/16 and 2016/17.

Figure B12. Heat map showing differences in probability of academic problems in adjacent time periods during the 2015/16 and 2016/17 school years, by student demographic characteristic, Propel Schools sample



PSSA is Pennsylvania System of School Assessment.

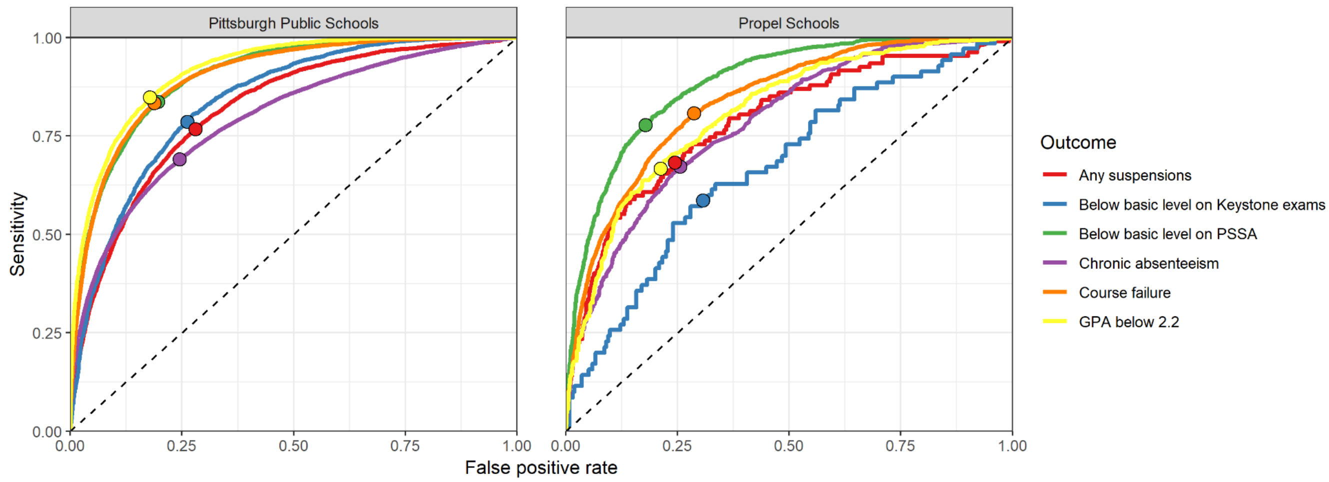
Note: See table 2 of the main report for definitions of outcomes. Saturation of color indicates the difference in probability of experiencing the outcome for students with and without the given predictor. Red indicates a positive relationship, blue indicates a negative relationship, and neutral colors indicate that larger values of the predictor are not, on average, associated with higher or lower likelihood of outcomes. See table B12 for the values used to generate the heat map.

Source: Authors' analysis of data from Propel Schools for school years 2015/16 and 2016/17 school years.

Predictive model performance

This section presents additional results related to the performance of the predictive models. Figure B13 displays the receiver operating characteristic (ROC) curves and optimal risk level cutoffs for each outcome for the Pittsburgh Public Schools and Propel Schools samples.

Figure B13. Performance of predictive models by outcome for the Pittsburgh Public Schools and Propel Schools samples, 2014/15–2016/17



PSSA is Pennsylvania System of School Assessment.

Notes: See table 2 of the main report for definitions of outcomes. The figure shows receiver operating characteristic (ROC) curves. Points on curves indicate the optimal risk level cutoff for maximizing sensitivity and minimizing the false positive rate, chosen as the point on the curves that are closest to the upper left-hand corner (the point of perfect prediction).

Source: Authors' analysis of data from Pittsburgh Public Schools and Propel Schools for school years 2014/15–2016/17.

Table B1 shows the optimal risk score cutoffs for each outcome for the Pittsburgh Public Schools and Propel Schools samples, along with the sensitivity and false positive rate at each optimal cutoff.

Table B1. Optimal risk score cutoffs by outcome for Pittsburgh Public Schools and Propel Schools samples, 2014/15–2016/17

	Pittsburgh Public Schools				Propel Schools			
Outcome	Prevalence of outcome	Optimal risk score cutoff	Sensitivity	False positive rate	Prevalence of outcome	Optimal risk score cutoff	Sensitivity	False positive rate
Chronic absenteeism	.264	.304	.755	.309	.191	.120	.739	.337
Suspensions	.050	.067	.711	.223	.010	.017	.757	.318
Course failure	.057	.065	.811	.166	.060	.073	.716	.200
Low grade point average	.322	.338	.822	.153	.378	.464	.805	.340
Score below basic level on PSSA test	.295	.347	.804	.162	.289	.372	.822	.221
Score below basic level on Keystone exam	.399	.418	.737	.208	.216	.323	.709	.400

PSSA is Pennsylvania System of School Assessment.

