# Expanding High-Dosage Tutoring

## Evidence from Saga's Live-Online Math Model

JUNE 2026

*Cynthia Miller*
*Claudia Solís-Román*
*Jean Grossman*

**Data from the Nation's Report Card for 2024 show that fewer than one-third of U.S. eighth-graders currently perform at the "proficient" level or higher in math.** Students' math performance fell dramatically after 2019, and scores have not yet recovered to their prepandemic levels.[1] Given the cumulative nature of math, poor performance in the early years can mean that students fall even further behind by high school. In 2024, just 22 percent of high school seniors were deemed proficient in math.[2] Scores tend to be lower for students from families with low incomes and for students attending economically disadvantaged schools, leading to achievement gaps that grow over time.[3]

Researchers and educators point to increased learning time as an important strategy to address low levels of math and reading proficiency, and there is encouraging evidence in support of its effects. A recent meta-analysis of nearly 100 tutoring programs found that tutoring led to large increases in reading and math performance, equivalent to gains of 3 to 15 months of additional learning.[4] A central finding from this research is that dosage, or the number of hours students receive tutoring, is a major determinant of effects.[5]

mdrc
BUILDING KNOWLEDGE TO IMPROVE SOCIAL POLICY

This brief examines the effects of Saga Education's *live-online* Algebra I tutoring model. The live-online model is the latest modification to Saga's original tutoring model, in which students are given personalized math instruction every day in small groups. In the model being studied here, two students alternate every other day between working directly with a tutor virtually and working independently on a computer-assisted learning platform. Enabling tutors to connect with their students virtually loosens the requirement that tutors need to be geographically close to schools, which could enlarge the pool of available tutors. Thus, the live-online model could be implemented in areas with a shortage of tutors. However, not all students may learn as well with a virtual tutor as with a live tutor. This study examines the effects of the live-online model and whether those effects vary across different types of students.

***Given the sequential nature of math, Algebra I stands as an important gatekeeper to success, not only because of the critical thinking it develops, but also because it is a prerequisite for higher-level math courses typically required for graduation and college enrollment.***

The new model was implemented in eight high schools in New York City between the 2021-2022 and 2023-2024 school years. The model was evaluated using quasi-experimental matching methods, in which ninth-grade students who participated in Saga were compared with a carefully matched group of similar students from other, similar New York City high schools who did not receive Saga tutoring. The students were then tracked through the end of the year to assess their math performance.

The findings indicate that the live-online model led to meaningful gains in math achievement. Students who received tutoring using the model were more likely to pass the standardized state Algebra I Regents Exam by the end of ninth grade (39 percent compared with 31 percent for the comparison group), they had higher Regents Exam scores, and they earned higher math grade point averages (GPAs)—all indicators of progress toward graduation. Given the sequential nature of math, Algebra I stands as an important gatekeeper to success, not only because of the critical thinking it develops, but also because it is a prerequisite for higher-level math courses typically required for graduation and college enrollment. In fact, failing Algebra I is often one precursor to dropping out.[6]

The average effects of this model are somewhat smaller than those found in earlier evaluations of Saga's original and in-person blended models (as described in the following section). One reason for this difference is that this model had weaker effects for boys and somewhat weaker effects for students who entered high school with very low math levels. These findings suggest that the mode of delivery may matter for certain types of students. Some students may benefit more from the structure and relationship building that come with in-person tutoring. Or perhaps the quality of tutoring is lower when delivered virtually.

Overall, the findings are encouraging and are consistent with the growing research showing that high-dosage tutoring models (those offering small-group tutoring at least three times a week) work. The live-online model represents a powerful way to make this type of tutoring available to more students by increasing the potential supply of tutors, but further research is needed into how students interact with and benefit from virtual tutoring. The findings can inform the efforts of

schools and districts to address declining math achievement and put students who are behind in math back on a path toward graduation.

## The Saga Model

Saga Education is a national nonprofit organization that works with public school districts to provide high-dosage math tutoring. Although it offers tutoring programs for middle school students, its traditional model focuses on ninth-and tenth-graders taking Algebra I, and it targets high schools in which a majority of students are from families with low incomes. In the **original model**, groups of two students are provided with 50 minutes of in-school tutoring per day with the same tutor, for one school year. Tutoring is delivered in person for the entire class period and replaces an elective course (for example, art or computer science) or occurs in a second hour of math. Tutors, most of whom are recent college graduates, are paid by Saga and screened for math skills and teaching ability.

