Symposium Title

Issues in Accountability Testing: State Standards, Alternate Assessments, and Opting Out

Symposium Justification

Standardized test scores are important indicators of educational opportunity and gatekeepers to future opportunity. For students, test scores shape how teachers and administrators view their academic potential, affect their access to advanced courses, and so may affect their college admissions and later career options. For schools and districts, test scores may affect their reputation, their resources, and their receipt of sanctions or rewards. But, who is tested, how they are tested and the standards by which they are evaluated impact our perceptions of student performance and may unfairly penalize certain students and the districts serving them. Poor and minority students may have less support or advocacy from parents to alter their test taking conditions – including opting out of tests or gaining access to special testing conditions – and changes to the standards may differentially influence their scores.

The four papers in this symposium expand our understanding of the changing nature of test taking behavior – specifically, "opting out" and testing with accommodations – and state standards. The first paper analyzes the growing opt-out movement, describing patterns of opt out and characterizing what types of students are opting out and from which types of school districts across the nation. This work sheds light on how high rates of non-random opt out limit the effectiveness of accountability policies intended to close achievement gaps, and may add noise to signals about the quality of schools. The second paper in the panel focuses on testing with accommodations – answering the question of whether affluent students' families may be better able to advocate for accommodations, gaining an advantage over poorer students. The last two papers in the panel explore how state standards have changed in the past few years, and how The Common Core State Standards (CCSS) and CCSS-aligned tests have influenced test performance across school districts.

All four papers provide population-based descriptive evidence using state accountability testing behavior, standards, and outcomes. We believe that this symposium will be of interest to SREE members who study educational inequality and accountability testing.

Paper 1 Abstract Title Page

Title:

Opt Out in U.S. School Districts

Authors and Affiliations:

Erin M. Fahle, Stanford University
Sean F. Reardon, Stanford University
Sam Trejo, Stanford University

Abstract Body

Background & Objectives:

Unprecedented numbers of parents chose not to allow their children to participate in state accountability testing in the 2014-15 school year – they "opted out." Thirteen states failed to meet the 95% federal participation requirement according to the Department of Education (Bennett, 2016). However, little is known about the motivation behind this growing opt out movement. There is some evidence that students who opt out of tests tend to be white and live in affluent school districts (Bennett, 2016; Chingos, 2015; Pizmony-Levy & Saraisky, 2016). However, a systematic study is needed to identify the extent and impact of opt out in the U.S.

Our study provides population-based analyses of the development of opt out in U.S. schools and districts from the 2012-13 to 2014-15 school years. We investigate the association between district opt out rates district characteristics, and examine how opt out varies across the gender, racial and socioeconomic subgroups served within schools and districts. We seek to answer the following questions:

(1) Where are and how many students are opting out across the nation?
(2) How have the characteristics of the students and districts who opt out changed in recent years?
(3) Can the bias in district test score means caused by opt out be corrected using bounding or adjustment methods?

Evidence from this analysis can inform policy by quantifying how opt out affects the comparisons of district or state performance from federal accountability testing, as well as how opt out affects the use of accountability data as a benchmark for closing achievement gaps within districts.

Data & Research Design:

We estimate opt out rates for nearly every school district in the U.S. using state accountability participation and assessment data from the ED*Facts* database, a Department of Education data collection initiative. The ED*Facts* data include counts of students who took the assessments, as well as counts of students who were currently enrolled in each school by grade, subject, and year. We have these counts for all students, as well as by the following student subgroups: gender, race, and economic disadvantage. In total, we estimate overall and subgroup-specific opt out rates for approximately 11,000 districts from the 2012-13 through 2014-15 school years, in the third through eighth grades in mathematics and ELA.

We match this data to school-district characteristic data from the American Community Survey and Common Core of Data in order to understand what types of districts have the highest opt out. Our characteristics include: the racial composition of the district, the percentage of students

receiving free lunch, segregation of poor and non-poor students, and the socioeconomic status of families with children in public schools.

We use state-fixed effects models to estimate the relationship between opt out rates and district characteristics within states. The inclusion of state fixed effects eliminates any variation in states due to differences in policy or testing. Additionally, because opt out is near zero in most U.S. districts, we analyze the characteristics of districts in different quantiles of the opt out distribution.

Findings/Results:

In the 2014-15 school year, only about 3.3% of students nationally (pooled across subject and grades) opted out of state accountability tests. Consistent with prior evidence, we find that opt out is highest in later grades – the national opt out rate is nearly 4% in eighth grade compared to a little over 2% in third grade. Interestingly, the increase in national student opt-out is consistently seen across most gender, racial and educational subgroups, with the exception that it was highest among white students.

At the state-level, there are dramatic differences in the participation rates: states range from having overall participation rates as low as 75% (e.g., New York), well below the federal guideline of 95%, to nearly 100% (e.g., West Virginia). Within states, opt out rates were somewhat consistent across districts, but by 2015 there is substantial variation among districts suggesting that there may be characteristics of districts or the students they serve that are related to opt out. For example, Figure 1 shows the distribution of district participation rates on mathematics assessments in New York over grades and years. In New York, there were almost no districts with lower than 95% participation rates in 2012-13, but by 2014-15 there are a substantial number of districts with rates ranging between 10% and 95%.

Using the state-fixed effects models, we find districts with more free lunch students have lower opt out rates, while more affluent districts and districts with more economic school segregation have higher opt out rates. However, we find significant variability across states in the direction and magnitude of these correlations, suggesting that the mechanisms driving opt out and the students who are opting out are different across states. We also show that this is a dramatic change from the relationship between district socioeconomic status and opt out in 2013. In 2013, the districts with the highest opt out rates served the poorest students; however, by 2015, it is clearly the opposite – the districts serving the most affluent students have the highest rates of opt out (Figure 2).