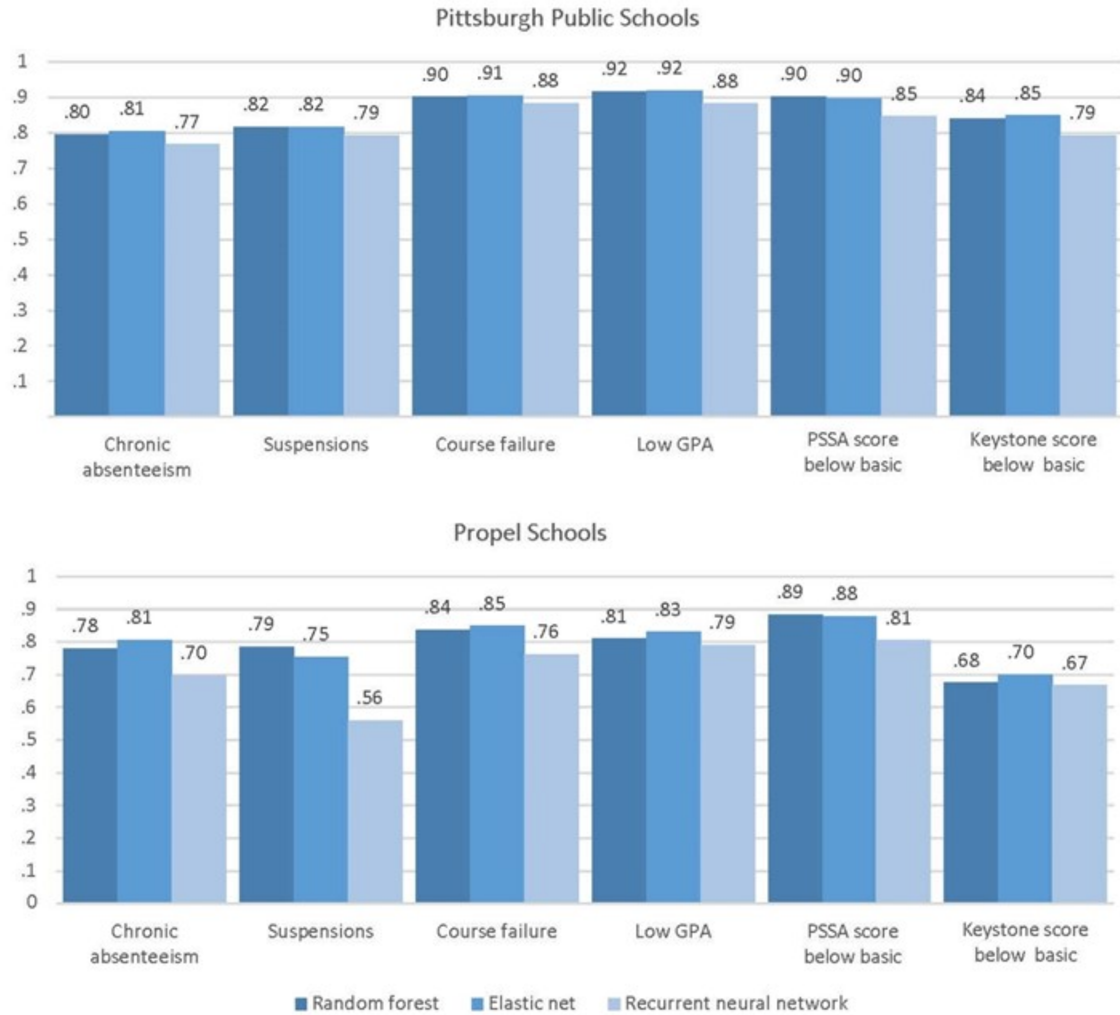
Note: See table 2 of the main report for definitions of outcomes.

Source: Authors' analysis of data from Pittsburgh Public Schools, Propel Schools, and the Allegheny County Department of Human Services for school years 2014/15–2016/17.

The study team compared the predictive performance of the random forest (RF) models to the predictive performance of two alternative approaches: elastic net logistic regression (EN) and a recurrent neural network (RNN). Appendix A includes a full description of each approach, including a discussion of the decision to highlight the findings of the RF models.

Figure B14 compares the predictive performance of the RF models to the predictive performance of the EN and RNN models, presenting the area under the ROC curves for each model, local education agency, and outcome.

Figure B14. Performance of alternative machine learning models for Pittsburgh Public Schools and Propel Schools samples, by outcome, 2014/15–2016/17 (area under receiver operating characteristic curves)



PSSA is Pennsylvania System of School Assessment.

Note: See table 2 of the main report for definitions of outcomes.

Source: Authors' analysis of data from Pittsburgh Public Schools, Propel Schools, and the Allegheny County Department of Human Services for school years 2014/15–2016/17.

Using the models run for all students in each local education agency, the study team compared how well the models predicted outcomes for subgroups of students defined by grade, gender, and race/ethnicity:

- There are no clear trends in the relationship between grade level and model performance that hold across all outcomes. Test scores show better predictive performance at the lower grade levels than in high school, indicating that Pennsylvania System of School Assessment (PSSA) scores are easier to predict than Keystone exam scores. This result is likely due to two factors. First, the sample of data for training the model was much larger for PSSA scores (see box 4 in the main report). Second, students take math and English language arts

PSSA tests each year in grades 3–8, so previous test performance is often available as a strong predictor, whereas each student usually takes Keystone exams (which are specific to a single course) only once, so prior results of the same test are generally not available.

- There are no clear trends in the relationship between gender and model performance, and differences by gender are relatively small (tables B2 and B3).
- There are some larger differences between White and Black students in the predictive performance of the models.[1] The performance of all models other than chronic absenteeism in PPS and Keystone exam scores in Propel Schools is better for White students than for Black students. This does not imply that the model exhibits any inherent racial/ethnic biases by systematically over- or under-estimating the risk level for either racial/ethnic group. Rather, it means that the relationships between students' predictors and outcomes are stronger among White students than among Black students. One reason for this is that the outcomes occur with higher frequency among Black students, so there might be more heterogeneity among at-risk Black students than among at-risk White students.

Table B2. Model performance by race and gender for Pittsburgh Public Schools sample, 2014/15–2016/17

Performance metric	Chronic absenteeism	Suspensions	Course failure	Low grade point average	Score below basic level on PSSA test	Score below basic level on Keystone exam
Area under the curve						
All students	.795	.817	.901	.917	.903	.840
Black	.803	.768	.874	.884	.878	.819
White	.777	.855	.935	.945	.924	.848
Female	.799	.837	.905	.919	.908	.848
Male	.791	.797	.894	.911	.897	.830
Sensitivity (at optimal cutoff)						
All students	.755	.711	.811	.822	.804	.737
Black	.725	.563	.738	.716	.724	.662
White	.791	.888	.906	.914	.893	.839
Female	.762	.769	.846	.852	.835	.765
Male	.750	.650	.777	.783	.772	.703
False positive rate (at optimal cutoff)						
All students	.309	.223	.166	.153	.162	.208
Black	.267	.183	.153	.133	.140	.175
White	.380	.381	.195	.204	.242	.343
Female	.306	.246	.195	.181	.182	.226
Male	.313	.205	.149	.134	.146	.193

PSSA is Pennsylvania System of School Assessment.
Note: See table 2 of the main report for definitions of outcomes.
Source: Authors' analysis of data from Pittsburgh Public Schools and the Allegheny County Department of Human Services for school years 2014/15–2016/17.

[1] The study did not examine differences for other racial/ethnic subgroups because these groups accounted for much smaller proportions of the student bodies in each local education agency. In PPS, 86 percent of students were White or Black; in Propel, 90 percent of students were White or Black.