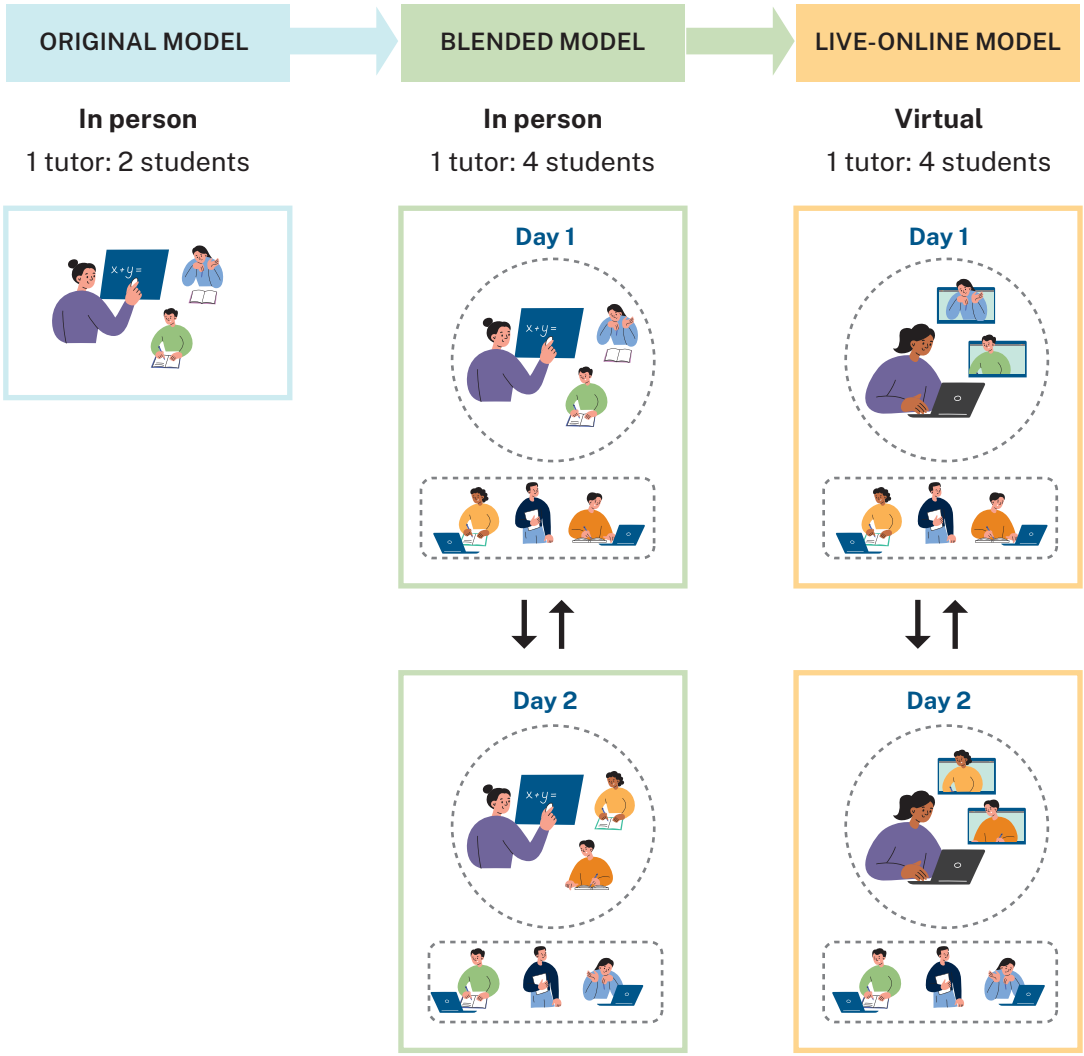
Students are typically selected for tutoring based on their eighth-grade math test scores, current math enrollment (typically Algebra I), and schedule availability (since tutoring makes up one class period). Some schools also make it a priority to offer tutoring to students with Individualized Education Program status and English Language Learner status.[7] Experimental evaluations of the original model found that it led to large gains in math achievement, equivalent to an additional year of learning.[8]

Saga later developed a lower-cost version of the original program, referred to as the **blended model**. In this model, a group of four students was divided into two pairs, and each pair alternated between days working with a tutor in person and days working individually in the tutoring room on a computer-assisted-learning platform, overseen by a staff member. An experimental evaluation of this program in schools in Chicago and New York City found that, despite being 30 percent less expensive per student, it had similarly large effects on math achievement.[9]

During and since the pandemic, Saga has tested an adaptation of the blended model: the **live-online model**, the focus of this study. In this model, students continue to alternate between days of tutoring and independent work in the tutoring room, but the tutors work with the pairs remotely. Tutors and students work together in Saga's online tutoring platform, called Saga Connect. This model has been implemented in eight schools in New York City since 2021. Figure 1 summarizes these three models.

*Experimental evaluations of the original model found that it led to large gains in math achievement, equivalent to an additional year of learning. An experimental evaluation of the blended model in schools in Chicago and New York City found that, despite being 30 percent less expensive per student, it had similarly large effects on math achievement. During and since the pandemic, Saga has tested the live-online model, the focus of this study.*

**Figure 1. Saga's Tutoring Models**




## Evaluation Questions and Methods

The evaluation was designed to address two main research questions: What are the effects of receiving the live-online program on math achievement at the end of the year? And do effects differ across student groups? Students in New York City public high schools do not take standardized state tests, unlike high school students elsewhere. Instead, to graduate, they are required to pass five Regents Exams (covering math, English, Science, and Social Studies) at any time during high school. The most commonly taken math Regents Exam is Algebra I, which can be taken in eighth, ninth, or tenth grade, most typically in ninth grade. This study assesses the effects of the live-online model on math achievement using students' Algebra I Regents Exam scores at the end of ninth grade, along with their math GPAs.

The study examines the outcomes of students in seven of the eight high schools in New York City that implemented Saga's live-online model at some point during the 2021-2022 through 2023-2024 school years. (Student data were not available for one high school.) Different sets of schools participated in each school year, with at least four schools participating each year. Students selected to participate in Saga at those schools met the general eligibility criteria outlined earlier, that is, enrollment in a relevant math class (Algebra I), low math test scores, and a schedule that allowed tutoring to fit into their day. Students participating in the program represent about half of the entering ninth-grade classes at these schools, although there is variation across high schools.