Conclusions:

Opt out has the potential to undermine the use of state accountability tests for monitoring education opportunity and inequality across districts. Although opt out increased for all student subgroups and grades, the differential rates among affluent students are likely to both bias average district test scores downward and underestimate the gap between poor and non-poor

students. States are already pushing back on student opt out through policy (Aragon, Rowland, & Wixom, 2015; Lorenzo, 2015); however, it is unclear whether that will be an effective strategy to reduce opt out rates.

Paper 3 Abstract Title Page

Title:

Trends in State Proficiency Standards in the Common Core Era

Authors and Affiliations:

Benjamin R. Shear, University of Colorado Boulder
Andrew D. Ho, Harvard University
Sean F. Reardon, Stanford University
Erin M. Fahle, Stanford University

Abstract Body

Background & Objectives:

Since the No Child Left Behind Act of 2001, federal accountability policies have required each state to set its own "challenging academic standards" in mathematics, reading or language arts, and science (ESSA, 2015; NCLB, 2002). Variability in standards across states is an educationally and politically relevant concern—the public should know whether educators and policymakers are holding students to higher standards in, for example, Massachusetts than in Mississippi. Methods for comparing performance standards across different tests are not straightforward and rest on particular assumptions. Each state operationalizes its proficiency standard by setting a "proficiency" cut score on state reading and math assessments during a standard setting session (Cizek, 2012); students must then score above the cut score to be considered proficient in a particular subject. Because accountability tests vary across states and grades, proficiency cut scores cannot be compared directly across states, grades, and subjects.

This paper estimates state proficiency cut scores on a common scale using a NAEP-based linking, and uses the results to study how state proficiency standards have changed over time from 2009-2015. We focus on how proficiency standards have changed for states adopting the Common Core State Standards (CCSS) and CCSS-aligned tests relative to other states. As the rigid formulas and impossible deadlines of early NCLB policies (Linn, 2003) gave way to the flexibility of waivers (Polikoff, McEachin, Wrabel, & Duque, 2014), states have been able to raise performance standards toward the aspirations of "career and college readiness" without relegating every school to restructuring. At the same time, the rise of the CCSS initiative inspired an initially large number of states to voluntarily adopt not only common academic content standards but, for an ultimately smaller number of states, common assessments and cut scores (through either the PARCC or SBAC assessment consortia).

In this paper, we describe whether and how much states raised standards and if these changes reduced variance in standards across states. This paper addresses the following research questions:

(1) How has the stringency of state proficiency standards changed from 2009-2015?
(2) How has the variability of stringency in proficiency standards across states changed from 2009-2015?
(3) Are there differences in the patterns of proficiency standard changes for states adopting the CCSS and CCSS-aligned tests?

Data & Research Design:

Heteroskedastic ordered probit (HETOP) models (Reardon, Shear, Castellano, & Ho, 2017) are used to estimate the location of each state's proficiency cut scores from aggregate mathematics and reading proficiency counts from every US school district in grades 3-8 in 2009-2015. Reliability-adjusted linear linking methods are then used to place each state's cut scores on the NAEP scale (Reardon, Kalogrides, & Ho, 2017).

We begin this process by fitting the HETOP model to estimate district test score distributions within each state s, grade g, year y, and subject b (math or reading). This provides estimates of state proficiency cut scores in a standardized distribution, $\hat{c}_{sygb}^{state}$. We then link these cut scores to the NAEP scale using the following equation:

$$\hat{c}_{sygb}^{naep} = \hat{\mu}_{sygb}^{naep} + \frac{\hat{c}_{sygb}^{state}}{\sqrt{\hat{\rho}_{sygb}^{state}}} * \hat{\sigma}_{sygb}^{naep}$$

where $\hat{\mu}_{sygb}^{naep}$ and $\hat{\sigma}_{sygb}^{naep}$ are the mean and standard deviation of NAEP scores and $\hat{\rho}_{sygb}^{state}$ is the estimated reliability of the state test, for state s in year y, grade g, and subject b. This process is repeated for cut scores separating additional performance levels. We use the observed NAEP means and standard deviations for each state in years and grades that NAEP is administered (4th and 8th grade in odd years) and use linearly interpolated means and standard deviations to estimate linked cut scores for additional grades and years. Reardon, Kalogrides and Ho (2017) show that linkages are robust when using linearly interpolated means and standard deviations for these additional years and grades.

We briefly contrast this approach, methodologically and practically, with the equipercentile linking method used to link state proficiency cut scores to the NAEP scales for a more limited set of grades and years in prior publications (e.g., Bandeira de Mello, Bohrnstedt, Blankenship, & Sherman, 2015; Braun & Qian, 2007). We use this comparison to evaluate whether caveats and limitations identified by prior linking evaluations apply to our results or can account for discrepancies (Ho & Haertel, 2007; Reardon, Kalogrides, et al., 2017; Thissen, 2007).

We then use these estimates of proficiency cut scores on a common scale to study trends in the stringency of proficiency standards. We begin by providing descriptive statistics about the distribution of proficiency standards for each grade and subject across time. We then use precision-weighted hierarchical linear models (HLM; Raudenbush & Bryk, 2002) to model the trend in proficiency standards over time for states that did and did not adopt the CCSS and CCSS-aligned tests, and to quantify changes in the variability of proficiency standards time.

