PIONEER
EDUCATION

An Analysis of How Massachusetts' "Student Growth" Model Limits Access to Charter Public Schools

By Dr. Cara Stillings Candal with Dr. Bruce Bean & Rebekah Paxton

WHITE PAPER
No. 197 | June 2019

PIONEER INSTITUTE
PUBLIC POLICY RESEARCH

Pioneer's Mission

Pioneer Institute is an independent, non-partisan, privately funded research organization that seeks to improve the quality of life in Massachusetts through civic discourse and intellectually rigorous, data-driven public policy solutions based on free market principles, individual liberty and responsibility, and the ideal of effective, limited and accountable government.

PIONEER EDUCATION

This paper is a publication of Pioneer Education, which seeks to increase the education options available to parents and students, drive system-wide reform, and ensure accountability in public education. The Center's work builds on Pioneer's legacy as a recognized leader in the charter public school movement, and as a champion of greater academic rigor in Massachusetts' elementary and secondary schools. Current initiatives promote choice and competition, school-based management, and enhanced academic performance in public schools.

PIONEER HEALTH

Pioneer Health seeks to refocus the Massachusetts conversation about health care costs away from government-imposed interventions, toward market-based reforms. Current initiatives include driving public discourse on Medicaid; presenting a strong consumer perspective as the state considers a dramatic overhaul of the health care payment process; and supporting thoughtful tort reforms.

PIONEER PUBLIC

Pioneer Public seeks limited, accountable government by promoting competitive delivery of public services, elimination of unnecessary regulation, and a focus on core government functions. Current initiatives promote reform of how the state builds, manages, repairs and finances its transportation assets as well as public employee benefit reform.

PIONEER OPPORTUNITY

Pioneer Opportunity seeks to keep Massachusetts competitive by promoting a healthy business climate, transparent regulation, small business creation in urban areas and sound environmental and development policy. Current initiatives promote market reforms to increase the supply of affordable housing, reduce the cost of doing business, and revitalize urban areas.

Pioneer Institute is a tax-exempt 501(c)3 organization funded through the donations of individuals, foundations and businesses committed to the principles Pioneer espouses. To ensure its independence, Pioneer does not accept government grants.

Table of Contents

Executive Summary

In 2010, the Massachusetts Legislature lifted the cap on the number of charter schools that the Board of Elementary and Secondary Education (BESE) can authorize in low-performing school districts. The "smart cap" requires the Department of Elementary and Secondary Education (DESE) to rank district performance (based on student outcomes) annually. Districts that fall into the lowest 10 percent of those rankings are eligible for an increase in the amount of tuition that they can pay to charter schools. Whereas the law limits district tuition payments to charter schools at 9 percent of net school spending, the smart cap raises that limit to 18 percent of net school spending in the Commonwealth's lowest performing districts.

Until 2015, the Commonwealth ranked district performance solely according to the number of students scoring proficient in core subject areas on the Massachusetts Comprehensive Assessment System (MCAS) tests. In 2015, BESE amended the smart cap regulations and began using an additional measure, the student growth percentile (SGP), as 25 percent of the formula for determining district rankings.

In 2015, BESE amended the smart cap regulations and began using an additional measure, the student growth percentile (SGP), as 25 percent of the formula for determining district rankings.

SGP tracks students' progress by comparing changes in an individual's MCAS score to those of students who scored similarly in prior years. The Commonwealth currently uses SGP for two main purposes: 1) as a factor in its overall system for holding districts accountable for student outcomes; and 2) to determine which school districts are eligible for an increase in the net school spending (NSS) cap on charter schools.

When used as one measure in a more holistic approach to assessing school outcomes, SGP can be informative. In districts that struggle to bring a majority of students to proficiency, it can provide information about whether teachers, schools, and districts are moving some students along the ladder of proficiency. However, researchers disagree as to the accuracy of the information that SGP scores provide. Like any measure of student achievement, SGP is subject to a degree of statistical error. Compared to measuring proficiency alone, SGP suffers from a large degree of error. This is leading a growing number of researchers to suggest that SGP may not be an appropriate metric for high-stakes policy decisions.

The stakes associated with using SGP as one part of the formula for determining districts that are eligible for an increase in the charter cap are high: many of the communities in which overall student proficiency is low also have long waiting lists for charter public schools. Until 2015, when proficiency alone was the measure that the Commonwealth used to determine the lowest performing districts, districts with low overall proficiency and longer waiting lists were more likely to become eligible for an increase in the NSS cap. This has changed since SGP was incorporated into the formula for determining the lowest performing districts.

The statistical error associated with SGP has a different impact on small (lower enrollment) and large (higher enrollment) districts. When a district enrolls higher numbers of students, its SGP is more likely to cluster around what researchers call a "typical mean," even if the district has individual schools with very low or very high SGPs. When a district enrolls fewer students, its SGP is more likely to lie outside of the typical mean. Smaller districts don't cluster around the typical mean in part because low enrollment makes their overall SGP rating more volatile.

Because low-enrollment districts have more volatile SGP scores, they tend to move in and out of the bottom ten percent at higher rates, sometimes displacing larger districts with SGP scores that cluster around a mean, even when those larger districts have low overall student proficiency. This means that some large districts where demand for charter schools is high become ineligible for an increase in the NSS cap, even when large numbers of students are still not proficient in core subjects. Conversely, the smaller districts that enter the bottom 10 percent due to volatile SGP scores are more likely to be districts where demand for charter schools is low. They may also have higher overall proficiency rates than the larger districts they displaced in the bottom 10 percent.

Researchers disagree as to the accuracy of the information that SGP scores provide. Like any measure of student achievement, SGP is subject to a degree of statistical error.

Using SGP as a factor in determining district eligibility for an increase in the charter school cap also has consequences for existing and prospective charter school operators. Including SGP in the formula has led to more movement in and out of the bottom 10 percent from year-to-year. Difficulty in predicting which districts will fall "in" or "out" of the bottom 10 percent annually means that existing operators have difficulty predicting enrollment and school budgets. For their part, prospective operators may feel discouraged from applying for charter schools in high demand areas that are on the cusp of the bottom 10 percent because it is unclear if the state will be able to authorize more charter schools from year to year.

The following paper presents data on the relationship between enrollment and SGP and its impact on determining the lowest performing districts in Massachusetts. Based on these data, the authors recommend that the Commonwealth stop using SGP as a factor in determining eligibility for an increase in the charter school cap and revert to a formula that uses absolute proficiency as the sole measure. The

Commonwealth should also separate determinations about increases in the net school spending cap from its overall accountability system. Other recommendations include mitigating the negative impacts of the current formula for existing and prospective charter operators and increasing transparency around the purpose and use of SGP more broadly.

The Smart Cap: An Overview

The Massachusetts Legislature established charter schools in 1993 as part of the Massachusetts Education Reform Act (MERA). Commonwealth charter schools have more autonomy in exchange for more accountability. With freedom from some of the bureaucratic constraints that can hinder district schools, the legislature hoped charters would innovate and provide new public-school options for students and families.[1]

To date, many Massachusetts charter schools have delivered on the promise of innovation. The majority have also proved to be very high-quality academic options, especially for poor and minority students and (increasingly) English language learners and students with disabilities.[2] Since 2009, several gold-standard studies have shown that Boston's charter schools, in particular, outperform their district counterparts in terms of standardized tests, graduation rates, number of Advanced Placement exams taken and passed, and college attendance.[3] Another study, by Stanford University's Center for Research on Educational Outcomes, described Boston charters as some of the highest performing public schools in the nation. It found that compared to students in surrounding district schools, Boston charter schools "added an additional 12 months of learning in reading and 13 months of learning in math each school year."[4]

To date, many Massachusetts charter schools have delivered on the promise of innovation. The majority have also proved to be very high-quality academic options, especially for poor and minority students

But in the beginning, the legislature could not predict whether charter schools would be successful. This is one reason why MERA capped the number of charter schools the state could authorize at 25.[5] Demand for these new options, especially in urban centers, outstripped supply within a few short years.[6] In response, the legislature modestly raised the statewide cap two more times over the next decade, but it did so in the midst of increasing antipathy toward charters, especially from the state's two powerful teachers unions.

Opposition to charter schools had existed from the start, but opponents moved to halt charter expansion at the same time it was becoming clear the schools were popular with parents. To quell fears that charter schools would "drain" district enrollments, in 1997 the legislature designed a second charter school cap. This cap limits the amount of money (charter school "tuition") that districts can send to charters when students choose to attend them.[7] This cap is also called "NSS".

