PACE
Policy Analysis for California Education

WORKING PAPER

Who's Missing? Exploring the Magnitude and Impact of Student Opt-Outs on School Accountability Systems

The number of students opting out of standardized tests has grown in recent years. This phenomenon represents a potential threat to our ability to accurately measure student achievement for schools and districts. This working paper documents the extent to which opting out is observed in the CORE districts. It then models the extent to which various accountability measures would be impacted by growth in the rate of opting out. The growth of opting out could significantly impact some accountability measures in use in California, but the CORE growth measure is largely unaffected. In contrast, accountability metrics that track student achievement by cohort are at risk of becoming biased even with relatively low absolute levels of opting out, and districts should consider explicitly adjusting for the characteristics of the students that actually sit for tests when designing school accountability systems.

Edward Cremata
Aspire Public Schools

VERSION: September 2019

Acknowledgements

This paper was produced as part of the CORE-PACE Research Partnership, which is focused on producing research that informs continuous improvement in the CORE districts (Fresno, Garden Grove, Long Beach, Los Angeles, Oakland, Sacramento City, San Francisco, and Santa Ana unified school districts) and policy and practice in California and beyond. This research was possible through the partnership support from the S. D. Bechtel, Jr. Foundation. Finally, we thank the CORE Districts for partnering with us in this research, providing the data for this study, and giving feedback on earlier drafts of this study. PACE working papers are circulated for discussion and comment purposes and have not undergone the peer-review process that accompanies official PACE publications.

The proportion of students opting out of standardized tests has grown substantially across the United States in recent years (Thompson, 2015). There is evidence that this phenomenon has seen a recent surge in popularity among parents for a variety of reasons, including concerns surrounding implementation of the common core standards and disillusionment with the federal emphasis on standardized testing and recent approaches to education reform more generally (US News, 2015; West et. al, 2017). According to media reports, states across the country have been faced with growing numbers of students choosing to opt out of standardized tests, which the federal government has warned could cost them millions of dollars in federal funding as a result of their failure to comply with ESSA-mandated student testing participation rates (Garcia, 2017).

The growth of the opt out movement could have a direct impact on the ability of districts and states to use student testing data in multiple important ways, including for school accountability systems populated by test scores in use in all fifty states and the District of Columbia (ECS, 2017). New York State in particular has struggled to evaluate schools using existing accountability rubrics as opt out rates have grown to over 20%, but New York is not alone; growth in the number of students skipping standardized tests has also been observed in large and diverse states like Florida, North Carolina, and Colorado (EdWeek, 2015). For policymakers in states that are largely unaffected at the moment, it is important to note that support for opting out among parents is much higher than the current observed levels of opting out across the country, as discussed on more detail below.

In this paper, I assess two important aspects to help ascertain the potential impact of the opt-out movement on district performance under state accountability systems that rely on standardized achievement tests to generate accountability metrics. As a test case for the real-world national implications of opting out, this evaluation investigates the magnitude and accountability system implications of student opt outs in the CORE collaboration, a consortium of multiple large California school districts[1].

One important way that students' failure to take standardized assessments may impact school districts is by biasing accountability systems intended to assess student achievement. Most state accountability systems compute average student proficiency or growth based on students' standardized test scores, and then use these measures as a measure of school quality, which is then used to assess school district achievement and progress. In California, as is the case across the country, student test results from the state's standardized testing regime (the California Assessment of Student Progress and Performance (CAASPP)) are currently used in school accountability metrics and may become unrepresentative of their schools and districts if a large enough proportion of students choose not to be tested. This is a particular concern given evidence that students are non-randomly opting out of testing, with minority students and students in poverty making up a smaller share of students opting out (Thompson, 2015). To assess the degree to which accountability metrics may be biased by opting out, I first ask: *What*

[1] The CORE districts include Fresno, Garden Grove, Los Angeles, Long Beach, Oakland, Sacramento, San Francisco, and Santa Ana.

is the current magnitude and demographic composition of students opting out in the evaluated California districts?

New school accountability measures in use across the United States typically include broader measures of student performance than previous No Child Left Behind and California Department of Education mandated school report cards. For example, the CORE districts used in this analysis use a measure of school performance called the School Quality Improvement Index (SQII)[2]. While the SQII includes a broad range of metrics to track student and school performance, it does include multiple performance measures that rely on standardized test score results, such as a student percentile growth model and the proportion of students meeting grade-level standards in math and ELA. As of 2016, 38 state's education accountability systems include measures of student test performance (Data Quality Campaign, 2016). Identifying and quantifying the myriad channels through which student opt outs may affect student and school accountability measures is of national importance given the prevalence of these systems and the weight placed on them when making critical policy and staffing decisions.

The potential for the opt out movement to bias school accountability systems is not simply a matter of academic interest. In the case of California districts, empowered by a law passed in California in 2013 (Lexis Nexis, 2013), evaluation rubrics will be used to set statewide performance goals, and elements of the new accountability system can be used to justify intervention into chronically underperforming schools and districts. Whether one believes in the use of student testing data primarily for the purposes of accountability or to develop and improve interventions designed to impact student achievement, if only a self-selected and non-random subset of students is used to represent the performance of a classroom, school, or district, this could hamper our ability to effectively implement, evaluate, and refine education policy and practice moving forward. To begin to assess the impact of student opt-out behavior on district performance on state accountability systems, I thus ask my second research question: *What is the impact of student opt outs on state accountability systems under a range of opt out growth scenarios similar to what has been observed in other states?*

Primary results suggest that, while the current incidence of opting out in the included CORE districts is fairly low (0.2% to 1%)[3], further growth could substantially impact districts' ability to accurately measure school performance and achievement gaps. This is due primarily to two factors; the non-representative demographic make-up of students opting out, and the concentration of opt out students by school. Regarding the potential for growth in opting out to impact school accountability systems, there is evidence that a school's absolute and ordinal (ranking) performance will be impacted by plausible opt out growth scenarios, assuming the

[2] See Figure A1 for a detailed breakdown of the measures used in SQII.