Preliminary Findings/Results:

We find that proficiency standards rose dramatically from 2009-2015, with the median state standard rising over .7 standard deviation units in reading/language arts and over .8 standard deviation units in mathematics. The increase was particularly substantial over the last year in the data, from 2014 to 2015. Convergence was less dramatic, with moderate decreases in variance in reading and only slight decreases in mathematics that are inconsistent across grades. Figure 3 displays quartiles of state proficiency cut scores over this 6-year period in mathematics to illustrate the kinds of findings we will present and discuss. Our preliminary results also suggest that proficiency standards increased more for states adopting the CCSS.

Conclusions:

We find clear evidence that state proficiency standards have changed dramatically during the CCSS era, and that there is substantial variability in the meaning of "proficient" across states, grades and years. We discuss the policy implications of these increases in performance standards. We also emphasize the caveats associated with comparing proficiency standards across states, and with comparing achievement outcomes or achievement gaps across states in light of such great variability in the meaning of "proficient."

Paper 4 Abstract Title Page

Title:

Trends in Within-State Achievement Inequality in the Common Core Era

Authors and Affiliations:

Sean F. Reardon, Stanford University
Rebecca Hinze-Pifer, Stanford University

Abstract Body

Background & Objectives:

This decade has witnessed a broad education standardization movement – the Common Core State Standards (CCSS), adopted by 45 states and DC at its peak – followed by a partial transition to CCSS-aligned standardized testing in roughly half of states. In a loosely-coupled educational sector with a diverse set of organizational actors of varying capacity, resources, and goals, the implications of the CCSS and related testing regimes for educational inequality are not immediately apparent.

The potential impacts of CCSS adoption, implementation, and transitions to new testing regimes may vary across levels of the educational system. Greater commonalities in academic content and instructional methods among states adopting the CCSS could potentially produce narrowed differences in academic achievement among these CCSS-adopting states. In contrast, the processes through which schools, districts, and states align instructional practices, materials, and testing instruments with the CCSS are unfolding in different ways and at different paces in each state.

Likewise, at the district level, it is unclear whether state-level CCSS adoption will shift disparities in academic performance. If the success of CCSS adoption relies on district resources and capacity, then adoption may lead to widening disparities among school districts and by socioeconomic status. Conversely, if a state's CCSS adoption does more to improve the quality of instruction and curricular coherence in less-advantaged schools and districts (because more advantaged districts already deployed CCSS-aligned practices, or because state agencies employ compensatory actions during implementation), then the CCSS initiative might narrow within-state achievement gaps by differentially raising performance in less-advantaged districts. It is far from clear which of these outcomes will result (if either), and how any outcomes will differ across states and districts.

Guided by a broad set of questions about the relationship between current efforts at standards-based reform and patterns of inequality in students' academic achievement, this paper assesses the ways that patterns of academic achievement and inequality have unfolded with the transition to CCSS-aligned testing. Specifically, we examine the following:

(1) As states transition to CCSS-aligned tests, to what extent do patterns of between-district inequality change within states? Are changes consequentially different from those observed during other periods of testing change?
(2) Comparing states adopting the two consortium-developed CCSS-aligned tests (PARCC and Smarter Balanced), are there differences in the magnitude or direction of changes in measured inequality?
(3) Within districts, how do gaps between lower and higher SES students shift with the transition to CCSS-aligned testing? Do these patterns differ for states using PARCC compared with those using Smarter Balanced?

Data & Research Design:

The core data for this analysis comes from the Stanford Education Data Archive (SEDA), which includes data on test results for every state accountability test taken in grades 3-8 in the U.S. over the period 2009-2015. Using these and other data on state and district characteristics, we examine patterns of academic achievement among CCSS-adopting and non-adopting states and among school districts with different levels of socioeconomic status and capacity. Although standardized test scores are at best a partial measure of the goals of public schooling, they are predictive of later outcomes, including college enrollment and subsequent earnings, and serve as a proxy measure for broader goals of schooling.

We examine three specific outcomes:

(a) the between-district proportion of test score variance in each state-year (the intraclass correlation, or ICC);
(b) the between-district association between average district achievement and district socioeconomic status; and
(c) within-district achievement gaps between poor and non-poor students.

We compare trends in these outcomes among four categories of states:

(a) CCSS-adopting states which transitioned to PARCC assessments in 2015,
(b) CCSS-adopting states which transitioned to Smarter Balanced assessments in 2015;
(c) CCSS-adopting states that used a different test; and
(d) states that did not adopt the CCSS.

We use a difference-in-differences design to compare changes in each of the outcome measures from 2014 to 2015 in states that used new tests in 2015 to changes in the outcomes in states whose tests did not change. We use outcome trends from 2009-2015 to control for prior state trends in the outcomes.

Preliminary Findings/Results:

We have preliminary results for some of the analyses that use the ICC and the achievement-SES gradient as the outcome.

Examination of the intraclass correlation coefficient (ICC) within states reveals a clear pattern. Figure 4 shows the time trend in the average annual state-level intraclass correlation coefficient (ICC) for our four categories of states. The transition to CCSS-aligned tests (the PARCC and SBAC) corresponds to a widening of between-district differences in states from 2014-2015. On average, the ICC increased by 15-20% when states switched to the PARCC or SBAC. The ICC did not change significantly from 2014-2015 in states that did not adopt the CCSS or in states that did, but that did not use the PARCC or SBAC assessments. The differences in the trends between

PARCC- and SBAC-adopting states and other states, controlling for prior trends, are statistically significant ($p<.001$).