In Massachusetts, when a student leaves a district for a charter school, the per-pupil allocation that she would have received in the district—from both the state and local sources—follows her to the charter. The law limits the amount districts can send to charters to 9 percent of net school spending.[8]

Opposition to charter schools had existed from the start, but opponents moved to halt charter expansion at the same time it was becoming clear that the schools were popular with parents.

In 2019, there is still room to establish charter schools under the statewide cap of 120 schools, but some large urban centers and "gateway" cities like Boston, Springfield, Lawrence, and Everett have reached the NSS cap. In these and other cities, there are tens of thousands of students waiting for new charter seats to become available. The statewide total of individual students on charter school waitlists was more than 25,000 as of March 2018.[9]

Massachusetts Charter Schools

Authorizer: Board of Elementary and Secondary Education (takes recommendations from Department of Elementary and Secondary Education)

Operating Commonwealth Charter Schools:
74

Unique students on waitlists as of May 2019:
25,308

Demographics

	Charters	State
English Language Learner	14.1%	10.5%
Special Education	15.5%	18.1%
Economically Disadvantaged	41.5%	31.2%

SOURCES: http://www.doe.mass.edu/charter/enrollment/fy2019/updated-waitlist.html; http://www.doe.mass.edu/charter/about.html

The NSS cap has changed slightly over time. In 2010, responding to incentives from the federal *Race to the Top* initiative, the legislature instituted a "smart cap," under which school districts whose performance on the Massachusetts Comprehensive Assessment System places them in the bottom 10 percent statewide are subject to an increase in the amount of charter school tuition they can pay. Now, when districts perform in the bottom 10 percent, their charter tuition cap

increases from 9 to 18 percent of net school spending. The law states that any new charter school seats awarded under the smart cap must go to "proven providers"—operators with a proven track record of helping students achieve strong academic outcomes.[10]

The smart cap solved two problems in 2010: it opened up new seats in high-demand urban centers, and increased the probability that students entering new charter schools would receive a high-quality education. But since that time, it has become clear that in some districts, the smart cap was only enough to make a small dent in charter school waiting lists.[11] Moreover, because of the proven provider clause, charter schools in low-performing districts have been replications of existing programs, with few new (or innovative) options entering the market.

Using growth as even a small factor when determining which districts are eligible for an increase in the charter school cap runs counter to the spirit of the 2010 law: that students in districts where overall proficiency is low should have access to additional charter school options.

These two consequences of the cap are concerning.[12] Also concerning, however, are state regulations that dictate how the DESE determines which districts fall into the bottom 10 percent.

For the first three years of the smart cap, the state determined the bottom 10 percent based on MCAS proficiency scores alone: It ranked districts according to the percentage of students meeting and not meeting proficiency thresholds on MCAS math, English language arts, and science tests. But in 2014 the Board of Elementary and Secondary Education (BESE – the Commonwealth's sole charter authorizer) determined that student growth scores on standardized assessments should be factored into district rankings. This was part of a broader move to incorporate growth scores into the state's entire accountability system—a system that is distinct from the ranking process the department had been using to determine which districts are eligible for an increase in the charter cap.

Since that time, the state has used a district's median "student growth percentile", in conjunction with the percentage of students who meet proficiency thresholds on each test, to determine which districts fall into the bottom 10 percent of performance. ***The current formula weights SGP at 25 percent and proficiency at 75 percent.***[13]

BESE's desire to consider growth as a factor in determining the lowest performing districts was well-intentioned. The board believed that districts should be rewarded when they help otherwise low-performing groups of students make progress, even when overall proficiency is low. Former BESE Vice Chair Harneen Chernow noted at the time:

> *"MCAS [test score] data almost always corresponds (negatively) with the socio-economic status of the district… but we are looking at improvement and innovation and change and where districts are doing good things… our goal should be to support and reinforce those outcomes."*[14]

The idea that growth should be encouraged might be important in the context of a broader system of district accountability. However, using growth as even a small factor when determining which districts are eligible for an increase in the charter school cap runs counter to the spirit of the 2010 law: that students in districts where overall proficiency is low should have access to additional charter school options.

Data show that some districts exit the bottom 10 percent of performance because their growth scores seem strong, even when overall student proficiency remains low. In such cases, the ranking is a false boon for districts (which still need ample support to improve) and a real loss for parents and students seeking access to charter school opportunities that will not be available.

When BESE made the decision to merge the formula for determining the lowest 10 percent of districts with the state's overall system of district accountability, it might not have fully understood the various impacts of using growth measurements for high-stakes policy decisions. When some BESE members questioned why SGP shouldn't receive more weight in the formula for determining district performance, then-Commissioner of Elementary and Secondary Education Mitchell Chester warned of factoring student growth too heavily: "I am not recommending a larger increase in the weighting of growth, because it would start to distort the identification of schools and districts most in need of our assistance."[15]

In this statement, Chester captured the concerns of many charter advocates. By rewarding the lowest performing districts for high growth scores, charter advocates feared two things: 1) Many districts would still be able to fail large swaths of students who were not proficient simply because they moved some students from the very "bottom" of proficiency categories to the middle or top of the bottom; and 2) Including growth as a factor would likely allow some districts where demand for charters is highest to exit the bottom 10 percent. Districts like Boston, for example, have had consistently low MCAS scores over time (especially for low-income and minority students, students with disabilities, and English language learners), but have come very close to exiting the bottom 10 percent because their SGP scores have impacted overall rankings.

Data show that some districts exit the bottom 10 percent of performance because their growth scores seem strong, even when overall student proficiency remains low. In such cases, the ranking is a false boon for districts.

On both counts, charter advocates have been right. The City of Lynn, Massachusetts provides one example. Between 2015 and 2018 Lynn has held various rankings within the bottom 10 percent, climbing slightly in some years due to a steadily increasing median SGP.[16] In the 2018–19 school year, Lynn exited the bottom 10 percent mainly due to a sizeable increase in SGP, although it is still the 24th lowest performing district in the Commonwealth in terms of absolute proficiency on MCAS.[17]

In 2018, only 33 percent of Lynn's fourth graders and 31 percent of eighth graders reached proficiency on MCAS English language arts assessments. In math, only 33 percent of fourth graders and 23 percent of eighth graders reached proficiency. For English language learners and students with disabilities, the data are even more discouraging: on grade 4 math assessments, only 23 percent of English language learners and 7 percent of students with disabilities scored proficient on MCAS. These numbers indicate that Lynn is not even close to helping a majority of students achieve basic proficiency.[18]

As of 2018 there are 1,920 unique students on charter school waiting lists in Lynn;[19] 12 percent of the district's student population is seeking access to charter schools that are, on average, higher performing than the district. However, access to additional charter school seats will not be available in the coming year, because the district's NSS cap will be reduced from 18 to 9 percent. ***Parents in Lynn who desire access to charter public schools, will have to wait for the district to fail even more students before the spending cap will increase again.***

Percentage of Students Not Meeting Expectations, Grade 8 MCAS English Language Arts and Math, Lynn School District, KIPP Charter Public School, Lynn, MA, State

	Lynn (District)	KIPP Lynn (Charter)	State
English	20	11	15
Math	18	7	12

SOURCE: http://profiles.doe.mass.edu

And while reduced access to high-performing schools is the most important problem the Lynn situation illustrates, the volatility of the net school spending cap creates another issue, too. In a given year a worthy charter operator could be poised to open a new school under and 18 percent NSS cap but ultimately be unable to do so because the district exits the bottom 10 percent before the school can open. Acting Commissioner Jeff Wulfson outlined this problem in a February 2018 memo to the board. Discussing a proposal for a new charter school in Lynn he wrote:

> *"…the application for Equity Lab Charter School substantially met the criteria for approval. However, I was unable to recommend this school for a charter because of upcoming changes in the NSS cap for Lynn."*[20]

A problem similar to what Wulfson described exists in other communities that are on the cusp of exiting the bottom 10 percent, such as Everett. In 2015 and 2016 Everett was ranked among the lowest performing districts in the Commonwealth. As of 2017 it exited the bottom 10 percent, in large part due to its SGP. This may seem like a win for Everett families, but it's not.

In 2015 and 2016 Everett was ranked among the lowest performing districts in the Commonwealth. As of 2017 it exited the bottom 10 percent, in large part due to its SGP.