[3] These figures are based on the indicator provided by the districts, and is likely biased downwards because it only captures instances where parents formally requested and signed a waiver allowing their students to opt out of standardized testing. California is one of the few states that allows students to opt out of state-mandated standardized testing without the need to submit a formal request (CTA, 2017).

current demographics of opt out students sustain. In contrast, measures that explicitly control for the composition on test takers, such as the SQII Growth Index, are far less affected by opting out.

Under a range of reasonable opt out growth scenarios, specific test-score based components of the analyzed district's school-level accountability systems may be altered in meaningful ways. If the proportion of students opting out grows, as the experience of other states suggests is plausible, while demographic patterns remain similar, dozens of schools could see themselves fall below district-established benchmarks of performance. This occurs due to the aforementioned distribution of student's choosing to opt out, as students more likely to opt out of testing (e.g. White and non-poverty students) also tend to be those with higher test scores on average. This suggests that even "holistic" accountability policies designed to reflect broader school-level contributions to student growth are susceptible to opting out, so long as they include test scores or other correlated measures.

The following section reviews the existing literature on the composition of students choosing to opt out and the potential implications for accountability systems currently in use across the country. The next section describes the data used in this analysis, and is followed by an explanation of the methods used to address each research question. The proceeding section presents results for each of the research questions, and the final section addresses implications of these findings for school and district personnel.

Literature Review

In this section I review the existing literature on the driving factors behind the growing opt out movement, and follow up with a review of the literature on the potential impacts of opting out on school and teacher accountability systems.

How Common is Opting Out?

The rate of opting out is highly variable across the country (Ujifusa, 2015). According to a 2015 report issued by the U.S. Department of Education, 13 states missed the 95% student participation requirement in the 2014-15 school year, in part due to the growth in opting out (Ujifusa, 2015). Certain states, including California (3%), Idaho (2%), and Oregon (4%), reported relatively low levels of opting out in 2015. In contrast, states like Colorado (11%), Rhode Island (11%), and New York (20%) experienced rates of opting out in 2015 far above the federal threshold of 95% test compliance (Bennet, 2016). It is important to note, however, that even at relatively low rates of opting out, such as those seen in California, opting out could still pose a problem if the incidence of opting out is clustered within particular classrooms, schools, or districts. Evidence of wide variation in opt out rates by district can be found in New York state, where Chateaugay, Rocky Point, and Onteora school districts reported rates of opting out of 90%, 80%, and 66%, respectively (Bennet, 2016). In contrast, opt out rates in New York City

were only 1.4% in 2015, consistent with the average opt out rates observed in the largest 66 urban school systems of approximately 1% (CGCS, 2015).

Who Opts Out of Tests and Why?

Parents and students are driven by a range of beliefs leading them to opt out of standardized tests. These include the belief that education reform has come to focus too strongly on standardized tests, as well as opposition to adoption of the common core standards in most states (Bidwell, 2015; Strauss, 2015). Other parents are motivated by concerns over student privacy and federal standards regarding the storage and sharing of sensitive student information online (USDOE, 2014; Thompson, 2015). The average parent who has their students opt-out does not reflect the average public school population. According to a recent national survey on the beliefs of parents and students in the opt out movement, the average opt out activist is a "highly educated, white, married, politically liberal parent whose children attend public school and whose household median income is well above the national average." (Pizmony-Levy & Saraisky, 2016) In addition, most members of the opt out movement have joined in the last four years, with nearly half (49%) first deciding to opt out in the last two years, suggesting that this phenomenon is accelerating in growth. Similar to other narrower studies of opt out activists, Pizmony-Levy and Saraisky find that common parental motivations include opposition to high-stakes testing as well as broader opposition to current educational reform efforts (Pizmony-Levy & Saraisky, 2016).

The results from national surveys are also corroborated by student testing data; the demographics of students opting out of standardized tests are not representative of their schools and districts (Harris, 2015). For example, in New York state students without a valid reason to miss assessments (a proxy for opting out) are much more likely to be White, and to not be economically disadvantaged or English language learners (NYSED, 2015; Ujifusa, 2015). Similarly, students opting out in Colorado were also more likely to be White and less likely to be eligible for free and reduced price lunches (CDOE, 2015). The lower number of economically disadvantaged students among those opting out has also been noted in Washington State (Parr & Teed, 2015).

High Levels of Support for Opting Out Among Those that Don't Opt Out

Of potential concern to policymakers and testing advocates, support for opting out among parents is far higher than the current levels of opting out across the country. National survey data shows that 44% of White respondents support the rights of parents to opt their students out of testing, compared to 35% of Hispanics and only 28% of Blacks supporting opting out (PDK/Gallup, 2015). When asked if they would exclude their own student from standardized testing, 34% of Whites, 28% of Hispanics, and 21% of Blacks said they would choose to do so (PDK/Gallup, 2015). These results display not only the wide demographic differences in support for opting out but also the potential for opt out rates to continue to grow, with nearly 30% of

parents expressing a desire to opt their own students out of testing, far above current levels (PDK/Gallup, 2015).