A different pattern is evident in the trends in the Achievement-SES gradient (Figure 5). On average, the association between district average academic achievement and district SES has been increasing over time, but the rate of increase is not significantly different among our four categories of states. There is also no significant change in the trends from 2014-2015, relative to prior years.

Conclusions:

The introduction of new CCSS-aligned tests appears to increase measured between-district test score variance, but this increase does not appear to be correlated with districts' SES. In other words, between-district variation in test scores grows even among districts of the same SES, suggesting that the causes of this increase are not due to factors correlated with district SES. The patterns we observe may reflect true changes in educational inequality, differences in the content measured by assessments, differences in who takes the tests (opt-out), and/or short-term adjustments as districts learn how to align teaching to tested standards. We plan to investigate these patterns further by comparing them to changes in test score variation that result from other (non-CCSS-related) changes in testing.

References

Ahram, R., Fergus, E., & Noguera, P. (2011). Addressing racial/ethnic disproportionality in special education: Case studies of suburban school districts. *Teachers College Record*, *113*(10), 2233–2266.

Aragon, S., Rowland, J., & Wixom, M. A. (2015). *Assessment opt-out policies: State responses to parent pushback*. Education Commission of the States.

Bandeira de Mello, V., Bohrnstedt, G., Blankenship, C., & Sherman, D. (2015). *Mapping state proficiency standards onto NAEP scales: Results from the 2013 NAEP reading and mathematics assessments* (No. NCES 2015-046). US Department of Education, Washington, DC: Institute of Education Sciences.

Bennett, R. E. (2016). Opt Out: An Examination of Issues. *ETS Research Report Series*. https://doi.org/10.1002/ets2.12101

Braun, H. I., & Qian, J. (2007). An enhanced method for mapping state standards onto the NAEP scale. In N. J. Dorans, M. Pommerich, & P. W. Holland (Eds.), *Linking and aligning scores and scales* (p. 313). New York, NY: Springer.

Chingos, M. M. (2015). *Who opts out of state tests?* Washington, D.C.: Brown Center on Education Policy at Brookings.

Cizek, G. J. (Ed.). (2012). *Setting performance standards: Foundations, methods, and innovations* (2nd ed.). New York, NY: Routledge.

Elliott, S. N., Kratochwill, T. R., & McKevitt, B. C. (2001). Experimental analysis of the effects of testing accommodations on the scores of students with and without disabilities. *Journal of School Psychology*, *39*(1), 3–24. https://doi.org/10.1016/S0022-4405(00)00056-X

Hibel, J., Farkas, G., & Morgan, P. L. (2010). Who is Placed into Special Education? *Sociol Educ*, *83*(4), 312–332. https://doi.org/10.1177/0038040710383518.WHO

Ho, A. D., & Haertel, E. H. (2007). *Apples to apples? The underlying assumptions of state-NAEP comparisons*. Washington, D.C.: Council of Chief State School Officers.

Lewin, T. (2002). Abuse Feared as SAT Test Changes Disability Policy.

Linn, R. L. (2003). Accountability: Responsibility and reasonable expectations. *Educational Researcher*, *32*(7), 3–13. https://doi.org/10.3102/0013189X032007003

Lorenzo, S.-J. (2015). *Opt out policies by state*. National Association of State Boards of Education.

Losen, D. J., & Orfield, G. (Eds.). (2002). *Racial Inequity in Special Education*. Cambridge, MA: Harvard Civil Rights Project.

Mitchell, H. (2012). Faking ADHD Gets You Into Harvard.

Morgan, P. L., Farkas, G., Hillemeier, M. M., & Maczuga, S. (2012). Are Minority Children Disproportionately Represented in Early Intervention and Early Childhood Special Education? *Educational Researcher*, *41*(9), 339–351. https://doi.org/10.3102/0013189X12459678

Morgan, P. L., Farkas, G., Hillemeier, M. M., Mattison, R., Maczuga, S., Li, H., & Cook, M. (2015). Minorities Are Disproportionately Underrepresented in Special Education: Longitudinal Evidence Across Five Disability Conditions. *Educational Researcher*, *44*(5), 1–15. https://doi.org/10.3102/0013189X15591157

Morgan, P. L., Staff, J., Hillemeier, M. M., Farkas, G., & Maczuga, S. (2013). Racial and ethnic disparities in ADHD diagnosis from kindergarten to eighth grade. *Pediatrics*, *132*(1), 85–93.