The Pioneer Charter School of Science (PCSS) operates three campuses in Revere, Everett, and Chelsea, Massachusetts. As of January 2019, the organization reported 435 individual students on waitlists from Everett alone.[21] Because Everett's net school spending cap on charter schools was cut in half in 2018, PCSS's Everett campus can only admit siblings (who have preference) of enrolled students from its waitlist, which continues to grow.

PCSS's waitlisted students aren't the only people affected by the uncertainty associated with Everett's ranking as a low-performing district. Enrolled students suffer as well. With the reduction in Everett's NSS cap, PCSS feared in 2018 that it could lose up to $200,000 due the forced decline in the number of students it can enroll. Such financial uncertainty is difficult for any school, because enrollment is the major budget driver, dictating everything from staffing to course and extra-curricular offerings. To its credit, DESE realized the severity of the financial blow PCSS might take and made adjustments, mainly to the sibling reimbursement formula, to protect the organization. But this scenario illustrates yet another negative impact of the regulations that dictate how the department determines the NSS cap.

The lowering of the charter school cap in communities like Everett and Lynn gets to the heart of a very politicized charter school debate. Advocates of including SGP in the formula for determining low performance do not want to "punish" districts that have demonstrated progress with student populations (mainly low-income) that they deem "harder to educate." For their part, charter advocates see the inclusion of SGP as a mechanism to deny students access to charter public schools, even though the traditional district is failing them.

But both these narratives are simplistic, and neither poses the most meaningful questions: ***What measure or measures will provide the most reliable information about which districts are eligible for an increase in the net school spending cap? What measure or measures will provide additional opportunities for***

families, aligned with the spirit of the 2010 law?

Absolute proficiency may not be the best indicator of whether a school or district is making progress educating students, but it does provide an indication of whether schools and districts are able to educate students to a minimum standard. Student growth percentiles are, on the other hand, a much more complicated measure. Parents, policymakers, and even educators assume they are a reliable measure of how much a teacher, school, or district has "moved" a student up the ladder of proficiency. But this is not an accurate description of SGP. According to an increasing number of psychometricians, SGP is widely misunderstood and widely misused.

For their part, charter advocates see the inclusion of SGP as a mechanism to deny students access to charter public schools, even though the traditional district is failing them.

In Massachusetts, reliance on SGP for making high-stakes policy decisions has made access to charter schools more unpredictable. SGP causes districts to enter and exit the bottom 10 percent with more frequency than they would if absolute proficiency remained the only measure of performance. This is because SGP measurements include a large degree of statistical error.

This paper discusses the reliability and impact of SGP as a factor in determining whether Massachusetts school districts are subject to an increase in the charter school cap. To illustrate the point, the authors use publicly available data to demonstrate how factoring SGP into the formula for determining the NSS cap moves districts with higher enrollment out of the bottom 10 percent. This doesn't happen because these districts are showing dramatic achievement growth; rather, when enrollment is high, SGP scores are more likely to cluster around a "typical mean." Conversely, districts with lower enrollments are more likely to have volatile SGP scores, which can move them into the bottom 10 percent even when absolute achievement is (comparatively) better than it is in high-enrollment districts.

The Student Growth Percentile and Accountability

In recent years, states have begun to use a student growth percentile as one way of reporting student, school, and district performance on standardized tests.[22] In the decade following implementation of the *No Child Left Behind Act*, which required states to employ high-stakes standardized assessments in return for federal funding, some stakeholder groups were concerned that reporting student test scores in terms of proficiency alone (whether a student meets a pre-defined standard on an assessment) was misleading.

A common argument against relying on proficiency alone is that it does not help teachers, students, parents, or policymakers understand how much progress a student has made over time. A student who enters school far behind her peers may progress rapidly with the help of an effective teacher but still be unable to pass a standardized test. Conversely, a student who enters with grade-level skills could make comparatively little academic progress over the same amount of time but still reach proficiency on an examination. In this scenario, it is easy to see why measuring growth, instead of just proficiency, is desirable.

In its explanation of this "growth model," DESE goes to great lengths to assure stakeholders that SGP is just one part of a more holistic approach to how it understands student, school, and teacher performance.[27] However, there are high stakes attached to SGP scores.

Measuring growth is particularly compelling for schools and districts that serve concentrated populations of low-income students and/or those that see wide achievement gaps when students enter school. Because proficiency rates correlate with socio-economic status (low-income children are more likely to struggle to meet proficiency), measuring growth feels more just to some educators and policymakers. Critics of measuring proficiency alone also point out that the pressure associated with helping students score proficient on tests incents educators to give disproportionate attention to "bubble kids," those who can pass the test with enough help.[23]

This was, in part, what the DESE reasoned when it first began reporting growth in 2011. The Department's *MCAS Student Growth Percentiles Interpretive Guide* from that year states:

> *Measuring student performance relative to standards specific to each grade level is useful in determining whether a student has met the standards for that grade. There are, however, several obstacles to using this approach to measure students' academic growth.* ***This is why we have developed "student growth percentiles," a measure of student progress that compares changes in a student's MCAS scores to changes in MCAS scores of other students with similar scores in prior years****. A student growth percentile measures student progress by comparing one student's progress to the progress of other students with similar MCAS performance histories. We refer to students with similar score histories as "academic peers."*[24]

Along with other measures, DESE reports student-level growth scores on parent/guardian reports. These individual student growth percentiles are also aggregated to the classroom, school, and district levels. Since 2011, DESE has used the median of all aggregated individual scores to report school and district-level SGP.[25] Under its new accountability system,

piloted in 2018, aggregate SGP scores will be reported as the mean of individual scores.[26]

In its explanation of this "growth model," DESE goes to great lengths to assure stakeholders that SGP is just one part of a more holistic approach to how it understands student, school, and teacher performance.[27] However, there are high stakes attached to SGP scores. Aggregated SGPs *may* be used at the school and district level for accountability purposes (including teacher evaluations), and as of 2018 the Commonwealth formally relies upon SGP as one part of a formula to rate districts and hold them accountable for performance."[28] SGP scores can influence which districts and schools might be candidates for state intervention or "turn-around" and which districts and schools no longer need intervention and/or support.

Massachusetts isn't alone in its use of SGP as a factor in determining school performance. When it was first introduced around 2009, the federal government quickly embraced the idea, encouraging states to incorporate some measure of student growth in *Race to the Top* applications and plans for *The Every Student Succeeds Act* (ESSA).[29]

Massachusetts isn't alone in its use of SGP as a factor in determining school performance. When it was first introduced around 2009, the federal government quickly embraced the idea, encouraging states to incorporate some measure of student growth in *Race to the Top* applications and plans for *The Every Student Succeeds Act* (ESSA).[29] As of 2018, a majority of states include some measure of student growth in their federal ESSA plans, though not all use SGP as the reporting tool.[30]

But in Massachusetts and nationwide, the rapid rise of SGP as a reporting tool and a mechanism for high-stakes policy decisions is troubling. A growing body of research suggests that practitioners and policy makers should better understand the utility, risks, and benefits of this sophisticated measure.

Sireci, Wells & Keller of the Center for Educational Assessment at the University of Massachusetts at Amherst summarize several problems with SGP, ranging from the practical to the very technical. They argue that SGP is a widely misunderstood reporting tool; parents don't understand what a "growth percentile" is, and teachers don't know how to use SGP to inform practice. They also contend that SGP is an unreliable measure, with "no validity evidence" to support its use. Although the authors limit their discussion of SGP to the student and classroom levels, their rationale also applies to SGP use at the district level. They forcefully argue that states should abandon SGP for both reporting and evaluation/accountability purposes.[31]

Some researchers disagree with Sireci, Wells, & Keller. They acknowledge that SGP is not a perfect way to measure student growth but argue that the information it can provide is useful enough to warrant its continued use. Andrew Ho of Harvard (who also advises DESE on its use of SGP) responded to Sireci, Wells, & Keller, writing:…" the sufficiency of SGP reliability (or any score reliability) depends upon the intended interpretations and uses of SGPs." In the view of Ho and others, if we are transparent about what SGP is and thoughtful with how we use it, SGP can complement other measures of student achievement.

Assessments of SGP reliability derive from an increasingly large body of literature, which notes that SGP scores suffer from a wide margin of error, especially at the most granular (student and teacher) reporting levels. Several studies in the past few years report that a 95 percent confidence interval for SGP scores is roughly 50 points.[32] In lay terms, if a school's *reported* SGP median score is 50, researchers can be 95 percent confident that the school's *actual* SGP median score is somewhere between 25 and 75. Five percent of the time, even this wide range will be incorrect.