Differences in the proportion of parents supporting student opt outs are also related to general levels of support for standardized testing by race and income (NORC, 2015; Tompson et. al, 2013). White parents generally place the least amount of faith in test scores as a measure of school quality compared to Black and Hispanic parents (PDK/Gallup, 2015). With regard to household income, 85% of parents earning less than 50,000 dollars per year said that regular student assessment is either important or extremely important, compared to 63% of parents earning over 100,000 dollars per year (Tompson et. al, 2013). The demographic differences in opt out prevalence suggest the existence of a positive relationship between test scores and the likelihood of opting out, but the evidence for this presumption is not so clear cut, as presented in the next section.

What Is the Impact of Opting Out on Accountability Systems?

As discussed above, students that opt out of testing tend to come from districts and demographic groups associated with higher levels of performance. However, there is evidence that there are countervailing forces that lead the relationship between test taking and performance to be more complicated than the demographic proportions would suggest (Ujifusa, 2015). For example, in New York state students opting out of tests in 2015 were slightly more likely to not have achieved proficiency in the previous year (NYSED, 2015). When controlling for demographics, a negative relationship between proficiency and test taking was reported in 648 of the 695 schools districts in New York (Chingos, 2015). In contrast, opting out was still found to be less common in economically disadvantaged districts, complicating the picture with respect to the potential implications for accountability systems (Chingos, 2015). The results from New York suggest that student-level analyses of academic achievement may be biased upward by the absence of students opting out (as they typically control for student demographics), while more aggregated accountability metrics (e.g. school and district measures), particularly those that do not control for race and income, are likely to be biased down as the proportion of students opting out of tests continues to grow.

Given the ambiguous relationship between opting out and measured proficiency, it is important to estimate the actual impact of student opt outs on accountability measures. A recent contribution to this literature ran multiple simulations to estimate the impact of opting out on the Pennsylvania school accountability system (Research for Action, 2014). According to their analysis, as few as 11 high performing students (those scoring proficient or above) opting out of test taking in a school of 200 students could push their school into an unsatisfactory performance bracket under the Pennsylvania System of School Assessment. Furthermore, as few as 20 low-performing students (those scoring basic or below) opting out of testing could cause the same school to artificially move up one level in the same rating system (Research for Action, 2014). As the authors point out, movement in either direction could complicate existing

accountability systems as well as federal funding formulas which take into account school performance (ESSA, 2015).

Teacher accountability systems appear to be less susceptible to opting out for a variety of reasons. Among these reasons are the fact that the construction of teacher accountability metrics, such as value-added measures (VAMs) of teacher contribution to student learning, explicitly take into account the demographic make-up of their classrooms. As discussed above, controlling for the particular types of students self-selecting out of testing may help to eliminate bias, particularly the inclusion of prior test score results and measures of race and family income. By adjusting for changing demographics in a teacher's classroom (e.g. race, free-reduced price lunch status, special education status, etc.), teacher VAMs are adjusted for the types of students being taught, limiting the impact of opting out on their results (Chingos, 2015). According to simulations run in North Carolina, average teacher VAMs change very little if five students per class opt out of taking tests, which accounts for 1/5th of the typical class in that state. In fact, VAMs calculated before and after the removal of five students per class showed a correlation of 0.97, demonstrating the success of controlling for student demographics to account for changes in the composition of test takers (Chingos, 2015). However, it should be noted that these students were randomly dropped from being included in teacher VAM calculations, which is not reflective of the reality of opting out in every state that has been evaluated. That said, teacher accountability systems typically have very low thresholds for performance before disciplinary action is taken (e.g. 1.5 standard deviations below the mean of expected performance), minimizing the real world impact of opting out on teacher-level accountability systems. It is for these reasons, as well as the abandonment entirely of teacher VAMs in many states and districts across the country (ASA, 2014), that this analysis is focused on the impact of opting out on school accountability systems. The next section presents the data used to address the research questions presented above.

Data

This analysis uses student level administrative data from multiple CORE districts. Student test score results on the Smarter Balanced Assessment Consortium-aligned California Assessment of Student Performance and Progress (CAASPP) are available for the academic years 2014-15 and 2015-16, as well as student demographic and enrollment information. School-level demographic information is supplemented where necessary using publicly available data from the Common Core of Data, available through the National Center for Education Statistics (NCES, 2017). Table 1 provides student demographic information for the students included in this analysis.

In total, there are 146,131 students in 2014-15 and 169,124 students in 2015-16 across all included districts. In 2014-15, 695 students are indicated as having a parent or guardian exempt them from testing, growing to 841 students in 2015-16, representing under 1% average opt out rates across all districts in each year. A majority of students in the included CORE districts under study are under-represented minorities (Black and Hispanic) (74%), while an average of 34% are

designated English language learners (ELL) and 10% have an individualized education program (IEP) or special education designation. In contrast, among opt out students, only 59% are under-represented minorities, 21% are designated ELL, and 9% receive special education services. These findings are broadly consistent with the demographics of opt out students found in other states (e.g. Harris, 2015; NYSED, 2015; Ujifusa, 2015). In 2014-15, but not in 2015-16, there are large differences between tested students and those that opted out in the number of reported absences, with opt out students reporting an average of 15 days absent compared to only eight for tested students. The additional seven days absent recorded for opt out students may reflect in part their absence during the five-day testing period in 2014-15, although the actual amount of time dedicated to testing varies by district. Finally, we see that the incidence of opting out varies by year and district, from a low of 0.1% in 2014-15 to a high of 1.1% in 2015-16, with some schools reporting opt out rates of greater than 10%.