Evaluating the program retrospectively, the study used propensity score matching to investigate the effects of Saga's live-online model on math achievement. This method creates a comparison group for the Saga students by statistically matching each Saga student to a similar student from a similar high school in New York City that did not participate in Saga. The latter group consisted of students attending Title 1 high schools in New York City: those with at least 40 percent of their students living in families with incomes below the federal poverty level. The city's nine specialized high schools, which require testing for admission, were not included in the pool. The analysis then estimated each effect by modeling students' outcomes as a function of whether they participated in Saga, along with all the student- and high school–level variables that were used to match students and create the comparison group.

The starting sample consisted of all students offered the Saga program at the participating high schools. Several additional criteria were then applied to obtain the sample used for analysis. The students had to be present in the New York City Public Schools data for eighth and ninth grade, for example, and could not be included in the study if they took Algebra I before ninth grade. About 2,200 students met these two criteria. In addition, because math test scores are a variable used in the matching process, the study sample was further restricted to students with at least one math test score in the data from sixth, seventh, or eighth grade. Test scores from years earlier than eighth grade were used for the ninth-graders who participated in Saga during the 2020-2021 school year because the COVID-19 pandemic led the school system to suspend testing for their eighth-grade year. In fact, testing did not resume fully for two or three years after the pandemic. Test score data were available for 1,691 students.

Students were matched on their individual characteristics and the characteristics of the schools they attended. Individual characteristics included demographics and measures of past performance — race/ethnicity, gender, past attendance, and past math test scores — helping to ensure that the selected comparison students provide a valid picture of what might have happened with Saga students' math outcomes in the absence of the program. School characteristics included factors such as the racial and ethnic composition of the student body, the percentage of students receiving free or reduced-price lunches, and the average math test scores from previous years, to ensure the comparison group members were enrolled in similar schools that taught math similarly well. Matching was conducted within each school year by selecting a student who was the closest match to each Saga student across all matching characteristics. Close matches were obtained for 87 percent of the Saga students, for a final analysis sample of 1,475 Saga students and 1,475 comparison stu-

dents.[10] After matching, the Saga and comparison groups were assessed on their similarity across a range of student and school characteristics. Differences for all the student characteristics and all but three of the high school characteristics were within the range of what is considered acceptable for the analysis by the What Works Clearinghouse.[11]

Propensity score matching creates a comparison group that is similar to the program group with respect to a large number of observable characteristics. But unlike a randomized controlled trial, it cannot ensure that the groups are matched on unobservable characteristics, such as motivation or level of family support. Nonetheless, research suggests that matching can serve as a valid alternative to random assignment, particularly when individuals are matched on preintervention values of the outcome in question, such as test scores.[12]

Data for the evaluation come from two primary sources. The first source is the Research Alliance for New York City Schools, which provided student-level administrative records from New York City Public Schools.[13] These records cover student- and school-level information on background characteristics, school enrollment, special services, attendance, achievement, and progress toward graduation. Records are available for all years in which a student was enrolled in the New York City public school system, which allows the analysis to match students on prior academic performance. This data source provides variables used for matching, for covariates, and for assessing academic outcomes. The second data source is the Saga program itself. For all students offered Saga, these data include a New York City student ID (necessary for linking students to the administrative records data), hours of Saga tutoring offered, and hours of Saga tutoring received.

## The Students and Schools

The sample of Saga students used for the evaluation comes from seven New York City high schools that implemented the program at some point between the 2021-2022 and 2023-2024 school years. The schools were large, with an average of over 1,600 students each. Most students in these schools were Black or Latino and were eligible for free or reduced-priced lunches, and the students in the sample reflect these school demographics: The majority of students in the sample were Black or Latino (83 percent), and 9 in 10 were eligible for free or reduced-price lunch. Sample students' average attendance rate in eighth grade was 87 percent, or below the benchmark for being considered "chronically absent" in New York City: Students who attend less than 90 percent of the time are considered chronically absent. Finally, reflecting eligibility criteria for Saga, sample students' average scores on math standardized tests in grades six, seven, or eight were below the New York City average and at about the thirty-fourth percentile of all scores.

## Saga Participation

All students in the Saga group received at least one day of the program. They could have received well over 100 hours of personalized instruction over the course of the entire school year, combining

time with a virtual tutor and independent work on the computer-assisted platform. For a variety of reasons, including changes in student schedules, students received less tutoring than that. On average, students received about 69 hours of tutoring over the course of the year.

There was notable variation around this average. Dividing the sample into four, for example, shows that the bottom 25 percent received less than 33 hours of tutoring, the next 25 percent received 33 to 79 hours of tutoring, the next 25 percent received 80 to 103 hours of tutoring, and the top 25 percent received 104 to 130 hours of tutoring. Thus, a sizable percentage of students received 80 or more hours of tutoring over the year. Most of the students in the bottom 25 percent received less than 10 hours of tutoring. It is not clear why some students received so few hours. However, schedule shifts at the beginning and middle of the school year are common, so tutoring may no longer have fit into those students' school day.

All students who received any Saga tutoring are included in the estimates of program effects, even if they received very few hours. The reason for including them is that unobservable factors may have influenced the amount of tutoring students ultimately received. Students who were more motivated, for example, might have been inclined to attend tutoring more often. Dropping students who received few hours from the sample, then, could leave a sample of Saga students who were relatively highly motivated. The matching method cannot match on motivation, and the selected comparison group will include both less and more motivated students. It is therefore preferable to use the full sample of Saga students for the analysis.