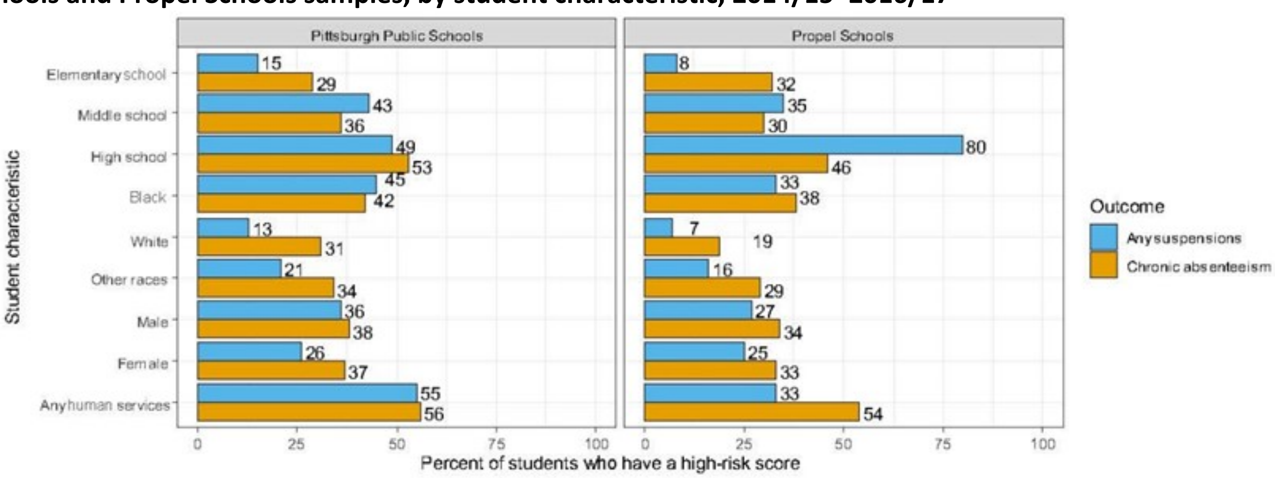
Table B3. Model performance by race and gender for Propel Schools sample, 2014/15–2016/17

Performance metric	Chronic absenteeism	Suspensions	Course failure	Low grade point average	Score below basic level on PSSA test	Score below basic level on Keystone exam
Area under curve						
All students	.782	.786	.838	.810	.885	.676
Black	.766	.749	.815	.806	.869	.684
White	.830	.944	.892	.820	.913	.583
Female	.791	.823	.844	.803	.887	.624
Male	.774	.759	.827	.806	.883	.718
Sensitivity (at optimal cutoff)						
All students	.739	.757	.716	.805	.822	.709
Black	.693	.685	.657	.796	.783	.742
White	.864	.924	.878	.860	.898	.556
Female	.749	.749	.762	.846	.848	.718
Male	.729	.751	.667	.751	.791	.696
False positive rate (at optimal cutoff)						
All students	.337	.318	.200	.340	.221	.400
Black	.311	.326	.187	.331	.207	.403
White	.439	.091	.296	.391	.273	.400
Female	.337	.244	.234	.409	.243	.485
Male	.338	.355	.177	.289	.199	.297

PSSA is Pennsylvania System of School Assessment.
Note: See table 2 of the main report for definitions of outcomes.
Source: Authors' analysis of data from Propel Schools and the Allegheny County Department of Human Services for school years 2014/15–2016/17.

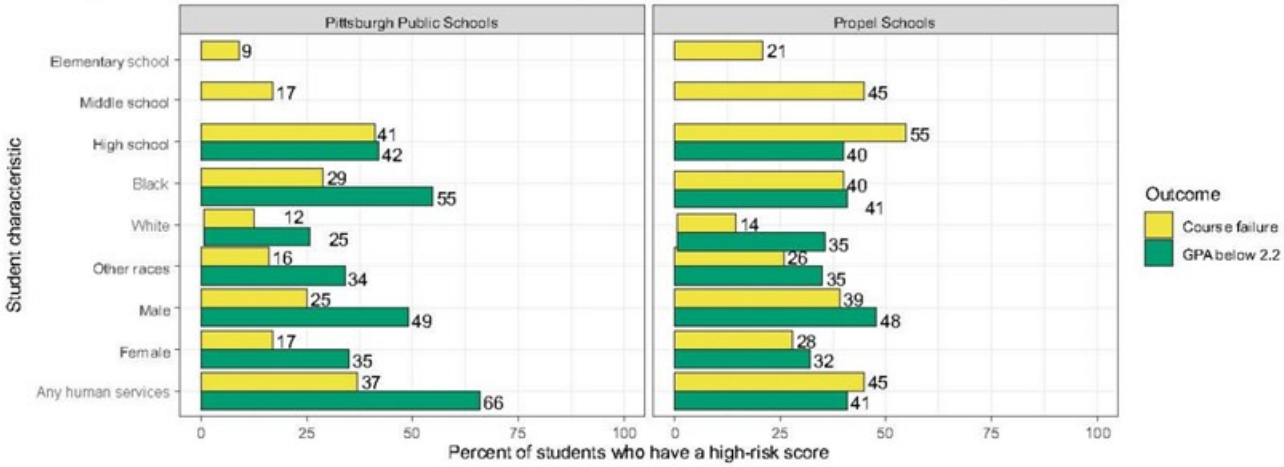
Figures B15–B17 present a breakdown of the proportion of students, by subgroup, who are predicted to have each outcome. Table B4 demonstrates that the models have similar levels of performance with and without Allegheny County Department of Human Services (DHS) predictors.