Table 1. Summary Statistics for Students & Schools in Included Districts, 2014-15 & 2015-16

		2014-15		2015-16		2014-15 & 2015-16	
	All Students	Tested	Opt Outs	Tested	Opt Outs	% Opting Out	Achievement Level (1-4)
	472,819	146,131	695	169,124	841	0.3%	2.09
Student Demographics							
White	23%	22%	41%	23%	41%	0.4%	2.35
Black	16%	14%	16%	14%	10%	0.3%	1.88
Hispanic	59%	62%	43%	62%	36%	0.2%	1.95
Asian	11%	11%	8%	11%	9%	0.2%	2.42
Native American	1%	1%	2%	1%	2%	0.6%	2.15
Pacific Islander	2%	1%	1%	1%	1%	0.1%	2.05
Filipino	2%	2%	1%	2%	1%	0.3%	2.66
Female	50%	50%	48%	50%	47%	0.2%	2.22
Free/Reduced Price Lunch	77%	76%	50%	78%	54%	0.2%	1.96
Special Education	10%	9%	9%	12%	9%	0.7%	1.46
English Language Learner	36%	36%	16%	35%	27%	0.0%	1.70
Student Attendance							
Avg # Days Enrolled	173	176	174	171	155	n/a	n/a
Avg # Days Absent	7	8	15	7	8	n/a	n/a
Avg # Days Attended	166	168	159	164	147	n/a	n/a
Avg # Suspensions	1	1	0	2	2	n/a	n/a

Note: Student ethnicity in one large district over-identified American Indian students far beyond what is reported in publicly available sources. This was primarily an issue for students identified as having multiple ethnicities. To address this issue, students are only treated as American Indian when that is their only racial designation. This reduces the number of American Indian students from 20% to 1% in the pooled data set.

This analysis investigates the impact of opting out on each element of the SQII that could be affected by changes in the proportions of students sitting for tests. These include the

Performance Index, a normalized 1-10 rating assigned to each school based on the proportion of students proficient in math and ELA. The school's raw proficiency rate is also analyzed along with the normalized proficiency index. The Growth Index is a student growth model, which compares the actual achievement of students with their expected achievement based on their demographics. Another component of SQII captures the extent to which students are on track to graduate, based on their GPA, completion of required courses, and chronic absences. The decision to opt out of standardized tests does not directly impact a student's GPA or ability to graduate (which make up the majority of this indicator), however it possible that a student's identification as chronically absent[4] (which is a part of the high school readiness measure) could be impacted by a student's decision to opt out of testing without an excused absence.

In addition to the specific components of SQII, this analysis also estimates the impact of opting out on other school-level measures of academic achievement. One measure investigated is the proportion of schools that move in the distribution of final school-level rankings (1 - 10) based on different opt out scenarios. The impact of opting out on measured school-wide average proficiency rates (combined math and ELA proficiency) is also estimated, as this is a common component of other state's and district's school accountability systems (Data Quality Campaign, 2017). Finally, the average standardized math and ELA performance of the included districts is calculated under the same range of opt out scenarios. To accurately simulate the impact of opt our growth, it is important to identify the types of students currently opting out of testing in the analyzed districts.

Identifying the magnitude and demographics of the opt out movement is aided by the inclusion of an indicator in the test score data marking student records for whom a parent or guardian waived their participation in testing in each year. By connecting this data with student, school, and district demographic data, I can identify the magnitude of opting out, both overall and by district, as well as the demographics of students that did and did not opt out of testing in 2014-15 and 2015-16. However, the opt out indicator provided by the districts likely represents a lower bound estimate of the true opt out rate, as it only reflects the proportion of students whose parents submitted a formal request to opt their child out of standardized tests to their school. A separate opt out measure was generated using existing enrollment and administrative student data, but is not included here because there is an insufficient level of data granularity available to be confident that this measure is accurate enough for inclusion in this analysis[5]. Specifically, data are not yet available identifying students' granted excused absences that happened to coincide with the testing period.

Using this method, the estimated opt out rate across all districts is 3.6%, which is substantially higher than that suggested by the "parent test waiver" indicator used in the data. This is also in line with other estimates of the incidence of opting out in California. Despite the

[4] As defined by being absent greater than 30 days within the district.

[5] The process involved identifying students as "opt outs" only if they are enrolled in a tested grade with good attendance records, do not have a potential waiver to participate in testing based on a disability, and are present in the district during the period of evaluation.

fact that I can only simulate the impact of opting out based on the demographics of students flagged with the provided "lower bound" opt out estimate, there is reason to believe these findings will be similar to the results had we been able to use the "upper bound" estimate as well, This is due to the similar demographic proportions of students identified as opting out under each flag. While the exclusion of the upper bound estimate means that this analysis under-reports the true rate of opting out in the included districts, simulations of the impact of future growth in opting out are likely only minimally affected, as it is the proportions of different types of students indicated as having opted out that drives those results. The specific process by which these simulations are conducted is discussed in detail in the next section.