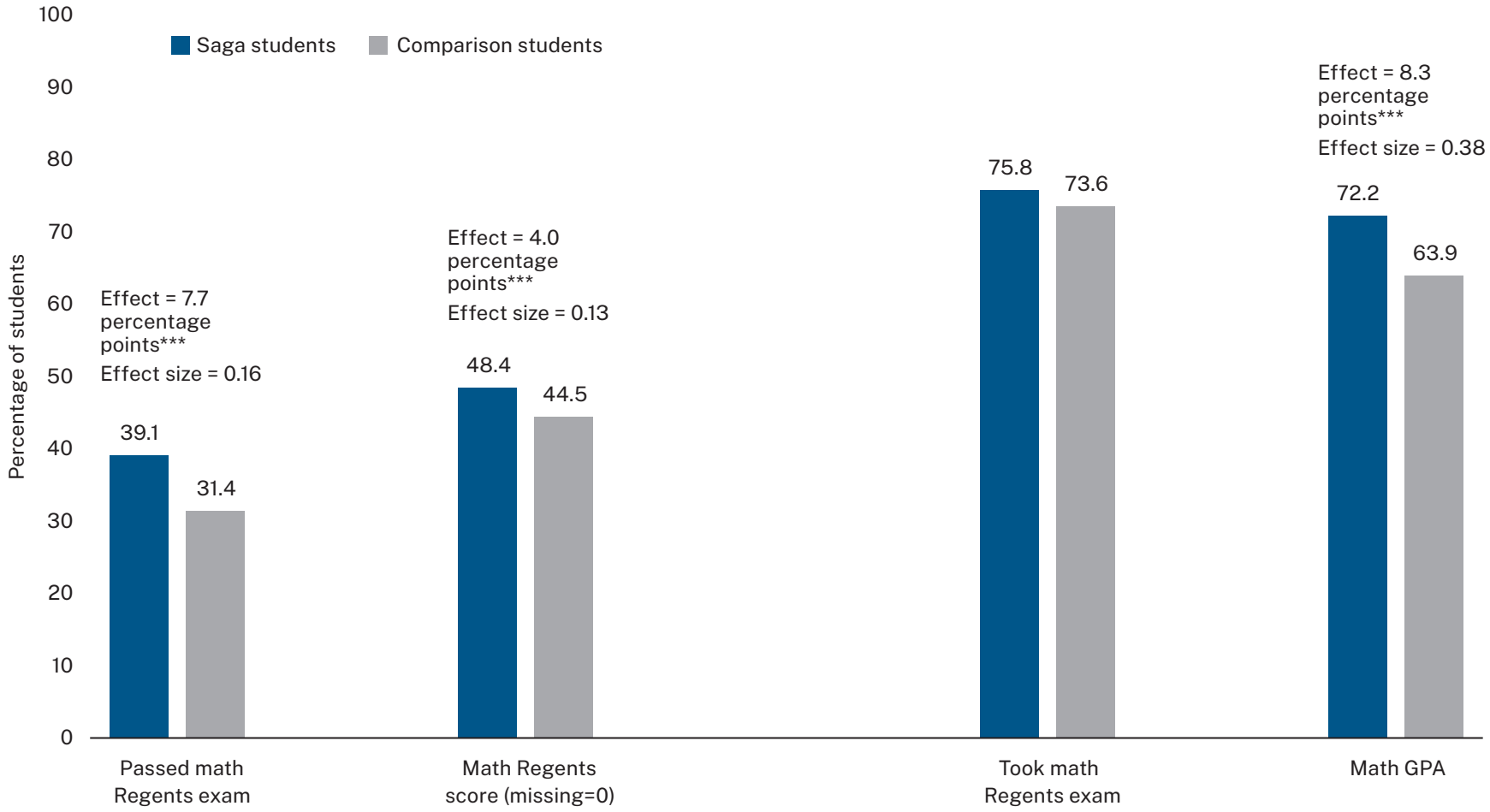
# Effects

## Effects Among the Full Sample

The primary measures of math achievement are Algebra I Regents Exam scores as of the end of ninth grade and pass rates (or achieving a score on that exam of 65 or higher) at the end of ninth grade. As noted, students are required to pass five Regents Exams to graduate from high school. Passing at least one Regents Exam by the end of ninth grade is considered an indicator of being on track for an on-time graduation.[14]

Figure 2 presents the estimated effects. Differences between the Saga and comparison groups that are statistically significant are denoted with asterisks. A statistically significant difference is one that is unlikely to be observed by chance if a program has no true effect. About 39 percent of Saga students passed the Regents Exam, compared with just over 31 percent of comparison students, a statistically significant effect of 8 percentage points. Note that the pass rate and the average score both include zeros for students who did not take the test, as restricting the analysis only to test takers could yield biased results. To be a valid estimate, the comparison in outcomes must be over the full Saga group and the full comparison groups. About 76 percent of ninth-graders took the test at the end of the year, with no statistically significant difference in test-taking rates between the Saga and comparison students.[15]

**Figure 2. Saga Increased Students' Math Performance**




SOURCES: Student records from New York City Public Schools and Saga Education program data.

NOTES: Math GPA is measured on a 100-point scale. Differences were tested for statistical significance: *** indicates significance at the 1 percent level, ** indicates significance at the 5 percent level, and * indicates significance at the 10 percent level. Differences that are not statistically significant are not reported above the corresponding bars.

Saga students also achieved slightly higher average scores on their Regents Exams (48 versus 45), a difference that is statistically significant. Average scores are relatively low because they include zeros for students who did not take the test. Among those who did take the test, average scores were in the low 60s.