With SGPs reported from 1–99 percent for students, schools, and districts, 50 points is a very large margin of error. Put differently, if one school received an SGP of 30 and another a 70, researchers couldn't be confident that the school with the higher score actually helped students grow more than the school with the lower score.

Researchers have been looking for ways to mitigate this bias, and there is evidence that aggregating SGP scores (for example, to the district level) makes them somewhat more reliable.[33] Other research suggests that using the mean, as opposed to the commonly-used median, to measure SGP could mitigate reliability issues. Castellano & Ho[34] find that the mean has greater sampling variability, making it "a more attractive aggregation function." This is one reason why the Commonwealth will begin to calculate SGP using the mean as opposed to the median.

Assessments of SGP reliability derive from an increasingly large body of literature, which notes that SGP scores suffer from a wide margin of error, especially at the most granular (student and teacher) reporting levels.

But this move will only mitigate some of the error inherent in SGP. This measure, like any other, is not perfect. Given this, ***DESE should be more transparent about the limits of SGP as an evaluation and reporting mechanism. Both the department and the board should also be wary of overemphasizing SGP as one part of a larger system of accountability and of using SGP for high-stakes policy decisions.***

Using SGP as part of the formula that determines which districts are eligible for a charter cap increase is a very

high-stakes policy decision, and over the years the consequences of incorporating SGP into the formula to determine the lowest performing 10 percent of districts have become clear. Including SGP in the formula distorts growth in some school districts.

The Impact of SGP on Calculating the Lowest Performing Districts in MA

In any given year since 2011, a cursory scan of school districts that fall into the bottom 10 percent on MCAS performance reveals a handful of districts—many of them "gateway districts"—that have difficulty helping students reach proficiency in English, math, and science. Districts with very low absolute performance on MCAS almost always fall into the bottom of the bottom 10 percent.

The "Smart" Cap

When districts fall into the lowest 10 percent, as measured by proficiency on standardized tests of English, math, and science (75 percent) and SGP (25 percent), the amount of tuition the district can send to charter schools rises from 9 percent to 18 percent of net school spending. Any new charter school seats in these districts must be awarded to "proven providers."

But in 2015, with the introduction of SGP as 25 percent of the formula for determining the lowest performing school districts, an interesting pattern began to emerge. Some of the largest Massachusetts school districts—ones with very low proficiency scores but middling to high growth scores—were steadily moving up in the overall rankings. Some moved out of the bottom 10 percent altogether. Curiously, districts that moved into the bottom, mainly due to very low SGP scores, tended to be very small, enrolling comparatively few students.

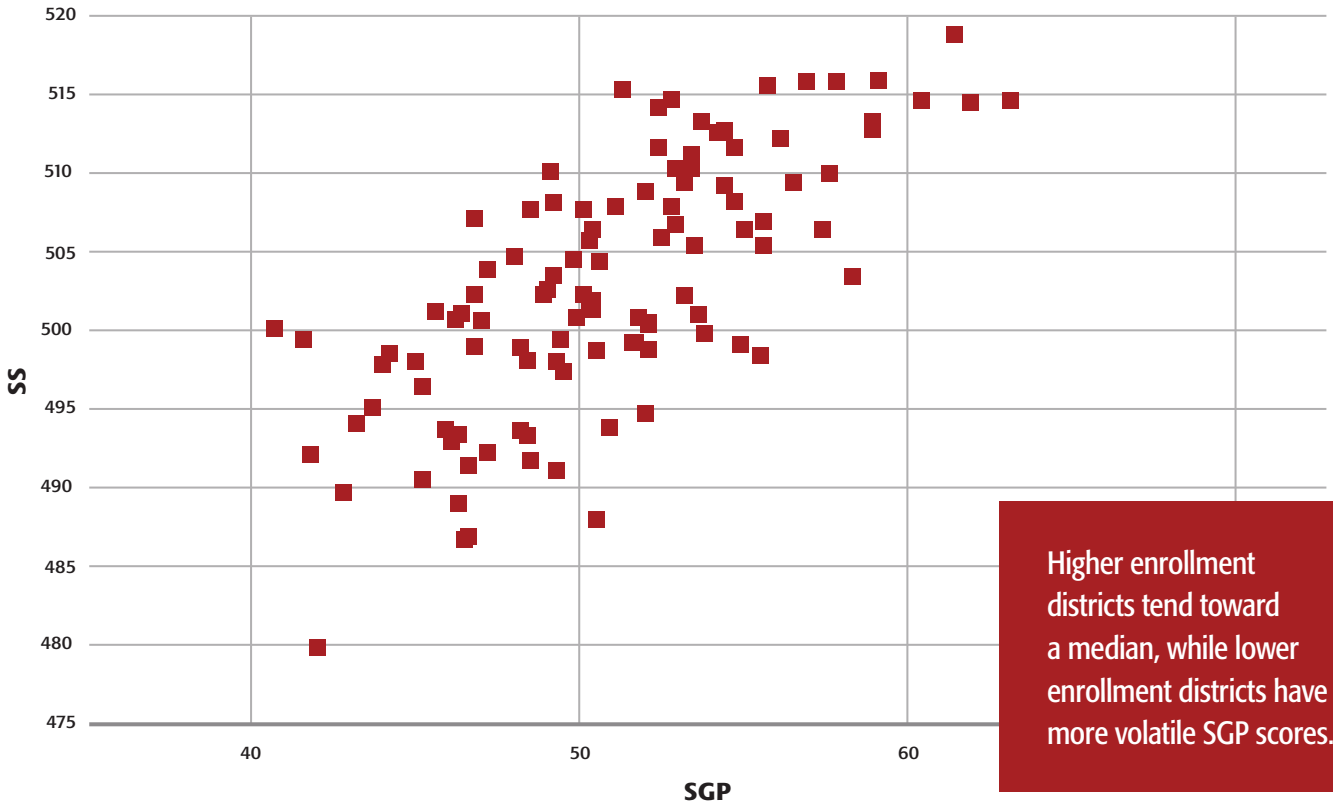
A high-level analysis of what happens when SGP is and is not included as a factor for determining the lowest performing districts suggests that there is a relationship between SGP calculations and the size of school districts. Put another way, school districts with comparatively low enrollment (in most cases below or far below 2,500) are more likely than their larger counterparts to have a median SGP that falls outside a "typical" range.

Existing literature on SGP suggests a relationship between sample size and the precision of aggregate SGP estimates. Culbertson finds that SGP estimates are less precise for high- and low-achieving students than for those with average achievement when the total sample size is small. He also notes that some steps states take to reduce this error, such as categorizing students into performance bands based on prior achievement before estimating SGP, may reduce error slightly but have other trade-offs, such as "reducing the similarity of students whose growth is compared."[36]

The authors tested the observation of a relationship between non-typical SGP and low enrollment by identifying high- and low-enrollment districts in the Commonwealth and plotting their SGP and proficiency scores (represented as the "scaled score" reported by DESE) to see if a pattern emerged.[36] The authors identified Massachusetts districts where fewer than 1,000 students were included in the 2018 Next Generation MCAS as low enrollment and districts where more than 1,000 students were included in the exam as high enrollment. Neither of these sample sizes is enough to sufficiently reduce the wide margin of error inherent in SGP, but the analysis below shows that districts with higher enrollment are more likely to cluster around what DESE refers to as a "typical" median of 40–60 SGP. [37]

