# High School Accountability:

# Early Evidence from Florida's Broward County Public Schools

Patrice Iatarola
Florida State University

Niu Gao
Public Policy Institute of California

Paper presented at the Association for Education Finance and Policy
Washington, D.C.
February, 2015

Correspondence on this article should be addressed to Patrice Iatarola, Department of Educational Leadership & Policy Studies, Florida State University, Tallahassee, FL 32306. Contact: piatarola@fsu.edu

The National Center for Scaling Up Effective Schools (NCSU) is a national research center and development center that focuses on identifying the combination of essential components and the programs, practices, processes and policies that make some high schools in large urban districts particularly effective with low income students, minority students and English Language Learners. NCSU and this research are funded by the Institute of Education Sciences (R305C10023). The opinions expressed in this article are those of the authors and do not necessarily represent the views of the sponsor.

Abstract

In 2009, Florida adopted the Differentiated Accountability (DA) plan, making it among the first to specifically incorporate into its existing school grading scheme college readiness targets. In this paper we use a rich panel of data on high school students in Broward County (Ft. Lauderdale) Public Schools to present early evidence of the impact of these changes to the high school accountability grades on students' participation and performance in accelerated coursework. Using a difference-in-difference approach, we find that DA increases students' probability of taking AP tests in English, Science and Social Science. However, it does little to improve students' performance in these tests and the effect, if any, is concentrated among white students and among students in high-performing schools. The treatment heterogeneity is likely due to the difference in instructional resources across schools. DA also has far-reaching impact as it elicits efforts from schools to prepare students for accelerated coursework. For instance, it increases students' probability of taking benchmark courses on time. Last, DA effect persists and accumulates over time.

## 1. Introduction

In 1999 with its *A+ Plan for Education*, Florida was among the first states to institute a standards-based accountability system that included a school grading system and that offered schools incentives and sanctions in relation to their school grades. Scores on the Florida Comprehensive Assessment Test (FCAT), which was expanded to grades three through ten in 1999, were central to the calculation of school grades, initially measured by performance level with performance gains subsequently incorporated into the school grade. Individual students were also held accountable for their performance on FCAT exams, for example high school students had to pass both the reading and math $10^{th}$ grade FCAT in order to graduate from high school, effective for the graduating class of 2003. The 2002 federal *No Child Left Behind* (NCLB) accountability system also relied on standardized test scores to hold schools accountability and was overlaid on to the existing *A+ Plan*.

Over the past decade, states across the U.S. implemented their own accountability systems that varied in a number of respects but all ultimately used standardized test scores, as required by NCLB, to hold schools and students accountable. The theory of action is that by introducing a direct incentive structure into public education and by holding schools accountable for student achievement, over time student learning will improve, school capacity will grow and achievement gaps will decrease. To date, the evidence on the effectiveness of accountability systems is mixed and with respect to high schools is very thin (Hanushek & Raymond, 2005; Carnoy & Loeb, 2002; Dee & Jacob, 2011). Accountability is also fraught with problems and among the often cited critiques are: distortion of the curriculum, teaching to the test, increased drop-out rate for students at risk of failing graduation tests, disproportionate burdens to minority and low-income students, failing too many schools without providing enough support, and gaming to the test (Darling-Hammond, 2004; Neal & Schanzenbach, 2010; Kain & Staiger, 2002; Reback, 2008; White & Rosenbaum, 2007; Booher-Jennings, 2005; U.S. Department of Education, 2009).

It is within this broader policy context that we examine Florida's revamping of its statewide accountability system to align with the college and career readiness goals and requirements that the President Obama laid out in his incentive-laden Race-to-the-Top initiative that included differentiated accountability. As discussed in detail in the background section of the paper that follows this introduction, the state reconfigured its grading formula for high schools to include components of accelerated course participation and performance, college readiness, and both four- and five-year graduation rates, in addition to pre-existing assessment components. We do so within the context of Broward County Public Schools, the sixth largest district in the nation and one that in key respects well represents a diverse urban/suburban context, racially/ethnically, socio-economic status as well as performance. With more than a decade of student-level data on hand for Broward County Public School students, we examine whether or not the change in the accountability grading scheme had an impact on students' likelihood of taking accelerated courses[1] and their performance on related exams and graduating within four years. We test the sensitivity of our results by examining, as well, students' early taking of benchmark courses that would track students to college-ready course levels. Further, we examine whether or not the impact, if any, differs or depends on students' socio-demographic characteristics and educational settings. In the sections that follow, we provide a background of the educational policy context, including a description of how the state grades schools, review the literature on accountability as well the conceptual basis and theory of change that are at the center of such systems. We then describe our methodological approach and report on our

[1] More specific definition is included in the next section of the paper.

findings.  We conclude with reflections on what we find and what the implications may be for policy and practice.

**2. Background**

Florida is once again at the forefront of educational accountability with an encompassing system that brings high schools more fully and rigorously into the accountability system by incorporating measures of college readiness and advanced course taking as part of high schools' grades.[2] This was done as part of Florida's efforts in establishing a Differentiated Accountability (DA) system, being one of six states selected by the U.S. Department of Education in 2008 to pilot DA. Policymakers hoped DA would allow for the alignment and integration of the federal (NCLB) and state (*A+*) accountability systems generally, intending as well to give states more flexibility in the implementation of school improvement strategies. In 2009, the Florida legislature mandated that DA be the state's official accountability system and to go into effect for the 2009-10 school year.

DA differs from previous models of accountability in two significant ways. First, it differentiates between schools in need of intensive intervention and those closer to meeting accountability goals and allows states to vary the intervention and assistance based on the status of schools. For instance, DA groups schools based on a 'school in need of improvement' (SINI) status, school grades, and AYP criteria (whether or not it has been met) and focuses improvement progressively with more support going to schools in need of intensive interventions. Second, DA is the first accountability policy that specifically incorporates college readiness targets in addition to the traditional benchmarks on standardized tests and graduation rates. Under DA, as initially proposed, 100 percent of the grades for elementary and middle

[2] The prior system used FCAT scores to hold high schools accountable.  These are notoriously limited measures given that the FCAT exam tests students through 10th grade.

schools is based on FCAT performance and learning gains, which is same as before. At the high school level, however, 50 percent of the grades of high schools are based on FCAT components and 50 percent on non-FCAT components, including graduation rate, graduation rate for at-risk students, accelerated coursework participation, performance, college readiness, and annual growth or decline in those measures (Florida Department of Education, 2008).