*The effect on this measure of performance suggests that tutoring is not simply "teaching to the test" but increasing general math proficiency.*

Another measure of math performance is end-of-year math GPA. Math GPA was also used as an outcome in the previous studies of Saga's model. It provides an assessment of every sample member's math ability at the end of the year, not just those who took the Regents Exam in ninth grade. Math GPA is measured over all math classes taken that year, though the majority of students took only Algebra I in ninth grade. Saga students ended the year with a math GPA of 72 on a 100-point scale compared with 64 for the comparison group, a statistically significant difference of 8 points. Note that GPA is expected to vary across schools and teachers due to differences in grading practices. It also probably reflects more than just math ability, since teachers may also value effort and attendance. The effect on this measure of performance aligns with findings from the earlier Saga studies and suggests that tutoring is not simply "teaching to the test" but increasing general math proficiency.

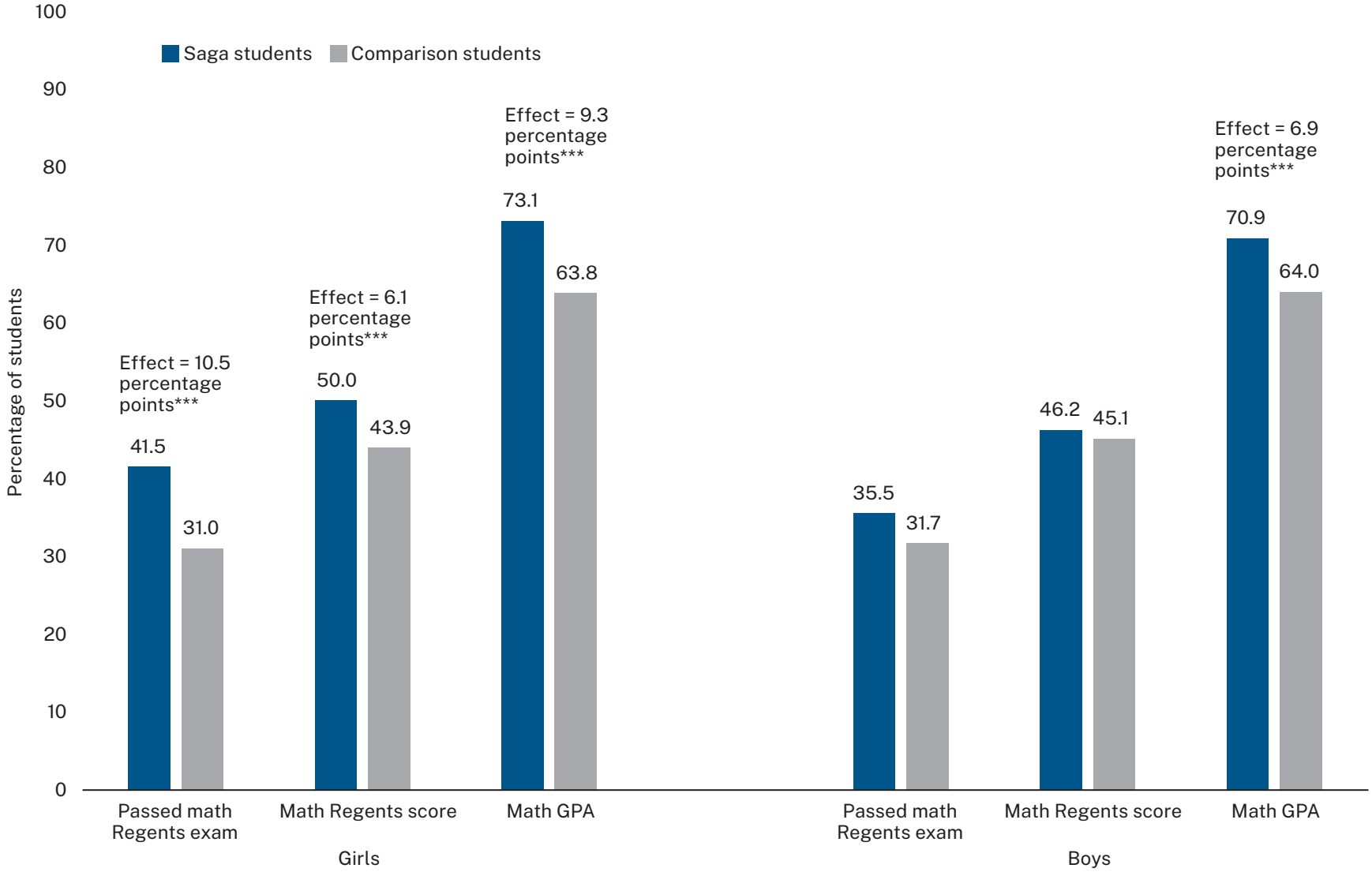
## Effects Among Subgroups

In-person Saga tutoring has been found to be effective for all types of students. The studies of Saga's other models did not find evidence of different effects for boys and girls, for students of different races/ethnicities, or for students with different prior math scores. However, different types of students may engage differently with virtual instruction, so this study still examines effects for subgroups defined in these ways and by family poverty status.

Figures 3 and 4 present the findings for subgroups defined by gender and math performance. Effects did not differ to a statistically significant degree for students of different races/ethnicities or for students with different family poverty statuses. The live-online tutoring model was more effective for girls than for boys. Girls who participated in the program had an 11 percentage point higher Regents pass rate and earned higher Regents scores, on average, than their comparison group counterparts. The effects for boys, in contrast, were much smaller. The increase in pass rates for boys was 4 percentage points (which is not statistically significant), and average test scores were very similar between the Saga and comparison groups. The differences in effects between girls and boys are statistically significant.