https://doi.org/10.1542/peds.2012-2390
National Research Council. (2002). *Minority Students in Special and Gifted Education*. (S. Donovan & C. T. Cross, Eds.). Washington, D.C.: National Academy Press.
Parrish, T. (2002). Racial disparities in the identification, funding, and provision of special education. In D. J. Losen & G. Orfield (Eds.), *Racial inequity in special education* (pp. 15–35). Cambridge: Harvard Education Press.
Pizmony-Levy, O., & Saraisky, N. G. (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.
Polikoff, M. S., McEachin, A. J., Wrabel, S. L., & Duque, M. (2014). The waive of the future? School accountability in the waiver era. *Educational Researcher*, *43*(1), 45–54. https://doi.org/10.3102/0013189X13517137
Ramey, D. M. (2015). The social structure of criminalized and medicalized school discipline. *Sociology of Education*, *88*(3), 181–201. https://doi.org/10.1177/0038040715587114
Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.
Reardon, S. F., Kalogrides, D., & Ho, A. D. Linking U.S. school district test score distributions to a common scale (2017). Working Paper, Stanford Center for Education Policy Analysis.
Reardon, S. F., Shear, B. R., Castellano, K. E., & Ho, A. D. (2017). Using heteroskedastic ordered probit models to recover moments of continuous test score distributions from coarsened data. *Journal of Educational and Behavioral Statistics*, *42*(1), 3–45. https://doi.org/10.3102/1076998616666279
Skiba, R. J., Artiles, A. J., Kozleski, E. B., Losen, D. J., & Harry, E. G. (2016). Risks and Consequences of Oversimplifying Educational Inequities: A Response to Morgan et al. (2015), (April), 221–225. https://doi.org/10.3102/0013189X16644606
Skiba, R. J., Simmons, A. B., Ritter, S., Gibb, A. C., Rausch, M. K., Cuadrado, J., & Chung, C. (2008). Achieving Equity in Special Education: History, Status, and Current Challenges. *Exceptional Children*, *74*(3), 264–288. https://doi.org/10.1377/hpb2011.8
Sullivan, A. L., & Bal, A. (2013). Disproportionality in Special Education: Effects of Individual and School Variables on Disability Risk. *Exceptional Children*, *79*(4), 475–494. https://doi.org/10.1177/001440291307900406
Tapper, J., Morris, D., & Setrakian, L. (2006). Does Loophole Give Rich Kids More Time on SAT?
Thissen, D. (2007). Linking assessments based on aggregate reporting: Background and issues. In N. J. Dorans, M. Pommerich, & P. W. Holland (Eds.), *Linking and aligning scores and scales* (pp. 287–312). New York, NY: Springer.

Figures

Figure 1. Distribution of District Participation Rates in New York, by Grade & Year in Mathematics.

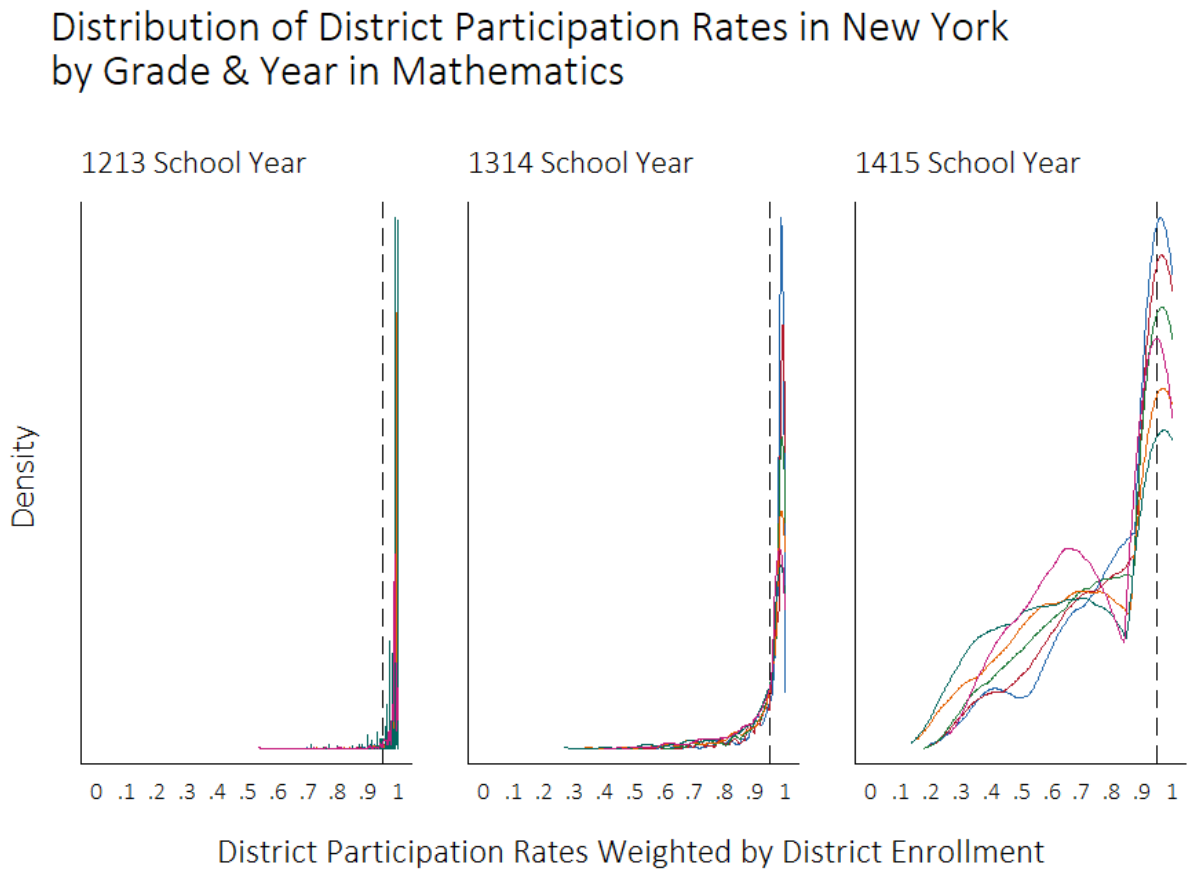


Figure 2. Average District Socioeconomic Status (SES) by Percentile Bands of the Opt Out Distribution.