Methods

To address the first research question, I compare the magnitude and proportions of students opting out of standardized tests across each of the districts in this study. Logistic regression models are used to identify which student characteristics are associated with a student's decision to opt out of standardized testing using the following approach:

(1) $OptOut_{\text{ist}} = \beta_0 + \beta_1 Student_{\text{ist}} + \beta_2 SchLevel_{st} + \varepsilon_{\text{ist}}$

where $OptOut_{\text{ist}}$ is a binary variable equal to 1 when student *i* opts out of testing in school *s* in year *t*, $Student_{\text{ist}}$ represents a vector of student demographics, and $SchLevel_{st}$ controls for school level (middle and high school), with elementary as the excluded category. All models are run with robust standard errors. Additional specifications are run that exclude school level as an independent variable, as well as models including student achievement as an additional control variable.

To answer research question two, I use the coefficients from equation 1 to predict each student's likelihood of opting out ($Predicted\_OptOut_{\text{ist}}$) and save these estimates. These predicted values are then used to inform the population of students excluded as higher opt out rates are simulated. For example, to simulate the impact of a 5% opt out rate, we want to remove from our data the 5% of students deemed most likely to opt out. Our best approximation of this scenario is to remove the 5% of students with the highest predicted likelihood of opting out, and to then recalculate accountability measures in their absence (because they would not have provided test results for inclusion). To simulate the impact of a 10% opt out rate, the 10% of students with highest predicted opt out probabilities are dropped, and so on up to a simulated 20% total opt out rate. The simulated opt out rates are chosen to reflect the experiences of a broad range of states, including states with relatively low opt out rates (e.g. Idaho – 2%, California – 3%, Oregon – 4%) to states with moderate (e.g. Colorado – 11%, Rhode Island – 11%) and high (e.g. New York – 21%) opt out rates. Regarding the demographic composition of those opting out, it is important to note that just because the demographics of those currently opting out are comparable within California and follow the patterns observed in other states, it is possible that the demographics of students opting out

could change as the incidence grows, in which case the estimates included here may not be valid.

To identify the impact of opting out on school-level accountability measures broadly, I simulate the value of each element of the SQII that uses test scores assuming opt out rates of 0%, 5%, 10%, 15%, and 20% levels. These levels are chosen to represent the current reasonable range of opt out rates experienced across the country, from the current low levels observed in California on up to the high levels seen in New York. These include the performance and growth indices for math and ELA (which represent the performance of each school's students and is scaled from 1-10), as well as the raw proficiency rates in math and ELA. I also calculate the proportion of schools that would fall into a lower level of performance tier (ranges from 1 – 5), and the average school-level drop in standardized achievement in math and ELA under the same opt out scenarios. The results of this evaluation are discussed in the next section.

Results

Research Question 1 – What Predicts Opting Out in the Included Districts?

As can be seen in Table 1, a small number of students currently opt out of standardized tests in the analyzed California districts according to district reports, equal to 0.1% to 1% in each district and year. Simple demographics comparisons suggest that, similar to findings across the country, students that opt out in California are more likely to be White, less likely to be Hispanic or English Language Learners (ELL), and less likely to qualify for free or reduced price lunches. When I predict each student's propensity to opt out as a logistic function of their demographics, as is seen in Table 2, we see that Asian (Odds Ratio (OR) = 0.49, $p < 0.001$), Pacific Islander (OR = 0.31, $p < 0.001$), Black (OR = 0.78, $p < 0.01$), and Hispanic (OR = 0.33, $p < 0.001$) students are all significantly less likely to opt out of standardized tests compared to their White counterparts. Controlling for school level does not meaningfully alter these results. Interestingly, while qualifying for free and reduced lunches remains negatively associated with the likelihood that a student opts out (OR = 0.51, $p < 0.001$), once we control for other student demographics being an English Language Learner becomes positively associated with opting out (OR = 1.27, $p < 0.01$).

Students identified as having a disability are substantially more likely to opt out (OR = 4.89, $p < 0.001$), which raises the concern that these students may be driving the results. To test this hypothesis, the specific disability flag for each students is identified, which can be seen in Appendix Table 1. As can be seen in Appendix Table 1, the majority of students identified as disabled are noted as having a speech or language impairment (18%), an emotional disturbance (3%), a specific learning disability (54%), or a non-specific health impairment (12%). We can also see that the incidence of opting out varies widely among identified disability flags, with students identified as having an emotional disturbance in particular being much more likely to opt out of testing, representing 3% of disabilities among all tested students but 25% of opt out

students with disability flags. Fortunately, as can be seen in Table 3, the results are not especially sensitive to the inclusion of our disability indicator.

The relationship between student demographics and the likelihood of opting out is also robust to the exclusion of indicators for free and reduced price lunch qualification and identification as an English Language Learner (Table 3), as well as to the inclusion of student's prior achievement as an explanatory variable (Table 4). The positive relationship between a student's prior achievement and their likelihood of opting out is interesting, as it suggests that opting out isn't being driven primarily by school administrators "pushing out" low performing students to deliberately alter the composition of tested students, as has been alleged[6]. Based on the robustness of the relationship between student demographics and the propensity to opt out across multiple specifications, the baseline estimates from column 2 in Table 2 (logistic regression with all student controls) are used to conduct the simulations discussed below.

[6] There is little formal evidence on this phenomenon, but media reports have identified isolated cases of school leaders informing low-performing students of their right to opt out to manipulate their test results (e.g. Balingit, 2016).