Figure B15. Proportion of students identified as high risk for absences and suspensions, Pittsburgh Public Schools and Propel Schools samples, by student characteristic, 2014/15–2016/17



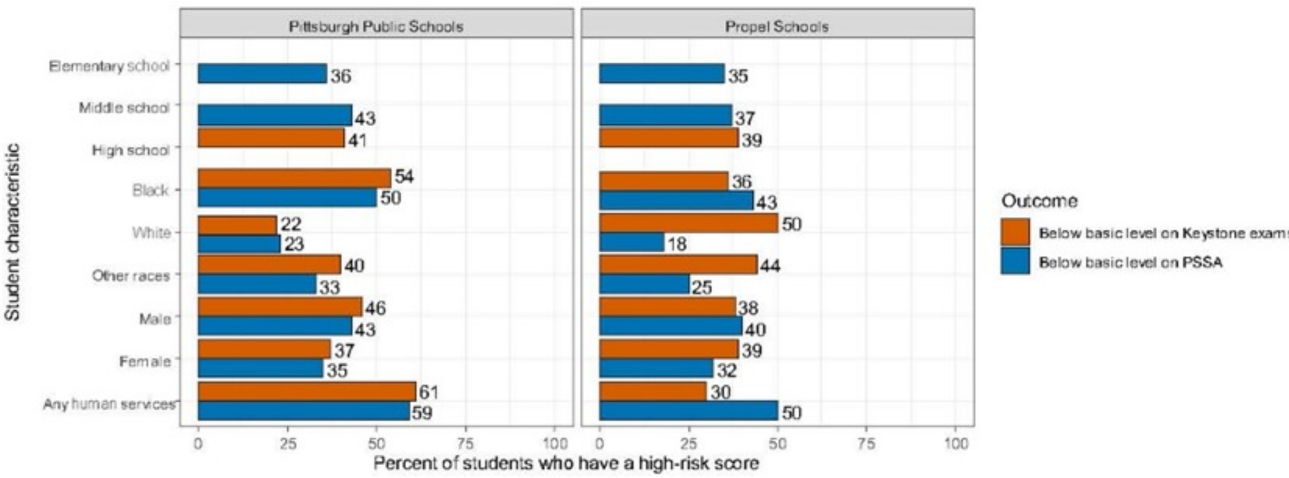
Note: See table 2 of the main report for definitions of outcomes. Bars represent percentage of students in each subgroup defined by student characteristic that have a high-risk score at any point during the 2016/17 school year.
Source: Authors' analysis of data from Pittsburgh Public Schools and Propel Schools for school years 2014/15–2016/17.

Figure B16. Proportion of students identified as high risk for course failure and low grade point average, Pittsburgh Public Schools and Propel Schools samples, by student characteristic, 2014/15–2016/17



Note: See table 2 of the main report for definitions of outcomes. Bars represent the percentage of students in each subgroup defined by student characteristic that have a high-risk score at any point during the 2016/17 school year.
Source: Authors' analysis of data from Pittsburgh Public Schools and Propel Schools for school years 2014/15–2016/17.

Figure B17. Proportion of students identified as high risk for below basic level scores on state tests, Pittsburgh Public Schools and Propel Schools samples, by student characteristic, 2014/15–2016/17



DHS the Allegheny County Department of Human Services. PSSA is Pennsylvania System of School Assessment.
Note: See table 2 of the main report for definitions of outcomes. Bars represent the percentage of students in each subgroup defined by student characteristic that have a high-risk score at any point during the 2016/17 school year.
Source: Authors' analysis of data from Pittsburgh Public Schools and Propel Schools for school years 2014/15–2016/17.

Table B4. Model performance with and without Allegheny County Department of Human Services variables as predictors, Pittsburgh Public Schools and Propel Schools samples, 2014/15–2016/17

	Pittsburgh Public Schools		Propel Schools	
Outcome	AUC with DHS variables as predictors	AUC without DHS variables as predictors	AUC with DHS variables as predictors	AUC without DHS variables as predictors
Chronic absenteeism	.795	.799	.782	.776
Suspensions	.817	.819	.786	.774
Course failure	.901	.905	.838	.842
Low GPA	.917	.916	.810	.816
Score below basic level on PSSA test	.900	.903	.885	.880
Score below basic level on Keystone exam	.839	.840	.676	.667

AUC is area under receiver operating characteristic curve; DHS is the Allegheny County Department of Human Services; PSSA is Pennsylvania System of School Assessment.

Note: See table 2 of the main report for definitions of outcomes. Both models still included in-school and demographic predictors. DHS predictors include data on child welfare involvement, housing and family support services, behavioral health services, and justice system involvement.

Source: Authors' analysis of data from Pittsburgh Public Schools, Propel Schools, and the Allegheny County Department of Human Services for school years 2014/15–2016/17.

Differences in probabilities corresponding to heat map figures

Tables B5–B12 show the differences in probabilities for each of the descriptive analysis relationships shown in the heat maps in the report (figures 1 and 2 in the main report and figures B1–B12).

Table B5. Differences in probability of academic problems for students with prior in-school events in adjacent time periods, Pittsburgh Public Schools sample, 2015/16 and 2016/17

	Outcome in following period												
	Elementary school				Middle school				High school				
Predictor from previous period	Chronic absenteeism	Any suspensions	Course failure	Score below basic level on state tests	Chronic absenteeism	Any suspensions	Course failure	Score below basic level on state tests	Chronic absenteeism	Any suspensions	Course failure	Low grade point average	Score below basic level on state tests
Academic problems													
Number of days absent													
Unexcused	.253	.016	.030	.147	.304	.055	.095	.173	.415	.061	.193	.323	.194
Total	.309	.021	.027	.116	.370	.069	.094	.167	.462	.067	.187	.313	.182
Suspensions													
Number of suspensions	.052	.094	.020	.117	.132	.137	.052	.143	.145	.144	.079	.174	.130
Any suspension	.139	.260	.044	.232	.228	.241	.086	.229	.283	.268	.140	.338	.221
For disruption	.152	.271	.051	.267	.240	.269	.088	.239	.286	.317	.149	.360	.230
For drugs	na	na	.037	na	.479	.024	.165	.418	.371	.077	.198	.295	.175
For fighting	.109	.245	.035	.160	.214	.200	.078	.215	.250	.200	.109	.273	.198
For weapons	.141	.281	.040	.168	.246	.038	.076	.142	.247	.009	.047	.039	.049
For other reasons	.169	.270	.045	.268	.199	.264	.119	.161	.292	.271	.144	.365	.312
Grades and test scores													
Number of courses failed	.081	.029	.074	.155	.167	.079	.170	.130	.267	.091	.289	.446	.205
Any core course failed	.196	.093	.157	.360	.301	.151	.258	.242	.344	.121	.300	.600	.265
Term grade point average below 2.2	na	na	na	na	na	na	na	na	.310	.120	.249	.636	.283
Cumulative grade point average	na	na	na	na	na	na	na	na	–.220	–.084	–.151	–.360	–.257
Score below basic level on state test	.046	.056	.019	.402	.110	.078	.039	.480	.162	.067	.063	.234	.363
Standardized test score	–.108	–.059	–.027	–.494	–.154	–.093	–.049	–.517	–.253	–.093	–.107	–.331	–.480
School mobility													
Withdrew from a school	–.077	–.009	.002	.232	–.024	–.028	.010	.250	–.026	–.022	.008	–.003	.210
Number of schools enrolled	.021	.014	.006	.052	.039	.004	.019	.064	.039	.007	.023	.042	.039

na is not applicable.