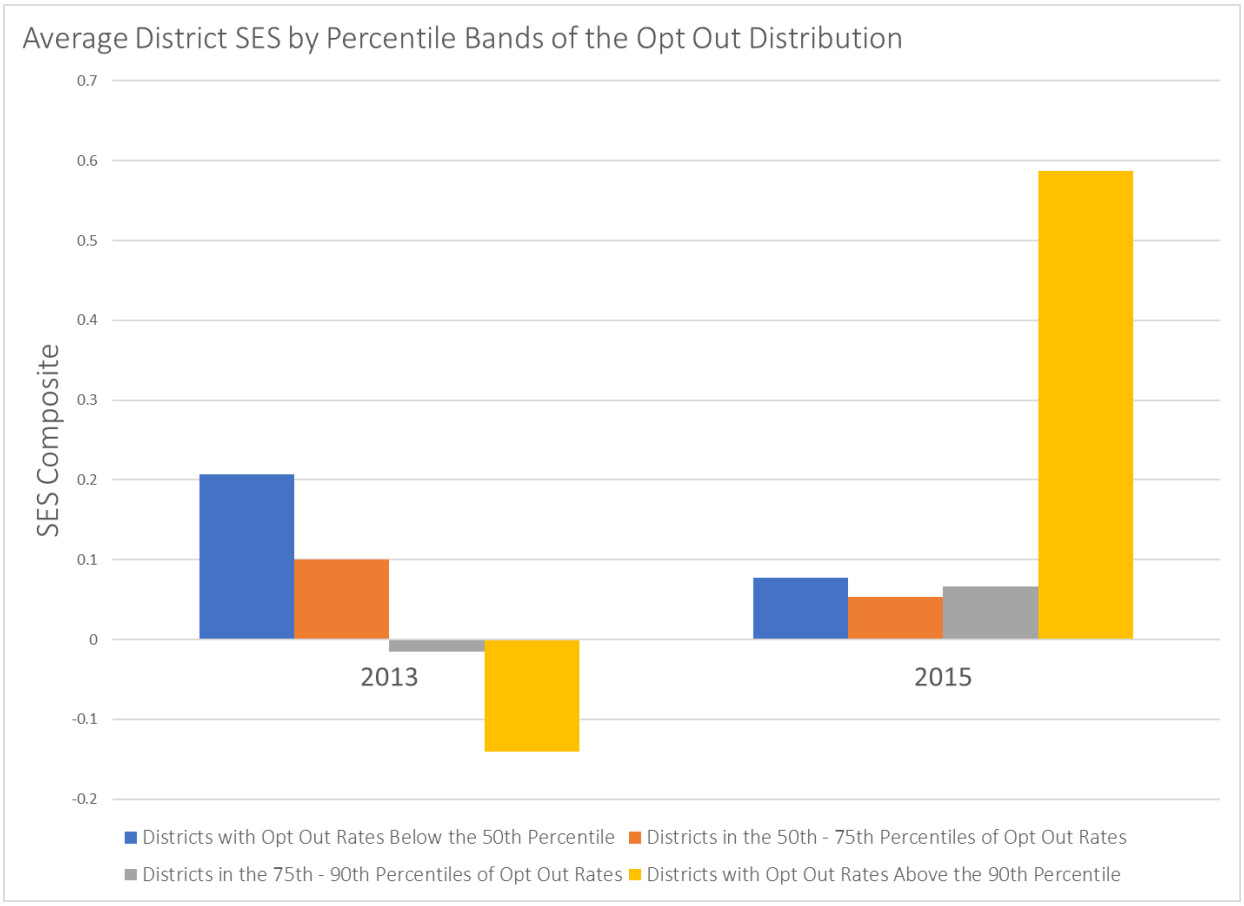


Figure 3. State Mathematics Proficiency Cut Scores Standardized Relative to 2009, by Grade.