Our focus, in this study, is on the second component – the school grade calculation, particularly as it relates to the incentives imbedded in the non-FCAT components for high schools. We are interested to see if the inclusion of non-FCAT components spurred increases in students' likelihood of taking accelerated courses and their performance on related exams, students' early taking of benchmark courses that would track students to college-ready course levels and students' likelihood of graduating within four years. As noted above, prior to 2009-10, school grades were based only on FCAT components. The new grading scheme reduced the FCAT components to 50% of the overall grade with the non-FCAT components comprising the other 50%. Within the other components accelerated course participation[3] accounts for 25% of the 'other components' and 12.50% of the total grade. Accelerated course performance[4] accounts 12.50% of the 'other components' and 6.25 of the total grade. Four-year graduation rates account for 25% of the component (12.50% overall) and at-risk four-year graduation accounts for 12.5% (6.25% overall). College readiness[5] in reading and math accounts for 25% of

[3] Specifically, participation means students' enrollment in Advanced Placement (AP), International Baccalaureate (IB), Advanced International Certificate of Education (AICE), industry certification, and dual enrollment courses as a share of the students enrolled in 11th and 12th grades. Students who take accelerated courses in 9th and 10th grades are eligible to be counted in the numerator if they are successful on the respective AP, IB, and AICE exams or if they earn a passing grade in dual enrollment courses. Please note, for this draft of the paper, we only examine AP courses and tests, future analysis will include IB, AICE courses and exams and dual enrollment courses and course grades.

[4] Specifically, performance is measured by students' success on AP, IB, and AICE exams and passing grades in dual-enrollment courses as a share of accelerated course participants.

[5] Specifically, college readiness is measured as the number of on-time graduates who score 'ready' on ACT, SAT or Common Placement Test (CPT) as a share of all on-time graduates.

the component (12.5% overall). In subsequent years, the specific weights for participation and performance in accelerated coursework were tweaked in 2010-11 and 2011-12 (See Appendix Table 1 for the specific grade components from 2009-10 through 2012-13), in each instance giving more weight to performance and less weight to participation. Also, in 2011-12, the weights for four-year graduation overall and for at risk students were split with half of the weight going toward four-year graduation and half going toward five-year graduation.

DA was not the only policy change made that would impact students' course taking and likelihood of graduating. At the same time, in 2010, the state legislature increased graduation requirements, phasing in changes over a six-year period starting with entering 9$^{th}$ graders in 2010-11 and ending with entering 9$^{th}$ graders in 2015-16. The key changes were focused on math and science coursetaking and intended to increase college readiness. For example, in 2010-11 entering 9$^{th}$ grades would be required to earn credits in Algebra I and Geometry along with passing an Algebra I end-of-course (EOC) exam in order to graduate. In the following year, they were also required to pass an EOC in Geometry and earn credit and pass an EOC in Biology (Florida Association of District School Superintendents, Undated). By 2015-16 entering 9$^{th}$ graders, in order to graduate, would need to pass EOCs in Algebra I, Geometry, and Biology, earn credits in Algebra II and pass a Physics, Chemistry or similarly rigorous course as well as earn credit in U.S. History with an EOC incorporated into the course grade. In 2013 the legislature pulled back on some of these requirements, as it was clear that districts and schools were facing significant challenges in helping students meet these standards. While students entering 9$^{th}$ grade in 2013-14 would still be required to take and earn credit in Algebra I and Geometry and pass the Algebra I EOC, the Geometry EOC would count only as part of their course grade (counts as well in Algebra I). In terms of science, students would still be required to have at least two of their three science credits in a lab course and take Biology with an EOC incorporated into their course grade (Florida Statutes 1003.4282, 2014).

The state legislature and department of education were not the only ones implementing changes. In response to the state's changes to high school accountability grades and graduation requirements, Broward County Public Schools, the sixth largest school district in the U.S. and the subject of our study, adopted a computer based course-assignment matrix to assign students into courses. The computer algorithm takes into account students' test scores and course-takings in previous grades and predicts which courses students should take. For instance, if a student (1) scored between 275 and 343 points in $7^{th}$ grade FCAT math, (2) took advanced middle/junior high math 3, and (3) earned a B or above in that course, the computer will place the student into Algebra 1 in $9^{th}$ grade. The computerized matrix took out the subjectivity in course placement decisions and pushed for more rigor in courses taken by students. Early evidence, however, is mixed on whether or not the matrix leads to more rigor and better performance in courses (Iatarola, Rutledge, Kim and Brown, 2015).

The DA and the course-assignment matrix offer a nice gateway to examine the effects of high school accountability policies on a variety of outcomes. Given that Florida is among the first states to implement DA and to include more expansive measures of high school performance, this study will provide critical early evidence of the policy's impact. This is particularly important as it may offer lessons for scaling up this policy across the fifteen other states that include accelerated course-taking and test-taking data on school report cards or other accountability rating system (Education Commission of the States, 2011) as well as scaling up lessons for Florida itself as the state puts in place a new college readiness assessment that will be administered by districts.

In this study we focus on accelerated coursework participation and performance, which are good predictors of college readiness and post-secondary success. We are interested in whether the introduction of DA affects students' participation and performance in advanced coursework. We also explore the mechanism through which DA affects accelerated coursework. For instance, are schools working to make sure that students are "on track" for accelerated coursework? Are schools encouraging students to take certain benchmark courses on time (e.g, taking geometry or higher by 10$^{th}$ grade)? Last, given that certain groups of students are traditionally under-represented in accelerated coursework, we explore the heterogeneity in the policy effect. Specifically, the following research questions and sub-questions guide our efforts:

- Does accountability that includes college readiness and rigorous course taking as outcome measures have an effect on students' advanced course taking and performance?
- Does the impact depend on the quality of the school as measured by school grades or value-added scores?
- Are the effects, if any, different for sub-groups of students (e.g., socio-economic status, race/ethnicity)?