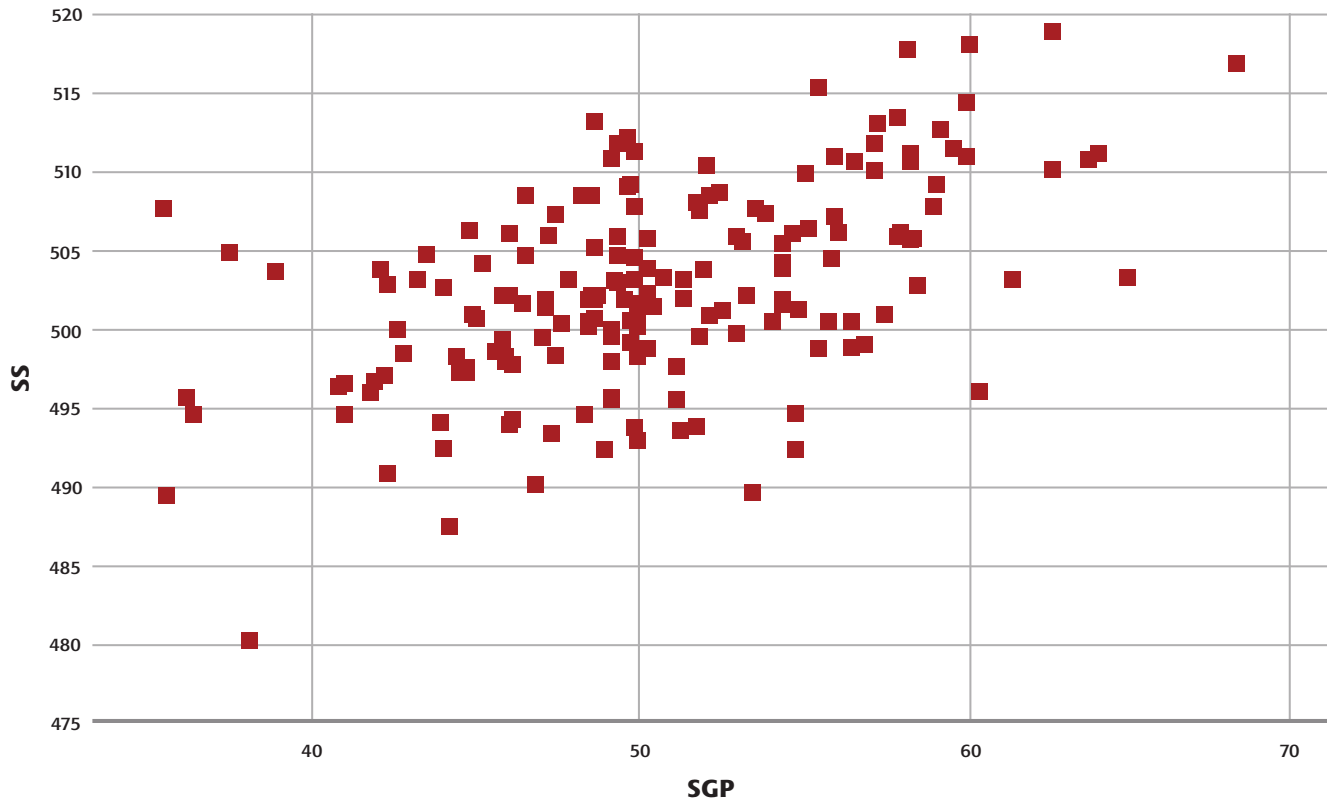
A high-level analysis of what happens when SGP is and is not included as a factor for determining the lowest performing districts suggests that there is a relationship between SGP calculations and the size of school districts.

Graph 1: 2018 MCAS ELA, SGP and Scaled Scores for High Enrollment Districts*



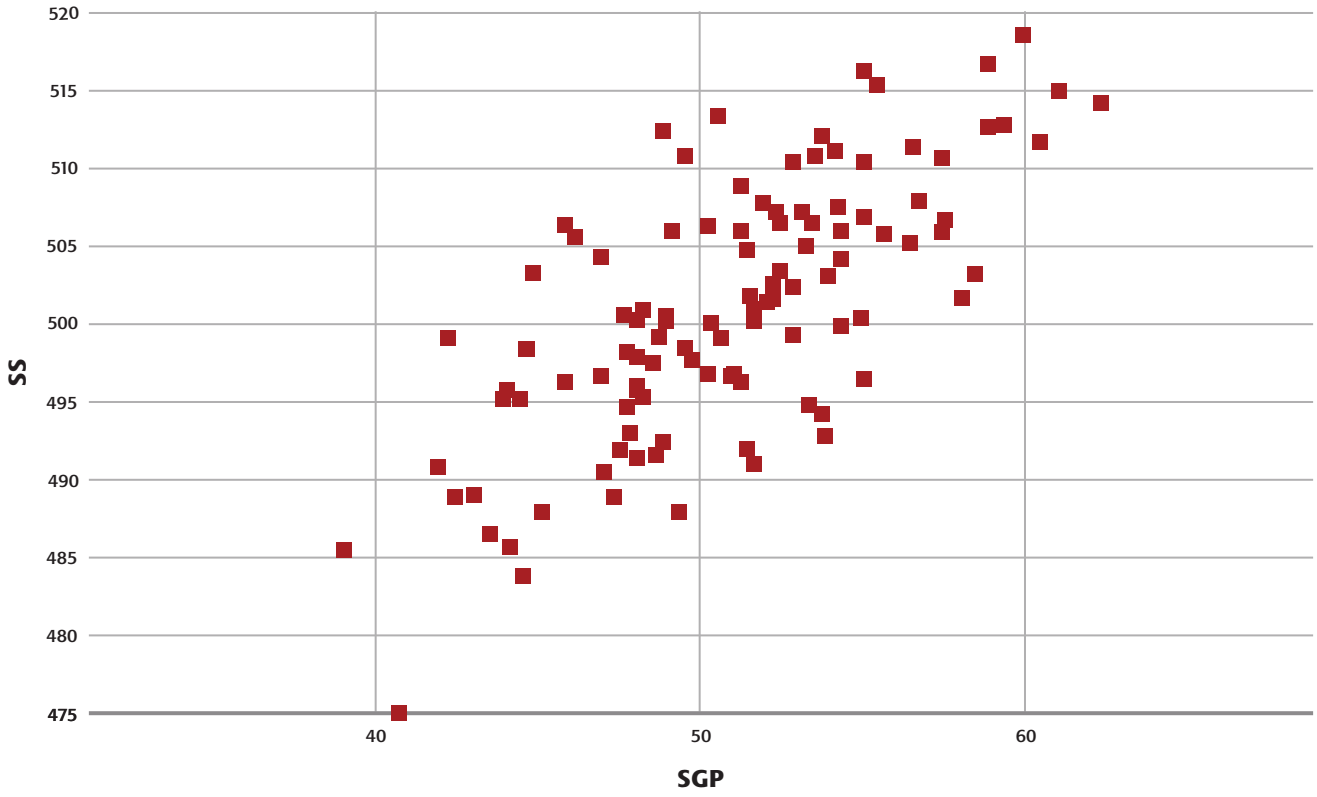
Higher enrollment districts tend toward a median, while lower enrollment districts have more volatile SGP scores.