Note: See table 2 of the main report for definitions of outcomes. For binary predictors, the values indicate the difference in probability of experiencing the outcome for students with and without the given predictor. For continuous predictors (such as the number of days absent), the values indicate the difference in probability of the outcome for two students who differ by two standard deviations. These results correspond to the heat maps in figures 1 and B1.

Source: Authors' analysis of data from Pittsburgh Public Schools and the Allegheny County Department of Human Services for school years 2015/16 and 2016/17.

Table B6. Differences in probability of academic problems for students with child welfare events, behavioral health, and housing services in adjacent time periods, Pittsburgh Public Schools sample, 2015/16 and 2016/17

	Outcome in following period													
	Elementary school				Middle school				High school					
Predictor from previous period	Chronic absenteeism	Any suspensions	Course failure	Score below basic level on state tests	Chronic absenteeism	Any suspensions	Course failure	Score below basic level on state tests	Chronic absenteeism	Any suspensions	Course failure	Low grade point average	Score below basic level on state tests
Child welfare services and removals													
Removals													
Removal episode occurred	.175	.032	.017	.014	.342	.240	.143	.250	.467	.119	.229	.372	.094
Total number of days in removal	.000	.004	.000	.038	.030	.017	.008	.029	.062	.025	.031	.056	.037
Services started			.					.					.
Any placement	.162	.024	.024	.110	.342	.199	.127	.254	.449	.150	.240	.402	.225
Kinship foster	.207	.032	.014	.014	.314	.140	.102	.318	.431	.148	.148	.302	.207
Nonkinship foster	.080	–.015	.052	.344	.226	.233	.054	.007	.317	.172	.110	.389	.260
Residential placement	–.027	.057	–.018	.129	.538	.288	.252	.331	.471	.162	.363	.514	.194
Services ongoing													
Any placement	–.025	.029	–.002	.256	.114	.067	.021	.189	.227	.090	.113	.199	.195
Kinship foster	–.003	.031	.003	.214	.125	.092	.011	.215	.244	.114	.090	.199	.158
Nonkinship foster	–.094	.028	–.014	.384	.036	.042	–.027	.043	–.023	.002	.082	.100	.230
Residential placement	.291	–.034	–.018	–.271	.246	–.087	.298	.643	.260	.064	.277	.381	.194
Nonplacement services	.030	.010	.006	.047	.074	.032	.022	.054	.093	.038	.057	.090	.053
Services ended													
Any placement	.086	.049	.033	.164	.332	.194	.138	.143	.423	.146	.250	.428	.230
Kinship foster	.108	.058	.042	.100	.322	.125	.136	.143	.418	.197	.252	.429	.139
Nonkinship foster	.007	.024	.034	.284	.337	.186	.068	–.071	.242	.005	.097	.230	.165
Residential placement	.347	.077	–.018	.229	.413	.277	.206	.377	.462	.143	.304	.497	.022
Behavioral health services	.												
Outpatient	.035	.040	.005	.127	.060	.052	.005	.120	.064	.025	.030	.059	.057
Counseling	.013	.016	.000	.047	.041	.023	.013	.050	.014	.003	.014	.027	.029
Inpatient	.011	.007	.000	.000	.020	.014	.000	.032	.019	.004	.010	.022	.018

Predictor from previous period	Outcome in following period: Elementary school				Middle school				High school				
	Chronic absenteeism	Any suspensions	Course failure	Score below basic level on state tests	Chronic absenteeism	Any suspensions	Course failure	Score below basic level on state tests	Chronic absenteeism	Any suspensions	Course failure	Low grade point average	Score below basic level on state tests
Housing and family support services													
Any homeless service received	.104	.014	.009	.089	.164	.056	.013	.152	.206	.062	.045	.186	.102
Any homeless service started	.167	.038	.006	.144	.269	.117	.032	.306	.310	.109	.044	.192	.062
Supportive services	.128	.009	.024	–.036	.231	.165	.021	.155	.209	.098	.021	.131	.061
Bridge and transitional housing services	.375	–.034	–.018	na	.246	–.087	–.046	na	.284	–.078	–.105	.175	na
Permanent housing	.082	.007	.007	–.121	–.028	–.041	–.015	.024	.164	.025	.007	.101	.142
Emergency shelter for 7 days or more	.268	.000	.018	.139	.399	.076	–.014	.268	.360	–.050	.067	.160	.260
Rental assistance and prevention	.013	.006	.001	.018	.035	.010	.003	.026	.026	.009	.007	.016	.007
Head Start	–.002	–.001	.003	.014	.003	.010	.007	na	.012	.007	.002	.007	.014
Other housing supports (public benefits)[a]													
Low-income public housing	.011	.007	.009	.084	.000	.018	.009	.062	.069	.024	.031	.075	.076
Start of low-income public housing	.092	.005	.016	.213	.159	.063	.020	.199	.184	.111	.061	.142	.048
Section 8 voucher	.033	.020	.005	.089	.061	.037	.013	.107	.082	.044	.031	.105	.081
Start of Section 8 voucher	.084	.034	.005	.133	.141	.070	.030	.189	.149	.055	.066	.182	.038

na is not applicable.

Note: See table 2 of the main report for definitions of outcomes. For binary predictors, the values indicate the difference in probability of experiencing the outcome for students with and without the given predictor. For continuous predictors (such as total number of days in removal), the values indicate the difference in probability of the outcome for two students who differ by two standard deviations. These results correspond to the heat maps in figures 2, B2, and B4.

a. Services not provided by the Allegheny County Department of Human Services.