**3. Methodology**

Since students' participation and performance in accelerated coursework enter directly into the calculation of school grades as do graduation rates, it is possible that, over time, schools learned their way around the system such that any policy effects we estimate are biased. More schools are earning A's than before (see Appendix Table 2). Schools are either really getting better or they are gaming the system. To address this issue, we include two additional sets of outcomes, which are not used to calculate school grades: whether or not students take certain benchmark courses at certain grades (e.g., taking Algebra I or higher in $9^{th}$ grade, taking geometry or higher in $10^{th}$ grade, and taking algebra II or higher in $11^{th}$ and 12th grade).[6]

### 3.1 Identification

Our data include twelve cohorts of students with the first cohort entering high school in 2000-01 school year and the last cohort entering in 2012-13 school year. For cohorts that have graduated and left school before 2009-10 school year, or for cohorts that have entered high schools after 2009-10 school year, they are either never exposed or always exposed to the implementation of DA and thus do not contribute to our identification. Instead, our identification is based on these cohorts that have entered high schools between 2006-07 and 2008-09 school year, and thus are enrolled in schools both before and after DA. Analytically, let $i$ , $s$ , and $t$ index student, school, and school year respectively, the model takes the following form:

$$P_{ist} = \Phi(\alpha + \beta A_{i8th} + \gamma S_{ist} + \delta DA_{ist} + \phi R_{st} + g_{ist} + d_t) \qquad (1)$$

where $P_{ist}$ represents student outcomes, such as the probability of taking/passing AP tests in Math, English, Science or Social Science, the probability of taking Algebra I or higher in $9^{th}$ grade and on the probability of receiving a high school diploma within four years; $A_{i8th}$ is normed student test scores in $8^{th}$ grade FCAT, $S_{ist}$ is a vector of student characteristics that includes race/ethnicity, socio-economic status (proxied by eligibility for free/reduced price lunch), educational need (having learning disabilities, or limited English proficiency), and

[6] We leverage the longitudinal nature of our data and use panel data methods to estimate the impact of the grading change on these sets of outcomes. We are using the student as our level of analysis because we are ultimately interested in the impact of the change in accountability components on student outcomes, which is ultimately the goal of holding schools accountable through 'school grading'.

discipline records;[7] $DA_{ist}$ takes the value of zero if students are enrolled in schools before DA was introduced in 2009-10 school year and unity if students are enrolled in schools after DA has been implemented.[8] $R_{st}$ includes school characteristics , such as average teacher experience, teacher credentials, school compositions, and school quality proxied by school grade in previous school year; $g_{ist}, d_t$ are grade and time effects respectively, and finally $\Phi$ the CDF of a normal distribution. We include pre-DA time trend to control for any pre-existing trend in accelerated coursework. We also include post-time trend to control for any state-wide policy changes that affect students and schools (e.g., changes in graduation requirements in for students entering 9th grade in 2010-11). The parameter of interest is $\delta$ which measures the effects of DA on AP participation (performance) and benchmark course-takings[9]. This set up is essentially a difference-in-difference estimator with the first difference being student-level gain (since we controlled for prior achievement) and the second difference being pre- and post-DA policy change (Wooldridge, 2002; Jacob, 2005). Because of the difference in participation and passing rate across subjects, we estimate equation (1) separately for Math, English, Science and Social Science. We also adjust the standard errors to account for school-year level clustering. Finally, to explore the heterogeneities by student characteristics (i.e., race/ethnicity, and socio-economic status) and by school contexts (i.e., high performing schools,

---

[7] Disciplinary behavior may serve as a proxy for attitude, motivation, etc., and will help us account for heterogeneities across students.

[8] If DA encourages more students to take advanced courses, we would expect the effects to accumulate over time: the longer DA was put in place, the more students who signed up for or were assigned to advanced courses. Thus, as a sensitivity analysis, we will include indicators for each year after the implementation of DA. If the effects accumulate, then the coefficients should point to the same direction.

[9] Due to the longitudinal nature of our data, including student fixed effects seems appealing. We did so in a linear probability model and the results are qualitatively similar to the one that uses 8th grade test scores. Since it is computationally challenging to include student fixed effects in non-linear models, we choose to use the model specification that includes 8th grade test scores

and low performing schools), we also run the model separately for different subgroups and for different schools.

There is an important caveat on the proposed identification and estimation strategy and that is the possible simultaneous effect of DA on both the supply (number offered) and demand (number taken) for advanced courses. However, we argue, with supportive qualitative and quantitative evidence that this is more of a demand side story. First, the purpose of including accelerated coursework participation in school grades is to increase the percentage of students taking advanced courses (Florida Department of Education, 2009). Second, qualitative work from the National Center for Scaling Up Effective Schools[10] suggested that much effort has been focused on enrolling students, especially low-performing, and or minority students in advanced courses (Iatarola, Rutledge, Kim & Brown, 2015). Third, the number and share of schools offering advanced coursework have remained relatively stable while the number and share of students taking advanced courses have increased substantially. As a matter of fact, schools' willingness to offer an advanced course is more related to student demand but resource constraint such as number of instructional staff (Iatarola, Conger & Long, 2011).

**3.2. Data & Sample**

In order to provide early evidence on the effect of DA, we rely on a decade plus of data on individual students and schools in Broward County Public School District. Our sample includes 350,588 high school students in 38 high schools from 2000-01 through 2012-13 school years. A descriptive summary of student outcomes and characteristics is included in Table 1. The rate of students taking benchmark math courses in 9th, 10th and 11th or 12th grades is higher in the post-DA period as compared to the pre period. AP participation rates increased in all

[10] This study is a supplemental study of the National Center, thus we have insights from the qualitative data collection and analysis that was conducted in four study schools in Broward County.

subjects, most notably in English that nearly doubled and social sciences that did double. Performance on AP exams, however, remained the same or slightly lower as did the average four-year graduation rate. Most of the characteristics and educational needs resemble those in a typical large urban school district. For instance, 36% of students are Black and 24% percent of students are Hispanic. By and large student characteristics remain stable during the time periods with the exception of free/reduced price lunch eligibility. The increase in free/reduced lunch eligible students is more dramatic since 2008, and may be picking up the fact that more families are going through economic hardship as a result of the recent financial recession and may be driven as well by better efforts to sign up families. About a third of schools earn an A letter grade and over time more schools are earning an A or B grade.