Table 2. Relationship between Opting Out & Student Characteristics - Estimates from Linear Probability and Logistic Regression Models

	Student Demographics		Students Demographics w/ School Level Indicator	
	LPM	**LOGIT**	**LPM**	**LOGIT**
Student characteristics				
Asian	-0.003***	0.49***	-0.003***	0.50***
	0.0004	0.07	0.0004	0.07
Pacific Islander	-0.004***	0.31**	-0.004***	0.31**
	0.0009	0.13	0.0009	0.13
Filipino	-0.001	0.8	-0.001	0.81
	0.001	0.16	0.001	0.16
Black	-0.002***	0.78**	-0.002***	0.79**
	0.0004	0.08	0.0004	0.081
Hispanic	-0.004***	0.33***	-0.004***	0.33***
	0.0003	0.04	0.0003	0.04
Parent Not High School Grad	-0.0001	0.92	-0.0001	0.92
	0.0002	0.11	0.0003	0.11
Male	-0.0005**	0.86**	-0.0004**	0.86**
	0.0002	0.06	0.0002	0.06
Free/Reduced Price Lunch	-0.003***	0.51***	-0.002***	0.51***
	0.0003	0.04	0.0003	0.04
Homeless	-0.001**	0.61**	-0.001**	0.61**
	0.0004	0.12	0.0004	0.12
English Language Learner	0.0003	1.27**	0.0001	1.23**
	0.0003	0.13	0.0003	0.13
Special Ed	0.008***	4.89***	0.008***	4.79***
	0.0003	0.36	0.0003	0.36
Middle School			-0.0008***	0.76**
			0.0002	0.06
High School			-0.0008**	0.77**
			0.0003	0.08
R^2	0.004	0.07	0.004	0.07
N students	266,641	266,641	266,641	266,641

Note: Results for logistic regression models are expressed as odds ratios for ease of interpretation. *** p < 0.01, ** p < 0.01, * p < 0.05.

Table 3. Relationship between Opting Out & Student Characteristics – Additional Specifications Excluding Select Covariates

	All Variables	No FRL Indicator	No Disability Indicator	No ELL Indicator
	LOGIT	LOGIT	LOGIT	LOGIT
Student characteristics				
Asian	0.49***	0.42***	0.41***	0.53***
	0.07	0.06	0.06	0.07
Pacific Islander	0.31**	0.25**	0.26**	0.31**
	0.13	0.1	0.11	0.13
Filipino	0.8	0.76	0.68	0.79
	0.16	0.16	0.14	0.16
Black	0.78**	0.60**	0.87	0.76
	0.08	0.06	0.09	0.08
Hispanic	0.33***	0.26***	0.30***	0.39***
	0.04	0.03	0.03	0.04
Parent Not High School Grad	0.92	0.77**	0.89	1.03
	0.11	0.09	0.10	0.11
Male	0.86**	0.86**	1.04	0.88
	0.06	0.06	0.07	0.06
Free/Reduced Price Lunch	0.51***		0.54***	0.50***
	0.04		0.05	0.04
Homeless	0.61**	0.58**	0.61**	0.58**
	0.12	0.11	0.12	0.11
English Language Learner	1.27**	1.11	1.42**	
	0.13	0.11	0.14	
Special Ed	4.89***	4.73***		4.63***
	0.36	0.35		0.34
R^2	0.07	0.06	0.03	0.06
N students	266,641	266,641	266,641	266,641

Note: Results for logistic regression models are expressed as odds ratios for ease of interpretation. *** $p < 0.01$, ** $p < 0.01$, * $p < 0.05$.

Table 4. Relationship between Opting Out & Student Characteristics Including Student Achievement

	Student Demographics		Student Demographics w/ School Level Indicator	
	LPM	**LOGIT**	**LPM**	**LOGIT**
Student characteristics				
Average Scale Score	-0.0001***	0.18***	-0.0001***	0.16***
	0.00002	0.06	0.00002	0.06
Asian	-0.003***	0.49***	-0.003***	0.50***
	0.0004	0.07	0.0004	0.07
Pacific Islander	-0.004***	0.31**	-0.004***	0.31**
	0.0009	0.13	0.0009	0.13
Filipino	-0.001	0.8	-0.001	0.81
	0.001	0.16	0.001	0.16
Black	-0.002***	0.78**	-0.002***	0.79**
	0.0004	0.08	0.0004	0.081
Hispanic	-0.004***	0.33***	-0.004***	0.33***
	0.0003	0.04	0.0003	0.04
Parent Not High School Grad	-0.0001	0.92	-0.0001	0.92
	0.0002	0.11	0.0003	0.11
Male	-0.0005**	0.86**	-0.0004**	0.86**
	0.0002	0.06	0.0002	0.06
Free/Reduced Price Lunch	-0.003***	0.51***	-0.002***	0.51***
	0.0003	0.04	0.0003	0.04
Homeless	-0.001**	0.61**	-0.001**	0.61**
	0.0004	0.12	0.0004	0.12
English Language Learner	0.0003	1.27**	0.0001	1.23**
	0.0003	0.13	0.0003	0.13
Special Ed	0.008***	4.89***	0.008***	4.79***
	0.0003	0.36	0.0003	0.36
Middle School			-0.0008***	0.76**
			0.0002	0.06
High School			-0.0008**	0.77**
			0.0003	0.08
R^2	0.004	0.07	0.004	0.07
N students	266,641	266,641	266,641	266,641

Note: Results for logistic regression models are expressed as odds ratios for ease of interpretation. *** p < 0.01, ** p < 0.01, * p < 0.05.