Graph 2: 2018 MCAS ELA, SGP and Scaled Scores for Low Enrollment Districts*

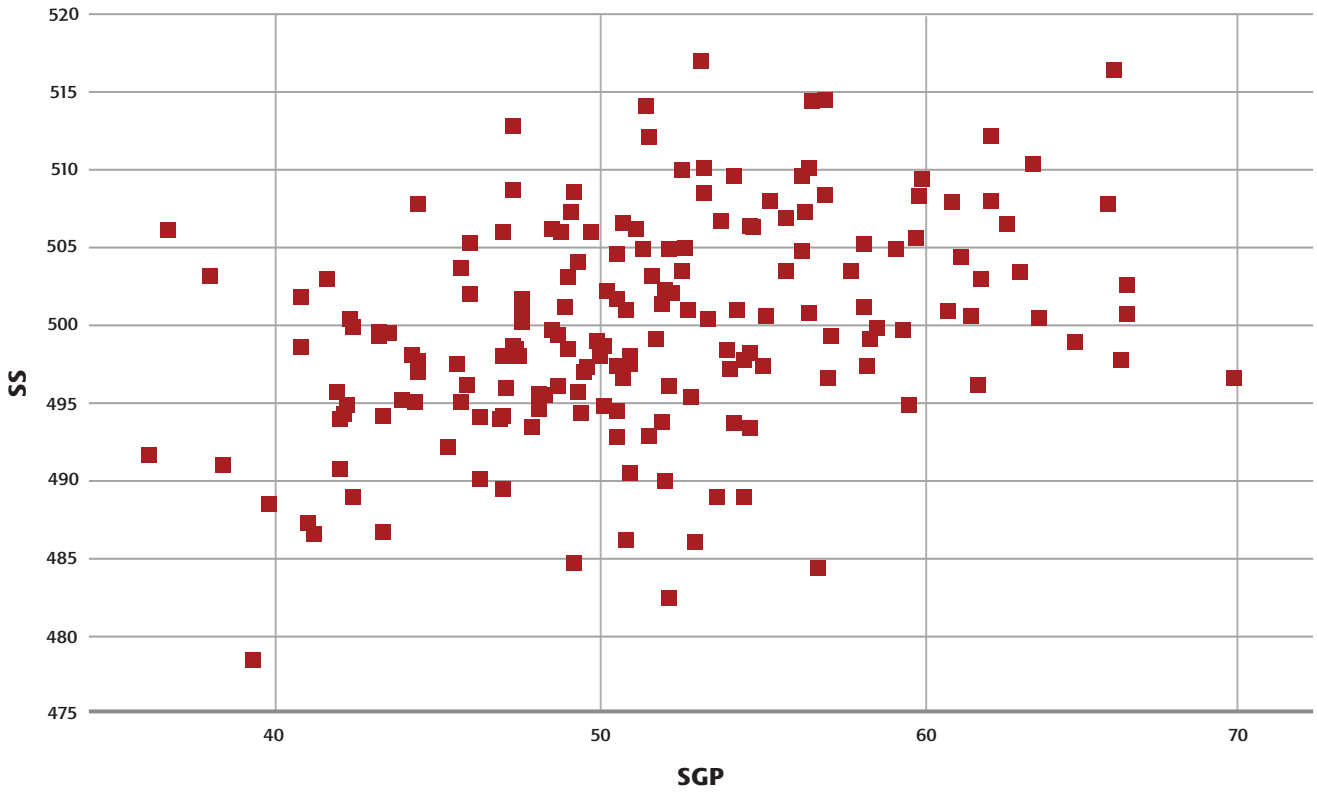


* Publicly available data provided by Massachusetts Department of Elementary and Secondary Education: authors' analysis

Graph 3: 2018 MCAS Math, SGP and Scaled Scores for High Enrollment Districts*



Graph 4: 2018 MCAS Math, SGP and Scaled Scores for Low Enrollment Districts*



* Publicly available data provided by Massachusetts Department of Elementary and Secondary Education: authors' analysis

These visuals suggest a correlation between a district's enrollment and typical (between 40 and 60) or non-typical (less than 40 or greater than 60) SGP. It indicates that districts with higher enrollment may receive a higher (or "better") SGP score because higher enrollment means they are more likely to have an aggregate SGP score close to the typical median of 50 (even though this is not a "true" SGP). Districts with lower enrollment are more volatile, and therefore more likely to show either very high or very low growth, and less likely to cluster around the typical median.

These visuals suggest that the already large amount of random error inherent in SGP calculations is exacerbated when sample size is insufficient. Because any relationship between a measure of district academic performance and something unrelated, such as sample size, is spurious, this correlation suggests that DESE and outside researchers should undertake further study.[38]

This relationship also begs two very important policy questions: ***Are larger districts that have low proficiency but high growth exiting the bottom 10 percent because their student outcomes are actually improving? Or, are smaller districts entering the bottom 10 percent, essentially displacing those larger districts that tend to cluster around a typical median SGP, because their SGP scores are even more unreliable than the SGP scores of the larger districts?*** Either way, a high-stakes policy that rewards and punishes districts based in part on SGP, is advancing a disingenuous proposition that any of these districts are performing either "better" or "worse" than others.

Using the most recent DESE data (2018), Chart 2 below shows the difference between districts that fall into the bottom 10 percent based on proficiency alone and districts that fall into the bottom 10 percent when SGP comprises 25 percent of the formula. It reveals that many low-enrollment districts that enter the bottom 10 percent based on low growth scores rank slightly higher when growth is not a factor.

Chart 2

Lowest Performing Districts by Achievement Only (2018)

Overall Rank (1–289)	District	Overall Achievement Rank out of 289	Overall Growth Rank out of 289
2	Holyoke	1	9
2	Southbridge	2	6
1	Webster	3	1
7	Chelsea	4	26
5	Springfield	5	11
14	Brockton	6	73
4	New Bedford	7	3
6	Gardner	8	5
26	Lawrence	9	132
12	Athol-Royalston	10	41
27	Orange	11	127
10	Fitchburg	12	30
13	Boston	13	41
8	North Adams	14	4
37	Worcester	15	136
19	Wareham	16	91
28	Adams-Cheshire	17	111
9	Winchendon	18	2
57	Randolph	19	201
11	Taunton	20	8
34	Fall River	21	110
43	Everett	22	139
40	Lowell	23	116
51	Lynn	24	159
21	North Brookfield	25	71
31	Gill-Montague	26	88
15	Ware	27	24
18	Haverhill	28	49
49	Salem	29	138

Lowest Performing Districts by Achievement and SGP (2018)

Overall Rank (1–289)	District	Overall Achievement Rank out of 289	Overall Growth Rank out of 289
1	Webster	3	1
2	Holyoke	1	9
2	Southbridge	2	6
4	New Bedford	7	3
5	Springfield	5	11
6	Gardner	8	5
7	Chelsea	4	26
8	North Adams	14	4
9	Winchendon	18	2
10	Fitchburg	12	30
11	Taunton	20	8
12	Athol-Royalston	10	41
13	Boston	13	41
14	Brockton	6	73
15	Ware	27	24
16	Greenfield	31	15
17	Pittsfield	30	25
18	Haverhill	28	49
19	Wareham	16	91
20	Easthampton	43	13
21	North Brookfield	25	71
22	Marlborough	34	46
23	Rockland	38	39
24	Leicester	45	19
25	Palmer	41	35
26	Lawrence	9	132
27	Orange	11	127
28	Adams-Cheshire	17	111
29	Hawlemont	51	11

Another way to view the data is to ask which districts exit the bottom 10 percent because of higher growth scores and which take their place due to lower growth scores. Chart 3 below shows districts that exit the bottom 10 percent when growth is a factor and the districts that enter to the bottom 10 percent to take their place. The chart also shows total enrollment for each district as well as the number of students that, according to the state, are on charter school waitlists in each district. Importantly, the total enrollment provided here is a larger number than that which the state would use to calculate SGP, because not all students in the districts will take standardized examinations that count toward the aggregate SGP score.

Of the districts that exit the bottom 10 percent when SGP is a factor, half have total student enrollment of more than 5,000

Chart 3 below gives a clearer picture of the influence enrollment likely has on SGP. Of the districts that exit the bottom 10 percent when SGP is a factor, half have total student enrollment of more than 5,000 and all but one have enrollment greater than 2500 students. In fact, four of the districts listed here (Lynn, Lowell, Worcester, and Fall River) are among the largest in the state. Of those that enter the bottom 10 percent, only one—Pittsfield—has total enrollment greater than 5,000 (and it is very likely that the number of students included in the state's SGP calculation for Pittsfield is not close to 5,000). ***Perhaps more telling, 75 percent of the districts that enter the bottom 10 percent when growth is a factor have total student enrollments below 2,500. At least five of these districts (Hawlemont, Easthampton, Greenfield, Palmer, and Leicester) are among the smallest in the state.***

Another indicator is the difference in the range of growth scores between the larger districts that exit the bottom 10 percent and smaller districts that enter the bottom 10 percent. As the literature suggests, larger districts that exit the bottom 10 percent cluster around the "typical" median, with a range of SGP percentile scores between 46.1 and 52 (6 points). The range among districts that enter the bottom 10 percent is considerably wider, between 36.2 and 56.4 (20 points).

And a closer look at SGP within districts reveals more wide differences. In Worcester, the largest district in this analysis, school-level aggregate SGP percentiles range from 22.3 on the low end to 72.7 on the high end. In Easthampton, the smallest district for which enough data are available to show a range of aggregate SGP percentiles between schools, the percentiles range from 40.8 on the low end to 63.5 on the high end. This suggests that, despite having a median SGP within the typical range, a large district like Worcester is home to many schools that theoretically have much lower SGP scores (we don't know if they are "true" SGP) but will nonetheless exit the bottom 10 percent because the entire district skews toward the median.

In fact, four of the districts listed here (Lynn, Lowell, Worcester, and Fall River) are among the largest in the state.