**Table 1. Descriptive Summary of Outcome Variables and Student Characteristics, 2001-2013**

| | All Years | Pre 2000-01 through 2008-09 | Post 2009-10 through 2012-13 |
|---|---|---|---|
| **Outcomes** | | | |
| Took Geometry or higher in 10th grade | 0.69 | 0.65 | 0.80 |
| Took Algebra II or higher in 11th or 12th grade | 0.63 | 0.58 | 0.76 |
| Took Geometry of higher in 9th grade | 0.25 | 0.24 | 0.27 |
| Took Algebra II or higher in 10th grade | 0.28 | 0.25 | 0.33 |
| AP Participation (11th and 12th graders) | | | |
| Math | 0.05 | 0.04 | 0.07 |
| English | 0.10 | 0.08 | 0.14 |
| Science | 0.06 | 0.05 | 0.09 |
| Social Science | 0.15 | 0.11 | 0.23 |
| AP Performance (scoring above proficiency) | | | |
| Math | 0.59 | 0.58 | 0.60 |
| English | 0.53 | 0.53 | 0.53 |
| Science | 0.39 | 0.40 | 0.38 |
| Social Science | 0.48 | 0.48 | 0.47 |
| Average four-year cohort graduation rate | 73% | 73% | 72% |
| | | | |
| **Students - socio-demographics, educational needs and performance** | | | |
| Asian | 0.03 | 0.03 | 0.04 |
| Black | 0.36 | 0.36 | 0.37 |
| Hispanic | 0.24 | 0.23 | 0.28 |
| White | 0.34 | 0.36 | 0.29 |
| Free/Reduced Price Lunch | 0.37 | 0.30 | 0.52 |
| Exceptional Education | 0.09 | 0.10 | 0.09 |
| Limited English Proficiency | 0.07 | 0.08 | 0.06 |
| FCAT Math (8th grade), normed | 0.07 | 0.07 | 0.08 |
| FCAT English (8th grade), normed | 0.06 | 0.06 | 0.07 |
| | | | |
| **Schools - characteritics and performance** | | | |
| Teacher Experience (years) | 18 | 19 | 16 |
| Teacher with an advanced degree (master or higher) | 0.49 | 0.49 | 0.47 |
| Distribution of School Accountability Grades | | | |
| A | 0.31 | 0.21 | 0.54 |
| B | 0.18 | 0.15 | 0.26 |
| C | 0.33 | 0.39 | 0.18 |
| D | 0.15 | 0.21 | 0.02 |
| F | 0.03 | 0.04 | 0.01 |
| | | | |
| N Students | 350,588 | - | - |
| N Schools | 38 | - | - |

Over time, we see some interesting changes in AP participation and performance by student race/ethnicity. For instance, over time the participation gap between Asian/White and Hispanic/Black students has grown modestly (Figure 1). In terms of performance, while there wass a quite substantial gap in AP science passage rate in 2001, by the end of 2013 the passage rate is about the same for Asian, Hispanic and White students. The average passage for Black students, however, declined significantly (Figure 2).

Figure 1. Changes in AP Participation among 11th and 12th graders, by subject and race.

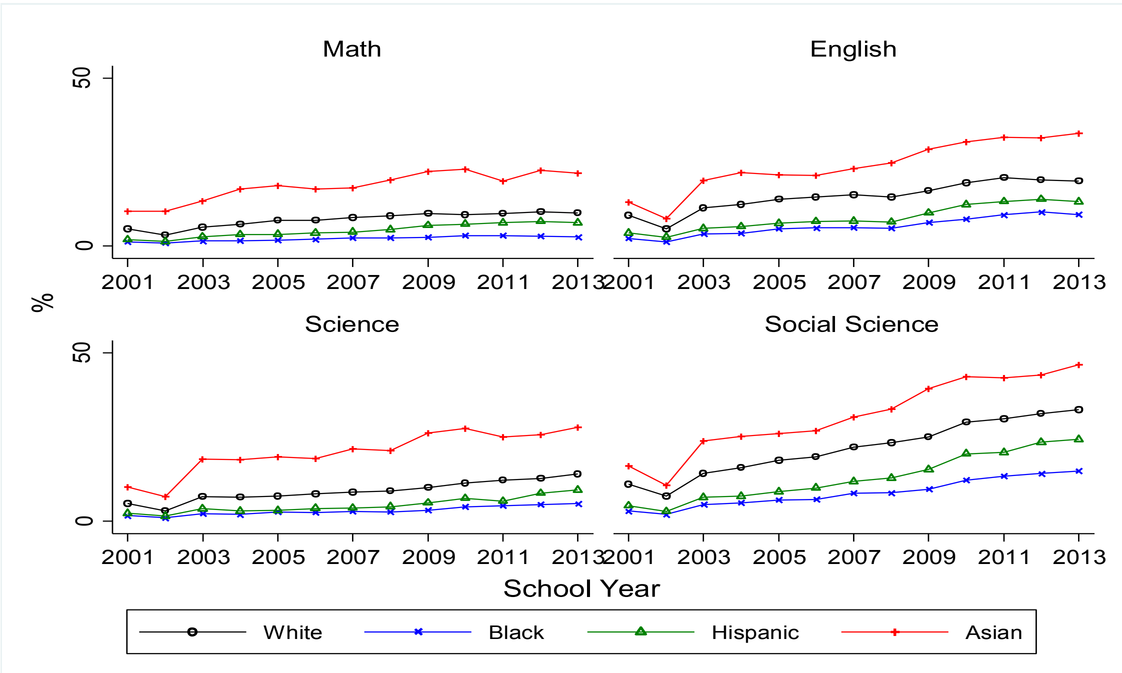
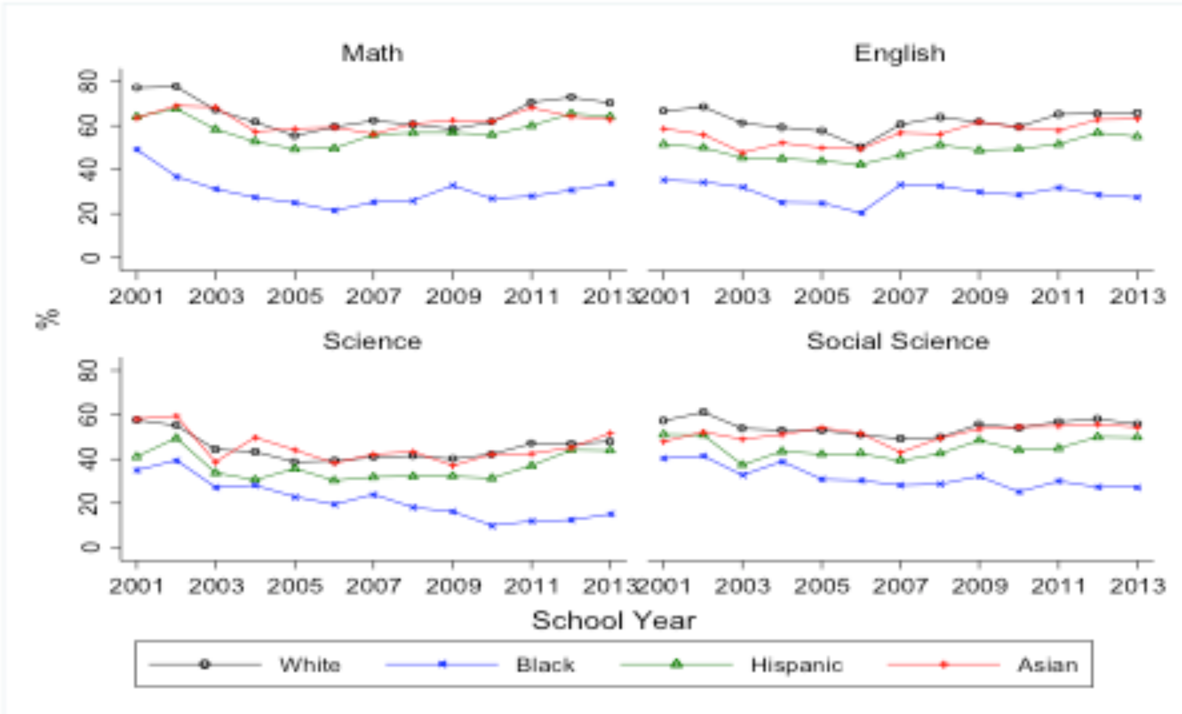