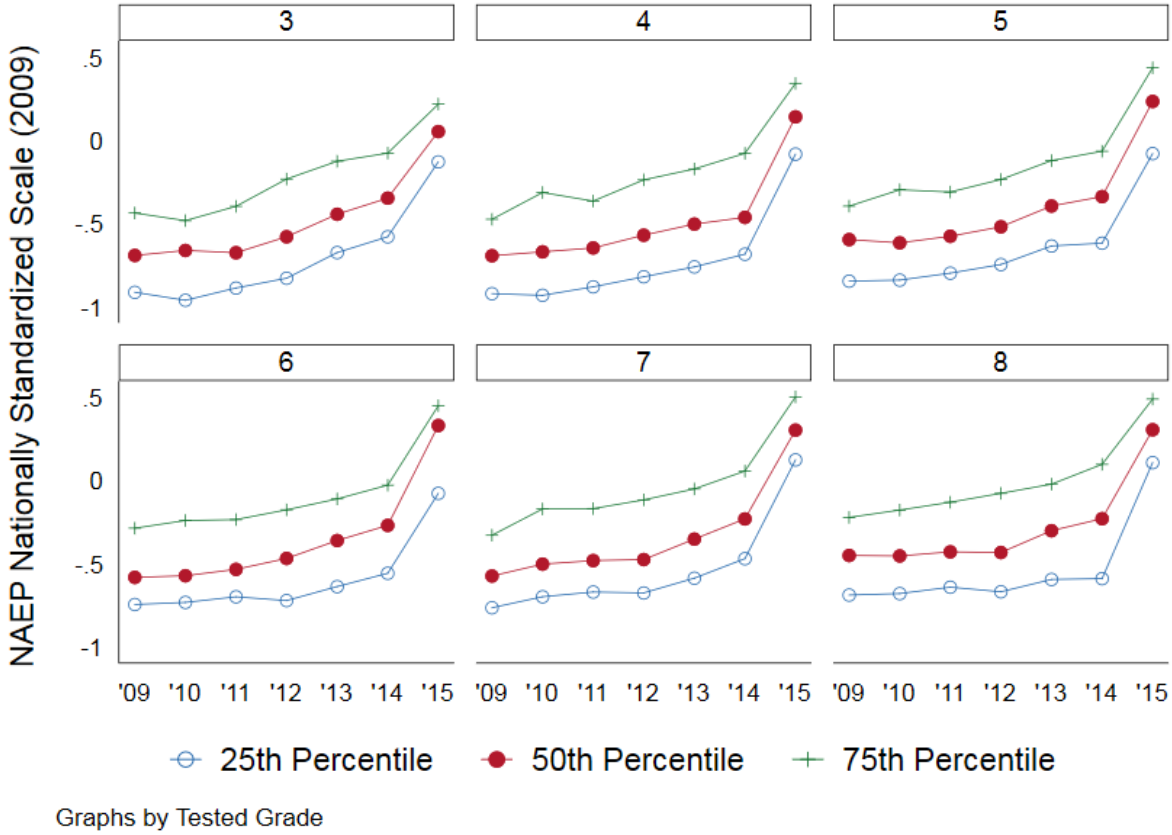


Figure 4. Trend in Between-District Intraclass Correlation Coefficient (ICC), by State Standards/Test, 2008-2015.

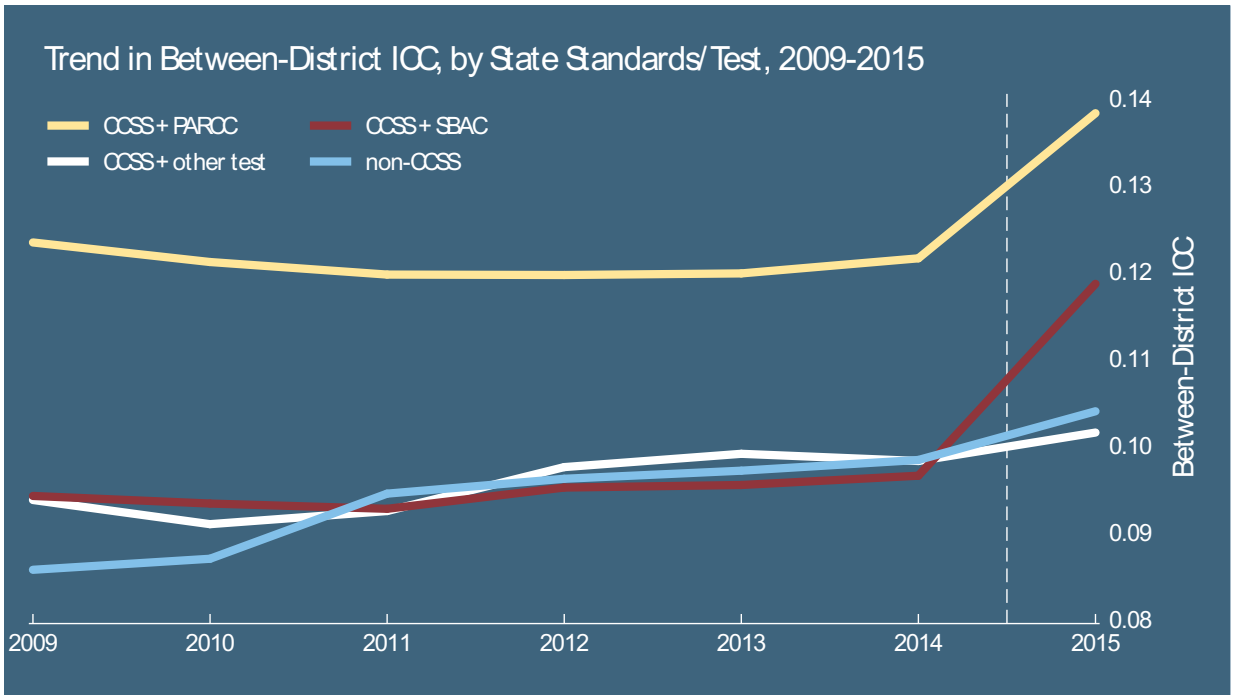
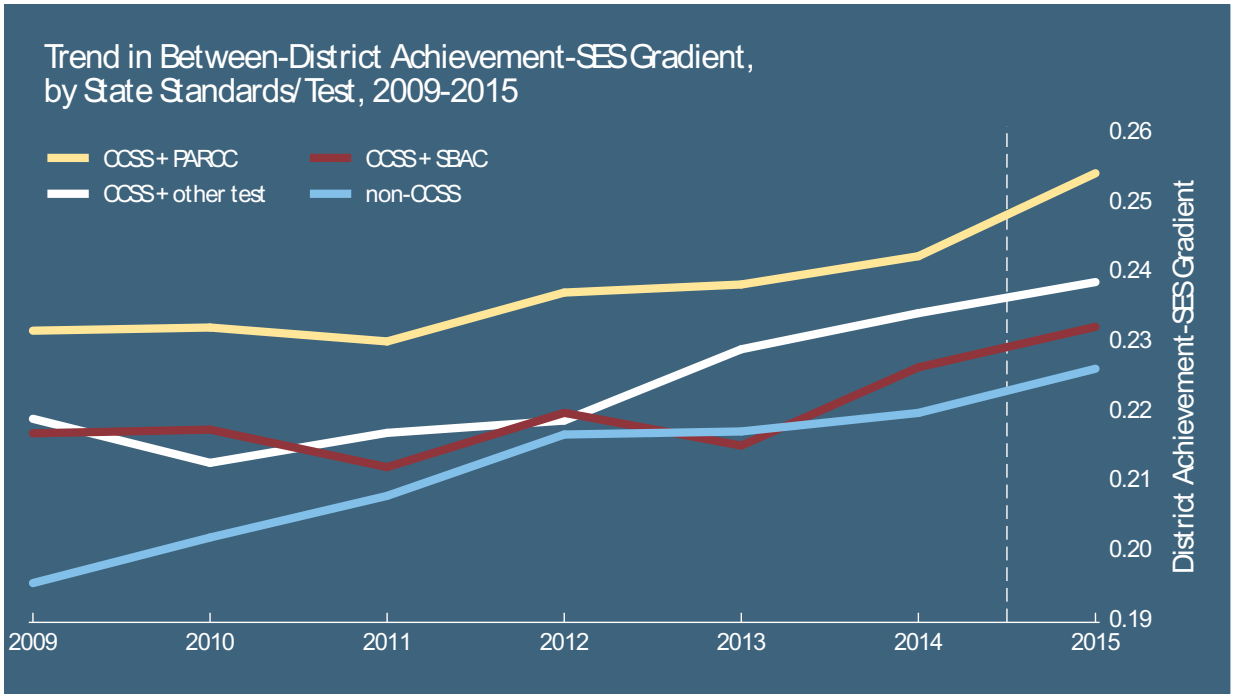