This analysis does not suggest that districts that enter the bottom 10 percent are high performing; all the districts presented in this analysis have comparatively low overall proficiency rates. Instead, it shows that districts with higher enrollment are more likely to have a typical SGP score, regardless of whether they are helping students achieve or

Chart 3: Districts that Exit and Enter The Bottom 10% with SGP In Formula (2017–18 MCAS Data)

Exit Bottom 10% With SGP in Formula			
	Total Enrollment	SGP %	Charter Waitlist
Salem	3,694	52	192
Lynn	15,517	47.2	1,464
Lowell	14, 436	46.1	481
Everett	7,068	48.4	868
Randolph	2,823	51.1	427
Worcester	25,306	46.2	874
Fall River	10,128	46.6	594
Gill-Montague	976	49.9	0
SGP % Range		**46.1–52**	

Enter Bottom 10% with SGP in Formula			
	Total Enrollment	SGP %	Charter Waitlist
Hawlemont	163	36.2	0
Rockland	2,193	56.4	56
Marlborough	4,575	50.9	133
Easthampton	1,541	46	95
Pittsfield	5,464	41.8	0
Greenfield	1,699	46.1	46
Palmer	1,400	49.1	12
Leicester	1,569	49.8	7
		36.2–56.4	

grow toward proficiency. Lower-enrollment districts, on the other hand, are less likely to show a typical SGP score and will therefore displace higher enrollment districts in the bottom 10 percent. If these data are correct, students are not well served by the state's formula for determining which districts need additional support.

By law, one form of intervention and/or support districts become eligible for when they enter the bottom 10 percent is additional charter school seats. Factoring demand for charters into this analysis is important, because a number of the districts that exit the bottom 10 percent when SGP is a factor have comparatively high demand (in the form of waitlists) for charter schools. Conversely, many of the districts that enter the bottom 10 percent when growth is a factor have few if any students on charter waitlists. This could indicate a total lack of desire for charter schools, or that there are few if any charter school options in these communities.

Lower-enrollment districts, on the other hand, are less likely to show a typical SGP score and will therefore displace higher enrollment districts in the bottom 10 percent.

The 2010 law didn't provide communities with high charter demand more access to these public schools, but in the years immediately after the law was enacted, the same communities that became eligible for an increase in the NSS cap also saw high demand for charter schools. It was not until the board voted to incorporate SGP into the formula for determining the lowest performing districts that communities with limited demand for charter schools became eligible for an increase in the cap. The 2010 law does not require that BESE use the same metrics for the state's overall accountability system and for determining the lowest performing districts that will become eligible for an increase in the NSS cap.

Theoretically, BESE could employ two systems: one that conservatively factors SGP into a district's accountability rating, and one that uses absolute proficiency alone (a more reliable and stable measure), to determine which districts are the lowest performing for purposes of the NSS cap. Using two systems would not only yield useful, comparative information for the state, it would address a serious equity issue by granting increased access to charter schools in the communities where parents most want them.

Moreover, using absolute proficiency as the only measure for determining which communities are eligible for an increase in the charter school cap ***is better aligned with the spirit of the law than a policy that considers both growth and proficiency. The intent of the 2010 law was to provide students and families in districts where schools struggle to help students meet proficiency with different public school options. It is time to realize the legislature's vision.***

Conclusions and Recommendations

Massachusetts has some of the highest performing charter schools in the country, but district net school spending caps inhibit access and innovation. Under the law, districts must abjectly fail before the NSS cap increases in high-demand districts. This policy frames charter schools as an escape valve, which is unhealthy for the public-school community (charter public schools included).

More importantly, regulations governing how the state determines the NSS cap have created a situation in which access to charter schools in some communities exists for only a short window of time. This deters prospective charter operators, creates financial and other uncertainties for existing providers, and confuses and frustrates parents and students on charter school waiting lists.

The analysis presented in this paper suggests that the current formula for determining districts that are eligible for an increase in the charter school cap is unreliable and volatile. The authors find a relationship between SGP and district enrollment, which suggests that statistical error influences district SGP scores.

When SGP accounts for 25 percent of the formula for determining the lowest performing districts, those with comparatively high enrollment are more likely to exit the bottom 10 percent when SGP is high but overall proficiency is low. Conversely, districts with comparatively low enrollment are more likely to enter the bottom 10 percent because their SGP scores are less likely to fall within a typical median. This can happen even when overall proficiency scores in these districts is comparatively higher. Because SGP is a less reliable measure than proficiency alone, districts that may be due access to more charter schools under the spirit of the law are being denied such access.

BESE could employ two systems: one that conservatively factors SGP into a district's accountability rating, and one that uses absolute proficiency alone (a more reliable and stable measure), to determine which districts are the lowest performing for purposes of the NSS cap.

The analysis has limitations. Notably, the authors do not have access to information about how the state's formula for determining SGP may mitigate (or attempt to mitigate) the random error inherent in SGP calculations and/or the error that can accompany small sample size. Moreover, the conclusions the authors draw about the relationship between SGP and enrollment derive from a very high-level analysis of the publicly available data, with limited information about the precise sample size included in DESE's actual SGP

calculations. The findings of this paper suggest that the state and other independent researchers should further investigate the relationship between aggregate SGP (whether median or mean) and enrollment.

Recommendations

Revert to a formula that uses proficiency in core subjects to determine which districts are subject to an increase in the charter school cap.

The growing body of literature on the student growth percentile suggests it not a reliable enough measure to use for high-stakes policy decisions. Proficiency on standardized tests should not be the only measure a teacher or school uses to understand what students can do or how they are progressing. However, proficiency scores in core subjects are more reliable indicators of how a majority of students are faring in a district and whether that district and its students need more support.

The legislature designed the "smart cap" specifically to provide more charter public schools to families who live in districts where student performance is low. Altering the original formula to include SGP has distorted our understanding of which districts are most in need. Reverting to a formula that uses absolute proficiency as the only measure of achievement can alleviate that distortion. Importantly, identification of the lowest performing districts does not have to exist as part of the state's overall district accountability system. The state could continue to hold districts accountable using both proficiency and growth, while basing eligibility for an increase in the NSS cap solely on proficiency.

Both of the Commonwealth's charter school caps drive potentially high-quality operators out of the state, but the department may be able to proactively redirect prospective operators in the future.

Mitigate the negative impacts of current regulations as much as possible.

In 2014, then-Commissioner Mitchell Chester warned that overemphasizing SGP could distort our understanding of which districts are most in need of support. Since that time, other negative impacts of the new formula for determining the lowest performing 10 percent have become clear, and DESE has tracked those impacts and worked to mitigate them, where possible. The department should continue to proactively assess the financial impact decreased enrollments can have on charter schools that draw students from districts on the cusp of the bottom 10 percent. It should mitigate negative financial impacts as much as possible, as declining enrollment that is a result of the reduction in the NSS cap is out of a school's control. DESE should also work closely with prospective charter operators to identify pockets of need and demand for charters that are less likely to be affected by a volatile NSS cap. Both of the Commonwealth's charter school caps drive potentially high-quality operators out of the state, but the department may be able to proactively redirect prospective operators in the future.

Increase transparency around the meaning and use of SGP.

There is consensus in the research literature that SGP measurements at all levels contain a degree (in some cases a large degree) of statistical error. Some argue that the amount of error is reason to abandon the measure. Others claim that all statistical measures are vulnerable to error and they would rather use SGP, however imprecise, as one of many indicators that can paint an overall picture of student performance for educators and policymakers.

There is consensus in the research literature that SGP measurements at all levels contain a degree (in some cases a large degree) of statistical error. Some argue that the amount of error is reason to abandon the measure.

As this paper points out, SGP may be an informative measure, but it should not be used for high-stakes policy decisions. No matter how it is used, the state should be transparent with all stakeholders about its limitations. SGP alone cannot tell policymakers whether students are receiving an adequate education or equitable educational opportunities. SGP alone cannot accurately tell parents how much knowledge or skill their children have gained in the course of a year. SGP is an estimate—a comparative tool—and DESE should work hard to educate all stakeholders about what it can and cannot tell them about student, school, and district performance.