Source: Authors' analysis of data from Pittsburgh Public Schools and the Allegheny County Department of Human Services for school years 2015/16 and 2016/17.

Table B7. Differences in probability of academic problems for students with justice system involvement in adjacent time periods, Pittsburgh Public Schools sample, 2015/16 and 2016/17

	Outcome in following period												
	Elementary school				Middle school				High school				
Predictor from previous period	Chronic absenteeism	Any suspensions	Course failure	Score below basic level on state tests	Chronic absenteeism	Any suspensions	Course failure	Score below basic level on state tests	Chronic absenteeism	Any suspensions	Course failure	Low grade point average	Score below basic level on state tests
Juvenile justice													
Start of case	.398	.238	.133	.364	.317	.143	.145	.276	.283	.114	.122	.273	.171
Case adjudicated delinquent, assigned placement	.291	.466	.232	na	.517	.271	.231	.322	.364	.204	.178	.391	.285
Case adjudicated delinquent, assigned day treatment	.625	.632	.107	.729	.524	.305	.184	.444	.382	.188	.169	.337	.337
Case adjudicated delinquent, assigned probation	.506	.323	.060	.443	.424	.197	.135	.310	.311	.132	.113	.270	.287
Consent decree, assigned nonplacement	.412	.311	.108	.229	.396	.242	.114	.306	.306	.140	.157	.328	.314
Consent decree, assigned day treatment	na	na	na	na	.590	.444	.258	.371	.406	.195	.220	.400	.241
Active case	.359	.290	.078	.352	.363	.177	.114	.283	.282	.124	.130	.295	.269
Jail and adult probation													
Time spent in county jail	.010	.007	.006	–.005	–.005	–.002	.001	–.004	.016	.005	.014	.016	–.004
Jail booking occurred	.007	.007	.006	na	–.005	–.002	.000	–.007	.010	.002	.009	.012	–.002
Adult probation	.006	.000	.000	na	.006	–.001	.003	–.007	.005	.001	.008	.010	.008
Family court													
Active case	.020	.007	.005	.027	.070	.028	.035	.055	.081	.032	.035	.064	.057
Delinquency event	.019	.009	.005	.021	.068	.025	.034	.057	.078	.031	.034	.062	.060
Custodianship event	.008	.000	.001	.372	.022	.016	.008	.007	.024	.007	.010	.019	–.107
Service coordination event	na	na	na	na	na	na	na	.643	na	na	na	na	na

na is not applicable.

Note: See table 2 of the main report for definitions of outcomes. Adjudicated delinquent is analogous to a "guilty" verdict for an adult and describes the sentence; consent decree is the settlement between the court and the juvenile that typically describes any required community service, day treatment or nonplacement services. The values indicate the difference in probability of experiencing the outcome for students with and without the given predictor. These results correspond to the heat map in figure B3.

Source: Authors' analysis of data from Pittsburgh Public Schools and the Allegheny County Department of Human Services for school years 2015/16 and 2016/17.

Table B8. Differences in probability of academic problems in adjacent time periods, Pittsburgh Public Schools sample, by student demographic characteristic, 2015/16 and 2016/17

	Outcome inn following period														
	Elementary school				Middle school				High school					
Predictor from previous period	Chronic absenteeism	Any suspensions	Course failure	Score below basic level on state tests	Chronic absenteeism	Any suspensions	Course failure	Score below basic level on state tests	Chronic absenteeism	Any suspensions	Course failure	Low grade point average	Score below basic level on state tests
Gender and race/ethnicity													
All males	.003	.027	.009	.045	.010	.019	.026	.061	–.035	.023	.043	.124	.069
Black males	.025	.049	.018	.178	.038	.067	.038	.194	.067	.066	.068	.228	.165
Black females	.016	–.004	.000	.082	.027	.039	–.001	.077	.132	.033	.009	.050	.071
American Indian/Alaska Native	.011	–.008	.019	.033	.079	–.027	.074	na	.004	–.008	.051	.136	–.081
Asian	–.106	–.033	–.014	–.085	–.087	–.062	–.042	–.032	–.135	–.051	–.048	–.157	–.007
Black	.032	.036	.014	.207	.052	.084	.029	.211	.151	.075	.058	.208	.194
Hispanic	–.056	–.025	–.009	–.028	–.060	–.043	–.024	–.015	–.004	–.036	.005	.043	.081
White	–.025	–.031	–.014	–.198	–.046	–.077	–.025	–.214	–.141	–.071	–.059	–.206	–.212
Disadvantage													
Economic disadvantage[a]	.088	.024	.012	.158	.128	.062	.033	.181	.183	.058	.070	.194	.165
Homeless	.285	.026	.025	.135	.267	.090	.064	.156	.284	.092	.151	.287	.159
School services eligibility													
Special education	.049	.023	–.001	.347	.096	.049	–.003	.358	.029	.003	–.006	.010	.300
Gifted	–.085	–.023	–.017	–.265	–.107	–.014	–.029	–.286	–.078	–.061	–.072	–.223	–.239
English as a second language services	–.059	–.028	–.006	.280	–.083	–.046	–.041	.302	–.079	–.007	–.012	–.026	.309

na is not applicable.

Note: See table 2 of the main report for definitions of outcomes. The values indicate the difference in probability of experiencing the outcome for students with and without the given predictor. These results correspond to the heat map in figure B5.

a. Based on eligibility for the national school lunch program.

Source: Authors' analysis of data from Pittsburgh Public Schools and the Allegheny County Department of Human Services for school years 2015/16 and 2016/17.