Figure 2. Changes in AP Performance among 11th and 12th graders, by subject and race.




**4. Results**

**4.1. Impact of DA on Students Probability of Taking and Passing at least one AP Exam**

Table 2 summarizes the impact of DA on students' probability of taking at least one AP exam. Each cell represents a separate Probit model with covariates specified in equation (1). DA increases the probability of participating in AP English and Science by 0.0385 and 0.0225, respectively. The point estimate is considerably larger in Social Science (0.0777), which is not surprising given that Social Science may be easier and does not require rigorous prerequisite coursework.

When we break the estimates down by student race/ethnicity and socio-economic status, there are some variations among subgroups. While on average students see a small increase in their probability of taking AP English, Science, and Social Science exams, the effect is larger for white students. When we look across schools, most of gains in English and Science are concentrated in relatively high performing schools (i.e., those that have earned an A or B letter grade in previous school year). Since these AP exams require rigorous prerequisite coursework and have lower passage rates, low performing schools probably are better off concentrating on easier to pass subjects such as Social Science, though the effects are still largest in the higher performing schools.

In more recent years, schools and districts are encouraging students to take AP exams in early grades. For instance, students are taking AP social sciences in grade 9 or 10 and as a matter of fact, in 2012-13 school year, 9th and 10th graders make up 47% of all AP Social Science takers.[11] To see if DA has any potential impact on these students, we run the same analysis and the results are qualitatively similar: DA increased 9th and 10th graders' probability of taking at least one AP Social Science exam and the effect is by and large the same among subgroups.[12]

**Table 2. Impact of DA on Students' Probability of Taking AP Exam by Subject and by Student Socio-demographics and School Accountability Grade.**

| | Full Sample | Black | Hispanic | White | FRPL | A | B | C | D |
|---|---|---|---|---|---|---|---|---|---|
| Math | 0.0009 | 0.0010 | 0.0006 | 0.0011 | 0.0011 | 0.0011 | 0.0050 | -0.0031 | 0.0040** |
| | [0.0022] | [0.0014] | [0.0026] | [0.0049] | [0.0015] | [0.0051] | [0.0033] | [0.0030] | [0.0019] |
| English | 0.0385*** | 0.0215** | 0.0344** | 0.0506** | 0.0272*** | 0.0501** | 0.0189 | 0.0495 | 0.0072 |
| | [0.0147] | [0.0107] | [0.0145] | [0.0246] | [0.0096] | [0.0244] | [0.0225] | [0.0306] | [0.0117] |
| Science | 0.0225*** | 0.0116** | 0.0208*** | 0.0381** | 0.0126*** | 0.0369** | 0.0161* | -0.0003 | 0.0071 |
| | [0.0084] | [0.0047] | [0.0070] | [0.0160] | [0.0048] | [0.0156] | [0.0086] | [0.0081] | [0.0094] |
| Social Science | 0.0777*** | 0.0445*** | 0.0794*** | 0.1118*** | 0.0520*** | 0.0922*** | 0.0916*** | 0.0855*** | 0.0374** |
| | [0.0182] | [0.0126] | [0.0176] | [0.0358] | [0.0112] | [0.0327] | [0.0294] | [0.0249] | [0.0176] |

1. Sample restricted to 11th and 12th graders
2. Each cell represents a Probit estimation with convariates including student race/ethnicity, educational need (ESE, LEP status), family background (Free lunch status), grade level, 8th grade FCAT test scores (standardized), school characteristics including teacher experience, teacher credentials, and last year's school grade, and school year fixed effects.
3. Standard errors in parenthesis and adjusted for school-year clustering.
4. *** p<0.01, ** p<0.05, * p<0.1

[11] This mostly happens in AP Social Sciences. For Math, English and Science, 9th and 10th graders make up less than 8 percent of all exam takers.
[12] These results are available upon request.

Table 3 summarizes the impact of DA on AP performance. On average DA does not appear to have any impact on students' performance in AP Math, English or Social Science, though it increases the probability of passing AP Science by 0.0862. The performance gains are concentrated among White students and among high performing schools (e.g., earning an A school grade). Students in low performing schools (e.g., earning a D letter grade) however experienced a modest decline in AP performance as a result of DA. One possible explanation is the difference in instructional resources (e.g., teacher quality) across schools, particularly with respect to teachers assigned to teach AP courses as experience and education levels may matter (Iatarola & Kim, 2015). Anther possible, though not likely explanation is that low performing schools are over-matching students who are not academically prepared to take these AP exams.