Girls and boys took the Regents Exam at roughly the same rate (about 75 percent), although the program increased girls' test-taking rates by a statistically significant 4 percentage points (not shown). Finally, both girls and boys experienced statistically significant increases in math GPA from Saga, but the effects for girls were larger by a statistically significant amount (a 9 percentage point increase for girls compared with a 7 percentage point increase for boys).

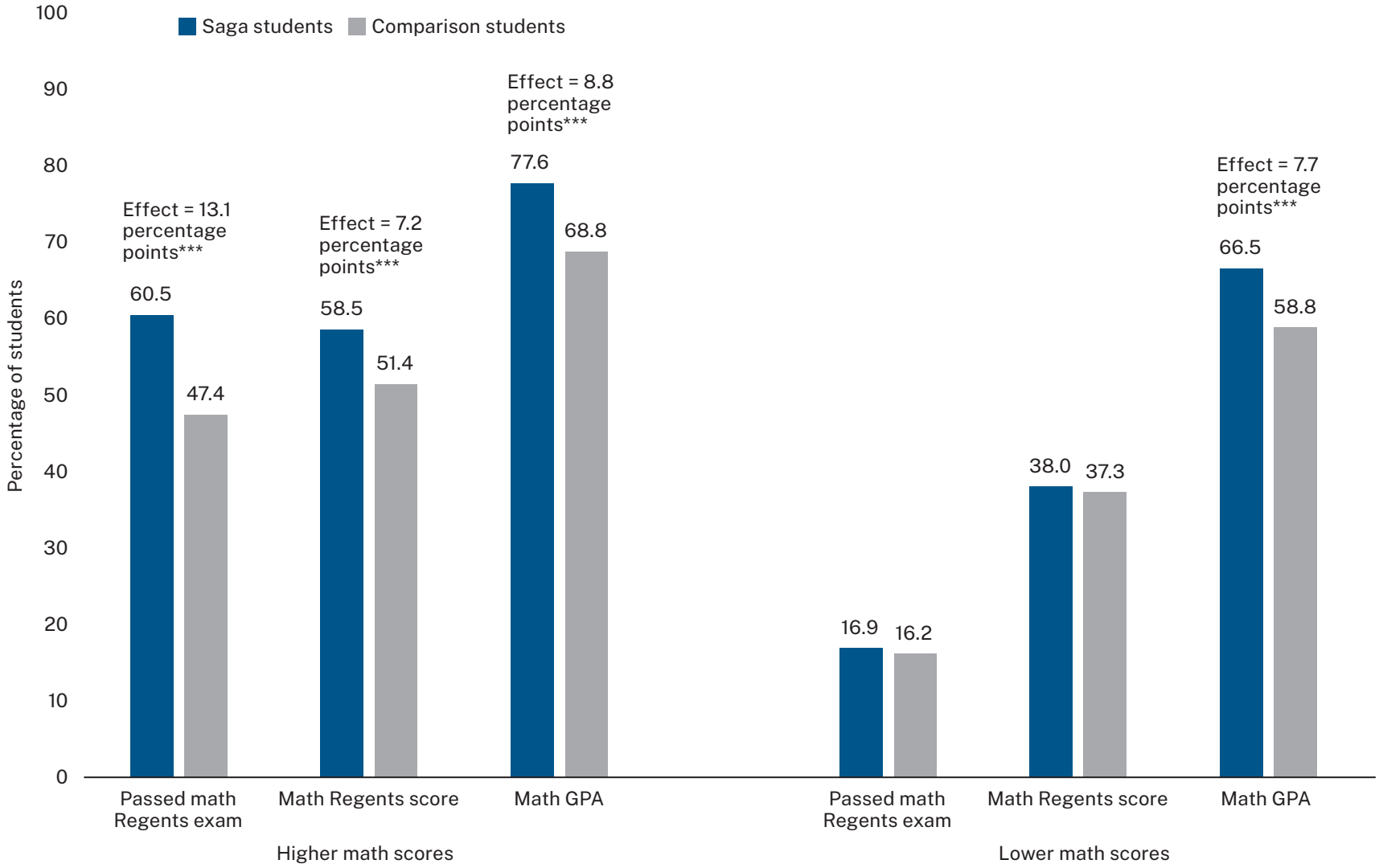
**Figure 3. Saga Had Larger Effects for Girls than for Boys**




SOURCES: Student records from New York City Public Schools and Saga Education program data.

NOTES: Math GPA is measured on a 100-point scale. Differences were tested for statistical significance: *** indicates significance at the 1 percent level, ** indicates significance at the 5 percent level, and * indicates significance at the 10 percent level. All differences in effects between subgroups are statistically significant at the 10 percent level. Differences that are not statistically significant are not reported above the corresponding bars.

**Figure 4. Saga Had Larger Effects for Students with Higher Past Math Scores**




SOURCES: Student records from New York City Public Schools and Saga Education program data.

NOTES: Math GPA is measured on a 100-point scale. Differences were tested for statistical significance: *** indicates significance at the 1 percent level, ** indicates significance at the 5 percent level, and * indicates significance at the 10 percent level. Differences in effects between subgroups are statistically significant at the 10 percent level for the outcomes "Passed Math Regents Exam" and "Math Regents Score." Differences that are not statistically significant are not reported above the corresponding bars.