Table B9. Differences in probability of academic problems for students with prior in-school events in adjacent time periods, Propel Schools sample, 2015/16 and 2016/17

	Outcome in following period													
	Elementary school				Middle school				High school					
Predictor from previous period	Chronic absenteeism	Any suspensions	Course failure	Score below basic level on state tests	Chronic absenteeism	Any suspensions	Course failure	Score below basic level on state tests	Chronic absenteeism	Any suspensions	Course failure	Low grade point average	Score below basic level on state tests
Academic problem													
Number of days absent													
Unexcused	.272	.007	.043	.142	.246	.011	.124	.100	.405	.056	.088	.185	.125
Total	.310	.005	.035	.103	.299	.005	.094	.082	.483	.038	.058	.119	.101
Suspensions													
Number of suspensions	.028	.010	.000	.033	.052	.021	.029	.015	.071	.040	.073	.101	na
Any suspension	.067	.073	.004	.192	.120	.076	.114	.031	.125	.106	.192	.284	na
For disruption	na	na	na	na	na	na	.042	na	.024	na	na	na	na
For drugs	na	na	na	na	.040	na	na	na	.010	.051	.060	.068	na
For fighting	.010	.005	.000	.015	.035	.015	.003	.015	.098	.017	.016	.033	na
For other reasons	.020	.011	.001	.027	.018	.013	.040	na	na	.013	.059	.083	na
Grades and test scores													
Number of courses failed	.093	.002	.111	.175	.144	.018	.227	.158	.063	.062	.286	.420	.105
Any core course failed	.141	.003	.174	.307	.152	.018	.250	.195	.058	.060	.268	.472	.169
Term grade point average below 2.2	na	na	na	na	na	na	na	na	.099	.052	.230	.544	.201
Cumulative grade point average	na	na	na	na	na	na	na	na	–.058	–.008	–.026	–.061	–.079
Score below basic level on state test	.037	.009	.037	.393	.076	.008	.086	.441	.076	–.001	.059	.140	.215
Standardized test score	–.074	–.008	–.049	–.479	–.079	–.010	–.098	–.444	–.011	–.011	–.068	–.164	–.313
School mobility													
Withdrew from a school	.119	.000	.005	na	na	.009	.032	na	na	.018	–.015	–.040	na
Number of schools enrolled	.000	–.001	.000	.039	.029	.002	–.008	.029	.062	–.024	–.015	–.019	na

na is not applicable.

Note: See table 2 of the main report for definitions of outcomes. For binary predictors, the values indicate the difference in probability of experiencing the outcome for students with and without the given predictor. For continuous predictors (such as the number of days absent), the values indicate the difference in probability of the outcome for two students who differ by two standard deviations. These results correspond to the heat maps in figures B6 and B8.

Source: Authors' analysis of data from Propel Schools and the Allegheny County Department of Human Services for school years 2015/16 and 2016/17.

Table B10. Differences in probability of academic problems for students with child welfare events, behavioral health, and housing services in adjacent time periods, Propel Schools sample, 2015/16 and 2016/17

	Outcome in following period												
	Elementary school				Middle school				High school				
Predictor from previous period	Chronic absenteeism	Any suspensions	Course failure	Score below basic level on state tests	Chronic absenteeism	Any suspensions	Course failure	Score below basic level on state tests	Chronic absenteeism	Any suspensions	Course failure	Low grade point average	Score below basic level on state tests
Child welfare services and removals													
Removals													
Removal episode occurred	.479	–.006	.113	.306	na	–.016	.406	.204	.145	–.038	.056	.011	–.231
Total number of days in removal	.028	–.001	.004	.014	–.009	.008	.011	.038	.028	–.006	.021	.034	–.004
Services started													
Any placement	.594	–.006	.078	.273	–.131	–.016	.006	.204	.313	–.038	.075	.039	–.231
Kinship foster	.479	–.006	.078	.384	na	–.016	–.094	.204	–.356	–.037	.097	.111	–.231
Nonkinship foster	.879	–.006	–.036	.116	na	–.016	.072	na	–.356	–.037	.041	–.056	–.231
Residential placement	na	na	na	na	–.131	–.016	.406	na	.647	–.037	.541	.611	na
Services ongoing													
Any placement	.379	–.006	.010	–.062	na	.151	.210	.504	na	–.038	.125	.211	.104
Kinship foster	–.121	–.006	.022	.002	na	.234	.343	.504	na	–.038	.041	.111	.104
Nonkinship foster	.879	–.006	–.036	–.284	na	–.016	–.094	na	na	–.037	–.126	–.390	na
Residential placement	na	na	na	na	na	–.016	–.094	na	na	–.037	.275	.611	na
Nonplacement services	.063	–.001	.006	.017	.007	.004	.006	.030	.061	–.004	.031	.044	–.027
Services ended													
Any placement	.279	–.006	.111	.002	–.132	–.016	–.023	.204	.247	–.038	.097	.111	–.231
Kinship foster	.279	–.006	.095	–.117	–.131	–.016	–.094	.204	.647	–.037	.182	.278	na
Nonkinship foster	na	–.006	.146	.717	–.131	–.016	.406	.204	–.356	–.037	.541	.611	–.231
Residential placement	na	–.006	na	na	–.131	–.016	.072	na	.313	–.037	–.063	–.056	na
Behavioral health services													
Outpatient	.036	.014	.004	.036	.028	–.001	.013	.026	.061	–.007	.018	.017	–.021
Counseling	–.013	.003	.001	.012	.029	–.002	.012	.021	.017	–.002	–.019	–.041	.041
Inpatient	.009	.014	.005	.031	.012	–.001	.004	.007	.052	–.003	.013	.024	.015

Predictor from previous period	Outcome in following period: Elementary school: Chronic absenteeism	Elementary school: Any suspensions	Elementary school: Course failure	Elementary school: Score below basic level on state tests	Middle school: Chronic absenteeism	Middle school: Any suspensions	Middle school: Course failure	Middle school: Score below basic level on state tests	High school: Chronic absenteeism	High school: Any suspensions	High school: Course failure	High school: Low grade point average	High school: Score below basic level on state tests
Housing and family support services													
Any homeless service received	.118	.003	.007	.075	.183	.031	.052	.057	.397	.000	.003	.093	na
Any homeless service started	.130	.010	.041	.250	.120	.049	.056	.072	na	.029	–.005	.078	na
Supportive services	.237	.012	.007	–.006	.871	.047	.156	.061	.647	–.038	.035	.326	na
Bridge and transitional housing services	–.121	–.006	na	na	–.131	na	na	na	–.356	na	na	na	na
Permanent housing	–.122	–.006	–.008	–.284	na	–.016	–.094	–.297	.646	–.037	–.014	–.190	na
Emergency shelter for 7 days or more	.380	–.006	–.011	–.284	.370	–.016	.156	.204	na	na	na	na	na
Rental assistance and prevention	.018	.002	.003	.002	.016	–.002	.001	–.012	–.044	.004	.000	.048	na
Head Start	.007	–.002	–.001	na	–.008	.005	–.004	na	–.004	–.006	–.009	.012	na
Other housing supports (public benefits)[a]													
Low-income public housing	.018	–.001	.011	.031	.064	.006	.020	.044	–.009	.011	–.002	.013	.033
Start of low-income public housing	.056	–.006	.028	–.117	.183	.082	.009	.370	.219	–.038	.021	–.057	na
Section 8 voucher	.013	.004	.016	.122	.035	.014	.021	.080	.039	.007	.055	.081	.055
Start of Section 8 voucher	.047	–.002	.037	–.159	–.001	.031	.037	–.074	.092	.052	.006	.056	na

na is not applicable.