**Table 3. Impact of DA on the Probability of Passing AP Exam (conditional on taking the exam), by Student Socio-demographics and School Accountability Grade.**

| | Full Sample | Black | Hispanic | White | FRPL | A | B | C | D |
|---|---|---|---|---|---|---|---|---|---|
| Math | 0.0819 | -0.0366 | -0.0008 | 0.1553* | -0.0037 | 0.1187 | 0.0750 | 0.0669 | -0.3790*** |
| | [0.0900] | [0.0834] | [0.1161] | [0.0860] | [0.0839] | [0.1147] | [0.1606] | [0.1188] | [0.0646] |
| English | 0.0002 | -0.0607* | 0.0160 | 0.0304 | -0.0409 | 0.0054 | 0.0357 | 0.0313 | -0.2179*** |
| | [0.0364] | [0.0329] | [0.0573] | [0.0383] | [0.0332] | [0.0468] | [0.0909] | [0.0649] | [0.0338] |
| Science | 0.0862* | -0.0080 | 0.0611 | 0.1051** | 0.0466 | 0.1127** | 0.0924 | 0.0401 | -0.2242*** |
| | [0.0450] | [0.0342] | [0.0660] | [0.0514] | [0.0355] | [0.0519] | [0.1134] | [0.0854] | [0.0363] |
| Social Science | 0.0010 | -0.0651 | -0.0187 | 0.0275 | -0.0343 | 0.0193 | 0.0212 | -0.1334* | -0.1729*** |
| | [0.0571] | [0.0422] | [0.0850] | [0.0608] | [0.0441] | [0.0810] | [0.1015] | [0.0724] | [0.0546] |

1. Sample includes 11th and 12th graders who take AP exams
2. Each cell represents a Probit estimation with convariates including student race/ethnicity, educational need (ESE, LEP status), family background (Free lunch status), grade level, 8th grade FCAT test scores (standardized), school characteristics including teacher experience, teacher credentials, and last year's school grade, and school year fixed effects.
3. Standard errors in parenthesis and adjusted for school-year clustering.
4. *** p<0.01, ** p<0.05, * p<0.1

The results so far suggest that DA improves students' participation in AP exams but not performance on AP exams. For instance, while students in all schools see an increase in their AP participation, those in low-performing schools experienced a decrease in AP performance. Since we can not compute individual teacher value-added out of data constraint, we look at the distribution of teacher experience, which is a reliable predictor of teacher quality. Average teacher experience is 11 in D schools, 2 years fewer than the number in A or B schools. This echoes our work by Iatarola and Kim (2015) on peer and teacher effects, though their study is does not provide causal findings. Another possible explanation is over-matching where schools are maximizing their participation rate for accountability purposes. However, this is not likely based on our qualitative work that suggests schools are adhering to Broward's course assignment matrix (Iatarola, Rutledge, Kim & Brown, 2015).

### 4.2. Impact of DA on Students Likelihood of Graduating within Four Years

Table 4 summarizes the impact of DA on the probability of graduating from high school within four years. DA decreases the probability of receiving a standard high school diploma by 0.0255. There are some issues with our data, so these results are very preliminary, moreover graduation, at least, in this study is a distal outcome as only one cohort graduated, so it is possible that we are picking up some odd things about this cohort or the current issues with our data on this particular outcome.

**Table 4. Impact of DA on Likelihood of Graduating within Four Year by Student Socio-demographics and School Accountability Grade.**

| | Full Sample | Black | Hispanic | White | FRPL | A | B | C | D |
|---|---|---|---|---|---|---|---|---|---|
| 4-Year Graduation | -0.0255*** | 0.0250*** | -0.0041 | -0.0011 | -0.0219** | -0.0105 | -0.0582*** | -0.0190 | 0.0081 |
| | [0.0081] | [0.0084] | [0.0065] | [0.0055] | [0.0092] | [0.0076] | [0.0122] | [0.0232] | [0.0179] |

1. Each cell represents a Probit estimation with convariates including student race/ethnicity, educational need (ESE, LEP status), family background (Free lunch status), grade level, 8th grade FCAT test scores (standardized), school characteristics including teacher experience, teacher credentials, and last year's school grade, and school year fixed effects.

2. Standard errors in parenthesis and adjusted for school-year clustering.

3. *** p<0.01, ** p<0.05, * p<0.1

## 5. Robustness Check

### 5.1 Impact of DA on Benchmark Course-takings

Since student participation and performance in accelerated coursework is directly related to school grades, one concern is that schools are maximizing their participation rate for accountability purposes (without real improvement in student achievement). We address this issue by looking at the impact of DA on students' benchmark course-takings, which prepare students ready for accelerated coursework but do not affect school grades directly. We focus on Math and look at two types of readiness measures: "on track" and "above track". A student is considered "on track" if (s)he takes Algebra I or higher in 9$^{th}$ grade, takes Geometry or higher in 10$^{th}$ grade, and takes Algebra II or higher in 11$^{th}$ and 12$^{th}$ grade and considered "above track" if (s)he takes Geometry or higher in 9$^{th}$ grade and takes Algebra II or higher in 10$^{th}$ grade. As seen in Table 5, DA increases students' probability of being "on track" and the effect is observed for all subgroups of students and for all types of schools. For instance, DA increases students' probability of taking Algebra I or higher in 9$^{th}$ grade by 0.3996 and the effect is the largest for low-income students. It increases students' probability of taking Geometry of higher in 10$^{th}$ grade by 0.3144 and the point estimate is the largest in low performing schools. The results seem to suggest that disadvantaged students and low performing schools are "catching up".