Research Question 2: What is the Impact of Opt Out Growth?

Table 5 provides the estimated performance of schools on all academic components of the SQII under a range of scenarios up to 20% total opt out rates, based on each student's predicted likelihood of opting out from equation 1 without school-level controls (i.e. only student demographics). In the first and final two rows, we see that the performance measures based on proficiency are most strongly affected by growth in opting out. Specifically, the Performance Index falls up to 0.5 points in math and 0.6 points in ELA on a 10 point scale, while the proportion of students proficient at grade level across all districts falls from 48% to 43% in math and from 42% to 38% in ELA assuming opt out rates of 20%. The growth measures included in SQII are impacted as well, although not to the same degree. As explained above, the SQII uses a normalized index capturing student growth, which accounts for student demographics, minimizing (but not eliminating) the impact of opting out on this measure. We see that the growth index, which ranges from 1 – 10, falls from 5.3 to 5.2 in math and actually holds steady at 4.7 ELA assuming opt out rates of up to 20%.

Table 5. Simulated School Quality Improvement Index (SQII) Ratings Based on Various Opt Out Growth Scenarios at Current Demographics

SQII Academic Components	Estimated Performance Based on Opt Out Rates of …				
	0%	5%	10%	15%	20%
Performance Index					
Math Performance Index Level (1-10)	5.9	5.7	5.6	5.5	5.4
ELA Performance Index Level (1-10)	4.8	4.6	4.4	4.3	4.2
Growth Index					
Math Growth Index (1 - 10)	5.3	5.4	5.3	5.2	5.2
ELA Growth Index (1 - 10)	4.7	4.7	4.6	4.8	4.7
On Track to Graduate					
High School Readiness Index	7.0	7.0	7.0	7.0	7.0
Proficiency Rate					
% Proficient in Math	48%	46%	45%	45%	43%
% Proficient in ELA	42%	41%	39%	39%	38%

Continued growth in the proportion of students opting out could also impact school level measures of achievement, as can be seen in Table 6. Given that higher performing students are more likely to opt out of standardized tests, their absence can hurt school-level average performance measures. For example, as can be seen in the first row of Table 6, using the same method to simulate opt outs as before, up to 4% of schools could see their average performance fall into a lower tier due solely to changes in the composition of test takers in their school. School-wide average proficiency rates could also fall from the current level of 41% to as low as 36% assuming 20% opt out rates at current demographics. As can be seen in rows 3 and

4 of Table 6, Standardized school-wide math and ELA performance may also fall as much as 10% of a standard deviation in math and 5% of a standard deviation in reading. These results make clear the potential impact of opting out on accountability measures currently in use in the included districts, and the implications of these results are discussed in the next section.

Table 6. Simulated School Level Performance Based on Various Opt Out Growth Scenarios at Current Demographics

School Level Academic Measures	Estimated Performance Based on Opt Out Rates of …				
	0%	5%	10%	15%	20%
Performance					
% Schools that move into Lower Performance Tier	n/a	1%	3%	3%	4%
School-Wide Proficiency Rate	41%	39%	38%	38%	36%
Avg. School-Level Math Score (s.d.)	0	-0.04	-0.06	-0.07	-0.1
Avg. School-Level ELA Score (s.d.)	0	-0.01	-0.02	-0.04	-0.05

Implications

Based on the evidence presented in this paper, continued growth of the opt out movement could have a direct impact on the ability of the CORE districts to accurately use student testing data in their school accountability systems. The results suggest that, while the current incidence of opting out in the included California districts is low according to district reports (< 1%), further growth could substantially impact their ability to accurately measure school performance and achievement gaps. This is due primarily to two factors; the demographic make-up of students opting out, and the concentration of opt out students by school. Similar to what has been found in previous literature (CDOE, 2015; NYSED, 2015; Ujifusa, 2015), the California districts included in this analysis show that White students are much more likely to opt out of standardized testing compared to minority students on average, as are students who do not qualify for free or reduced price lunches.

Under a range of plausible opt out growth scenarios based on the experience of other states, up to 4% percent of schools would see their performance fall into a lower tier of their district's accountability system, average math/ELA school-wide proficiency rates would fall from 41% to 36% (a similar fall to the district-wide average drop shown in Table 5), and average standardized school-wide performance would fall more than 5% of a standard deviation in ELA and 10% of a standard deviation in math. These results suggest that school-level accountability measures need to account for the types of students actually taking tests, lest they be biased substantially downward based on the under-representation of White students and students not in poverty in particular. All academic measures included in SQII that use test score data would also be affected, particularly those based on proficiency as opposed to growth, by a large number of higher performing students opting out of testing altogether.

The implications of the finding that the SQII's Growth Index is largely unaffected by even a substantial increase in opting out has broader implications for the ideal construction of accountability systems moving forward. By explicitly accounting for the types of students sitting for tests, accountability systems can control to a large extent for the bias introduced through growth in non-random student opt outs. However, the ability to adjust for larger and larger proportions of students opting out of testing has limits. One limitation, as is currently being experienced in multiple school districts in New York State, is that a sufficient number of students still need to sit for tests for plausibly reflect the performance of their district. In cases where only ten or twenty percent of students are used to reflect the performance of a school or district, the ability rescue a meaningful measure of quality using statistical adjustments is limited due to the likelihood of unrepresentativeness of the tested students and lack of common support among the included covariates to effectively account for the absence of certain subgroups. Another limitation is that the act of adjusting based on expected performance implicitly lowers the bar or different groups of students. The benefit of including uncorrected proficiency rates in accountability systems is that is provides a way to hold all schools accountable, regardless of their demographic makeup, to certain externally benchmarked expectations for future education and career success. In contrast, adjusting for the demographic composition of a school removes the ability to hold all students to the same level of expected achievement, replacing it with a unique metric based on performance relative to expected achievement for each student. In other words, the trading of growth models for proficiency models is not a value-neutral proposition, and it will not necessarily reflect the values of those tasked with creating accountability systems in education.