Note: See table 2 of the main report for definitions of outcomes. For binary predictors, the values indicate the difference in probability of experiencing the outcome for students with and without the given predictor. For continuous predictors (such as total number of days in removal), the values indicate the difference in probability of the outcome for two students who differ by two standard deviations. These results correspond to the heat maps in figures B7, B9, and B11.

a. Not provided by the Allegheny County Department of Human Services.

Source: Authors' analysis of data from Propel Schools and the Allegheny County Department of Human Services for school years 2015/16 and 2016/17.

Table B11. Differences in probability of academic problems for students with justice system involvement in adjacent time periods, Propel Schools sample, 2015/16 and 2016/17

	Outcome in following period												
	Elementary school				Middle school				High school				
Predictor from previous period	Chronic absenteeism	Any suspensions	Course failure	Score below basic level on state tests	Chronic absenteeism	Any suspensions	Course failure	Score below basic level on state tests	Chronic absenteeism	Any suspensions	Course failure	Low grade point average	Score below basic level on state tests
Juvenile justice													
Start of case	na	na	na	.383	.069	–.016	.081	–.197	.148	.040	.099	.150	–.030
Case adjudicated delinquent, assigned placement	na	na	na	na	.869	–.016	.406	–.297	na	–.037	na	na	na
Case adjudicated delinquent, assigned day treatment	na	na	na	na	na	na	na	na	.646	–.037	–.126	–.390	na
Case adjudicated delinquent, assigned probation	na	na	na	na	.369	–.016	.406	–.297	.145	.296	.375	.611	na
Consent decree, assigned nonplacement	.879	–.006	–.036	na	.370	–.016	.195	.260	.145	.046	.003	.312	–.121
Consent decree, assigned day treatment	na	na	na	na	na	na	na	na	na	–.037	–.126	–.390	na
Active case	.379	–.006	–.036	.383	.256	.010	.190	.083	.092	.066	.137	.263	–.049
Jail and adult probation													
Time spent in county jail	na	na	na	na	na	–.016	.906	na	na	na	na	na	na
Jail booking occurred	na	na	na	na	na	.000	.026	na	na	na	na	na	na
Adult probation	–.004	.000	–.001	na	–.007	.000	–.003	na	na	na	na	na	na
Family court													
Active case	.024	.008	.009	.043	.052	–.002	.028	–.005	.012	.004	.036	.346	–.030
Delinquency event	.004	.216	.137	.028	.315	–.016	.247	–.003	–.023	–.038	.142	.362	.271
Custodianship event	.479	–.006	.099	.431	.369	–.016	.656	–.297	.313	.463	.375	.111	–.231
Service coordination event	na	na	na	na	–.131	na	na	na	na	–.037	.875	.611	na

na is not applicable.

Note: See table 2 of the main report for definitions of outcomes. Adjudicated delinquent is analogous to a "guilty" verdict for an adult and describes the sentence; consent decree is the settlement between the court and the juvenile that typically describes any required community service, day treatment or nonplacement services. The values indicate the difference in probability of experiencing the outcome for students with and without the given predictor. These results correspond to the heat map in figure B10.

Source: Authors' analysis of data from Propel Schools and the Allegheny County Department of Human Services for school years 2015/16 and 2016/17.

Table B12. Differences in probability of academic problems in adjacent time periods, Propel Schools sample, by student demographic characteristic, 2015/16 and 2016/17

	Outcome in following period												
	Elementary school				Middle school				High school				
Predictor from previous period	Chronic absenteeism	Any suspensions	Course failure	Score below basic level on state tests	Chronic absenteeism	Any suspensions	Course failure	Score below basic level on state tests	Chronic absenteeism	Any suspensions	Course failure	Low grade point average	Score below basic level on state tests
Gender and race/ethnicity													
All males	.004	.003	.012	.024	–.036	–.004	.048	.050	–.090	.014	.059	.136	.064
Black males	.017	.007	.025	.101	–.012	–.001	.063	.144	–.071	.022	.050	.141	.109
Black females	.029	.001	.006	.079	.030	.011	–.021	.046	.050	–.003	–.042	–.092	–.056
American Indian/Alaska Native	na	na	–.035	na	na	na	.072	na	na	na	na	na	na
Asian	na	–.006	–.021	–.187	na	na	–.080	na	na	na	–.099	na	na
Black	.042	.007	.030	.177	.021	.011	.040	.185	–.020	.025	.006	.053	.075
Hispanic	.024	–.006	–.001	–.099	–.004	–.016	.026	.140	na	–.037	–.023	–.097	na
White	–.059	–.006	–.028	–.185	–.032	–.013	–.057	–.184	.003	–.016	.006	–.053	–.079
Disadvantage													
Economic disadvantage[a]	.095	.005	.032	.170	.085	.015	.057	.170	.172	.008	.041	.111	.024
Homeless	.045	.000	.000	–.001	.018	.003	.001	.002	.127	–.007	.021	.039	.033
School services eligibility													
English as a second language services	na	na	.024	na	na	na	na	na	na	na	–.098	na	na

na is not applicable.

Note: See table 2 of the main report for definitions of outcomes. The values indicate the difference in probability of experiencing the outcome for students with and without the given predictor. These results correspond with the heat map in figure B12.

a. Based on eligibility for the national school lunch program.

Source: Authors' analysis using data from Propel Schools and the Allegheny County Department of Human Services from the 2015/16 and 2016/17 school years.