Figure 5. Trend in Between-District Achievement-SES Gradient, by State Standards/Test, 2009-2015.



Tables

Table 1. District Level Predictors of Special Education

	(1)	(2)	(3)
Percent Native American	0.0160***	0.0077***	-0.0078
	(0.0007)	(0.0007)	(0.0174)
Percent Asian	0.0205***	0.0036**	0.0004
	(0.0012)	(0.0012)	(0.0143)
Percent Hispanic	-0.0733***	-0.0696***	-0.0394***
	(0.0003)	(0.0003)	(0.0051)
Percent Black	-0.0213***	-0.0400***	-0.0306***
	(0.0004)	(0.0004)	(0.0060)
SES Composite	-0.0090***	-0.0117***	-0.0085***
	(0.0001)	(0.0001)	(0.0014)
Student-Teacher Ratio		-0.0006***	-0.0008**
		(0.0000)	(0.0002)
Per Pupil Expenditure		0.0000***	0.0000***
		(0.0000)	(0.0000)
Mean Reading/Language Arts Achievement		0.0000***	-0.0002***
		(0.0000)	(0.0001)
Mean Math Achievement		-0.0001***	-0.0003***
		(0.0000)	(0.0000)
N	393033	374713	374713
Adjusted R-Squared	0.147	0.249	0.142
State and Grade Fixed Effects	No	No	Yes

Notes: Percent White is the reference category. SES composite is calculated as the first principal component factor score of median income, percent with a bachelor's degree or higher, poverty rate, Supplemental Nutrition Assistance Program (SNAP) rate, single mother headed household rate, and unemployment rate. Standard errors in parentheses. * p<.05, ** p<.01, *** p<.001

Table 2. District Level Predictor of Rates of Assessment Types by Subject

	Reading/English Language Arts				*Mathematics*			
	Regular Assessments	Regular Assessment w/ Accommodations	Alternate Assessment	Alternate Assessment w/ Modifications	Regular Assessments	Regular Assessment w/ Accommodations	Alternate Assessment	Alternate Assessment w/ Modifications
Percent Native American	-0.0099*	-0.0045	-0.0141	0.0205***	-0.0084	-0.0065	-0.0131	0.0191***
	(0.0043)	(0.0051)	(0.0081)	(0.0051)	(0.0045)	(0.0049)	(0.0081)	(0.0053)
Percent Asian	-0.0198**	0.0293***	-0.0052	-0.0116	-0.0223**	0.0366***	-0.0035	-0.0040
	(0.0068)	(0.0078)	(0.0117)	(0.0065)	(0.0071)	(0.0076)	(0.0117)	(0.0067)
Percent Hispanic	-0.0216***	0.0200***	-0.0138***	-0.0033	-0.0313***	0.0340***	-0.0112**	-0.0029
	(0.0022)	(0.0025)	(0.0040)	(0.0021)	(0.0023)	(0.0024)	(0.0040)	(0.0022)
Percent Black	0.0323***	-0.0252***	-0.0062	0.0041	0.0379***	-0.0300***	-0.0057	0.0045
	(0.0023)	(0.0025)	(0.0039)	(0.0024)	(0.0023)	(0.0025)	(0.0039)	(0.0025)
SES Composite	0.0045***	-0.0052***	-0.0050***	-0.0042***	0.0039***	-0.0044***	-0.0053***	-0.0042***
	(0.0005)	(0.0006)	(0.0009)	(0.0005)	(0.0005)	(0.0005)	(0.0009)	(0.0006)
N	69577	56626	40012	20030	69555	60055	39166	19741
Adjusted R-Squared	0.027	0.043	0.002	0.020	0.029	0.048	0.002	0.019

Notes: All models include state and grade fixed effects, as well as controls for student-teacher ratio, per-pupil expenditure, and mean achievement in math and reading/ela. Percent White is the reference category. SES composite is calculated as the first principal component factor score of median income, percent with a bachelor's degree or higher, poverty rate, Supplemental Nutrition Assistance Program (SNAP) rate, single mother headed household rate, and unemployment rate. Standard errors in parentheses. * p<.05, ** p<.01, *** p<.001