**Table 5. Impact of DA on Benchmark Course-takings by Student Socio-demographics and School Accountability Grade.**

| | Full Sample | Black | Hispanic | White | FRPL | A | B | C | D |
|---|---|---|---|---|---|---|---|---|---|
| **'On Track'** | | | | | | | | | |
| Probability of taking Algebra I or higher in 9th grade | 0.3996*** | 0.2344*** | 0.2656*** | 0.2096*** | 0.4983*** | 0.3014*** | 0.4541*** | 0.4436*** | 0.3353*** |
| | [0.0388] | [0.0358] | [0.0334] | [0.0217] | [0.0455] | [0.1083] | [0.1296] | [0.0537] | [0.0805] |
| Probability of taking Geometry or higher in 10th grade | 0.3144*** | 0.1534*** | 0.1387*** | 0.0995*** | 0.0426 | 0.0880*** | 0.0029 | 0.1244 | 0.5637*** |
| | [0.1008] | [0.0229] | [0.0170] | [0.0120] | [0.0466] | [0.0340] | [0.0708] | [0.1646] | [0.0264] |
| Probability of taking algebra II or higher in 11/12th grade | 0.1963*** | 0.1953*** | 0.1497*** | 0.1452*** | 0.2151*** | 0.1972*** | 0.1828** | 0.1616*** | 0.1761*** |
| | [0.0349] | [0.0214] | [0.0167] | [0.0180] | [0.0398] | [0.0452] | [0.0814] | [0.0584] | [0.0584] |
| **'Above Track'** | | | | | | | | | |
| Probability of taking Geometry or higher in 9th grade | 0.0951*** | 0.0051 | 0.0033 | 0.0239 | 0.0874*** | 0.0578 | -0.1068 | 0.0676* | 0.0503 |
| | [0.0353] | [0.0112] | [0.0167] | [0.0224] | [0.0297] | [0.0660] | [0.0943] | [0.0364] | [0.0441] |
| Probability of taking Algebra II or higher in 10th grade | 0.2145 | 0.0795*** | 0.0847*** | 0.1150*** | -0.0405 | 0.0504 | -0.0785 | 0.0323 | 0.8894*** |
| | [0.2202] | [0.0201] | [0.0296] | [0.0307] | [0.0297] | [0.0669] | [0.0856] | [0.2339] | [0.0189] |

1. We exclude 2001 school year because no 8th grade FCAT scores were available.
2. Each cell represents a Probit estimation with convariates including student race/ethnicity, educational need (ESE, LEP status), family background (Free lunch status), grade level, 8th grade FCAT test scores (standardized), school characteristics including teacher experience, teacher credentials, and last year's school grade, and school year fixed effects.
3. Standard errors in parenthesis and adjusted for school-year clustering.
4. *** p<0.01, ** p<0.05, * p<0.1

When we look at "above track" measures, we find that DA increases the probability of taking geometry or higher in 9$^{th}$ grade by 0.0951 and again low-income students are benefiting the most. It also increase the probability of taking Algebra II or higher in 10th grade among major race/ethnicity groups and in low performing schools.[13] Again, some evidence that disadvantaged students and low performing schools are "catching up".

### 5.2. Treatment Effect by Year

In this section we are interested in whether DA effect persists. If it does, we would expect the effect to accumulate over time and it appears so. Table 6 shows the estimates of treatment effect by year. Students see a 0.0572 increase in their probability of taking AP Social Science first year after DA and the effect increases to 0.0872 by the end of the fourth year. There are later year effects in English and Science as well. The delay in effect may be due to the fact that students in subjects other than Social Sciences, may need more prior preparation in their coursework or at least a higher track of prior coursetaking.[14] Though DA do not have any short-term effect on AP performance, by the end of the fourth year, DA increases students' probability of passing AP Science by 0.0901.

[13] The point estimate is unusually large in D schools. It is likely because a very small number of students taking algebra II or higher in 10th grade in these schools.

[14] Another potential threat to our identification is the simultaneous effect of DA on course offerings and course-takings. The number of AP courses being offered, however, does not change significantly within and across schools over time. Our regression analyses are available upon request.

**Table 6. Treatment Effect, by Years in Treatment.**

| | AP Participation | | | |
|---|---|---|---|---|
| | AP Math | AP English | AP Science | AP Social |
| 1 year after DA | 0.0010 | 0.0190 | 0.0106 | 0.0572*** |
| | [0.0021] | [0.0148] | [0.0076] | [0.0157] |
| 2 years after DA | 0.0009 | 0.0369** | 0.0106 | 0.0573*** |
| | [0.0020] | [0.0167] | [0.0082] | [0.0175] |
| 3 years after DA | 0.0029 | 0.0408** | 0.0197* | 0.0714*** |
| | [0.0024] | [0.0177] | [0.0104] | [0.0191] |
| 4 years after DA | 0.0010 | 0.0433** | 0.0264** | 0.0872*** |
| | [0.0023] | [0.0182] | [0.0112] | [0.0225] |
| | AP Performance | | | |
| | AP Math | AP English | AP Science | AP Social |
| 1 year after DA | 0.0703 | 0.0395 | 0.0451 | 0.0498 |
| | [0.0770] | [0.0349] | [0.0437] | [0.0568] |
| 2 years after DA | 0.0975 | 0.0311 | 0.0530 | 0.0331 |
| | [0.0752] | [0.0367] | [0.0431] | [0.0590] |
| 3 years after DA | 0.1090 | 0.0417 | 0.0887* | 0.0404 |
| | [0.0773] | [0.0372] | [0.0486] | [0.0593] |
| 4 years after DA | 0.0811 | 0.0002 | 0.0901* | 0.0010 |
| | [0.0879] | [0.0364] | [0.0489] | [0.0571] |

## 6. Conclusions and Implications

Using administrative student level data from Broward County public high schools from 2000-01 to 2012-13 school years, we estimate the impact of Differentiated Accountability (DA) on students' participation and performance in accelerated coursework, which is a reliable measure of college readiness and post-secondary success. We find that DA increases students' probability of taking AP exams in English, Science and Social Science, but it does not have any impact on Math, which requires more rigorous prerequisite coursework. While all students are benefiting from DA, the effect is the largest among low-income students and among White students. Across schools most of the gains in AP English and Science are driven by the gains in high performing schools. In terms of AP performance, DA does not appear to have any impact on likelihood of passing AP exams, except for a few cases where white students and students in high performing schools see an improvement in their performance in AP science. There is also evidence suggesting that low performing schools see a decline in their AP performance across all subjects. The impact of DA is broad. It elicits efforts from schools to prepare students for accelerated work by taking benchmark courses on time (or ahead of time). The impact of DA is also persistent and over time the positive effect accumulates.

These findings have significant implications. First, it highlights the challenges increasing the likelihood of students in low-performing schools to take AP courses and be successful on exams, which may stem from differences in the quality of instruction between low- and high-performing schools. Schools can nudge students into taking more AP exams, but without quality instruction, students are not likely to reap the benefit of accelerated coursework. Second, the impact, if any, is smaller for Hispanic and Black students, which may reflect lower overall rates of taking AP courses. In future analyses we plan on incorporating a broader definition of accelerated course work that would include IB, AICE and dual enrollment. It may be that these other accelerated pathways, especially dual enrollment, are ones that are more likely to be taken by students who do excel, but who may not meet the standards set for course assignment to AP and IB or AICE courses. It will also give us a better sense of the wider range of course strategies the district and schools are using to ensure that they achieve in areas that are recognized and rewarded by the school grading system.