If the proportion of students opting out grows in the analyzed districts, as the experience of other states suggests is plausible, while demographic patterns remain similar, dozens of schools could see themselves fall below district-established benchmarks of performance. This suggests that even when districts go to great lengths to develop "holistic" accountability policies to reflect broad-based school-level contributions to student performance they are still susceptible to opting out, so long as they include test scores or other correlated measures. Additional analysis identifying the specific "near threshold" schools most at risk of falling into lower performance brackets in their district's accountability systems are also calculated for each district.[7] As the evidence presented here suggests, the phenomenon of opting out can have substantial impacts on our ability to measure the performance of schools and districts, and the experience of other states shows that it can rapidly grow from a relative non-issue to a direct threat to the basic legitimacy of accountability systems.

[7] These analyses are omitted from this paper as all publicly available documents must not allow for the identification of specific districts or schools.

References

Angrist, J. & Pischke, J. (2008) Mostly Harmless Econometrics; An Empiricist's Companion. 1st Edition.

Aragon, S., Rowland, J., & Wixom, M. A. (2015). Assessment Opt-Out Policies: State Responses to Parent Pushback. ECS Education Trends. Education Commission of the States.

Brandimarte, P. (2014). Handbook in Monte Carlo simulation: applications in financial engineering, risk management, and economics. John Wiley & Sons.

Bennett, R. E. (2015). The changing nature of educational assessment. Review of Research in Education, 39(1), 370-407.

Bennett, R. E. (2016). Opt out: An examination of issues. ETS Research Report Series, 2016(1), 1-16.

Council of Great City Schools (2015). Student Testing in America's Great City Schools: An inventory and preliminary analysis. Retrieved from http://www.cgcs.org/cms/lib/DC00001581/Centricity/Domain/87/Testing%20Report.pdf

Doherty, K. M., & Jacobs, S. (2015). State of the States 2015: Evaluating Teaching, Leading and Learning. National Council on Teacher Quality.

Education Commission of the States (2017). 50-State Comparisons - Accountability & Reporting: Current System. Retrieved from: http://ecs.force.com/mbdata/mbQuest6S?rep=SA171

Flores Aguilar, Y. (2009), "Public School Choice: A New Way at LAUSD", Motions/resolutions presented to the Los Angeles City Board of Education for consideration, 25 March, available at: http://laschoolboard.org/

Foster, J. B. (2016). The opt out revolt: Democracy and education. Monthly Review, 67(10), 1.

Fruchter, N. & McAlister, S. (2008). School governance and accountability: outcomes of mayoral control of schooling in New York City. Providence, RI: Annenberg Institute for School Reform, Brown University.

Hart, R., Casserly, M., Uzzell, R., Palacios, M., Corcoran, A., & Spurgeon, L. (2015). Student Testing in America's Great City Schools: An Inventory and Preliminary Analysis. Council of the Great City Schools.

Harris, D. & Larsen, M. (2015). What do families want and why? New Orleans families and their school choices before and after Hurricane Katrina. Education Research Alliance for New Orleans, policy brief.

Harris, E. A. (2015). 20% of New York State students opted out of standardized tests this year. New York Times.

Hill, P. T., & Campbell, C. (2011). Growing number of districts seek bold change with portfolio strategy. Seattle, WA: Center on Reinventing Public Education, University of Washington Bothell. Retrieved from: http://www.crpe.org/cs/crpe/download/csr_files/Brief_PSDP_SchoolDistricts_June11.pdf

Kappa, P. D. Gallup.(2015). The 47th PDK/Gallup Poll of the public's attitudes toward the public schools: Testing doesn't measure up for Americans. Phi Delta Kappan, 97(1).

Pizmony-Levy, O. & Saraisky, N. (2016). Who Opts Out and Why? Results from a national survey on opting out of standardized tests. Research Report. New York; Teachers College, Columbia University.

Tompson, T. N., Benz, J. K., & Agiesta, J. (2013). Parents' attitudes on the quality of education in the United States. Chicago, IL: Associated Press-NORC Center for Public Affairs Research.

Ujifusa, A. (2015). Amid cries of overtesting, a crazy quilt of state responses. Education Week.

Ujifusa, A. (2015). Education Department asks 13 states to address low test-participation rates. Education Week.

Ujifusa, A. (2015). New York's student opt-out rate for testing hits 20 percent. Education Week, 35(1), 4.

West, M., Peterson, P., & Barrows, S. (2017). What Do Parents Think of Their Children's Schools? EdNext Survey. Retrieved from: http://educationnext.org/files/2016ednext-parents-poll.pdf.

West, M. et. al (2017). The 2017 EdNext Poll on School Reform. EdNext, Winter 2018, vol. 18 (1).

Appendix

Figure A1. Components of the School Quality Improvement Index used by the Included Districts