Endnotes

1. Massachusetts Education Reform Act, Chapter 71, section 89.
2. INSERT DEMOGRAPHIC DATA and CITE RCTS, including Setren (with explanation).
3. Abdulkadiroglu et. al (2009) "Informing the debate: Comparing Boston's charter, pilot, and traditional schools," The Boston Foundation; Angrist et. al (2016) "Stand and Deliver: Effects of Boston's Charter High Schools on College Preparation, Entry, and Choice," *Journal of Labor Economics* 34(2); Center for Research on Educational Outcomes (CREDO) 2013) "National Charter School Study," Stanford University; Setren, Elizabeth (2015) Special Education and English Language Learner Students in Boston Charter Schools: Impact and Classification, Massachusetts Institute of Technology.
4. Center for Research on Educational Outcomes (CREDO) 2013).
5. Massachusetts Education Reform Act, Chapter 71, section 89.
6. "Schools at the Top of the Hill," *The Economist*, February 22, 1997.
7. Massachusetts Education Reform Act, Chapter 70, Section 8.
8. *Ibid*
9. Massachusetts Department of Elementary and Secondary Education, "FY-18 Updated Waitlist Appendix," **http://www.doe.mass.edu/charter/enrollment/FY2018/Waitlist.html**.
10. Massachusetts Education Reform Act, Chapter 12, Section 7(3).
11. Within the first year of the smart cap, the State had filled almost all available charter school seats in the City of Boston, which was subject to a lift of the cap to 18 percent of net school spending. Demand was still so high, that then-Commissioner Mitchell Chester placed a moratorium on the authorization of new charter school seats, so as not to give up all new seats in the first two years of the cap-lift. See: "Massachusetts Department of Elementary and Secondary Education Charter School Fact Sheet."
12. For further opinion on this, see Candal, Cara (2018) *The Fight for the Best Charter Public Schools in the Nation,* Pioneer Institute, Chapter 6.
13. Memo to the Board of Elementary and Secondary Education from Mitchell D. Chester, Commissioner of Education, Massachusetts, June 20, 2014. Note: proficiency is determined using the "average ranks for each district's English language arts, mathematics, and science composite performance index; the percentage of students scoring warning or failing in English language arts, mathematics, and science; the percentage of students scoring advanced in English language arts, mathematics, and science" Massachusetts Education Law and Regulations, Charter Schools, 603 CMI 1.04 (9).
14. Interview with Harneen Chernow, Massachusetts Board of Elementary and Secondary Education, October 5, 2014.
15. Memo to the Board of Elementary and Secondary Education from Mitchell D. Chester, Commissioner of Education, Massachusetts, June 20, 2014
16. Massachusetts Department of Elementary and Secondary Education, "Districts Subject to Increases in the Charter School Cap," "Lowest 10 Percent District Rankings from Prior Application Cycles," **http://www.doe.mass.edu/charter/enrollment/CapIncrease/.**
17. Massachusetts Department of Elementary and Secondary Education, "Districts Subject to Increases in the Charter School Cap," "SY 2019 Analysis"
18. Massachusetts Department of Elementary and Secondary Education, School and District Profiles, Lynn, **http://profiles.doe.mass.edu**
19. Massachusettes Department of Elementary and Secondary Education, "FY-18 Initial Waitlist Appendix," **http://www.doe.mass.edu/charter/enrollment/FY2018/Waitlist.html**.
20. "Charter schools: charter authorizing in Lynn under Net School Spending Caps," Memorandum from Acting-Commissioner Jeff Wulfson to the Board of Elementary and Secondary Education, February 16, 2016.
21. Massachusetts Department of Elementary and Secondary Education, Initial Report on Charter School Waitlists as of March 5, 2018.
22. Wells, Craig S., Sireci, Stephen G., & Bahry, Louise M. 2014 "Estimating the Amount of Error in Student Growth Percentiles," Center for Educational Assessment Research Report N.o. 869, University of Massachusetts at Amherst.
23. Polikoff, Morgan (2017) "Proficiency vs. Growth: Toward a Better Measure," Future Ed Blog, **https://www.future-ed.org/work/proficiency-vs-growth-toward-a-better-measure/**
24. Massachusetts Department of Elementary and Secondary Education (2011) *MCAS Student Growth Percentiles Interpretive Guide*, p. 1.
25. *ibid*
26. Massachusetts's Next-Generation Accountability System, Summer 2018 (Presentation), **www.doe.mass.edu/accountability/accountability-2018.pptx**
27. Massachusetts Department of Elementary and Secondary Education (2011) *MCAS Student Growth Percentiles Interpretive Guide*
28. Massachusetts's Next-Generation Accountability System, Summer 2018 (Presentation), **www.doe.mass.edu/accountability/accountability-2018.pptx**
29. *Race to the Top* was a federal grant program introduced under the Obama administration. States competed for funds based on their willingness to incorporate certain reforms, including charter schools, teacher evaluations, and alternative methods to measure student progress on standardized tests. The Every Student Succeeds Act is the name of the law that governs state access to federal education money. States develop "plans" describing how they will approach certain reforms required in the law and the federal Department of Education must approve those plans for states to receive federal money for schools.
30. Education Commission of the States, "Accountability and Reporting, ESSA Plans," May, 2018, **http://ecs.force.com/mbdata/mbQuest5E?rep=SA172**
31. Sireci, Stephen G., Wells, Craig, S., & Keller, Lisa A. (ND) "Why We Should Abandon Student Growth Percentiles," Research Brief 16.1, Center for Educational Assessment, University of Massachusetts at Amherst, p. 2.
32. Wells, Serici & Bahri (2014); Dimmick, J., Lawrence, B., Mooney, F., Quinn, E, Soraci, A., & Yarmalowicz, M. (2016) "Student Growth Percentiles as Conditional Probabilities", Poster presented at the 30th New England Statistics Symposium, Yale School of Public Health, New Haven, CT; Lash, A., Makkonen, R., Tran, L., & Huang, M. (2016). *Analysis of the stability of teacher-level growth scores from the student growth percentile model* (REL 2016–104). Washington, DC:

U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory West. Retrieved from: **http://ies.ed.gov/ncee/edlabs**;

33. Castellano, Katherine E., and Lockwood, J.R.(2017) "Estimating Pure Student Growth Percentile Distributions Using Latent Regression Multidimensional IRT Models, *Educational and Pyschological Measurement*, 77(6) 917–944.
34. Castellano, K. E., and A. D. Ho. 2015. "Practical differences among aggregate-level conditional status metrics: From median student growth percentiles to value-added models." *Journal of Educational and Behavioral Statistics* 40 (1): 35–68.
35. Culbertson, Michael J. (2016) "Precision of Student Growth Percentiles with Small Sample Sizes," Regional Educational Laboratory Central.
36. In the analyses below the authors use data publicly available from the Massachusetts Department of Elementary and Secondary Education.
37. All of the data specific to Massachusetts provided in this report come from data published by the Massachusetts Department of Elementary and Secondary Education, specifically its School and District Profiles and "SY 2019 Analysis of Districts Eligible for an Increase in the Charter School Cap."
38. This statement derives from a conversation with Lisa Keller, Ed.D and Craig Wells, Ph.D, December 18, 2018.

About the Author

Dr. Cara Stillings Candal is Director, Educational Opportunity at the Foundation for Excellence in Education and a senior fellow at Pioneer Institute. She was formerly research assistant professor and lecturer at the Boston University School of Education and a founding team member of the National Academy of Advanced Teacher Education. Candal holds a B.A. in English literature from Indiana University at Bloomington, an M.A. in social science from the University of Chicago, and a doctorate in education policy from Boston University. She is the author/editor of several books, including *The Fight for the Best Charter Public Schools in the Nation*.

Dr. Bruce Bean has been involved in various forms of data analysis for over forty years, creating custom software for many uses. Working as an accountant, he created a complete custom software general ledger package and other software programs. As the founder of his own economic statistics software company, he created custom software that contained over 14,000 economic and financial time series and was used nation wide. For the past twenty-one years, he has been the data manager of The Community Group in Lawrence, MA, where he has been involved in analyses of MA Department of Elementary and Secondary Education and local school data, both for in-house and consulting. He earned a Ph.D. in Physics from the University of New Hampshire.

Rebekah Paxton is a Research Assistant at Pioneer Institute. She first joined Pioneer in 2017 as a Roger Perry intern, writing about various transparency issues within the Commonwealth, including fiscal policy and higher education. Since then, she has worked on various research projects under PioneerPublic and PioneerOpportunity, in areas of state finance, public policy, and labor relations. She recently earned an M.A. in Political Science and a B.A. in Political Science and Economics, from Boston University, where she graduated summa cum laude.

About Pioneer

Pioneer Institute is an independent, non-partisan, privately funded research organization that seeks to improve the quality of life in Massachusetts through civic discourse and intellectually rigorous, data-driven public policy solutions based on free market principles, individual liberty and responsibility, and the ideal of effective, limited and accountable government.

PIONEER INSTITUTE
PUBLIC POLICY RESEARCH

185 Devonshire Street, Suite 1101 Boston MA 02110 617.723.2277 617.723.1880
www.pioneerinstitute.org Facebook.com/PioneerInstitute Twitter.com/PioneerBoston