To examine effects among subgroups defined by past math performance, the study divided students into those who were below and above the median baseline test score. All students offered Saga's program had low scores, but some were further behind than others before ninth grade. Saga led to a large effect for the higher-performing students: About 61 percent of higher-performing students who participated in Saga passed the Regents Exam, compared with 47 percent of their comparison counterparts, an increase in pass rates of 13 percentage points. Effects for the lower-performing subgroup were small and not statistically significant.

Regents pass rates may not be a good marker of math progress for the lower-performing group, however, as it is probably very difficult for these students to achieve a passing score by the end of ninth grade, even if they do become somewhat more proficient. In fact, only about 16 percent of comparison students in the lower-performing subgroup passed the Regents Exam, despite the fact that about 74 percent of them took the exam.

Effects on Regents Exam scores and math GPAs provide a more nuanced picture for this lower-performing group. On the one hand, there is no evidence that Saga led to an increase in Regents Exam scores (with an average score of about 38 percent for the Saga subgroup with lower prior math performance and 37 percent for the comparison students). On the other hand, the lower-performing group did experience an increase in math GPA of about 8 points on a 100-point scale. The effect is similar in size to the effect for the higher-performing group. One benefit of assessing GPA is that it is available for all students, not just those who took the Regents Exam. GPA also captures not just math ability but attendance and effort as well, reflecting both an important marker of progress and a potential predictor of longer-term success. It may also be the case that gains in math ability for these students captured in GPA are not yet enough to increase performance on the formal Regents Exam.

## Conclusions

In an effort to promote the expansion of high-dosage tutoring, Saga Education developed a live-online version of its blended model, in which students alternate between tutoring and self-directed work but meet with tutors virtually rather than in person. The findings presented here show that this virtual model led to an increase in Regents Exam pass rates of 8 percentage points, or an increase of 25 percent. Students receiving Saga also had higher Regents Exam scores and higher math GPAs at the end of ninth grade.

How do the findings compare with earlier evaluations of Saga's programs? Since the central outcome measured varies across studies, it is easiest to consider effects in terms of effect sizes (effects as a proportion of the standard deviation of the outcome among the control or comparison group). Among students who received at least one hour of tutoring, the original model led to an increase in scores on a standardized math test taken by all students of 0.28 standard deviations. The blended model, in which students alternate every other day between being tutored and using a computer-assisted platform, led to an increase in math test scores for those who received at least one hour

of tutoring of 0.23 standard deviations. The live-online model, in turn, led to an increase in Regents pass rates of 0.16 standard deviations and an increase in average scores of 0.13 standard deviations. To put these effect sizes into context — to illustrate how much learning they represent — research comparing students' standardized test scores in the spring of eighth grade with scores in the spring of ninth grade shows that ninth-grade students gain, on average, 0.22 standard deviations in math learning over a year.[16]

However, Regents measures cannot provide an assessment of students who did not take the exam, and New York City high school students do not all take standardized tests. The outcome used in the current study that does assess the math ability of all sample members is math GPA. For this outcome, the live-online model led math grades to increase by 0.38 standard deviations. This shift is equivalent to moving the average math GPA of these Saga students up 14 points of the comparison group's GPA distribution on the New York City 100-point GPA scale. In other words, a student with a GPA at the fiftieth percentile would move up to the sixty-fourth percentile. High school GPA has been found to be an important predictor of success beyond high school.[17]

One reason the live-online model had smaller effects on Regents Exam outcomes is that the program does not appear to have benefited all students equally. Although earlier studies found no differences in effects across types of students, the effects of the live-online model were much smaller for boys (although they were still positive) and also smaller for students who entered ninth grade with very low levels of math proficiency.

Overall, the findings are encouraging, as the model led to a sizable increase in math GPA and in Regents pass rates for Algebra I, which is considered a gateway to higher-level math and to graduation. Although it is not clear yet whether this model is less costly than the blended model, as the cost study will be presented in a later report, the live-online model does address staffing issues that are a major component of expanding tutoring. More research is needed on how tutoring is delivered virtually and how different types of students interact with it.