References


Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes? A cross-state analysis. *Educational Evaluation and Policy Analysis,* 24(4), 305-331.

Conger, D., Long, M. C., & Iatarola, P. (2009). Explaining race, poverty, and gender disparities in advanced course-taking. *Journal of Policy Analysis and Management, 28*(4), 555-576.

Darling-Hammond, L. (2004). Standards, accountability, and school reform. *The Teachers College Record,* 106(6), 1047-1085.

Dee, T., & Jacob, B. (2011). The impact of No Child Left Behind on student achievement. *Journal of Policy Analysis and Management,* 30(3), 418-446.

Education Commission of the States. (2011). *State Notes: Advanced Placement.* Denver, CO.

Firestone, W. A. (2004). *The ambiguity of teaching to the test: Standards, assessment, and educational reform.* Mahway, NJ: Laurence Erlbaum Associates.

Florida Association of District School Superintendents (Undated). *2013 Legislative Platform*. http://www.fldoe.org/board/meetings/2013_03_19/fadss.pdf

Florida Department of Education. (2008). *Differentiated Accountability – Proposal to the U.S. Department of Education.* Tallahassee, FL. https://www2.ed.gov/admins/lead/account/differentiatedaccountability/fldap.pdf

Florida Statutes (2014). 1003.4282 *Requirements for a Standard High School Diploma.* http://www.leg.state.fl.us/Statutes/index.cfm?App_mode=Display_Statute&Search_String=&URL=1000-1099/1003/Sections/1003.4282.html

Hanushek, E.A., & Raymond, M. E. (2004). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management,* 24(2), 297-327.

Iatarola, P., Conger, D., & Long, M.C. (2011), Determinants of high schools' advanced course offerings. *Educational Evaluation and Policy Analysis*, 33(3)*,* 340-359.

Iatarola, P. & Kim, T.K.(2015). *Expansion of AP – Course, Classroom and Outcome Implications: The Case of Broward County Public High Schools.* Unpublished manuscript, National Center on Scaling Up Effective Schools. Manuscript in preparation.

Iatarola, P., Rutledge, S.A., Kim, T.K., & Brown, S. (2015). *Assigning Rigor, High School Courses: A Mixed Methods Analysis.* Unpublished manuscript, National Center on Scaling Up Effective Schools. Manuscript in preparation.

Kane, T. J., & Staiger, D.O. (2002). The Promise and pitfalls of using imprecise school accountability measures. *Journal of Economic Perspectives,* 16(4), 91-114.

Long, M. C., Conger, D., & Iatarola, P. (2012). Effects of high school course-taking on secondary and postsecondary success. *American Educational Research Journal, 49*(2), 285-322.

McDonnell, L. M. & Elmore, R. (1991). Getting the job done: Alternative policy instruments. Education Policy Implementation. A. Odden. Albany, NY, SUNY Press: 157-183.

McLaughlin, M. W. (1987). The Rand change agent study revisited: Macro perspectives and micro realities. *Educational Researcher,* 19(9), 11-16.

Neal, D., & Schanzenbach, D.W. (2010). Left behind by design: Proficiency counts and test-based accountability. *The Review of Economics and Statistics,* 92(2), 263-283.

Reback, R. (2008). Teaching to the rating: School accountability and the distribution of student achievement, *Journal of Public Economics,* 92 (5-6), 1394-1415.

U.S. Department of Education. (2009). *ESEA Reauthorization: A Blueprint for Reform.* Washington, DC: U.S. Government Printing Office.

# APPENDIX TABLES

**Appendix Table 1. Florida School Grade Calculations 2009-10.**

| | | | 2009-10 | | | 2010-11 | | | 2011-12 | | | 2012-13 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Points | % of Compo-nent | % Overall | Points | % of Compo-nent | % Overall | Points | % of Compo-nent | % Overall | Points | % of Compo-nent | % Overall |
| Assessment Components 50% | Performance | Reading | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 |
| | | Math (Alg 1) | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 |
| | | Writing | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 |
| | | Science (Bio) | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 |
| | Learning Gains All Students | Reading | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 |
| | | Math (Alg 1) | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 |
| | Learning Gains Lowest 25% | Reading | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 |
| | | Math (Alg 1) | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 |
| | | Subtotal | 800 | 100 | 50 | 800 | 100 | 50 | 800 | 100 | 50 | 800 | 100 | 50 |
| Other Components 50% | Acceleration | Participation | 200 | 25.00 | 12.50 | 175 | 21.88 | 10.94 | 150 | 18.75 | 9.38 | 150 | 18.75 | 9.38 |
| | | Performance | 100 | 12.50 | 6.25 | 125 | 15.63 | 7.81 | 150 | 18.75 | 9.38 | 150 | 18.75 | 9.38 |
| | Graduation | Overall-4 year | 200 | 25.00 | 12.50 | 200 | 25.00 | 12.50 | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 |
| | | Overall-5 year | 0 | 0.00 | 0.00 | 0 | 0.00 | 0.00 | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 |
| | | At Risk-4 year | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 | 50 | 6.25 | 3.13 | 50 | 6.25 | 3.13 |
| | | At Risk-5 year | 0 | 0.00 | 0.00 | 0 | 0.00 | 0.00 | 50 | 6.25 | 3.13 | 50 | 6.25 | 3.13 |
| | College Readiness | Reading | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 |
| | | Math | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 | 100 | 12.50 | 6.25 |
| | | Subtotal | 800 | 100 | 50 | 800 | 100 | 50 | 800 | 100 | 50 | 800 | 100 | 50 |
| | | TOTAL | 1600 | | | 1600 | | | 1600 | | | 1600 | | |

Note: Additional points on 'other components' if demonstrate annual increase in component. Points deducted if decrease in 'other components'.
Source: http://schoolgrades.fldoe.org/reports/index.asp

**Appendix Table 2. Changes in the distribution of school grades from year t to year t+1.**

| | School Grade in Year t+1 | | | | | |
|---|---|---|---|---|---|---|
| School Grade | A | B | C | D | F | Total |
| A | 93 | 22 | 2 | 0 | 0 | 117 |
| B | 25 | 26 | 20 | 3 | 0 | 74 |
| C | 14 | 25 | 73 | 19 | 4 | 135 |
| D | 3 | 5 | 23 | 33 | 3 | 67 |
| F | 0 | 0 | 3 | 4 | 0 | 7 |
| Total | 135 | 78 | 121 | 59 | 7 | 400 |