## Notes and References

1. National Assessment of Educational Progress, "NAEP Report Card: Mathematics" (website: https://www.nationsreportcard.gov/reports/mathematics/2024/g4_8/?grade=8, 2024). Students scoring at or above 299 out of 500 total points are determined to be *NAEP Proficient*, demonstrating solid academic performance and competency over challenging subject matter.
2. National Assessment of Educational Progress, "NAEP Report Card: Grade 12 Mathematics" (website: https://www.nationsreportcard.gov/reports/mathematics/2024/g12/, 2024).
3. National Assessment of Educational Progress, "NAEP Report Card: Grade 12 Mathematics, Performance by Student Groups" (website: https://www.nationsreportcard.gov/reports/mathematics/2024/g12/performance-by-student-group/#group-population-percentages, 2024).
4. Andre Nickow, Philip Oreopoulos, and Vincent Quan, "The Impressive Effects of Tutoring on PreK-12 Learning: A Systematic Review and Meta-Analysis of the Experimental Evidence," NBER Working Paper 27476 (National Bureau of Economic Research, 2020).
5. Monica P. Bhatt, Terence Chau, Barbara Condliffe, Rebecca Davis, Jean Grossman, Jonathan Guryan, Jens Ludwig, Matteo Magnaricotte, Shira Kolnik Mattera, Fatemeh Momeni, Philip Oreopoulos, and Greg Stoddard, *Personalized Learning Initiative Interim Report: Findings from 2023-24* (University of Chicago Education Lab, 2025); Scaling HighDosage Tutoring: Evidence from Saga's LiveOnline Math Model. Carly D. Robinson, Matthew A. Kraft, Susanna Loeb, and Beth E. Schueler, *Accelerating Student Learning with High-Dosage Tutoring* (EdResearch for Recovery Project, 2021).
6. Ron Schachter, "Solving Our Algebra Problem: Getting All Students Through Algebra I to Improve Graduation Rates," *District Administration* 49, 5 (2013): 43–46; Jordan Rickles, Jessica Heppen, Suzanne Taylor, Nicholas Sorensen, and Kirk Walters, "Course Progression for Students Who Fail Algebra I in Ninth Grade" (American Institutes for Research, 2017).
7. Individualized Education Programs are education plans for students with disabilities who are eligible for special education.
8. Jonathan Guryan, Jens Ludwig, Monica P. Bhatt, Philip J. Cook, Jonathan M.V. Davis, Kenneth Dodge, George Farkas, Roland G. Fryer Jr., Susan Mayer, Harold Pollack, Laurence Steinberg, and Greg Stoddard, "Not Too Late: Improving Academic Outcomes Among Adolescents," *American Economic Review* 113, 3 (2023): 738–765.
9. Monica P. Bhatt, Jonathan Guryan, Salman A. Khan, Michael LaForest-Tucker, and Bhavya Mishra, "Can Technology Facilitate Scale? Evidence From a Randomized Evaluation of High Dosage Tutoring," NBER Working Paper 32510 (National Bureau of Economic Research, 2024)
10. The method used was "greedy caliper" matching, which selects a student for the comparison group who is the closest match of a given Saga student in terms of propensity score, within a specified maximum distance (caliper). Matching was conducted with replacement, and 23 students were matched to two different Saga students.
11. What Works Clearinghouse, *What Works Clearinghouse Procedures and Standards Handbook, Version 5.0* (U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance [NCEE], 2022). This report is available on the What Works Clearinghouse website at https://ies.ed.gov/ncee/wwc/Handbooks. The three high school–level characteristics with differences larger than the recommended threshold are the size of the student body, the percentage of students who are Asian, and the percent of students who are White. The main school-level variable that is likely to affect a student's math achievement is test scores for earlier years of ninth-graders, for which the differences between Saga and comparison schools are within the recommended threshold.

12. Kelly Hallberg, Thomas D. Cook, Peter M. Steiner, and M. H. Clark, "Pretest Measures of the Study Outcome and the Elimination of Selection Bias: Evidence from Three Within Study Comparisons," *Prevention Science* 19, 3 (2018): 274–283.
13. Administrative records are data collected in the normal course of administering a program.
14. James J. Kemple, Micha D. Segeritz, and Nickisha Stephenson, "Building On-Track Indicators for High School Graduation and College Readiness: Evidence from New York City," *Journal of Education for Students Placed at Risk* 18, 1 (2013): 7–28.
15. Many of the students who did not take the exam in ninth grade are likely to take it in tenth grade.
16. See Appendix C in Howard S. Bloom, Carolyn J. Hill, Alison Rebeck Black, and Mark W. Lipsey, *Performance Trajectories and Performance Gaps as Achievement Effect-Size Benchmarks for Educational Interventions* (MDRC, 2008).
17. Brian M. Galla, Elizabeth P. Shulman, Benjamin D. Plummer, Margo Gardner, Stephen J. Hutt, J. Parker Goyer, Sidney K. D'Mello, Amy S. Finn, and Angela L. Duckworth. "Why High School Grades Are Better Predictors of On-Time College Graduation Than Are Admissions Test Scores: The Roles of Self-Regulation and Cognitive Ability," *American Educational Research Journal* 56, 6 (2019): 2,077–2,115.

**Acknowledgments**

Many thanks to the Research Alliance for New York City Schools (RANYCS) for providing student records data from Saga Education and the New York City Public Schools, and to the Overdeck Family Foundation for supporting that work. Thanks to Alan Safran, Chris Dupuis, and Cathryn Cook at Saga Education, Jim Kemple at RANYCS, and William Corrin at MDRC for comments and review of an earlier draft. Thanks also to Pei Zhu and Rebecca Unterman for review of the analytic methods. Joshua Malbin edited the document and Carolyn Thomas prepared it for publication.

Funding for this study was provided by the Michael & Susan Dell Foundation.

The following organizations support dissemination of MDRC publications and our efforts to communicate with policymakers, practitioners, and others: Arnold Ventures, Ascendium Education Group, Yield Giving/MacKenzie Scott, and earnings from the MDRC Endowment. Contributors to the MDRC Endowment include Alcoa Foundation, The Ambrose Monell Foundation, Anheuser-Busch Foundation, Bristol-Myers Squibb Foundation, Charles Stewart Mott Foundation, Ford Foundation, The George Gund Foundation, The Grable Foundation, The Lizabeth and Frank Newman Charitable Foundation, The New York Times Company Foundation, Jan Nicholson, Paul H. O'Neill Charitable Foundation, John S. Reed, Sandler Foundation, and The Stupski Family Fund, as well as other individual contributors.

The findings and conclusions in this report do not necessarily represent the official positions or policies of the funders.

**For information about MDRC and copies of our publications, see our website: www.mdrc.org.**



**New York**
200 Vesey Street, 23rd Flr.
New York, NY 10281
Tel: 212 532 3200

**Washington, DC**
750 17th Street, NW
Suite 501
Washington, DC 20006

**Oakland**
475 14th Street, Suite 750
Oakland, CA 94612
Tel: 510 663 6372